

# 자율주행차량 환경에서 SDV 차량용 소프트웨어 비교 연구

김현지\*, 윤철희\*\*, 이근희\*\*\*

## Comparative Study of Software for SDV Vehicles in Autonomous Vehicle Environments

Kim Hyun Ji\* and Cheol Hee Yoon\*\*

### 요약

SDV는 차량 기능을 소프트웨어로 정의하고 업데이트함으로써 자율주행 구현의 핵심 기반으로 자리 잡고 있으며, 다양한 소프트웨어 플랫폼(Android Automotive OS, QNX, Tesla 독자 플랫폼)을 통해 사용자 맞춤형 서비스와 구독형 비즈니스 모델 등 새로운 가치를 창출하게 한다. 자율주행 차량 환경에서 소프트웨어 중심 자동차(SDV)의 역할과 중요성을 분석하고, 주요 SDV 소프트웨어 플랫폼을 비교하여 기술적 과제 평가를 시도하였다. SDV가 완전 자율주행을 실현하려면 사이버 보안, 고성능 데이터 처리, OTA(무선 업데이트) 안정성 등 다수의 기술 과제를 해결해야 하는데 본 논문은 이러한 과제들을 살펴보고 SDV가 미래 차량 산업의 혁신적 발전을 주도할 필수적인 인프라로서 자율주행 Level 5 구현을 위한 SDV의 잠재적 역할을 고민하였다.

### Abstract

SDV is positioned as a core foundation for implementing autonomous driving by defining and updating vehicle functions in software, and enables the creation of new values such as customized services and subscription business models through various software platforms (Android Automotive OS, QNX, Tesla's proprietary platform). This paper analyzed the role and importance of software-centric vehicles (SDVs) in the autonomous vehicle environment and attempted to evaluate technical challenges by comparing major SDV software platforms. In order for SDVs to realize fully autonomous driving, they must solve various technical challenges such as cybersecurity, high-performance data processing, and OTA (over-the-air update) stability. This paper examined these challenges and considered the potential role of SDVs in implementing Level 5 autonomous driving as essential infrastructure that will lead the innovative development of the future automobile industry.

### Key words

Software-Defined Vehicle (SDV), Autonomous Driving Technology, Over-the-Air (OTA) Updates,

---

\* 건국대학교 컴퓨터공학부, dolcei6966@skku.edu,

\*\* 경찰대학교, bertter@police.ac.kr,

\*\*\* 연세대학교, gunhee.lee@yonsei.ac.kr

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임.  
(No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

## I. 서 론

오늘날 자동차 산업은 자율주행, 전기차, 커넥티드 카 등의 기술 발전과 동시에 급격한 변화를 맞이하고 있다. 소프트웨어 중심 자동차(SDV, Software-defined Vehicle)의 개념이 대두되고 있으며, 전통적인 하드웨어 중심의 차량 설계가 소프트웨어 기반으로 전환되는 중요한 패러다임 변화가 일어나고 있다. SDV는 소프트웨어를 통해 차량의 주요 기능을 정의하고 업데이트할 수 있는 개념으로, 이는 자율주행 기술의 상용화와 차량의 서비스화에 있어서 핵심적인 요소가 되었다. SDV는 기존 차량과 달리 차량 기능이 하드웨어 설계에 고정되지 않고 소프트웨어를 통해 확장하고 변형할 수 있는 혁신적인 방식을 제공하며, 과거의 자동차와 달리 새로운 기능을 추가하기 위해 하드웨어 변경이나 모듈 교체를 하지 않고, 무선 소프트웨어 업데이트(OTA)를 통해 자동차의 소프트웨어의 개선을 지속하고 새로운 기능을 추가할 수 있다. 이를 통해 자율주행 기능과 운전자 지원 시스템(ADAS)을 보다 유연하게 적용할 수 있으며, 차량의 자율성과 사용자 경험을 지속적으로 향상시킬 수가 있다. 본 논문은 자율주행차량 환경에서 SDV 소프트웨어의 발전 과정을 분석하고, 주요 SDV 소프트웨어 플랫폼을 비교하여 각 플랫폼의 장점과 한계를 평가하는 것을 목표로 하였다. 그리고, SDV 소프트웨어가 직면한 기술적 과제와 향후 발전 가능성을 모색함으로써 자율주행 차량의 미래 방향성을 제시하였다.

## II. 본 론

### 2.1 SDV 소프트웨어 개요

SDV는 차량 기능이 소프트웨어로 정의되고 제어되는 차량을 의미한다. 이는 차량이 단순한 기계적 하드웨어 장비를 넘어, 소프트웨어에 의해 성능과 기능이 결정되고 발전될 수 있는 존재로 변화하는 것을 뜻한다. 특히 자율주행 기술의 발전과 SDV의 도입은 매우 밀접하게 연관되어 있으며, 자율주행의 모든 단계(Level 0에서 Level 5)에 걸쳐 SDV의

역할이 중요해진다. 기존의 차량 설계는 하드웨어 중심으로 기능이 제한되었지만, SDV는 자율주행 단계가 높아질수록 요구되는 다양한 소프트웨어 기능과 성능을 차량 생애 주기 동안 지속적으로 제공할 수 있는 점이 특징이다. 특히, ADAS(Advanced Driver Assistance Systems)같은 운전자를 지원하는 시스템은 SDV의 핵심 소프트웨어 기능 중 하나로, 차량 주행 중 발생할 수 있는 위험 요소를 감지하고 경고하는 역할을 한다. 즉, SDV 소프트웨어는 차량의 기능과 성능을 지속적으로 개선하며 자율주행 기술이 요구하는 다양한 기능을 충족시킬 수 있다. 또한, SDV 소프트웨어 플랫폼은 자율주행 기능 구현에 있어서 중요한 역할을 하며, 각 플랫폼은 설계와 기능 면에서 차별화된 특징을 가진다. 대표적인 SDV 소프트웨어 플랫폼으로는 Google의 Android Automotive OS, Blackberry의 QNX, Tesla의 독자 소프트웨어가 있다.

#### a. Android Automotive OS

Google이 주도하는 Android Automotive OS는 차량의 인포테인먼트 기능을 중심으로 한 플랫폼으로, 다양한 어플 그리고 서비스를 통해 사용자의 경험을 강화하는 데 중점을 둔다. 이 OS는 사용자 친화적인 인터페이스와 클라우드 기반 서비스와의 호환성이 뛰어나 차량 내 다양한 기능을 통합할 수 있다. 그러나 자율주행과 관련한 고도의 실시간 반응성과 보안성에서는 다소 제한적이라는 평가를 받고 있으며, 비자율주행 기능에 초점을 맞추고 있다.

#### b. QNX

Blackberry의 QNX는 차량 내 실시간 데이터 처리가 필요한 자율주행 환경에서 안정적이고 보안성이 강한 소프트웨어 플랫폼으로 평가받고 있다. QNX는 실시간 운영체제(RTOS)의 특성을 활용하여 ADAS 및 자율주행과 같은 핵심 기능을 안정적으로 지원하며, 다양한 Tier-1 부품사와 협력하여 자율주행 소프트웨어의 표준 플랫폼으로 자리 잡고 있다. 이는 자율주행의 안전성과 신뢰성을 확보하는 데 중요한 기여를 한다.

### c. Tesla 독자 플랫폼

Tesla는 독자적인 소프트웨어 플랫폼을 통해 자율주행 기술을 실시간으로 개선하고 OTA 업데이트 기능을 통해 지속적인 성능 개선을 제공한다. Tesla 소프트웨어는 대규모 데이터 수집과 인공지능 기반 자율주행 알고리즘 개선에 중점을 두며, 사용자 맞춤형 자율주행 경험을 제공한다. 다만, 독자적인 플랫폼 사용으로 인해 다른 제조사와의 호환성은 낮다.

## 2.2 SDV 소프트웨어 아키텍처의 구성 요소와 역할

SDV 소프트웨어 아키텍처는 자율주행 기능을 포함한 차량의 제어와 관리를 위해 필수적인 구성 요소들로 이루어져 있다. 주요 구성 요소는 E/E(Electrical/Electronic) 아키텍처, 실시간 운영체제(RTOS), 중앙집중식 컴퓨팅 시스템으로 나뉜다.

### a. E/E 아키텍처

E/E 아키텍처는 차량 내 다양한 도메인(인포테인먼트, ADAS, 차체 제어 등)을 통합하고 관리하는 역할을 한다. SDV는 차량 내 하드웨어와 소프트웨어를 유연하게 연결하는 아키텍처를 통해 차량의 다양한 기능을 소프트웨어만으로 제어할 수 있다.

### b. 실시간 운영체제(RTOS)

자율주행 차량은 실시간 반응성이 중요하기 때문에 SDV는 실시간 운영체제를 통해 안정적인 자율주행 기능을 제공한다. RTOS는 특히 ADAS와 같은 실시간 반응이 필요한 자율주행 소프트웨어에서 중요한 역할을 하며, 차량의 안전과 관련된 기능을 신속하고 안정적으로 처리할 수 있게 한다.

### c. 중앙집중식 컴퓨팅 시스템

SDV에서는 차량 내 센서와 기기들이 방대한 양의 데이터를 생성하므로, 이를 효과적으로 처리하기 위한 중앙집중식 컴퓨팅 시스템이 필요하다. 중앙 컴퓨터는 차량 내 데이터를 통합적으로 관리하고 분석하여 자율주행의 의사결정 속도를 높이며, 여러

시스템 간의 데이터를 유기적으로 연계하여 최적화된 자율주행 환경을 제공한다.

### 다. SDV 소프트웨어의 주요 기술적 과제와 해결 방안

SDV의 성공적 구현에는 여러 기술적 도전 과제가 있다. 특히, 보안성과 데이터 관리 및 처리 능력이 중요하며, OTA 업데이트 기능이 중요하다. 아래와 같이 상세 설명이 가능하다.

#### a. 보안성

차량이 디지털화됨으로 인해 사이버 보안 위협이 나날이 상승하고 있다. 자율주행 차량은 외부 네트워크와 상시 연결되는 경우가 많아 해킹의 위험이 높다. SDV 소프트웨어는 보안 위협에 대응하기 위해 네트워크 보안을 강화하고, 강력한 암호화와 방화벽을 사용하여 보안성을 높이고 있다.

#### b. 데이터 관리 및 처리 능력

자율주행 차량은 무수히 많은 양의 데이터를 실시간으로 처리해야 함으로 고성능 데이터 처리 시스템이 필요하다. 이를 위해 엣지 컴퓨팅과 클라우드 컴퓨팅이 효과적으로 활용되고 있으며, 차량에서 생성되는 데이터를 효율적으로 처리할 수 있도록 고성능 컴퓨팅 플랫폼이 필요하다.

#### c. OTA 업데이트 기능

OTA 업데이트는 SDV의 핵심 기능으로, 소프트웨어의 개선과 새로운 기능 추가를 위해 필요하다. 이를 통해 차량은 운행 중에도 소프트웨어가 자동으로 업데이트될 수 있으며, 차량의 기능적 수명을 연장할 수 있다. OTA는 다양한 부품에 최신 소프트웨어를 전달할 수 있어, 자율주행 기능의 안전성과 신뢰성을 높이는 데 기여한다.

## III. 결 론

소프트웨어 중심 자동차(SDV)는 자율주행 시대에 차량의 혁신을 주도하는 핵심 인프라로 자리 잡고 있으며, 차량 기능을 소프트웨어로 정의하고 지

속적으로 업데이트할 수 있게 한다. SDV는 Android Automotive OS, QNX, Tesla 독자 플랫폼 등 다양한 소프트웨어 플랫폼을 통해 자율주행 기능뿐 아니라 사용자 맞춤형 서비스, 구독형 비즈니스 모델 등을 지원하여 차량의 가치를 높이고 있다. SDV는 자율주행 차량의 완전한 상용화를 위한 핵심 인프라로 자리 잡고 있으며, 미래에는 보다 높은 수준의 자율성이 있을 것으로 기대한다. 즉, 완전 자율주행 (Level 5) 단계에서는 차량의 모든 의사결정이 소프트웨어에 의해 이루어지며, 사람의 개입 없이도 주행이 가능하기 때문에, SDV의 미래 발전은 차량 데이터의 양과 복잡성이 더 증가하는 방향으로 나아가며, 차량 내 고성능 컴퓨팅 시스템의 통합이 가속화될 것이다.

SDV가 완전한 자율주행을 실현하려면 보안성, 고성능 데이터 처리, OTA 업데이트의 안정성 등 수많은 기술적 과제를 해결해야 할 것이다. 사이버 보안 강화, 엣지와 클라우드 컴퓨팅의 효과적 활용, OTA 시스템의 신뢰성 확보가 필수적이다. 인공지능, 클라우드 컴퓨팅, 사물인터넷(IoT) 등 최신 기술이 더 밀접하게 결합될 것이며, SDV는 차량의 기능을 개인화하여 사용자 맞춤형 경험을 제공할 수 있다. 이를 통해 차량의 서비스화, 즉 서비스형 차량 (MaaS, Mobility as a Service) 확산을 촉진할 것이다. 향후 SDV는 단순한 이동 수단을 넘어, 맞춤형 디지털 경험을 제공하는 고도화된 플랫폼으로 발전할 것이며, 자율주행차 시대에 차량 산업의 변화를 주도할 핵심 요소로 기대된다.

### Acknowledgement

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

### 참 고 문 헌

[1] H. Lee, J. Park, "Trends and Challenges in

Software-Defined Vehicles for Autonomous Driving Systems," *Journal of Intelligent and Connected Vehicles*, 6(3), P123-134, 2022.

- [2] Google, "Android Automotive OS Overview," Retrieved from <https://developer.android.com/>, 2021.
- [3] Blackberry Limited, "QNX for Autonomous Vehicles: A Guide to Real-Time Operating Systems in Self-Driving Cars," Waterloo: Blackberry QNX, 2020.
- [4] Tesla Inc., "Tesla's Approach to Full Self-Driving and Software Updates," Retrieved from <https://www.tesla.com/autopilot>, 2021.
- [5] SBD Automotive, "The Evolution of E/E Architecture in Software-Defined Vehicles," *Automotive Engineering Review*, 58(2), P67-81, 2022.
- [6] D. Kim, S. Choi, "Cybersecurity Challenges in Software-Defined Vehicles: An Analysis and Future Directions," *IEEE Transactions on Vehicular Technology*, 72(4), P2056-2067, 2023.
- [7] Fujikimera Research Institute, "Predicted ECU Adoption in SDVs and Future Automotive Trends," Japan: Fujikimera Research Institute, 2022.
- [8] SAE International, "J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," SAE Standard, 2023.
- [9] M. Johnson, K. Park, "High-Performance Computing Architectures for Autonomous Vehicles: A Comprehensive Review," *IEEE Transactions on Intelligent Transportation Systems*, 25(1), pp. 215-232, 2023.
- [10] UNECE, "UN Regulation No. 156 - Software Update and Software Update Management System," United Nations Economic Commission for Europe, 2021.
- [11] ISO/SAE, "ISO/SAE 21434:2021 Road vehicles - Cybersecurity engineering," International Organization for Standardization, 2021.

# SDV 기술개발 이슈에 따른 방안 연구

이도연\*, 김남선\*\*, 신동현\*\*

## A Study on Technical Solutions and Strategies for Software Defined Vehicle Development Issues

Lee Do Yeon\*, Kim Nam Sun\*, and Shin Dong Hyoun\*\*

### 요 약

자동차 산업의 새로운 패러다임으로 부상하고 있는 소프트웨어 정의 차량(Software Defined Vehicle, SDV)의 기술개발 과정에서 발생하는 주요 이슈들을 심층적으로 분석하고 이에 대한 체계적인 해결방안을 제시하였다. SDV는 차량의 모든 기능을 소프트웨어로 정의하고 제어하는 새로운 개념으로, 자동차 산업을 디지털로 바꾸는 핵심 요소가 되어 가고 있다. 본 논문에서는 문헌 연구와 산업 동향 분석을 통해 SDV 기술개발의 주요 이슈를 소프트웨어 아키텍처, 협력체계, 사업모델, 규제대응 측면에서 분석하고, 각 이슈별 해결방안을 도출하였다. 본 논문의 연구 결과는 SDV 기술개발 전략 수립에 실질적인 시사점을 제공할 것으로 기대된다.

### Abstract

This paper deeply analyzes the major issues that arise in the technology development process of Software Defined Vehicle (SDV), which is emerging as a new paradigm in the automobile industry, and presents systematic solutions to them. SDV is a new concept that defines and controls all functions of a vehicle with software, and is becoming a key element in digitalizing the automobile industry. In this paper, through literature research and industry trend analysis, the major issues of SDV technology development are analyzed in terms of software architecture, cooperation system, business model, and regulatory response, and solutions are derived for each issue. The research results of this paper are expected to provide practical implications for establishing SDV technology development strategies.

### Key words

Software Defined Vehicle, Automotive Software Architecture, Connected Car

---

\* 경찰대학 dyl9898@naver.com,

\*\* 경찰대학 nsk@police.go.kr

\*\*\* 연세대학 1001sdh@yonsei.ac.kr

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임.  
(No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

## I. 서 론

자동차 산업은 현재 전례 없는 기술적 변혁기를 맞이하고 있다. 전통적인 하드웨어 중심의 제조 패러다임에서 소프트웨어 중심의 플랫폼 패러다임으로 빠르게 전환되고 있으며, 이러한 변화의 중심에 소프트웨어 정의 차량(SDV)이 있다. SDV는 단순히 소프트웨어가 탑재된 차량을 의미하는 것이 아니라, 차량의 모든 기능과 서비스가 소프트웨어를 통해 정의되고 제어되는 새로운 개념의 자동차를 의미한다. 최근의 연구에 따르면, 2025년까지 신차의 약 70%가 SDV 형태로 출시될 것으로 예측되고 있으며, 이는 자동차 산업이 더 이상 전통적인 제조업의 영역에 머무르지 않고 있다는 것을 의미하고 있다. 즉, 전통 자동차 제조 플랫폼이 소프트웨어 플랫폼 산업으로 진화하고 있음을 보여주고 있으며, 완성차 제조사들은 이러한 변화에 대응하기 위해 대규모 투자 및 새로운 기술개발 체계를 구축하고 있다. 본 논문의 목적은 SDV 기술개발 과정에서 발생하는 주요 이슈들을 체계적으로 분석하고, 이에 대한 실효성 있는 해결방안을 제시하는 것으로, 특히 소프트웨어 아키텍처의 복잡성, 이해관계자 간 협력체계, 새로운 사업모델 발굴, 규제대응 등 핵심 이슈들에 초점을 맞추어 연구를 진행하였다.

## II. 본 론

SDV는 하드웨어 중심의 기존 차량들과 달리, 소프트웨어가 핵심이 되어 차량의 기능을 제어하는 신개념의 자동차를 지칭한다. 전통적인 자동차가 하드웨어 사양에 의해 기능이 고정되는 것과 달리, SDV는 지속적으로 소프트웨어 업데이트를 지원하기 때문에, 차량의 기능을 시간이 지나도 확장할 수 있다. 이는 스마트폰의 운영체제가 지속적인 업데이트를 통해 새로운 기능을 제공하는 것과 유사한 개념이다. SDV의 핵심적인 특징은 다음과 같다. 첫째, 무선 업데이트(OTA)를 이용하여 기능의 지속적인 개선이 가능하다. 둘째, 하드웨어와 소프트웨어의 분리(Decoupling)를 통해 유연한 기능 확장이 가능하다. 셋째, 차량 데이터의 실시간 수집과 분석을

통해 개인화된 서비스 제공이 가능하다. 넷째, 서비스 지향 아키텍처(SOA)를 통해 모듈화된 기능 구현이 가능하다. 또한, SDV를 구현하기 위한 핵심 기술 요소는 크게 네 가지로 구분된다. 첫째, 중앙 컴퓨팅 플랫폼은 고성능 프로세서와 대용량 메모리를 탑재하여 차량의 모든 기능을 통합 제어한다. 둘째, 차량용 운영체제(Vehicle OS)는 다양한 소프트웨어 애플리케이션의 실행 환경을 제공한다. 셋째, 서비스 지향 아키텍처는 차량의 다양한 기능을 서비스 단위로 모듈화하여 관리한다. 넷째, 클라우드 플랫폼은 차량과 외부 시스템 간의 데이터 연동 및 서비스 제공을 지원한다. 이러한 기술 요소들은 상호 유기적으로 연계되어 작동하며, 각 요소의 성능과 안정성이 SDV의 전체적인 성능을 좌우하게 된다. 특히 차량용 운영체제는 실시간성, 안정성, 보안성이 모두 보장되어야 하는 복잡한 요구사항을 충족해야 한다.

### 2.1 SDV 기술개발의 주요 이슈

#### a. 소프트웨어 아키텍처 측면

현대의 자동차에는 평균적으로 1억 줄 이상의 소프트웨어 코드가 탑재되어 있으며, 이는 여객기의 소프트웨어 코드량보다도 많은 수준이다. 이러한 방대한 양의 소프트웨어를 효율적으로 관리하고 업데이트하는 것이 SDV 개발의 핵심 과제이다. 특히 현재의 분산형 ECU 구조에서는 한 가지 기능을 구현하기 위해 여러 ECU가 상호작용해야 하며, 이로 인해 소프트웨어 통합과 업데이트가 매우 복잡해진다. 예를 들어, 차선이탈방지 기능 하나를 업데이트하기 위해서는 카메라 ECU, 조향 ECU, 제동 ECU 등 여러 ECU의 소프트웨어를 동시에 수정해야 한다. 그리고 실시간 제어가 필요한 기능(예: 제동, 조향)과 일반적인 기능(예: 인포테인먼트)이 공존하는 환경에서 시스템의 안정성과 실시간성을 동시에 보장하는 것이 중요한 과제이다. 이는 단순히 기술적인 문제를 넘어서 안전과 직결되는 문제이기에 보다 접근이 신중할 필요가 있다. 소프트웨어 아키텍처의 복잡성을 해결하기 위해서는 우선 중앙집중식 컴퓨

팅 아키텍처로의 전환이 필요하다. 현재의 분산형 ECU 구조에서 도메인 중심의 통합 컨트롤러 구조로 전환함으로써, 소프트웨어 관리와 업데이트의 복잡성을 크게 줄일 수 있다. 서비스 지향 아키텍처(SOA)의 도입도 중요한 해결방안이다. SOA를 통해 차량의 기능들을 독립적인 서비스 모듈로 구현함으로써, 기능의 추가와 수정이 용이해지고 재사용성도 높아진다. 또한 실시간성이 요구되는 기능과 일반 기능을 명확히 분리하여 관리함으로써, 시스템의 안정성과 실시간성을 동시에 확보할 수 있다.

#### b. 이해관계자 협력체계 측면

SDV 개발은 전통적인 자동차 제조사와 부품업체 뿐만 아니라, 소프트웨어 기업, 반도체 기업, 통신 기업 등 다양한 이해관계자들의 협력이 필요한 분야이다. 이러한 다양한 이해관계자들 간의 효과적인 협력체계 구축이 중요한 이슈로 대두되고 있다. 특히 전통적인 자동차 산업의 수직적이고 폐쇄적인 협력구조로는 빠르게 변화하는 소프트웨어 기술 발전에 대응하기 어렵다. 소프트웨어 기업들이 가진 민첩하고 유연한 개발 문화와 자동차 산업의 안정성 중심 문화 간의 조화도 중요한 과제이다. 효과적인 협력체계 구축을 위해서는 개방형 혁신(Open Innovation) 모델의 도입이 필요하다. 자동차 산업 생태계 전반의 참여자들이 상호 연계된 통합 개발 인프라를 구현하고, 표준화된 인터페이스를 정의함으로써 효율적인 협력이 가능하다. 애자일(Agile) 개발방법론의 적용으로 전통적인 워터폴(Waterfall) 방식에서 벗어나, 빠른 프로토타이핑과 지속적인 피드백을 통한 개발 방식을 도입함으로써 개발 효율성을 높일 수도 있다.

#### c. 사업모델 혁신 측면

SDV는 기존 자동차 산업의 사업모델을 근본적으로 변화시키고 있다. 전통적인 일회성 판매 모델에서 벗어나, 소프트웨어 업데이트와 새로운 기능 추가를 통한 지속적인 수익 창출이 가능한 구독 모델로의 전환이 필요하다. 또한, 차량에서 생성되는 데이터를 활용한 새로운 서비스 개발과 수익화 방안도 중요한 과제이다. 새로운 사업모델 구축을 위해

서는 소프트웨어 중심의 가치 창출 체계를 확립을 하고, 기본적인 차량 기능 외에 추가적인 성능 향상과 기존에 없던 기능을 추가하는 등 유료 서비스로 제공하는 방식등을 고려해야 한다. 즉, 차량 데이터 기반의 새로운 서비스 개발과 차량 운행 데이터를 기반으로 한 예측 정비 서비스, 맞춤형 보험 서비스, 차량 관리 서비스 등 다양한 부가가치 서비스를 제공 해야한다.

## 2.2 SDV 기술개발 규제대응

SDV 개발과 상용화를 위해서는 다양한 규제적 이슈들을 해결해야 한다. 특히 소프트웨어 업데이트에 대한 안전성 인증, 차량 데이터의 수집과 활용에 관한 개인정보보호 규제, 자율주행 관련 법제도 등이 중요한 과제로 대두되고 있다. 더불어 국가별로 상이한 규제 환경에 대응하기 위한 전략 수립도 필요하다. 유럽의 GDPR, 미국의 자율주행차 규제, 중국의 데이터 보안법 등 각국의 규제 환경을 고려한 맞춤형 대응을 고려해야 한다.

## III. 결 론

본 논문은 SDV 기술개발 과정에서 나타나는 주요 이슈들을 분석하였고 이에 대한 해결방안을 제시하였다. 소프트웨어 아키텍처의 복잡성, 이해관계자 간 협력체계, 사업모델 혁신, 규제대응 등 다양한 측면에서의 과제들이 존재하며, 이러한 과제들은 상호 연관되어 있어 통합적인 접근이 필요하다. 첫째, SDV 개발을 위해서는 소프트웨어 중심의 개발 체계로의 전환이 필수적이다. 이는 단순한 기술적 변화가 아닌 조직문화, 그리고 현재의 업무방식의 근본적인 변화를 수반한다. 둘째, 다양한 이해관계자들 간의 효과적인 협력체계 구축이 성공의 핵심 요소이다. 특히 자동차 산업과 소프트웨어 산업 간의 문화적 차이를 극복하고 시너지를 창출할 수 있는 협력 모델 구축이 중요하다. 셋째, 소프트웨어 중심의 새로운 사업모델 발굴이 필요하다. 전통적인 하드웨어 판매 중심의 사업모델에서 벗어나, 소프트웨어와 서비스 기반의 지속적인 수익 창출 모델을

구축해야 한다. 넷째, 안전성과 보안성에 대한 체계적인 접근이 필요하다. SDV의 특성상 소프트웨어 결합이 직접적인 안전사고로 이어질 수 있기 때문에, 엄격한 검증체계와 보안대책 수립이 필수적이다. 향후 연구에서는 각 해결방안의 실효성을 검증하기 위한 다양한 이해관계자들이 참여하는, 실제 SDV 개발 프로젝트를 통해 제시된 해결방안들의 현장 적용과 검증을 기대한다.

### Acknowledgement

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

### 참 고 문 헌

- [1] 산업통상자원부, "국내외 커넥티드기반 스타크 및 SDV기술개발 동향과 사업전략," 한국산업기술평가원, P1-158, 2024.
- [2] 한국자동차연구원, "소프트웨어 중심 자동차(SDV) 전모와 향후 전망," 자동차산업동향, Vol. 12, No. 1, P23-45, 2024.
- [3] M. Fern, J. Smith, K. Brown, "The Software-Defined Vehicle Revolution: Industry Analysis and Future Prospects," Morgan Stanley Research Institute, Vol. 8, No. 4, P156-178, 2023.
- [4] R. Anderson, P. Kumar, "Software-Defined Vehicles: The Next Big Thing in Automotive," Boston Consulting Group Technical Report, TR-2023-089, P1-42, 2023.
- [5] S. Johnson, L. Zhang, "Mastering Automotive Software Excellence: A Comprehensive Study," McKinsey Quarterly, Vol. 45, No. 2, P78-95, 2023.
- [6] D. Miller, T. Wilson, "Market Guide for Software-Defined Vehicles," Gartner Research Report, GR-2023-567, P1-67, 2023.
- [7] IEEE Vehicular Technology Society, "Standardization Roadmap for Connected and Automated Vehicles," IEEE Standards Association, Vol. 5, P234-289, 2023.
- [8] SAE International, "Guidelines for Safe OTA Software

Updates in Automotive Applications," SAE Technical Paper Series, 2023-01-0234, P1-28, 2023.

- [9] H. Park, J. Lee, "Global Automotive Software Trends: Executive Survey Analysis," KPMG Research Journal, Vol. 15, No. 3, P45-67, 2023.
- [10] L. Chen, M. Davis, "Future of Automotive Software and Electronics: Technical Assessment," Deloitte Research Series, Vol. 7, P123-156, 2023.



# SDV 기술 전환을 위한 대응 방안에 대한 연구

오수진\*, 윤철희\*\*, 박광민\*\*\*

## A Study on Strategic Approaches for Software-Defined Vehicle Technology Transformation

Cheolhee Yoon\*, Soojin Oh\*\*, and Kwangmin Park\*\*\*

### 요약

SDV(Software-defined Vehicle)가 자동차 산업의 핵심 주제로 등장하면서 그 중요성이 더욱 증대되고 있다. SDV는 향후 모빌리티의 핵심 경쟁력으로서 차량에 탑재되는 소프트웨어의 개발과 통합에 초점을 맞추고 있다. 본 연구에서는 SDV의 발전 단계별 특성을 분석하고, 차량 가치 향상을 위한 소프트웨어 업데이트의 장점과 함께 시스템 복잡성 증가에 따른 과제들을 검토하였다. 이를 통해 기존 시스템과 차세대 시스템 아키텍처의 상호 연동성 확보, 데이터 관리 체계 구축, 보안 강화 등 SDV의 성공적인 기술 전환을 위한 전략적 대응 방안을 제시하였다.

### Abstract

As SDV(Software-defined Vehicle) emerges as a key topic in the automobile industry, its importance is increasing. SDV focuses on the development and integration of software installed in vehicles as a core competitiveness of future mobility. In this study, we analyzed the characteristics of SDV by development stage and reviewed the advantages of software updates for improving vehicle value and the challenges due to increased system complexity. Through this, we presented strategic response measures for successful technology transition of SDV, such as securing interoperability between existing systems and next-generation system architecture, establishing a data management system, and strengthening security.

### Key words

SDV, Software-defined Vehicle, autonomous driving, autonomous vehicle

## 1. 서론

현대 자동차 산업은 급격한 패러다임 전환을 맞

이하고 있다. 특히 첨단 모빌리티 시장에서 소프트웨어가 핵심 경쟁력으로 부상하면서, 이를 기반으로 한 새로운 개념인 SDV(Software-defined Vehicle,

\* 경찰대학 ohsoojin@police.ac.kr,

\*\* 경찰대학 bertter@police.ac.kr,

\*\*\* 연세대학 pkm992025@yonsei.ac.kr

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임.  
(No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

SDV)가 주목받고 있다. SDV는 단순한 운송 수단을 넘어, 소프트웨어를 통해 차량의 기능과 성능을 유연하게 확장하고 개선할 수 있는 진보된 모빌리티 플랫폼이다.

최근의 차량 개발 동향을 살펴보면, 신규 도입되는 차량 기능의 90% 이상이 전기전자 시스템을 기반으로 구현되고 있으며, 이러한 시스템의 핵심은 소프트웨어 기술에 있다. SDV는 이러한 추세를 한층 더 발전시켜, 기존의 하드웨어 중심 차량 구조에서 벗어나 소프트웨어와 컴퓨팅 파워를 중심으로 한 새로운 아키텍처를 제시한다. 이를 통해 차량의 기능과 성능을 실시간으로 제어하고 최적화할 수 있으며, 사용자의 니즈에 따라 지속적인 기능 개선과 확장이 가능하다. SDV는 새로운 기능에 대한 실시간 접속이 가능하며, 차량 수명 전체에 걸쳐 새로운 기능을 실시간으로 추가함으로써 지속적으로 차량을 '업그레이드'할 수 있다. 그리고, 차량에서 수집한 대량의 데이터는 SDV가 제공하는 인공지능에 도입되어, 차내 안전성이 강화되는데 이는 특정 부품의 마모 및 고장을 실시간으로 찾아내어 수리의 필요성을 예측하거나 고장을 직접 수정할 수 있는 기능으로 개선되었다. 그리고, SDV 기술을 사용하여 클라우드를 통해 자동차를 업그레이드하면 차량 자체의 가치가 향상되는데, 이유는, 시간이 지날수록 새로운 기능이 추가된 차량에 비해 부족한 기능을 보완하여 차량의 가치 저하를 방지할 수 있기 때문이다. 이처럼 SDV는 자동차 업계의 핵심 주제로 등장했으며 자동차를 제어하고 통제하는 소프트웨어의 중요성은 더욱 높아질 것이 분명하다.

## II. 본 론

### 가. SDV 개요

SDV는 다양한 측면에서 정의될 수 있다. 차량의 주요 기능과 특성이 소프트웨어에 의해 결정되는 차세대 자동차, 소프트웨어를 기반으로 지속적인 발전이 가능한 스마트 모빌리티, 하드웨어 인프라를 소프트웨어 중심으로 제어하는 지능형 운송 수단 등이 그 예이다. 이러한 SDV의 핵심은 전통적인 자동차 구조에 고도화된 소프트웨어 시스템과 컴퓨팅

기술을 접목한 것으로, 차량의 핵심 요소들을 소프트웨어로 통합 관리한다. 이를 통해 차량의 핵심 성능, 사용자 편의성, 안전 기능, 감성적 요소에 이르기까지 다양한 측면을 실시간으로 최적화하고 업데이트를 시킬 수 있다는 것이 SDV의 차별화된 특징이다. SDV는 내비게이션과 접속성의 강화, 실시간 애플리케이션 업데이트, 안전성과 지속가능성 강화 등 자동차가 그 수명 안에서 진화할 수 있는 가능성을 제공하며, 운전 체험을 개선하기 위해 이루어진다. 앞으로도 자동차는 점점 더 똑똑해지고 더 많이 연결될 것이므로, 소프트웨어는 그 설계와 진화 모두에 있어 더욱 중요한 역할을 하게 된다. SDV가 부상한 배경은 자동차의 가치 창출이 기존의 가치인 하드웨어에서 소프트웨어라는 새로운 가치로 이동하고 있기 때문이다. 자동차 산업의 새로운 변화를 이끄는 CASE(Connected, Autonomous, Shared, Electric) 트렌드는 기존 자동차 산업의 근본적인 혁신을 촉진하고 있다. SDV의 등장 배경에는 현대 소비자들의 디지털 경험에 대한 기대가 크게 작용했다. 소비자들은 일상적으로 사용하는 스마트폰의 직관적인 인터페이스와 다양한 기능들을 차량에서도 경험하기를 원했으나, 기존의 차량 시스템은 이러한 기대를 충족시키지 못했다. 이러한 시장의 요구를 해결하기 위해 완전한 소프트웨어 프로그래밍이 가능한 SDV가 개발되었다.

SDV 기술은 크게 네 단계의 발전 과정을 거치며 진화하고 있다. 첫 번째 연결(Connected) 단계에서는 기본적인 네트워크 연결성을 바탕으로 주차 보조 등 단순화된 안전 기능과 제한된 소프트웨어 기능이 구현되었다. 두 번째 적응(Adaptive) 단계는 OTA(Over-The-Air) 기반의 소프트웨어 업데이트가 가능해지는 단계로, 현재 많은 기업들이 독자적인 구현 방식을 채택하면서 시스템 복잡도가 증가하는 특징을 보이고 있다. 세 번째 역동(Dynamic) 단계는 개방형 소프트웨어 플랫폼과 차세대 전자제어 아키텍처를 기반으로 혁신적인 개발 환경이 구축되는 시기이다. 마지막 몰입(Immersive) 단계에서는 완전 자율주행 기술과 AI, 가상현실 기술이 통합되어 새로운 모빌리티 경험을 제공하고 다양한 수익 모델이 창출될 것으로 예상된다.

표 1. SDV 레벨 분류

Table 1. SDV Level Classification

구분	0	베이직	모던	어드벤스드	최첨단	미래지향
레벨	0	1	2	3	4	5
연결	없음	3G/4G	4G	4G	5G	5G/6G
디스플레이	Non-touch	작은스크린 (논터치 or 터치)	중간사이즈 터치스크린	대형 터치스크린	멀티대형 터치스크린	모든 승객에 터치스크린 제공
자율주행	Level 0	Level 1	Level 1/Level 2	Level 2	Level 3	Level 3/4+

이러한 발전 단계는 차량에 탑재되는 소프트웨어 기능의 수와 복잡도에 따라 Level 0부터 5까지의 세부 단계로도 구분된다. 각 단계의 진화는 자동차 제조사들의 전략적 판단과 기술 구현 능력에 따라 다양한 형태로 전개되고 있으며, 이는 SDV 생태계에서 각 기업의 경쟁력을 결정하게 되는 중요한 점이 되고 있다.[표 1]. 일반적으로 SDV 레벨 분류는 지속적인 셀룰러 연결(5G/6G), 크고 터치가 가능한 화면, 차량의 구성요소에 연결되는 강력한 중앙 컴퓨팅 시스템 또한 많은 SDV가 다양한 외부 서비스 플랫폼과의 연계 및 차량 내 결제 인프라 도입을 통한 사용성 향상하는지 여부 등에 따라 더욱 진화된 SDV 형태로 분류한다.

#### 나. SDV 전환을 위한 대응 방안

SDV는 소프트웨어를 주기적으로 업데이트하기 때문에 차량의 가치를 높일 수 있다는 장점이 있는 동시에 복잡성 증가에 따른 각종 비용이 발생할 수 있다. 따라서, 새로운 물리적, 논리적 시스템 아키텍처 구현과 잘 설계된 개념이 SDV의 성공적인 확장 운영을 좌우하는 핵심 요소가 될 것이다. 개발 부서 간 소프트웨어 호환성을 높일 필요가 있으며, 이전 시스템과 차세대 시스템 아키텍처의 상호 연동이 필요하다. 또한, SDV OS는 데이터 관리와 보안에 핵심적 역할을 하는데, OEM은 OS를 통제해 SDV 차량에서 생성된 데이터를 보호하고 데이터 프라이버시 관련 정책을 준수할 수 있다. 자동차 산업에서 데이터의 역할이 커지는 가운데, OS를 통제할 데이터 소유권을 보호와 보안 강화는 더욱 중요하다. 또한, 데이터 중심 커넥티드 서비스 개발이 필요하다. 차량 성능 보정, 운전자 지원 시스템 추가, 보안

기능 업데이트 및 개인화된 앱 설치와 같은 서비스들은 곧 OTA 방식으로 출시될 업데이트의 변화로 SDV에는 스마트폰처럼 개개인의 데이터가 무수히 축적될 것으로 예상된다. 자율주행 기술과 첨단 운전자 지원 시스템(ADAS)의 고도화를 위해서는 방대한 양의 데이터 확보와 체계적인 관리 체계 구축이 필수적인 상황이다. 자동차 제조사들은 실시간 데이터의 효과적인 수집과 분석 역량을 바탕으로 급성장하는 모빌리티 데이터 시장에서 우위를 차지할 수 있을 것으로 보인다.

아울러 차세대 구동 시스템 개발도 중요한 과제로 대두되고 있다. 자동차 제조사들은 포괄적인 파워트레인 혁신 전략을 수립해야 하는데, 이는 AI 기술을 활용한 배터리 성능 예측과 클라우드 기반 제어 시스템을 갖춘 전기 구동계 개발부터, 지속 가능한 에너지 공급 및 충전 시설 네트워크 구축까지 광범위한 기술 영역을 포함해야 한다.

### III. 결 론

자동차 산업의 SDV 전환은 기존의 '하드웨어 중심 설계에 소프트웨어를 적용하는 방식'에서 '소프트웨어 아키텍처를 중심으로 차량을 구현하는 방식'으로의 근본적인 패러다임 전환을 의미한다. 이러한 변화의 핵심적인 특징은 차량 운행 과정에서 생성되는 방대한 데이터의 활용에 있다. 이를 통해 단순한 차량 제어를 넘어 운전자의 행동 패턴과 선호도를 심층적으로 이해하고 분석할 수 있게 되었다.

SDV의 성공적인 구현을 위해서는 첨단 소프트웨어 개발 역량과 통합 컴퓨팅 플랫폼이 필수적이며, 여기에 클라우드 컴퓨팅, 빅데이터 분석, 인공지능 등 최첨단 기술의 융합이 요구된다. 이러한 SDV로

의 기술 혁신은 자동차 산업의 디지털 전환을 가속화할 뿐만 아니라, 혁신적이고 사용자 중심적인 새로운 모빌리티 서비스 창출로 이어질 것으로 전망된다.

### Acknowledgement

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

### 참 고 문 헌

- [1] S. Kim, J. Lee, "Software-Defined Vehicle Architecture and Implementation Strategies," IEEE Transactions on Vehicular Technology, Vol. 72, No. 3, P124-139, 2023.
- [2] W. Zhang, H. Chen, M. Liu, "Current Trends in Software-Defined Vehicle Development," Automotive Innovation Journal, Vol. 7, No. 1, P15-32, 2024.
- [3] M. Fern, J. Smith, K. Brown, "The Software-Defined Vehicle Revolution: Industry Analysis and Future Prospects," Morgan Stanley Research Institute, Vol. 8, No. 4, P156-178, 2023.
- [4] R. Anderson, P. Kumar, "Software-Defined Vehicles: The Next Big Thing in Automotive," Boston Consulting Group Technical Report, TR-2023-089, P1-42, 2023.
- [5] L. Wang, P. Chen, "Cloud Infrastructure Design for Software-Defined Vehicles," IEEE Access Technical Journal, Vol. 12, No. 1, P15678-15692, 2024.
- [6] Y. Kim, S. Park, "Development Trends and Future Prospects of Software-Defined Vehicles," Journal of Korean Institute of Electronic Engineering, Vol. 60, No. 4, P56-68, 2023.
- [7] M. Anderson, S. Wilson, "Security Challenges in Software-Defined Vehicle Updates," Journal of Cybersecurity, Vol. 12, No. 3, P178-195, 2023.
- [8] Y. Liu, X. Chen, "Integration Technologies for Autonomous Software-Defined Vehicles," Transportation Research Journal, Vol. 148, No. 2, P103-118, 2024.
- [9] C. Lee, D. Kang, "Digital Transformation in Automotive Industry: SDV-Centered Innovation," Journal of Korean Automotive Engineering, Vol. 31, No. 3, P89-98, 2023.
- [10] S. Martinez, D. Brown, "Electric Powertrain Systems in Software-Defined Vehicles," Energy and Transportation Systems, Vol. 15, No. 2, P234-249, 2024.

# 자율주행 차량의 SDV 전환을 위한 시스템 아키텍처 대응 방안

조민제\*, 강장묵\*\*

## A Study on System Architecture Response Plans for SDV Transition of Autonomous Vehicles

Cho MIn Je\*, Kang Jang Mook\*\*

### 요 약

본 연구는 자율주행 차량의 Software Defined Vehicle(SDV) 전환에 따른 시스템 아키텍처 대응 방안을 제시한다. 자동차 산업이 소프트웨어 중심으로 재편됨에 따라, 차량의 기능이 하드웨어에서 소프트웨어로 이동하는 변화가 가속화되고 있다. 특히 자율주행 기술의 발전은 차량 아키텍처의 근본적인 변화를 요구하고 있다. 본 연구에서는 SDV 전환을 위한 시스템 아키텍처의 발전 단계를 분석하고, 핵심 요소 기술과 과제를 도출하였다. 연구 결과, SDV 시스템 아키텍처는 분산 ECU 아키텍처에서 도메인 제어 아키텍처를 거쳐 중앙집중식 아키텍처로 발전해야 함을 확인하였다. 이를 위해 차량용 OS, E/E 아키텍처, 소프트웨어 플랫폼의 통합적 발전이 필요하며, 보안성과 실시간성이 보장되어야 한다. 본 연구는 자율주행 차량의 SDV 전환을 위한 체계적인 아키텍처 프레임워크를 제시함으로써, 향후 자동차 산업의 소프트웨어 전환에 기여할 것으로 기대된다.

### Abstract

This study presents system architecture response plans for the Software Defined Vehicle (SDV) transition of autonomous vehicles. As the automotive industry is being reorganized around software, the shift of vehicle functionality from hardware to software is accelerating. In this study, we analyzed the development stages of system architecture for SDV transition and derived core technical elements and challenges. The research findings confirmed that SDV system architecture should evolve from distributed ECU architecture through domain control architecture to centralized architecture. This requires integrated development of vehicle OS, E/E architecture, and software platforms, while ensuring security and real-time performance. This study also suggests that systematic approaches to cyber security and functional safety are essential for successful SDV implementation.

### Key words

Software Defined Vehicle (SDV), Autonomous Vehicle, System Architecture, Vehicle Operating System

---

\* 경찰대학교, [epiwolrd@police.go.kr](mailto:epiwolrd@police.go.kr)

\*\* 동국대학교, [honukang@gmail.com](mailto:honukang@gmail.com)

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임.  
(No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

## I. 서 론

자동차 산업은 CASE(Connected, Autonomous, Shared, Electric)로 대변되는 급격한 패러다임 변화를 겪고 있다. 특히 자율주행 기술의 발전과 동시에 차량의 소프트웨어화가 가속화되면서, Software Defined Vehicle(SDV)이 미래 자동차의 핵심 화두로 부상하고 있다. SDV는 차량의 기능을 소프트웨어로 정의하고 제어하는 개념으로, 자동차의 성능과 기능이 소프트웨어에 의해 결정되는 것을 의미한다.

현재 자동차 업계는 전례 없는 기술적 변혁기를 맞이하고 있다. 2020년을 기점으로 전기차 시장이 본격적으로 성장하기 시작했으며, 자율주행 기술은 레벨 3을 뛰어넘고 레벨 4를 향해 빠르게 진화하고 있다. 이러한 변화의 중심에는 소프트웨어가 있다. 2010년 자동차 한 대당 약 300개의 반도체가 사용되었던 것에 비해, 레벨 3 이상의 자율주행차량에서는 약 2,000개의 반도체가 요구될 것으로 예측된다. 이는 단순한 부품 수의 증가가 아닌, 차량의 패러다임이 하드웨어 중심이었던 것에서 소프트웨어 중심으로 바뀌고 있다는 것을 의미한다. 특히 주목할 점은 차량용 소프트웨어의 복잡도와 규모가 기하급수적으로 증가하고 있다는 것이다. 현대의 프리미엄 차량에는 1억 라인 이상의 소프트웨어 코드가 탑재되어 있으며, 이는 여객기의 소프트웨어 코드량을 훨씬 상회하는 수준이다. 더욱이 자율주행 기술이 고도화될수록 인공지능, 센서 퓨전, 고정밀 맵핑 등 복잡한 소프트웨어의 도입이 필수적이다. 이러한 변화는 기존의 분산형 ECU 기반 아키텍처로는 더 이상 효과적으로 관리할 수 없는 상황을 초래하고 있다. 자동차의 SDV 전환은 단순한 기술적 진화를 넘어 비즈니스 모델의 혁신도 동반한다. 테슬라의 경우 2027년까지 차량 사용자당 월평균 100달러의 소프트웨어 관련 수익을 창출할 것으로 예상되며, 스텔란티스, 포드, GM 등 전통적인 자동차 제조사들도 2030년까지 소프트웨어 서비스에서 각각 200억 달러 이상의 수익 창출을 목표로 하고 있다. 이는 자동차가 더 이상 단순한 이동 수단이 아닌, 소프트웨어 플랫폼으로 진화하고 있음을 보여준다. 그러나 SDV로의 전환 과정에서 해결해야 할 과제도 많다.

실시간성과 안전성의 보장, 사이버 보안 위협 대응, 소프트웨어 플랫폼의 표준화, 개발 인력 확보 등 다양한 도전 과제가 존재한다. 특히 자율주행 차량의 경우, 소프트웨어 결함이 직접적인 안전 위협으로 이어질 수 있기 때문에, 더욱 체계적이고 신중한 접근이 필요하다.

이러한 배경에서 본 연구는 자율주행 차량의 SDV 전환을 위한 시스템 아키텍처 대응 방안을 제시하고자 한다. 특히 현재의 분산형 아키텍처에서 중앙집중식 아키텍처로의 전환 과정에서 필요한 핵심 기술 요소와 구현 전략을 분석하고, 실시간성, 안전성, 보안성을 확보하기 위한 구체적인 방안을 도출하고자 한다. 이를 통해 자동차 산업의 SDV 전환을 위한 체계적인 로드맵을 제시하고, 향후 연구 방향을 제안하고자 한다.

## II. 본 론

### 2.1 SDV 시스템 아키텍처 현황과 과제

자동차 산업에서 SDV로의 전환은 시스템 아키텍처의 근본적인 재구성을 요구한다. 현재의 자동차 아키텍처는 6080개의 독립적인 ECU가 각각의 기능을 담당하는 분산형 구조를 가지고 있으며, 이는 5,000만8,000만 라인의 소프트웨어 코드로 구성되어 있다. 이러한 분산형 아키텍처는 새로운 기능 추가 시 ECU를 증설해야 하는 구조적 한계를 가지고 있으며, 소프트웨어 업데이트와 유지보수가 복잡하다는 문제점을 안고 있다. 특히 자율주행 기술의 발전에 따라 차량의 컴퓨팅 요구사항은 급격히 증가하고 있다. 레벨 3 이상의 자율주행에서는 카메라, 라이다, 레이더 등 다양한 센서의 데이터를 실시간으로 처리해야 하며, 이는 초당 테라바이트 규모의 데이터 처리를 필요로 한다. 기존의 분산형 ECU 구조로는 이러한 대용량 데이터의 실시간 처리가 어려우며, ECU 간 통신 지연으로 인한 성능 저하 문제도 발생한다. 이러한 한계를 극복하기 위해 도메인 제어 아키텍처가 도입되고 있다. 도메인 제어 아키텍처는 유사한 기능들을 도메인 단위로 통합하여 관리하는 방식으로, ADAS/자율주행, 파워트레인, 차

체, 인포테인먼트 등의 주요 도메인별로 통합 제어기를 둔다. 각 도메인 컨트롤러는 고성능 프로세서를 탑재하고 있으며, 도메인 간 통신은 기가비트 이더넷을 통해 이루어진다. 이를 통해 시스템의 복잡도를 낮추고 소프트웨어 업데이트의 효율성을 높일 수 있다.

## 2.2 SDV 전환을 위한 핵심 기술 요소와 구현 전략

SDV 구현의 핵심 기술 요소는 크게 세 가지 축으로 구성된다. 첫째, 차량용 운영체제는 실시간성과 안정성을 보장하면서도 다양한 애플리케이션의 실행을 지원해야 한다. 현재 시장에서는 QNX, Android Automotive OS, AGL(Automotive Grade Linux) 등이 경쟁하고 있다. QNX는 높은 신뢰성과 실시간성으로 ADAS/자율주행 시스템에서 주로 사용되며, Android Automotive OS는 풍부한 생태계를 바탕으로 인포테인먼트 시스템에서 강세를 보이고 있다. 차량용 운영체제의 발전 방향은 크게 두 가지로 나눌 수 있다. 하나는 실시간성과 안전성이 중요한 도메인에서 사용되는 RTOS(Real-Time Operating System)이며, 다른 하나는 사용자 경험과 서비스 확장성이 중요한 도메인에서 사용되는 GPOS(General Purpose Operating System)이다. 특히 AUTOSAR Adaptive Platform은 이 두 가지 요구사항을 모두 만족시키기 위한 표준화된 소프트웨어 플랫폼을 제공한다. 둘째, E/E(Electric/Electronic) 아키텍처의 혁신이 필요하다. 현재의 E/E 아키텍처는 도메인 제어 아키텍처를 거쳐 중앙집중식 컴퓨팅 아키텍처로 진화하고 있다. 중앙집중식 컴퓨팅 아키텍처에서는 고성능 중앙처리장치가 차량의 모든 기능을 통합 제어하며, 실시간성이 요구되는 기능은 전용 프로세서를 통해 처리된다. 특히 TSN(Time Sensitive Networking) 기술이 적용된 차량용 이더넷의 도입은 실시간 데이터 전송을 가능하게 하는 핵심 요소이다. 셋째, 서비스 지향 아키텍처(SOA)를 기반으로 한 소프트웨어 플랫폼의 구축이다. SOA는 차량의 다양한 기능을 독립적인 서비스 단위로 모듈화하여 제공함으로써, 유연한 기능 확장과 업데이트를 가능

하게 한다. 특히 마이크로서비스 아키텍처의 도입은 서비스 간 독립성을 보장하고, API를 통한 표준화된 인터페이스를 제공함으로써 소프트웨어의 재사용성과 유지보수성을 높인다.

## 2.3 보안 및 안전성 확보를 위한 아키텍처 설계방안

SDV 환경에서 보안과 안전성은 가장 중요한 요구사항이다. 특히 자율주행 차량의 경우, 사이버 보안 위협이 직접적인 안전 위협으로 이어질 수 있다. UN Regulation No. 155의 발효로 사이버 보안 관리 시스템(CSMS)의 구축이 의무화되었으며, ISO/SAE 21434는 자동차 사이버 보안 개발 프로세스에 대한 국제 표준을 제시하고 있다.

보안 아키텍처는 다층 방어(Defense-in-Depth) 전략을 기반으로 설계되어야 한다. 하드웨어 수준에서는 HSM(Hardware Security Module)을 통한 암호화 키 관리와 보안 부팅이 필요하며, 네트워크 수준에서는 도메인 간 통신을 제어하는 중앙 게이트웨이와 네트워크 세그멘테이션이 필요하다. 소프트웨어 수준에서는 코드 서명, 런타임 무결성 검증, 안전한 업데이트 메커니즘이 구현되어야 한다. OTA(Over-The-Air) 업데이트 시스템은 특별한 주의가 필요하다. 업데이트 패키지의 무결성 검증, 롤백 메커니즘, 단계적 배포 등의 전략이 필요하며, 업데이트 실패 시에도 차량의 안전한 동작을 보장할 수 있어야 한다. 또한 업데이트 프로세스 자체가 새로운 공격 벡터가 되지 않도록 엄격한 보안 통제가 필요하다.

## III. 결 론

본 연구는 자율주행 차량의 SDV 전환을 위한 시스템 아키텍처 대응 방안을 분석하고 제시하였다. 자동차 산업이 소프트웨어 중심으로 재편되는 가운데, 시스템 아키텍처의 혁신은 성공적인 SDV 전환을 위한 핵심 요소임을 확인하였다. 특히 현재의 분산형 ECU 아키텍처에서 도메인 제어 아키텍처를

거쳐 중앙집중식 컴퓨팅 아키텍처로의 진화는 필연적인 흐름이며, 이 과정에서 실시간성, 안전성, 보안성의 확보가 중요한 과제임을 도출하였다. 연구 결과, SDV 시스템 아키텍처의 성공적인 구현을 위해서는 다음과 같은 요소들이 필수적임을 확인하였다. 첫째, 차량용 운영체제는 실시간성과 안전성을 보장하면서도 풍부한 서비스 생태계를 지원할 수 있어야 한다. AUTOSAR Adaptive Platform과 같은 표준화된 소프트웨어 플랫폼의 도입은 이러한 요구사항을 충족시키는 핵심 전략이 될 수 있다. 둘째, E/E 아키텍처는 고성능 중앙처리장치와 실시간 네트워크를 기반으로 재구성되어야 하며, 특히 TSN 기술의 도입은 실시간 데이터 전송을 위한 필수 요소이다. 셋째, 서비스 지향 아키텍처와 마이크로서비스 설계 패턴의 도입을 통해 소프트웨어의 모듈성과 유연성을 확보해야 한다. 또한, 사이버 보안과 기능 안전성의 통합적 접근이 필요함을 확인하였다. UN Regulation No. 155와 ISO 26262의 요구사항을 만족시키면서도, 효율적인 소프트웨어 업데이트와 기능 확장이 가능한 아키텍처의 설계가 중요하다. 특히 OTA 업데이트 시스템의 보안성 확보와 함께, 하드웨어 보안 모듈, 보안 부팅, 런타임 무결성 검증 등 다층적 보안 메커니즘의 구현이 필수적이다. 향후 연구 과제로는 다음과 같은 사항들이 고려되어야 한다. 첫째, 인공지능 기술의 발전에 따른 시스템 아키텍처의 최적화 방안 연구가 필요하다. 자율주행의 고도화에 따라 AI 연산량이 급증할 것으로 예상되며, 이를 효율적으로 처리할 수 있는 아키텍처의 연구가 요구된다. 둘째, 클라우드-엣지-차량 간의 효율적인 컴퓨팅 분산 전략에 대한 연구가 필요하다. 제한된 차량 내 컴퓨팅 자원을 효율적으로 활용하면서도 클라우드의 장점을 최대한 활용할 수 있는 방안이 모색되어야 한다. 셋째, 다양한 도메인 간의 효율적인 데이터 공유와 서비스 통합을 위한 표준화된 인터페이스의 연구가 필요하다.

SDV로의 전환은 자동차 산업의 패러다임을 근본적으로 변화시키고 있다. 본 연구에서 제시한 시스템 아키텍처 대응 방안은 이러한 변화에 대응하기 위한 기술적 프레임워크를 제공한다. 향후 자율주행 기술의 발전과 함께 SDV의 중요성은 더욱 커질 것

으로 예상되며, 이에 따른 지속적인 연구와 기술 혁신이 필요할 것이다. 특히 산업계와 학계의 긴밀한 협력을 통해 표준화된 아키텍처의 개발과 검증이 이루어져야 하며, 이는 자동차 산업의 미래 경쟁력 확보를 위한 핵심 과제가 될 것이다.

## 참 고 문 헌

- [1] 산업통상자원부, "소프트웨어 중심 자동차(SDV) 전모와 향후 전망", 산업통상자원부 정책보고서, P1-156, 202
- [2] J.H. Kim, S.M. Park, C.W. Lee, "A Study on System Architecture for Software Defined Vehicles", Journal of Automotive Engineering, 45(2), P123-134, 2023
- [3] BCG, "The Software-Defined Vehicle Revolution: How to Prepare for the Next Evolution of Automotive Architecture", BCG Research Report, P1-48, 2023
- [4] S.W. Lee, J.H. Choi, "Development of Security Architecture for Connected and Autonomous Vehicles", IEEE Transactions on Vehicular Technology, 72(3), P2876-2889, 2023
- [5] Markets&Markets, "Automotive Operating Systems Market - Global Forecast to 2030", Market Research Report, 2023
- [6] Y.H. Park, D.W. Kim, "Implementation of Real-time Operating System for Autonomous Vehicles", International Journal of Automotive Technology, 24(4), P567-578, 2023
- [7] H.S. Jung, K.T. Min, "A Comprehensive Study on E/E Architecture Evolution for Software Defined Vehicles", SAE Technical Paper, 2023-01-0234, 2023
- [8] S.J. Kang, M.H. Lee, "Analysis of Automotive Grade Linux and AUTOSAR Adaptive Platform for SDV", IEEE Access, Vol. 11, P45678-45690, 2023
- [9] T.H. Kim, J.W. Lee, "Cyber Security Management System for Software Defined Vehicles: Challenges



and Solutions", Information Security Journal, 33(2),  
P234-245, 202

- [10] W.S. Cho, B.H. Kim, "Development of Domain Control Architecture for Next Generation Vehicles", International Journal of Automotive Engineering, 14(3), P345-356, 2023
- [11] Deutsche Bank Research, "Software-defined Vehicles: The Next Big Thing in Automotive", Market Analysis Report, P1-35, 2024
- [12] ISO 26262, "Road vehicles - Functional safety", International Organization for Standardization, 2018

# 생성형 AI를 활용한 교통사고 조사에 관한 연구

나승균\*, 윤철희\*\*, 이봉규\*\*\*

## Traffic Accident Investigation Using Generative AI: Enhancing Accuracy and Efficiency in Accident Analysis

Ra Seung Kyun\*, Cheol Hee Yoon\*\*, and Lee Bong Gyou\*\*\*

### 요약

교통사고는 전 세계적인 문제로, 매년 135만 명 이상의 사망자가 발생하고 있다(WHO, 2018). 목격자 진술, 현장 재구성, 수동 데이터 분석에 의존하는 전통적인 사고 조사 방식은 시간이 많이 소요되고, 비용이 크며, 인적 오류가 발생하기 쉽다. 본 연구는 교통사고 조사의 정확성과 효율성을 향상시키기 위해 생성형 AI(GAN)의 적용 방안을 탐구한다. 대규모 데이터셋을 분석하고, 실제와 유사한 사고 시나리오를 생성하며, 정밀한 사고 현장 재구성을 통해 AI가 어떻게 사고 분석을 혁신할 수 있는지 보여준다. 이를 통해 보다 효과적인 예방 및 대응 전략을 수립하여 교통사고로 인한 사망자와 경제적 손실을 줄이는 것이 본 연구의 목표이다.

### Abstract

Traffic accidents are a global issue, resulting in over 1.35 million deaths annually (WHO, 2018). Traditional methods of accident investigation, relying on eyewitness accounts, scene reconstruction, and manual data analysis, are time-consuming, costly, and prone to human error. This study explores the application of Generative AI (GAN) to improve the accuracy and efficiency of traffic accident investigations. By analyzing large datasets, generating realistic accident scenarios, and reconstructing accident scenes with precision, the study demonstrates how AI can revolutionize accident analysis. The goal is to reduce fatalities and economic losses from traffic accidents through more effective prevention and response strategies.

### Key words

Traffic accidents, Generative AI (GAN)

---

\* 연세대학, rokmc724@hanmail.net

\*\* 경찰대학교, bertter@police.ac.kr

\*\*\* 연세대학, bglee@yonse.ac.kr

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임.  
(No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

## I. 서론

교통사고로 인한 피해를 줄이기 위해 다양한 사고 조사 기법이 활용되고 있다. 주로 사고 현장의 물리적 증거를 기반 차량의 블랙박스, 목격자의 진술, 사고 현장 사진, 영상 자료를 수집하여 사고의 원인을 분석하고 책임 소재를 판단하는 방법을 선택하는데, 시스템 되지 않은 상황에서의 조사 및 주관적 판단에 의해 결과의 정확성이 저하될 소지가 있다. 최근 인공지능(AI) 기술의 발전으로, 교통사고 조사 및 예방 전략에 대한 새로운 접근법이 주목받고 있으며, 특히 생성형 적대 신경망(GAN: Generative Adversarial Networks) 기법으로 교통사고 데이터를 학습하면, 과거 데이터를 기반으로 사고 시나리오를 예측하고, 다양한 조건에서 사고 분석을 할 수가 있다.

## II. 관련 연구

블랙박스 영상에 기반한 AI 사고 분석 기술은 사고 상황을 빠르고 객관적으로 판단하는 방법으로 주목된다. AI는 3D 합성곱 신경망(Convolutional Neural Network, CNN)을 활용하여 차선, 차량의 위치, 시간 흐름 등을 동시에 분석하고, 교통사고 과실 비율을 더 공정하고 정확하게 평가할 수 있다. 블랙박스 영상을 AI로 분석하여 주관성을 배제하고 차도, 차선, 차량 등 사고 상황을 공정하게 판단할 수 있는 기술을 개발하였다. 이를 통해 기존의 수작업 평가 대비 사고 조사에 필요한 시간과 인력을 줄일 수 있는 가능성을 제시하였다 (김동진, 2023). 또한, 교차로와 횡단보도 인근에서 발생하는 사고를 줄이기 위해 AI와 엣지 컴퓨팅을 결합한 교통 사고 예방 시스템이 도입하였다. 교차로 근방에서 속도를 조절하고 위험 요소를 실시간으로 모니터링함으로써 교통사고 발생 가능성을 예측하고 경고 메시지를 제공한다. 부산시에서 AI 기반 돌발 상황 탐지 솔루션을 도입하여 사고 예방을 위한 실시간 도로 위험 요소 감지를 시행하고 있다. 이 시스템은 엣지 컴퓨팅을 이용해 교차로 근방에서 속도 조절과 실시간 경고 기능을 제공함으로써 사고 가능성을 줄

이는 데 기여한다 (Jang et al., 2024).“ 그리고 기존 확률 통계 분석에 비해 딥러닝을 도입함으로써 예측 오차를 평균 4km/h 내외로 줄이는 성과를 보였고, 이 기술은 실제 교통상황 관리에 유용하게 활용될 수 있음을 알리고 있다.(Park et al., 2022).

## III. 본론

### 가. 데이터 수집과 전처리

공공 데이터베이스, 교통 감시 시스템, 기상 데이터를 사용하여, 교통사고 발생 시간, 위치, 사고 유형, 피해 규모 등의 데이터를 확인할 수 있다. 그리고 교통 감시 시스템에서는 CCTV, 차량 블랙박스, 도로 센서 등 다양한 교통 감시 장치에서 데이터를 수집이 가능하며 기상 데이터는 기상청 데이터를 통해 사고 당시의 날씨 조건(비, 눈, 안개 등)을 수집이 가능하다. 수집된 데이터는 누락된 값, 중복 데이터, 오류 데이터를 제거를 통해 특징 추출, 사고 유형, 시간대, 차량 속도, 도로 상태 등 주요 변수를 식별하고 분석에 필요한 형태로 변환이 가능하다.

### 나. GAN 적용

GAN 모델은 생성자(Generator)와 판별자(Discriminator)로 구성되는데, 생성자는 실제 교통사고 데이터를 기반으로 새로운 사고 시나리오를 생성하고, 판별자는 생성된 시나리오가 실제 사고 데이터와 얼마나 유사한지를 평가하게 된다. 이 두 네트워크는 상호 경쟁하면서 학습을 통해 점점 더 현실적인 데이터를 생성하게 된다. 모델 성능을 평가하기 위해 정확도(Accuracy), 정밀도(Precision), 재현율(Recall) 등의 지표를 사용하는데, 생성된 시나리오가 실제 사고와 얼마나 유사한지를 평가가 가능하며, 교차 검증 기법을 활용하여 모델이 과적합되지 않도록 고려해야 한다. 이후 학습 데이터와 테스트 데이터를 분리하여 모델의 일반화 성능을 평가한다.

표 1. GAN 모델 구조  
Table 1. GAN model structure

구성 요소	설명
생성자 (Generator)	실제 교통사고 데이터를 학습하여 새로운 사고 시나리오를 생성하는 역할을 수행
판별자 (Discriminator)	생성된 시나리오가 실제 사고와 얼마나 유사한지 평가하여, 점점 더 현실적인 데이터를 생성하게 하는 역할

#### 다. 사고 시나리오 생성

GAN 모델을 사용하여 사고 시나리오를 생성한다. 생성된 시나리오는 다양한 조건(예: 기상 상태, 도로 상황, 차량 속도 등)을 반영이 되며, 이러한 시나리오를 통해 사고 재구성이 가능하다. 사고 현장과 차량 경로를 CCTV 영상, 차량 블랙박스, 센서 데이터를 통해 조합하여 사고 당시 차량의 움직임과 사고 직전의 상황을 복원이 가능하다. GAN 모델을 통해 생성된 사고 시나리오는 실제 사고와 매우 유사한 형태로 나타났으며, 악천후 시나리오에서는 비가 내리는 상황에서 브레이크 미작동으로 인해 발생한 다중 충돌 사고를 재현할 수 있고, 이를 통해 운전자 행동 및 도로 안전 개선 방안을 고려할 수 있다. 또한, 고위험 지역에서의 사고 발생 시나리오를 시뮬레이션함으로써, 해당 지역에서의 교통 안전 개선 필요성을 제안할 수 있다.

#### 라. 사고 시나리오 검증

생성된 사고 시나리오의 신뢰성과 정확성을 확보하기 위해 체계적인 다단계 검증 과정이 필요하다. 교통사고 조사 전문가와 현직 교통경찰의 정성적 평가를 통해 생성된 시나리오의 현실성과 타당성을 검토하고, 이 과정에서 전문가들은 실제 사고 조사 경험을 바탕으로 시나리오의 현실 적용 가능성과 사고 원인 분석의 정확성을 평가할 수 있다. 실제 발생했던 유사 사고 데이터와의 비교 분석을 통해 시나리오의 정확도를 평가하는 과정도 수행하는데, 이때 차량의 이동 궤적, 충돌 각도, 속도 변화, 제동 거리, 노면 마찰계수 등 주요 물리적 특

성들이 실제 사고 데이터의 패턴과 일치하는지 검증할 수가 있다. 특히 날씨, 도로 상태, 시간대 등 환경 조건에 따른 변수들이 실제 사고 사례와 얼마나 유사한지 중점적으로 분석하게 된다.

이후, 사고 유형별 특성이 실제 사고 분포와 얼마나 부합하는지 확인하는 절차도 진행하는데, 이를 통해 GAN이 생성한 시나리오가 실제 교통사고의 통계적 특성을 잘 반영하고 있는지 검증하며, 특히 사고 다발 구간과 시간대별 사고 패턴이 실제 데이터와 일치하는지 면밀히 분석할 수가 있다. 그리고, 차량 블랙박스와 CCTV 영상을 활용한 검증 단계에서는 실제 사고 영상과 생성된 시나리오를 비교 분석할 수 있다. 딥러닝 기반의 영상 분석 기술을 활용하여 차량의 움직임, 사고 발생 과정, 충돌 순간의 세부적인 특징들이 얼마나 정확하게 재현되었는지 검증하며, 이 과정에서 특히 사고 직전 운전자의 회피 동작이나 긴급 제동과 같은 중요한 세부 사항들이 현실적으로 반영되었는지 확인할 수 있다. 모든 검증 과정을 종합하여 각 시나리오별 신뢰도 점수를 산출하는 것으로 검증 프로세스를 진행하며 이때 전문가 평가(30%), 실제 사고 데이터 비교(25%), 통계적 검증(25%), 영상 분석 결과(20%)를 종합적으로 고려하여 100점 만점의 신뢰도 점수를 부여하고, 85점 이상의 시나리오만을 최종 분석에 활용함으로써 AI 기반 사고 분석의 객관성과 신뢰성을 확보할 수는 방안을 고려하고 있다.

#### IV. 결론

본 논문을 통해 교통사고 예방 및 대응 전략 수립에서 AI가 실질적인 사고 예방과 검증에 활용될 수 있는지를 확인하려 하였다. 또한, 결과적으로 인공지능이 교통사고 예방에 많은 변화를 가져올 수 있음을 시사하고자 하였다. 사고 위험이 높은 지역과 시간대를 인공지능 알고리즘을 통해 식별하여 예방하고, 검증 단계에서 구체적이고 실효성 있는 확인 프로세스로 인공지능을 통한 교통 정책 수립과 사고 분석에 실질적인 활용을 기대한다.

## 참 고 문 헌

- [1] Kim, D., & Lee, J. (2018). "Traffic accident analysis using machine learning approaches." IEEE Big Data and Smart Computing.
- [2] Goodfellow, I., et al. (2014). "Generative Adversarial Networks." Neural Information Processing Systems.
- [3] World Health Organization (2018). "Global Status Report on Road Safety."
- [4] 김동진, "교통사고 예방부터 과실 비율 판단까지...AI 역할 확대," IT동아, 2023년 10월 3주차.
- [5] AJang et al., "엣지 컴퓨팅을 활용한 교통사고 예방 시스템," 한국교통안전공단, 2024.
- [6] Park et al., "실시간 교통상황 예측을 위한 AI 기반 도로 관리 시스템," 한국교통공학회, 2022.

# 자율주행 차량 SDV 원격시동 보안 방안 연구

심윤보\*, 윤철희\*\*

## A Study on response to development of SDV(Software-defined Vehicle) technology

Shim Yun Bo\*, Choel hee Yoon\*\*

### 요 약

최근 자동차 산업에서 커넥티드 카 서비스의 중요성이 증가함에 따라, 차량 제어 모바일 애플리케이션의 보안 강화가 주요 과제로 대두되고 있다. 본 연구에서는 H사의 APK를 대상으로 앱의 보안 메커니즘을 분석하고 잠재적 취약점을 식별하였다. Jadx-gui, Frida, rootAVD 등의 분석 도구를 활용하여 앱의 Entry Point 분석, 동적 코드 로딩 기법 분석, 앱 프로텍터 분석, 그리고 루팅 탐지 우회 가능성 검증을 시도하였다. 분석 결과, 앱에서 사용되는 다양한 보안 기술들의 실제 구현 현황과 한계점을 확인하였으며, 이를 바탕으로 보안 강화를 위한 실질적인 개선 방안을 제시하였다. 본 연구의 결과를 통해 자율주행 차량의 SDV 원격 시동 제어 애플리케이션의 보안 설계 및 구현에 있어 중요한 연구가 될 것으로 기대된다.

### Abstract

As the importance of connected car services in the automotive industry increases, strengthening the security of vehicle control mobile applications has emerged as a major challenge. This study analyzes the security mechanisms and identifies potential vulnerabilities in H's APK. Using analysis tools such as Jadx-gui, Frida, and rootAVD, we examined the app's entry point, dynamic code loading techniques, app protector analysis, and rooting detection bypass possibilities. The analysis revealed the current implementation status and limitations of various security technologies used in the app, based on which we proposed practical improvement measures for security enhancement. The results of this study are expected to serve as an important reference for the security design and implementation of vehicle control applications.

### Key words

Mobile App Security, APK Analysis, Vehicle Control Application, Dynamic Code Loading, App Protector, Rooting Detection, BlueLink

### 1. 서 론

현재 모바일 앱은 일상생활의 필수적인 부분이

되었으며, 특히 자동차 관련 모바일 앱은 차량 제어, 상태 모니터링, 원격 시동 등 다양한 기능을 제공하며 급속도로 발전하고 있다. 이러한 발전과 함

\* 동국대학교, helloong@gmail.com

\*\* 경찰대학교, bertter@police.ac.kr

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임.  
(No.2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

게 모바일 앱의 보안 중요성 또한 크게 부각되고 있다. 특히 차량 제어 관련 앱의 경우, 보안 취약점이 발견될 경우 사용자의 안전과 직결될 수 있어 높은 수준의 보안이 요구된다. H사의 앱은 차량 원격 제어, 차량 상태 확인, 주행 기록 관리 등 다양한 서비스를 제공하는 대표적인 커넥티드 카 플랫폼으로 이러한 앱은 사용자 인증, 데이터 암호화, 앱 무결성 검증 등 다양한 보안 메커니즘을 적용하고 있으나, 이러한 보안 메커니즘의 실효성과 견고성에 대한 분석이 필요한 시점이다.

본 연구에서는 H사의 APK를 대상으로 앱의 구조와 보안 메커니즘을 분석 후 잠재적인 취약점을 식별하며, 이에 대한 대응 방안을 제시하였다. 특히 동적 코드 로딩, 앱 프로텍터, 루팅 탐지 등의 보안 기술이 실제로 어떻게 구현되어 있는지 심층적으로 분석하고, 이러한 보안 메커니즘의 우회 가능성을

검토함으로써 더욱 강화된 보안 방안을 모색하였다.

## II. 본 론

### 2.1 분석 환경 설정

#### a. 안드로이드 에뮬레이터 분석 환경 설정

분석도구로는 Jadx-gui 1.5, frida 16.5.6, frida-tools 13.6.0, firda-dexdump, rootAVD를 사용하였다. 앱 실행을 위한 안드로이드 에뮬레이터[AVD]는 설치되어 있다고 가정하고, 개발자 메뉴를 활성화시킨후 rootAVD를 이용하여 설치된 AVD 환경을 확인하고 에뮬레이터 타입 및 API 버전에 맞는 Rooting을 실행하였다. Rooting이 완료되면 Magisk 앱 설치 확인 및 Zygisk 활성화하고 앱을 설치한다. AVD 에뮬레이터 내 Frida Server 실행한 다음 Frida Server Process를 확인한다.

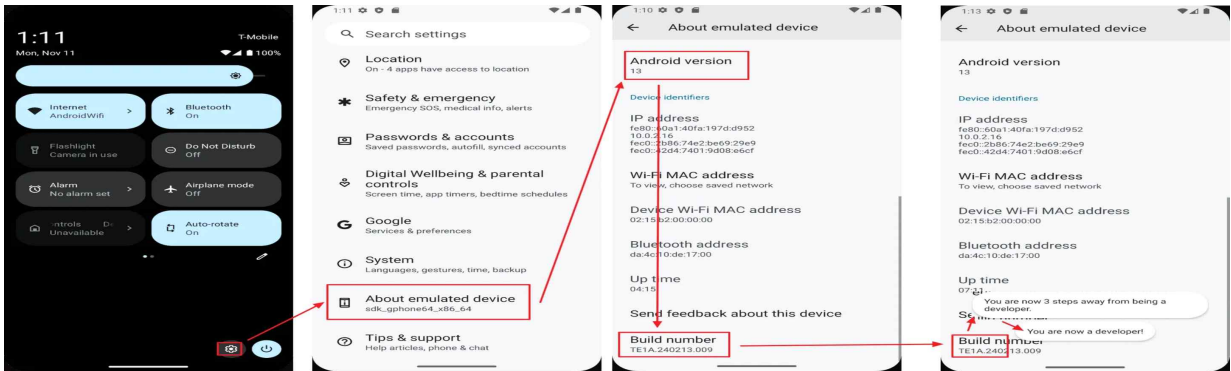


그림 1. 개발자 메뉴 활성화

Fig 1. Activate developer menu

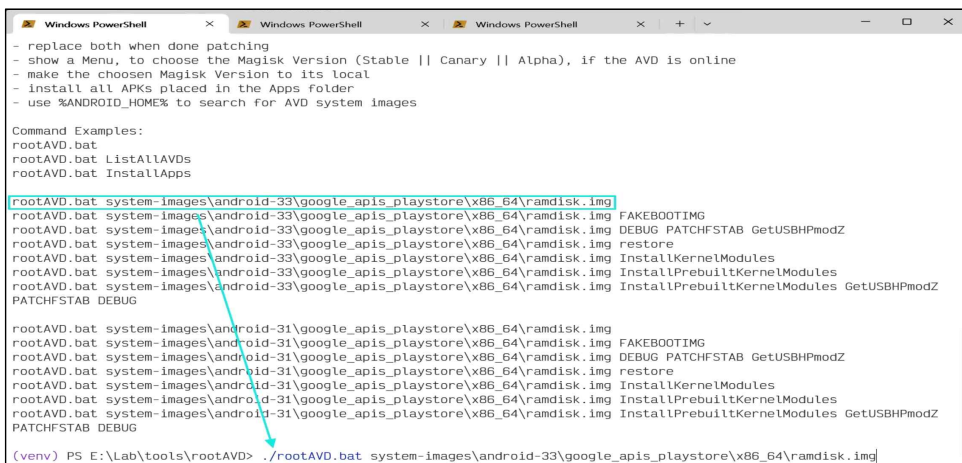


그림 2. 에뮬레이터 타입 및 API 버전에 맞는 Rooting 실행

Fig. 2. Execute rooting appropriate for emulator type and API version

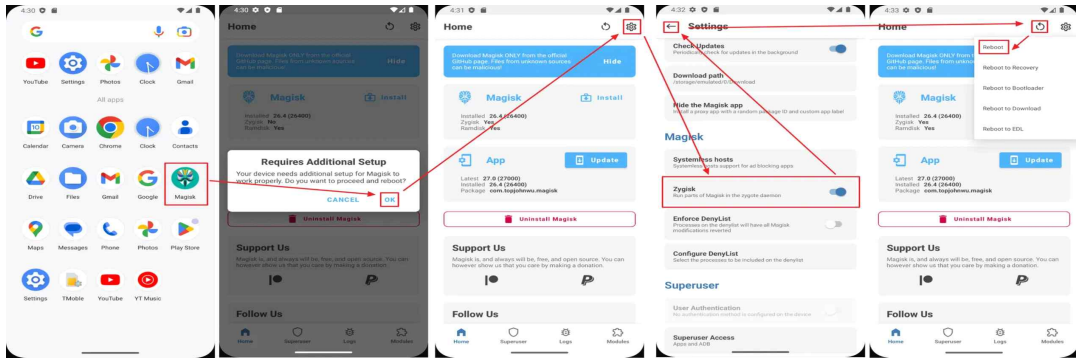


그림 3. Magisk 앱 설치 확인 및 Zygisk 활성화  
Fig. 3. Verify Magisk app installation and activate Zygisk

```
(venv) PS E:\Lab\apks> adb devices
List of devices attached
emulator-5554 device

(venv) PS E:\Lab\apks> adb install .\blueink_3.880.apk
Performing Streamed Install
Success
```

그림 4. 앱 설치  
Fig. 4. install app

```
(venv) PS E:\Lab\tools> adb devices
List of devices attached
emulator-5554 device

(venv) PS E:\Lab\tools> adb push .\frida-server-16.5.6-android-x86_64 /data/local/tmp
.\frida-server-16.5.6-android-x86_64: 1 file pushed, 0 skipped. 219.3 MB/s (11347474 bytes in 0.494s)
```

그림 5. Frida Server 파일 로컬 → 에뮬레이터 장치로 복사  
Fig. 5. Copy Frida Server files local → to emulator device

```
(venv) PS E:\Lab\tools> adb shell
emulator64_x86_64_arm64:/ # su -
emulator64_x86_64_arm64:/ # cd /data/local/tmp
emulator64_x86_64_arm64:/data/local/tmp # ls -al
total 221712
drwxrwx-x 2 shell shell 4096 2024-11-10 16:56 .
drwxr-x-x 5 root root 4096 2024-10-06 00:27 ..
-rwxrwxrwx 1 shell shell 60880 2022-04-25 07:47 Gdump
-rwxr-rw- 1 shell shell 113462264 2024-10-06 00:39 frida-server-16.5.2-android-x86_64
-rwxr-rw- 1 shell shell 113474744 2024-11-10 07:55 frida-server-16.5.6-android-x86_64

emulator64_x86_64_arm64:/data/local/tmp # ls -al
total 221712
drwxrwx-x 2 shell shell 4096 2024-11-10 16:56 .
drwxr-x-x 5 root root 4096 2024-10-06 00:27 ..
-rwxrwxrwx 1 shell shell 60880 2022-04-25 07:47 Gdump
-rwxr-rw- 1 shell shell 113462264 2024-10-06 00:39 frida-server-16.5.2-android-x86_64
-rwxr-rw- 1 shell shell 113474744 2024-11-10 07:55 frida-server-16.5.6-android-x86_64
emulator64_x86_64_arm64:/data/local/tmp # |
```

그림 6. 복사한 파일 실행권한 부여  
Fig. 6. Grant execution permission to copied files

```
emulator64_x86_64_arm64:/data/local/tmp # ./frida-server-16.5.6-android-x86_64 &
[1] 5277
emulator64_x86_64_arm64:/data/local/tmp # |
```

그림 7. AVD 에뮬레이터내 Frida Server 실행  
Fig. 7. Run Frida Server in AVD emulator

```
emulator64_x86_64_arm64:/data/local/tmp # ps -ef | grep -v grep | grep frida
root 5277 5028 1 17:06:01 /debug_ramdisk/.magisk/pts/0 00:00:00 frida-server-16.5.6-android-x86_64
emulator64_x86_64_arm64:/data/local/tmp # |
```

그림 8. Frida Server Process 확인  
Fig. 8. Check Frida Server Process

## 2.2 앱 분석

### a. APP Entry Point 및 Dex dynamic code loading

AVD에 설치된 앱을 실행하면 앱 스플래시 화면과 검은색 화면이 반복적으로 전환되어 나타나며 실행되지 않는다. 이부분에서 App 프로텍터의 동작 여부를 추정 할 수 있다. 실제 분석을 통해 프로텍터 존재 여부를 확인하고자 앱 APK 파일을 jadx-gui로 실행한다. APK 진입점(Entry Point) 확인을 위해 왼쪽 패널의 “리소스”탭의 “AndroidManifest.xml”을 클릭하여 확인한다.

“AndroidManifest.xml” 파일에서 MAIN과 LAUNCHER 키워드로 해당 APK의 Entry Point에 해당하는 시작 Activity를 확인한다. 확인 된 패키지 경로 “com.velox.hkmc\_tmlk.activities.auth”와 “SDMSplashActivity”를 소스코드 트리에서 검색하면 해당 소스 패키지 경로가 소스코드상에 존재하지 않음을 확인 할 수 있다. 해당 APK 파일은 일부 코드들이 앱 기동시 로딩되는데 대표적으로 사용되는 동적 코드 로딩 기법에는 크게 3가지 메소드(DexClassLoader, PathClassLoader, InMemoryDexClassLoader)가 사용된다.

메소드들이 존재하는지 전체 코드를 검색하면 PathClassLoader를 사용하는 코드가 존재하는 것을 알 수 있고, 해당 코드 상단에 선언된 내용을 보면 BaseDecClassLoader 또한 선언되어 있으며 아래 코드들에서 사용되고 있는 것을 확인할 수 있다.



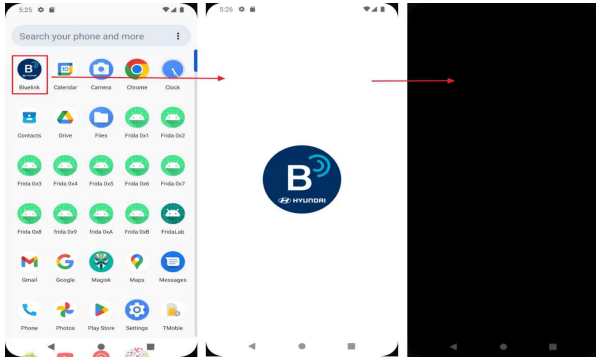


그림 9. 앱 실행  
Fig. 9. Run App

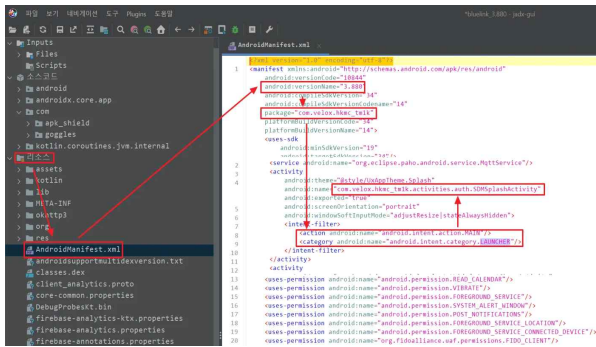


그림 10. APK 진입점(Entry Point) 확인  
Fig. 10. Check APK Entry Point

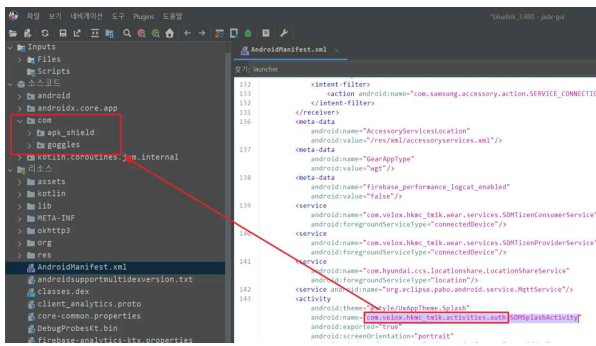


그림 11. 해당 소스 패키지 경로가 소스코드상에 존재하지 않음  
Fig. 11. The source package path does not exist in the source code

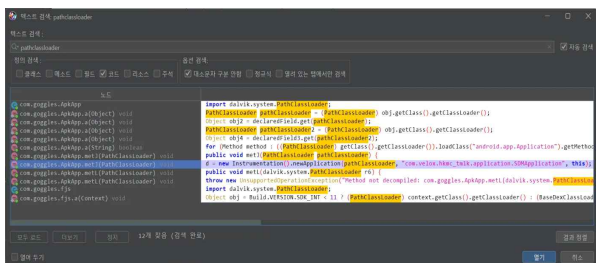


그림 12. 전체 코드 검색  
Fig. 12. Full code search

### b. APP Protector 확인

분석 전, lib 폴더에 native 라이브러리들이 존재하는지 확인하고 여러 개의 so 파일들이 존재함을 확인한다. 존재하는 코드상에서 의심스러운 부분을 확인할 때, 왼쪽 패널의 소스코드 트리에 노출된 com.apk\_shield 카테고리가 확인된다. 해당 카테고리 내에는 2개의 클래스와 하나의 인터페이스가 존재하는데 그중 skdb라는 클래스가 코드량이 방대하며 무언가 암호화 또는 인코딩된 문자열들이 보이고 루팅에 관련된 상수정의를 하는 것처럼 보인다. 끝이 “=” 또는 “=”으로 끝나는 문자열은 Base64로 인코딩되어 있음을 추정할 수 있다. 이를 디코딩해보면 루팅에 사용되는 su, supersu 등의 바이너리나 루팅 또는 디버깅에 관련된 여러 메시지를 인코딩해놓은 것을 확인할 수 있으며, 앱 프로텍터 클래스인 것을 알 수 있다. 코드를 확인하면 수많은 리소스 ID와 같은 상수 선언들이 이어지다 코드의 마지막 부근인 약 4만라인 부근에서 함수와 코드 로직이 나타나는 것과 중간지점에 native 메소드를 선언 또는 콜하는 것을 알 수 있다. 해당 함수들이 실제로 앱 기동시 콜 되는지 간단하게 Frida의 아래 코드를 사용해서 확인할 수 있으며, frida script를 명령어로 실행했을 때 감박임이 사라진 것을 알 수 있다. 그러나, 감박임은 사라져도 기기 루팅으로 인한 앱 실행이 제한되며 Confirm 또는 확인 버튼을 누르면 앱이 Pause 상태로 들어간다.

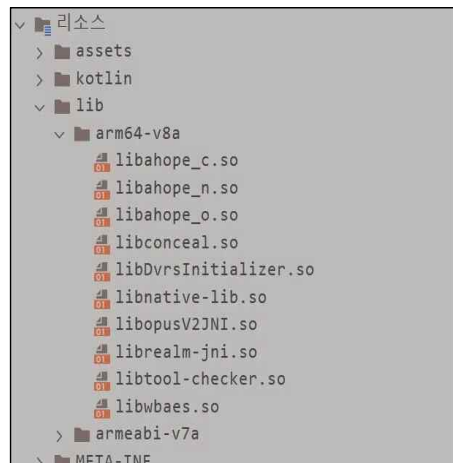


그림 13. lib 폴더상 native 라이브러리 존재 확인  
Fig. 13. Check the existence of native library in the lib folder

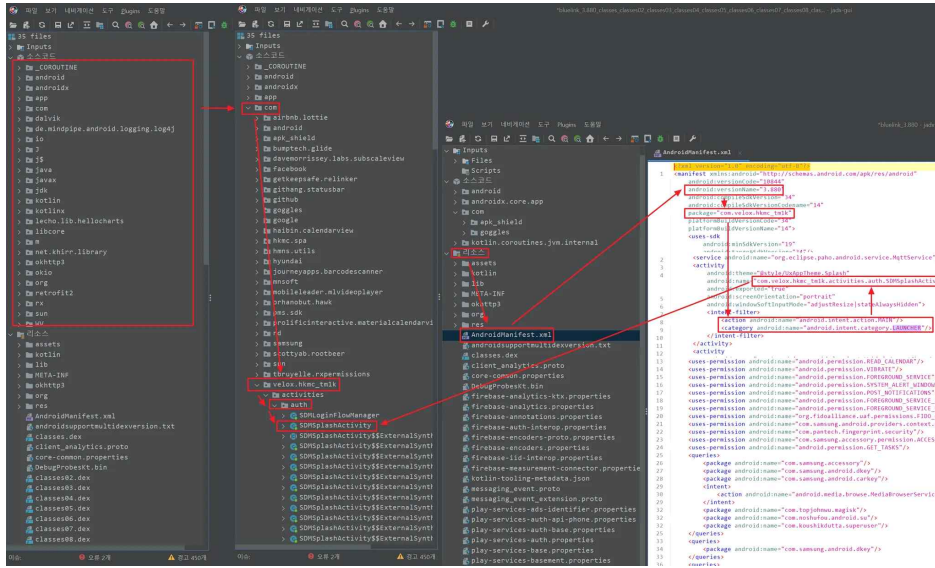


그림 14. 앱 시작점 코드 존재 확인  
Fig. 14. Check for existence of app entry point code

### c. Dex Dump와 Code Recovery

Rooting Bypass 전에 동적 코드 로딩 기법으로 숨겨진 코드를 메모리 덤프(frida-dexdump)를 통해 복원하고 pip를 통해 frida-dexdump를 설치한다. frida-dexdump 설치 후 앱 실행 및 F Attach 모드로 frida-dexdump를 실행한다. 덤프를 실행한 후, 해당 경로 하위에 bluelink라는 폴더와 함께 덤프된 dex 파일들이 생성된다. dex-dump에서 덤프된 파일을 jadx-gui에 apk가 로드된 상태에서 추가로 로드하면 동적 로드된 코드들이 보여지는데 처음 로드했던 apk 소스보다 많은 양의 코드가 추가되었음을 알 수 있으며, 또한 위에서 확인했던 AndroidManifest.xml 파일에서의 앱 시작점 코드가 존재하는 것을 확인할 수 있다.

### d. Rooting Bypass

루팅을 탐지하는 방법에는 여러가지가 있지만 대부분 su, superuser, magisk 등 슈퍼유저 권한을 부여하는 권한 상승 바이너리가 존재하는지 검사하는 키워드들이 존재하는 경우가 많다. 여기서 루팅시 사용했던 rootAVD에서 설치되는 magisk 바이너리를 검사하는 로직이 있는지 확인한 결과, 전체 검색창에서 “magisk”를 검색하면 키워드 검색 결과가 나온다.



그림 15. “magisk” 키워드 검색 결과  
Fig. 15. “magisk” keyword search results

scottyab.rootbeer이라는 루팅 탐지 패키지를 통해서 루팅 검사를 하는 것을 알 수 있다. 해당 클래스의 가장 의심스러운 isRooted() 메소드 코드를 더블 클릭하여 열면 isRooted() 메소드는 많은 로컬 메소드들을 콜하여 그 반환 값이 참 거짓인지 논리연산을 하고 있는 것을 볼 수 있다. 또한, 해당 클래스를 살펴보면 여러가지 루팅 탐지 로직들이 존재함을 알 수 있다. 루팅된 기기에서 자주 발견되는 바이너리 경로와 파일명 등을 상수로 선언하고 그 내역을 순회하며 검색하는 로직들도 보여진다. RootBeer 클래스내의 여러가지 루팅 탐지 메소드들을 frida로 재작성하여 루팅 탐지를 우회하면, 앱이 실행은 되지만 루팅이 탐지되어 확인 버튼을 누르면 앱이 종료된다. 확인한 isRooted()에 나열된 메소드들 전부 frida 스크립트를 사용하여 재작성한다.

이때 주의할 점은 현재 동적 코드 로딩으로 RootBeer 클래스가 로딩됨으로 해당 클래스가 로드 될 때까지 계속해서 메모리 영역을 검색하고, 클래스가 발견되어 지면 발견 플래그를 업데이트하여 더이상 수행되지 않도록 변경 한 후 해당 클래스의 메소드들을 재작성한다. 해당 스크립트를 포함하여 재실행하면 깔끔하게 루팅 탐지가 우회되어 루팅된 장비에서 실행된 모습을 확인할 수 있다.

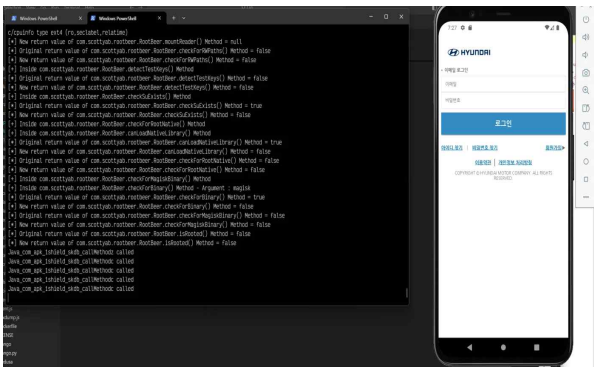


그림 16 루팅된 장비에서 실행 확인  
Fig. 16. Confirm running on rooted device

### III. 결 론

자율주행 차량 sdv의 모바일 앱 보안은 현대 디지털 환경에서 필수적인 요소로 자리잡고 있으며, 이를 강화하기 위한 다양한 기술과 방법이 개발되고 있다. 비록 완벽한 보안은 달성하기 어렵지만, 지속적인 노력과 최신 보안 기술의 도입을 통해 보안 위협을 효과적으로 관리할 수 있다. 특히, 민감한 정보를 다루는 앱 개발자들은 보안에 대한 높은 의식을 가지고, 사용자 데이터를 안전하게 보호하기 위한 최선의 노력을 기울여야 할 것이다.

### 참 고 문 헌

[1] Y. H. Kim, J. S. Park, "A Study on Security Vulnerabilities and Countermeasures of Connected Car Applications," Journal of the Korea Institute of Information Security & Cryptology, 31(4), pp.789-798, 2021  
[2] S. J. Lee, H. K. Kim, "Analysis of Android

Application Security: Focusing on Runtime Protection Techniques," Journal of the Korea Institute of Information Technology, 19(12), pp.157-168, 2021  
[3] David Woods, Timothy Mattern, "Smartphone Security: Analysis of Mobile Anti-Debug Techniques and Countermeasures," IEEE Access, vol. 9, pp. 123456-123470, 2023  
[4] M. K. Park, S. H. Lee, "Security Analysis of Vehicle Control Applications: A Case Study of Remote Start Systems," International Journal of Automotive Technology, 24(5), pp.891-904, 2023  
[5] Zhen Liu, Yan Chen, "A Comprehensive Survey of Android Application Security: From Traditional Protection to AI-based Detection," ACM Computing Surveys, 55(2), Article 35, 2023  
[6] John Smith, Maria Garcia, "Software-Defined Vehicles: Security Challenges and Solutions in Connected Car Platforms," IEEE Transactions on Vehicular Technology, 72(8), pp.234-245, 2023  
[7] H. J. Choi, K. S. Park, "Research on Security Enhancement of Vehicle Control Mobile Applications using Dynamic Code Analysis," Journal of KIISE, 48(6), pp.678-689, 2021



# 데이터 도메인별 드리프트 감지 기술의 최신 동향

김범윤\*, 전병우\*\*

## Trends in Data Drift Detection Techniques Across Different Data Domains

Bumyoon Kim\* and Byeungwoo Jeon\*\*

### 요약

데이터 드리프트 또는 개념 드리프트가 발생하는 경우 네트워크 모델의 성능이 저하될 가능성이 매우 크다. 따라서 이런 드리프트를 감지하는 것이 모델 성능을 유지하는데 중요한 역할을 한다. 이미 여러 분야에서 이러한 중요성을 인식하고 드리프트 감지에 대해 활발히 연구가 진행되고 있다. 본 논문에서는 여러 데이터 도메인에서 드리프트를 감지하는 기술의 최신 동향에 대해 알아보하고자 한다.

### Abstract

When data drift or concept drift occurs, the performance of network models is likely to decline. Detecting these drifts, therefore, plays a crucial role in maintaining model performance. Recognizing the importance of drift detection, extensive research is already underway across various domains. This paper aims to explore recent trends in techniques for detecting drifts across multiple data domains.

### Key words

data drift, drift detection, various data domain, image data drift, text data drift, audio data drift

## 1. 서론

기계학습의 발전으로 많은 분야에서 인공지능망 등을 활용하여 종래 풀기 힘든 여러 문제들을 해결할 수 있게 되었다. 인공지능망 모델은 학습에 사용된 데이터와 실제 환경에서의 데이터가 유사할수록 좋은 성능을 낸다. 실제 환경에서 접하는 데이터 환경과 모델이 학습한 데이터 환경 간 차이가 발생하

는 것을 데이터 드리프트 또는 개념 드리프트라고 하는데 (이하 드리프트), 드리프트가 발생하는 경우 모델의 성능이 기대에 못 미치게 된다 [1]. 따라서 드리프트를 감지하여 성능저하를 저감시키는 것은 모델의 유지관리 측면에서 매우 중요하다고 볼 수 있다. 본 논문에서는 다양한 데이터 도메인에서 드리프트를 감지하는 기술들의 최신 동향에 대해 알아보하고자 한다.

\* 성균관대학교 전자전기컴퓨터공학과, kbumyoon@skku.edu,

\*\* 성균관대학교 전자전기컴퓨터공학과, bjeon@skku.edu

※ 이 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. RS-2024-00337489, 분석 모델의 성능저하 극복을 위한 데이터 드리프트 관리 기술 개발)

## II. 데이터 도메인별 드리프트 감지 기술

## 참 고 문 헌

영상도메인에서는 드리프트를 감지하기 위하여 네트워크 모델의 출력 값이나 특성의 변화를 관찰하는 기술이 연구되었다[2]. 이러한 연구들은 드리프트가 발생하면 모델의 출력 값들의 확률 분포에 변화가 생기는 것을 기반으로 하였다. 또한 별도의 네트워크 모델을 사용하여 영상으로부터 드리프트와 밀접한 특징을 추출하고 비교 분석을 수행하여 드리프트를 감지하는 연구도 수행되었다 [3][4]. 이때 추출된 특징들의 비교는 주로 공간상의 거리를 측정하는 코사인 유사도 등을 활용하여 수행된다.

텍스트와 오디오 도메인에서도 영상 도메인에서와 마찬가지로 입력 데이터로부터 특징을 추출하여 분석함으로써 드리프트를 감지하는 연구들이 많이 수행되고 있다[5][6][8][9]. 주로 시계열 특징을 통해 드리프트를 감지한다. 또한 데이터의 통계적인 특성이 변하는 것을 직접적으로 감지하는 연구들 또한 많이 수행되고 있다 [7][10].

드리프트를 감지하는 연구들은 크게 데이터의 통계적 특징을 활용하거나, 네트워크 모델을 통한 특징을 활용하는 두 가지 방법으로 나뉜다. 데이터의 통계적 특징을 활용하는 경우 실시간성이 보장되거나 계산 복잡도가 상대적으로 낮다. 네트워크를 통해 특징을 추출하는 경우 특징을 추출하는데 있어 계산 비용이 높고, 시간도 오래 걸린다는 단점이 있지만, 복잡한 환경 또는 정확하게 드리프트를 감지하여야 하는 상황에 적절하다.

## III. 결 론

본 논문에서는 데이터 드리프트를 감지하기 위한 기술들의 동향에 대해 조사하였다. 주로 데이터를 고차원 공간으로 임베딩하여 이를 통해 드리프트를 감지하는 기술들이 주를 이루고 있다. 이를 토대로, 멀티 모달 네트워크 사용 환경에서 서로 다른 도메인의 데이터를 같은 공간으로 임베딩하여 동시에 드리프트를 감지한다면 더욱 효율적으로 멀티 모달 모델을 유지관리할 수 있을 것임을 기대할 수 있다.

- [1] F. Bayram, B. S. Ahmed, and A. Kassler. "From concept drift to model degradation: An overview on performance-aware drift detectors", *Knowledge-Based Systems*, 2022.
- [2] A. Prenner and B. Kainz, "Bias Assessment and Data Drift Detection in Medical Image Analysis: A Survey," 2024, *arXiv:2409.17800*.
- [3] Y. Wu et. al, "Efficient Data-Sketches and Fine-Tuning for Early Detection of Distributional Drift in Medical Imaging," 2024, *arXiv:2408.08456*.
- [4] W. Li et al., "A drift detection method for industrial images based on a defect segmentation model," *Knowledge-Based Systems*, vol. 301, Oct. 2024.
- [5] R. Mfondoum et. al., "Outlier Detection in Streaming Data for Telecommunications and Industrial Applications: A Survey," *Electronics*, vol. 13, no. 16, Aug. 2024.
- [6] T. Ganitidis et. al., "Sustaining model performance for covid-19 detection from dynamic audio data: Development and evaluation of a comprehensive drift-adaptive framework," 2024, *arXiv:2049.19300*.
- [7] H. Mehmood et. al., "A novel Edge architecture and solution for detecting concept drift in smart environments," *Future Generation Computer Systems*, vol. 150, pp. 127-143, Jan. 2024.
- [8] S. Greco et. al., B. Vacchetti, D. Apiletti, and T. Cerquitelli, "Unsupervised Concept Drift Detection from Deep Learning Representations in Real-Time," 2024, *arXiv:2406.17813*.
- [9] J. Li, H. Yu, Z. Zhang, X. Luo, and S. Xie, "Concept Drift Adaptation by Exploiting Drift Type," *ACM Transactions on Knowledge Discovery from Data*, vol. 18, no. 4, pp. 1-22, Feb. 2024.
- [10] C. Garcia, A. Koerich, A. Britto, and J. Barddal, "Methods for Generating Drift in Text Streams," 2024, *arXiv:2403.12328*.

# 데이터 드리프트 감지를 위한 오픈소스 라이브러리 분석 비교

예성진\*, 김범윤\*\*, 전병우\*\*\*

## Comparative Analysis of Open-Source Libraries for Data Drift Detection

Seongjin Ye\*, Bumyoon Kim\*\*, and Byeungwoo Jeon\*\*\*

### 요약

데이터 드리프트는 머신러닝 모델의 성능을 저하시킬 수 있는 요소로, 이를 조기에 감지하지 못할 경우 예측 정확도가 감소할 수 있다. 이러한 문제를 해결하기 위해 다양한 오픈소스 라이브러리들이 개발되었으며, 각 라이브러리는 다양한 드리프트 감지 방법을 제공한다. 본 논문에서는 Evidently AI, Alibi Detect, NannyML, Deepchecks 네 가지 주요 라이브러리를 비교하여 주요 기능과 드리프트 감지 기법을 분석하였다.

### Abstract

Data drift is a factor that can degrade the performance of machine learning models. Failure to detect it early can lead to a decrease in prediction accuracy. To address this issue, various open-source libraries have been developed, each providing diverse methods for drift detection. This paper compares four major libraries—Evidently AI, Alibi Detect, NannyML, and Deepchecks—and analyzes their key features and drift detection techniques.

### Key words

data drift detection, machine learning, open-source libraries for drift detection

## 1. 서론

머신러닝 모델의 학습과정과 실제 운영시의 데이터 간에 차이가 발생할 수 있는데, 이를 일컫는 데이터 드리프트는 모델 성능을 저하시킬 수 있다. 실시간 스트림 데이터에는 드리프트가 자주 발생하는데, 이를 감지하지 못할 경우 예측의 정확도가 감소할 수 있어 [1], 머신러닝 모델의 성능 유지를 위해

데이터 드리프트의 조기 감지가 필요하다. 이러한 필요성에 따라, 데이터 드리프트 감지를 위한 다양한 라이브러리가 개발되고 있다. 각 라이브러리에서 지원하는 감지 방법 데이터의 유형에 따라 다를 수 있다 [1]. 본 논문은 Evidently AI [2], Alibi Detect [3], NannyML [4], Deepchekcs [5]의 네 가지 오픈소스 라이브러리가 제공하는 드리프트 감지 방법을 분석한다.

---

\* 성균관대학교 전자전기공학부, dolcei6966@skku.edu,  
\*\* 성균관대학교 전자전기컴퓨터공학과, kbumyoon@skku.edu,  
\*\*\* 성균관대학교 전자전기컴퓨터공학과, bjeon@skku.edu

※ 이 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. RS-2024-00337489, 분석 모델의 성능저하 극복을 위한 데이터 드리프트 관리 기술 개발)

## II. 드리프트 감지 라이브러리 분석

Evidently AI[2]는 모델의 예측 값에서 발생하는 데이터 분포의 변화를 감지하기 위한 데이터 드리프트 감지 알고리즘을 제공한다. 데이터 유형에 적합한 여러 감지 방법을 지원하며, 데이터의 분포를 여러 피처로 분석한 시각 자료를 제공한다. 전체 피처의 50% 이상 드리프트가 발생한 경우에 데이터 세트 전체에 드리프트가 발생했다고 판단한다. DatasetDriftPreset() 함수를 활용하여 드리프트 판단 기준을 포함한 기본 설정을 임의로 변경할 수 있다.

Alibi Detect[3]는 여러 감지 방법을 지원하고 Spot-The-Difference Test라는 라이브러리 고유의 감지 방법을 사용한다. 해당 방법은 입력 데이터의 개별 피처가 아니라, 데이터 자체의 드리프트를 감지한다는 목적이 있다. 따라서, 드리프트 여부를 판단할 때 소요되는 시간이 적지만 드리프트에 영향을 끼치는 데이터의 피처는 구분하지 못한다. preprocess\_fn() 함수를 활용하여 데이터의 차원 축소나 텍스트 임베딩의 전처리 기능을 지원한다.

NannyML[4]은 오픈소스 버전과 확장된 기능을 제공하는 NannyML Cloud 버전이 있다. NannyML은 데이터에 따라 변하는 모델의 성능을 실시간으로 모니터링하며 드리프트를 추정한다. 모델의 성능 저하의 원인이 되는 주요 특징을 식별하고 이를 바탕으로 모델의 재훈련 필요성을 평가한다. Cloud 버전은 자동화된 모니터링 데이터 수집 기능, 기 설정된 재훈련 파이프라인을 통한 모델 재훈련 기능을 제공한다.

Deepchecks[5]는 데이터 드리프트 감지와 모델 성능 모니터링을 지원하는 파이썬 라이브러리로, 모델이 프로덕션에 들어가기 전 데이터와 모델의 일관성을 검증하는 데 중점을 둔다. 모델이 학습한 데이터와 테스트 데이터를 기반으로 모델 성능을 평가하며, 대시보드에서 주요 성능 지표를 시각적으로 확인할 수 있는 기능을 제공한다. 해당 라이브러리는 full\_suite() 함수를 통해 Confusion Matrix, 모델 비교, 데이터 드리프트 여부 등의 결과를 하나의 리포트로 제공한다.

표 1은 각 라이브러리에서 지원하는 데이터 드리프트 감지 방법을 비교한 것이다.

표 1. 각 라이브러리의 데이터 드리프트 감지 방법

드리프트 감지 방법	Evidently AI	Alibi Detect	NannyML	Deepchecks
Chi-Squared Test	○	○		○
CBPE			○	
Context-Aware MMD		○		
CVM Test	○	○		○
KS Test	○	○	○	○
M-CBPE			○	
MMD		○		
PSI	○			
Spot-The-Difference		○		
Wasserstein Distance	○		○	

## III. 결 론

데이터 드리프트는 머신러닝 모델의 성능을 저하시킬 수 있는 중요한 요소로, 다양한 드리프트 감지 라이브러리를 통해 이를 효과적으로 관리할 수 있다. 본 연구에서는 Evidently AI, Alibi Detect, NannyML, Deepchecks 네 가지 주요 오픈소스 라이브러리를 비교하여, 각 라이브러리가 제공하는 드리프트 감지 기법과 특징을 분석하였다.

## 참 고 문 헌

- [1] G. I. Webb, R. Hyde, H. Cao, H. L. Nguyen, and F. Petitjeon. "Characterizing concept drift," Data Mining and Knowledge Discovery, 2016.
- [2] Evidently AI Software repository. Accessed: Nov 2 024. [Online]. Available: <https://github.com/evidentlyai/evidently>.
- [3] Alibi Detect Software repository. Accessed: Nov 2 024. [Online]. Available: <https://github.com/SeldonIO/alibi-detect>.
- [4] NannyML Software repository. Accessed: Nov 2024. [Online]. Available: <https://github.com/NannyML/nannyml>.
- [5] Deepchecks Software repository. Accessed: Nov 2024. [Online]. Available: <https://github.com/deepchecks/deepchecks>.



# 마이크로서비스 아키텍처 기반 데이터 드리프트 서비스 실증 시나리오에 관한 연구

이지현\*, 박종빈\*\*

## A Study on Data Drift Service Implementation Scenarios Based on Microservices Architecture

Lee Ji Hyun\*, and Jongbin Park\*\*

### 요 약

본 논문은 데이터 드리프트 문제를 해결하기 위해 마이크로서비스 아키텍처를 기반으로 한 서비스 실증 시나리오 및 구현 방안을 제안한다. 다양한 산업 분야에서 기 학습된 모델의 성능이 현장에서 크게 저하되거나 잘못된 예측을 초래하는 경우가 빈번한데 이는 데이터 드리프트 문제로 해석하고 대응할 수 있다. 이러한 데이터 드리프트 대응 기술은 클라우드 뿐만 아니라 다양한 온프레미스 환경, 클라우드-온프레미스 연동 환경으로 확대 적용할 필요가 있다. 그러나 구현하는 측면에서는 개발 복잡도가 크게 증가하고 유지 관리 측면에서도 어려움이 발생한다. 또한 기술 개발 요구사항을 충실히 반영하면서 개발 결과물을 효과적으로 검증할 수 있는 실증 시나리오 역시 중요하다. 따라서 본 논문에서는 사회적, 경제적, 기술적으로 파급력이 높은 실증 영역을 설정하고 마이크로 서비스 아키텍처 방식을 활용하여 실증 시나리오를 구체적으로 구현하는 방법을 소개한다.

### Abstract

This paper proposes a service implementation scenario and a solution based on a microservice architecture to address the issue of data drift. In various industries, the performance of pre-trained models often deteriorates significantly or leads to incorrect predictions in real-world environments, which can be attributed to data drift. Techniques for handling data drift need to be applied not only to cloud environments but also to various on-premises and hybrid cloud-on-premises environments. However, implementing such solutions increases development complexity and introduces challenges in maintenance. Moreover, it is crucial to establish proof-of-concept scenarios that can effectively verify development outcomes while faithfully reflecting the technical requirements. Therefore, this paper introduces a method for concretely implementing demonstration scenarios using a microservice architecture in areas with significant social, economic, and technological impact.

### Key words

data drift, data drift service, microservice architecture, implementation scenarios

---

\* 한국전자기술연구원, leejihyun@keri.re.kr,

\*\* 한국전자기술연구원, jpark@keti.re.kr(교신저자)

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. RS-2024-00337489, 분석 모델의 성능저하 극복을 위한 데이터 드리프트 관리 기술 개발)

## 1. 서론

최근들어 공공(교통, 기상, 스마트시티 등)과 민간(이커머스, 금융 등)영역에서 다양하고 많은 양의 데이터를 생성하고 있으며, 이 데이터를 분석하고 예측 및 결정을 지원하기 위한 효율적인 데이터 분석 및 머신러닝 기술이 필요하다[1]. 시간이 지남에 따라 데이터의 분포나 관계가 변하는 현상을 의미하는 데이터 드리프트는 학습 모델의 성능을 낮추는 주요 요인으로 작용하는 것으로 보고되고 있다 [2]. 이로 인해 새로운 제품, 새로운 시장 등 급격한 변화는 데이터 드리프트 문제를 야기할 수 있다. 이러한 데이터 드리프트 문제를 해결하기 위해서는 학습 모델의 예측 성능 변화를 지속적으로 추적 및 관찰하여 드리프트 발생 여부를 확인하고 원인을 분석하는 과정이 선행되어야 하며, 필요할 경우 학습 데이터의 구성을 바꾸어 모델을 재학습 시키는 과정이 요구된다.

이와 같은 데이터 드리프트 대응 기술은 종래의 클라우드 환경뿐만 아니라 다양한 온프레미스 환경, 클라우드-온프레미스 연동 환경으로 확대 적용할 필요가 있다. 그러나 이를 효과적으로 구현하기 위해서는 개발 복잡도가 크게 증가할 뿐만 아니라 서비스 유지 관리 측면에서도 다양한 어려움이 발생할 수 있다. 또한 적절한 실증영역을 확보하고 이를 현장에 적용하여 실질적으로 발생하는 문제점을 적극적으로 파악하고 이를 기술 개발에 반영하는 노력도 필요하다. 따라서 본 논문에서는 클라우드와 온프레미스 등의 다양한 환경에서 데이터 드리프트 문제 해결이 필요한 실증 서비스 영역을 정의한다. 그리고 각 서비스를 독립적으로 관리하고 마이크로서비스 아키텍처 방식을 활용하여 요구사항 변화에 유연하게 대응할 수 있는 동시에 안정적인 개발과 개발과 유지보수의 효율성을 높이는 방안을 제안한다.

## II. 본론

### 1) 데이터 드리프트 문제 해결 실증 시나리오

데이터 드리프트 대응 기술은 여러 분야에서 활용할 수 있으나 본 논문에서는 공공 및 민간 분야에서 사회적, 경제적, 기술적 파급력을 고려하여 표 1 과 같이 실증 서비스 영역을 정의한다. 이는 실시간으로 변화하는 데이터에 신속하게 대응할 수 있는 전략적 적용 시나리오에 해당하며 시나리오 별 사례의 특성과 요구사항에 맞추어 선정하였다.

표 1. 데이터 드리프트 관리 실증 시나리오 예시  
Table 1. Example of data drift management implementation scenarios

분야		실증 시나리오 예시
공공	교통	신형 번호판 도입 이후 모델의 번호판 인식 정확도가 급격히 떨어지면, 시스템이 이를 감지하고 데이터 드리프트로 인한 성능 저하 경고 및 재학습 권유
	민생	새로운 피싱 패턴이 등장하거나 피싱 메시지에 사용되는 단어나 문구의 빈도가 실시간으로 변할 때, 이를 탐지하고 패턴 변화에 따라 알림 제공
민간	마케팅	새로운 트렌드에 따라 특정 상품의 수요가 급증하거나 급감하면, 이를 드리프트로 감지하여 모델이 이를 반영하지 못하는 상황을 조기에 확인

### 2) 마이크로서비스 아키텍처

마이크로서비스 아키텍처(Microservice Architecture)는 클라우드 컴퓨팅 서비스 등의 대규모 소프트웨어 시스템의 복잡성을 효과적으로 관리하기 위해 등장하였다[3]. 마이크로서비스 아키텍처는 소프트웨어가 잘 정의된 API를 통해 통신하는 독립적인 소규모 서비스들로 구성된 소프트웨어 개발을 위한 아키텍처 및 조직적 접근 방식이다. 이러한 마이크로서비스 아키텍처는 서비스 지향 소프트웨어 산업에서 주요 아키텍처 스타일로 자리 잡았다. 마이크로서비스는 시스템을 작고 가벼운 서비스로 나누어, 각 서비스가 매우 응집력 있는 비즈니스 기능을 수행하도록 설계되었다[4].

### 3) 마이크로서비스 아키텍처 기반 데이터 드리프트 관리 기술

점차 기존의 전통적인 컴퓨팅 시스템인 온프레미스 방식과 다양한 클라우드 시스템을 연동하여 통합 운영·관리하는 하이브리드 환경이 구축되고 있다[5]. 이러한 환경에서 다중 플랫폼 간의 통합과 상호 운용성이 중요한 요소로 작용한다. 마이크로서비스 아키텍처는 각 서비스가 독립적으로 개발되어 다양한 플랫폼에 배포될 수 있기 때문에 클라우드와 온프레미스 환경을 아우르는 확장성과 유연성을 제공할 수 있다. 특히 데이터 드리프트 문제는 데이터의 종류에 따라 클라우드 또는 온프레미스 환경에서 개별적으로 발생할 수 있으므로 마이크로서비스 아키텍처는 이러한 이질적인 환경에서 발생하는 데이터 드리프트 문제를 통합적으로 관리하고 대응할 수 있는 적합한 아키텍처라고 할 수 있다.

데이터 드리프트 문제 해결을 위한 기술 개발에 있어 중요한 요소 중 하나는 실증 시나리오를 통해 제안된 기술의 유효성을 검증하는 과정일 것이다. 이때, 마이크로서비스 아키텍처는 서비스 간의 결합도가 낮아 독립적인 기능별 검증이 용이하다는 장점을 갖고 있다. 데이터 드리프트 대응 기술이 특정 데이터나 특정 단계에 한정되어 있을 경우, 해당 기술만 독립적으로 모듈화하여 검증할 수 있는 환경을 제공한다. 이는 전체 시스템을 대상으로 하는 통합 테스트보다 더 세밀하고 빠른 실증을 가능하게 한다. 이러한 측면에서 마이크로서비스 아키텍처는 다양한 검증 요구사항을 충족하기에 적합한 구조적 기반을 제공한다.

그림 1은 본 논문에서 제안하는 마이크로서비스 아키텍처 방식을 적용한 데이터 드리프트 관리 서비스 구조를 설명한다. 데이터 드리프트 관리를 위한 요소 기술을 컨테이너 방식의 마이크로서비스 모듈로 구성하고 이를 조합하여 실증 서비스 환경에 배포된 요소 기능들을 나누어 분석하는 방식이다. 구체적으로는 데이터 드리프트 진단부, 데이터 재구성부, 분석 모델 재학습부를 포함한다. 예를 들어 그림 1에서 서비스 A는 기능 a, b, c, d를 포함하는데 데이터 드리프트 진단부가 각 기능을 분석하여 기능 b에 문제가 있다고 판단되면 데이터 재구성 기능을 추가로 적용하여 문제를 해결한다.

구체적인 예로 표 1에 제시된 교통 분야의 경우,

표 2에 나와있는 “번호판 영역 검출 기능”, “번호판 글씨 인식 기능”, “번호판 종류 분류 기능”, “차량 분류 기능”이 데이터 드리프트 관리 서비스 적용이 가능한 기능으로 판단된다. 이를 통해 각 기능별로 발생하는 데이터 드리프트 현상을 파악하고 적절히 처리할 수 있다.

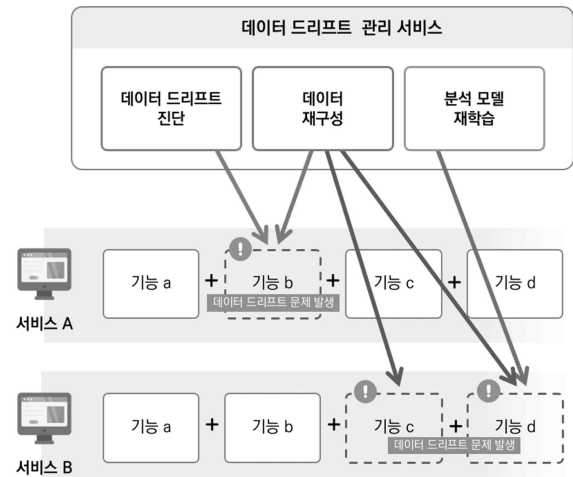


그림 1. 마이크로서비스 아키텍처 기반 데이터 드리프트 서비스 실증 시나리오 개념도

Fig. 1. Conceptual diagram of data drift service implementation scenario based on microservice architecture

표 2. 실증 시나리오별 데이터 드리프트 관리가 필요한 요소 기능 예시

Table 2. Example of functional elements requiring data drift management by implementation

분야		실증 시나리오에 따른	
		발생 가능한 데이터 드리프트 환경 예시	데이터 드리프트 관리가 필요한 요소 기능 예시
공공	교통	- 신규 번호판 등장 - 신규 차종 등장 - 신규 교통법 등장	- 번호판 영역 검출 기능 - 번호판 글씨 인식 기능 - 번호판 종류 분류 기능 - 차량 분류 기능
	민생	- 신규 피싱 패턴 등장 - 신규 첨부파일 및 링크 유형 등장	- 피싱 메시지 패턴 인식 기능 - 이메일 발신자 도메인 검출 기능 - 첨부파일 및 링크 유형 검출 기능
민간	마케팅	- 신상품 등장 - 신규 트렌드 등장 - 신규 채널 등장	- 제품 및 서비스 선호도 변화 분석 기능 - 마케팅 채널별 성과 분석 기능 - 경쟁사 활동 변화 분석 기능

### III. 결 론

본 논문에서는 공공 및 민간 영역에서 데이터 드리프트 관리 기술의 적용이 시급하고 파급효과가 큰 실증 영역을 설정하고, 각 모듈이 독립적으로 기능을 수행하면서도 서비스 간의 유기적인 연계를 통해 데이터 드리프트 문제를 신속히 감지하고 해결할 수 있는 마이크로서비스 아키텍처 기반의 데이터 드리프트 관리 기술을 제안하였다. 아울러 각 모듈은 데이터 드리프트를 감지, 분석, 대응 등 필요한 기능을 분리하여 수행하는데 이는 다양한 시나리오에서 시스템의 확장성과 유연성을 극대화할 수 있을 것으로 기대된다.

### 참 고 문 헌

[1] Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G., "Learning under concept drift: A review". IEEE transactions on knowledge and data engineering 31.12, 2346-2363, 2018.

[2] 이예은, 이태진. "MLOps를 위한 효율적인 AI 모델 드리프트 탐지방안 연구." 인터넷정보학회논문지 24(5), 17-27, 2023.

[3] 신성환, 박현찬, 김영필. "컨테이너 기반 마이크

로서비스 모니터링 주요 사례 연구". 전자공학회논문지 61, no.9, 37-45. 2024.

[4] Alshuqayran, Nuha, Nour Ali, and Roger Evans. "A systematic mapping study in microservice architecture". 2016 IEEE 9th international conference on service-oriented computing and applications (SOCA), IEEE, 2016.

[5] 김성운, "온프레미스 스토리지와 퍼블릭 클라우드 스토리지간 데이터 통합 관리 및 신뢰성 보장 기술 개발", 한국전자통신연구원, 2020.

# 데이터 드리프트 관리 알고리즘을 통한 조작된 영상데이터 영역 검출에 대한 고찰

이성희\*, 황재훈\*, 윤철희\*

## Analysis of Data Drift Management Algorithms for Detecting Manipulated Areas in Video Data

Sung hee Lee\*, Jaehun Hwang\*, and Cheol Hee Yoon\*

### 요 약

AI 기술의 발달에 따라 유명인 또는 지인을 사칭한 딥페이크 범죄가 발생하고 있고 그 빈도와 피해가 지속적으로 증가할 것으로 예상된다. 그동안 딥페이크를 판별하기 위한 다양한 연구들이 있었으며 많은 성과가 있었지만, 다양한 단서와 특징을 기반으로 가짜 콘텐츠를 식별하지만 관련된 변화에 대응하기 위해 지속적인 데이터 학습과 보정이 필요하기 때문에 본 논문에서는 데이터 드리프트 관리알고리즘을 통해 딥페이크 검출 기술과 딥페이크(Deepfake) 탐지 기술의 지속적 개선 방안을 연구하고자 하였다.

### Abstract

With the development of AI technology, deepfake crimes impersonating celebrities or acquaintances are occurring, and the frequency and damage are expected to continue to increase. There have been various studies to identify deepfakes and many achievements have been made. However, although fake content is identified based on various clues and characteristics, continuous data learning and correction are required to respond to related changes. Therefore, this paper proposes a data drift management algorithm. Through this, we attempted to study deepfake detection technology and ways to continuously improve deepfake detection technology.

### Key words

data drift detection, machine learning, open-source libraries for drift detection

## 1. 서 론

최근의 인공지능은 대규모언어모델(LLM, Large-scale Language Model)을 기반으로 한 ChatGPT와 1980년대부터 개발되어 온 적대적 이미지 생성

\* 경찰대학교, dolcei6966@police.ac.kr

\*\* 경찰대학교, kbummyoon@police.ac.kr

\*\*\* 경찰대학교, bertter@police.ac.kr

※ 이 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. RS-2024-00337489, 분석 모델의 성능저하 극복을 위한 데이터 드리프트 관리 기술 개발)

모델(GAN, Generative Adversarial Network)이 생성형 AI 큰 틀을 마련하고 있으며, 2023년 이후부터는 ‘AI 대변혁기’라고 할 수 있을 만큼 하루가 다르게 다양하고 새로운 기술이 개발되고 있다. 특히, 대규모 데이터(Hyper-scale Data) 학습을 통해 인간의 영역이라고 할 수 있는 창작의 영역까지 진화하는 시점이다.

본 논문은 최신 생성형 AI 기술 기술에 대응하기 위해 최근에 떠오르고 있는 국내외 데이터 드리프트 대응 관리 알고리즘 분석을 통해, 영상 왜곡에 대한 기술 조사 분석과 취약점 분석, 2차 영상 처리를 위한 영상 개선 기술 관련 조사와 조작된 영상 검출 방향을 고찰하고자 하였다.

## II. 본 론

### 2.1 이미지 생성형 AI

2014년 GAN(적대적 생성 신경망, Generative Adversarial Network)의 등장으로 생성형 AI가 인기를 끌기 시작했다. GAN은 생성하는 모델과 이를 판단하는 모델이 서로 경쟁하면서 실제와 가까운 이미지나 동영상, 음성등을 자동으로 만들어 내는 모델이다. GAN은 생성 모델의 성능이 좋고 학습데이터가 적어도 생성 모델을 활용하여 보충이 가능하며 분류 모델을 더 강인하게 학습할 수 있다. GAN을 이용한 영상은 생성자와 원본을 구분할 수

없을 정도로 정교한 데이터를 생성할 수 있는 능력을 갖추었다. GAN은 이미지 생성 분야에서 가장 혁신적인 아이디어로 꼽히고 있다[1].

이미지 생성형 AI 모델은 오픈 AI(OpenAI)의 달리(DALL-E), 구글의 딥드림, 미드 저니(Midjourney), 스테이블 디퓨전(Stable Diffusion) 등이 있다. 현재 이미지 생성형 AI는 예술(그림, 영상, 디자인, 문학), 공학(프로그래밍, 생명공학, 화학, 금융공학 등), 커뮤니케이션(음성, 문자 등) 등의 다양한 영역에서 AI 기술을 접목해 활용하고 있으며, 앞으로도 관련 서비스가 더욱 발전하고 있다[1].

### 2.2 영상 왜곡 감지용 데이터 드리프트 관리 알고리즘

딥페이크(deepfake)란 실제로 행동하지 않았거나 말하지 않은 것을 한 것처럼 조작한 이미지 또는 영상을 말하며, 딥페이크(deepfake)는 "deep learning"과 "fake"의 합성어로, 인공지능 기술을 사용하여 현실에 존재하지 않는 것이나 발생한 적이 없는 사건을 이미지, 비디오, 오디오 등을 합성하여 만들어 내는데, 주로 Multi-attentional Deepfake Detection를 통해 여러 공간적 주의집중 헤드를 사용하여 다양한 로컬 부분에 집중하고 미세한 아티팩트를 확대하는 텍스처 특징 강화 블록을 이용한 새로운 다중 주의집중 딥페이크 감지가 가능하다. [2]

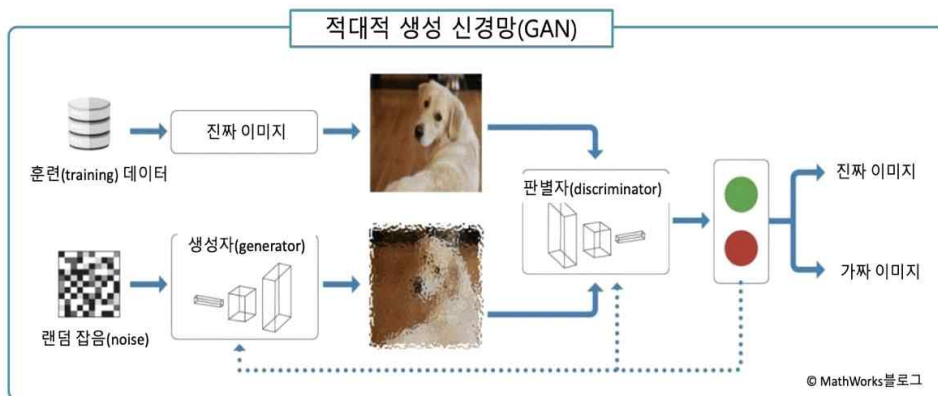


그림 1. 적대적 생성 신경망의 기본 구조  
Fig. 1. Generative Adversarial Network process

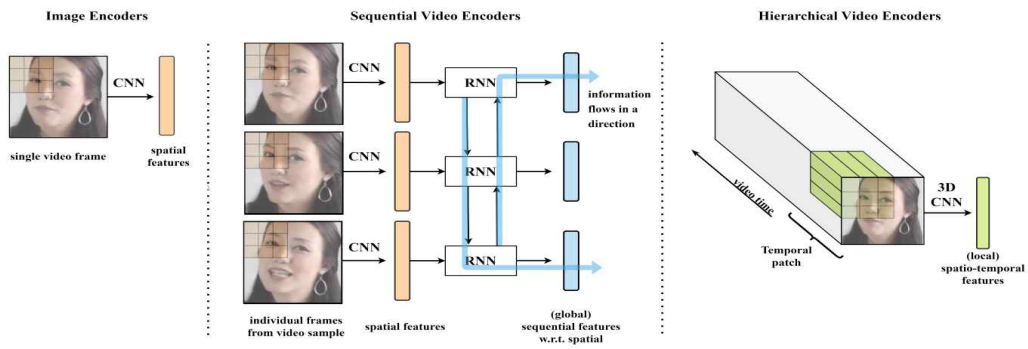


그림 2 이미지 비디오 인코더 차이점  
Fig. 2. Image Video Encoder Differences

[그림 2]는 이미지 인코더, 순차 비디오 인코더 및 계층적 비디오 인코더의 차이점을 설명한 그림. 주황색과 초록색 사각형은 2D 및 3D 커널의 사용. 기술은 오락적인 목적으로 활용될 수 있으나, 악의적인 정보 조작에도 사용될 위험이 있다. 현재 대부분의 딥페이크 탐지 방법은 이미지 기반의 분류기 (Convolutional Neural Networks, CNNs)를 이용해 단일 이미지에서 딥페이크 특성을 탐지하는 방식에 의존하고 있지만, 시공간적(spatio-temporal) 특징을 활용한 3D CNN 모델을 통해 [2]. 딥페이크 탐지 모델이 새로운 방법에 대해 더 잘 일반화되려면 다양한 딥페이크 알고리즘과 과제를 포함한 학습 데이터가 필요하며, 이때 데이터드리프트 관리 기법이 수행된다. 특정 유형의 딥페이크만 학습할 경우, 새로운 유형의 딥페이크가 등장했을 때 탐지 성능이 크게 저하될 수 있기 때문에 다양한 딥페이크 알고리즘의 학습이 반드시 필요하다[2].



그림 3. R3D 모델과 Grad-CAM 기법을 활용  
Fig. 3. Utilizes R3D models and Grad-CAM techniques

### a. 딥페이크 검출 기술: 구조적인 리뷰

- **데이터 수집 (Data Collection)** 딥페이크 탐지의 첫 번째 단계는 원본 데이터와 딥페이크 데이터를 수집하고 정리하는 것이다. 이 과정에서는 가공되지 않은 원본 이미지나 비디오뿐만 아니라 딥페이크로 조작된 데이터를 확보한다. 이러한 데이터를 통해 탐지 모델이 학습하고, 원본과 딥페이크를 구분하는 능력을 배양할 수 있게 된다. **얼굴 탐지 (Face Detection)**: 수집된 이미지나 비디오에서 얼굴을 탐지하는 과정이다. 이 단계에서는 얼굴의 위치를 식별하여 나이, 성별, 감정 등의 특징을 분석할 수 있는 영역을 찾아낸다. 얼굴 탐지 알고리즘을 통해 비디오나 이미지에서 주목해야 할 부분을 파악하며, 딥페이크와 관련된 특정 표정을 추출하는 데 사용된다. **특징 추출 (Feature Extraction)**: 얼굴 탐지가 완료되면, 탐지기(detector)를 위한 다양한 특징들을 얼굴 영역에서 추출한다. 이 단계에서 추출된 특징은 후보 특징들로, 나중에 딥페이크 탐지 모델에 사용될 수 있다. 여기에는 얼굴의 형태, 텍스처, 색상 등의 물리적 특징이 포함되며, 딥페이크 영상의 인공적인 흔적을 찾기 위한 중요한 정보로 활용된다. **모델 선택 (Model Selection)**: 적절한 분류 모델을 선택하는 단계다. 이 단계에서는 딥러닝 기반 모델, 머신러닝 기반 모델, 통계 모델 등 여러 모델들 중에서 탐지 성능이 가장 좋은 모델을 선택한다. 딥페이크 탐지에 자주 사용되는 모델로는 CNN(Convolutional Neural Networks), RNN(Recurrent Neural Networks) 등이 있으며, 데이터의 특성과 목

적에 맞게 최적의 모델을 선택한다. **모델 평가 (Model Evaluation):** 마지막 단계에서는 선택된 모델을 다양한 성능 평가 지표를 사용하여 평가한다. 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1 스코어(F1 Score) 등의 지표를 통해 모델이 딥페이크를 얼마나 효과적으로 탐지하는지 평가하고, 이를 바탕으로 모델을 개선한다[3].

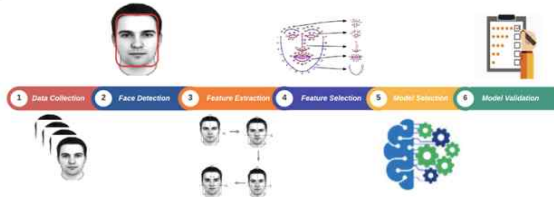


FIGURE 6. Steps of Deepfake detection.

그림 4. 딥페이크 검출 단계  
Fig. 4. Deepfake detection steps

b. 딥페이크(Deepfake) 탐지 기술의 분류 체계

매체에 따른 분류 (Media)는 딥페이크 탐지 알고리즘은 처리하는 매체에 따라 분류될 수 있다. 탐지 대상이 이미지인지 비디오인지, 혹은 이미지와 비디오를 동시에 처리할 수 있는지에 따라 다음과 같이 나뉜다. 이미지 기반 탐지(Image-based Detection): 단일 이미지에서 딥페이크 여부를 탐지하는 방법이다. 비디오 기반 탐지(Video-based Detection): 연속적인 프레임을 처리하여 시간적 정보를 포함한 딥페이크를 탐지하는 방법이다. 또한, 이미지 및 비디오 기반 탐지 (Image and Video-based Detection)를 통해 이미지와 비디오 데이터를 모두 탐지할 수 있다. 사용된 특징(Features)에 따른 분류는 탐지 알고리즘이 사용한 특징에 따라 분류할 수 있다. 딥페이크 탐지에 사용되는 대표적인 12가지 특징들이 있으며, 이러한 특징은 얼굴의 형태, 텍스처, 색상, 움직임의 불일치 등과 같은 물리적 또는 디지털적 단서를 포함한다. 탐지 모델은 이러한 특징을 기반으로 딥페이크 여부를 판단한다. 탐지 방법에 따른 분류 딥페이크 탐지 방법은 사용된 알고리즘의 종류에 따라 분류된다. 이들 중 주요 탐지 방법은 다음과 같다.

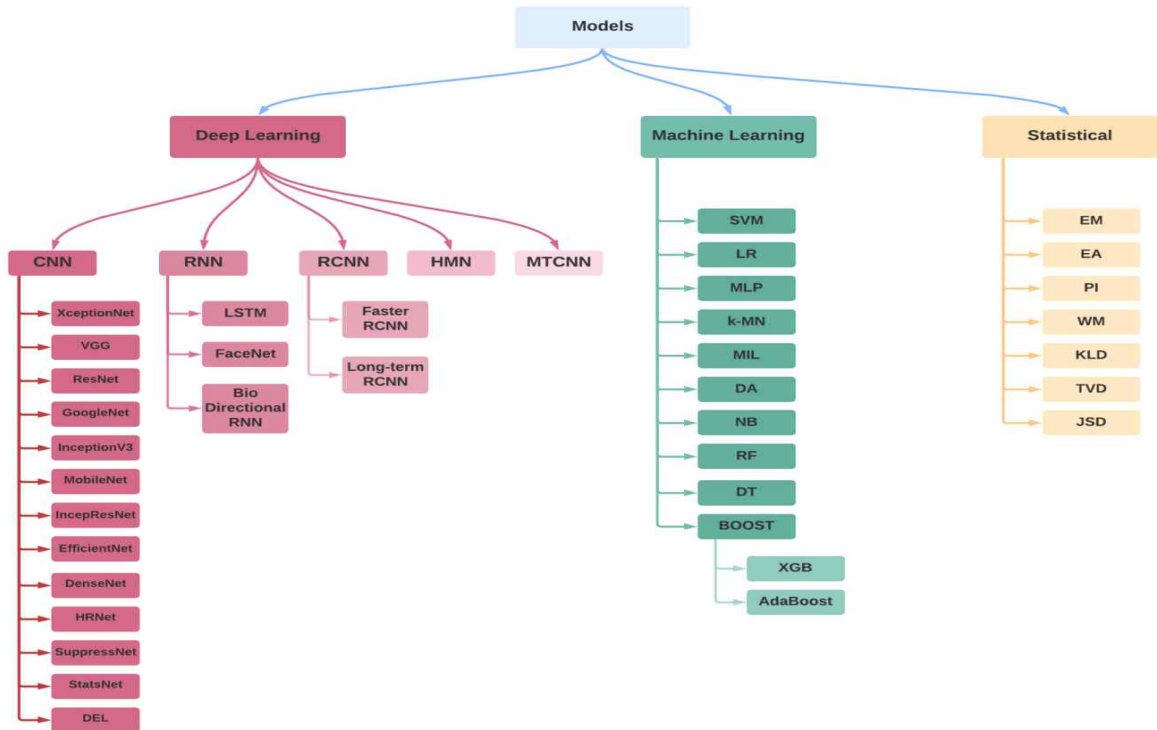


FIGURE 8. The list of Deepfake detection models.

그림 5. 딥페이크 검출 방법들  
Fig. 5. Deepfake detection steps



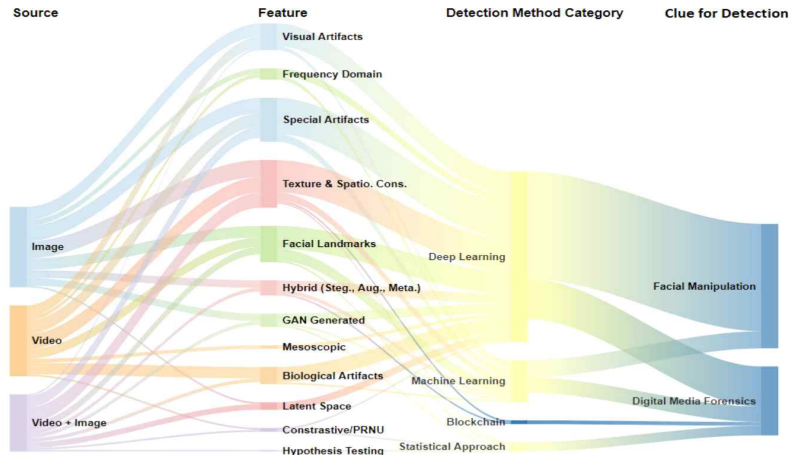


그림 6 딥페이크(Deepfake) 탐지 기술의 분류 체계  
 Fig. 6. Classification of Deepfake Detection Technologies

딥러닝 기반 방법으로 CNN, RNN과 같은 딥러닝 모델을 활용한 탐지 방법이다. 그리고, 머신러닝 기반 방법은 전통적인 머신러닝 알고리즘을 사용하여 특징을 분류하는 방법이다. 또한, 탐지 단서에 따른 분류 (Clue for Detection)를 통해 탐지 알고리즘은 특정 단서를 기반으로 딥페이크를 탐지한다. 이러한 단서는 얼굴 조작(Facial Manipulation) 즉, 얼굴에서 나타나는 조작된 흔적을 기반으로 탐지한다. 디지털 미디어 포렌식(Digital Media Forensics)은 디지털 미디어의 메타데이터, 압축 아티팩트, 화질의 비일관성 등을 분석하여 딥페이크 여부를 판별하는 방법이다. [4]

딥페이크 탐지 모델이 딥페이크 콘텐츠를 식별하는데 사용하는 단서와 특징들은 딥페이크 비디오의 조작된 부분에서 발생하는 미세한 불일치와 인위적 변화를 기반으로 한다. 이러한 단서와 특징들은 시각적, 시간적 정보를 결합하여 딥페이크 탐지 모델이 학습하고 탐지할 수 있게 한다

### III. 결 론

데이터드리프트 관리 알고리즘이 필요한 딥페이크 탐지 모델은 다양한 단서와 특징을 기반으로 가짜 콘텐츠를 식별하지만 관련된 변화에 대응하기 위해 지속적인 데이터 학습과 보정이 필요하다. 주로 사용된 딥페이크 탐지 모델은 공간적 아티팩트 (Spatial Artifacts), 생리적 신호(Biological and Physiological Signs), 오디오-비주얼 불일치(Audio-

Visual Inconsistencies), 시간적 불일치(Temporal Inconsistencies)와 GAN 생성 흔적(GAN Traces)을 분석하는 방법, 연속된 프레임 간의 불일치나 GAN 생성 과정에서 발생하는 인위적인 흔적을 탐지하여 딥페이크 방법이 있고, 이는 얼굴의 기하학적 정보, 좌우 대칭성의 불일치, 그리고 딥페이크 생성 과정에서 발생하는 업샘플링 흔적, 시각적 아티팩트는 딥페이크 비디오 흔적, 눈 깜박임의 비정상적 패턴 피부색의 미세한 변화(혈류 변화) 또한, 오디오와 영상의 타이밍 차이나 입술 움직임의 불일치 탐지 등으로 판별한다. 이때 사용된 알고리즘은 모두 상황별 데이터드리프트 관리 알고리즘을 통해서 지속적으로 고도화해야 한다.

### 참 고 문 헌

- [1] [ProGAN: How NVIDIA Generated Images of Unprecedented Quality, Sarah Wolf, Toward Data Science, <https://towardsdatascience.com/progan-how-nvidia-generated-images-of-unprecedented-quality-51c98ec2cbd2>
- [2] Ganiyusufoglu, Ipek, et al. "Spatio-temporal features for generalized detection of deepfake videos." arXiv preprint arXiv:2010.11844 (2020).
- [3] RANA, Md Shohel, et al. Deepfake detection: A systematic literature review. IEEE access, 2022, 10: 25494-25513.
- [4] AI-based datalift-based data rescue algorithm analysis technology research service research

# 데이터 드리프트 관리를 통한 로봇전이기술에 대한 연구 - 영상 기술을 중심으로

이도연\*, 윤철희\*\*

## A Study on Robot Transfer Technology through Data Drift Management: Focusing on Video Technology

Do Yeon LEE\*, Cheol Hee Yoon\*\*

### 요 약

인공지능과 로봇 기술의 급속한 발전으로 데이터 기반 학습이 로봇 시스템의 핵심 요소로 자리잡았다. 그러나 시간의 흐름으로 인해 입력 데이터의 특성이 변화하는 데이터 드리프트 현상은 로봇 시스템의 성능을 저하하게 하는 주된 원인이 되고 있다. 본 연구는 로봇 시스템에서 발생하는 데이터 드리프트 현상을 체계적으로 분석하고, 이를 효과적으로 관리하여 로봇의 전이학습 성능을 향상시키는 방법을 제안한다. 특히 비디오 기반 학습을 통한 로봇 조작 기술의 향상과 데이터 드리프트 관리 방안을 중점적으로 연구하였으며, 실제 환경에서의 적용 가능성을 검토하였다. 본 연구의 결과는 로봇 시스템의 지속적인 성능 유지와 향상을 위한 기초 연구로서 의미를 가진다.

### Abstract

With the rapid development of artificial intelligence and robot technology, data-based learning has become a core element of robot systems. However, data drift, which is a phenomenon in which the characteristics of input data change over time, is a major cause of deterioration in the performance of robot systems. This study systematically analyzes the data drift phenomenon occurring in robot systems and proposes a method to effectively manage it to improve the transfer learning performance of robots. In particular, we focused on improving robot manipulation skills through video-based learning and data drift management methods, and examined the applicability in real environments. The results of this study are meaningful as basic research for maintaining and improving the continuous performance of robot systems.

### Key words

SData Drift Detection, Robot Transfer Learning, Video-based Learning

---

\* 경찰대학교, dy19898@naver.com

\*\* 경찰대학교, bertter@police.ac.kr

※ 본 연구는 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No. RS-2023-00225661, 디지털 증거의 증명력 제고를 위한 인과관계 추론 및 표현 기술 개발)

## I. 서론

최근 인공지능과 로봇 기술의 발전은 로봇의 자율성과 적응력을 획기적으로 향상시켰다. 특히 비디오 데이터를 활용한 학습 방법은 로봇이 인간의 행동을 관찰하고 모방하는 능력을 크게 발전시켰다. 생성형 AI 기술의 발전과 함께, 로봇의 학습 능력은 더욱 정교해지고 있으며, 이는 산업 현장에서의 활용 가능성을 크게 높이고 있다. 그러나 실제 환경에서 운용되는 로봇 시스템은 시간이 지남에 따라 입력 데이터의 특성이 변화하는 데이터 드리프트 현상에 직면하게 된다. 이러한 드리프트는 환경 변화, 센서 노후화, 작업 조건의 변경 등 여러가지의 요인들로 인해 발생할 수 있고, 학습된 모델의 성능을 저하시키는 주요 원인이 된다. 데이터 드리프트 현상은 시간에 따른 데이터 분포의 변화를 포함하고, 이러한 변화는 단순한 데이터 분포의 이동뿐만 아니라, 데이터의 품질 저하, 새로운 패턴의 출현, 기존 패턴의 소멸 등 다양한 형태로 나타날 수 있다. 특히 로봇 시스템에서는 이러한 드리프트가 작업 수행 능력의 직접적인 저하로 이어질 가능성이 존재하기에, 이에 대한 관리와 대응이 필수적이다.

본 연구는 로봇 시스템에서 발생하는 데이터 드리프트를 효과적으로 관리하고, 이를 통해 로봇의 전이학습 성능을 향상시키는 것을 목적으로 한다. 특히 비디오 기반 학습에서 발생하는 드리프트 현상에 중점을 두어, 드리프트의 감지 및 분류, 대응 전략 수립, 그리고 이를 활용한 효과적인 전이학습 방법의 개발을 포함한다. 이를 통해 로봇 시스템의 지속적인 성능 유지와 향상을 위한 기반을 마련하고자 한다.

## II. 본론

데이터 드리프트는 시간이 지날 때마다 데이터의 통계적 특성이 변화하게 되는 현상을 일컫는다. 크게 세 가지 유형으로 분류되는데 첫째, 입력 데이터의 분포가 변화하는 공변량 드리프트는 센서의 노후화나 환경 조건이 변화됨으로 인해 발생할 수 있다. 둘째, 입력과 출력 간의 관계가 변화하는 개념 드리프트는 작업의 특성이나 목표가 변경될 때 발

생한다. 셋째, 출력 변수의 분포가 변화하는 사전 확률 드리프트는 작업 환경의 전반적인 변화로 인해 나타날 수 있다. 특히, 로봇 시스템에서 이러한 드리프트는 특히 심각한 문제를 발생시킬 수 있다. 비디오 기반 학습에서는 조명 조건의 변화, 카메라 설정의 변경, 환경의 물리적 변화 등이 복합적으로 작용하여 드리프트를 발생시킨다. 이는 단순한 성능 저하뿐만 아니라, 로봇의 안전한 작동과 직결되는 문제로 이어질 수 있다. 특히 생성형 AI 기술과 결합된 로봇 시스템에서는 이러한 드리프트가 더욱 복잡한 형태로 나타날 수 있어, 체계적인 관리와 대응이 필수적이다. 또한, 로봇의 비디오 기반 학습 경우 다양한 접근 방법을 통해 이루어지고 있는데, 특징 추출 방법에서는 CNN을 활용한 시각적 특징 추출과 자세 추정, 키포인트 감지 등이 핵심적인 역할을 하고 있다. 이러한 기술들은 로봇이 인간의 동작을 이해하고 모방하는데 필수적인 요소이고, CNN과 강화학습의 결합을 통해 로봇이 환경과의 상호작용을 통해 학습하도록 한다. 그리고, 모방 학습 방법은 멀티모달 접근을 통해 더욱 효과적인 학습이 가능하는데 이는 비디오 데이터뿐만 아니라 다양한 센서 데이터를 통합적으로 활용하여, 보다 정교한 동작 학습을 가능하게 한다.

데이터 드리프트 체계적인 관리를 위한 통합 관리 시스템을 고려할 수 있다. 멀티모달 데이터 처리 시스템과 적응형 학습 구조를 포함하여 멀티모달 데이터 처리 시스템은 비디오, 센서 데이터 등 다양한 입력 소스를 통합적으로 처리하여 데이터 드리프트 현상을 포괄적으로 분석 후 대응하는 프로세스를 자동화 할 수가 있다. 이런 자동화 프로세스의 경우 단일 데이터 소스에 의존할 때보다 더 안정적인 데이터 드리프트 감지를 가능하고, 변화에 능동적으로 대응할 수 있다. 또한, 적응형 학습 구조는 데이터 분포의 변화에 동적으로 대응할 수 있는 메커니즘을 제공하기 때문에, 이는 온라인 학습과 점진적 학습 방법을 결합하여 지속적인 모델 갱신을 가능하게 한다. 특히 실시간 데이터 스트림에서 발생하는 드리프트에 대응하기 위해서는 변화 감지와 모델 적응이 동시에 이루어지도록 설계되어야 한다. 데이터 드리프트 관리는 전이학습의 개선을 통해 실시간 모니터링 시스템, 적응형 학습 메커니즘, 그

리고 이상불 기반 접근을 통해 이루어질 수 있다. 즉, 실시간 모니터링 시스템은 데이터 분포의 변화를 지속적으로 관찰하고 분석하여, 드리프트의 조기 감지와 대응을 가능하게 한다. 이는 다양한 통계적 지표와 머신러닝 기반의 이상 감지 알고리즘을 활용하여 구현이 된다.

Quarterly, Vol. 31, No. 2, P234-256, 2023.

### III. 결 론

본 논문을 통해 데이터 드리프트 관리를 통한 로봇 전이학습의 새로운 접근 방법을 제시하려 하였다. 제안된 방법으로 데이터 드리프트 현상에 대한 체계적인 분석과 관리 방안을 제공하며, 로봇 학습 시스템의 성능 향상 가능성을 고찰하였다. 향후 연구에서는 더욱 복잡한 실제 환경에서의 적용성 검증을 통해, 다양한 로봇 플랫폼에 대한 적용성 검증 그리고 새로운 형태의 데이터 드리프트에 대한 대응 방안을 연구할 것이다. 마지막으로 실시간 처리 능력의 향상과 함께, 더욱 효율적인 드리프트 관리 시스템의 개발도 중요한 연구 과제가 될 것으로 기대한다.

### 참 고 문 헌

- [1] C. Eze, C. Crick, "Learning by Watching: A Review of Video-based Learning Approaches for Robot Manipulation," IEEE Robotics and Automation Letters, Vol. 9, No. 2, P45-67, 2024.
- [2] T. R. Hoens, R. Polikar, N. V. Chawla, "Learning from Streaming Data with Concept Drift and Imbalance: A Comprehensive Overview," Journal of Artificial Intelligence Research, Vol. 12, No. 1, P89-101, 2012.
- [3] A. Mallick, S. Lee, H. Wang, "Data Drift Mitigation Strategies in Large-scale Machine Learning Systems," Journal of Machine Learning Applications, Vol. 15, No. 3, P77-94, 2022.
- [4] Y. Yang, C. Farrar, D. Mascarenas, "Advanced Techniques in Video Motion Analysis and Manipulation," International Journal of Digital Imaging, Vol. 24, No. 4, P112-134, 2017.
- [5] K. Park, M. Davis, "Real-time Data Drift Detection in Robotic Systems," Robotics and Autonomous Systems

## 로봇 및 제조 AI모델 개발 및 기술지원

김남규\*, 서동우\*\*, 황규현\*\*\*, 허태상\*\*\*\*, 정대용\*\*\*\*\*

### AI model development and technical support for Robotics and Manufacturing

Nam-gyu Kim\*, Dongwoo Seo\*\*, Gyuhyun Hwang\*\*\*, Tae-Sang Huh\*\*\*\*, Daeyoung Jung\*\*\*\*\*

KISTI는 공공 및 산업분야에서 AI모델 개발을 통한 문제해결을 통해 기술지원을 수행해왔다. 특히 산업분야에서는 스마트 팩토리의 핵심장비인 산업용 로봇의 예지보전 및 제조분야의 기본장비인 CNC머신의 절삭공구 마모예측에 대한 AI모델을 개발하였다. 우선 산업용 로봇은 자동차 제작 시 많이 수행되는 실링공정을 대상으로 로봇의 각 축의 RV감속기에 음향 센서를 부착, 센서 데이터를 수집 후 딥러닝 모델을 개발하였다. 또한 RV 감속기에 대해 구조동역학 시뮬레이션을 수행 후 이를 기반으로 4개 고장모드에 대한 차수감수모델을 개발하였다. 이 AI 모델과 차수감수모델을 연계할 수 있도록 플랫폼에 탑재하여 산업용 로봇의 이상을 탐지하는 기술을 개발하였다. CNC머신의 경우 공구 마모가 생산성과 품질에 중요한 영향을 미치므로, CNC머신의 상태 데이터를 활용하여 공구 마모를 예측하는 AI모델을 개발하였다. 이를 위해 CNC머신 데이터 수집하고 전처리한 후 피쳐 엔지니어링을 수행하였다. 그 후 다양한 머신러닝 앙상블 모델로 학습한 후 각 모델별로 예측성능을 측정하였다. 최종적으로 XGBoost 모델이 가장 높은 예측 정확도를 달성하였다.

#### Key words

artificial intelligence, robotics, manufacturing, machine learning

#### 참고문헌

- [1] M. Berg, "Preventive maintenance policy for units subject to intermittent demand," Operations Research, Vol. 32, pp. 584-595, June 1984.
- [2] D. Hartmann, M. Herz, U. Wever, "Model order reduction a key technology for digital twins," Reduced Order Modeling for Simulation and Optimization, pp. 167-179, 2018.
- [3] F. Klocke, T. Krieg, "Coated Tools for Metal Cutting - Features and Applications," CIRP Annals, Vol. 48, No. 2, pp. 515-525, 1999.
- [4] T. Chen, C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794, 2016.

---

\* 한국과학기술정보연구원, ssgyu@kisti.re.kr, \*\*한국과학기술정보연구원, seodongwoo@kisti.re.kr(교신저자),  
\*\*\* 한국과학기술정보연구원, ghhwang@kisti.re.kr, \*\*\*\*한국과학기술정보연구원, tshuh@kisti.re.kr, \*  
\*\*\*\* 한국과학기술정보연구원, daeyoungjung@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24L2M1C5, (NTIS)2710018523)

# 공공분야에서 디지털트윈 기반 AI모델 개발 및 기술지원 사례

서동우\*, 김남규\*\*

## Digital Twin-Based AI Model Development and Technical Support in the Public Sector

Dongwoo Seo\*, Nam-gyu Kim\*\*

본 논문은 공공분야에서 디지털트윈 기반의 딥러닝 기술 개발과 지원 사례를 연구하는데 목적이 있다. 특히, 본 연구에서는 에너지, 안전, 수산자원 관리 등의 공공분야에서 딥러닝 기술을 활용한 디지털트윈 기술과 사례를 다룬다. 에너지 분야에서는 진동수주형 파력발전소 에너지 효율증대를 위한 진동수주실 수위 예측, 원전 취수구 해양생물 모니터링을 통한 원전 중단 예방을 위한 생물 밀집도 분석 기술을 연구하였으며 해양 안전분야에서는 대형 트롤어선에서의 조업자 안전성 증대를 위한 영상 및 음성 기반의 위험 방지 기술을 연구하였다. 마지막으로 해양수산자원 관리 분야에서는 정치망 조업에서 어류 등의 수산자원의 효율적인 관리를 위한 불륨 기반의 조업량 참다량어 추정 기술을 연구하였다. 해당 기술을 개발하기 위해서는 데이터 및 사례 확보가 관건인데 이를 해결하기 위해 현장과 긴밀하게 협력하여 다양한 센서 및 장비를 설치하고 운영함으로써 양질의 데이터를 확보하였으며 이를 통해 현장에서 실 운영 가능한 딥러닝 모델을 연구 및 개발하고 실증하였다. 특히, 본 연구에서는 가상현실로 다양한 환경과 상황을 모델링 및 시뮬레이션하여 원전, 해양 등의 특수한 환경 특성상 획득 및 재현이 어려운 데이터를 생성하는 기술로 인공지능 모델의 정확도를 향상시켰다. 추후에는 본 연구의 기술 개발 결과와 사례를 바탕으로 딥러닝 모델을 지속적으로 개선하고자 한다.

### Key words

Digital Twin, Deep Learning, Virtual Reality

### 참고문헌

- [1] Dongwoo Seo, Taesang Huh, Myungil Kim, Jaesoon Hwang and Daeyong Jung, 2021. Prediction of Air Pressure Change Inside the Chamber of an Oscillating Water Column-Wave Energy Converter Using Machine-Learning in Big Data Platform, *Energies*, 14(11), 241–247, 2016
- [2] 김남규 서동우 허태상 황규현 손영진, 이미지/영상 처리 인공지능 기술 활용의 현황과 전망, KISTI 이슈브리프 제 73호, <https://repository.kisti.re.kr/bitstream/10580/19181/1/KISTI%20%EC%9D%B4%EC%A%88%EB%B8%8C%EB%A6%AC%ED%04%84%20%EC%A0%9C73%ED%08%B8.pdf>

---

\*한국과학기술정보연구원, seodongwoo@kisti.re.kr, \*\*한국과학기술정보연구원, ssgyu@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24L2M1C5, (NTIS)2710018523)

# HEMOS-Cloud 및 NEURON을 이용한 인공지능 모델 개발 환경 구축

황규현\*, 정대용\*\*, 김남규\*\*\*

## Building an AI Model Development Environment using HEMOS-Cloud and NERON

Gyuhyun Hwang\*, Daeyong Jung\*\*, and Namgyu Kim\*\*\*

본 논문에서는 인공지능 모델 개발에 필수로 요구되는 데이터 수집 및 전처리, 성능 테스트 및 안정화, 컨테이너 기반 다중 GPU 분산 환경 등의 기능[1, 2]을 효율적으로 지원하기 위해 구축한 인공지능 개발 환경에 관해 설명한다. 효과적인 인공지능 모델 개발 환경을 구축하기 위해 자체 보유 계산 자원인 HEMOS-Cloud[3]와 NEURON을 데이터 가공과 학습 및 검증 등에 활용하였고, 컨테이너 기반 개발 환경 지원을 위해 HEMOS-DLP를 사용하였다. 고성능 GPU가 탑재된 클라우드 시스템인 HEMOS-Cloud는 주로 인공지능 모델의 초기 개발과 성능 테스트 등에 사용되며, 다양한 형태로 저장된 방대한 크기의 학습 데이터를 효과적으로 저장하고 관리할 수 있도록 웹 인터페이스 기반 모니터링 및 관리 기능을 지원한다. HEMOS-Cloud에 수집된 데이터는 HPC 자원인 NEURON에서 학습 가능한 형태로 전처리·가공·검증 할 수 있으며, 수집된 데이터의 이상치(outlier) 탐지 및 제거 과정 수행을 통해 학습용 데이터의 품질을 개선할 수 있다. HEMOS-DLP를 이용해 인공지능 모델 개발에 필요한 다양한 버전의 딥러닝 프레임 워크를 간편히 구성할 수 있는 기능과 멀티 GPU를 사용한 분산 학습 및 대규모 데이터의 전처리가 가능하도록 환경을 구축하였다. HEMOS-Cloud를 활용해 개발된 인공지능 모델을 고성능 계산 자원인 NEURON에서 대규모 데이터 학습 및 검증과 하이퍼 파라미터 조정 (hyper parameter tuning)을 할 수 있도록 개발 환경을 설계하였고, 이를 이용해 개발된 인공지능 모델의 고도화와 성능 최적화를 할 수 있도록 하였다.

### Key words

artificial intelligence, ai platform, high performance computing

### 참고문헌

- [1] L. E. Lwakatare, I. Crnkovic and J. Bosch, "DevOps for AI - Challenges in Development of AI-enabled Applications," 2020 SoftCOM, pp. 1-6, 2020.
- [2] 클라우드와 AIaaS 서비스 동향, "클라우드와 AIaaS 서비스 동향", ETRI Insight, 2022년 9월.
- [3] HEMOS-Cloud, <https://hemos-cloud.kisti.re.kr>, 2024.

---

\*한국과학기술정보연구원, ghhwang@kisti.re.kr, \*\*한국과학기술정보연구원, daeyongjung@kisti.re.kr, \*\*\*한국과학기술정보연구원, ssgyu@kisti.re.kr(교신저자)

※ 이 논문은 2024년도 한국과학기술정보연구원의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24L2M1C5, (NTIS)2710018523)

# 인공지능 온습도 예측모델을 통한 조선산업의 이산화탄소 저감 연구

이재섭

## Study on Carbon Dioxide Reduction in the Shipbuilding Industry Using an AI-Based Temperature and Humidity Prediction Model

Jae-Sum Lee

본 기술보고서는 지구 온난화로 인한 기후 변화가 조선산업의 각종 공조설비 가동과 압축에어 생산을 위한 에너지 소비 증가와 이산화탄소 배출을 가중시키는 상황에서, 조선소의 환경에 최적화된 에너지 관리 솔루션을 제안한다. 도장 공정, VOCs 저감 설비, 압축에어 콤프레셔 가동으로 인해 발생하는 연간 수십만톤의 이산화탄소 배출을 감소시키기 위해, 인공지능 기반의 온습도 예측 모델을 설계하고 이를 적용하여 에너지 사용을 최적화하고 배출량을 줄이는 방안을 모색했다.

울산 지역의 특정 조선소를 대상으로 한 사례 분석을 통해 제습기, 가스히터, 콤프레셔 등의 주요 설비가 연간 약 129,000톤의 이산화탄소를 배출하고 있음을 확인하였고, 친환경 도료 사용 증가와 고내화학적 특수 도장 공법으로 인해 온습도 및 이슬점 관리를 위한 에너지 사용이 더욱 증가하고 있는 실정이다. 이를 해결하기 위해, 본 보고서는 조선소의 혹독한 작업 환경에서도 견딜 수 있는 방폭형 센서와 안정적인 통신망을 도입하여 실시간 온습도와 이슬점 데이터를 확보하는 방안을 제안한다.

수집된 데이터를 기반으로 인공지능 예측 모델을 적용하여 조선소 내 도장공장, Cargo Hold, 압축 에어배관의 온습도와 이슬점을 예측하고, 최적의 에너지 사용 전략을 도출함으로써 주요 설비의 에너지 소비를 효율화할 수 있는 방안을 검토하였다. 본 기술보고서에서 제안하는 접근 방식은 공장 에너지 관리시스템(FEMS), 모델 예측 제어(MPC), 클라우드 기반 에너지 관리 시스템(EMS) 등 선행 연구의 성과를 바탕으로 조선산업에 특화된 솔루션을 설계하고, 실질적인 에너지 절감 및 이산화탄소 배출 절감 효과를 입증하는 것을 목표로 한다.

본 연구의 결과는 조선산업뿐만 아니라 대규모 제조 공정에서 온습도 기반의 에너지 절감 기술을 실현하기 위한 실용적 방안을 제시하며, ESG 경영의 실천을 통한 친환경 제조공법 도입의 필요성을 강조한다.

### Key words

big data, prediction, CO<sub>2</sub>, FEMS, MPC, EMS, ESG

### 참고문헌

- [1] Seok-Ju. Lee and Van. Quan. Dao., "Energy-Efficient Operation Simulation of Factory HVAC System based on Machine Learning", Journal of Korea Society of Industrial Information Systems, Vol. 29, pp. 47-54, April 2024.
- [2] Hyo-Bok. Lee., "The Improvement of Steam Energy-Efficiency through FEMS based on IT", pp. 29-57, December 2013.
- [3] Jong. Hoon. Sung and Yeong. Sik. Cho., "Machine Learning Approach for Pattern Analysis of Energy Consumption in Factory", KIPS Trans. Comp. and Comm. Sys, vol. 8, pp. 87-92, April 2019.

---

\* HD현대미포, 336688@hmd.co.kr



# 계측시스템의 AI 적용을 통한 성능 개선

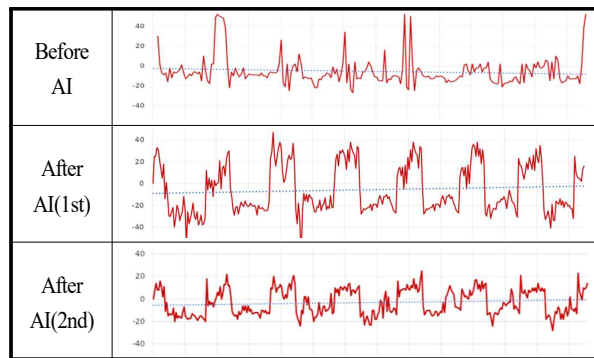
김재현\*, 최우석\*\*, 백승훈\*\*\*

## Improving through AI application in measurement systems

Kim Jae-Hyun\*, Choi Woo-Suk\*\*, and Baek Seung-hoon\*\*\*

본 연구는 시설분야와 다양한 산업군에서 다양하게 활용되고 있는 경사나 거리, 형태를 측정하고 시간의 변화에 따른 수치의 변화를 측정하는 계측기술에 대한 연구로서 과거의 손계측에서 전기전자방식의 부착계측기가 활용되고 있으며 최근에는 비부착, 비접촉 방식의 계측방식이 산업에 다양한 형태로 운용 중이다. 디지털기술의 발전으로 레이저, 초음파, 이미지 등은 이미 산업군에서 다양하게 적용되어 시설물의 안정성 평가를 위한 변위 계측 방식으로 이미지 데이터를 분석하는 비접촉방식의 계측기를 개발하여 4년여 운용하였다. 대상 계측기는 1m 거리에서 0.1mm미만(0.04mm)의 분해능을 가지는 것으로 시험기관을 통하여 성능을 검증하였다. 성능확인 후 계측현장에서 계측기를 설치하고 이미지 분석 과정을 수년간 진행한 결과, 일정한 조도환경에서 성능은 실내환경의 결과를 보이고 있으나 일출, 일몰, 야간 등의 조도저하 또는 변화되는 시간대의 수집 이미지에서 빛에 의한 왜곡과 안개시에는 대상 이미지의 인식을 저하로 엣지탐지(edge detection)에 어려움이 확인되었다. 범용의 파이썬(Phthon)으로 개발된 계측 소프트웨어는 현장 운용 기간동안 조도의 차이에 대해서 일부 개선하였으나, 근본적으로 조도변화와 이미지의 훼손에 대해 개별적 보정이 현실적으로 불가능하였다.

최근 수집된 자료에 대한 딥러닝(Deep learning)을 실시하고 학습된 AI를 적용할 경우 실외의 다양한 환경변화에 대응이 가능할 것으로 판단하여 일반 주야간의 이미지 6286건과 안개(Fog)이미지 226건을 생성하여 에포크(Epoch) 20으로 학습을 진행하였다. 또한 학습손실(Training Loss)과 검증손실(Validation Loss)로 AI의 적정성을 확인하고 평가지표(Metrics)등을 통하여 정밀도를 확인하였다. 계측시스템에 AI를 적용한 결과, 일반 이미지의 학습으로 이미지 데이터의 특이값의 감소를 확인하였고, 추가 안개(Fog)이미지 학습을 통하여 이미지 판독과 엣지탐지(edge detection)를 위한 데이터의 상하한 편차감소를 확인하였다.



Key words : AI, deep learning, non-attachment, non-contact measurement technology

\* 시설물안전연구원(주) 본부장, isriofficemail@gmail.com(교신저자), \*\* 연구소장, \*\*\* 시설물안전연구원(주) 연구원

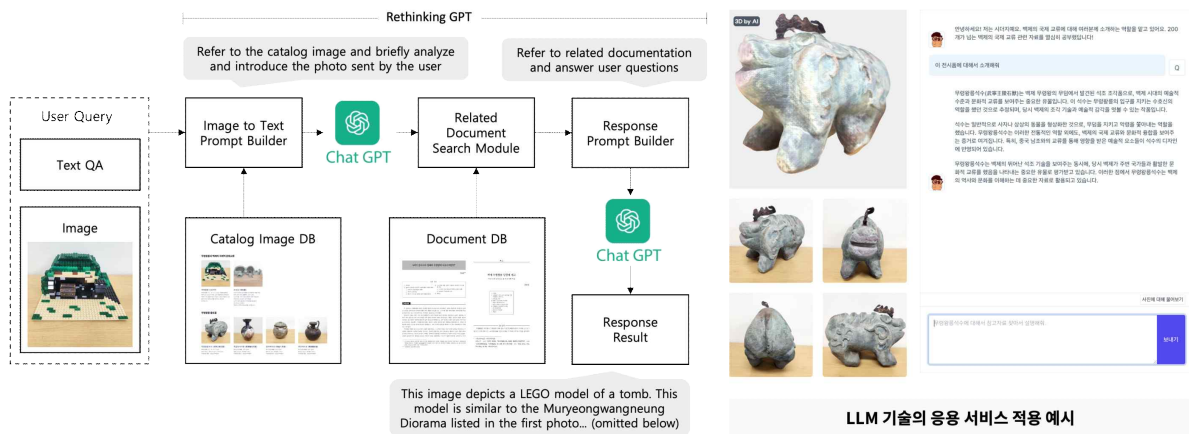
# 산업 분야 적용을 위한 Lowcode 기반의 LLM 최적화 방법

채영훈\*, 김채린\*\*, 정초롱\*\*\*

## Low-Code Based LLM Optimization Methods for Industrial Applications

Yeonghun Chae\*, Chaerin Kim\*\*, Chorong Jeong\*\*\*

본 연구는 대규모 언어 모델(LLM)의 산업 분야 적용을 위한 구조적 접근을 제안하며, Low-code 기술인 DIZEST를 활용해 복잡한 로직이 포함된 LLM 시스템을 효과적으로 관리하고 사용자 친화적 인터페이스를 제공하는 방법을 탐구한다. 기존 LLM의 응용 시스템 구축 시 발생하는 높은 기술적 복잡성과 프롬프트 엔지니어링의 난이도를 해결하기 위해, 본 연구에서는 DIZEST를 활용한 직관적인 모델 관리 방식을 도입하였다. DIZEST는 복잡한 정책 구조와 데이터 처리를 요구하는 다양한 산업 분야에서 LLM을 활용할 수 있도록 지원하며, LangChain과 RAG(Retrieval-Augmented Generation) 기술을 결합하여 대화형 서비스의 응답 신뢰도를 높이는 것을 목표로 한다. 연구의 주요 결과는 LLM 시스템의 구축 및 응용 시 발생하는 복잡한 프로세스를 Low-code 방식으로 간소화함으로써, 비전문가도 모델의 설정 및 조작이 가능하도록 지원하는 것이다. 또한, 이미지와 텍스트 정보를 결합하여 사용자 맞춤형 답변을 생성하는 응답 모듈을 설계하여, 실시간으로 산업 환경의 요구에 맞춘 정보를 제공하는 인터페이스를 구현하였다. 본 연구에서 제안하는 LLM 응용 시스템은 문서 데이터베이스와 이미지 카탈로그를 연결하여, 사용자가 제출한 이미지를 분석하고 관련 정보를 제공하는 프로세스를 포함한다. 이를 통해 사용자는 텍스트 기반 질의응답 외에도 시각적 자료와의 통합된 응답을 받을 수 있어, 실제 산업 현장에서 LLM의 실용성을 크게 향상시킬 수 있다.



Key words : LLM, LangChain, Retrieval-Augmented Generation (RAG), Low-code AI solution, Industrial Application

\*주식회사 시즌, proin@season.co.kr, \*\*주식회사 시즌, cherry@season.co.kr, \*\*\*주식회사 시즌, jeongcr@season.co.kr

※ 본 연구는 2024년도 중소벤처기업부의 “창업성장기술개발사업(TIPS)” 지원을 받아 수행되었음(RS-2024-00441746)

# 실시간 설비 데이터 계측 시스템 구성 및 설계에 대한 연구

조정훈\*, 이준영, 성시영, 이신혜, 이선희\*\*

## A Study on the Configuration and Design of Real-time facility data measurement system

Jung-Hoon Cho\*, Joon-Young Lee, Si-Young Sung, Shine-Hye Lee, and Sun-Hwi Lee\*\*

### 요 약

본 논문에서는 회전기계에서 발생할 수 있는 고장에 대한 안전진단 사례 연구이다. 시간 기반 유지보수(TBM) 방법은 과잉 정비가 될 확률이 높고, 정기보수를 위한 설비 정지는 생산성이 줄어드는 문제가 발생한다. 이러한 문제를 해결하기 위해 산업 현장에서는 사물 인터넷(IoT) 기술을 통해 설비의 상태를 모니터링하고 수집된 데이터를 이용해 상태 기반 정비를 필요로 한다. 인공지능 기술의 급격한 발전으로 상태 기반 유지보수(CBM) 방법이 주목받고 있으며 누적된 데이터를 이용해 학습한 인공지능 모델은 산업계에 도입되고 있다.

### Abstract

In this paper, we study case for rotary machine with a breakdown analysis. The time-based maintenance(TBM) method is likely to be over-maintained, and facility suspension for maintenance causes a problem of reduced productivity. To solve the problem, industrial sites require monitoring the status of facilities through Internet of Things(IoT) technology and status-based maintenance using collected data. Due to the rapid development of artificial intelligence technology, condition-based maintenance(CBM) methods are attracting attention, and Artificial Intelligence models learned using accumulated data are being introduced into the industry.

### Key words

Rotary machine, TBM, Maintenance, Anomaly detection, IoT, Artificial Intelligence, CBM

## 1. 서 론

제조, 발전, 충전 산업 등에서 사용되는 모터, 펌프, 팬, 기어박스, 컴프레셔와 같은 설비는 예상하지 못한 고장으로 인적/물적 피해가 발생할 수 있다. 대상 설비에서 돌발 고장에 의한 가동 정지가 발생

하기 전에 정비 혹은 교환을 해야 인적 또는 물적 피해를 막을 수 있다. 대상 설비의 이상은 사전에 발견하기 어렵고, 고장이 발생하여도 원인을 찾기가 매우 어려워 적절한 조치까지 많은 시간이 소요되며, 생산 또는 설비의 가동 정지에 따른 막대한 손실이 발생하게 된다.

\*퓨처메인(주), jhcho@futuremain.com, \*\*퓨처메인(주), shl@futuremain.com

※ 이 연구는 2024년 산업통상자원부 및 한국산업기술기술평가원(KEIT) 연구비 지원에 의한 연구임(RS-2024-00444913)



그림 1. 설비 가동 중단으로 인한 생산성 손실  
Fig. 1. Loss of productivity due to disruption of facilities

기계장치는 항상 안정적이고 모든 기능이 정상적인 상태에서 운전되는 것이 바람직하나 시간의 경과에 따라 마모에 의한 열화가 일어나고 운전 중에 발생하는 기계적, 열적 스트레스에 의해 결함을 일으키게 된다. 시간 기반 유지보수(TBM) 방식은 주기를 정하고 실행하는 시간 기준 보전이라 동일 기기에는 같은 보전 기준에 따라 점검과 수리를 시행한다. 동일한 설비의 플랜트에서는 신뢰성이 높고, 위험도도 최소화되지만 보전 비용이 많이 든다. 따라서 과잉 정비가 될 수 있고, 정기적인 보수를 위한 설비 정비는 생산성을 감소시키는 문제가 발생한다. 상태 기반 유지보수(CBM) 방식은 열화 정도를 평가하고, 점검 주기의 연장을 목표로 하고 있으며, 설비 상태에 기반한 예지보전 방식이다. 시간 기반 유지보수 방법보다 비용이 매우 낮다[1]. 최근 많은 기업에서 Internet of Thing(IoT)를 통한 설비의 상태 기반 예지보전 방식을 목표로 준비하고 있다.

## II. 본 론

일반적인 회전 설비에 이상이 생기면 진동이 발생하고, 이를 유발시키는 힘은 기계 부품의 회전 운동에 의해 발생하게 된다. 이 힘은 주로 부품의 회전 속도에 따라 방향과 진폭이 변화한다. 대부분의 진동 문제들은 회전 속도에 영향을 받는다[2].

실시간 설비 데이터 계측 시스템은 변위, 속도, 가속도의 기본 원리에 의해 결함을 검출할 수 있다. 변위는 시간에 대한 속도의 적분량으로 고정된 기준에 대한 상대위치 변화로서 측정하고, 속도는 시간에 대한 가속도의 적분량 또는 시간에 대한 변위의 미분량으로 피로 파손과 직접적인 연관이 있어 예지 정비에 적합하며 불균형, 축 정렬 불량 등 기

계적 결함을 감지한다. 가속도는 시간에 대한 속도의 미분량으로 힘과 직접적인 연관이 있고 베어링 결함, 기어비와 같은 고주파 특성에 민감하다[3].

표 1. 진동의 파라미터

Table 1. Vibration Parameters

	최저점	원점	최고점
변위(Displacement)	최저	0	최고
속도(Velocity)	0	최고	0
가속도(Acceleration)	최고	0	최저

가속도와 속도, 변위에는 상관관계가 작용하고 일정한 주기로 샘플링된 이산적인 데이터의 경우 변위 데이터를 통해 속도 및 가속도를 산출할 수 있다.

$$Velocity = \frac{a}{2\pi f} (= \int a dt) \quad (1)$$

$$Displacement = \frac{a}{4\pi^2 f^2} (= \int v dt) \quad (2)$$

물리 현상인 진동은 연속해서 값을 갖는 아날로그 그 현상으로 가속도센서의 출력은 아날로그 신호이다. FFT 및 설비 데이터 계측 시스템은 컴퓨터의 고속화 대응량화에 따라 아날로그 신호를 디지털로 변환하는 A/D변환을 통해 데이터를 디지털로 변환한 후 신호 처리한다. 고속 푸리에 변환(FFT, Fast Fourier Transform)은 시간 영역 데이터로부터 주파수 스펙트럼을 추출하는 기법이다[4]. 시간 영역 데이터를 주파수 영역의 데이터로 변환할 때, 실제로는 시간 영역 데이터는 시간 간격  $\Delta t$  마다의 이산 데이터가 되고, 그 개수는 유한개인  $N$ 이 된다. 이것은 연속 시간의 데이터를 측정시간  $T = N \Delta t$  동안 잘라낸 것과 같다. 샘플링 주파수의 절반을 나이퀴스트 차단 주파수라고 한다. 푸리에 변환할 수 있는 최대 주파수이다.

진동 파형의 주파수 정보를 정확하게 얻기 위해서는 1주기당 2배 이상 샘플링해야 한다. 신호에서 특정 주파수 성분을 걸러내기 위해 필터를 사용한다. 필터는 주파수에 따라 통과시키는 주파수 대역(통과대역, Pass Band)과 감쇠시켜서 통과를 저지하는 주파수 대역(저지대역, Stop Band)의 두 특성을 가진 전자 회로이다.

필터의 특성으로 차단 주파수를 경계로 신호를 완벽하게 통과시키는 부분과 저지하는 부분으로 나누어야 이상적이지만 실제로는 천이대역(Transition Band)이 있고, 그 좌우(통과대역과 저지대역)로 크기(Gain)가 변동하는 맥동(Ripple)이라 불리는 특성이 있다. 형식에 따라 천이 대역의 폭이나 맥동의 크기가 다르다[5].

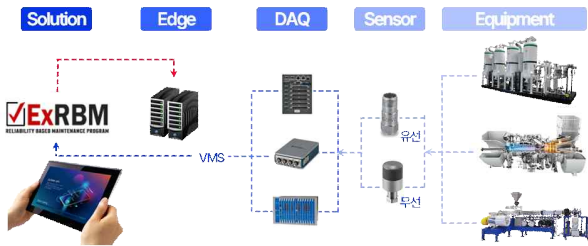


그림 2. 측정 개념도  
Fig. 2. Measurement concept

### III. 구 현

실시간 설비 데이터 계측 시스템의 하드웨어 구성은 진동 신호를 가속도센서로 받아서 DAQ를 통해 아날로그 신호를 디지털로 변환한다. 수집된 데이터는 Edge(PC)에서 라벨링 및 분석된다. Edge의 사양은 Intel® Core™ i7, DDR4 SODIMM 16GB, M.2 NVMe SSD 256GB와 별도의 SSD 1TB로 구성된다. 가속도센서와 DAQ로 측정된 Raw Data는 신호 처리와 분석을 통해 결과를 User Interface(UI)로 보여준다.



그림 3. 실시간 설비 데이터 계측 시스템  
Fig. 3. Real-time facility data measurement system

설비의 진동 특성 파라미터를 통해 결함을 추출하고 이를 기반으로 설비를 실시간 자동으로 진단하여 결함의 유형과 원인, 고장을 예측하여 설비 관리자들에게 통보함으로써 신속하게 설비를 유지 보수할 수 있다. 설비 상태를 ISO 기준에 따라 4단계인 양호, 장시간 운전, 불량, 극도 불량으로 분리하여 운전 시 설비로부터 진동 데이터를 수집하고, 실시간으로 설비 상태를 표시하고 있다. 베어링 특성만을 관리함으로써 미세한 진동량 변화에도 베어링 결함을 검출할 수 있다. 설비 데이터 계측 시스템을 이용하여 터빈의 정상상태 및 비정상 상태의 데이터를 측정 및 수집하여 분석 프로그램(Data Viewer)을 사용하여 시계열의 시간 영역과 주파수 스펙트럼 결과를 보여준다.

터빈의 정상상태의 데이터를 분석하면 정상상태일 때 진동량이 낮은 상태에서 회전 성분과 함께 2X ~ 3X 성분 및 배수 성분들이 전반적으로 낮은 값을 갖고 나타나고 있다.

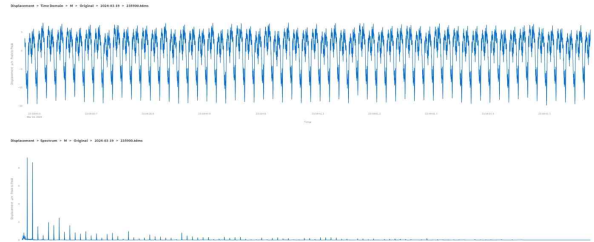


그림 31. 터빈의 정상상태 데이터  
Fig. 4. Turbine steady state data

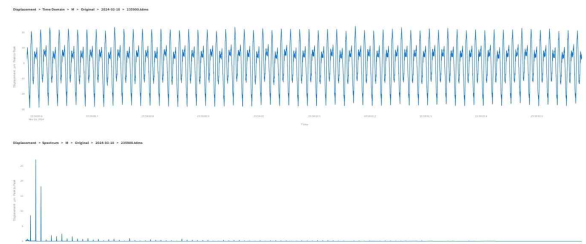


그림 32. 터빈의 비정상 상태(축정렬 불량) 데이터  
Fig. 5. Turbine Abnormal status data(Misalignment)

터빈이 비정상 상태일 때 2X RPM 및 3X RPM 성분이 매우 높게 발생을 하며, 전체 진폭의 크기가 2배 이상 커지는 것을 볼 수가 있다. 그림 6과 7에서는 설비의 비정상 상태의 하나인 질량불평형일 때의 진동 데이터를 보여주고 있다. 설비의 회전 성

본인 1X RPM 성분이 매우 높게 발생하고 있음을 알 수 있다. 이를 기반으로 설비를 실시간 자동으로 진단하여 설비에서 발생하는 결함의 유형을 검출할 수 있으며 이에 따른 적절한 대처 방법을 세울 수 있다.

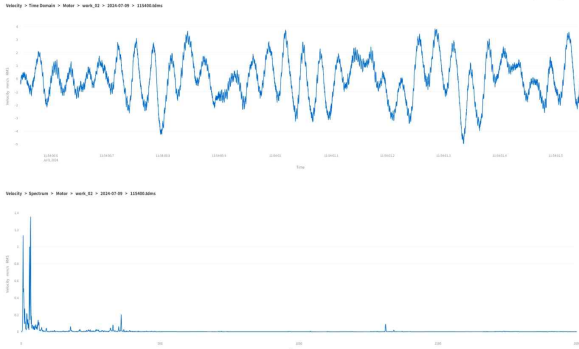


그림 33. Fan 구동 Motor의 정상상태 데이터  
Fig. 6. Fan Motor steady state data

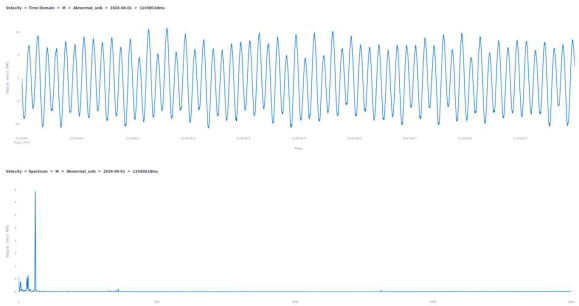


그림 7. Fan 구동 Motor의 비정상 데이터  
Fig. 7. Fan Motor Abnormal status data(Unbalance)

#### IV. 결론 및 향후 연구 방향

본 논문에서는 산업 현장에서 운전되고 있는 회전기계의 설비 상태를 실시간으로 계측하고 분석할 수 있는 시스템을 구성하고 실제 운영 중인 설비에서 진동 데이터를 측정 및 수집하였다. 측정된 데이터의 Raw Data는 확장자가 TDMS 파일로 저장되며, 저장된 데이터는 LabView로 작성된 프로그램을 통해 주파수를 분리하고, 알고리즘을 통해 여러 가지 고장에 대한 실시간 진단 및 예측을 수행한다. 이를 최종적으로 사용자에게 JSP를 통해 Web UI로 표현한다. FFT 분석을 통해 여러 가지 고장에 대한 실시간 진단 및 예측을 수행함으로써 현장에서 운영하는 불특정 회전기계에 대한 실시간 모니터링 구

축이 가능하다.

향후 데이터 계측 및 분석 진단기를 Edge(PC)에서 경량화 및 저전력의 NPU를 사용하여 고장의 근본 원인을 온-디바이스 AI 환경에서 딥러닝 모델의 연산 효율성 증대 및 메모리 사용량 절감을 위해 데이터 정밀도를 줄이면서도 모델 정확도를 유지하는 경량화 신경망 구조 탐색 기반 최적 기술의 알고리즘을 개발하여 설비의 선제적 고장감지/예측/대응 등의 예지보전 경쟁력을 강화하고자 한다.

#### 참 고 문 헌

- [1] 이종원, "ISO 표준에 따른 기계설비의 상태감시와 진단(진동, 영역 IV)", 한국소음진동공학회, pp. 159, January 2019.
- [2] Lee, B. J., "Vibration and Maintenance Handbook", KEPCO, pp. 18-36, 1998.
- [3] J. H. Cho, "Development of Smart IIoT sensor for predictive maintenance of high-speed rotary machine and facilities", Summer Annual Conference of IEIE, VOL 45 NO. 01, pp. 1581-1583, June 2022.
- [4] 정재천, "인공지능 시대의 회전기기 진단", 인터비전, pp. 100, February 2020.
- [5] 이종원, "ISO 표준에 따른 기계설비의 상태감시와 진단(진동, 영역 IV)", 한국소음진동공학회, pp. 89-104, January 2019.

#### ACKNOWLEDGMENTS

본 논문은 정부(산업통상자원부)의 지원으로 전자부품산업기술개발 - 주력산업 IT 개발사업 고장도수율이 높은 설비 대상 연합학습 추론 모델 기반 온-디바이스 AI 자가 진단 및 실증 기술개발 과제 의 지원을 받아 수행된 연구임. (RS-2024-00444913)

# 이종 설비의 잔여 수명 예측을 위한 연합학습 연구 동향

김성래\*, 차민혁, 이태경, 이종혁, 남건우, 강대영, 윤상일, 김준영\*\*

## A Survey on Federated Learning for Remaining Useful Life Prediction of Heterogeneous Equipment

Seongrae Kim\*, Minhyeok Cha, Taekyong Lee, Joong-hyeok Lee, Keonwoo Nam, Daeyoung Kang, SangIl Yoon, and Joon-Young Kim\*\*

압축기, 펌프, 팬, 모터와과 같은 다양한 설비는 플랜트 산업을 비롯한 여러 산업 분야에서 필수적으로 활용되며, 각 설비는 제조사, 작동 원리, 운영 조건에 따라 특성이 크게 달라진다. 이로 인해 동일한 목적을 수행하는 설비라도 현장에 따라 성능 및 수명에 차이가 발생하며, 잔여 수명(Remain Useful Life)을 예측하기 위한 데이터 또한 이질적인 특성을 지니게 된다. 이러한 데이터의 이질성으로 인해, 잔여 수명 예측 모델은 설비별 고유의 특징을 충분히 반영할 필요가 있으며, 고장 유형에 따라 변화하는 잔여 수명을 정확히 추정하는 데 어려움이 따른다. 더불어, 설비 운영사들은 데이터 프라이버시 보호와 보안상의 이유로 운영 데이터를 외부로 반출하지 않으면서도 잔여 수명 예측 모델을 구축하려는 요구가 있다. 본 논문에서는 이러한 문제를 해결하기 위해 데이터 프라이버시를 보장하면서도 이종 설비의 다양한 데이터를 효과적으로 학습할 수 있는 연합학습(Federated Learning) 접근 방식을 제안한다. 연합학습의 대표적인 기법인 FedAvg<sup>[1]</sup>, FedProx<sup>[2]</sup>, FedMA<sup>[3]</sup>, FedBN<sup>[4]</sup> 등을 분석하고, 이들이 잔여 수명 예측 모델에 적용될 때의 각 기법의 특징 및 적용 가능성을 평가한다. 이를 통해 이질적인 설비 데이터의 특성을 반영하면서도 산업 현장에서의 실질적인 활용 가능성을 논의하고자 한다.

### Key words

Federated Learning, Data Privacy, Remain Useful Life

### 참고문헌

- [1] McMahan, Brendan, Moore, Eider, Ramage, Daniel, Hampson, Seth, y Arcas, Blaise Aguera. "Communication-efficient learning of deep networks from decentralized data." *Artificial Intelligence and Statistics*, 1273-1282, 2017.
- [2] Li, Tian, Sahu, Anit Kumar, Zaheer, Manzil, Sanjabi, Maziar, Talwalkar, Ameet, and Smith, Virginia. "Federated Optimization in Heterogeneous Networks." *Proceedings of Machine Learning and Systems*, vol. 2, pp. 429-450, 2020
- [3] Wang, Hongyi, Yurochkin, Mikhail, Sun, Yuekai, Papailiopoulos, Dimitris, Khazaeni, Yasaman. "Federated learning with matched averaging." *arXiv preprint arXiv:2002.06440*, 2020.
- [4] Li, Xiaoxiao, Jiang, Meirui, Zhang, Xiaofei, Kamp, Michael, Dou, Qi. "Fedbn: Federated learning on non-iid features via local batch normalization." *arXiv preprint arXiv:2102.07623*, 2021.

\* 고등기술연구원 AI/DX센터 산업인공지능연구그룹, clipboard@iae.re.kr,

\*\* 고등기술연구원 AI/DX센터 산업인공지능연구그룹, joonykim@outlook.com

※ 이 연구는 2024년 산업통상자원부 및 한국산업기술기술평가원(KEIT) 연구비 지원에 의한 연구임.(RS2024-00444913)

# 고장도수율이 높은 설비를 위한 딥러닝 기반 잔여 수명 예측 모델 개발

차민혁\*, 김성래, 이태경, 이종혁, 강대영, 남건우, 윤상일, 김준영\*\*

## Development of Deep Learning-Based Remaining Useful Life Prediction Models for High-Failure-Rate Equipment

Minhyeok Cha\*, Seongrae Kim, Taekyong Lee, Joong-hyeok Lee, Daeyoung Kang, Keonwoo Nam, Sang-il Yoon, and Joon-Young Kim\*\*

압축기, 터빈, 펌프와 같은 회전 설비는 고장이 빈번하게 발생하며, 단 한 번의 고장으로 전체 공장 및 플랜트의 가동이 중단되는 심각한 피해를 초래할 수 있다. 이러한 고장 다빈도 설비들은 가동 중단으로 인한 생산 손실을 최소화하기 위해 기존의 시간 기반 모니터링(Time Based Monitoring, TBM) 방식에 의한 예방적 유지보수 방법을 채택해 왔다. 그러나 시간 기반 모니터링 방식은 고정된 주기에 따라 유지보수를 수행하기 때문에, 과잉 정비 또는 적절한 보전 시기 누락으로 인한 비효율이 발생해 설비의 생산성 저하의 위험이 있다. 이를 해결하기 위해, 최근에는 설비 상태 정보를 기반으로 잔여 수명을 예측하여 최적의 유지보수를 수행하는 상태 기반 모니터링(Condition Based Monitoring, CBM) 방식의 필요성이 대두되고 있으며, 관련 연구가 활발히 이루어지고 있다. 본 연구는 고장도수율이 높은 설비의 효율적 유지보수 시점 예측, 범용적 적용 및 모델 간 성능 평가를 목표로 한다. NASA의 C-MAPSS(Commercial Modular Aero-Propulsion System Simulation) 데이터셋을 학습 데이터로 활용하였으며, Bi-LSTM(Bidirectional Long Short-Term Memory)[1], LSTM(Long Short-Term Memory)[2], CNNLSTM(Convolutional Neural Network Long Short-Term Memory)[3], Transformer[4]를 활용하여 딥러닝 기반 잔여 수명 예측 모델을 개발하였다. 또한, 모델들의 성능 최적화를 위해 격자 탐색(Grid Search) 기법을 사용한 하이퍼파라미터 튜닝을 수행하였다. 아울러, 예측 성능 평가를 위해 RMSE(Root Mean Square Error)와 Asymmetric Score 지표를 사용하여 C-MAPSS 데이터셋의 네 개 서브 데이터셋에 대한 모델들의 성능을 비교 분석하였다. 본 연구는 각 모델의 장단점을 기반으로 최적의 활용 방안을 제시하여, 설비의 비가동 시간 감소와 유지보수 비용 절감을 위한 상태 기반 유지보수 시스템의 설계에 실질적인 기여를 할 것으로 기대된다.

### Key words

Predictive Maintenance, Remain Useful Life, Condition Based Monitoring

### 참고문헌

- [1] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, 1997.
- [2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [3] X. Shi and Z. Chen, H. Wang, D. Y. Yeong, W. K. Wong, W. C. Woo, "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting," in *Advances in Neural Information Processing Systems*, vol. 28, pp. 802-810, 2015.
- [4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is All You Need," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998-6008, 2017.

\* 고등기술연구원 AI/DX센터 산업인공지능연구그룹, pro0776@iae.re.kr,

\*\* 고등기술연구원 AI/DX센터 산업인공지능연구그룹, joonykim@outlook.com

※ 이 연구는 2024년 산업통상자원부 및 한국산업기술기술평가원(KEIT) 연구비 지원에 의한 연구임.(RS2024-00444913)



# 설비 유지보수체계 지능화를 위한 스마트MRO 플랫폼 구조 개발

정성관\*, 김준익\*\*, 안경익\*\*

## Development of a SmartMRO Platform Architecture for Intelligent Facility Maintenance Systems

Sungkwan Jung\*, Junik Kim\*\*, and Kyungik An\*\*

설비 상태에 맞춰 적시에 정비 수행을 실시하여 불필요한 가동중지 및 비용손실을 최소화 할 수 있는 상태기반정비(CBM: Condition-based Maintenance, 또는 예지정비) 개념의 발전과 이를 지원하는 플랫폼들의 다양화로 산업 현장에서 새로운 정비개념의 이점을 활용하고자 하는 요구는 나날이 증가하고 있다[1]. 하지만, 현장의 작업관행과 익숙하지 못함에 따른 장애발생의 위험, 새로운 시스템 도입의 비용 부담 등으로 선뜻 적용하지 못하고 있으며, 특히 기존 정비 개념 (고장정비 및 예방정비)과 상태기반정비 개념을 적절하게 조율하여 적용하기 위한 방법의 부재로 실질적인 활용이 극히 제한되는 경우가 빈번하다[2].

이 논문은 산업 현장에서 설비 유지관리를 위해 고장정비 또는 예방정비 방식을 사용하고 있던 상황에서 설비의 상태 진단에 따른 상태기반정비(예지정비) 방식을 유기적으로 연계 적용할 수 있는 방법에 대해 다루고 있다. 이를 위해, 설비의 정비 체계를 기반으로 정비계획을 관리하는 MRO 정보시스템(MIS), 설비 상태 진단 및 잔여수명을 예측하는 설비상태분석시스템(PHM) 그리고 설비 운영계획 정보 연계를 위한 기업정보시스템(ERP, PLM, MES 등) 연계 모듈로 구성된 'SmartMRO' 플랫폼 구조를 제안한다.

SmartMRO 플랫폼의 MRO정보시스템은 일반적인 예방정비시스템과 유사하게, 설비의 신뢰성 유지 목표에 따라 설정된 예방정비 주기를 기반으로 기업정보시스템의 장비운영계획을 참고하여 정비 일정을 수립하고 이의 수행 기록을 관리한다. 또한, 설비상태분석시스템에서 발견한 상태이상 정보를 기반으로 추가적인 정비계획을 수립하고 이의 수행을 기록한다. 이 과정에서 MRO정보시스템은 정비 수행 직전의 설비 상태 정보를 설비상태 분석시스템을 통해 수집할 수 있는데, 이 상태분석 및 잔여수명예측 정보를 기반으로 계획정비시점에서의 성능 목표 충족 여부를 평가할 수 있다. 이러한 데이터 주기적으로 분석하여 대상 설비의 달성신뢰도를 점검하고 이에 따라 예방정비 주기 재설정 작업을 수행할 수 있다. 이러한 SmartMRO 플랫폼 구조의 적용을 통해 현장에서는 예방정비 체계를 기반으로 상태기반정비 체계의 장점을 활용할 수 있으며, 이를 통해 현장 업무 혼선 및 부담을 경감할 수 있을 것으로 기대한다.

### Key words

condition-based maintenance (CBM), smart MRO platform, predictive maintenance, maintenance integration

### 참고문헌

- [1] A. K. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical systems and signal processing*, vol. 20, issue 7, pp. 1483-1510, 2006.
- [2] H. N. Teixeira, I. Lopes and A. C. Braga, "Condition-based maintenance implementation: a literature review", *Procedia Manufacturing.*, vol. 51, pp. 228-235, June 2021.

---

\*시버리솔루션스(주), sk.jung@seaburysolutions.com(교신저자), \*\*시버리솔루션스(주)

※ 이 연구는 2024년 산업통상자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임.(RS2024-00444913)

# Llama 3 KoBLLa: 한국어 이중언어 모델

김병천\*

## Llama 3 KoBLLa: A Korean Bilingual Large Language model

Byung Chun Kim\*

### 요약

본 연구는 Meta의 Llama 3 8B 모델을 기반으로 구축된 한국어 이중언어 대형 언어 모델, KoBLLa를 소개한다. KoBLLa는 최신 한국어 위키백과 데이터와 이중언어 정렬 데이터셋인 KoBLIMA를 활용하여 한국어와 영어 모두에서 우수한 성능을 보일 수 있도록 설계되었다. 특히 KoBLIMA는 모델이 한국어와 영어의 언어적 공간을 상호 참조하며 학습하도록 유도한다. 이를 통해 KoBLLa는 두 언어의 지식과 맥락을 자연스럽게 통합하는 능력을 갖추게 되었으며, 명령 수행 능력과 언어적 이해도를 동시에 강화하였다.

### Abstract

This study introduces KoBLLa, a **Korean Bilingual Large Language** model based on Meta's Llama 3 8B. KoBLLa leverages recent Korean Wikipedia data and the bilingual alignment dataset, KoBLIMA, to excel in both Korean and English. KoBLIMA encourages cross-referencing between Korean and English language spaces, thus enhancing KoBLLa's ability to integrate knowledge and context from both languages seamlessly. This dual alignment strengthens the model's instruction-following capabilities and linguistic comprehension.

### Key words

Large Language Model, Transformer, Llama3, Bilingual, Korean-English

### 1. 서론

현재 한국어 대형 언어 모델에 대한 수요는 꾸준히 증가하고 있으며, 다양한 한국어 모델들이 출시되고 있다. 그러나 한국어와 한국 문화에 대한 깊이 있는 이해와 한국어로 된 지시를 정확히 따를 수 있는 능력을 갖춘 모델은 여전히 부족하다. 이러한 요구에 부응하기 위해서는 단순히 한국어 텍스트를 학습하는 것만으로는 한계가 있으며, 한국어의 문화

적 맥락과 한국어 사용자의 특성까지 이해할 수 있는 새로운 접근법이 필요하다.

KoBLLa 모델은 기존의 한국어 모델과 차별화된 접근 방식을 취해, 언어적 이해와 명령 수행 능력을 동시에 갖춘 다재다능한 모델을 목표로 한다. 이를 위해 KoBLLa는 한국어와 영어 모두에서 명확하고 정확한 지시 수행 능력을 잃지 않으면서도 한국어

\*(주)매쓰에이아이, bckim@maiclass.com

※ 이 논문은 2024년 산업통상자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임(RS-2024-00444913)

를 심도 있게 이해할 수 있도록 설계되었다. 특히, 한국어 사용자가 직면하는 다양한 분야에 맞춤형 배경 지식을 제공함으로써 실제 활용도가 높은 한국어 이중언어 모델을 지향한다.

기존 한국어 모델들은 주로 한국어 텍스트를 학습하여 언어 이해 능력을 키우는 데 중점을 두어왔으나, 그 결과 명령 수행 능력에는 한계가 존재했다. 또한, 고품질의 한국어 명령 데이터는 구하기 어려울 뿐 아니라, 상업적 활용에는 라이선스 문제가 따를 수 있다. 이러한 문제를 해결하기 위해, KoBLLa는 한국어를 외국어로서 깊이 있게 이해하면서도 영어 지시 수행 능력을 유지하는 모델을 구축하는 것을 목표로 한다.

## II. 모델

KoBLLa 모델은 Meta의 Llama 3 8B 모델을 기반으로 한국어 이해와 문화적 맥락, 특정 분야의 지식 습득 능력을 향상시키기 위해 추가 훈련된 이중언어 모델이다. Meta의 Llama 3는 80억 파라미터를 가진 대형 언어 모델로, 월간 7억 명 이하의 활성 사용자 수를 가진 소규모 기업이 자유롭게 사용할 수 있다. Llama 3는 기본적으로 영어 기반 모델이지만, 한국어를 포함한 다양한 외국어 데이터로 학습되었고, 비알파벳 언어를 다룰 수 있는 충분한 어휘 범위를 갖추고 있다. KoBLLa는 이러한 Llama 3 8B 모델을 활용하여 한국어에 특화된 추가 학습을 수행함으로써 한국어에 대한 깊이 있는 이해를 도모하였다.

또한, KoBLLa의 인스트럭트(Instruction) 모델은 영어 지시를 잘 따르도록 훈련된 Llama 3 8B-Instruct 모델을 기반으로 한다. 최근 연구에 따르면, 언어 모델의 지시 수행 능력은 단순한 지식 습득만이 아니라, 상황에 맞게 의사소통 스타일을 유연하게 전환하는 능력에도 좌우된다. 이를 위해 KoBLLa는 Llama 3의 기본 모델에서 한국어 언어, 문화, 지식을 텍스트 데이터로 학습하고, 인스트럭트 모델(또는 대화 모델)에서는 상대적으로 소규모의 지시 데이터로 추가 훈련을 수행하였다. 이러한

접근은 LIMA[1] 모델의 아이디어를 참고하였다.

인스트럭트 모델의 훈련 과정에서 모델의 전체 파라미터에 큰 변화를 주지 않기 위해 LoRA(Low-Rank Adaptation, [2])와 Chat Vector[3] 개념을 도입하였다. 이 방식은 모델의 특정 작은 부분에 집중하여 변화를 주는 것으로, 인스트럭트 모델의 가중치와 기본 모델의 가중치 간 차이를 계산해 ‘Chat Vector’라 명명하였다. Llama 3의 인스트럭트 모델과 기본 모델부터 추출된 Chat Vector는 한국어로 추가 학습된 기본 모델에 적용되며, 이를 통해 KoBLLa의 인스트럭트 모델을 미세조정하기 위한 초기값으로 사용된다.

## III. 학습데이터

KoBLLa 모델의 학습에는 최신 한국어 데이터셋을 사용하여 한국어에 대한 깊이 있는 이해를 강화하였다. 이를 위해 두 가지 주요 데이터셋을 사용하였다.

### Ko-wikitext (20240501)

KoBLLa 기본모델을 학습하기 위해 최신 한국어 위키백과 덤프 데이터를 기반으로 구축하였다. 이 데이터셋은 한국어 위키백과에서 제공하는 정보를 최신 버전으로 반영하기 위해 lovit/kowikitext 저장소[4]의 스크립트를 수정하여 .csv 형식으로 저장하였다. Ko-wikitext는 한국어 위키백과와 동일한 CC-BY-SA 3.0 라이선스를 따르며, 자세한 라이선스 정보는 Creative Commons에서 확인할 수 있다.

### KoBLIMA

KoBLIMA는 KoBLLa 인스트럭트 모델을 훈련시키기 위해 Google의 Gemini 1.5 Flash를 사용하여 생성된 한국어-영어 이중언어 버전의 LIMA (Less Is More for Alignment, [1]) 데이터셋이다. KoBLIMA는 적은 양의 데이터로도 효율적인 언어 정렬을 달성하는 LIMA[1]의 접근법을 한국어 데이터에 맞게 확장한 것으로, KoBLLa 모델의 한국어-영어 지시 수행 능력을 높이는 데 기여한다. 데이터셋은 한국어 질문과 영어 대답, 영어 질문과 한국어 대답으로 이

루어져 있어 영어와 한국어의 언어공간을 상호 참조 하도록 유도되어 있다. 상세한 정보는 관련 저장소에서 확인할 수 있다.[5]

이와 같이 최신의 한국어 데이터와 최적화된 정렬 방식의 이중언어 데이터셋을 바탕으로, KoBLLa 모델은 한국어와 영어 모두에서 우수한 성능을 발휘하도록 학습되었다.

#### IV. 결론

본 연구에서는 Meta의 Llama 3 8B 모델을 기반으로 한국어와 영어 모두에서 우수한 성능을 발휘할 수 있는 한국어 이중언어 모델, KoBLLa를 구축하였다. KoBLLa는 최신 한국어 위키백과 코퍼스[4]와 한국어 이중언어 정렬 데이터셋인 KoBLIMA[5]를 활용하여, 한국어의 언어적 이해와 문화적 맥락을 포괄적으로 학습하였으며, 영어 지시 수행 능력 또한 유지하는 방향으로 훈련되었다. 특히, KoBLIMA[5] 데이터셋은 모델이 영어와 한국어의 언어 공간을 상호 참조하며 학습하도록 설계되었다. 이를 통해 KoBLLa는 두 언어의 지식과 맥락을 자연스럽게 통합하는 능력을 갖추게 되었다.

기존의 한국어 모델들이 언어 이해에 초점을 맞춘 반면, KoBLLa는 텍스트 생성뿐만 아니라 다양한 분야의 지시 수행 능력을 강화하여 실제 응용 가능성을 높이는 것을 목표로 하였다. 이를 위해 LoRA[2] 및 Chat Vector[3]와 같은 효율적 파라미터 조정 방식을 도입함으로써 모델의 학습 효율성과 비용 절감 효과를 동시에 달성하였다.

KoBLLa는 한국어와 영어를 유연하게 다루며, 학술 연구, 비즈니스 응용, 대화형 AI 등 다양한 분야에서의 활용 가능성을 보여준다. 향후 KoBLLa의 성능을 더욱 발전시키기 위해 보다 정교한 한국어 데이터와 다양한 언어적 맥락을 포함한 훈련을 통해 언어적 통합 능력을 강화할 수 있을 것으로 기대한다.

#### 참 고 문 헌

[1] Chunting Zhou, Pengfei Liu, Puxin Xu, Srinu Iyer, Jiao Sun, Yuning Mao, Xuezhe Ma, Avia Efrat, Ping Yu, Lili Yu, Susan Zhang, Gargi Ghosh, Mike Lewis, Luke Zettlemoyer, Omer Levy, "LIMA: Less Is More for Alignment", arXiv:2305.11206 [cs.CL], 18 May. 2023.

[2] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, "LoRA: Low-Rank Adaptation of Large Language Models", arXiv:2106.09685 [cs.CL], 17 Jun. 2021.

[3] Shih-Cheng Huang, Pin-Zu Li, Yu-Chi Hsu, Kuang-Ming Chen, Yu Tung Lin, Shih-Kai Hsiao, Richard Tzong-Han Tsai, Hung-yi Lee, "Chat Vector: A Simple Approach to Equip LLMs with Instruction Following and Model Alignment in New Languages", arXiv:2310.04799 [cs.CL], 7. Oct. 2023.

[4] <https://github.com/lovit/kowikitext/>

[5] <https://github.com/bckim-mathai/KoBLIMA/>

#### ACKNOWLEDGMENTS

본 논문은 정부(산업통상자원부)의 지원으로 전자부품산업기술개발 - 주력산업 IT 개발사업 고장도수율이 높은 설비 대상 연합학습 추론 모델 기반 온-디바이스 AI 자가 진단 및 실증 기술개발 과제 의 지원을 받아 수행된 연구임. (RS-2024-00444913)

# 미래 융합 IT인문학을 통한 지속 가능한 실천교육 리빙랩 기반의 유니베르시타스 자생플랫폼 구축 방안

김미진\*, 주연수\*, 최근배\*\*, 박남제\*\*\*

## A plan to build a self-sustaining platform for Universitas based on a living lab for sustainable practical education through future convergence IT humanities

Mijin Kim\*, Yeonsu Joo\*, Keunbae Choi\*\*, and Namje Park\*\*\*

미래형 지능 정보화 사회의 시대에서 다양한 학문과 기술들이 융합되어 새로운 지식과 가치를 창출하는 가능성이 갈수록 주목받고 있으며, 변화를 수용하고 주체적인 태도로 미래를 개척하는 인재를 양성하기 위한 'IT융합 교육'이 떠오르고 있다. 시대의 주체인 인간과 미래 융합 IT기술이 적용된 사회에 대한 제 현상에 대응하는 선제적 연구로서, 데이터를 비판적으로 성찰하는 인문학적 창의성을 적용하는 연구를 진행하였다[1]. 제주대학교 융합과학기술사회연구소는 지난 2019년부터 1단계 연구를 통해 초·중등을 위한 미래융합 IT 인문학 교육 모델 연구 및 핵심원리 교재의 개발, 현장적용, 고도화, 효과성분석, 교재 및 교수학습 자료의 검증, 보급, 활성화, 확산을 위해 연구하였다. 2단계에서는 3년동안 글로벌 리빙랩 메이커스를 위한 초·중등 인문소양의 동서양 고전 속 미래융합IT인문학 및 新 교수·학습법 실증을 연구하고 있다. 2단계의 1차년도에 미래 융합 IT인문학 거버넌스 기반 확산형을 통한 전국화, 2차년도에 글로벌 교육 협업 네트워크 다각화를 통한 국제화, 3차년도에 지속 가능한 실천교육 리빙랩의 적용에서 나타난 자생플랫폼의 선도적 다각화 수행을 목표로 하고 있다. 미래 융합 IT인문학의 연구 결과물을 확산하고, 지능정보기술과 인문학의 심도깊은 연구와 지속 가능한 연구 정착을 위하여 비즈니스 모델링을 수행할 예정이다. 특별 캠프, 교원 연수, 경연 대회, 선도형 방과후 수업의 참가비를 받거나, 교재, 교구, 교수·학습자료, MOOC 동영상 등의 미래융합IT인문학 연구 결과물을 대중에게 판매하여 수익을 창출하는 구조의 자생플랫폼을 구축할 계획이다. 또한, 글로벌 협의체 기반 리빙랩을 적용한 미래융합IT인문학 자생플랫폼을 구축하고, 이를 활성화를 위한 연구소 법인 설립 수행할 예정이다. 지속 가능한 실천교육 리빙랩의 협의체를 넘어서 하나의 목적을 지닌 공동체인 '유니베르시타스(Universitas)'를 구체화하기 위해 법인을 설립하고, 글로벌 협의체 기반 리빙랩을 적용한 미래융합IT인문학 자생플랫폼의 다각화할 예정이다. 이러한 인문사회연구소의 발전 방안은 자생플랫폼의 다각화로 여러 분야에서 광범위한 효과를 기대할 수 있을 것으로 사료된다.

### Key words

Universitas, practical education, IT humanities, self-sustaining platform, living lab

### 참고문헌

- [1] Eunsun Choi, Namje Park, "Validation of Differences in Creativity and Core Competencies according to the Application of Creative Information Programs for Elementary and Secondary Students", Journal of The Korean Association of Information Education, 27(3), pp.269-279, June 2023.

\*제주대학교 융합과학기술사회연구소, \*\*제주대학교 교육대학 초등수학교육전공, \*\*\*제주대학교 교육대학 초등컴퓨터교육전공, 교신저자(namjepar@jejunu.ac.kr)

※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269).

# 늘봄학교 초등학생 대상 생성형 AI를 활용한 예술융합 프로그램과 예술융합 교재 개발

임현정\*, 박남제\*\*

## Development of Art Fusion Program and Art Fusion Textbook Using Generative AI for Elementary School Students at Neulbom School

Hyunjeong Lim\*, and Namje Park\*\*

2024학년도부터 시행된 늘봄학교는 돌봄 대기 문제를 100% 해소하고, 80% 이상의 학부모가 높은 만족도를 보이는 등 긍정적인 성과를 나타내었다. 이에 따라 교육부와 시도교육청은 1학기 운영 경험을 토대로 2학기부터 모든 초등학교에 늘봄학교를 안정적으로 도입할 예정이다(2024 교육부 발표). 제주대학교 늘봄학교사업단은 미래의 창의적 융합 인재 양성을 목표로 다양한 분야와 융합하여 초등 1학년 맞춤형 프로그램을 개발해 주도하고 있다. 그중 생성형 AI를 활용한 미디어 아트, 인공지능 창작곡 만들기, AI와 함께하는 화가, AI 음악가와 리듬 라이즈 등을 개발하였다. 생성형 AI를 활용한 예술융합교육 프로그램은 학생들에게 급변하는 미래 기술을 탐구할 기회를 제공하며, 창의력과 비판적 사고능력을 기르는 데 중점을 둔 커리큘럼 프레임워크이다. 이는 기술을 활용하여 예술적 해결을 함으로써 미래 사회에서 직면할 도전과 문제 해결 능력, 독창적인 아이디어를 기르는 초학문적 접근법이다. 제주대학교 늘봄학교사업단은 늘봄학교 예술융합 프로그램을 위해 예술하는 어린이 예술융합 교재를 개발했다. 이 교재는 예술융합의 창의적인 결과물을 보여주며, 초등 저학년 예술융합 교육의 중요성과 가치를 사회적, 문화적, 개인적 차원에서 깊이 있게 분석할 수 있는 도구이다. 초등 저학년 예술융합교재의 주요 함의는 다음과 같다. 창의적인 사고와 자기 표현 능력 발달에 중요한 역할을 한다. 예술 활동을 통해 자신의 감정을 표현하고 상상력을 발휘할 수 있다. 사회적 상호작용 및 협력을 통해 다양한 예술적 표현을 통해 타인의 시각과 경험을 이해하고 존중하는 태도를 기르며 그룹 활동을 통해 팀워크와 협력 능력을 배우고, 서로 다른 의견을 존중하며 협력하는 능력을 기른다. 자신의 작품을 만들고 발표하는 과정에서 성취감과 자신감이 증진하여 이러한 경험들로 학생들의 자아 존중감과 자기 효능감을 높이는 데 중요한 역할을 한다. 또한, 일상 생활에서 스트레스를 줄이고 정서적으로 안정된 상태를 유지하는 데 도움이 된다. 본 논문은 늘봄학교의 초등 저학년 대상 맞춤형 생성형 AI를 활용한 예술융합 프로그램과 교재 개발의 중요성을 강조한다. 프로그램은 학생들의 창의력, 비판적 사고, 사회적 상호작용 능력, 정서적 안정에 기여하며, 미래 사회에서 필요한 다양한 역량을 함양할 수 있도록 도울 것이며 예술하는 어린이 예술융합교재는 이러한 교육적 가치를 구체화하고, 학생들의 전인적 성장을 지원하는 데 중요한 도구가 될 것이다.

### Key words

Art Fusion Program, Art Fusion Textbook, Generative AI, Elementary School Students, Neulbom School

### 참고문헌

- [1] Kim, Y., Park, N. (2012). The Effect of STEAM Education on Elementary School Student's Creativity Improvement. In: Kim, Th., et al. Computer Applications for Security, Control and System Engineering. Communications in Computer and Information Science, vol 339. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-35264-5\\_16](https://doi.org/10.1007/978-3-642-35264-5_16).

\*제주대학교 일반대학원 융합정보보안학협동과정 박사과정, \*\*제주대학교 교육대학 초등컴퓨터교육전공 교수, 교신저자(namjepark@jejunu.ac.kr)

※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269).

# 신임 디지털포렌식 수사관을 위한 디스크 포렌식 교안 개발

곽창현\*, 박남제\*\*

## A Development on Disk Forensic Curriculum for Educating New Digital Forensic Investigators

Changheon Kwag\*, and Namje Park\*\*

본 논문에서는 경찰의 신임 디지털포렌식 분석관을 대상으로 한 디스크 포렌식 교안을 개발한다. 경찰 교육 기관에서는 현장에 투입할 신임 디지털포렌식 분석관을 대상으로 기초 디지털포렌식 교육을 진행하고, 이후 디지털포렌식 경력자들을 대상으로 디스크 포렌식, 모바일 포렌식, 영상복구 포렌식 등의 순서로 심화 교육을 진행한다. 경찰 교육기관의 해당 교육은 이를 담당하는 교수요원들의 개인적인 노력으로 과정이 설계되고 진행되어왔다. 경찰 수사교육기관은 수 년 내 그 외형적인 규모를 확장할 계획이고, 이에 맞춰 디지털포렌식 교육 커리큘럼도 체계적으로 재편될 시기가 되었다. 경찰은 20년 전부터 매년 민간 IT 인력을 경력채용하여 사이버 수사과 디지털포렌식 분석관으로 임용하고 있다. 신임 분석관들은 대부분 이렇게 채용된 사이버 수사관으로서 각자의 민간 경력에 따라 여러 IT 분야의 지식과 경력을 가지고 있다. 신임 사이버 경력채용자 모두가 디지털포렌식 영역의 경력이나 지식을 가지고 있지는 못하다. 경찰의 신임 디지털포렌식 분석관 교육의 대상은 이 IT 기본 지식과 경력을 가진 사람을 상정한다. 그리고 현장 실무의 요청에 의해 신규 분석관에게는 현실적으로 2주, 10일 가량의 디지털포렌식 교육 시간이 주어진다. 이 과정에는 파일시스템과 OS 아티팩트를 포함한 디스크 분석, 스마트폰 등 모바일 기기 분석, 관련 법과 절차, 법정에 제출할 보고서 작성 등의 내용이 포함된다. 이 과정에서 디스크 분석은 실습을 위주로 약 4일이 할당된다. 전국의 경찰 디지털포렌식 수사관들을 대상으로 한 설문을 통해 이 디스크 분석 교육 부분에 반드시 포함되어야 하는 내용을 선별하고, 이를 이용한 교안과 디스크 이미지를 개발하고자 한다. 본 논문 이후 해당 과정에 추가로 필요한 법과 절차, 모바일 분석, 보고서 등 다른 분야도 같은 연구 방식을 이용하여 교안을 개발하려 한다.

### Key words

Digital Forensic, Digital Forensic Image, Digital Forensic Curriculum, Digital Forensic Training

### 참고문헌

- [1] Myeonggil Choi, "A Study on Curriculum Design for Educating Digital Forensic Experts", Journal of Information Technology Applications & Management, Vol.30, No.6, pp. 113-142, December 2023.
- [2] Sojung Oh, Taegi Lee and Gibum Kim, "A Study on Digital Forensic Image Production Model", Journal of Digital Forensics, Vol.15, No.2, June 2021.

---

\*제주대학교 일반대학원 융합정보보안학협동과정 박사과정, \*\*제주대학교 교육대학 교수, namjepark@jejunu.ac.kr(교신)  
※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2022S1A5C2A04092269).

# 양자 컴퓨터 시뮬레이션 기반 초등 융합예술교육의 혁신적 교육 모형 개발 및 비전

임현정\*, 김진수\*, 박남제\*\*

## Development of Innovative Education Model and Vision for Elementary Convergence Art Education Based on Quantum Computer Simulation

Hyunjeong Lim\*, Jinsu Kim\*, and Namje Park\*\*

양자 컴퓨터 시뮬레이션을 활용한 초등 융합예술교육의 혁신적 교육 모형은 전통적인 교육 방식을 넘어 학생들이 과학적 사고와 예술적 감수성을 동시에 기를 수 있는 새로운 학습 환경을 조성하는 것을 목표로 한다. 양자 컴퓨터 시뮬레이션은 복잡한 양자역학적 현상을 시각적, 청각적 형태로 표현함으로써 학생들이 추상적인 과학적 개념을 직관적으로 이해할 수 있도록 돕는다. 이를 초등 예술교육에 도입하면 학생들은 양자 세계의 복잡한 원리를 체험하고, 이를 바탕으로 자신만의 예술 작품을 창작하는 과정에서 과학과 예술의 융합적 사고를 발달시킬 수 있다. 이러한 교육 모형은 학생들이 예술적 표현과 과학적 탐구를 동시에 경험할 수 있도록 설계되었으며, 이를 통해 학생들은 기존의 교육 방식을 뛰어넘는 창의적 문제 해결 능력과 융합적 사고력을 기를 수 있다. 양자 컴퓨터 시뮬레이션으로 생성된 다양한 시각적, 청각적 데이터를 바탕으로 학생들은 자신만의 예술적 해석을 추가하여 독창적인 작품을 완성하게 되며, 이는 그들이 예술적 감수성과 과학적 이해를 동시에 확장하는 데 중요한 역할을 한다. 이러한 교육 모형의 비전은 학생들이 미래 사회가 요구하는 창의적이고 융합적인 인재로 성장할 수 있도록 돕는 데 있으며, 이를 위해 양자 컴퓨터 시뮬레이션과 예술 창작을 결합한 다양한 교육 프로그램과 교재 개발, 교사 역량 강화, 그리고 학습 환경 조성 등의 전략이 필요하다. 양자 컴퓨터 시뮬레이션 기반 융합예술교육은 학생들에게 단순한 기술 습득을 넘어, 자신만의 예술적 언어와 과학적 이해를 바탕으로 새로운 문제를 창의적으로 해결할 수 있는 능력을 배양하게 하며, 이는 기존의 과학 교육이나 예술 교육의 틀을 넘어서서 학생들이 주체적으로 학습하고, 다차원적 사고를 발전시킬 수 있는 환경을 제공한다는 점에서 교육적 혁신을 가져올 수 있다. 나아가 이러한 교육 모형은 학생들이 양자역학의 복잡한 개념을 이해하고, 이를 예술적 창작을 통해 체화함으로써, 추상적 개념에 대한 깊은 이해와 창의적 표현 능력을 동시에 발달시키는 데 기여할 수 있으며, 이는 학생들이 다양한 문제 상황에서 창의적 해결책을 찾고, 과학적, 예술적 사고를 융합하여 새로운 가치를 창출하는 데 중요한 밑거름이 될 것이다. 양자 컴퓨터 시뮬레이션 기반 초등 융합예술교육은 학생들이 새로운 지식과 개념을 단순히 수용하는 데 그치지 않고, 이를 바탕으로 자신만의 독창적 사고와 표현을 발전시킬 수 있도록 하는 데 목표를 두고 있으며, 이는 그들이 미래 사회에서 요구되는 창의적 인재로 성장할 수 있도록 하는 교육적 비전을 제시한다.

### Key words

Creative Education, Arts Education, Quantum Computer

### 참고문헌

- [1] E. Choi and N. Park, "A study on the elementary informatics curriculum design through future competency analysis", Journal of The Korean Association of Information Education, vol. 25, no. 2, pp. 249-264, April 2021.

\*제주대학교 일반대학원 융합정보보안학협동과정 박사과정, \*\*제주대학교 교육대학 초등컴퓨터교육전공 교수, 교신저자(namjepark@jeju.ac.kr)

※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269).



# 제주도의 중소기업 영업비밀 유출 예방 방안

최익서\*, 박남제\*\*

## Measures to prevent trade secret leakage of small and medium-sized businesses in Jeju Island

Ikseo Choi\*, and Namje Park\*\*

본 논문에서는 중소기업 영업비밀 유출 예방 방안을 전국적인 차원에서 살펴보는 것이 아니라 제주도라는 섬에서 가지고 있는 기업체의 특수성과 정보보호 인식을 확인하고 유출 현황을 살펴보고, 영업비밀 유출 예방안을 마련하고자 한다. 산업기술은 국가기관이 산업 또는 경제에 영향을 미칠 수 있는 기술 중 지정, 고시 등을 통해 산업기술이라고 특정한 기술을 말하고, 영업비밀은 공공연히 알려지지 아니하고 독립된 경제적 가치를 경영상 또는 기술상의 정보 등을 말한다[1][2]. 제주도에 산업기술을 가진 기업체가 존재하기는 하지만 2023년 기준으로 그 숫자가 10개 미만이고, 제주도 내 기업체가 독립적으로 산업기술을 보유하기 보다는 육지에 있는 기업이 보유하고 있는 산업기술을 공동으로 보유하면서 이를 활용하는 실정이다. 그 보유 업체도 환경 관련 업체가 대부분이고, 국가에서 공고하는 환경 관련 사업을 낙찰받기 위한 가산점 확보 차원에서 활용하는 것이 현실이다. 또한 제주도는 대기업이 거의 없고 1인 기업 또는 가족기업이 다수를 차지하고 있어서 산업기술 또는 영업비밀을 보호해야 한다는 인식이 부족하고, 정보보호는 비용이 필요한 귀찮은 일이라는 인식이 육지에 비해서 다소 강한 편이다. 2023년 부정경쟁방지 및 영업비밀에 관한 법률이 개정되면서 기업이 보안조치를 취하지 않으면 처벌을 받을 수 있음에도 불구하고 이러한 특별법이 있음을 인식하지 못하며, 산업기술 또는 영업비밀의 보호 방법도 알지 못하고 있는 실정이다. 그리고 제주도에서는 육지부 기업들에 비해 규모가 작은 소규모 기업들이 많아 별도로 영업비밀 등에 관한 교육을 실시 하기가 어렵기 때문에 도청 등에서 경제적 지원 등을 하는 사업설명회에서 모인 기업들을 대상으로 산업기술 및 영업비밀 관리 방안 등에 대해 교육을 병행해서 실시해야 한다. 현재 일부 교육이 이러한 방법으로 실시되기도 하지만 영업비밀 유출 예방 방안에 관한 교육은 거의 없는 실정이라 추가할 필요가 있다. 그리고 교육 횟수도 연간 몇 회가 되지 않고 있으며 실제로 그러한 교육을 이수하는 기업체도 전체 기업수에 비해 아주 미비한 실정이다. 따라서 제주도청 등 관련 기관에서는 사업설명회뿐만 아니라 창업, 창직 관련 교육을 실시하고 있는 더큰내일센터, 창조경제혁신센터, 제주테크노파크 등에서 의무적으로 교육시간을 배정하고, 이들 기관 평가와 설명회에 참석한 기업의 평가 항목에 포함시켜 평가해야 한다. 또한 실제로 산업기술 등 유출 수사를 하고 있는 제주경찰청 소속 산업기술보호수사팀을 통해 제주에서 실제 발생하고 있는 다수의 영업비밀 유출 사건에 대해 구체적인 사례 등을 제주도청 관계자를 포함한 기업 등에게 교육하고 함께 공유할 필요가 있다. 이와같이 제주도는 별도의 집합교육이 어렵기 때문에 도청, 경찰청, 중소벤처기업진흥공단 등 모든 기관이 함께 관심을 갖고 예방 교육을 실시하여야 하고, 제주도 소재 기업은 반드시 관련 교육을 받아서 영업비밀 보호하기 위한 법률상 제도상 각종 조치를 취해야 한다.

### Key words

Industrial Technology, Trade Secret Leakage, Information Protection, Education

### 참고문헌

- [1] Strengthening The Capability of The Police Investigation in The Industrial Technology Protection Field. pp. 13-14, August 2020,
- [2] Korea Industrial Security Research, 01/02/2024, Vol. 14, p. 55-83.

\*제주대학교 일반대학원 융합정보보안학협동과정 박사과정, \*\*제주대학교 교수, namjepark@jejunu.ac.kr(교신)  
※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2022S1A5C2A04092269).

# 제주대학교 융합과학기술사회연구소의 필란트로피를 향한 인문학과 기술의 공존

정유진\*, 김미진\*, 주연수\*, 최근배\*\*

## The Coexistence of Humanities and Technology for Philanthropy

Yujin Jung\*, Mijin Kim\*, Yeonsu Joo\*, and Keunbae Choi\*\*

교육부는 한국연구재단과 함께 인문사회 기초학분분야 및 사회문제해결형 연구, 대학교육과 연계 등을 통해 연구소의 사회적·교육적 기능을 확대하여 인문사회연구소의 자생을 돕는 사업을 추진하고 있다. 제주대학교 융합과학기술사회연구소는 2012년 과학기술사회연구센터로 출범하여 본 사업을 2019년에 유치하고 5년간 130억원의 지원금으로 연구센터의 자생을 도모하고 있다. 지난 5년간 센터는 초·중등 인문소양의 동서양 고전 속 미래 융합IT인문학 및 신(新)교수·학습법을 적용한 실증 리빙랩 기반 구축을 위한 연구를 수행하기 위해 제주대학교에 미래융합IT인문학 교양과목을 개설 하여 운영하고 있다. 미래기술IT인문학 과목은 미래 사회가 다양한 학문과 기술이 융합하여 새로운 지식과 가치를 창출하는 미래형 지능 정보화 사회에 살아갈 디지털 네이티브(digital native)세대의 미래사회 핵심역량을 키워줄 수 있는 교사의 역량 함양을 목표로 하고 있다. 미래 사회가 요구하는 융합적이고 경쟁력 있는 인재를 양성하기 위해서는 학교 교육이 보다 유연하고 다양한 교육과정이 도입 되어야 하고 이는 학생들의 잠재력과 바람직한 가치관을 찾고 키워줄 수 있는 융합·인성 교육이 중요하다는 것을 골자로 한다. IT-인문학이라는 신 개념을 도입하고 연구하여 기반을 다져가기 위해 국내·외 유수의 연구기관과 협력 네트워크를 구축하고 IT인문학의 교육 모델을 제주에서 전국으로, 전국에서 해외로 글로벌라이제이션(glocalization)을 실천해 나아가고 있다. 학구적 협력 노력의 구체화와 가시화를 위해 지난 2022년부터 융합과학기술사회연구소 학술논문지를 출간하며 센터의 명칭을 연구소로 격상하고 자생화 노력을 순차적으로 진행하고 있다. 그동안 융합과학기술사회연구소는 초·중등생 대상 뿐만 아니라 지역 사회에 인간성에 대한 디지털 온톨로지 확립하고 미래융합IT의 기술을 인문학 연구의 기술적 보조도구로서 인문학 영역 확대에 기여할 수 있도록 공익법인의 설립을 목전에 두고 있다. 연구소가 지역사회의 필란트로피(Philanthropy)를 위하여 개인보다는 인류(humanity)라는 공동체 차원의 인류 발전을 위한 서비스를 구축해 나아가기 위한 법인 설립으로, 연구·개발의 대의를 세우기 위한 첫걸음을 내딛기 위하여 또한 본 인문사회연구소 지원사업의 방침을 짚고자 최종 목표를 설정하였고 마무리 단계에 접어 들고 있다. 본 연구소의 '필란트로피를 향한 인문학과 기술의 공존'은 인간 중심적이고 지속 가능한 사회를 만들기 위해 인문학적 가치와 첨단 기술이 상호 보완적으로 작용하는 것을 의미한다. 필란트로피, 즉 공공의 선을 위한 기부와 봉사 활동은 전통적으로 인류애와 윤리적 관점을 중시했으며, 이러한 활동의 기저에는 인문학적 통찰이 자리해 왔고 여기에 기술이 결합되면, 단순히 자원을 나누는 것을 넘어 사회문제에 대한 접근 방식을 확장하고, 더욱 효율적이고 혁신적인 해결책을 마련할 수 있습니다.

또한 미래융합IT인문학 과목은 제주대학교에 개설된 교양과목에 지나지 않고 국가평생교육진흥원에서 운영하는 한국형 온라인 공개강좌인 K-MOOC에 동 주제로 '인문학으로 만나는 디지털 혁명과 미래기술'로 운영하고 있으며, 대한민국 최초의 사이버대학교인 한국열린사이버대학교(OCU)에 미래기술인문학 과목을 운영하여 해마다 300명 이상의 이수자를 배출하고 있다. 앞으로 융합과학기술사회연구소는 IT인문학을 새로운 학문으로 인정받아 미래기술과 인류를 위한 필란트로피를 구축해 나가고자 한다.

### Key words

Digital Education Innovation, IT Humanities, Philanthropy, Advanced Future Technology, The Coexistence

\*제주대학교 융합과학기술사회연구소 \*\*제주대학교 교육대학 초등수학교육전공, kbchoi@jejunu.ac.kr(교신저자)

※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269).

# 공공기관 개인정보파일의 체계적 관리를 위한 개인정보보호수준평가 지표 개선 방안

현광남\*, 박남제\*\*

## Measures to Improve Personal Information Protection Level Evaluation Indicators for Systematic Management of Public Institution Personal Information Files

KwangNam Hyeon\* and Namje Park\*\*

본 논문에서는 공공기관의 개인정보파일의 체계적인 관리를 위해 「개인정보 보호수준 평가」의 지표 개선을 제안한다. 공공기관은 개인정보보호법에 따라 공공기관에서 운용되고 있는 개인정보파일을 개인정보보호 종합포털 등록을 의무화하고 있다. 개인정보파일은 공공기관에서 개인정보를 어떻게 처리하고 있는지를 국민에게 공개하여 개인 정보의 정정·삭제, 처리 정지 등을 요구할 수 있도록 개인정보자기결정권을 보장하려는 제도이다. 그러나 지금까지 형식적인 개인정보파일의 정비에 따른 개인정보취급자와 개인정보처리자의 관심 및 개선의 지 부족, 개인정보파일 등록 기준의 불명확성, 개인정보파일을 등록하기 위한 절차적 복잡성으로 인해 개인정보파일 관리에 한계가 있었다[1]. 이에 따라, 2024년부터 공공기관에서 진행되는 개인정보보호 수준 평가를 통해 개인정보파일 등록·관리를 질적 향상을 제안해보고자 한다. 개인정보 보호수준 평가는 개인정보보호법 개정(2023. 9.)에 따라 중앙행정기관, 지방자치단체, 공공기관, 지방공사, 시도 교육청 및 교육지원청 등 1,400여개 기관의 참여하게 된다. 법적 의무사항 준수 여부 등 정량 지표의 자체평가(60)와 개인정보 보호 업무 추진 내용의 적절성·충실성 등을 반영한 정성 지표 평가(40)로 이루어지며 24년의 평가결과는 25년 4월에 발표할 예정이다. 오랜 기간(24년~25년)에 걸쳐 진행되는 평가로 인해 공공기관 개인정보취급자의 전보 등 업무 변경 등이 발생하게 되고, 해당 업무에 대한 연속성이 단절되어 개인정보파일에 대한 전문성이 제한될 수 있는 단점이 있다. 그래서 각 기관에서 진행되고 있는 개인정보보호수준 평가 전·후의 각 기관의 개인정보파일을 비교 분석하고, 개인정보취급자들이 공통적으로 개선해야 되는 부분에 대한 선택적 교육을 해당연도에 진행하고, 그 결과를 토대로 개인정보 파일을 심층적으로 추가 정비·개선한 결과를 평가 지표에 반영하여, 개인정보파일의 통일성, 정확성을 확대해 나가는 방안을 제안하고자 한다.

Key words

Personal Information, Information Protection, Personal Information Protection Act,  
Public Institution, Right to Self-information Decision.

### 참고문헌

- [1] Yong-Hak Kim, Yong-Sun Ko, Joong-soo Bang, & Sam Hyun Chun. "A Study on the Improvement of Personal Information File Management in Public Institutions". Journal of The Korea Society of Information Technology Policy & Management, vol. 11(1), 1135-1140. January 2019.

\*제주대학교 일반대학원 융합정보보안협동과정 박사 과정, \*\*제주대학교 교수, namjepark@jejunu.ac.kr(교신)

※ 이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269)

# 초등학생의 디지털 리터러시 함양을 위한 딥페이크 및 가짜 뉴스 판별 교육 프로그램 개발 및 적용 방안

임현정\*, 박남제\*\*

## The Development and Application of Deepfake and Fake News Discrimination Education Programs to Cultivate Digital Literacy of Elementary School Students

Hyunjeong Lim\*, and Namje Park\*\*

최근 딥페이크 기술의 발전으로 인해 초등학생들이 진짜처럼 보이는 가짜 영상과 이미지를 쉽게 접할 수 있게 되면서, 정보의 진위를 판단하지 못하고 잘못된 믿음을 가질 위험이 커지고 있다. 예를 들어, 유명 인물이 하지 않은 발언이 딥페이크 기술을 통해 유포되어 많은 사람들이 이를 사실로 믿는 사례가 있었다. 이러한 상황에서 초등학생들에게 딥페이크와 가짜 뉴스의 위험성을 이해시키고, 이를 판별할 수 있는 능력을 길러주는 교육이 절실히 필요하다[1]. 본 연구는 초등학생들이 디지털 환경에서 딥페이크와 가짜 뉴스를 비판적으로 이해하고 판별할 수 있는 능력을 기르기 위한 교육 프로그램을 개발하고, 이를 효과적으로 적용할 방안을 모색하는 데 목적이 있다. 연구는 디지털 리터러시, 딥페이크, 가짜 뉴스 판별에 관한 선행 연구를 바탕으로 초등학생의 발달 단계와 교육 환경에 맞춘 구체적인 5차시 교육 프로그램을 설계하였다. 첫 번째 차시에서는 딥페이크와 가짜 뉴스의 정의와 종류를 소개하며, 간단한 애니메이션과 사례를 통해 학생들이 두 개념을 쉽게 이해할 수 있도록 한다. 두 번째 차시에서는 정보의 출처와 신뢰성 확인하기를 주제로 뉴스나 영상의 출처를 확인하고 신뢰할 수 있는 정보인지 판단하는 방법을 학습한다. 세 번째 차시에서는 딥페이크와 가짜 뉴스를 구별하는 법을 배우며, 학생들이 딥페이크 영상과 이미지를 분석하고, 과장된 표현이나 불일치하는 부분을 찾는 활동을 진행한다. 네 번째 차시에서는 가짜 뉴스와 진짜 뉴스를 구별하는 퀴즈 활동을 통해 다양한 뉴스를 읽고 진위를 판단하는 연습을 한다. 마지막 다섯 번째 차시에서는 디지털 환경에서의 안전한 행동 수칙을 주제로, 딥페이크와 가짜 뉴스의 위험성에 대해 토론하고, 안전하게 정보를 활용하는 방법을 다짐으로 정리한다. 이 교육 프로그램은 초등학생들이 이해하기 쉬운 시청각 자료와 게임형 학습 활동을 포함하여, 학생들이 흥미를 가지고 적극적으로 참여할 수 있도록 구성되었다. 연구는 개발된 프로그램의 구조와 내용을 이론적으로 검토하고, 교육 전문가와 초등학교 교사를 대상으로 한 설문 조사 및 인터뷰를 통해 프로그램의 타당성과 현장 적용 가능성을 평가하였다. 결과적으로, 본 프로그램은 초등학생들의 디지털 리터러시 함양과 정보 감별 능력 향상에 긍정적인 영향을 미칠 수 있음을 확인하였다. 향후 연구에서는 실제 초등학교 수업에 본 프로그램을 적용하여 학생들의 학습 효과를 측정하고, 교육 환경에 맞게 프로그램을 지속적으로 개선할 필요가 있다. 또한, 교사 연수 프로그램과 학부모를 위한 가정 연계 학습 자료를 제공하여, 프로그램의 실질적인 활용도를 높이하고자 한다. 본 연구는 초등학생들이 디지털 환경에서 안전하게 정보를 이용하고, 딥페이크와 가짜 뉴스를 효과적으로 구별할 수 있는 기초적인 능력을 기르는 데 기여할 수 있으며, 초등학교 교육과정에서 디지털 리터러시 교육의 중요성을 강조하는 데 의의가 있다.

### Key words

Digital Literacy, Deepfake Detection, Fake News Education, Elementary School Students

### 참고문헌

- [1] J. Hwang, E. Choi, and J. Han. "Development and application of ethics education convergence projects using deepfake apps." Fact check trend report "FACT" -.5 (2022): 23-27.

\*제주대학교 일반대학원 융합정보보안학협동과정 박사과정, \*\*제주대학교 교육대학 초등컴퓨터교육전공 교수, 교신저자(namjepark@jejunu.ac.kr)

※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269).

# 제주대학교의 제주도 지역형 특화 늘봄학교 운영 모델과 선도 교육 콘텐츠 다양화 사례

이채린\*, 김이경\*\*, 박남제\*\*\*

## Jeju National University AI and Future Technology Experience Neulbom School Program Model

Chaelin Lee\*, Leekyeong Kim\*\*, and Namje Park\*\*\*

제주대학교는 초등학교 학생들에게 창의적 사고와 미래 기술을 체험할 기회를 제공하기 위해 AI를 활용과 미래 기술을 체험할 수 있는 늘봄학교 프로그램을 운영하고 있다. 이 프로그램은 4차 산업혁명 시대에 필수적인 기술인 인공지능(AI), 블록체인, 항공우주 기술 등을 학생들이 친숙하게 접할 수 있도록 구성되었다. 주요 프로그램으로는 "놀이로 배우는 첨단 미래 신기술"과 "꿈속의 미래 우주 여행"이 있다. 첫 번째 프로그램인 "놀이로 배우는 첨단 미래 신기술"은 AI와 블록체인의 기본 원리를 초등학교 1학년 학생들이 쉽게 이해할 수 있도록 기획되었다. 학생들은 AI 머신러닝 개념을 담은 동화책을 통해 AI의 개념을 자연스럽게 접하고, 보드게임을 통해 블록체인 기술의 기본 원리를 체험하며 첨단 기술에 대한 이해를 넓힌다. 또한, 이 프로그램은 학생들이 놀이를 통해 비행기 제작, 인공지능을 활용한 미술 작품 생성, 사진과 라이트 아트를 이용한 예술 활동 등 다양한 체험을 제공함으로써 기술의 원리와 활용 방식을 학습하고, 창의적 사고와 문제 해결 능력을 키울 수 있도록 돕는다. VR 가상현실 체험도 포함되어 있어 학생들이 AI와 같은 미래 신기술에 대해 흥미를 가지고 탐구할 수 있는 기회를 제공한다. 두 번째 프로그램인 "꿈속의 미래 우주 여행"은 항공우주 과학을 중심으로 학생들의 탐구력을 자극하고자 한다. 이 프로그램은 항공기와 우주에 대한 기초 개념을 학생들에게 전달하며, 글라이더와 오르니토퍼 등 다양한 비행체를 만들어 날려보는 활동을 통해 항공 과학을 경험할 수 있게 한다. 또한, 광섬유를 활용해 별자리를 만드는 활동을 통해 우주와 천문학에 대한 관심을 높이고 탐구 정신을 배양할 수 있다. 이러한 활동을 통해 학생들은 항공과 우주에 대한 호기심을 키우고, 생활 속에서 미래 기술이 적용된 사례를 살펴봄으로써 자신의 일상에서 과학적 탐구의 흥미를 더할 수 있다. 이와 같은 늘봄학교 프로그램은 제주대학교가 지역사회와 협력하여 기획된 창의적 교육 프로그램으로, 미래 신기술을 놀이와 체험을 통해 자연스럽게 습득할 수 있도록 한다. 이러한 교육은 학생들이 놀이와 경험을 통해 AI와 미래 기술에 대한 기본 지식과 흥미를 키우며, 앞으로의 변화된 사회에서 필요한 창의적 문제 해결력을 기르는 데 중점을 둔다.

### Key words

AIDT(Artificial Intelligence Digital Textbook), Future Technology, Neulbom School, Aerospace Science, Experiential Learning

### 참고문헌

- [1] Ministry of Education, "2024 Noolbom School Implementation Plan," Press Release on Ministry of Education Website(<https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=294&boardSeq=97914&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=020402&opType=N>), February, 2024.

\*제주대학교 창의교육거점센터, \*\*제주대학교 융합과학기술사회연구소, \*\*\*제주대학교 교육대학 초등컴퓨터교육전공, 교신저자(namjepark@jejunu.ac.kr)

※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269).

# 제주대학교 융합과학기술사회연구소의 필란트로피를 향한 인문학과 기술의 공존

정유진\*, 김미진\*, 주연수\*, 최근배\*\*

## The Coexistence of Humanities and Technology for Philanthropy

Yujin Jung\*, Mijin Kim\*, Yeonsu Joo\*, and Keunbae Choi\*\*

교육부는 한국연구재단과 함께 인문사회 기초학분분야 및 사회문제해결형 연구, 대학교육과 연계 등을 통해 연구소의 사회적·교육적 기능을 확대하여 인문사회연구소의 자생을 돕는 사업을 추진하고 있다. 제주대학교 융합과학기술사회연구소는 2012년 과학기술사회연구센터로 출범하여 본 사업을 2019년에 유치하고 5년간 130억원의 지원금으로 연구센터의 자생을 도모하고 있다. 지난 5년간 센터는 초·중등 인문소양의 동서양 고전 속 미래 융합IT인문학 및 신(新)교수·학습법을 적용한 실증 리빙랩 기반 구축을 위한 연구를 수행하기 위해 제주대학교에 미래융합IT인문학 교양과목을 개설 하여 운영하고 있다. 미래기술IT인문학 과목은 미래 사회가 다양한 학문과 기술이 융합하여 새로운 지식과 가치를 창출하는 미래형 지능 정보화 사회에 살아갈 디지털 네이티브(digital native)세대의 미래사회 핵심역량을 키워줄수 있는 교사의 역량 함양을 목표로 하고 있다. 미래 사회가 요구하는 융합적이고 경쟁력 있는 인재를 양성하기 위해서는 학교 교육이 보다 유연하고 다양한 교육과정이 도입 되어야 하고 이는 학생들의 잠재력과 바람직한 가치관을 찾고 키워줄 수 있는 융합·인성 교육이 중요하다는 것을 골자로 한다. IT-인문학이라는 신 개념을 도입하고 연구하여 기반을 다져가기 위해 국내·외 유수의 연구기관과 협력 네트워크를 구축하고 IT인문학의 교육 모델을 제주에서 전국으로, 전국에서 해외로 글로벌라이제이션(glocalization)을 실천해 나아가고 있다. 학구적 협력 노력의 구체화와 가시화를 위해 지난 2022년부터 융합과학기술사회연구소 학술논문지를 출간하며 센터의 명칭을 연구소로 격상하고 자생화 노력을 순차적으로 진행하고 있다. 그동안 융합과학기술사회연구소는 초·중등생 대상 뿐만 아니라 지역 사회에 인간성에 대한 디지털 온톨로지 확립하고 미래융합IT의 기술을 인문학 연구의 기술적 보조도구로서 인문학 영역 확대에 기여할 수 있도록 공익법인의 설립을 목전에 두고 있다. 연구소가 지역사회의 필란트로피(Philanthropy)를 위하여 개인보다는 인류(humanity)라는 공동체 차원의 인류 발전을 위한 서비스를 구축해 나아가기 위한 법인 설립으로, 연구·개발의 대의를 세우기 위한 첫걸음을 내딛기 위하여 또한 본 인문사회연구소 지원사업의 방점을 찍고자 최종 목표를 설정하였고 마무리 단계에 접어 들고 있다. 본 연구소의 '필란트로피를 향한 인문학과 기술의 공존'은 인간 중심적이고 지속 가능한 사회를 만들기 위해 인문학적 가치와 첨단 기술이 상호 보완적으로 작용하는 것을 의미한다. 필란트로피, 즉 공공의 선을 위한 기부와 봉사 활동은 전통적으로 인류애와 윤리적 관점을 중시했으며, 이러한 활동의 기저에는 인문학적 통찰이 자리해 왔고 여기에 기술이 결합되면, 단순히 자원을 나누는 것을 넘어 사회문제에 대한 접근 방식을 확장하고, 더욱 효율적이고 혁신적인 해결책을 마련할 수 있습니다.

또한 미래융합IT인문학 과목은 제주대학교에 개설된 교양과목에 지나지 않고 국가평생교육진흥원에서 운영하는 한국형 온라인 공개강좌인 K-MOOC에 동 주제로 '인문학으로 만나는 디지털 혁명과 미래기술'로 운영하고 있으며, 대한민국 최초의 사이버대학교인 한국열린사이버대학교(OCU)에 미래기술인문학 과목을 운영하여 해마다 300명 이상의 이수자를 배출하고 있다. 앞으로 융합과학기술사회연구소는 IT인문학을 새로운 학문으로 인정받아 미래기술과 인류를 위한 필란트로피를 구축해 나가고자 한다.

### Key words

Digital Education Innovation, IT Humanities, Philanthropy, Advanced Future Technology, The Coexistence

\*제주대학교 융합과학기술사회연구소 \*\*제주대학교 교육대학 초등수학교육전공, kbchoi@jejunu.ac.kr(교신저자)

※이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022S1A5C2A04092269).

# 공공데이터 오류 유형 분류를 위한 BERT 파인튜닝 방법 연구

유희정\*, 김장원†

## A Study on BERT Fine-Tuning Methods for Classification of Public Data Error Types

Heejung Yu\*, and Jangwon Gim†

### 요 약

AI 시장 규모가 커짐에 따라 AI 학습에 필요한 고품질 데이터의 수집·생성·관리의 중요성이 더욱 부각되고 있다. 한편, 공공데이터포털은 약 9만여 건의 공공데이터를 개방하고 있으나, 일부 데이터는 AI 학습데이터로 활용하기에 품질상 한계가 있다는 지적이 있다. 본 논문에서는 이러한 공공데이터의 오류 유형을 분석하고, BERT 기반 모델을 활용하여 오류 유형을 자동으로 분류하는 방법을 제안한다.

### Abstract

As the AI market continues to grow, the importance of collecting, generating, and managing high-quality data for AI training is becoming increasingly emphasized. Meanwhile, the Public Data Portal provides access to approximately 90,000 public datasets; however, some of these datasets have quality limitations, posing challenges for their use as AI training data. This paper analyzes error types within such public datasets and proposes a method to automatically classify these error types using a BERT-based model.

### Key words

public open data, data analysis, error type classification, BERT

## 1. 서 론

지난 3월 글로벌 리서치회사인 마켓츠앤마켓츠 (MarketsandMarkets)는 전 세계 인공지능(AI) 시장 규모가 2023년 1502억 달러(약 200조 원)에서 2030년에는 1조3452억 달러(약 1800조 원)로 9배가량 성장할 것으로 예측했다[1]. 이러한 AI는 데이터로부터 학습하기 때문에, 고품질 데이터를 수집·생성·관

리하는 것의 중요성 역시 커지고 있다.

한편, 공공데이터포털은 2013년 공공데이터법 시행과 함께 서비스를 개시했으며, 현재까지 약 9만여 건의 공공데이터가 포털에서 개방되고 있다[2]. 그렇지만, 공공데이터는 88% 이상의 높은 완결성을 보이지만, 상대적으로 낮은 정확성을 가지므로 데이터의 품질 개선이 필요하다[3]. 즉, 정확한 데이터를 기반으로 AI 모델 학습을 위해서는 품질 향상을 위

\* 국립군산대학교 소프트웨어학과 yu001206@kunsan.ac.kr

† 국립군산대학교 소프트웨어학과 (교신저자) jwgim@kunsan.ac.kr

한 추가적인 데이터 처리가 필요하다.

따라서 본 논문에서는 공공데이터의 오류 유형을 분석하고, BERT 파인튜닝 기반의 유형 분류 방법을 제시한다.

## II. 관련 연구

### 2.1 공공데이터 오류 유형 분석

공공데이터에 포함된 데이터 중에서 주소 데이터의 오류 유형을 분석한 연구가 수행되었다[4]. 주소 데이터 정제 알고리즘을 통해 주소 누리집 도로명주소 검색 API 결과와 실제 데이터세트에 포함된 값을 비교하고 이때 오류 유형별 출현 규칙에 따라서 유형 분석을 수행하였다[4,5]. 그렇지만 전통적인 방법인 규칙 기반 오류 유형 분류는 사전에 정의된 규칙에 종속적이므로 데이터세트에 포함된 값들 중에서 예기치 않은 구문 오류, 복잡한 의미상의 오류 유형을 식별하고 분석에 한계가 있다.

### 2.2 BERT

BERT(Bidirectional Encoder Representations from Transformers)는 Google AI Language 팀에서 개발한 딥러닝 기반 언어 모델이다. BERT는 기존 모델들과 달리 학습 시 텍스트의 양방향 문맥을 동시에 고려하여 단어와 문장의 의미를 더욱 정확하게 파악할 수 있다. 또, 파인튜닝 과정에서 최소한의 추가 파라미터만 학습하여 광범위한 자연어처리 작업에서 우수한 성능을 보인다.

## III. BERT 파인튜닝 기반의 오류 분류

공공데이터 오류 유형 자동 분류를 위한 BERT 파인튜닝 방법은 그림 1과 같다. 수집한 공공데이터의 오류 유형을 분석하고 정의한 뒤, 데이터 전처리와 BERT 파인튜닝을 통해 최종적으로 오류 유형을 자동 분류한다.

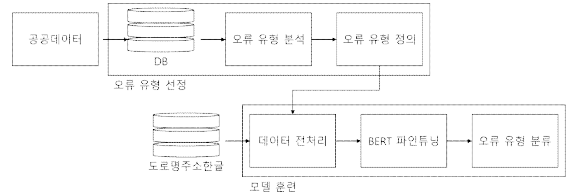


그림 35. 공공데이터 오류 유형 분류를 위한 BERT 파인튜닝 방법

Fig 1. Fine-Tuning Approach for Classifying Public Data Error Types Using BERT

### 3.1 오류 유형 분석 및 정의

공공데이터에 데이터세트에 포함된 다양한 항목 중에서 도로명주소는 다양한 형태의 오류가 발생하고 있다[4]. 따라서 본 논문에서 제안하는 방법을 적용하는 실험 데이터로 데이터세트에 포함된 도로명주소 항목을 선택하여 다양한 오류 유형을 분석한다. 따라서, 공공데이터포털에 공개된 표준데이터 세트 중에서 도로명주소 칼럼이 포함된 100건을 랜덤으로 선택하고 수작업으로 검토한 오류 유형은 표 1과 같다. 공공데이터세트에 포함된 다른 항목들의 오류보다 다양한 유형의 오류가 나타나고 있다.

표 1. 공공데이터 오류 유형 분석 예시

Table 1. Example of Public Data Error Type Analysis

번호	오류 이름	설명
1	공백문자 2개 이상	공백문자가 2개 이상인 경우
2	띄어쓰기 누락	띄어쓰기가 누락된 경우
3	붙여쓰기 안 함	붙여 써야 하는 도로명을 띄어 쓴 경우
4	도로명/건물번호 띄어쓰기 안 함	도로명주소와 건물번호(부번) 사이에 띄어쓰기가 안 된 경우
5	특수문자 포함	주소에 포함되는 [, ( ) - . ] 를 제외한 특수문자가 포함된 경우
6	다중 값	주소 성분이 2개 이상인 경우
7	건물부번 오류	건물부번이 0인 데도 표기할 경우
8	건물번호 하이픈 누락	건물번호와 부번을 잇는 '-' 누락



특히 주소 정보의 경우 단순한 문자열이 아닌 주소를 구성하는 계층적인 구성요소가 포함되어 있으므로 이를 구분하는 것은 주소 구성요소를 식별하는 것과 함께 데이터의 연계 및 통합을 위해 중요하다.

### 3.2 데이터 전처리

오류 유형 분류를 위한 학습 데이터는 특정 오류 유형에 편향되지 않도록 각 유형별 데이터 개수를 균일하게 맞춘다. 이후 BERT 파인튜닝을 위해 각 인스턴스에 포함된 토큰을 구분하고, 토큰임베딩 과정을 수행하여 파인튜닝의 입력 데이터 형태로 구축한다.

### 3.3 BERT 파인튜닝

BERT 파인튜닝은 학습 데이터셋을 이용해 오류 유형 분류 모델을 최적화하는 과정이다. 먼저, 사전 학습된 BERT 모델을 불러와 출력 레이어를 오류 분류에 맞게 조정한다. 이 출력 레이어는 3.1에서 정의한 오류 유형에 대한 다중 클래스 분류를 위해 최적화될 수 있다. 모델 훈련이 완료된 후, 검증 데이터셋을 사용해 각 오류 유형에 대한 분류 정확도와 F1-score를 평가함으로써 모델의 분류 성능을 측정할 수 있다.

## IV. 결론

본 논문에서는 고품질의 AI 학습용 데이터 구축을 위해 공공데이터에 포함된 오류 유형 분류 연구를 수행하였다. 분석된 오류 유형을 자동화된 기법으로 분류하기 위한 BERT 파인튜닝 방법을 제안하였다. 파인튜닝 기반의 오류 식별 정확도를 향상시킴으로써 보다 고품질의 AI 학습용 데이터 구축에 기여할 수 있을 것으로 기대한다.

## 참 고 문 헌

- [1] MarketsandMarkets, “Artificial Intelligence (AI) Market by Technology (ML, NLP, Context-aware AI, Computer Vision), Software (Type (Discriminative AI, Generative AI), No-code AI, Low-code AI), Hardware (Accelerators, Processors, Memory, Networking) - Global Forecast to 2030”, 2024.
- [2] 김준영, “행정안전부, 9만여 건 공공데이터 공개로 국민의 삶 변화시킨 ‘공공데이터포털’ AI 기술 활용으로 재탄생”, 정필, 2024,10.16, <https://www.jeongpil.com/1806238>
- [3] H. L. Kim, “Quality Evaluation of the Open Standard Data”, The Journal of the Korea Contents Association, vol. 20, no.9, August, 2020.
- [4] J. W. Lee, S. H. Kim, D. W. Kim and S. E. Lee, “Errors in Address Data for Address Based Integration of Building Open Data Using Open API - A Case Study Pertaining to the Address List of Korean Public Facilities -”, Journal of the Architectural Institute of Korea, vol. 38, no. 4, pp. 25-34, April, 2022.
- [5] J. W. Lee, S. H. Kim, S. E. Lee and D. W. Kim, “Address-based Integration of Building Open Data Using OpenAPI - A Case Study of the Korean Public Institution Building List -”, Journal of the Architectural Institute of Korea, vol. 37, no. 4, pp. 11-22, April, 2021.

# 지식그래프를 활용한 공공데이터 연계 방법

홍현석\*, 김장원†

## The Processes for Public Data Linkage Using Knowledge Graphs

Hyunseok Hong\*, and Jangwon Gim†

### 요약

공공데이터는 다양한 분야에서 활용 가치가 높지만, 서로 다른 기관에서 제공된 데이터의 연계와 통합에는 어려움이 따른다. 본 논문은 이를 해결하고 상호운용성을 강화하기 위해, 도로명주소의 계층적 정보를 활용한 지식그래프 기반 연계 방안을 제안한다. 데이터 정제와 검증을 포함한 프로세스를 통해 다양한 데이터의 연계와 분석을 가능하게 하여 새로운 가치를 창출할 수 있을 것으로 기대된다.

### Abstract

Public data holds substantial value across various fields, yet integrating and linking data provided by different institutions remains challenging. This paper proposes a knowledge graph-based linkage approach utilizing the hierarchical structure of road name addresses to enhance interoperability. By incorporating processes for data cleansing and validation, the proposed method facilitates comprehensive data integration and analysis, contributing to the generation of new insights.

### Key words

knowledge graph, data cleansing, public open data, data integration

## 1. 서론

우리나라는 ‘공공데이터의 제공 및 이용 활성화에 관한 법률’[1]에 따라 정부, 지방자치단체, 공공기관이 보유한 데이터를 개방하여 다양한 분야에서 새로운 가치 창출을 지원하고 있다. 그러나 서로 다른 출처의 데이터를 연계하여 유의미한 정보를 추출하기 위해서는 데이터의 상호운용이 필수적이며, 이는 여전히 도전 과제로 남아 있다.

지식그래프는 데이터 내 개체와 그들 간의 의미적 관계를 명확히 정의하여 데이터의 상호운용성을 높이는 효과적인 도구로 주목받고 있다[2].

현재 공개되고 있는 다수의 공공데이터에는 도로명주소가 포함되어 있다. 이는 시도, 시군구, 읍면, 도로명, 건물 번호로 구성되어 있으며 건물의 고유 식별자 역할을 한다[3]. 또한, 도로명주소는 ‘도로명주소법’[4]에 의해 국가에서 체계적으로 관리되어 표준화된 주소 정보를 제공하므로, 다양한 분야의 공

\* 국립군산대학교 소프트웨어학과 (hsuk@kunsan.ac.kr),

† 국립군산대학교 소프트웨어학과 (교신저자) (jwgim@kunsan.ac.kr)

공공데이터를 연계하는 데 공통 항목으로 활용될 수 있다. 그러나 공공데이터로 제공되는 일부 데이터에서 도로명주소에 다양한 오류가 식별되고 있다. 오류가 있는 데이터는 지식그래프 생성 시 개체가 식별되지 않아 데이터의 누락으로 이어질 수 있으며 지식그래프의 불완전성을 증가시켜 추론과 분석의 정확도가 저하될 수 있다[5]. 따라서 본 논문에서는 공공데이터의 상호운용성을 위해 지식그래프를 활용한 공공데이터 연계 방법을 제안한다. 본 논문의 구성은 2장에서는 관련여구를 정리하고, 3장에서는 제안 방법에 대한 개요를 설명하고, 4장에서 결론을 맺는다.

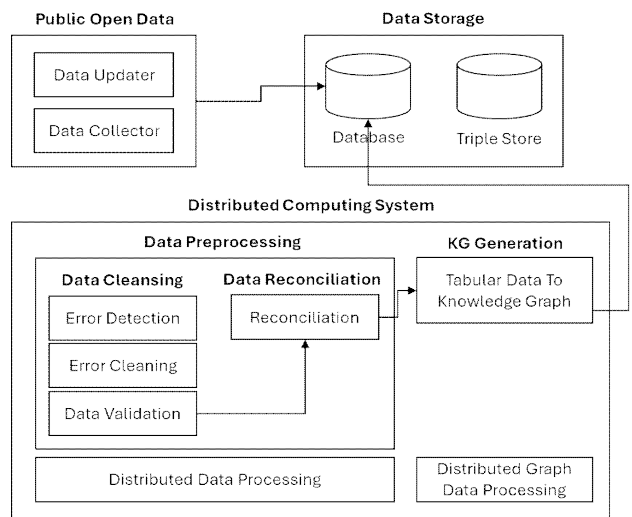
## II. 관련연구

다양한 출처의 데이터를 통합하고 상호운용성 향상 방안으로 지식그래프를 활용하는 연구가 여러 분야에서 활발히 진행되고 있다. 이 중 도서관 데이터의 경우, 데이터의 상호운용성을 개선하기 위해 표준 온톨로지 어휘를 활용하여 메타데이터를 설계해 도서관 지식그래프를 구축한 연구가 진행되었다 [2]. 현재 공개되고 있는 도서관 데이터에는 도서 대출 현황, 연구데이터 등 다양한 데이터가 LOD(linked open data) 형태로 제공되고 있으나 데이터 수준에서 연계의 목적이 실현되지 않고 있는 것이 현실이다. 따라서 이러한 분리된 데이터의 연결 및 통합을 위해 지식그래프를 적용하고 데이터의 분류체계와 메타데이터의 불일치를 해소하였으며 분산된 도서관 데이터 사이의 상호운용성을 지원한다. 건물의 고유 식별자 역할을 할 수 있는 도로명주소는 데이터를 연계할 수 있는 공통 항목으로써 활용될 수 있다. 또한 도로명주소는 공공데이터에서 광범위하게 활용되고 있으나 여전히 품질의 문제가 지속적으로 논의되고 있다. 도로명주소를 데이터 연계의 공통 항목으로 사용을 위해 데이터의 정제는 필수적이며 다양한 정제 방법이 연구되고 있다[6-7]. 따라서 본 논문은 공공데이터의 상호운용성 향상을 위한 방안으로 데이터 수집부터, 정제, 조정, 지식그래프 생성까지의 모든 프로세스를 포함한 공공데이터의 연계 및 정제 절차를 제안한다.

## III. 제안 방법

본 논문이 제안하는 공공데이터 연계 방법은 데이터 수집, 저장, 전처리, 지식그래프 생성의 네 가지 주요 단계로 구성된다. 공공데이터 연계에서 도로명주소는 표준화된 국가 관리 체계에 따라 시도, 시군구, 읍면 등 계층적 구조로 구성되어 있어 다양한 데이터 간 연계 시 공통 항목으로 활용될 수 있다는 점에서 핵심적 역할을 수행한다. 특히 도로명주소는 전국적으로 광범위하게 사용되어 데이터 출처 간 통합성 확보에 유리한 특성이 있으며, 최신 상태를 유지함으로써 데이터의 신뢰성과 활용도를 높일 수 있다.

그림 1은 제안 방법의 개요도를 보인다. 공공데이터에 대한 연계 방법의 첫 단계는 공공데이터와 도로명주소 데이터를 정기적으로 수집하여 최신 상태를 유지하고, 이를 파일과 데이터베이스에 저장해 관리하는 것이다. 저장된 데이터는 파일시스템, 데이터베이스, 트리플 저장소로 관리되며, 각각 정제 대상과 결과의 저장, 도로명주소 데이터 저장, 생성된 지식그래프의 RDF 트리플 저장한다. 수집된 도로명주소 데이터는 오류 추출, 정제, 검증의 전처리 과정을 거친다.



<그림 36. 지식그래프를 활용한 공공데이터 연계 방법의 개요도>

< Fig 1. Diagram of the process utilizing knowledge graphs for public open data linkage>

전처리 단계에서 오류 유형을 정의해 데이터를 분석하고, 중복 오류나 문법적 오류를 정제한다. 특히 데이터의 계층 구조를 반영한 트리 형태의 데이터 구조를 적용하여 검색 효율성을 높이고, 정방향 및 역방향 검색을 통해 누락된 속성이나 오타를 정제할 수 있다. 전처리된 도로명주소 데이터에는 계층적 식별번호를 부여하여 상위 개념과 하위 개념 간의 관계를 반영하며, 이를 바탕으로 지식그래프를 생성하고 RDF 형식으로 변환하여 트리플스토어에 저장한다. 제안된 연계 방법을 통해 공공데이터 간 상호연용성이 개선되며, 다양한 데이터 연계와 고차원적 분석이 가능하다.

#### IV. 결론

본 논문은 공공데이터의 상호연용성을 강화하기 위해 지식그래프 기반의 연계 방법을 제안한다. 데이터의 정제와 검증 과정을 통해 지식그래프 생성 프로세스를 정의하고, 계층적 정보를 활용하여 상위 개념 간의 데이터셋 연계를 통해 데이터 연결의 확장성을 높이며, 하위 계층 정보를 바탕으로 데이터셋 간 연계를 통해 데이터의 다양성을 확보할 수 있는 방안을 제시하였다. 이를 통해 공공데이터의 상호연용성을 높이고, 지식그래프 기반의 추론 및 분석을 통해 새로운 가치를 창출할 수 있을 것으로 기대한다. 향후 연구에서는 도로명주소뿐 아니라 공공데이터의 모든 속성을 통합할 수 있는 지식그래프 모델링을 통해 다양한 데이터 간의 연계를 확장하고, 이를 활용한 시스템의 실효성을 분석할 예정이다.

#### 참 고 문 헌

- [1] 행정안전부, “공공데이터의 제공 및 이용 활성화에 관한 법률“ 제 19408호, 2023.
- [2] 송채은, “도서관 데이터의 상호 운용상호연용성 개선을 위한 지식그래프 적용방안”, 고려대학교 석사학위논문, 2023.
- [3] 박하람 and 김학래, “공공데이터를 활용한 주소 기반 데이터 분석과 제안”, 디지털콘텐츠학회논문지, Vol. 25, No. 9, pp. 2617-2626, 2024.

- [4] 행정안전부, ”도로명주소법“, 제 17574호, 2020.
- [5] Destandau M. and Fekete J.-D., “The missing path: Analysing incompleteness in knowledge graphs.”, Information Visualization, Vol. 20, No. 1, pp. 66-82, 2021.
- [6] 박필민 and 임성수, “학습 데이터 자동 구축을 통한 양방향 게이트 수노한 유닛 기반 도로명주소 파싱 모델”, 한국측량학회지, Vol. 41, No. 5, pp. 301-310, 2023.
- [7] 송재용, “행정구역 위계정보와 단어 유사도를 이용한 도로명주소 변환 모델”, 고려대학교 석사학위논문, 2014.

# PPO 기반의 사전 캐싱 차량 선정 및 캐싱량 최적화 방안

남영주\*

## A PPO-Based scheme for Optimizing Precaching Vehicle Selection and Caching Amounts

Youngju Nam\*

### 요 약

자율 주행 차량의 보급으로 인해 차량 네트워크에서의 모바일 데이터 트래픽이 급증하고 있으며, 이는 제한된 무선 자원을 가진 도로변 기지국과 기지국에 과부하를 초래하고 있다. 이 문제를 해결하기 위해 사전 캐싱 전략이 도입되었으나, 동적인 차량 네트워크 환경에서 최적의 사전 캐싱 차량을 선택하는 것은 어려운 과제다. 본 논문에서는 PPO 기반의 강화 학습을 활용하여 사전 캐싱 차량의 선정과 사전 캐싱량을 최적화하는 방안을 제안한다. 향후 시뮬레이션하여, 제안된 방안이 기존의 방식보다 기지국의 통신 범위 밖에서 콘텐츠 전달 성공률을 향상시킴을 검증할 예정이며, 이는 네트워크 성능과 서비스 품질 개선에 기여할 것으로 기대된다.

### Abstract

Due to the widespread adoption of autonomous vehicles, mobile data traffic within vehicular networks has surged, causing overloads on roadside units and base stations with limited wireless resources. To address this problem, precaching schemes have been introduced; however, selecting the optimal precaching vehicles in dynamic vehicular network environments remain challenging tasks. In this paper, we propose a scheme to optimize the selection of precaching vehicles and the determination of precaching amounts by utilizing a Proximal Policy Optimization. We plan to conduct simulations to verify that the proposed method improves content delivery success rates outside the communication range of base stations compared to existing methods, which is expected to contribute to enhancing network performance and Quality of Service.

### Key words

Internet of Vehicle, Vehicular Network, V2V, Content Precaching, Reinforcement Learning, PPO

### 1. 서 론

자율 주행 차량의 발전으로 승객들은 이동 중에

도 고화질 비디오 스트리밍, 온라인 게임 등 다양한 콘텐츠를 즐길 수 있게 되었다. 이로 인해 차량 네트워크 내 모바일 데이터 트래픽이 급증하여, 제한

---

\*군산대학교 소프트웨어학과, imnyj@kunsan.ac.kr.

※ 지원기관표기(사사표기)

된 무선 자원을 가진 도로변 기지국 (RSU)과 기지국에 과부하를 초래하고 있다. 네트워크 과부하는 지연 시간 증가로 이어져 서비스 품질과 경험 품질을 저하시킬 뿐만 아니라, 안전에 중요한 응용 프로그램에도 부정적인 영향을 미칠 수 있다.

이러한 문제를 해결하기 위해 사전 캐싱 전략이 도입되었으며, 이는 차량의 이동 패턴을 기반으로 미래의 콘텐츠 요청을 예측하여 필요한 데이터를 미리 저장하는 방식이다 [1]. 특히 차량 간 사전 캐싱에서는 사전 캐싱 차량을 선정하여, RSU 범위 밖에서 요청 차량에게 콘텐츠를 전달한다. 그러나 동적인 차량 네트워크 환경에서 최적의 사전 캐싱 차량을 선택하고 적절한 사전 캐싱량을 결정하는 것은 여전히 어려운 과제다.

본 논문에서는 Proximal Policy Optimization (PPO) 기반의 강화 학습을 활용하여, 사전 캐싱 차량 선정과 사전 캐싱 양을 최적화하는 방법을 제안한다. 사전 캐싱 문제를 강화 학습 모델로 정의하여, RSU 범위 밖에서 콘텐츠 전달 성공률을 최대화하는 정책을 학습한다. 최종적으로 PPO 기반 학습 모델의 사전 캐싱 방안이 기존의 방식보다 콘텐츠 전달 성능을 향상시킴을 기대한다.

## II. PPO 기반 V2V 사전 캐싱 방안

본 연구에서는 Proximal Policy Optimization(PPO) 기반의 강화 학습을 활용하여, 동적인 차량 네트워크 환경에서 사전 캐싱 차량의 선정과 사전 캐싱량을 최적화하는 방법을 제안한다. PPO는 정책 그래디언트 방법 중 하나로, 정책의 업데이트 시 큰 변화로 인한 불안정성을 줄이면서도 효율적인 학습이 가능하다.

우선, 사전 캐싱 문제를 강화 학습 프레임워크로 모델링하였다. 에이전트는 RSU로 설정되며, 상태는 현재 네트워크의 정보로 구성된다. 여기에는 차량들의 위치, 속도, 요청 콘텐츠의 종류 등이 포함된다. 행동은 사전 캐싱 차량의 선정과 각 차량에 할당할 사전 캐싱량을 결정하는 것이다. 보상은 RSU 범위 밖에서 콘텐츠 전달 성공 여부와 네트워크 자원의 효율적인 사용을 반영하여 설계하였다.

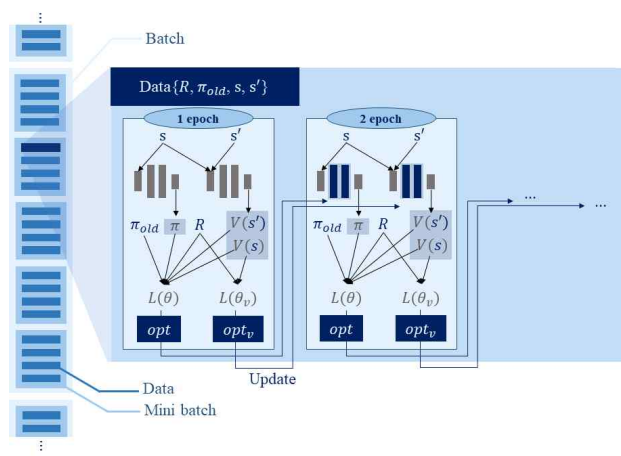


그림 1. PPO 모델의 학습 과정  
Fig. 1. Learning process of a PPO Model

PPO 알고리즘을 통해 에이전트는 그림 1과 같이 시뮬레이션 환경에서 반복적인 학습을 수행한다. 각 에피소드에서 에이전트는 현재 상태를 관찰하고 최적의 행동을 선택하여 실행한다. 그 결과로 얻어진 보상을 기반으로 정책을 업데이트하며, 이 과정에서 클리핑 기법을 사용하여 정책의 급격한 변화를 방지한다.

## III. 결 론

본 논문에서는 Proximal Policy Optimization(PPO) 기반의 강화 학습을 활용하여, 동적인 차량 네트워크 환경에서 사전 캐싱 차량의 선정과 사전 캐싱량을 최적화하는 방안을 제안하였다. RSU 범위 밖에서의 콘텐츠 전달 성공률이 향상되고 네트워크 지연 시간이 감소함을 기대한다. 이는 서비스 품질과 경험 품질의 향상으로 이어지며, 네트워크 자원의 효율적인 활용을 가능하게 한다. 향후 실시간 환경에서의 적용 및 다양한 네트워크 조건에서의 성능 검증이 기대된다.

## 참 고 문 헌

- [1] G. Yu, Y. He, J. Wu, Z. Chen and J. Pan, "Mobility-Aware Proactive Edge Caching for Large Files in the Internet of Vehicles," IEEE IoT J., vol. 10, no. 13, pp. 11293-11305, July 2023.

# 소프트웨어 비교과 프로그램 참여율 향상 방안 : 교육수요자를 중심으로

김정연\*, 이석훈\*\*, 정동원\*\*

## On Improvement of Participation Rates in Software Extracurricular Programs: Focusing on Education Consumers

Jungyeon Kim\*, Sukhoon Lee\*\*, and Dongwon Jeong\*\*

### 요 약

4차 산업혁명 시대에 맞추어 소프트웨어 중심으로 대학 교육체계를 혁신하고 AI 등 신기술 수요에 부합하는 SW전문·융합인재 양성을 위하여 소프트웨어 교육 및 사업이 활발히 운영되고 있다. 현재 대학에서는 다양한 소프트웨어 비교과 프로그램에 운영되고 있지만 학생들의 참여율을 저조한 것으로 보여진다. 이 논문에서는 대학에서 진행하고 있는 소프트웨어 비교과 프로그램의 참여율 향상을 위해 프로그램의 현황에 대하여 분석하고 프로그램 운영 방안을 제안하고자 한다.

### Abstract

In line with the era of the 4th Industrial Revolution, software education and projects are actively being operated to innovate the university education system centered on software and to foster SW professional and convergence talent that meets the demand for new technologies such as AI. Currently, universities are operating various software extracurricular programs, but student participation rate appears to be low. This paper analyzes the current status of the program and proposes an operation program to improve the participation rate in the software extracurricular program being conducted at the university.

### Key words

Kunsan National University, National Program of Excellence in Software, software education

---

\* 국립군산대학교 소프트웨어융합공학과, jjyeon@kunsan.ac.kr

\*\* 국립군산대학교 소프트웨어학과 {leha82, djeong}@kunsan.ac.kr(교신저자)

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업”을 지원 받아 수행되었음(2023-0-00065).

## I. 서 론

4차 산업혁명은 정보통신 기술(ICT)의 융합으로 이루어진 차세대 산업혁명으로, 4차 산업혁명의 주요 기술에는 빅데이터 분석, 사물인터넷, 인공지능(IoT), 나노기술(NT), 바이오기술(BT) 등이 있다. 주요 기술들은 각각의 기술로써 핵심 역할을 수행하는 것이 아닌 기술과 기술이 결합되어 융합된 새로운 기술 혁신이 일어나고 있다. 이것은 국내만 해당되는 것이 아니라 세계 주요 국가에서도 디지털 시대에 맞는 인재 양성을 위하여 다양한 노력을 하고 있다[1,2].

국내에서는 SW전문·융합인재 양성체계 구축을 위하여 2015년부터 소프트웨어 중심대학 사업은 진행하고 있다. 소프트웨어 중심대학 사업은 산업현장의 요구를 반영하기 위하여 소프트웨어 중심으로 대학 교육체계를 혁신하고 AI 등 신기술 수요에 부합하는 SW전문·융합인재 양성체계 구축을 목적으로 한다[3]. 국립순산대학교 소프트웨어 중심대학 사업에서는 4개 분야(활동 지원 프로그램, 융합교육 지원 프로그램, 대내·외 지원 프로그램, 역량강화 지원 프로그램)의 소프트웨어 비교과 프로그램을 운영하고 있다. 각 분야에서는 4개 이상의 세부 운영 프로그램이 개설되어 있으며, 재학생을 대상으로 활발하게 운영되고 있다.

소프트웨어 중심대학 사업에서는 다양한 프로그램을 진행하고 있지만 학생들의 참여율은 저조한 것으로 파악된다. 일부 분석 연구[4]가 진행되었으나 설문조사 등 교육수요자인 학생의 의견을 기반의 분석 결과를 제시하지는 않는다.

이 논문에서는 교육수요자에 대한 실제 설문조사 및 분석 연구를 수행한다. 분석 결과는 기존 비교과 프로그램의 개선 방안, 새로운 비교과 프로그램 개발 등 다양한 목적으로 활용될 수 있으며, 이 논문에서는 분석 결과를 토대로 소프트웨어 비교과 프로그램 참여율 향상 방안을 제안한다.

## II. 운영 프로그램 제안을 위한 응답 분석

설문조사를 위해 교육수요자를 저학년과 고학년으로 구분하며, 저학년은 1~2학년으로 고학년은 3~4학년으로 분류하여 설문조사 결과를 분석한다.

표 1은 설문조사를 위한 대상 그룹과 그룹별 설문 참여 인원을 보여준다.

표 1. 실험 데이터 세트  
Table 1. Experimental data sets

구분	조사대상	참여인원
저학년	1,2학년	45명
고학년	3,4학년	23명
총계		68명

분석을 위한 설문의 구성은 표 2와 같다. 표 2에서, 첫 번째 문항은 소프트웨어 비교과 프로그램 참여여부와 참여 횟수, 미참여시 그 이유를 알아보고 두 번째 문항은 참여하였던 프로그램과 만족도를 조사하였다. 세 번째 문항은 현재 진행 중인 프로그램 이외의 희망하는 프로그램이 있는지에 대한 질문이다.

표 2. 설문 구성  
Table 2. Survey composition

구분	내 용
A1	소프트웨어 비교과 프로그램에 참여한 경험이 있습니까?
A1-1	참여하였다면 몇 회 참여하였나요?
A1-2	미참여시, 그 이유는 무엇입니까?
A2	참여하였던 프로그램은 무엇인가요?
A3	희망하는 프로그램이 있다면 서술하십시오.

표 2에 기술된 설문 내용을 바탕으로 설문을 실시하였으며, 그 설문 응답 결과는 다음과 같다.

표 3은 <A1> 소프트웨어 비교과 프로그램에 참여한 경험이 있습니까?에 대한 응답 결과를 보여준



다. 응답 결과에서, 먼저 저학년의 경우 45명 중 63%가 참여한 경험이 있으며 37%가 참여한 경험이 없다고 응답하였다. 고학년의 경우 23명 중 98%가 참여한 경험이 있으며 5%가 참여한 경험이 없다고 응답하였다.

표 3. A1 설문 응답 결과

Table 3. A1 survey response results



표 4는 <A1-1> 참여하였다면, 몇 회 참여하였나요?에 대한 응답 결과를 보여준다. 응답 결과에서, 먼저 저학년의 경우 1회 참여한 경우가 47%로 가장 많았으며, 고학년의 경우 5회 이상 참여한 경우가 59%로 가장 많은 것으로 분석되었다.

표 4. A1-1 설문 응답 결과

Table 4. A1-1 Survey response results

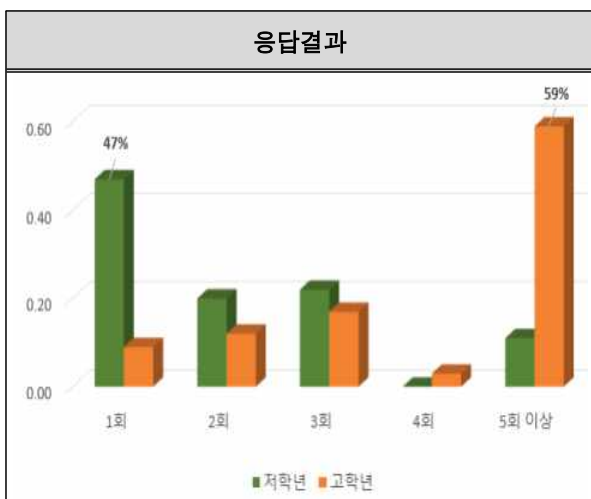


표 5는 <A1-2> 미참여시, 그 이유는 무엇입니까? 라는 설명에 대한 응답 결과를 보여준다 저학년은

프로그램 운영시간에 대한 이유가 75%로 가장 많았으며, 고학년의 경우 학업, 개인사정의 이유가 45%로 제일 많은 것으로 분석되었다.

표 5. A1-2 설문 응답 결과

Table 5. A1-2 Survey response results

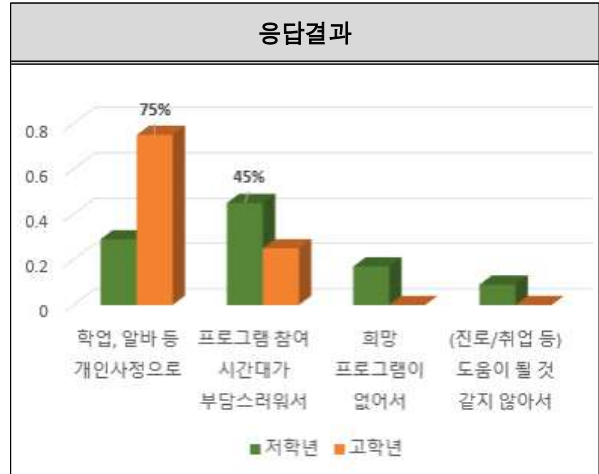
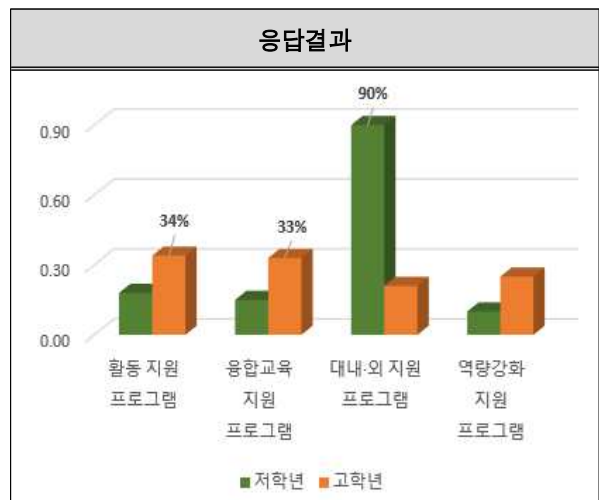


표 6은 <A2> 참여하였던 프로그램은 무엇인가요?에 대한 응답 결과를 보여준다. 본 문항은 참여한 학생들을 대상으로 진행하였다. 저학년의 경우 대내·외 지원 프로그램에 대부분의 학생들이 참여했으며, 고학년의 경우 활동 지원 프로그램과 융합 교육 지원 프로그램이 비슷하게 참여한 것으로 분석되었다.

표 6. A2 설문 응답 결과

Table 6. A2 survey response results



마지막으로, 표2 설문구성 구분의 <A3> 희망하는 프로그램이 있다면 서술하시오에 대한 응답에서, 역량강화 지원 프로그램으로 백엔드 관련 교육과 언리얼 심화 과정을 희망하는 것으로 나타났다.

본 설문에 대한 응답결과의 분석내용을 종합적으로 정리하면, 먼저 저학년의 경우 소프트웨어 비교과 프로그램에 참여한 경험을 가지고 있으나 운영시간이 부적절하여 참여하지 못한 학생들이 많은 것으로 보여진다. 따라서 저학년은 정규 교육과정 외의 시간을 파악하여 활동 지원 프로그램과 융합교육 지원 프로그램, 역량강화 지원 프로그램의 세부 프로그램을 계획하여야 하며, 표2에 기술된 설문구성 구분의 <A3> 문항에 대한 응답으로 백엔드 관련 교육과 언리얼 심화 과정은 고학년 대상으로 진행해야 하는 프로그램이며, 저학년 대상 프로그램 개설은 저학년 대상 추가 설문진행 및 분석을 통하여 개설되어야 한다.

고학년의 경우 소프트웨어 비교과 프로그램에 대다수 학생들이 5회 이상 참여한 경험을 가지고 있다. 그러나 일부 학생들은 학업, 알바, 개인사정 등으로 미참여하는 상황임에 따라 4개 분야의 소프트웨어 비교과 프로그램에서 상기 이유를 보완할 수 있는 프로그램 개설이 필요하다. 예를 들어 자격증 취득 연계 프로그램 또는 에너지&SW융합, 첨단수송&SW융합 교육 프로그램 등과 같이 여러분야와 융합된 프로그램을 운영하고 융합 프로젝트 등을 계획하여 지원하는 세부계획이 필요한 것으로 판단된다.

### III. 결 론

본 논문에서는 소프트웨어 비교과 프로그램 참여율 향상을 위한 방안을 제안하였다.

저학년과 고학년 모두 활동 지원 프로그램과 융합교육 지원 프로그램(에너지&SW융합, 첨단수송&SW융합 등) 운영 계획하여 운영한다면 국립군산

대학교 소프트웨어 중심대학 사업의 소프트웨어 비교과 프로그램 참여율이 보다 높아질 것으로 기대된다. 또한 재학생들의 만족도도 높아져 성공적인 소프트웨어 비교과 프로그램 운영이 될 것으로 보여진다.

향후 연구로는 소프트웨어 중심대학 사업 가치확산 프로그램 등에 대한 연구 수행이 필요할 것으로 생각되며 조사에서 취득된 데이터 분석을 통한 상황의 원인분석과 실시간 데이터 분석을 통한 트렌드 파악 등 다양한 관점에서의 분석과 결론 도출을 하는 것이 목표이다.

### 참 고 문 헌

- [1] 뉴스핌, "앞서간 해외 AI교육 현장, 한국은?", 2020, <https://www.newspim.com/>.
- [2] AI타임스, "핀란드는 유럽 AI의 중심... 헬싱키대학은 최적의 AI 학습 기관", 2024, <https://www.aitimes.com/>.
- [3] SW중심대학협의회, 2024, <https://www.swuniv.kr/>.
- [4] DongwJeong, Jangwon Gim, "On the National Program of Excellence in Software of Kunsan National University", Proceedings of the KIIT Summer Conference, 2024.

# On CRN-based Target Sensing in Farms

Ding Xutong\* and Dongwon Jeong\*\*

## Abstract

This paper discusses the performance testing of camera-radar networks and current mainstream 3D target detection algorithms (including camera-based, radar, LiDAR, and other methods) in different environments. The CRN detection method has high accuracy and efficiency in the farm environment. The study highlights the potential of CRN-based systems to improve autonomous agricultural robots, resulting in greatly increased operational efficiency while reducing reliance on human labor.

## Key words

3D Object Detection, Camera Radar Net, Agricultural Automation

## I. Introduction

With the rapid advancements in agricultural automation, 3D object recognition technologies are increasingly being applied in farm environments. Unlike traditional 2D recognition, 3D object detection provides superior capabilities in spatial positioning, distance perception, and depth information. Cameras provide high-resolution 2D images, capturing details such as the color and texture of object, while radar provides depth information, maintaining detection capabilities even in low-light conditions or when objects are obscured.

In a farm setting, the primary role of 3D object detection is to help robots identify paths and the surrounding plants and people, ensuring safe and accurate navigation. Recognizing paths is crucial for the robot to avoid obstacles, maintain a safe distance, and plan auxiliary picking routes. Additionally, farmers and other workers frequently appear in orchards,

making it important to identify and avoid humans. Through Camera-Radar networks[1], robots can distinguish between target objects such as fruits, trees, and people, enabling accurate object detection and scene understanding to ensure safe and efficient completion of orchard picking tasks.

## II. Related Works

Camera-based 3D object detection infers the 3D structure and location of objects in a scene by analyzing a single 2D image. This task relies on recovering depth information from the 2D image, which is generated through the perspective projection process. BEVFormer[2] uses spatiotemporal attention mechanisms to effectively convert multiple camera perspectives into a unified bird's-eye view. DETR3D[3] introduces a 3D-to-2D query mechanism, establishing associations between multi-view features and queries through a cross-attention mechanism, enabling precise

---

\* Department of Software Convergence Engineering, Kunsan National University, dxt2321@gmail.com

\*\* Department of Software Convergence Engineering, Kunsan National University(Corresponding author), djeong@kunsan.ac.kr

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065).

3D object localization. Camera-based systems provide rich color and texture information but struggle with depth estimation in complex environments and are susceptible to environmental influences.

LiDAR-based 3D object detection operate by emitting laser pulses and measuring the time delay of these pulses reflected off objects in the environment. This time delay is used to calculate the distance between the laser beam and the target, generating highly accurate 3D point cloud data. PV-RCNN[4] enhances detection performance by integrating point cloud features with voxel features. So some studies tend to take lidar methods and try to incorporate image information, such as BEVFusion[5].

### III. Experimental Design

To evaluate and compare the performance of various 3D object detection algorithms, including camera-based, Lidar-based, and camera-radar fusion networks (CRN), we will design a series of experiments focusing on accuracy, inference speed, and adaptability in diverse environmental conditions. This experimental design aims to analyze these algorithms under conditions relevant to agricultural and other dynamic environments where robust object detection is essential.

In the accuracy comparison part of the experimental scheme, this paper mainly compares the recognition accuracy performance of different 3D target recognition algorithms in a variety of typical environments (such as day, night, rain and fog weather, complex terrain, etc.). For this purpose, we chose nuScenes, a widely used standard dataset, to conduct the experiment.

mAP is the average of precision values at different recall thresholds, measuring the overall accuracy of object detection across various confidence levels. IoU measures the overlap between the predicted and ground truth bounding boxes, calculated as the area of

intersection divided by the area of union. It is used to assess the accuracy of object localization. The recognition accuracy of the algorithm is evaluated quantitatively by using the above indexes. The design of the experiment is shown in Fig. 1.

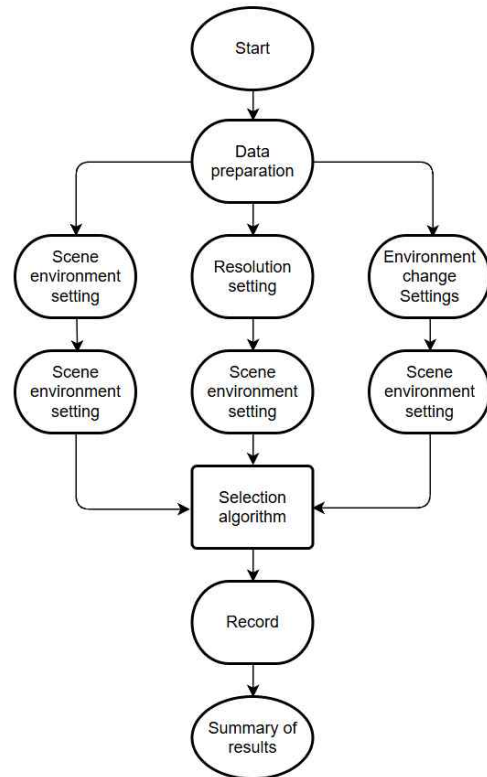


Fig. 43. Experimental Flowchart.

The purpose of the speed comparison experiment is to evaluate the inference speed of each algorithm under different resolution conditions, focusing on its performance in the application scenario with high real-time requirements. By comparing the inference speed of different algorithms at various resolutions, this experiment can reveal the computational complexity of the algorithm and help evaluate its applicability in resource-constrained environments. This is especially important for 3D object recognition tasks that require real-time processing, such as path planning and obstacle detection for agricultural robots on farms.

The evaluation indexes of this experiment are shown in Table. 1.

Table. 1. Evaluation Metrics.

Experiment Type	Metrics	Variables
Precision Comparison	mAP/loU	Scene types
Speed Comparison	FPS	Resolution
Scene Adaptability	FPR	Scene types

FPR means the proportion of false positives among all actual negatives, reflecting the algorithm's tendency to incorrectly identify non-target objects as targets. Scene types include day, night, rain, fog, complex terrain, etc.

The scene adaptability experiment aims to evaluate the performance and reliability of different algorithms in complex or extreme environments, such as low light, harsh weather, and partial occlusion. This experiment simulates various extreme scenario conditions to test the adaptability of each algorithm and records its false positive and false negative rates under these conditions. Specifically, controlled lighting and occlusion conditions are used to simulate lighting changes between day and night, as well as visibility obstructions in foggy or dusty environments.

#### IV. Conclusion

Through the design of the above experiments, this paper will systematically evaluate the performance of 3D target detection algorithms based on camera, Lidar and CRN multi-modal fusion in agricultural scenes, covering various key requirements in practical applications from detection accuracy, real-time performance, environmental adaptability and system stability. These experimental results will provide theoretical support and practical guidance for robot navigation, fruit picking and other agricultural tasks in agricultural automation scenarios.

Radar calculates the distance and velocity of a target by emitting radio waves and receiving the reflected waves. The millimeter-wave electromagnetic waves used by radar have lower atmospheric

attenuation and better penetration capabilities in rain, smoke, and dust. These characteristics allow radar sensors to operate in all-weather conditions, even under harsh weather. However, radar sensors have lower angular resolution, making it difficult to distinguish objects that are close to each other, leading to sparse point cloud distribution. This also provides a direction for future optimization.

#### References

- [1] Y. Kim, et al. "Crn: Camera radar net for accurate, robust, efficient 3d perception," Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
- [2] Z. Li, et al. "Bevformer: Learning bird's-eye-view representation from multi-camera images via spatiotemporal transformers." European conference on computer vision, Cham: Springer Nature Switzerland, 2022.
- [3] Y. Wang, et al. "Detr3d: 3d object detection from multi-view images via 3d-to-2d queries," Conference on Robot Learning, PMLR, 2022.
- [4] S. Shi, et al. "Pv-rcnn: Point-voxel feature set abstraction for 3d object detection," Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020.
- [5] Z. Liu, et al. "Bevfusion: Multi-task multi-sensor fusion with unified bird's-eye view representation," 2023 IEEE international conference on robotics and automation (ICRA), IEEE, 2023.

# Design and Implementation of QR-Based Shared Ledger System

Wenjie Zhang\* and Dongwon Jeong\*\*

## Abstract

This paper proposes a QR code-based shared ledger system, developed to provide individuals and teams with a convenient means of recording and sharing economic activities. Recognizing the limitations of traditional paper ledgers and spreadsheets, the system leverages QR technology for easy ledger sharing. Users can create and manage ledgers for expenses and incomes, inviting others to participate. The implementation utilized uni-app for frontend, Spring Boot for backend, and MySQL for data management. Evaluation results indicate that the QR-based shared ledger system is user-friendly and enhances collaborative financial management.

## Key words

QR code, shared ledger, financial management, mobile app, system design

## I. Introduction

The QR-based shared ledger system was developed to address the inefficiencies of traditional financial record-keeping, which often relies on manual methods or spreadsheets. Information asymmetry problems generally exist among different users of accounting information, and only information sharing among different users can solve the information asymmetry problem. Sharing financial information with others can provide better collaboration and understanding[1]. Shared ledger systems are designed for several reasonable needs and scenarios. First, they can be used for collaboration and cost sharing. Young people in modern society usually use Dutch pay when attending events together. When friends travel, eat together, or participate in

other activities, sharing ledgers can help track everyone's spending and determine what each person should pay. Sharing ledgers among roommates can help track common expenses such as rent, water and electricity bills, and food costs. This helps to share costs fairly and reduce disputes[2]. This paper focuses on the design and implementation of a mobile and web-based ledger system that facilitates quick sharing of financial records. The system allows users to maintain collaborative financial records, accessible via QR code scanning.

## II. Related Work

Research into financial management has highlighted the convenience of QR code technology. Mobile apps have become a key tool for financial management.

---

\* Department of Software Convergence Engineering, Kunsan National University, 2001877@kunsan.ac.kr

\*\* Department of Software Convergence Engineering, Kunsan National University (Corresponding Author), djeong@kunsan.ac.kr

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065)

Many financial applications have begun to use QR code technology to provide more efficient financial records and real-time sharing. This combination promotes financial transparency and coordination collaboration by allowing users to access their financial data anytime, anywhere[3]. Previous studies have shown that QR codes streamline the process of recording financial data and sharing it in real-time[4]. However, issues with data security and privacy remain critical and require further exploration.

### III. System Design

The core of the system involves generating QR codes that allow for quick and secure sharing of ledger entries. The system architecture includes a contributor, a database for storage, and access control mechanisms. Data synchronization is achieved using WebSocket technology, ensuring real-time updates across users. Fig. 1. shows the system function diagram reference.

The flow diagram illustrates the registration, login, ledger creation, and QR-based ledger sharing process.

Refer to Fig. 2 for the system flow chart. Fig. 3 illustrates a conceptual diagram of the sharing method. Each part of the figure is explained below.

- User A (Author): The author of an account project and can invite other users by generating a QR code containing project information.
- QR Code Generator: A tool that generates a QR code containing project information that User A commonly uses.
- Shared Project Information: Data of account items such as transaction records, expenses, and revenues.
- User B (Contributor): An invited user who can scan the QR code generated by User A to add to an account item and share account information with User A.
- Shared Project Data Store: This is a database where data of account items are stored[5].
- Data Permissions/Access Control: A mechanism that

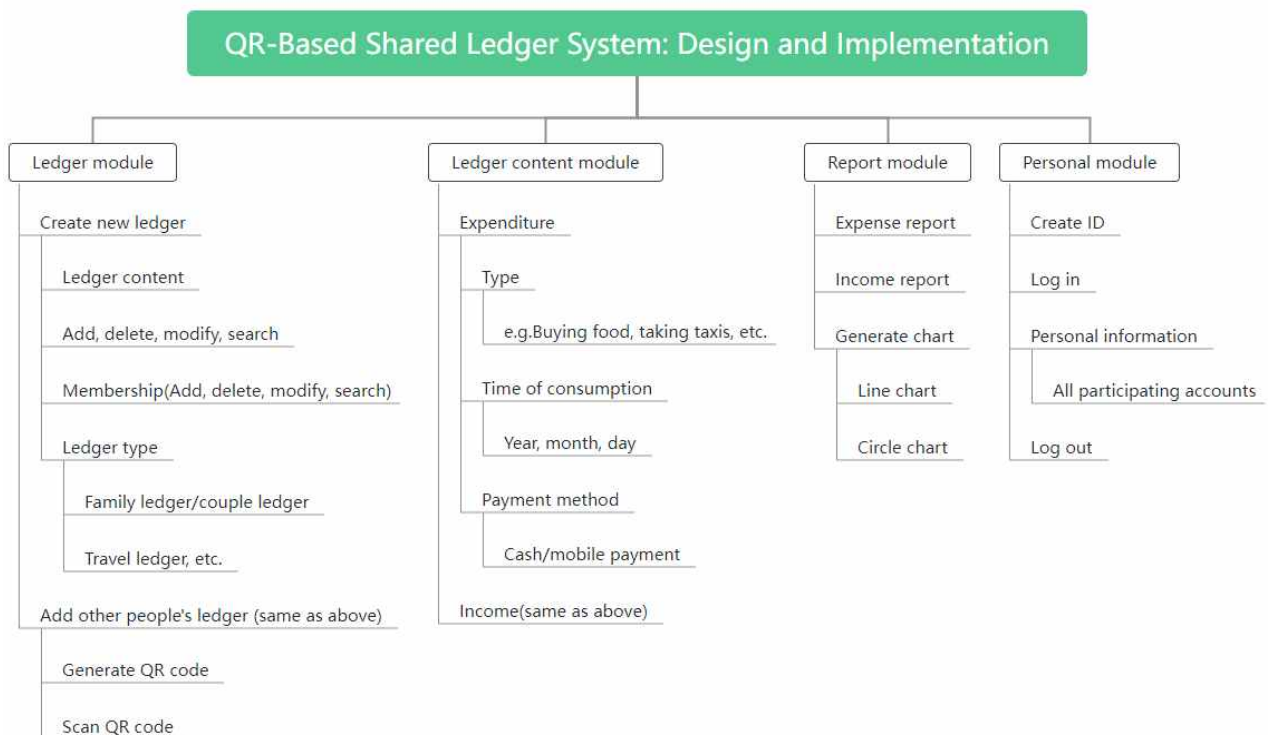


Fig. 1. System Function Diagram.

manages user permissions and access control for project data. Real-time Data Sync Mechanism: A technology that ensures real-time data synchronization between multiple users, using WebSocket.

- User C (Viewer): This is another invited user. Usually, they can only view information about account items, but cannot edit them.
- Mobile App (for User B and User C): This is the mobile app used by User B and User C. They can scan the QR code to access and participate in the account items. By scanning the QR code, User B and User C can easily participate in the shared account items and work with the author and other participants.

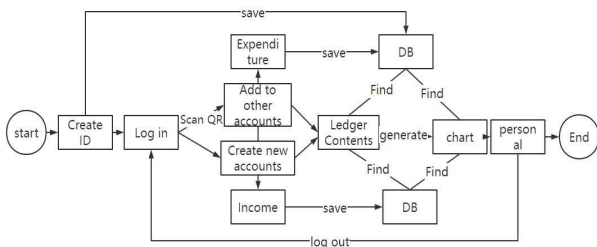


Fig. 2. System Flowchart.

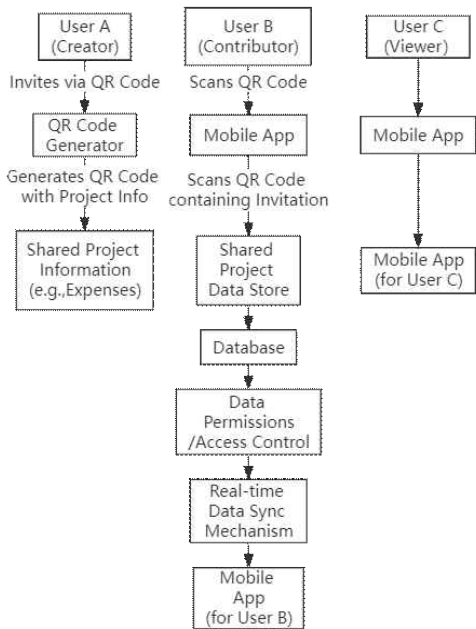


Fig. 3. Concept of Sharing Method.

## IV. Implementation

The system is developed using modern web technologies, integrating QR code generation for streamlined sharing. The key features include user registration, ledger creation, data synchronization, and financial analysis tools to aid in comprehensive financial management.

The implementation environment is as shown in Table 1. The operating system uses Windows. The database management tool uses Navicat Premium. In the integrated development environment, Eclipse is used for Java development, and HBuilder X is used for front-end development and mobile app development. LDPlayer is used to simulate an Android device and run Android applications.

Table. 1. Implementation Environment.

Feature	Specifications
OS	Windows 10 Pro
CPU	Intel(R) Core(TM) i7-6700 CPU @ 3.40GHz
RAM	16.0GB
DBMS	Navicat Premium 11.0.17
IDE	Eclipse 4.13.0
	HBuilder X 3.8.7
AVD	LDPlayer9

Fig. 4 (a) and (b) show the registration and login interfaces of this system, respectively.



Fig. 4. Implementation Interfaces: Registration and Login



Fig 5 includes the examples of the implementation for the system proposed in this paper

Fig. 5(a) shows the functions of the home page, and the descriptions for the functions is as follows. Click icon 1 to search for books related to the keywords you entered. Click icon 2 to create a new book. Click icon 3 to view all books. Click icon 4 to add specific income and expenses to the book. Click icon 5 to view the personal center.

Fig. 5(b) shows the interface when adding a new ledger, and Fig. 5(c) illustrates an interface for adding specific content to a ledger, including ledger name, ledger type, payment method, income and expense classification, date, amount, and notes.



Fig. 5. Implementation Interfaces.

## V. Conclusion

This paper demonstrated the feasibility of a QR-based shared ledger system that simplifies financial data sharing and collaboration. Future improvements will include enhanced security for QR codes and broader network compatibility, aiming to meet the financial collaboration needs in various scenarios.

## References

- [1] W. Liu, "Collaborative Sharing System of Accounting Information for Stakeholders Based on XBRL," 2012 Fifth International Conference on Business Intelligence and Financial Engineering, Lanzhou, China, pp. 597-600, 2012.
- [2] A. G. Anderson, "Conflicts of Interest: Efficiency, Fairness and Corporate Structure," 1997.
- [3] D. M. Salvioni, "Transparency culture and financial communication," 2002.
- [4] F. Liébana-Cabanillas, I. R. d. Luna, and F. J. Montoro-Ríos, "User behaviour in QR mobile payment system: the QR Payment Acceptance Model, Technology Analysis & Strategic Management," 2015.
- [5] L. Miller, "An Internet-Based Accounting Information Systems Project," American Journal Of Business Education, 2012.

# 딥페이크 탐지 모델을 위한 학습용 데이터의 전처리 기법에 관한 연구

최승호\*, 이석훈\*\*

## A Study on Preparation Techniques for Training Data for Deepfake Detection Models

Seungho Choi\* and Sukhoon Lee\*\*

### 요 약

딥페이크는 인공지능 기술의 발전으로 등장한 새로운 형태의 미디어로 다양한 분야에서 혁신적인 기능을 제공하지만 허위정보 유포와 프라이버시 침해 등의 문제를 야기하고 있다. 이러한 문제로 일반적인 영상 처리에서의 데이터 전처리는 모델의 성능을 향상시키기 위해 필수적이지만, 딥페이크 탐지에는 기존의 전처리 방법을 그대로 적용하기 어렵다. 본 논문에서는 딥페이크 탐지 모델을 위한 학습용 데이터의 전처리 기법들을 조사하고 다른 목적의 이미지 데이터 전처리 기법들과의 차이점을 분석한다.

### Abstract

Deepfake is a new form of media that has emerged with the development of artificial intelligence technology, and although it provides innovative functionality in various fields, it also causes problems such as spreading false information and infringement of privacy. Due to these problems, data preprocessing in general image processing is essential to improve the performance of the model, but it is difficult to apply the existing preprocessing method to deepfake detection. In this paper, we will explore and analyze various preprocessing techniques to enhance the accuracy and reliability of deepfake detection models and compare them with general image processing methods.

### Key words

Deepfake technology, Artificial Intelligence, Preprocessing Technique, False Information, Privacy Infringement

## 1. 서 론

딥페이크(Deepfake)는 인공지능 기술의 발전으로 인해 등장한 새로운 형태의 미디어로서, 현실과 매우 유사한 가짜 영상을 생성할 수 있게 되었다. 이

러한 딥페이크 기술은 다양한 분야에서 혁신적인 가능성을 열어주지만, 동시에 허위정보 유포, 프라이버시 침해 등 사회적 문제를 야기하고 있다[1][2]. 일반적인 영상 처리에서의 데이터 전처리는 조명, 색상, 해상도 등의 다양성을 통일시켜 모델이 데이

\* 국립군산대학교 소프트웨어융합공학과, csh005@kunsan.ac.kr

\*\* 국립군산대학교 소프트웨어학과, leha82@kunsan.ac.kr(교신저자)

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음(2023-0-00065).

터를 더 명확하게 구분하도록 돕는 것이 주된 목적이다.

그러나 딥페이크 이미지는 생성 과정에서 특정한 패턴이나 인공적인 흔적(artifact)을 포함하고 있어, 기존의 전처리 방법을 그대로 적용하면 이러한 중요한 특징을 손실할 수 있다는 문제가 있다[3]. 예를 들어, 딥페이크 이미지에서는 얼굴 주변에 미세한 왜곡이나 블러링이 발생할 수 있으며, 이러한 요소들은 딥페이크를 감지하는 데 중요한 단서가 된다. 그런데 전처리 과정에서 노이즈 제거나 이미지 보정을 무분별하게 적용하면 이러한 단서들이 제거되어 탐지 모델의 성능을 저하시킬 수 있다. 따라서 딥페이크 탐지를 위한 데이터 전처리 기법은 기존의 영상 처리와는 다른 접근이 필요하다.

본 논문에서는 딥페이크 탐지 모델을 위한 학습

용 데이터의 전처리 기법들을 조사하고, 다른 목적의 이미지 데이터 전처리 기법들과의 차이점을 분석하고자 한다.

## II. 딥페이크 데이터 전처리 기법

일반적으로 이미지 데이터의 전처리 과정은 이미지 품질 향상과 모델의 학습 효율 증대를 위해 사용되며, 주로 이미지 분류(classification), 객체 탐지(object detection), 이미지 향상(image enhancement) 등의 작업에 적용된다. 한편, 딥페이크 탐지에서의 데이터 전처리는 이미지 고유의 특징과 생성 과정에서 나타나는 인공적인 흔적을 보존하고 강조하는 것이 핵심이다. 표 1은 각 이미지 처리 기법에서 사용되는 이미지 데이터의 전처리 기법들을 보인다.

표 1. 딥페이크 탐지를 포함한 이미지 데이터 전처리 기법들

Table 1. Image data preprocessing techniques including deepfake detection

구분	기법	설명
이미지 분류	크기조정[1][2]	• 다양한 해상도의 이미지를 일정한 크기로 조정하여 모델 입력에 적합하게 변환
	정규화[1][2]	• 픽셀값을 일정한 범위로 스케일링하여 학습을 안정화
	데이터 증강[4][5]	• 회전, 이동, 반전, 확대, 노이즈 추가 등 다양한 변형을 적용하여 데이터 부족이나 과적합 방지
객체 탐지	노이즈 제거[1][2]	• 가우시안 블러, 미디언 필터 등을 사용하여 이미지의 노이즈 감소
	색상 조정[2]	• 조명 조건이나 색상 불균형을 보정하여 이미지 간의 일관성 유지
	평활화 및 샤프닝[2]	• 이미지의 경계를 부드럽게 하거나 강조하여 특정 특징 추출
이미지 향상	히스토그램 평활화[6]	• 명암비를 조정하여 이미지 디테일 개선
	색상 조정[2]	• 색상의 왜곡을 보정
딥페이크 탐지	인공 흔적 강조[7][8]	• 주파수 영역 분석 : 이산 푸리에 변환을 통해 이미지의 주파수 성분을 분석하고, 딥페이크와 실제 이미지 간의 차이를 부각 • 엣지 검출 : 소벨, 캐니 등의 엣지 검출 기법을 활용하여 이미지의 경계선을 강조하고 비정상적인 부분을 탐지
	프레임 간 일관성 분석[2]	• 시간적 이상 탐지 : 옵티컬 플로우 분석을 통해 프레임 간 움직임의 일관성을 검사하고 부자연스러운 변화를 감지
	안면 랜드마크 및 특징 분석[9]	• 특징점 추출 : 얼굴의 눈, 코, 입 등 주요 특징점을 추출하여 미묘한 위치 변화나 비율의 어색함을 탐지 • 텍스처 분석 : 피부의 질감이나 모공 등의 미세한 텍스처를 분석하여 합성 여부를 판단
	색상 및 조명 불일치 탐지[2]	• 색상 분포 분석 : 히스토그램 비교를 통해 얼굴과 배경 간의 색상 분포 차이를 분석 • 조명 방향 및 그림자 분석 : 조명 모델을 적용하여 얼굴과 배경의 조명 일관성을 검사
	잔차 이미지 활용[10]	• 잔차 계산 : 공간 리치 모델 필터 등을 통해 잔차 이미지를 생성 • 잔차 특성 학습 : 잔차 이미지에서 추출한 특징을 활용하여 모델이 딥페이크의 미세한 흔적을 학습

### III. 결 론

본 연구에서는 이미지 분류, 객체 탐지, 이미지 향상 등 목적별 일반적인 이미지 전처리 기법과 딥페이크 탐지를 위한 전처리 기법을 비교하였다. 일반적인 이미지 전처리 기법은 주로 이미지의 품질 향상과 모델의 학습 효율성을 높이기 위해 사용된다. 이러한 기법들은 해상도 조정, 노이즈 제거, 색상 보정, 데이터 증강 등을 포함하며, 모델이 다양한 환경에서 안정적으로 학습할 수 있도록 한다.

반면, 딥페이크 탐지를 위한 전처리 기법은 딥페이크 이미지의 고유한 특징을 보존하고 강조하는데 초점을 맞춘다. 주파수 영역 분석을 통한 인공 흔적 강화, 엣지 검출을 통한 비정상 경계선 탐지, 프레임 간 일관성 분석을 통한 시간적 이상 확인, 안면 랜드마크 및 텍스트 분석을 통한 미세한 변형 탐지 등이 그 예이다. 이러한 기법들은 딥페이크 이미지의 특이점을 부각시켜 모델이 합성 여부를 보다 정확하게 판단할 수 있도록 한다.

결과적으로 일반적인 이미지 처리 기법에서의 데이터 전처리 기법들은 이미지의 균일성과 일반화 능력을 향상시키는데 중점을 둔다. 반면, 딥페이크 탐지를 위한 데이터 전처리 기법은 인공적인 변형과 이상 패턴을 식별하고 보존하는데 주력하는 것을 알 수 있다.

본 연구를 통해 딥페이크 탐지 분야에서 데이터 전처리의 중요성과 차별성을 강조하였으며, 이는 향후 딥페이크로 인한 사회적 문제를 최소화하고 안전한 디지털 미디어 환경 조성에 기여할 것으로 기대된다[2].

### 참 고 문 헌

[1] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics++: Learning to detect manipulated facial images," Proc. IEEE International Conference on Computer Vision (ICCV), Seoul, Korea, pp. 1-11, Oct. 2019.

[2] L. Verdoliva, "Media forensics and deepfakes: an overview," IEEE Journal of Selected Topics in

Signal Processing, vol. 14, no. 5, pp. 910-932, Oct. 2020.

[3] R. Durall, M. Keuper, and J. Keuper, "Watch your up-convolution: CNN-based generative deep neural networks are failing to reproduce spectral distributions," Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, pp. 7890-7899, June 2020.

[4] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 60, pp. 1-48, Dec. 2019.

[5] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," arXiv preprint arXiv:1712.04621, 2017.

[6] J. S. Yun and J. H. Kim, "An improvement of histogram equalization using edge information of an image," Journal of Korea Multimedia Society, vol. 20, no. 2, pp. 188-195, Feb. 2017.

[7] X. Yang, Y. Li, and S. Lyu, "Exposing deep fakes using inconsistent head poses," Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, pp. 8261-8265, May 2019.

[8] J. Frank, T. Eisenhofer, L. Schönherr, A. Fischer, D. Kolossa, and T. Holz, "Leveraging frequency analysis for deep fake image recognition," Proc. IEEE International Conference on Machine Learning and Applications (ICMLA), Boca Raton, FL, USA, pp. 735-742, Dec. 2020.

[9] Y. Li, M.-C. Chang, and S. Lyu, "In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking," Proc. IEEE International Workshop on Information Forensics and Security (WIFS), Hong Kong, China, pp. 1-7, Dec. 2018.

[10] I. Amerini, R. Caldelli, L. Galteri, F. Picchioni, and A. Del Bimbo, "Deepfake video detection through image distortion analysis," Proc. IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, pp. 3561-3565, Sept. 2019.

# ANSYS 유한요소해석을 활용한 에스컬레이터 스텝롤러의 안전성 및 피로 수명 평가

이유선\*, 이상민\*\*

## Evaluation of Safety and Fatigue Life of Escalator Step Rollers Using ANSYS Finite Element Analysis

Yu-sun Lee\* and Sang-Min Lee\*\*

현재 대한민국 지하철역은 에스컬레이터 한 줄 서기 보행문화가 자리잡았다. 이는 에스컬레이터의 불균형한 하중을 유발하여 스텝롤러의 마모와 기계적 고장을 가속시키고 급정지 및 역주행의 원인으로 작용하여 부품의 조기 교체와 안전성 저하를 초래한다. 더불어 현행 승강기 유지관리 법령에서는 에스컬레이터 스텝롤러의 교체 주기에 대한 명확한 지침이 부족하고, 이에 따라 관련 업계 기업 또한 예방적 유지보수가 아닌 파손 시 교체가 일반적인 실정이다.

본 논문은 에스컬레이터 스텝롤러의 내구성과 안전성 강화를 목적으로, 에스컬레이터 사용 환경에서 발생하는 문제를 해결하고자 Ansys Workbench를 이용하여 스텝롤러의 동적·정적 시험과 피로 수명 분석을 진행하였으며, 반복적 하중과 충격이 스텝롤러의 수명에 미치는 영향을 평가하였다. 시뮬레이션을 통해 실제 지하철 사용 환경에서 발생하는 하중 조건을 반영하여, 에스컬레이터 보행 습관에 따른 부하를 재현하고 부품의 내구성을 객관적으로 측정할 수 있는 근거 자료를 제공하고자 하였다. 본 연구는 향후 스텝롤러의 예방적 유지보수 기준을 제안하고, 에스컬레이터의 안정성 향상과 장기적인 유지관리 비용 절감에 기여할 것으로 기대된다.

### Key words

Escalator Step Roller Fatigue Life, Safety Improvement, ANSYS Workbench, Elevator, Simulation

### 참고문헌

- [1] 행정안전부, “승강기 안전관리법 시행규칙” 제1장 제8조, 2024.
- [2] 행정안전부, “승강기 안전관리법 시행령” 제1장 제13조, 2024.
- [3] 국토교통부, “도시철도 정거장 및 환승·편의시설 설계 지침” 제3장, 20220
- [4] 국가기술표준원, “자율안전확인대상공산품의 안전기준” 제2조 부속서 62, 2018.
- [5] 국가기술표준원, “안전확인대상생활용품의 안전기준“ 제2조 제정·개정이유, 2019.
- [6] 한국철도공사, "KRCS H244 00 철도용품 공사규격서 에스컬레이터 스텝체인", 2021.
- [7] 이데일리, "특하면 멈추는 지하철 에스컬레이터... '한 줄 서기'가 원인", 2014.
- [8] YTN, “와르르 에스컬레이터 역주행, 유독 지하철만? 백화점엔 없는 '이것 때문'”, 2023

\* 국립한국교통대학교 글로벌융합대학원

\*\* 국립한국교통대학교 창의융합학부 부교수 (교신저자 : leesm@ut.ac.kr)

※ 2024년 국립한국교통대학교의 지원을 받아 수행하였음

※ 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0024169, 2024년 지역혁신클러스터육성)

# 도시 침수로 인한 승강기 피해 방지 IoT 플랫폼 구현

박건웅\*, 이상민\*\*

## Implementation of IoT platform to prevent elevator damage caused by urban flooding

Gun-Woong Park\*, Sang-Min Lee\*\*

최근 급격한 기후 변화로 도시 침수가 빈번해지며 승강기 안전 문제가 부각되고 있다. 승강기 구조상 안전장치 구획 및 유지보수 작업 공간으로 활용되는 피트(Pit)는 대부분 건축물의 최하층에 위치해 있어 침수 시 승강기 기능과 승객 안전에 직접적인 영향을 미치게 된다. 특히 2024년 7월 논산에서 발생한 엘리베이터 침수 사고는 침수로 인한 승강기 오작동으로 인명 피해가 발생한 사례다. 2024년 9월 기준, 국내 승강기 설치 대수는 약 860,060대이며 매년 증가하고 있지만, 승강기 침수에 대응하는 안전장치는 가격, 성능, 설치 요건 등의 문제로 상용화되지 못한 상황이다. 본 연구에서는 기존의 침수 방지 장치를 대체할 수 있는 경제적이고 효율적인 단일 보드 마이크로 컨트롤러 기반의 IoT 승강기 침수 방지 플랫폼을 제안한다. 이 플랫폼은 승강기 피트의 침수 감지, 데이터 분석, 원격제어를 통해 침수 사고에 선제적으로 대응하고 모니터링할 수 있도록 설계하였다.

### Key words

Urban Flooding, Flood Detection, Elevator safety, Data Analysis, Microcontroller, Internet of Things

### 참고문헌

- [1] Federal Emergency Management Agency, "Elevator installation for buildings located in special flood hazard areas in accordance with the National Flood Insurance Program", NFIP Technical Bulletin 4, U.S. Department of Homeland Security, June 2019.
- [2] 환경부(물재해대응과), "도시하천유역 침수피해 방지대책법", 법률 제19719호, 2023년 9월.
- [3] Tae-Gyu Koh and Won-Young Lee, "A Study on Urban Planning Techniques for Flood Reduction in the Lowlands: Focused on Seoul Metropolitan City in Korea", 서울도시연구 제13권 제4호, pp. 287-300, December 2012.
- [4] 행정안전부(승강기정책과), "승강기안전부품 안 전기준 및 승강기 안전기준", 행정안전부고시 제2022-18호, 별표 22 엘리베이터 안전기준 제 4조 제1호, 2022년 3월.
- [5] 행정안전부(승강기사고조사위원회), "승강기 사고 조사 보고서 충남 논산시 ○○○○○ 엘리베이터", 의결문 제2024-17호, 2024년 8월.
- [6] ㈜인에스코, "피트침수탐지시스템", <https://insko123.com/46>

\* 국립한국교통대학교 글로벌융합대학원

\*\* 국립한국교통대학교 창의융합학부 부교수 (교신저자 : leesm@ut.ac.kr)

※ 2024년 국립한국교통대학교 지원을 받아 수행하였음

※ 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임

(P0024169, 2024년 지역혁신클러스터육성)

# 스마트 IoT 안전모를 활용한 승강기 작업자 안전성 향상 및 위험요소 분석

백제현\*, 박건웅\*, 이상민\*\*

## Improvement of Elevator Workers' Safety and Analysis of Risk Factors Using Smart IoT Safety Helmet

Je-Hyeon Baek\*, Gun-Woong Park\*, Sang-Min Lee\*\*

최근 대한민국 승강기 신규 설치 및 유지 보수 산업의 증가와 함께 승강기 작업 중 발생하는 추락, 끼임, 감전 등과 같은 안전 문제가 부각되고 있다. 승강기 안전관리법에 따르면 노후 승강기의 경우 승강기 교체 권장 주기는 설치 이후로 15년이며, 노후화에 따른 엘리베이터 교체 및 유지보수의 수요는 해마다 증가할 것으로 예상된다. 최근 국가승강기정보센터에서 발표한 5년간 발생한 승강기 안전사고 중 승강기 작업자의 사망 및 부상에 대한 비율이 약 13.7%를 차지하였다. 안전사고 예방을 위한 대책으로는 승강기 검사 및 설치 작업 시 2인 1조 작업 방식의 대책안도 제시되었으나 사고재해는 매년마다 꾸준히 발생하고 있다. 이에 본 논문에서는 승강로 내 작업환경을 고려하여 센싱을 활용한 스마트 IoT 안전모를 제작하였으며, 물체감지센서, 조도센서, 온습도센서, 가스감지센서, 맥박센서 등 안전모에 부착된 센서 데이터는 실시간으로 모니터링되어 위험 상황 발생 시 즉각 경고를 제공하며, 사고 발생 후에는 데이터 분석을 통해 사고 원인과 위험요소를 파악하고 예방 대책을 수립하는 것을 목적으로 한다.

### Key words

Smart IoT Helmet, Elevator Safety, Worker Safety, Sensor-based Monitoring, Risk Factor Analysis

- [1] 행정안전부, “승강기 안전관리법 시행규칙” 2024.
- [2] 행정안전부, “승강기 안전관리법 시행령” 2024.
- [3] 산업안전보건연구원, “엘리베이터 설치와 유지보수 시 안전작업기준 및 전용발판 개발에 관한 연구” 2020.
- [4] 정완영, 구예진, 김종진. (2019). 착용형 센서와 헬멧을 이용한 작업자의 작업환경 모니터링. 융합신호처리학회 논문지, 20(2), 91-98.
- [5] Do hyung Kim, & Yeon-Chul Ha (2023). A Study on the Application of Smart Safety Helmets and Environmental Sensors in Ships. 융합신호처리학회 논문지, 24(2), 82-89.
- [6] Jinbae Kim, junho Jung, kyunggon Jo, & Jo jiho Jang (2015-06-24). Development for Smart helmet using the sensor and RMI technology. 한국정보과학회 학술발표논문집, 제주.

---

\* 국립한국교통대학교 글로벌융합대학원

\*\* 국립한국교통대학교 창의융합학부 부교수 (교신저자 : leesm@ut.ac.kr)

※ 2024년 국립한국교통대학교의 지원을 받아 수행하였음

※ 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0024169, 2024년 지역혁신클러스터육성)

# 지능형 스마트 안전관리를 위한 영상처리 기반 LTE모뎀 모니터링 제어 시스템 구현

이상민\*, 곽윤식\*\*

## Implementation of an LTE modem monitoring control system based on image processing for intelligent smart safety management

Sang-Min Lee\*, Yoonsik Kwak\*\*

지능형 정보기술사회가 되면서 사물인터넷(IoT), 인공지능(AI), 초연결 네트워크 고도화, 빅데이터(Bigdata), 딥러닝(Deep-Learning)등과 같은 기술이 우리의 일상생활 생태계 및 산업기술과 융합하여 차세대 기술 성장의 실증화 되고 스마트 안전관리에 대한 관심이 높아지고 있다. 이에 따라 AI 알고리즘 및 IoT 기반 실시간 예측 시스템 등을 통해 클라우드에서 분석하고, 머신 러닝 알고리즘을 통해 고장 가능성을 예측하는 등의 스마트 안전 관리 기술 개발들이 이루어 지고 있다. 본 논문에서는 사전에 학습된 객체 감지, 패턴 분석, 행동 인식 수행을 통해 인공지능(AI)과 머신러닝 기술이 결합한 딥러닝 모델을 통해 영상 데이터를 수집 처리하여 LTE모뎀을 활용한 수집데이터 처리 전송을 위한 제어를 통해 공장 자동화, 보안, 교통 관리, 의료, 스마트 시티 등 다양한 산업에서 활용가능한 영상처리 기반 모니터링 시스템을 제안하고자 한다.

### Key words

Object Detection, IoT, Big data, Deep learning, Smart Elevator, Safety, Analysis of the floating population

### 참고문헌

- [1] 이상민, "ICT융합과 미래를 여는 커넥티드 스마트승강기 동향," 주간기술동향, www.iitp.kr, 2021.11.3.
- [2] Kim, B. S., & Park, P., "Safety Management Improvement Plan for Elevator Worker Safety Accident Prevention," Korea Safety Management and Science, Vol. 22, No. 2, pp. 23-29, 2020.
- [3] Moon, H. C., "A Study on the Problem and Improvement of Elevator Safety System A Study on the Improvement of Legal Engineering between Elevator Engineering and Related Laws," Journal of The Korea Society of Computer and Information, Vol. 25, No. 10, pp. 221-230.
- [4] 박용문, "LTE통신 모뎀을 이용한 동보장치 이중화 제어시스템 개발", 한밭대학교 정보통신전문대학원 학위논문, 2023년 2월.
- [5] 이상민, "NB-IoT기반 스마트 승강기 원격 회생 제어 모니터링 시스템 구현", 한국승강기학회, 2023년 6월

---

\* 국립한국교통대학교 창의융합학부 부교수

\*\* 국립한국교통대학교 컴퓨터공학과 교수

※ 2024년 국립한국교통대학교 지원을 받아 수행하였음

※ 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0024169, 2024년 지역혁신클러스터육성)



# LLM을 활용한 문장 임베딩 대조학습 최적 샘플 선택 방법론

최성호\*, 이기주\*, 송민섭\*, 김남규\*\*

## Optimal Sample Selection in Contrastive Learning of Sentence Embeddings Using Large Language Models

Sungho Choi\*, Giju Lee\*, Minseob Song\*, and Namgyu Kim\*\*

### 요 약

대조학습(Contrastive Learning)은 문장 임베딩을 생성하는 매우 효과적인 방법으로, 대조학습에 사용되는 Positive 샘플과 Hard negative 샘플은 모델 성능에 큰 영향을 미친다. 최근에는 대형 언어 모델(LLM)을 활용하여 이러한 샘플을 생성하는 연구가 활발히 이루어지고 있지만, LLM이 생성한 샘플이 항상 최적의 샘플은 아닐 수 있다. 이에 본 연구에서는 LLM이 생성한 Positive 및 Hard Negative 샘플의 일관성을 평가하고, 의미적 및 구조적 유사도 기반의 평가 지표를 활용하여 적절한 샘플을 선택할 수 있는 새로운 방법론을 제안한다. 제안한 방법론을 통해 LLM이 생성한 Positive 및 Hard Negative 샘플의 적합성을 평가한 결과, Positive 샘플은 일관되게 생성된 반면 Hard Negative 샘플의 경우 샘플 간 편차가 커서 추가적인 검증이 필요함을 확인하였다. 이러한 검증을 통해 대조학습에 적합한 샘플을 사전에 선택함으로써, 대조학습을 통한 문장 임베딩의 성능을 향상시킬 수 있을 것으로 기대한다.

### Abstract

Contrastive learning is a highly effective method for generating sentence embeddings, and the selection of positive and hard negative samples used in this process significantly impacts model performance. Recently, research on generating such samples using large language models (LLMs) has become increasingly prevalent. However, the samples generated by LLMs may not always be optimal. In this study, we propose a novel method to evaluate the consistency of positive and hard negative samples generated by LLMs and to select appropriate samples based on semantic and syntactic similarity metrics. Experimental results show that while positive samples are generated consistently, hard negative samples exhibit significant variation. By utilizing the proposed method to select suitable samples from those generated by LLMs, it is expected that contrastive learning can improve the performance of sentence embeddings in tasks involving semantic textual similarity.

### Key words

sentence embedding, contrastive learning, large language model, positive sample, hard negative sample, evaluation metric

---

\* 국민대학교 비즈니스IT전문대학원 석사과정, Email: {unlver, kijoo0714, magnet9805}@kookmin.ac.kr.

\*\* 국민대학교 비즈니스IT전문대학원 교수(교신저자), Email: ngkim@kookmin.ac.kr

## 1. 서 론

문장 임베딩은 문장 간의 의미적 유사도를 정량적으로 평가하기 위해 필수적인 역할을 하며, 이를 효과적으로 학습하기 위한 방법으로 대조학습(Contrastive Learning)이 널리 사용되고 있다. 대조학습에서는 Positive 샘플과 Negative 샘플을 활용하여 유사한 문장 쌍은 임베딩 공간에서 가깝게, 다른 문장 쌍은 멀리 떨어지도록 학습한다. 이때 Positive 샘플은 학습하고자 하는 문장(Anchor Sentence)과 의미가 유사한 문장으로, Negative 샘플은 Anchor 문장과 다른 의미를 가지는 문장으로 정의된다[1]. 대조학습 과정에서 비교적 쉽게 학습되는 Negative 샘플에 비해, 모델이 구별하기 어려워 더 세밀한 학습을 유도할 수 있는 Hard Negative 샘플을 활용하는 방법에 대한 연구가 이루어지고 있다. Hard Negative 샘플은 Anchor 문장과 구조는 유사하지만, 의미는 다른 문장으로 정의된다[2].

대조학습을 통한 문장 임베딩 과정에서 적합한 Positive 샘플과 Hard negative 샘플을 선택하는 것은 모델 성능 향상에 중요한 영향을 미친다[3]. 최근에는 LLM(대형 언어 모델)을 통해 이러한 샘플을 생성하는 연구가 활발히 진행되고 있지만, LLM이 생성한 샘플의 적합성은 아직 충분히 연구되지 않았다[4]. 이에 따라 본 연구에서는 대조학습을 수행하기 전에 LLM이 생성한 샘플의 적합성을 평가하고, 이를 기반으로 최적의 샘플을 선택하는 새로운 검증 방법론을 제안한다.

## II. LLM을 이용한 샘플 생성 및 평가

본 장에서는 LLM을 활용하여 각 문장에 대해 Positive 및 Hard Negative 샘플을 생성한 후, 이 샘플들이 대조학습에 적합한지 검증하는 과정을 설명한다. 이를 위해 각 샘플의 의미적 유사도(Semantic Similarity)와 구조적 유사도(Syntactic Similarity)를 평가하고, 그 결과를 바탕으로 최적의 샘플을 선정하는 방법론을 제안한다. 이 과정은 대조학습을 수행하기 전에 생성된 샘플의 품질을 미리 평가하여 학습 성능을 향상시키는 것을 목표로 한다.

먼저 <그림 1>의 Phase 1에서는 NLI 데이터셋으로부터 Anchor 문장을 추출한 후, LLM을 활용하여 Positive 샘플과 Hard Negative 샘플을 각각 생성한다. Phase 2에서는 이전 단계에서 생성된 Anchor 문장과 Positive/Hard Negative 샘플 간의 의미적 유사도 및 구조적 유사도를 반영한 평가 수식을 계산하여 점수가 가장 높은 최적 샘플을 도출한다.

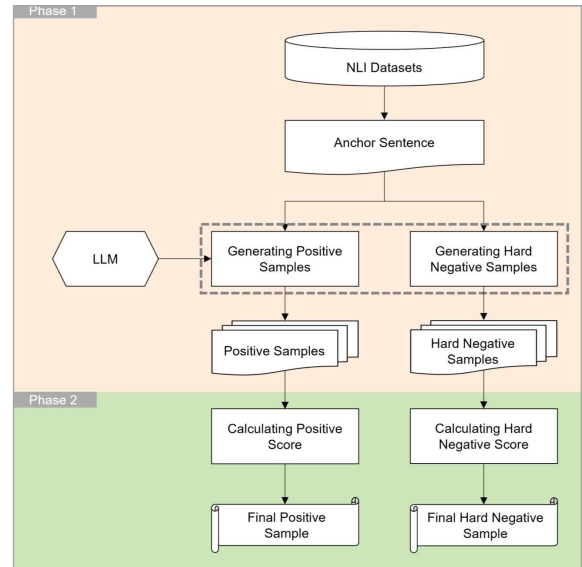


그림 1. 제안 방법론 개요  
Fig. 1. Overall Research Process

생성된 샘플의 의미적 유사도는 TF-IDF를 사용한 EMD(Earth Mover's Distance)를 통해 계산하며, 구조적 유사도는 POS 태그 및 의존 구문 분석을 통해 계산한다[5]. Positive 샘플과 Hard Negative 샘플을 평가하기 위해 의미적 유사도와 구조적 유사도의 가중치  $\alpha$ 와  $\beta$ 를 설정하며, 두 가중치의 합은 1로 고정된다. 의미적 유사도와 구조적 유사도는 각각 0에서 1 사이의 값을 가지며, Anchor 문장과 샘플 간의 유사성을 나타낸다. Positive Score는 수식 (1)과 같이 의미적 유사도와 구조적 유사도를 고려하여 0에서 1 사이의 값을 가진다. Positive 샘플은 의미적 유사도가 더 중요한 요소이므로,  $\alpha$ 의 가중치가 더 높게 설정된다.

$$positive\ score = \alpha * Semantic\ Similarity + \beta * Syntactic\ Similarity \quad (1)$$

Hard Negative Score는 수식 (2)와 같이 의미적 유사도가 음수로 적용되어 -1에서 1까지의 값을 갖는다. 이는 Hard Negative 샘플이 구조적으로 유사하지만 의미적으로 차별성이 있어야 함을 반영한다.

$$\text{hard negative score} = \alpha * (-\text{Semantic Similarity}) + \beta * \text{Syntactic Similarity} \quad (2)$$

이러한 평가 방식을 통해, LLM이 생성한 각 샘플의 적합성을 정량적으로 평가하여 최적의 Positive 및 Hard Negative 샘플을 선택할 수 있다.

또한, 제안 방법론의 필요성을 확인하기 위해 각 샘플의 평가 점수를 계산하고 비교하는 실험을 수행하였다. Positive 샘플은 점수 편차가 작은 것으로 보아 일관되게 생성되었음을 확인할 수 있었으며, 이러한 샘플은 대조 학습에 안정적으로 영향을 미칠 가능성이 크다. 반면, Hard Negative 샘플은 점수 편차가 큰 것으로 보아 일관성이 부족하다는 문제를 확인하였으며, 이는 다양한 Hard Negative 후보 중 적절한 샘플을 선택하는 과정이 필요함을 나타낸다. 제안 방법론을 활용해 적절한 샘플을 선택하여 문장 임베딩 대조학습을 진행한다면, 문장 임베딩의 성능을 향상시킬 수 있을 것으로 기대한다.

### III. 실험

본 연구에서는 제안 방법론의 필요성을 확인하기 위해 각 샘플의 의미적 유사도와 구조적 유사도를 기반으로 Positive Score와 Hard Negative Score를 계산하고, 그 결과를 분석하는 실험을 수행하였다. 구체적으로 LLM이 생성한 10개의 Positive 및 Hard Negative 샘플에 대해 각각의 점수를 계산하여 생성된 샘플의 일관성을 확인하였다.

실험 결과 Positive 샘플의 경우 최고 점수와 최저 점수 간의 차이가 0.068로 매우 적었으며, 대부분의 점수가 평균과 큰 차이를 보이지 않았다. 이는 Positive 샘플이 비교적 일관되게 생성되었음을 의미한다. 반면에 Hard Negative 샘플은 최고점과 최저점 간의 차이가 0.3953으로, Positive 샘플에 비해 편차가 훨씬 큰 것으로 나타났다. 이는 LLM이 생성한 Hard Negative 샘플들의 일관성이 부족하다는 점을 보여주며, Hard Negative 샘플에 대해서는 추가적인 샘플 선택 및 평가 과정이 필요하다는 결론을 도출하였다.

### IV. 결론

본 연구는 LLM을 활용해 생성된 Positive 샘플과 Hard Negative 샘플의 적합성을 평가하고, 이를 통해 최적의 샘플을 선택하는 방법론을 제안하였다.

### 참고 문헌

- [1] T. Gao, X. Yao, and D. Chen, "SimCSE: Simple Contrastive Learning of Sentence Embeddings," Proc. EMNLP, pp. 6894-6910, November 2021.
- [2] H. Wang, Y. Li, Z. Huang, Y. Dou, L. Kong, and J. Shao, "SNCSE: Contrastive Learning for Unsupervised Sentence Embedding with Soft Negative Samples," arXiv, March 2022.
- [3] J. Zhang, Z. Lan, and J. He, "Contrastive Learning of Sentence Embeddings from Scratch," Proc. EMNLP, pp. 3916-3932, December 2023.
- [4] N. M. An, S. Waheed, and J. Thorne, "Capturing the Relationship Between Sentence Triplets for LLM and Human-Generated Texts to Enhance Sentence Embeddings," Findings of the Association for Computational Linguistics: EACL 2024, pp. 624-638, March 2024.
- [5] S. Chauhan, R. Kumar, S. Saxena, A. Kaur, and P. Daniel, "SemSyn: Semantic-Syntactic Similarity Based Automatic Machine Translation Evaluation Metric," IETE Journal of Research, vol. 70, no. 4, pp. 3823-3834, April 2024.

# 명시적 규제를 활용한 오프라인 강화학습 기반 로봇 밀기 및 파지

신지섭\*, 유승환\*, 박주열\*, 박원우\*\*, 박형근\*\*, 홍승윤\*\*, 김병찬\*\*, 윤종완\*\*\*, 고병진\*\*\*\*, 박태준\*\*

## Offline Reinforcement Learning-Based Robotic Pushing and Grasping using Explicit Regularizer

Jisub Shin\*, Seunghwan Yu\*, Juyeol Park\*, Wonwoo Park\*\*, Hyoung-Geun Park\*\*, Seungyoun Hong\*\*,  
Byeongchan Kim\*\*, Jong-Wan Yoon\*\*\*, Byungjin Ko\*\*\*\*, and Taejoon Park\*\*

### 요약

물체가 밀집된 환경에서 로봇 기반 파지 작업은 해결해야 할 도전 과제이다. 이는 밀기와 같은 보조 행동을 통해 파지를 쉽게 수행하도록 도울 수 있으며, 이를 위해 로봇에게는 의사 결정 능력이 필요하다. 이러한 능력은 온라인 강화학습을 통해 학습할 수 있지만, 로봇이 지속적으로 환경과 상호작용하는 과정에서 비용이 크게 발생한다. 이에 대한 대안으로 오프라인 강화학습을 사용하여 상호작용 비용을 완화할 수 있지만, 확률적 경사하강법의 암묵적 규제가 오프라인 강화학습에 적용되면, 상태-행동 쌍들이 과도하게 유사해지는 문제가 있다. 본 논문은 명시적 규제를 적용한 오프라인 강화학습 기반의 밀기 및 파지 모델을 제안한다. 실험 결과, 제안한 모델이 밀집된 물체 환경에서 밀기를 통해 효율적인 파지 작업을 수행할 수 있음을 보인다.

### Abstract

Robot-based grasping tasks in dense object environments are a challenging issue that needs to be addressed. Auxiliary tasks, such as pushing, can facilitate grasping, and for this, the robot needs decision-making abilities. These abilities can be learned through online reinforcement learning, but high costs occur as the robot continuously interacts with the environment. As an alternative, offline reinforcement learning can be used to reduce interaction costs, but when stochastic gradient descent's implicit regularization is applied in offline reinforcement learning, a problem occurs where state-action pairs become overly similar. In this paper, we propose an offline reinforcement learning-based pushing and grasping model with explicit regularizer. Our experiments results show that the proposed model is suitable for pushing and grasping task in dense object environments.

### Key words

object grasping, robotic manipulation, offline reinforcement learning, simulation

\* 한양대학교 인공지능융합학과, mimak98@hanyang.ac.kr, sinanju06@hanyang.ac.kr, qkrwnduf1997@hanyang.ac.kr

\*\* 한양대학교 로봇공학과, neal0907@hanyang.ac.kr, bborang3505@hanyang.ac.kr, hong0807@hanyang.ac.kr,  
kchan124@hanyang.ac.kr, taejoon@hanyang.ac.kr(공동 교신저자)

\*\*\* 한양대학교 지능형로봇학과, jongwanyoon@hanyang.ac.kr(공동 교신저자),

\*\*\*\* 한양대학교 스마트융합공학부, byungjinko@hanyang.ac.kr(공동 교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업(IITP-2024-RS-2020-II201741, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 인공지능융합혁신인재양성사업(IITP-2024-RS-2022-00155885, 50%) 연구임

## I. 서론

픽 앤 플레이스는 물체나 부품을 파지하여 목표 위치로 옮기는 작업이다. 이는 물류 및 제조 산업의 자동화 공정에서 핵심적인 역할을 한다. 이 작업의 기본은 파지(Grasp)이며, 파지 성공률에 따라 자동화의 수준이 결정된다. 그러나 물체가 밀집된 환경에서는 그리퍼가 물체를 잡을 수 있는 공간이 제한되어 파지 성공률이 저하되는 문제가 발생한다. 이러한 문제를 해결하기 위해, 밀기와 같은 보조 행동을 통해 파지할 공간을 확보해야 한다[1]. 따라서 로봇이 픽 앤 플레이스를 효과적으로 수행하려면, 밀기 및 파지 전략에 대한 의사 결정 능력이 요구된다.

최근 온라인 강화학습(Online Reinforcement)은 물체 파지 문제에서 주목할 만한 성과를 보이며 활발히 연구되고 있다[2]. 온라인 강화학습은 복잡한 환경 속에서 시행착오를 통해 최적의 행동 전략을 학습한다. 그러나 에이전트가 환경과 지속적으로 상호작용하는 비용이 크게 발생하는 문제가 있다. 이를 해결하기 위한 대안으로 오프라인 강화학습(Offline Reinforcement Learning)이 주목받고 있다. 오프라인 강화학습은 이미 수집된 데이터를 기반으로 환경과 상호작용 없이 학습할 수 있다. 그러나 확률적 경사하강법의 암묵적 규제가 오프라인 강화학습에 적용되면, 상태-행동 쌍들이 과도하게 유사해지는 문제인 특징 적응(Feature co-Adaptation)이 발생한다. 이 문제는 상태-행동 쌍들을 구분하지 못해 에이전트의 의사 결정 과정에서 부정확한 행동 결정으로 이어질 수 있으며, 오프라인 강화학습에서 해결해야 할 중요한 과제이다.

본 논문에서는 명시적 규제를 적용하여 행동 결정 문제를 개선한 오프라인 강화학습 기반 밀기 및 파지 모델을 제안한다. 본 연구의 실험 결과는 명시적 규제를 추가한 밀기 및 파지 모델이 복잡한 환경에서의 파지 작업에 적합함을 보인다.

## II. 방법론

본 연구는 평행 그리퍼가 장착된 6축 매니퓰레이

터가 밀집한 환경의 물체를 밀기 및 파지하는 문제를 MDP(Markov Decision Process)로 정의한다. MDP는  $(S, A, T, r, \gamma)$ 의 튜플로 구성된다. 상태 공간  $S$ 는 밀집한 물체들을 Top Down View로 촬영한 RGB-D 이미지이다. 행동 공간  $A$ 는 밀기와 파지라는 두 가지의 행동 유형과 RGB-D 이미지의 픽셀 좌표를 기반으로 좌표계 변환을 통해 결정된 x-y-z 좌표이다. 이때  $22.5^\circ$ 의 배수로 yaw의 각도를 총 16가지의 방향으로 정의하며, 행동 유형에 따라 밀기는 그리퍼가 닫힌 상태로 16가지의 방향 중 한 방향으로 시작 지점인 x-y-z 좌표로부터 10cm만큼 선형으로 물체를 밀고, 파지는 시작 지점인 그리퍼의 중간 좌표인 x-y-z 좌표에서 16가지의 방향 중 한 방향으로 물체를 파지한다.  $T(s'|s, a)$ 는 현재 상태에서 특정 행동을 취했을 때 다음 상태로 전이될 확률을 나타내는 상태 전이 확률이다. 보상 함수  $R(s, a)$ 는 행동 유형에 따라 물체 파지를 성공했을 때 1의 보상, 물체 밀기를 성공했을 때 0.5의 보상, 그 외는 0의 보상을 제공한다. 밀기의 성공은 밀기 시도 전후의 Depth 이미지 간 픽셀 값 차이가 임계값 이상이면 성공이라 판단하며, 파지의 성공은 그리퍼가 물체를 들어 올린 후에도 물체와 그리퍼 간의 물리적 접촉이 유지되고 있으면 성공이라 판단한다.

Visual Pushing and Grasping(VPG)[5]은 상태  $s_t$ 인 RGB-D 이미지를 입력으로 받아, 상태와 동일한 크기의 해상도 이미지인 Q-Map을 출력한다. Q-Map은 각 픽셀에 대응하는 Q-값을 지니며, 이 값은 각 픽셀의 좌표계 변환을 통한 x-y-z 좌표에서 행동을 수행했을 때 받게 되는 기대되는 누적 보상을 의미한다.

VPG는 온라인 강화학습 기반으로 설계된 모델이기 때문에 오프라인 강화학습을 적용하기 위한 두 가지 종류의 변형 모델을 사용하였다. 첫 번째 모델은 오프라인 강화학습 알고리즘인 Conservative Q-Learning(CQL)[3]을 적용한 모델로, VPG+CQL이다. CQL은 오프라인 데이터에 없는 행동에 대한 Q-값을 감소시킨다. 이를 통해 데이터 수집 정책과 학습 정책 간의 분포 차이로 발생하는 OOD(Out-of-Distribution) 행동에 대한 Q-값 과대평가 문제를 완화할 수 있다. 두

번째 모델은 VPG+CQL+DR3로, VPG+CQL의 손실 함수에 DR3 명시적 규제[4]를 추가하여 상태-행동 쌍의 특징 벡터 간 내적을 최소화하는 모델이다. 이를 통해 상태-행동 쌍들이 과도하게 유사해지는 것을 방지하며, 특징 적응 문제를 완화할 수 있다.

본 연구에서는 온라인 VPG의 정책을 사용하여 오프라인 강화학습을 위한 데이터를 수집하였다. 그림 1과 같이 MuJoCO 시뮬레이터에서 수집된 데이터는 (상태, 행동, 보상, 다음 상태) 쌍으로 구성되었으며, 10개의 물체가 무작위 위치에 배치되었다. 이 과정을 반복하여 약 50,000개의 데이터를 수집하였다. 수집된 데이터는 VPG+CQL과 VPG+CQL+DR3 학습에 사용하였다.

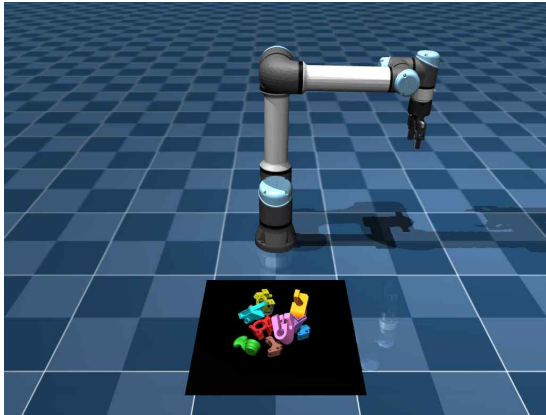


그림 1. 오프라인 데이터 수집 과정  
Fig. 1. Offline data collection process

### III. 실험 결과

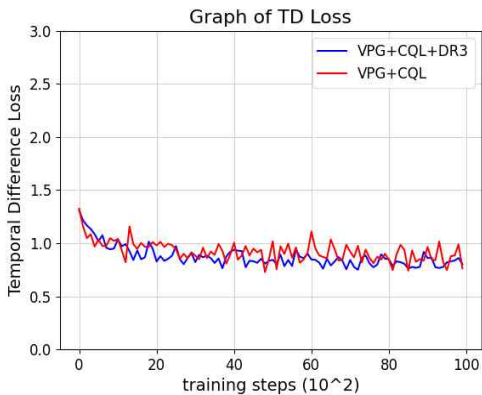


그림 2. VPG+CQL과 VPG+CQL+DR3의 TD loss 비교  
Fig. 2. Comparing TD loss of VPG+CQL and VPG+CQL+DR3

물체에 대한 행동 유형인 밀기 및 파지에 관한 VPG+CQL과 VPG+CQL+DR3의 학습 성능을 비교한 그림 2를 살펴보면, x축은 학습 횟수를, y축은 시간 차 오차인 Temporal Difference(TD) error를 나타낸다. 두 모델 모두 TD error 값이 0으로 수렴하지는 않지만, VPG+CQL+DR3이 VPG+CQL에 비해 다소 낮은 TD 값을 보인다. 이는 DR3의 명시적 규제가 오프라인 강화학습에 영향을 미친 것으로 해석할 수 있다.

표 1. 밀집한 배치에 대한 시뮬레이션 결과  
Table 1. Simulation results on dense arrangements

모델	작업	파지	행동
	완료율(%)	성공률(%)	효율성(%)
VPG+CQL	41.85	71.15	24.77
VPG+CQL+DR3	83.33	77.43	41.65

표 1에서는 학습이 완료된 VPG+CQL과 VPG+CQL+DR3을 비교한 결과로, 11개~12개의 물체가 밀집된 환경에서의 작업 완료율과 파지 성공률, 행동 효율성을 보여준다. 여기서 작업 완료율은 환경 내 물체 중 파지에 성공한 물체의 비율을, 파지 성공률은 전체 파지 시도 중 성공적으로 파지한 횟수의 비율을, 행동 효율성은 전체 밀기 및 파지 시도 중 성공적으로 파지한 횟수의 비율을 의미한다. 실험 결과, VPG+CQL+DR3이 VPG+CQL에 비해 작업 완료율이 41.48% 높았으며, 파지 성공률은 6.28%, 행동 효율성은 16.88% 더 높은 것을 볼 수 있다. 이는 밀집된 물체 환경에서 VPG+CQL+DR3이 VPG+CQL보다 적합함을 보여준다.

### IV. 결론

본 연구는 명시적 규제를 적용하여 행동 결정 문제를 개선한 오프라인 강화학습 기반 밀기 및 파지 모델을 제안한다. 실험 결과, VPG+CQL+DR3은 VPG+CQL에 비해 근소한 차이로 작은 오차를 보이며, 작업 완료율과 파지 성공률, 행동 효율성이 더 높았다. 이를 통해 본 연구가 밀집된 물체 환경에서의 파지 작업이 적합함을 입증한다.

참 고 문 헌

- [1] XU, Kechun, et al. Efficient learning of goal-oriented push-grasping synergy in clutter. IEEE Robotics and Automation Letters, 2021, 6.4: 6337-6344.
- [2] NEWBURY, Rhys, et al. Deep learning approaches to grasp synthesis: A review. IEEE Transactions on Robotics, 2023, 39.5: 3994-4015.
- [3] KUMAR, Aviral, et al. Conservative q-learning for offline reinforcement learning. Advances in Neural Information Processing Systems, 2020, 33: 1179-1191.
- [4] KUMAR, Aviral, et al. Dr3: Value-based deep reinforcement learning requires explicit regularization. arXiv preprint arXiv:2112.04716, 2021.
- [5] ZENG, Andy, et al. Learning synergies between pushing and grasping with self-supervised deep reinforcement learning. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018. p. 4238-4245.

# 불량 검출에 사용되는 임베딩 기반 이상 탐지 모델의 백본 네트워크와 이상 탐지 성능의 상관관계 분석

이승재<sup>1</sup>, 하헌명<sup>2</sup>, 오홍규<sup>3</sup>, 김명섭<sup>4</sup>, 김혜원<sup>5</sup>  
박태준<sup>6</sup>, 윤종완<sup>7</sup>, 고병진<sup>8</sup>

## Analyzing the Relationship Between Backbone Networks and Detection Performance in Embedding-Based Anomaly Detection for Defect Inspection

Lee SeungJae<sup>1</sup>, Ha Heonmyung<sup>2</sup>, Oh HongGyu<sup>3</sup>, Kim MyeongSeop<sup>4</sup>, Kim HyeWon<sup>5</sup>  
Park TaeJoon<sup>6</sup>, Yoon Jongwan<sup>7</sup>, Ko Byungjin<sup>8</sup>

### 요 약

본 논문은 산업 현장의 불량 검출을 자동화하기 위한 이상탐지 모델의 성능 향상 방안을 제안한다. 비지도 이상 탐지(UAD)에서는 ImageNet-1K Classification에서 우수한 성능을 보이는 ConvNext 같은 Backbone Network가 존재하지만, 대부분 ResNet을 사용한다. 본 연구에서는 이상 탐지 성능 향상을 위해 UAD 모델인 Reverse Distillation에 ConvNext를 대칭 및 비대칭 구조로 적용하였으나, Pixel-AUROC 성능이 0.012 ~ 0.08 하락하였다. 시각적 분석 결과, Backbone Network의 일반화 성능이 높을수록 모델의 압축 및 복원 격차가 줄어들어 성능이 저하됨을 발견하였다. 이를 통해 압축 및 복원에 적절한 균형을 가진 Backbone Network를 선택하면 이상 탐지 성능을 향상시킬 수 있음을 확인하였다.

### Abstract

This paper proposes a method to improve the performance of anomaly detection models used for automating defect detection in industrial environments. In Unsupervised Anomaly Detection (UAD), there are backbone networks like ConvNext, which show excellent performance in ImageNet-1K classification, but most models still rely on ResNet. In this study, ConvNext was applied in both symmetric and asymmetric structures to the UAD model, Reverse Distillation, aiming to enhance anomaly detection performance. However, Pixel-AUROC performance decreased by 0.012 to 0.08. Visual analysis revealed that as the generalization capability of the backbone network increases, the model's compression and reconstruction gap decreases, leading to a decline in performance. This finding suggests that choosing a backbone network with an appropriate balance between compression and reconstruction can improve anomaly detection performance.

### Key words

Anomaly Detection, Defect Inspection, Backbone Network, Deep Learning

<sup>1</sup>한양대학교, applejoker@hanyang.ac.kr, <sup>2</sup>한양대학교에리카, hhm107@hanyang.ac.kr, <sup>3</sup>한양대학교, omar0701@hanyang.ac.kr,

<sup>4</sup>한양대학교, kimas1215@hanyang.ac.kr, <sup>5</sup>한양대학교, khw9846@hanyang.ac.kr, <sup>6</sup>한양대학교에리카, taejoon@hanyang.ac.kr(교신),

<sup>7</sup>한양대학교에리카, jongwanyoon@hanyang.ac.kr(교신), <sup>8</sup>한양대학교에리카, byungjinko@hanyang.ac.kr(교신)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업(IITP-2024-RS-2020-II201741, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 인공지능융합혁신인재양성사업(IITP-2024-RS-2022-00155885, 50%) 연구임



## 1. 서 론

산업 분야에서 불량검출은 불량 제품의 출하율을 낮추기 위해 필수적인 분야이지만 자동화 연구가 미흡해 많은 산업에서 작업자의 육안으로 검사가 진행되고 있다. 이에 생산 효율성 증대를 위해 불량 검출에 딥러닝을 적용한 자동화 연구가 활발히 이루어지고 있다.

제조업 특성상 불량제품보다 정상제품이 월등히 많기에 데이터 불균형이 발생한다. 이러한 데이터 불균형 문제의 영향을 최소화하고 불량 영역을 효과적으로 인식하는 딥러닝 분야로는 Anomaly Detection (AD)이 있으며, 그 중 Unsupervised Anomaly Detection (UAD)는 정상데이터의 분포를 학습해 정상과 다른 표현 분포를 불량으로 인식함으로써 불량 데이터를 학습과정에 사용하지 않아 데이터 불균형 문제를 상당부분 극복하였다.

UAD분야는 현재 Feature-Embedding 기법의 모델 [1~3]이 상당부분 최고 성능 (State-of-the-Art, SOTA)를 차지하고 있으며, ImageNet[4]과 같은 대규모 데이터로 사전학습(Pretrained)된 Backbone Network가 주요 구성요소이다. 현재 SOTA인 대부분의 UAD 모델은 ImageNet Classification에서 더 높은 성능을 보여준 모델들이 있음에도 불구하고, 2015년에 발표된 Convolutional Neural Network(CNN)기반인 ResNet [5]을 Backbone Network로 사용하고 있다.

이에 본 논문에서는 모델의 최신화로 성능 향상을 이끌어내기 위해 UAD분야의 SOTA인 Reverse Distillation(RD)[1]의 Backbone Network인 ResNet을 대칭 및 비대칭 구조로 ConvNext[6]를 적용해 보았지만, Pixel-AUROC 성능이 0.012 ~ 0.08 하락하였기에 시각화를 통해 관련 원인을 분석한 결과, Backbone Network의 일반화 성능이 높은 경우 Feature-Embedding 기법의 압축 및 복원의 격차가 줄어들어 성능이 하락한다는 것을 발견하였다.

이를 통해 Feature-Embedding 기법이 적용된 AD 모델의 Backbone Network 선택시 정상 데이터의 압축-복원의 차이는 적으면서 불량 데이터의 압축-복원의 차이는 최대화 할 수 있는 Backbone Network를 선택하면 이상탐지 성능을 향상시킬 수 있는점

을 규명하였다.

## II. Reverse Distillation

RD모델은 UAD분야에서 SOTA인 모델로서 그림 1과 같이 Teacher, Student 그리고 Feature Concat Embedding Module로 구성되어 있다.

학습과정에서 오직 정상데이터만 입력으로 받아 사전학습된 ResNet을 Teacher로 사용해 3가지 해상도의 특징맵을 추출하고 이후 Feature Concat Embedding Module을 통해 3가지 해상도를 합치고 압축하여 ResNet으로 구성된 Student의 입력으로 사용한다. 이후 Student는 해당 입력을 사용해 3가지 해상도를 추출하게 된다.

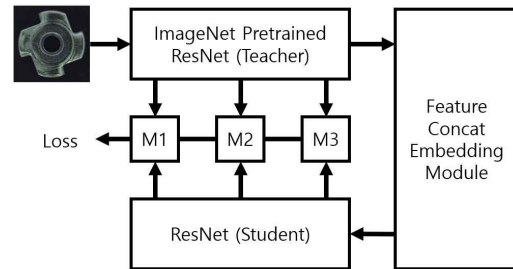


그림 1. RD모델 프레임워크  
Fig. 1. RD model Framework

이후 Teacher와 Student의 특징맵을 비교하여 Cosine Similarity(1)로 유사도를 측정하여 Loss를 산출한 후 역전파를 사용해 학습을 진행하며, 해당 방법으로 학습된 모델은 한 이미지에 대해 Teacher와 Student가 유사한 특징맵을 추출하게 된다.

$$\text{Cosine Similarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} \quad (1)$$

학습 이후 추론단계에서 정상데이터의 경우 두 특징맵의 차이가 없지만, 불량데이터의 경우 사전학습된 Teacher와 정상만 경험한 Student의 특징맵 차이가 발생하여 이를 점수화 및 시각화하여 불량을 검출할 수 있다.

그림 3. RDCN의 대칭 및 비대칭 구조  
Fig. 3. Symmetric and Asymmetric Structures of RDCN

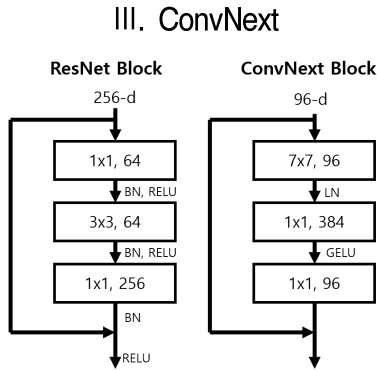


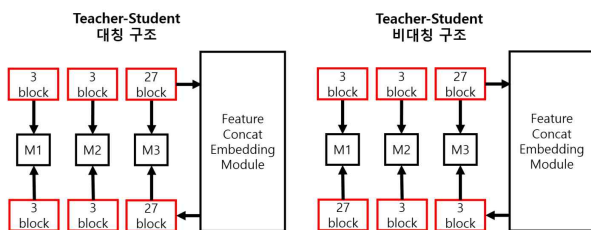
그림 2. ResNet과 ConvNext의 Block 구조 차이  
Fig. 2. Difference in Block Structure between ResNet and ConvNext

ConvNext는 Vision Transformer기반의 Backbone Network인 Swin Transformer[7]의 구조를 CNN 기반 모델에 접목시키고 다양한 최신화 기법을 이용했다.

해당 모델은 그림 2와 같이 7x7의 Convolution Filter를 채택하였으며, Optimizer로 기존 Adam의 발전형인 AdamW[8]를 채택하고 Activation Function으로 GELU[9]를 채택 및 Normalization으로 BatchNorm(BN)대신 LinearNorm(LN)을 채택하였다.

ConvNext는 ImageNet-1K Classification에서 Top1 ACC를 ResNet-50대비 3.2% 높은 82.0%로 달성하여 SOTA를 달성하였다. 이를 통해 ConvNext는 일반적인 딥러닝 태스크에서 ResNet대비 더 높은 일반화 및 정확도를 보여주는 것을 알 수 있다. 이를 통해 AD모델의 성능증대를 위해서는 Teacher의 특징맵을 정상 데이터에 대해서는 Student가 복원할 만큼의 성능이 있어야하며, 불량 데이터에 대해서는 복원하지 못할 만큼 성능이 낮아야 한다.

#### IV. RD모델 Backbone 최신화 실험



본 논문에서는 RD모델의 Backbone인 ResNet을 ConvNext로 교체한 Reverse Distillation using ConvNext (RDCN)모델을 새롭게 정의하여 성능향상을 시도하였다. ResNet과 ConvNext 모두 4 Stage로 이루어져 있으며, 순차적인 3개의 Stage로 Teacher를 구성하고 4번째 Stage로 Feature Concat Embedding Module을 일부 구성한다.

학습 및 추론은 RD모델의 과정을 그대로 준수하여 실험하였으며, 데이터셋은 MVTEC Dataset[10]의 Bottle, Metalnut, Hazelnut 클래스를 사용하였다.

실험 결과는 표 1에 나와있으며, ResNet에 비하여 Pixel-AUROC기준 대칭구조는 0.08 하락하였으며, 비대칭 구조는 0.012 하락하였다.

#### V. Backbone과 AD성능의 상관관계 분석

본 논문의 IV에서의 실험은 Backbone의 성능이 향상함에 따라 AD의 성능도 향상할 것으로 기대하였지만, 성능이 하락함에 따라 둘 간의 상관관계를 시각화를 통해 분석하였다.

표 1. Backbone에 따른 Pixel-AUROC 및 AUPRO  
Table 1. Pixel-AUROC and AUPRO by Backbone

Backbone	Class	Pixel AUROC	AUPRO
ResNet-50	Bottle	0.988	0.968
	Metalnut	0.974	0.925
	Hazelnut	0.990	0.958
ConvNext-Base (대칭 구조)	Bottle	0.98	0.957
	Metalnut	0.858	0.715
	Hazelnut	0.875	0.760
ConvNext-Base (비대칭 구조)	Bottle	0.973	0.929
	Metalnut	0.957	0.840
	Hazelnut	0.987	0.955

그림 4는 RD모델과 대칭구조의 RDCN의 Teacher와 Student간의 차이를 시각화한 것이다. 시각화는 [0,1]의 값으로 스케일링 되어 있으며, 붉을수록 복원 차이가 적은 것이고 푸를수록 복원 차이가 심한

것을 나타낸다.

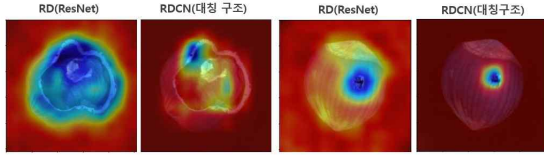


그림 4. RD와 RDCN의 시각화 차이  
Fig. 4. Difference in Visualization between RD and RDCN

그림 4의 왼쪽의 큰 불량을 시각화한 경우 ResNet 기반의 RD는 Student가 불량 영역 전반에 걸쳐 복원을 잘 못한 반면, RDCN은 극소부분을 제외하곤 상당부분 복원한 것을 알 수 있다. 오른쪽의 작은 불량의 경우에도 RD는 불량영역이 가장 덜 복원되었지만 정상영역의 복원도 일부 실패한 것을 알 수 있다. 하지만 RDCN의 경우 정상영역은 거의 완벽히 복원하여 차이가 없으며 작은 불량의 경우 복원에 실패하여 확실한 차이가 있어 기존 ResNet기반보다 높은 성능을 보임을 알 수 있다.

따라서 RDCN의 성능이 낮은 이유에 대해서는 Backbone의 높은 성능 때문에 복원하지 말아야 할 불량영역마저 복원을 성공하여 불량을 정상으로 오판독하는 것을 알 수 있다. 이를 통해 AD모델의 성능증대를 위해서는 Teacher의 특징맵을 정상 데이터에 대해서는 Student가 복원할 만큼의 성능이 있어야 하며, 불량 데이터에 대해서는 복원하지 못할 만큼 성능이 낮아야 한다.

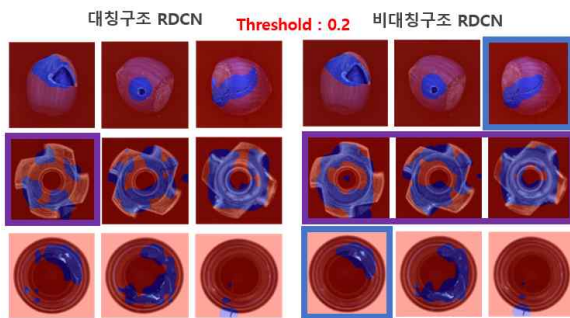


그림 5. RDCN의 대칭 및 비대칭 구조의 복원 차이  
Fig. 5. Restoration Differences Between Symmetric and Asymmetric Structures of RDCN

비대칭 구조의 RDCN은 이에 대한 근거를 뒷받침 할 수 있다. 그림 5는 대칭구조와 비대칭구조를

Threshold = 0.2로 이진시각화한 것으로, Metalnut 클래스에서 비대칭구조가 복원실패영역이 늘어난 것을 알 수 있다. 이는 Teacher의 특징맵을 비대칭 구조로 인해 특징맵 차이가 극대화 되어 Student가 복원에 실패한 것을 알 수 있다. 또한 비대칭 구조의 정보제한으로 인해 Student가 Overfitting되지 않아 Hazelnut과 Bottle의 일부는 정상영역을 성공적으로 복원한 것을 알 수 있다. 이는 표 1의 성능증가로 이어졌다.

## VI. 결 론

본 논문에서는 백본 네트워크와 이상 탐지(AD) 성능 간의 관계를 RD 모델을 통해 조사하였다. ResNet 백본을 ConvNext로 대체했을 때 성능이 저하되는 것을 발견하였으며, 이는 ConvNext의 뛰어난 일반화 능력이 정상과 불량 영역 구분을 방해했기 때문으로 분석되었다.

일반화 능력이 높은 ConvNext를 사용할 경우 복원 실패가 줄어들어 성능이 저하된다는 것을 확인하였으며, 이는 정상 데이터를 복원하면서도 불량 데이터는 복원하지 못하도록 하는 적절한 복원 능력을 갖춘 백본 선택이 중요함을 시사한다. 특히, RDCN의 비대칭 구조는 복원 실패 영역을 증가시켜 과적합을 방지하고 더 나은 이상 탐지 성능을 가능하게 했다.

## 참 고 문 헌

- [1] H. Deng and X. Li, "Anomaly Detection via Reverse Distillation From One-Class Embedding," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9737-9746, June 2022.
- [2] K. Roth, L. Pemula, J. Zepeda, B. Schölkopf, T. Brox, and P. Gehler, "Towards Total Recall in Industrial Anomaly Detection," arXiv preprint arXiv:2106.08265, 2021.
- [3] A. Mousakhan, T. Brox, and J. Tayyub, "Anomaly detection with conditioned denoising diffusion

- models", arXiv preprint arXiv:2305.15956, December 2023.
- [4] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 248-255, June 2009.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, June 2016.
- [6] Z. Liu, H. Mao, C. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11976-11986, June 2022.
- [7] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 10012-10022, October 2021.
- [8] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," International Conference on Learning Representations (ICLR), 2019.
- [9] D. Hendrycks and K. Gimpel, "Gaussian Error Linear Units (GELUs)," arXiv preprint arXiv:1606.08415, 2016.
- [10] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "MVTec AD: A comprehensive real-world dataset for unsupervised anomaly detection", Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9592-9600, June 2019.

# 위험 환경에서의 V-SLAM을 활용한 FireRep 로봇 에이전트 훈련

팔라시 잉글\*, 오세라\*, 김영갑\*\*

## FireRep Robotic Agents Training For Hazardous Environments Using V-SLAM

Palash Ingle\*, Se-Ra Oh\*, and Young-Gab Kim\*\*

### 요 약

정부와 지자체는 시민들의 안전한 일상을 위해 여러 방면에서 엄격한 조치를 시행하고 있으나, 여전히 화재나 유독가스 누출 같이 인적, 물적 피해를 주는 사고들이 발생하고 있다. 이 논문에서는 도시에서 발생하는 사고들에 효율적으로 대응하기 위한 V-SLAM 프레임워크를 제안한다. V-SLAM 프레임워크는 접근이 어려운 지역에서도 로봇 시스템이 독립적으로 내비게이션 지도를 만들어내며 다양한 센서를 활용해 최선의 결정을 내린다. 또한, 제안하는 모델은 인간의 개입 없이도 최적의 경로를 찾아 자율적인 탐사를 수행할 수 있다.

### Abstract

Government agencies and municipalities take strict measures to keep daily activities running smoothly. Despite the heavy surveillance, everyday hazardous events (e.g., fire, poison gas leakage, etc.) occur, causing human lives to be lost and massive damage to ecosystems. We propose an efficient V-SLAM framework, allowing the training of robotic systems to create V-SLAM maps for navigation in hard-to-reach areas independently; the robotic agents are comprised of multiple sensors to make appropriate decisions to respond to hazardous events. The model enables autonomous exploration, letting robots identify optimal paths and enhance their ability to operate without human intervention.

### Key words

detection, robotic agent, segmentation, slam, video surveillance

## I. Introduction

With modernization in intelligent cities, today's governments are heavily involved in the security and safety of their people. Artificial intelligence-based

---

\*세종대학교, palash@sju.ac.kr, sroh@intop.im

\*\*세종대학교, alwaysgabi@sejong.ac.kr(교신저자)

※ This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00419201).

video surveillance is adopted to identify suspected areas and alert citizens of any hazardous situation (e.g., fire and poison gas leakage) in advance. Predicting the hazardous situation in advance is highly difficult; upon identifying the hazardous situation, the time taken to mitigate it is crucial. The more time it takes to respond to a fire, the more live threats and heavy damage to the surroundings. The disaster and response management team tries to mitigate the fire situation; however, many terrains are difficult to reach out to in the practical world. Training a robotic agent for a hazardous environment is a difficult task, as in every hazardous environment, the terrain is unique and poses different challenges. In this study, we proposed a robotic agent equipped with state-of-the-art sensors, such as a Depth AI camera (detecting abnormal objects) and a LiDAR sensor (estimating distance), which intercepts the situation and takes corresponding action in a hazardous situation. The main feature of the proposed study is that the robotic agent accepts images as input to produce a V-SLAM environment, establishing an interactive virtual space for robotic agents. It enables seamless virtual interactions with real-world scenes, enhancing agent engagement in simulated settings.

## II. Related Work

There are several techniques for performing SLAM, such as the extended Kalman filter and Rotated-BRIEF SLAM [1]. A fundamental task in SLAM is estimating unknown variables based on known data, usually gathered from environmental sensors like RGB cameras, lidar, radar, and both monocular and stereo cameras. To process this data, feature extraction methods like Speeded Up Robust Features (SURF) identify key characteristics such as color intensity, edges, corners, and intersections. SLAM systems track the movement of these features within the input stream to estimate position changes and generate a 3D map, which is refined using processes like loop closure and bundle adjustment [2]. Visual SLAM supports constructing and refining agent interactions within virtual environments. For example, Faragher et al. developed a markerless indoor positioning system (IPS) with radio SLAM, using a point cloud generated from smartphone sensors and employing the DP-SLAM algorithm, which avoids the need for pre-existing floor plans.

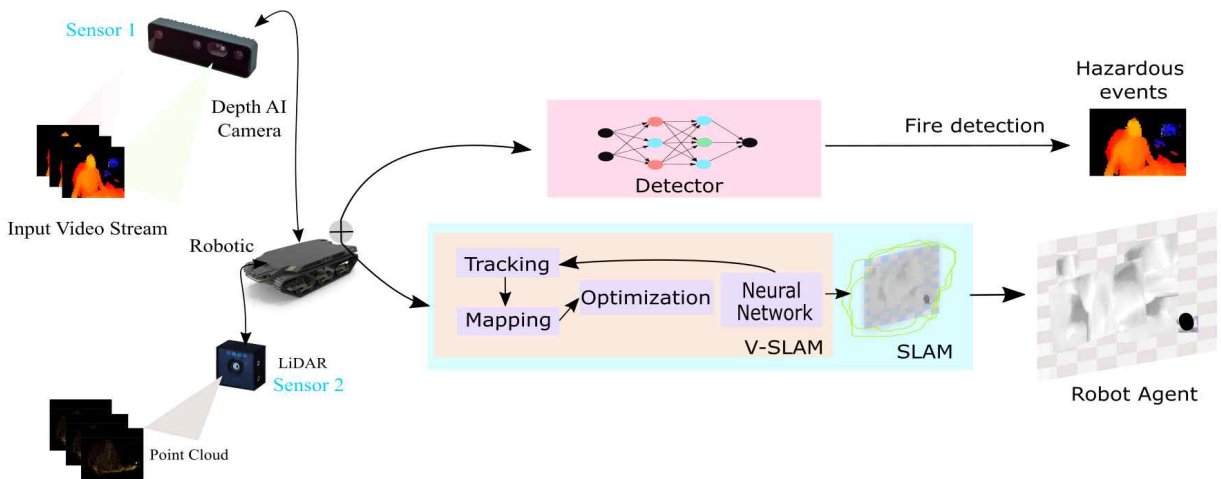


Fig. 1. Visualization of the FireRep Robotic Agent Framework

### III. FireRep Robotic Agent Framework

In the FireRep Robotic Agent Framework, image data is analyzed independently in 2D alongside LiDAR-generated point cloud data. To reduce computational load, the model uses only the point clouds within the bounding box area, employing wrapping and cropping methods to exclude boundary points. Fig. 1. provides a procedural overview of the FireRep Framework, showing the dynamic integration of sensory data. The video input includes a background frame and multiple foreground objects, with the foreground objects classified through a segmentation model. To do SLAM with a neural network, a CNN first extracts essential elements from the surroundings, such as edges and corners, using camera images. These characteristics are input into a pose estimation module, often using RNNs or pose regression to estimate the camera or robot's location and orientation. A mapping network processes this posture data and extracted features to create and continually update a map, often in the form of a grid or point cloud. A loop closure detection network is also used to identify previously visited sites, which helps to correct the map and prevent cumulative mistakes. A neural network tracks, maps, and optimizes the SLAM. The model is usually trained end-to-end on big datasets, allowing it to learn feature extraction, localization, and mapping.



Fig. 2. Visualization of sensory fused data

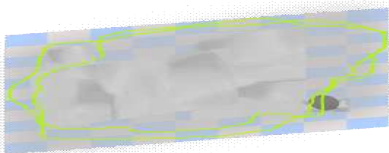


Fig. 3. Visualization of VSLAM with agent

### IV. Results

We tested the proposed FireRep robotic agent framework using the KITTI dataset [3]. The FireRep sensor fusion can be seen in Fig. 2. Based on the fusion between the camera and the point cloud data, Fig. 3 illustrates the VSLAM and the robotic agent interaction.

### V. Conclusion

The components of the FireRep Robotic Agents framework were evaluated by KITTI datasets. By harnessing the sensory data from the Robotic Agents, the model precisely extracts the fused segmented fire object and estimates the distance from the point of fire for projecting the due course of actions. Efficient 3D reconstruction is vital for autonomous agents conducting reconnaissance and assessment tasks. Training these agents in virtual environments enriched with real-world data enhances their navigation abilities, preparing them for effective operation in real-world settings.

### References

- [1] Raul Mur-Artal, Jose Maria Martinez Montiel, and Juan D Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE transactions on robotics*, 31(5):1147-1163, 2015.
- [2] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer Vision-ECCV 2006: 9th European Conference on Computer Vision*, Graz, Austria, May 7-13, 2006. *Proceedings, Part I* 9, pages 404-417. Springer, 2006.
- [3] Geiger, A., Lenz, P., Stiller, C. & Urtasun, R. Vision meets robotics: The kitti dataset. *The International Journal Of Robotics Research*. 32, 1231-1237 (2013).

# 협업 프로세스 구축을 위한 LLM 기반 전문가 추천시스템 개발

박상인\*,\*\*

## Development of an LLM-based expert recommendation system to build a collaboration process

Sangin Park\*,\*\*

현대 산업에서 전문가 추천시스템은 지식 공유와 협업을 촉진하는 중요한 도구가 될 수 있다. 본 논문에서는 대규모 언어 모델(LLM, Large Language Model)과 지식 증강 생성 기법(RAG, Retrieval Augmented Generation)을 활용하여 오픈 소스 LLM 기술의 약점을 보완한 사내용 전문가 추천시스템을 개발한 사례를 제시한다. 제안된 시스템은 LLM의 자연어 처리 능력과 RAG의 지식 검색 기능을 활용하여, 사용자의 질의에 대한 신뢰성과 정확성이 높은 사내 전문가를 추천한다. 구체적으로 시스템은 다음과 같은 절차를 따른다. 첫째, 사용자의 질의를 LLM을 통해 분석하고 관련된 정보를 추출한다. 이후, RAG 기법을 통해 대규모 지식 데이터 베이스에서 관련 문서를 검색하여 사용자의 질문과 일치하는 전문가 후보군을 형성한다. 마지막으로, 검색된 문서와 전문가의 프로젝트 수행 실적을 결합하여 최적의 인원을 추천한다. 또한, 고도화된 Retrieval Method가 LLM의 생성물에 미치는 영향을 파악하였고, 특정 방식들은 구체적이고 정확한 정보를 생성하는데 유리한 측면이 있음을 확인하였다. 서비스를 위한 실행 시간과 생성물의 품질을 고려한 최적의 Retrieval Method를 도출하였다. 본 시스템은 사내 지식 데이터 베이스를 적극적으로 활용하여 조직의 협업을 촉진하는 중요한 도구가 될 것으로 기대된다.

### Key words

Large Language Model, Retrieval-Augmented Generation, Recommendation System, Cooperation Process

### 참고문헌

- [1] M. Lewis, B. Oguz, and D. He, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459-9474, 2020.
- [2] Q. Chen, J. Huang, and Y. Chen, "Personalized Expert Recommendation Based on Deep Learning and Knowledge Graphs," *Expert Systems with Applications*, vol. 168, pp. 114201, 2021.
- [3] K. Guu, K. Lee, Z. Tung, et al., "REALM: Retrieval-Augmented Language Model Pre-Training," *Proceedings of the 37th International Conference on Machine Learning*, pp. 3929-3938, 2020.

---

\*고려대학교 SW·AI 융합대학원 석사과정, Email : qkrtkddls123@korea.ac.kr

\*\*현대자동차 AVP본부 연구원, Email : 6829656@hyundai.com(교신저자)



# Diffusion 기반 Inpainting을 활용한 컴퓨터 비전 기반 이상 탐지 모델의 성능 개선

김명섭\*, 오흥규\*, 이승재\*, 윤종완\*\*, 고병진\*\*\*, 박태준\*\*\*\*

## Improving the performance of computer vision-based anomaly detection models using Diffusion-based Inpainting

Myeongseop Kim\*, Honggyu Oh\*, Lee Seungjae\*, Jongwan Yoon\*\*, Byungjin Ko\*\*\*, and Taejoon Park\*\*\*\*

### 요 약

컴퓨터 비전 분야에서 이상 탐지 기술은 산업 결함 탐지, 의료 분야 등 다양한 분야에 적용되고 있다. 일반적으로 Reconstruction 기반 이상 탐지는 입력 이미지와 복원된 이미지를 비교함으로써 이상을 탐지하는 방법이 주로 사용된다. 하지만 이러한 기존 방법들은 입력된 정상 이미지를 다른 형태로 복원해 오판하는 문제가 있다. 이 문제를 완화하고 이상 탐지 성능 개선을 위해 이상 탐지 모델인 DDAD에 Diffusion 기반 Inpainting 기법을 추가한다. 본 연구는 Diffusion 기반 Inpainting을 통한 향상된 복원 품질을 이용해 Bottle, Hazelnut, Transistor와 같은 데이터셋에서 기존 방법론보다 개선된 성능을 확인할 수 있었다.

### Abstract

In the field of computer vision, anomaly detection technology is applied in various areas, such as industrial defect detection and the medical field. Generally, reconstruction-based anomaly detection is widely used, where anomalies are identified by comparing the input image with the reconstructed image. However, these conventional methods often misclassify normal images due to reconstructing them with a different form. To address this issue and improve anomaly detection performance, we incorporate a diffusion-based inpainting technique into the DDAD anomaly detection model. This study confirmed that, through enhanced reconstruction quality achieved by diffusion-based inpainting, the proposed approach outperformed conventional methods on datasets such as Bottle, Hazelnut, and Transistor.

### Key words

ComputerVision based Anomaly Detection, Deep learning, Unsupervised Learning, Diffusion, Inapinting,

---

\* 한양대학교 인공지능융합학과, kimas1215@hanyang.ac.kr, omar0701@hanyang.ac.kr applejoker@hanyang.ac.kr

\*\* 한양대학교 지능형로봇학과, jongwanyoon@hanyang.ac.kr(공동 교신저자),

\*\*\* 한양대학교 스마트융합공학부, byungjinko@hanyang.ac.kr(공동 교신저자),

\*\*\*\* 한양대학교 로봇공학과, taejoon@hanyang.ac.kr(공동 교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업(IITP-2024-RS-2020-II201741, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 인공지능융합혁신인재양성사업(IITP-2024-RS-2022-00155885, 50%) 연구임

## 1. 서 론

컴퓨터 비전 기반 이상 탐지(Anomaly Detection)란 한 이미지에서 이상을 식별하고 위치를 찾는 작업을 말한다. 보통 산업 제품의 결함 탐지, 의료 분야, 비디오 감시와 같은 여러 분야에서 다양하게 이용된다. 하지만 일반적으로 이상 정보에 대한 데이터셋이 부족하기 때문에 정상 이미지만 이용해 이상을 탐지하는 비지도 학습(Unsupervised Learning)이 많은 주목을 받았다.

그 중 Reconstruction 기반 이상 탐지 방법은 입력 이미지를 정상처럼 보이도록 복원한 후, 입력 이미지와 복원된 이미지를 비교함으로써 이상을 식별하고 위치를 파악할 수 있다. 이러한 방법에는 다양한 생성 모델이 사용되며, Auto Encoder [2], VAE(Variational Auto Encoder) [3], GAN(Generative Adversarial Networks) [4], Diffusion [5] 모델 등이 이상 탐지 모델에 도입되어 우수한 성능을 달성했다.

하지만 기존 Reconstruction 기반 이상 탐지 방법은 그림1과 같이 입력된 정상 이미지와 복원된 이미지의 형태가 서로 달라 정상 이미지를 이상으로 판단하는 문제가 있다. 이에 따라 이상 탐지 성능이 저하된다.



그림 1. 기존 Reconstruction 기반 방법의 문제점 예시  
Fig. 1. Examples of problems with existing Reconstruction-based methods

이 문제를 완화하기 위해 본 논문에서는 이미지의 일부를 가리고 복원하는 기법 중 하나인 Diffusion 기반의 Inpainting 기법을 채택했다[7]. 그리고 Diffusion 기반 Anomaly Detection 모델인 DDAD(Denoising Diffusion Anomaly Detection) [8]에 이 Inpainting 기법을 추가해 입력 이미지에 정상인

부분을 활용하고 이상 부분에 대해서는 무작위 마스크로 가려서 입력 이미지와 유사한 복원 이미지를 만든다.

실험은 MVTEC데이터셋 중 Bottle, Hazelnut, Transistor를 사용하였다[1]. 실험 결과, Bottle은 Image AUROC는 동일하게 100%를 달성하였으며, Pixel AUROC는 제안하는 모델이 1.0%의 향상되었다. Hazelnut의 Image AUROC는 0.1% 향상되었으며, Pixel AUROC는 0.4% 향상, Transistor는 각각 0.1% , 0.7%의 향상되었다.

결과적으로 본 연구에서는 DDAD 모델에 Inpainting 기법을 추가 하는 실험을 진행하였고, 실험 결과 Bottle, Hazelnut, Transistor와 같은 데이터셋으로 기존 방법론 보다 개선된 성능을 달성하였다.

## II. Inpainting 알고리즘을 활용한 Denoising Diffusion 모델

먼저 모델의 설계를 위해 DDAD모델을 메인 프레임워크로 사용하였다[8]. DDAD는 Reconstruction 기반의 이상 탐지 모델로, 세 단계로 나누어 학습을 진행한다. 첫 번째 단계에서는 각 클래스에 대한 정상 이미지들을 사용하여 Denoising U-Net을 학습한다. 두 번째 단계에서는, 첫 번째 단계에서 학습된 Denoising U-Net을 이용해 입력 이미지에 대한 복원을 진행한다.

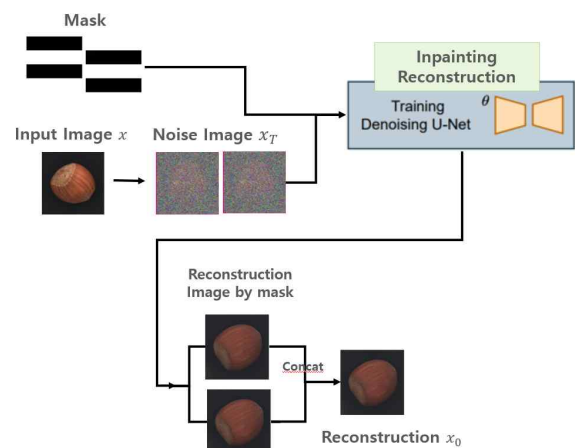


그림 2. 제안하는 Inpainting 기법을 활용한 Reconstruction 방식

Fig. 2. Reconstruction Method Using Proposed Inpainting Techniques

이때 입력 이미지뿐만 아니라 타깃 이미지를 함께 사용하여, 입력 이미지가 타깃 이미지처럼 복원되게 한다. 이러한 방식은 이상 탐지 시, 이상 부분을 정상적인 이미지로 복원하도록 유도한다. 세 번째 단계에서는, Feature-wise 비교를 위한 사전 학습된 Feature Extractor를 사용한다. 이때 Feature Extractor를 현재 Domain에 맞추기 위해 Domain Adaptation 기법을 적용한다. 이를 위해, 타깃 이미지와 두 번째 단계에서 만들어진 복원된 이미지를 사용하여 이용하여 두 이미지의 Feature 간 거리를 최소화하고, Distillation Loss를 추가해 기존 학습된 정보를 유지하면서도 새로운 Domain에 적응할 수 있도록 한다.

이상 탐지 시, 이상 이미지와 복원된 이미지를 통해 픽셀 수준과 특징 수준에서 이 둘을 비교한다. 그리고 그 차이를 통해 이상을 탐지하고 위치를 파악한다.

본 논문에서 제안하는 모델은 기존 DDAD 모델에 이상 탐지 과정에 Inpainting 기법을 활용한 Reconstruction 방식을 추가한 모델이다. 이 Inpainting 기법을 활용한 Reconstruction 방식의 구조는 그림2와 같다.

먼저, 제안하는 방식은 마스크의 개수만큼 입력 이미지에 대한 Noise Image를 생성한다. 그런 다음 Noise Image와 Mask를 결합하여 Inpainting 기법을 적용한다. 이렇게 복원된 이미지들의 평균을 구해 복원된 이미지를 생성한다.

이상 탐지 과정에서는 생성된 이미지를 기존 DDAD의 알고리즘에 타깃 이미지의 형태로 넣는다. 이를 통해 이상 이미지의 정상 정보를 일부 유지하면서, 이상이 없는 이미지로 복원되게 유도한다. 이렇게 생성된 최종 복원 이미지로 이상 탐지를 진행한다.

### III. 실험

실험은 MVTEC 데이터 셋 중 Bottle, Hazelnut과 Transistor를 사용하였다. 실험의 평가 지표로는 Image AUROC와 Pixel AUROC를 사용하였다. 실험 결과는 표 1과 같다.

표 1. 실험결과  
Table 1. Experimental Result

	DDAD		제안하는 모델	
	Image AUROC	Pixel AUROC	Image AUROC	Pixel AUROC
Bottle	100%	92.3%	100%	93.3% (+1.0%)
Hazelnut	99.7%	98.1%	99.8% (+0.1%)	98.5% (+0.4%)
Transistor	99.8%	93.8%	99.9% (+0.1%)	94.5% (+0.7%)

실험 결과 Bottle은 Image AUROC는 동일하게 100%를 달성하였으며, Pixel AUROC는 제안하는 모델이 1.0%의 향상되었다. Hazelnut의 Image AUROC는 0.1% 향상되었으며, Pixel AUROC는 0.4% 향상, Transistor는 각각 0.1% , 0.7%의 향상되었다.

실험 결과를 시각화한 그림은 그림 3과 같다. 시각화 결과, 이상 이미지에 대해서 정상처럼 보이게 복원된 것을 확인할 수 있었고 실제로 Heat Map에서도 이상 위치를 잘 파악된 것을 확인할 수 있다.

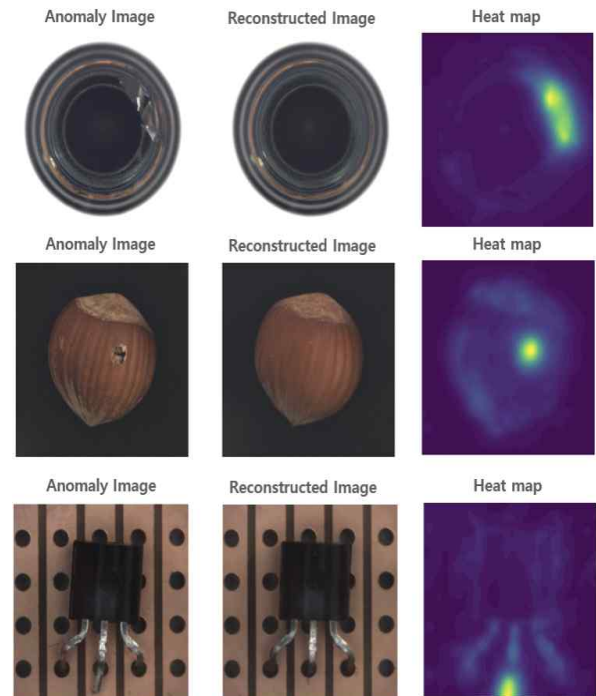


그림 3. 제안하는 모델의 실험 결과 시각화  
Fig. 3. Visualize the experimental results of the proposed model

#### IV. 결 론

본 논문에서는 기존의 Reconstruction 기반 이상 탐지 방법에서 정상 이미지와 복원된 이미지 형태가 다른 문제를 완화하기 위해 Diffusion 기반의 Inpainting 기법을 채택하였다. DDAD의 이상 탐지 단계에 Inpainting 기법을 추가 함으로써, 이상 이미지의 이상 부분은 무작위 마스크로 처리하고, 정상인 부분을 활용하였다. 제안하는 모델은 Bottle, Hazelnut, Transistor에서 기존 DDAD보다 개선된 성능을 달성했다.

#### 참 고 문 헌

- [1] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "MVTec AD: A comprehensive real-world dataset for unsupervised anomaly detection", Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9592-9600, June 2019.
- [2] Y. Bengio, L. Yao, G. Alain, and P. Vincent, "Generalized denoising auto-encoders as generative models", Proc. Advances in Neural Information Processing Systems (NIPS), pp. 899-907, 2013.
- [3] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes", Proc. International Conference on Learning Representations (ICLR), 2014.
- [4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", Proc. Advances in Neural Information Processing Systems (NIPS), pp. 2672-2680, 2014.
- [5] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models", Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 6840-6851, 2020.
- [6] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models", Proc. International Conference on Learning Representations (ICLR), 2020.
- [7] J. Lugmayr, M. Danelljan, A. Romero, F. Yu, R.

- Timofte, and L. Van Gool, "RePaint: Inpainting using denoising diffusion probabilistic models", Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11461-11471, 2022.
- [8] A. Mousakhan, T. Brox, and J. Tayyub, "Anomaly detection with conditioned denoising diffusion models", arXiv preprint arXiv:2305.15956, December 2023.

# 국방 SW 공급망 보안 강화를 위한 발전방향 연구

서영희\*

## Research on the development direction for strengthening the security of the defense SW supply chain

Younghee Seo\*

2020년 솔라윈즈와 2021년 Log4j 사고 등 2020년 이후 SW 공급망 공격 사례가 늘어나고 있으며 그 피해규모 역시 커지고 있어 체계적 대응의 필요성이 높아지고 있다. 2024년 9월 레바논에서 공급망 공격 사례인 무선 호출기·무전기 폭발사고로 인해 2,800여 명의 인명피해가 발생하기도 하였다. 이에 본 논문에서는 늘어나는 SW 공급망 공격에 대응하여 국방 영역의 SW 공급망 보안 강화 및 사이버 복원력을 갖추기 위해 SW 구성요소 정보를 관리하기 위한 방안을 모색하고자 한다. 미국은 SW 공급망 공격에 적극적으로 대응하기 위해 2021년 행정명령 EO-14028에서 연방 및 관련 기관에게 안전한 SW 개발체계를 갖추었다는 자체증명서(Self-attestation Form)의 제출과 SW 자재명세서인 SBOM(Software Bill of Materials) 제시 등 공급망 보안 관련 역할과 업무를 지시하였다. 이에 따라 2024년 10월 미 육군은 신규 도입되는 SW에 대한 계약에서 SBOM을 요구하는 지시를 하달하였다. 우리 정부도 과학기술정보통신부와 국정원을 중심으로 SW 공급망 보안 강화를 위한 지침을 발간하고 2025년 1월에 로드맵을 제시하고자 준비하고 있다. 따라서 우리 군도 도입되는 SW 구성요소를 관리를 통해 가시성을 제고하여 공급망 위협이 발생하였을 때 이를 빠르게 감지 및 대응하기 위한 대책 마련이 요구된다. 먼저 우리 군에 적합한 SBOM의 기본 요소를 지정하고 우선 적용이 필요한 대상을 지정하여 점진적인 확산이 필요하다. 새로운 기술이나 제도의 도입은 군의 조직적 저항이 상대적으로 낮은 분야부터 단계적으로 도입하는 방식이 유용할 수 있다. 도입을 확산하기 위해서는 실증 사업을 통해 SBOM을 통한 구성요소 정보 관리를 위한 효과성을 분석하여 이해관계자를 설득하는 과정이 필요하다. 이후 국방에 도입되는 SW에 대한 모든 구성 요소에 대한 관리를 체계적으로 수행해야 할 것이다. 중장기적으로는 중앙 리포지토리에 향후 도입되는 모든 국방 SW 내 디지털 자산 정보에 대한 파악을 위해 목록화 및 가시화하여 취득되는 자산에 대한 SW 구성요소 정보의 종합적으로 관리하기 위한 통합 관리체계 구축 및 국방 SW 공급망 보안 강화를 실행하기 위한 조직별 역할을 제안한다.

### Key words

Defense SW Supply Chain security, Military SW Supply Chain security, SBOM, Self-attestation

### 참고문헌

- [1] 과학기술정보통신부, "SW 공급망 보안 가이드라인", May 2024.
- [2] 윤정현, "국방 분야의 인공지능 활용성 제고 방안과 시사점", STEPI Future Horizon+, vol. 48, pp. 44-52, Dec. 2020.
- [3] 이성훈, 홍건식, "헤즈볼라 무선호출기·무전기 폭발 사태의 특징 및 시사점", 이슈브리프 606호, Sep. 2024.
- [4] Department of the Army, "Assistant Secretary of the Army SBOM Policy", Oct. 2024.
- [5] Executive Office of the President, "Improving the Nation's Cybersecurity", Executive Order 14028, May 2021.

\*한국국방연구원, yhseo@kida.re.kr

## 메신저 어플리케이션 보안성 분석

정다윗\*, 김준욱\*, 서효원\*, 정원준\*, 최석규\*, 최정원\*, 김재성\*\*, 이종호\*\*

### Analysis of Messenger Application Security

Da-Wit Jeong\*, Jun-Wook Kim\*, Hyo Won Seo\*, Won Jun Jeong\*, Seok Gyu Choi\*, Jeong Won Choi\*, Jae-Sung Kim\*\*, and Jong-Ho Lee\*\*

#### 요약

초기의 메신저는 텍스트 전송만 가능했으나 기술 발전으로 사진, 동영상 등 파일 전송이 가능해졌다. 스마트폰의 활성화 이후에는 쇼핑, 송금 등의 다양한 서비스까지 제공하는 플랫폼으로 성장했다. 또한, COVID-19 확산으로 비대면 소통의 중요성이 부각되면서, 원격 근무, 비즈니스 협업 등 다양한 분야에서 필수적인 존재가 되었다. 그러나 최근까지 지속적으로 메신저 어플리케이션에 대한 보안 취약점이 발견되고 있다. 따라서 본 논문에서는 국내외에서 많이 사용되고 있는 메신저 어플리케이션에 대한 보안 위협을 분석하고 안전성을 평가하였다.

#### Abstract

Early messaging applications were limited to text messaging, but technological advances have allowed them to transfer files such as photos and videos. Since the widespread adoption of smartphones, messaging platforms have expanded to provide various services, including shopping and remittances. In addition, messaging applications have become essential in fields such as remote work and business collaboration, as the spread of COVID-19 has highlighted the importance of contactless communication. However, security vulnerabilities of messaging applications continue to be identified. Therefore, in this paper, we analyze and evaluate the security threats of popular messaging applications used both domestically and internationally.

#### Key words

messenger application, vulnerability analysis, privacy protection, security risk assessment

#### 1. 서론

메신저는 1990년대 후반에 처음 등장하였다[1]. 초기의 메신저는 네트워크 속도, 기기의 성능, 등의

기술적 한계로 텍스트만 송수신할 수 있었다. 그러나, 기술의 발전과 사용자의 편의성 요구에 따라 사진, 동영상 등 파일 전송이 가능한 메신저가 등장하기 시작하였다. 스마트폰이 보급되기 시작한 2010년대부터는 모바일 메신저가 등장하였고 모바일 메신

\*KITRI 차세대 보안 리더 양성 프로그램(교육생), {djung0605, wqsaxz6996, shw03117, bob13th.pongchi, tjrrb0313, chjw1346}@gmail.com, \*\*KITIR 차세대 보안 리더 양성 프로그램(멘토), public4d@gmail.com, korea.hellsonic@gmail.com (교신저자)

※ 본 연구는 한국정보기술연구원이 주관하는 차세대 보안 리더 양성 프로그램 사업의 지원을 받아 수행된 연구임

제가 음성 통화, 문자 메시지 등을 대체하였다[2]. 또한, 카카오톡, 라인과 같은 메신저에서는 메신저 내에서 쇼핑, 송금, 택시 호출과 같은 다양한 서비스를 제공하면서 하나의 플랫폼으로 성장하였다[3]. 2019년 말에는 COVID-19의 확산으로 인해 사회적 거리 두기를 시행함에 따라 비대면 소통의 중요성이 부각되었다[4]. 메신저는 개인 간 소통 도구를 넘어서 원격 근무, 비즈니스 협업, 고객 서비스 등의 다양한 분야에서 필수적인 도구로 자리잡았다.

이처럼 메신저는 일상생활에 가장 밀접하게 접촉되어있는 만큼 프라이버시 보호가 중요하다. 그러나, 최근까지도 메신저에서 개인 정보 유출, 원격 코드 실행 등과 취약점이 지속적으로 발견되고 있다[5]. 최근 카카오톡에서 발생한 1-Click 취약점인 CVE-2023-51219[6]는 악의적인 링크를 클릭했을 때, 딥링크 유효성 검사 문제로 인해 WebView에서 임의의 JavaScript를 실행하여 액세스 토큰을 유출해 사용자의 계정을 탈취할 수 있는 취약점이다. 그 외에도 Telegram에서는 스크립트가 실행될 수 있는 파일 확장자에 대한 오타로 인해 경고창 없이 바로 스크립트가 실행되는 취약점이 존재했다. 이와 같은 보안 취약점은 검증 로직의 미흡함과 같은 프로그램의 결함에서 발생한다. 따라서 본 논문에서는 많은 수요를 가지고 있는 메신저 중 VoIP, 게이밍, 보안과 같은 특정 요소에 특화된 메신저들 대한 보안성을 분석하고 안전성을 검증하고자 한다.

## II. 메신저 어플리케이션 분석

본 논문에서는 국내·외에서 많은 수요를 가지고 있는 메신저 어플리케이션 중 Skype, Discord, Signal에서 발생할 수 있는 보안 위협에 대해 분석하고 안전성을 평가하였다.

### 2.1 Skype

Skype는 2003년 Skype Technologies에서 개발되었고 2011년 Microsoft에 인수되어 Microsoft가 서비스 중인 인스턴스 메신저이다[7]. Skype에서는 채팅, 음성/영상 통화, 파일 공유와 같은 기능만 제공한다. 따라서, 데이터 변조, 정보 유출과 보안 위협 벡터

에 대해 분석을 수행하였다. Skype 분석 환경은 Table 1. 와 같다.

표 1. 스카이프 분석 환경

Table 1. Experimental data sets

항목	내용
버전	8.130.0.205
플랫폼	Windows
기반 언어	Javascript

Skype는 Electron 기반 어플리케이션이기 때문에 어플리케이션에 필요한 리소스, 소스코드를 포함하고 있는 아카이브 파일인 ASAR이 존재한다. ASAR 파일은 Node.js 패키지인 asar 패키지를 통해 파일 추출이 가능하여 소스코드 확인이 가능하다.

#### 2.1.1 webPreferences

```
webPreferences: {
  preload: 1.join(__dirname, "preload.js"),
  nodeIntegration: !1,
  webviewTag: !0,
  webSecurity: !0,
  nodeIntegrationInSubFrames: !0,
  additionalArguments: ["--skype-process-type", "MainRootViewId"},
  sandbox: !1,
  spellcheck: !0,
  backgroundThrottling: !1,
  contextIsolation: !1
},
```

그림 1. Skype의 webPreferences 설정 값  
Fig. 1. Skype's webPreferences Settings

Fig. 1은 Skype의 main window에 대한 설정 값을 나타낸다. nodeIntegration이 false일 경우에는 Node.js API에 대한 접근이 차단되어 process 객체, require 함수 등을 사용할 수 없기 때문에 보안 위협이 발생하기 어렵다. contextIsolation는 Renderer process와 Electron 프레임워크의 Javascript context의 분리를 결정하는 옵션이다. 그러나 해당 옵션이 false로 설정되어 있어, 공격자가 XSS 등의 방식을 통해 Prototype Pollution을 수행하여 악성 행위를 유도할 수 있는 보안 위협이 존재한다[8].

#### 2.1.2 음성/영상 통화

Skype에서는 P2P 기술을 적용한 VoIP 방식을

사용하여 통화 기능을 제공한다. 이때 Skype는 STUN 프로토콜을 이용하여 P2P 연결을 수행하기 때문에 통화를 걸기만 해도 상대방의 IP 주소를 확인할 수 있는 보안 위협이 존재한다.

### 2.1.3 채팅

```

1  {
2    "clientmessageid": "123456789",
3    "composetime": "2023-10-27T10:00:00",
4    "content": "<a href='http://example.com'>http://example.com/a>",
5    "messagetype": "RichText",
6    "contenttype": "text",
7    "indisplayname": "John Doe",
8    "receiverdisplayname": "John Doe",
9    "properties": {
10     "urlpreviews": [{"key": "http://example.com", "value":
11     {"url": "http://example.com", "target_url": "http://example.com/",
12     "size": "1256", "status_code": "200", "content_type": "text/html",
13     "site": "example.com", "category": "generic", "title": "Example Domain",
14     "favicon": "https://urlp.asm.skype.com/v1/url/image/Favicon?
15     url=https%3A%2F%2Fth.bing.com%2Fth%2Fid%2F00F.
16     Wrg9YD5ieAHc3Y2D5Y5Rcgk3Fpid83dSkup%26w%3d128%26h%3d128%26dpr%3d1.5",
17     "favicon_meta": {"width": "80", "height": "80", "thumbnail": "",
18     "targetthumbnail": "", "user_pic": ""}}]}
19   }
20 }

```

그림 2. URL Link가 전송될 때 Json 데이터  
Fig. 2. Json data when URL Link is sent

Skype는 채팅을 통해 URL Link를 보냈을 때 Fig. 2.와 같은 형태로 API 서버로 전송한다. URL이 a 태그 형태로 content의 값으로 전송되고 있으며 properties의 urlpreviews 값을 통해 Favicon, Thumbnail 등을 API 서버를 통해 불러오는 것을 확인할 수 있다. API 내부에서 필터링이 수행될 것으로 추측되지만, 퍼징을 통한 방식으로 내부 검증 로직을 우회하게 된다면 XSS(Cross-Site, SSRF 등과 같은 취약점으로 연계될 수 있는 잠재적 보안 위협이 존재한다.

## 2.2 Discord

Discord는 2015년 Discord Inc.에서 개발된 인스턴트 메신저이다[9]. Discord에서는 채팅, 음성/영상 채팅, 화면 공유 등을 지원하는데 2024년 9월에 DAVE 프로토콜을 도입하여 중단간 암호화를 지원함으로써 중간자 공격으로부터 보안 위협이 발생하기 어렵다. 따라서, 내부 상호작용을 통해 발생할 수 있는 보안 위협에 대해 분석을 수행하였다. Discord 분석 환경은 Table 2. 와 같다.

Discord도 Skype와 같이 Electron 기반 어플리케이션으로 ASAR 파일을 언패킹하여 소스코드 확인이 가능하다.

표 2. 디스코드 분석 환경

Table 2. Discord Analysis Environment

항목	내용
버전	1.0.9168
플랫폼	Windows
기반 언어	Javascript

### 2.2.1 webPreferences

```

webPreferences: {
  blinkFeatures: 'EnumerateDevices,AudioOutputDevices',
  nodeIntegration: false,
  sandbox: false,
  preload: _path.default.join(__dirname, 'mainScreenPreload.js'),
  nativeWindowOpen: true,
  enableRemoteModule: false,
  spellcheck: true,
  contextIsolation: true,
  additionalArguments: ['--enable-node-leakage-in-renderers'],
  devTools: ENABLE_DEVTOOLS
}

```

그림 3. 디스코드의 webPreferences 설정 값  
Fig. 3. Discord's webPreferences Settings

Fig. 3은 Discord는 main Window에 대한 설정 값을 나타낸다. Discord에서는 nodeIntegration이 false로 설정되어 있어 Node.js API에 대한 접근이 차단되어 있다. 또한, contextIsolation이 true로 설정되어 있어 Prototype Pollution과 같은 보안 위협으로 안전하다.

### 2.2.2 IPC 모듈

Electron 기반 어플리케이션은 main 프로세스와 renderer 프로세스가 분리되어 있기 때문에 두 프로세스 간에 접근이 불가능하다. 그러나, 프로그램의 설계에 따라 각 프로세스 간에는 접근이 필요한 경우가 존재한다. 이때 사전에 정의된 IPC 이벤트를 통해 통신을 수행한다. IPC 이벤트에서는 과도한 접근 권한을 부여하거나 검증 로직의 미흡함으로 인해 보안 위협이 발생할 수 있다.

```

_ipcMain.default.on('OPEN_EXTERNAL_URL', (_event, externalUrl) => {
  (0, _securityUtils.saferShellOpenExternal)(externalUrl).catch(() => {
    console.error('Failed to open external URL', externalUrl);
  });
});

```

그림 4. OPEN\_EXTERNAL\_URL IPC 이벤트  
Fig. 4. OPEN\_EXTERNAL\_URL IPC Event

Fig. 4는 외부 URL에 접근할 수 있는 IPC 이벤



트를 나타낸다. IPC 모듈을 통해 OPEN\_EXTERNAL\_URL 이벤트를 externalUrl 인자와 함께 호출하면 saferShellOpenExternal 함수를 실행시킨다.

```
function saferShellOpenExternal(externalUrl) {
  if (shouldOpenExternalUrl(externalUrl)) {
    return _electron.shell.openExternal(externalUrl);
  } else {
    return Promise.reject(new Error("External url open request blocked"));
  }
}
```

그림 5. saferShellOpenExternal 함수 정의  
Fig. 5. Definition of saferShellOpenExternal Function

saferShellOpenExternal 함수 내에서는 전달받은 인자인 externalUrl을 인자로 하여 shouldOpenExternalUrl (Fig. 6) 함수를 실행시킨다. 인자로 전달된 URL의 Scheme에 대해 BLOCKED\_URL\_PROTOCOLS (Fig. 7)를 통해 블랙리스트 기반 필터링을 수행한다.

```
function shouldOpenExternalUrl(externalUrl) {
  let parsedUrl;
  try {
    parsedUrl = _url.default.parse(externalUrl);
  } catch (_) {
    return false;
  }
  if (
    parsedUrl.protocol == null ||
    BLOCKED_URL_PROTOCOLS.includes(parsedUrl.protocol.toLowerCase())
  ) {
    return false;
  }
  return true;
}
```

그림 6. shouldOpenExternalUrl 함수 정의  
Fig. 6. Definition of shouldOpenExternalUrl Function

필터링되는 Scheme은 로컬 파일에 접근할 수 있거나 스크립트를 실행시킬 수 있는 것들로 구성되어 있다. 이때 Windows에서는 절대 경로를 나타낼 때 "C:"와 같은 형태로 시작한다. 따라서, 공격자가 XSS를 통해 로컬 파일의 절대 경로에 대해 IPC 통신을 수행하면 로컬 파일에 접근할 수 있게 된다.

### 2.3 Signal

Signal은 2014년 Signal Messenger LLC에서 개발된 Electron 기반의 오픈소스 인스턴트 메신저이다 [10]. Signal에서는 Signal Protocol을 이용한 종단간 암호화를 지원하기 때문에 중간자 공격으로부터 안전하다. 또한, Signal은 짧은 주기로 업데이트를 수

행하며 인스턴스 메신저로써 채팅, 파일 공유, 음성 전화와 같은 최소한의 기능만 제공되기 때문에 비교적 보안 위협으로부터 안전하다.

표 3. 시그널 분석 환경

Table 3. Signal Analysis Environment

항목	내용
버전	v7.30.0
플랫폼	Windows
기반 언어	Javascript

#### 2.3.1 webPreferences

```
webPreferences: {
  ...defaultWebPrefs,
  nodeIntegration: false,
  nodeIntegrationInWorker: false,
  sandbox: false,
  contextIsolation: !isTestEnvironment(getEnvironment()),
  preload: join(
    dirname,
    usePreloadBundle
    ? '../preload.wrapper.js'
    : '../ts/windows/main/preload.js'
  ),
  spellcheck: await getSpellCheckSetting(),
  backgroundThrottling: true,
  disableBlinkFeatures: 'Accelerated2dCanvas, AcceleratedSmallC canvases',
},
```

그림 7. Signal의 webPreferences 설정 값  
Fig. 7. Signal's webPreferences Settings

Signal에서는 모든 window에 대하여 nodeIntegration이 false로 설정되어 있으며, contextIsolation이 테스트 환경을 제외하곤 true로 설정되기 때문에 Node.js API에 접근, Prototype Pollution과 같은 잠재적 보안 위협으로부터 안전하다.

#### 2.3.2 파일 업로드

```
const DANGEROUS_FILE_TYPES =
  /\. (ADE|ADP|APK|BAT|CAB|CHM|CMD|COM|CPL|DIAGCB8|DLL|DMG|EXE|HTA|INF|INS|ISP|
```

그림 8. 파일 전송 시 필터링되는 확장자  
Fig. 8. Filtered File Extensions During File Transfer

Fig. 8은 파일 전송하기 위해 파일 선택 시 필터링되는 확장자들이다. 주로 실행 가능한 실행 파일 이거나 스크립트 파일의 확장자로 총 47종류의 확장자에 대해 블랙리스트 기반 필터링을 수행한다. 그러나, 일부 실행 가능한 스크립트 확장자에 대해서는 필터링이 수행되고 있지 않다.

보안상 이유로 첨부 파일 유형이 허용되지  
않음

그림 9. 파일 업로드 필터링 되었을 때 메시지  
Fig. 9. Message when file upload is filtered

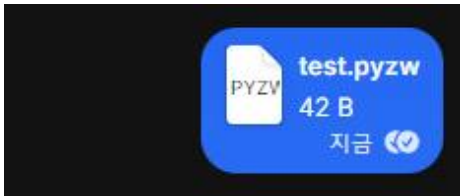


그림 10. 정상적으로 전송된 pyzw 파일  
Fig. 10. Successfully transferred pyzw file

블랙리스트에 존재하는 확장자의 파일을 업로드 하였을 때 Fig. 9와 같은 메시지가 뜬다. 그러나, Python 스크립트를 실행시킬 수 있는 zip 아카이브인 pyzw를 업로드하였을 때, 파일이 전송되는 것을 확인할 수 있다(Fig. 10). 이는 다른 취약점과 연계되어 0-Click 취약점으로 발전했을 때, 원격 코드 실행과 같은 치명적인 취약점으로 발전될 수 가능성이 높다.

### III. 결론

본 논문에서는 국내·외에서 많이 사용되는 메신저 중 Skype, Discord, Signal에 대한 보안 위협을 분석하고 안전성을 평가하였다. 대부분 기능에는 검증 로직이 잘 적용되어 있으나, 일부 기능에서 검증 로직의 미흡함으로 인해 다른 보안 위협과 연계되었을 때 치명적인 취약점으로 발전될 가능성이 있다. 또한, API 통신 부분에서 발생할 수 있는 보안 위협에 대해서는 벤더사의 정책상 이유로 분석을 수행할 수 없기 때문에 잠재적인 보안 위협이 더 존재할 가능성이 있다. 따라서, 이러한 연구 결과를 토대로 메신저 어플리케이션에 대한 보안에 대한 중요성을 알리는 계기가 될 것이며, 메신저 어플리케이션 보안 연구 발전에 기여할 것으로 기대한다.

### 참고 문헌

[1] K. Leurs, Instant Messengers, IRC, ICQ. In: K.

Harvey, (ed.) Encyclopedia of Social Media and Politics, SAGE Publications, London, UK, 2014. ISBN 9781452244716.

- [2] S. Park, K. Cho, and B.G. Lee, What makes smartphone users satisfied with the mobile instant messenger? Social presence, flow, and self-disclosure, International Journal of Multimedia and Ubiquitous Engineering, vol. 9, no. 11, pp. 315-324, 2014. <https://doi.org/10.14257/ijmue.2014.9.11.31>.
- [3] 허수정, 국내외 슈퍼앱의 현황과 활용 전망, KCISA, vol. 23, no. 6, 2023, [https://webzine.kcisa.kr/vol023/\\_file/issue\\_report.pdf](https://webzine.kcisa.kr/vol023/_file/issue_report.pdf).
- [4] K. Paerata, The use of workplace instant messaging since COVID-19, Telematics and Informatics Reports, vol. 10, p. 100063, 2023. <https://doi.org/10.1016/j.teler.2023.100063>.
- [5] G. Sheyla, and C. Boonthum-Denecke. Security Risks and Challenges of Remote Working and Learning, The 2023 ADMI Symposium, 2023.
- [6] Common Vulnerabilities and Exposures, "CVE-2023-51219"(Internet), <https://www.cve.org/CVERecord?id=CVE-2023-51219>
- [7] Skype Homepage (<https://www.skype.com>)
- [8] L. Carettoni, Electronegativity: A Study of Electron Security, Doyensec, <https://doyensec.com/resources/us-17-Carettoni-Electronegativity-A-Study-Of-Electron-Security-wp.pdf>
- [9] Discord Homepage (<https://discord.com>)
- [10] Signal Homepage (<https://signal.org>)

# AI 보안(AI Security) 정의의 통합적 고찰

진승헌\*, 김수형\*

## An Integrated Perspective on the Definition of AI Security

Seung-Hun Jin\*, and Soo-Hyung Kim\*

### 요 약

AI 기술은 빠르게 발전하고 있다. 2022년 ChatGPT의 출시 이후, AI는 전문가들만의 연구 대상이 아니라 일반인들도 쉽게 활용할 수 있는 기술로 자리 잡았으며, 이를 통해 텍스트, 이미지 등 다양한 콘텐츠를 생성하는데 활용되고 있다. 생성형 AI 시장은 빠르게 성장 중이며, 해외 보고서에 따르면 2032년에는 1조 3천억 달러 규모로 예상된다. 기업과 조직의 AI 도입률이 늘어나면서 AI 관련 위협도 함께 증가하고 있다. 이러한 상황에서 AI 보안(AI Security)에 관한 관심이 커지고 있지만, AI 보안에 대한 정의와 분류는 주체에 따라 달라 통일된 시각을 갖기 어렵다. 본 논문에서는 산/학/연에서 제시하고 있는 AI 보안의 정의와 특징을 분석하여 통합적 관점에서 AI 보안을 정의하기 위한 핵심요소를 제시하고자 한다.

### Abstract

AI technology is advancing rapidly. Since the launch of ChatGPT in 2022, AI has become a tool that is not only the subject of research for experts but also easily accessible for the general public, enabling the generation of various types of content such as text and images. The generative AI market is growing quickly, and according to overseas reports it is expected to reach a size of \$1.3 trillion by 2032. The adoption rate of AI among businesses and organizations is also increasing, leading to a rise in AI-related threats. In this context, interest in AI security is growing but definitions and classifications vary depending on the stakeholders, making it difficult to have a unified perspective. This paper seeks to examine the definitions and features of AI security as outlined in academia, industry, and research in order to propose the essential components of a thorough definition of AI security.

### Key words

AI Security, Safety, Trust, AI for Security, Security for AI

### 1. 서 론

AI 시대가 우리 눈 앞에 펼쳐지고 있다. 2016년

알파고의 등장은 바둑과 같은 복잡한 분야에서는 AI가 인간을 넘어서는 것은 시기상조라는 막연한 믿음을 깨뜨렸다. 이는 기술이 우리의 예상보다 빠

\* ETRI 사이버보안연구본부 인공지능데이터보안연구실, jinsh@etri.re.kr(교신저자), lifewsky@etri.re.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2023-00229400, 안전한 메타버스 환경을 위한 사용자 인증 및 프라이버시 보호 기술개발)

르게 발전하고 있음을 실감하게 해주었다. 당시만 해도 AI는 관련 전문가들의 연구 영역이라고 여겼지만, 2022년 11월 OpenAI가 ChatGPT를 선보이면서 이런 인식이 급격히 바뀌었다. 이로 인해 AI는 더 이상 특정 분야 전문가들만의 연구 대상이 아닌, 일반인들도 생활에서 쉽게 접하고 이용할 수 있는 기술로 자리매김하기 시작했다. ChatGPT나 미드저니 등과 같은 생성형 AI의 출현으로 사람들은 텍스트, 이미지, 음악, 코딩 등 다양한 콘텐츠를 손쉽게 만들어내고 있다. 이에 따라 AI의 경제적 영향력도 크게 확대되고 있다. 블룸버그 보고서에 따르면, 2022년 400억 달러 규모였던 생성형 AI 시장이 2032년까지 1조 3천억 달러로 증가할 것으로 예상된다. 또한, AI 도입 비율도 빠르게 증가하고 있어서 2018년부터 2023년까지 기업 및 조직의 AI 도입 비율이 약 50%에 머물렀으나, 2024년에는 72%로 증가할 것으로 예상했다.[1] 이제 AI는 전문가들만의 연구 대상이 아닌 우리 삶의 일부가 되어가고 있으며, 이를 통해 우리의 삶은 더욱 나아지고 생산성은 크게 향상될 것으로 보인다. 하지만 AI 시대의 도래는 긍정적인 측면과 아울러 부정적인 면도 함께 가져오고 있다. 대표적으로, 딥페이크 기술을 악용한 성범죄와 같은 심각한 문제가 증가하고 있다.[2] 이 외에도 소수자 차별, 안전사고 등 다양한 AI 관련 위협들이 발생하고 있다.[3] 이에 이러한 위협에 대응할 수 있는 AI 보안에 관한 관심이 높아지고 있다. 하지만 AI 보안에 대한 정의와 분류는 각 주체의 관점에 따라 다르게 정의하고 있어 전체적인 시각을 가지기 어렵다. 본 논문은 산업계, 학계, 연구기관의 다양한 문헌을 분석하여 AI 보안의 정의와 주요 특징을 정리하고, 이를 통해 AI 보안의 정의를 전체적인 관점에서 고찰하는 것을 목적으로 한다.

## II. AI 시대의 보안 위협 분석

AI 기술의 발전은 우리의 삶과 산업에 긍정적인 변화를 가져왔으나, 동시에 새로운 형태의 보안 위협을 만들어냈다. 이러한 위협은 딥페이크 기술을 악용한 성범죄나 AI를 이용한 자동화된 사이버 공

격 등으로 대표된다. 본 장에서는 AI 기술 발전에 따라 발생한 주요 사고사례를 통해 AI 시대의 보안 위협을 분석한다.

Google의 AI 영상 분석 시스템 오류 사례[4], Microsoft Tay AI 챗봇 사례[5], Amazon AI 채용 시스템 편향 사례[6] 등은 AI가 부적절하게 학습되면 심각한 윤리적 문제와 편향성으로 인한 공정성 문제 등을 일으킬 수 있다는 것을 보여주었다. Facebook AI 추천 시스템의 극단적 콘텐츠 확산 사례[7], GPT-3 악용 사례[8] 등에서는 AI가 잘못 사용될 경우 허위 정보의 확산을 조장하고 사회 혼란을 일으킬 수 있음을 보여주었다. 또한, Tesla 자율주행 차량 사고[9] 등을 통해 AI의 오류가 인간의 안전과 생명에 심각한 영향을 미칠 수 있음을 시사하였다. 또한, 딥페이크 기술은 점차 고도화되면서 성범죄, 금융 사기, 정치적 허위 정보 확산 등 다양한 범죄에 악용되고 있다.[13] 이러한 국내외 사례들은 딥페이크 기술이 강력한 도구로서 생산성과 창의성을 높이는 동시에, 악용된다면 개인과 국가사회 및 경제에 심각한 피해를 줄 수 있음을 보여준다.[14]

AI 기술이 여러 분야에서 혁신적인 변화를 가져왔지만, 이와 동시에 발생하는 보안 위협은 위에 열거된 만큼 심각하다. AI 시스템이 부적절한 데이터 학습으로 인한 윤리적 문제, AI 시스템의 인식 오류로 인한 인명 사고, AI 채용 시스템에서 나타나는 편향성, 그리고 딥페이크 기술의 악용은 AI의 오용 또는 취약점이 개인의 인권과 사회적 안전에 중대한 영향을 미칠 수 있음을 보여준다.

## III. AI 보안(AI Security)의 정의와 핵심요소

다양한 형태의 AI 위협이 발생하면서 AI 시스템과 데이터를 보호하고, AI를 활용하여 보안을 강화하는 AI 보안이 주목받고 있다. 하지만 산업계, 학계, 연구기관의 관점에 따라 정의와 강조점이 상이하여, 통일된 정의가 부재한 상황이다. 이는 단일한 정의가 꼭 필요하다는 것을 의미하는 것은 아니지만, 다양한 관점을 이해하는 것이 AI 보안의 통합적 시각을 형성하는 데 중요하다고 생각한다. 본 장

에서는 산업체[15~18], 학계[19~21], 연구기관[22][23]이 제시하고 있는 AI 보안의 정의와 특징을 분석하여 AI 보안 정의의 핵심요소를 도출하였다.

AI 보안 정의와 주요 내용을 주체별 특징으로 정리하면 다음 표와 같다.

표 1. AI 보안 정의의 주체별 특징

주체	특징
산업체	- 실용적 측면에 초점, AI 시스템의 무결성과 보안을 강화하는 기술 - IBM, Microsoft, Google은 제품과 서비스의 보안 강화를 중심으로 AI를 사용하며, AI의 책임 있는 사용을 강조
학계	- 이론적 연구와 실증적 접근을 통해 AI 시스템의 위험 식별 및 완화로 정의 - Stanford와 MIT는 AI의 신뢰성과 견고성을 확보하여 사회적 채택과 안전한 활용에 중점
연구기관	- 표준화와 거버넌스를 통한 정책 개발을 다루며, AI와 사이버보안의 융합에 초점을 맞춤 - NIST와 ENISA는 AI 보안의 수명주기 전반에 걸친 위협 탐지와 완화를 강조

각 주체별 AI 보안의 정의와 주요 특징을 종합하여 AI 보안을 정의할 때에 고려할 3가지 핵심요소를 아래와 같이 도출하였다.

- 1) **삼원적 정의** : 일반적으로 AI 보안은 두 가지 측면에서 정의되고 있다. 첫 번째는 ‘Security for AI’로, AI 시스템과 그 내부 데이터의 무결성, 가용성, 기밀성을 보호하는 것이다. 두 번째는 ‘AI for Security’로, AI를 활용해 보안 시스템을 개선하고 강화하는 것을 의미한다. 그러나 최근에는 개인정보보호, 윤리, 신뢰성 등 사회적 요구를 포함하여 더욱 광범위한 정의가 필요하다는 의견이 증가하고 있다. 예를 들어, AI 시스템의 예측이 투명하고 편향되지 않도록 하는 신뢰성은 AI 보안의 핵심 과제 중 하나다. 이러한 맥락에서, AI Security는 Security for AI, AI for Security, Safe Use of AI라는 세 가지 축으로 확대하여 정의하는 것이 적절하다. 여기서 ‘Safe Use of AI’는 AI의 윤리적 사용, 개인정보보호, 신뢰성 확보 등 책임 있는 AI 활용을 포함한다.

- 2) **투명성과 책임성** : AI 보안에서 투명성과 책임성은 신뢰 확보의 핵심요소다. 점점 더 많은 국가가 AI의 투명성과 책임성을 강화하는 법규를 제정하고 있으며, 이 과정에서 데이터 거버넌스와 AI 시스템의 설명 가능성이 중요한 역할을 한다. 예를 들어, 유럽연합의 AI 법안은 고위험 AI 시스템에 대해 투명성을 의무화하며, 기업과 조직은 규제 준수를 위해 시스템의 내부 작동 방식을 공개하고 감사 가능한 구조를 마련해야 한다. 미국과 한국을 비롯한 여러 국가에서도 AI 관련 법적 프레임워크를 정비하고 있으며, 이러한 법제화는 투명한 데이터 활용과 AI 모델의 예측 가능성에 대한 사회적 요구를 반영하고 있다.

- 3) **포괄적 접근** : AI 보안 문제는 단순히 기술적 오류를 해결하는 것에 그치지 않는다. 법적, 사회적, 윤리적 책임까지 함께 고려해야 한다. 예를 들어, 자율주행차 사고 발생 시에 책임 소재를 명확히 규명하려면 기술적 문제뿐만 아니라 법적·정책적 검토도 필수적이다. 이를 위해 다양한 이해관계자(정부, 기업, 시민사회 등)가 참여하는 포괄적 접근이 필요하다. 특히 AI 보안과 관련된 법적 규제는 기술의 발전과 함께 성장해야 하며, 기술, 법, 정책, 서비스 간 다학제적 협업과 균형이 중요하다.

## IV. 결론

AI 시대가 도래하면서, 사용자들은 생산성의 급격한 향상 등과 같은 혁신적인 경험을 하고 있다. 반면에 딥페이크 악용으로 인한 성범죄 등과 같은 피해에 노출되기도 했다. 본 고에서는 AI 오남용 또는 취약성으로 인하여 발생한 사례를 알아보고, 산업체, 학계, 연구계의 다양한 문헌과 관점을 분석해 AI 보안의 정의와 특징을 정리하고, AI 보안 정의 시에 고려해야 할 3가지 핵심요소를 제시하였다. 향후에는 다양한 추가 문헌 분석을 통하여 AI 보안에 대한 분류체계를 수립하고 통합적 보안 프레임워크를 정립하여, AI 위협과 우려에 대한 대응을 효과적으로 수행하는 데 이바지하고자 한다.

## 참 고 문 헌

- [1] 마케팅 GPT·코파일럿 챗...미국 기업의 혁신을 이끄는 생성형 AI. (2024). [https://dream.kotra.or.kr/kotranews/cms/news/actionKotraBoardDetail.do?SITE\\_NO=3&MENU\\_ID=180&CONTENTS\\_NO=1&bbsGbn=243&bbsSn=243&pNttSn=218851](https://dream.kotra.or.kr/kotranews/cms/news/actionKotraBoardDetail.do?SITE_NO=3&MENU_ID=180&CONTENTS_NO=1&bbsGbn=243&bbsSn=243&pNttSn=218851).
- [2] 딥페이크 성범죄' 10대 증가 추세"...경찰, 집중 단속. (2024). <https://news.kbs.co.kr/news/pc/view/view.do?ncd=8044872>.
- [3] AI도 실수 많이한다...AI가 실패한 최신 사건 10 가지. (2021). <https://www.aitimes.com/news/articleView.html?idxno=135396>.
- [4] Google apologises for Photos app's racist blunder . (n.d.). <https://www.bbc.com/news/technology-333478866>.
- [5] 인공지능 세뇌의 위험...MS 채팅봇 '테이' 차별 발언으로 운영중단. (2016). <https://www.yna.co.kr/view/AKR20160325010151091>.
- [6] AI 채용' 차별 논란에... 뉴욕 "성별-인종 편향 공개하라", 첫 규제. (2023). <https://www.donga.com/news/Inter/article/all/20230707/120119104/1>.
- [7] 가짜 뉴스로 흥한 자가 왕이 될 것이다?"...美대선 판도 최대 변수될 수도. (2024). <https://www.mk.co.kr/news/it/11126911>.
- [8] Malicious Prompt Engineering With ChatGPT . (2023). <https://www.securityweek.com/malicious-prompt-engineering-with-chatgpt/>.
- [9] 테슬라, 오토파일럿 사망사고 결함 인정..."유족과 합의. (2024). <https://www.hankyung.com/article/2024040992315>.
- [13] 성범죄에 보이스피싱까지...딥페이크의 진화 . (2024). <https://www.mediatoday.co.kr/news/articleView.html?idxno=320744>.
- [14] 인공지능의 악용 사례, 딥페이크 기술과 과제 . (2022). [https://spri.kr/posts/view/23469?code=data\\_all&study\\_type=industry\\_trend](https://spri.kr/posts/view/23469?code=data_all&study_type=industry_trend).
- [15] What is AI security?. (n.d.). <https://www.ibm.com/think/topics/ai-security>.
- [16] Basic concepts of AI security. (n.d.). <https://learn.microsoft.com/en-us/training/modules/fundamentals-ai-security/2-basic-concepts-of-ai-security>.
- [17] Introducing Google's Secure AI Framework . (n.d.). <https://blog.google/technology/safety-security/introducing-googles-secure-ai-framework/>.
- [18] AI 보안 (AI Security). (n.d.). <https://www.fasoo.com/glossary/a/ai-%EB%B3%B4%EC%95%88-ai-security>.
- [19] Stanford University(Stanford Center for AI Safety). (n.d.). <https://aisafety.stanford.edu/>.
- [20] MIT(AI Risk Repository) . (n.d.). <https://airisk.mit.edu/>.
- [21] UC Berkeley(AI Security Initiative) . (n.d.). <https://cltc.berkeley.edu/program/ai-security-initiative/>.
- [22] AI Risk Management Framework. (n.d.). <https://www.nist.gov/itl/ai-risk-management-framework>.
- [23] Cybersecurity of AI and Standardisation . (n.d.). <https://www.enisa.europa.eu/publications/cybersecurity-of-ai-and-standardisation>.

# 무결성 검증을 위한 DNN 기반 특징추출

강윤희\*, 권태언\*\*

## DNN based feature extraction for verifying integrity

Yunhee Kang\* and Taeun Kwon\*\*

데이터 무결성 확인은 영상데이터의 변경 여부를 확인하기 위한 필수적인 기능이다. 무결성 확인은 SVD의 특이값을 구성하여 수행할 수 있으나 특이값을 얻기 위해서는 높은 시간복잡도로 인해 계산 비용의 제약점을 갖는다[1,2]. 이 논문에서는 무결성 검사를 위해 영상데이터의 추론과정에서 특징 추출한 중간값을 활용하는 기법을 제안한다. 이를 위해 추론을 위해서는 기학습 모델을 양자화(quantization)하여 최적화한 DNN(Deep Neural Network)을 구성하여 사용한다. 양자화는 모델의 파라미터를 작은 비트로 표현함으로써 계산과 메모리 접근속도를 높이는 경량화 기법이다[3]. CNN(Convolutional Neural Network)은 입력 데이터에 대해 계층적으로 학습하는 DNN으로서 입력 데이터의 특징을 추출하기 위해 Convolution, Pooling, Activation의 기본블록을 순차적으로 수행한다[4]. 제안한 특징추출 기법은 영상은 DNN에 입력 후 DNN의 수행과정에서 네트워크계층의 중간결과물을 핑거프린트 구성에 활용한다. 이를 위해 DNN 네트워크의 수행 행태의 모니터링을 위해 hook 을 구성하여 핑거프린트 구성을 위한 중간결과물을 활용한다. 실험은 영상분류에 ResNet18을 사용하며 Jetson Orin Nano와 PyTorch 환경에서 특징추출 모듈을 개발한다. Hook을 사용한 특징추출은 정의된 계층에서 수행된 활성화함수의 수행결과를 영상의 특징벡터로 구성하여 사용한다. 무결성 검증에 필요한 특징에 대한 핑거프린트는 특징벡터에 대해 암호화해시함수인 SHA-256을 적용하여 생성한다.

### Key words

data integrity, DNN, feature extraction, hook, finger print

### 참고문헌

- [1] Shen, Jianhong (Jackie) (2001). "On the singular values of Gaussian random matrices". *Linear Alg. Appl.* 326 (1-3): 1-14.
- [2] Bosagh Zadeh, Reza; Carlsson, Gunnar (2013). "Dimension Independent Matrix Square Using MapReduce"
- [3] Benoit Jacob, Skirmantas Kligys, Bo Chen, Menglong Zhu, Matthew Tang, Andrew Howard, Hartwig Adam, Dmitry Kalenichenko; *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 2704-2713
- [4] Y. LeCun, K. Kavukcuoglu and C. Farabet, "Convolutional networks and applications in vision," , Paris, France, 2010, pp. 253-256, doi: 10.1109/ISCAS.2010.5537907.

---

\* 백석대학교, : yhkang@bu.ac.kr(교신저자표시), \*\* (주)하스퍼, peterkwon@harsper.co.kr

※ 본 논문은 중소벤처중소벤처기업부(중소기업기술정보진흥원, RS-2023-00225234) 2023 년도 산학연 CollaboR&D 사업의 산업현장의 디지털 영상데이터의 AI 기반 무결성 및 검증 기술 개발과제의 지원을 받아 수행된 연구임

# 동영상 관심영역 암호화 시 견고한 스트림 암호화 적용 방법

김덕한\*, 김영갑\*\*

## A Robust Stream Cipher Method for Video Region of Interest Encryption

Deok-Han Kim\* and Young-Gab Kim\*\*

### 요 약

동영상 암호화는 비밀키를 통해 특정 사용자만 영상을 복호화할 수 있게 하며, 동영상 관심영역 암호화는 민감한 부분만 선택적으로 암호화해 보안성을 높인다. 그러나 기존 동영상 관심영역 암호화 방식은 키 스트림(Key Stream)을 반복 사용해 보안 취약성을 초래할 수 있다. 본 논문에서는 타일마다 다른 초기화 벡터(IV; Initial Vector)를 적용하여 키 스트림의 중복을 방지하고, 보안성을 강화하는 방법을 제안한다.

### Abstract

Video encryption uses a secret key to ensure that only a specific user can decrypt the video, while region-of-interest encryption selectively encrypts sensitive parts to increase security. However, existing video interest region encryption schemes use the key stream (KS) repeatedly, which can lead to security vulnerabilities. In this paper, we propose a method to apply a different initialization vector (IV) for each tile to prevent the redundant use of key streams and enhance the security of video interest region encryption.

### Key words

video encryption, video roi encryption, stream cipher

## 1. 서 론

최근 동영상 콘텐츠의 종류와 양이 증가하면서, 민감정보를 보호하기 위한 동영상 비식별화 기술이 중요해지고 있다. 동영상 암호화 기술은 비밀키를 가진 특정 사용자에게 한해서 복원할 수 있도록 동영

상을 비식별화 처리할 수 있는 특징을 가진 동영상 비식별화 기술 중 하나이다[1,2,3]. 동영상 관심영역(ROI; Region of Interest) 암호화 기술은 민감정보를 포함한 영역을 선택적으로 암호화하여, 비민감한 부분의 활용성을 유지하면서 개인정보와 같은 민감정보를 보호할 수 있다[4,5].

\* 세종대학교, lpsi4862@sju.ac.kr,

\*\* 세종대학교, alwaysgabi@sejong.ac.kr(교신저자)

※ 이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2019-II190231, 공공 인프라 안전을 위한 인공지능 기반 영상보안 기술 및 시스템 개발)



기존의 동영상 암호화 기술은 인코딩 중 구문 요소를 순차적으로 암호화하며, 복호화를 위해 구문 요소의 암호화 및 복호화가 동일한 순서로 이루어져야 한다. 동영상 관심영역 암호화 기술은 오류 전파(error propagation) 현상을 방지하기 위해 타일 단위로 인코딩 및 암호화를 수행하지만, 이 경우, 구문 요소의 순차적인 암호화가 보장되지 않을 수 있다. 현재의 동영상 관심영역 암호화 기술은 이러한 문제를 해결하기 위해 각 프레임의 타일마다 동일한 키 스트림을 사용하지만, 이는 키 스트림의 재사용으로 인한 보안 취약성을 초래할 수 있다.

본 논문에서는 이러한 문제를 해결하기 위해, 각 프레임의 타일마다 다른 초기화 벡터(IV; Initial Vector)를 사용하여 키 스트림이 반복 사용되지 않도록 하는 방법을 제안한다.

## II. 본 론

### 2.1 문제 정의

동영상 암호화 기술은 암호화 과정에서 계산 오버헤드를 줄이고 처리시간을 단축하기 위해, 인코딩 수행 중 시각적으로 중요한 일부 구문 요소만을 선택적으로 암호화하는 방식으로 수행된다. 구문 요소를 암호화하는 방식에는 두 가지 주요 접근 방식이 있다. 첫 번째는, 구문 요소들을 블록으로 구성하여 순차적으로 암호화하는 블록 암호화 방식이고, 두 번째는 의사 난수 생성기(PRNG; Pseudo-Random Number Generator)를 통해 생성된 키 스트림과 구문 요소들을 순차적으로 XOR 연산하여 암호화하는 스트림 암호화 방식이다. 두 방식 모두 인코딩 시 구문 요소의 암호화 순서가 디코딩 시의 복호화 순서와 일치해야만 정상적인 복호화가 가능하다.

동영상 관심영역 암호화는 민감한 정보를 보호하기 위해 동영상의 일부 영역만 선택적으로 암호화하는 기술이다. 암호화는 실시간 처리, 구문 준수를 위해 일반적으로 인코딩 과정 중 수행된다. 그러나 이 과정에서 암호화된 영역이 암호화되지 않은 영역에 영향을 미치는 오류 전파 현상이 발생할 수 있다. 이러한 오류 전파는 암호화되지 않은 영역이

예측 과정에서 암호화된 영역을 참조할 때 발생할 수 있으며, 이는 예상하지 않은 시각적 왜곡을 발생시켜 동영상의 활용도를 저하시킬 수 있다.

HEVC(High Efficiency Video Coding) 동영상 관심영역 암호화 시, 오류 전파 현상을 방지하기 위해 타일 단위로 인코딩을 수행한다. 타일은 HEVC에서 병렬처리를 지원하기 위해 도입된 기술로[6], 각 타일은 독립적으로 인코딩되며 예측 범위가 타일 경계까지 제한된다. 타일 기반 인코딩 시, 병렬 처리를 통해 인코딩 속도를 향상시키고, 오류 전파 현상을 타일의 경계로 제한할 수 있다. 그러나 타일들이 병렬 처리되기 때문에 각 타일의 구문 요소가 인코딩되는 순서를 보장할 수 없다.

현재의 HEVC 동영상 관심영역 암호화 기술은 이러한 문제를 해결하기 위해, 각 프레임의 타일마다 동일한 키 스트림을 사용하여 구문 요소를 암호화하고 있다. 그러나 같은 키 스트림을 여러 타일에서 반복적으로 사용하는 경우, 보안 문제가 발생할 수 있다. 예를 들어, 공격자가 입력 값을 임의로 조작할 수 있고, 인코딩에 사용된 인코더의 구조를 알고 있으며, 인코딩 및 암호화된 동영상에 접근할 수 있는 상황이라면, 특정 타일에서 사용된 키 스트림을 추측할 수 있다. 공격자가 이 키 스트림을 알아낸다면, 이를 사용해 암호화된 전체 동영상을 복호화할 수 있게 된다. 따라서, 키 스트림이 타일마다 중복되지 않도록 하는 방법이 필요하다.

### 2.2 제안 방법

동영상 암호화에서 각 타일의 키 스트림의 중복 사용을 방지하기 위해서는, 타일마다 서로 다른 초기화 벡터를 사용이 필요하다. 각 프레임의 타일별 초기화 벡터는 다음 두 가지 조건을 만족하며 초기화되어야 한다.

- **조건 1.** 프레임의 각 타일을 암호화할 때 사용된 초기화 벡터를 복호화 시점에 알아야 한다.
- **조건 2.** 동일한 키 스트림이 생성되는 것을 방지하기 위해, 한 번 사용된 초기화 벡터는 재사용되지 않아야 한다.

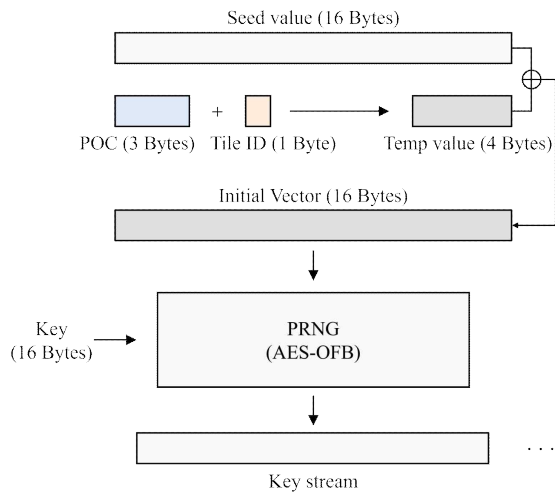


그림 1. 각 프레임의 타일별 IV 초기화 방법  
 Fig. 1. IV initialization method for each tile in a frame

본 논문에서는 AES-OFB 모드를 사용하여 128비트 키를 기반으로 의사 난수 생성기(PRNG; Pseudo Random Number Generator)를 구현한 경우를 설명한다. 제안된 방법은 그림 1과 같이 시드 값(seed value, POC(Picture Order Count), 타일 ID를 결합하여 각 타일마다 고유한 초기화 벡터를 생성하는 방식으로 수행된다. 먼저, 16바이트 크기의 임의의 시드 값을 준비한다. 이 값은 암호화 과정에서 고유한 키 스트림을 생성하는 데 중요한 역할을 한다. 이후 POC 및 타일 ID를 준비한다. POC는 프레임 번호를 나타내며, 3바이트 크기로 준비한다. 이는 초당 60 프레임으로 24시간 이상의 동영상을 처리할 수 있는 크기이다. 타일은 대부분의 경우 하나의 프레임에 255개 이상의 타일이 존재하지 않기 때문에, 타일 ID를 1바이트 크기로 준비한다. 이후 POC와 타일 ID를 연결(concatenation)하여 4바이트 크기의 임시 값(temp value)을 생성한다. 마지막으로, 16바이트의 시드 값과 4바이트의 임시 값을 XOR 연산하여 각 프레임의 타일마다 고유한 초기화 벡터를 생성한다. 이 초기화 벡터는 각 프레임의 각 타일마다 고유하며, 동일한 동영상에서 중복되지 않는다.

이 방식으로 생성된 초기화 벡터는 타일마다 고유하므로, 키 스트림의 중복 사용을 완전히 방지할 수 있다. 또한, POC와 타일 ID는 디코더에서도 접근 가능한 정보이기 때문에, 암호화 및 복호화 과정에서 시드 값만 공유되면 인코더와 디코더가 동일

한 초기화 벡터를 사용할 수 있다. 이러한 방법을 통해 각 타일별로 고유한 키 스트림을 생성하여 보안성을 강화하고, 더 안전한 동영상 관심영역 암호화를 실현할 수 있다.

### III. 결 론

본 논문에서는 HEVC 동영상 관심영역 암호화시 키 스트림의 중복 사용을 방지하기 위한 방법을 제안하였다. 제안한 방법은 POC 및 타일 ID를 사용하여 초기화 벡터를 타일마다 초기화하여, 키 스트림의 중복 사용을 방지한다.

제안한 방법은 키 스트림의 재사용으로 인한 민감정보의 노출 위험을 줄이고, 동영상 관심영역 암호화의 보안성을 강화하는 데 기여한다.

### 참 고 문 헌

- [1] A. Massoudi et al., "Overview on selective encryption of image and video: challenges and perspectives", *Eurasip Journal on Information Security*, vol. 2008, no. 1, p. 179290, 2008.
- [2] F. Liu and H. Koenig, "A survey of video encryption algorithms", *Computers & Security*, vol. 29, no. 1, pp. 3-15, 2010.
- [3] K. M. Hosny et al., "Multimedia security using encryption: A survey", *IEEE Access*, vol. 11, pp. 63027-63056, 2023.
- [4] M. Farajallah et al., "ROI encryption for the HEVC coded video contents", in *2015 IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 1080-1084.
- [5] J.-Y. Yu and Y.-G. Kim, "Coding unit-based region of interest encryption in HEVC/H.265 video", *IEEE Access*, vol. 11, pp. 47967-47978, 2023.
- [6] K. Misra et al., "An overview of tiles in HEVC", *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 969-977, Dec. 2013.

# URL 데이터를 활용한 피싱 웹사이트 탐지 모델 개발

이원정\*, 함종수\*\*

## Development of a phishing website detection model using URL data

Wonjung LEE\*, and Jongsoo Ham\*\*

### 요 약

피싱 웹사이트는 전 세계적으로 심각한 사이버 보안 위협 중 하나로, 그 기법이 점점 더 정교해지고 있어 다양한 연구들이 진행되고 있다. 본 연구에서는 URL만으로 피싱 사이트 탐지가 가능하도록 기존 연구보다 다양한 데이터 특성을 학습하는 머신러닝 모델을 제안한다. 7가지 머신러닝 모델을 기존 연구에 사용되던 특성만으로 학습한 결과와 새롭게 추가한 특성까지 같이 학습한 결과를 비교하고, 새로운 특성이 실제 성능 향상에 도움이 되었는지 특성 중요도를 확인하였다. 그 결과 재현율(Recall) 97.11%, F1-Score 97.34%, 정밀도(Precision) 97.58%로 Random Forest 모델이 가장 우수한 성능을 보였으며, 기존 연구에 사용되던 특성만으로 학습한 결과보다 재현율(Recall)은 4.1%, F1-Score는 4.5%, 정밀도(Precision)는 4.9% 향상된 것을 확인하였다.

### Abstract

Phishing websites are one of the serious cybersecurity threats worldwide, and various studies are being conducted as their techniques are becoming increasingly sophisticated. In this study, we propose a machine learning model that learns more diverse data features than existing studies so that phishing site detection is possible with URLs alone. We compared the results of learning seven machine learning models with only the features used in existing studies and the results of learning newly added features together, and confirmed the feature importance to see if the new features actually helped improve performance. As a result, the Random Forest model showed the best performance with a recall of 97.11%, an F1-Score of 97.34%, and a precision of 97.58%. It was confirmed that the recall, F1-Score, and precision improved by 4.1%, 4.5%, and 4.9%, respectively, compared to the results learning only with the features used in existing studies.

### Key words

Phishing, Machine Learning, URL, Random Forest, Feature

---

\*프리랜서, jjung2516@naver.com, \*\*한양대학교 인공지능융합학과, hamjongsoo0724@gmail.com(교신저자)

## I. 서 론

피싱 웹사이트는 전 세계적으로 심각한 사이버 보안 위협 중 하나로, 사용자로부터 중요한 정보를 탈취하거나 악성 소프트웨어를 설치하는 등의 방법으로 피해를 유발하고 있다. 최근 들어 피싱 공격은 그 기법이 점점 더 정교해지고 있으며, 새로운 유형의 피싱 웹사이트 탐지는 기존의 탐지 시스템으로는 대응이 어려워지고 있다. 이러한 정교한 설계로 인해 사용자가 쉽게 구별하기 어려운 피싱 사이트를 탐지하기 위한 다양한 연구들이 진행되어왔다.

기존 연구들은 주로 URL 길이, 도메인 정보, HT TPS 사용 여부, IP 주소 등과 같은 기본적인 URL 및 도메인 관련 특징을 분석하여 피싱 여부를 탐지하는 방식에 집중해왔다.

본 연구에서는 URL 만으로 피싱 사이트 탐지가 가능하도록, 기존 연구보다 다양한 데이터 특성을 활용하여 저비용 환경에서도 탐지가 가능한 머신러닝 모델을 활용하고자 한다. 특히, 본 연구에서는 URL 단축 서비스 사용 여부, 파일 확장자, IP 주소 사용 여부와 같은 기존 연구들에서 다룬 요소뿐만 아니라, URL의 구조적 특성(예: URL 경로 길이, URL 엔트로피, 서브 도메인의 길이, 'www.'의 포함 여부 등)을 추가 분석함으로써 피싱 사이트 탐지 모델의 성능을 향상시키고자 한다. 또한, 본 연구에서 제안하는 데이터 특성은 피싱 사이트가 URL에서만 보이는 특정 패턴들을 보다 세밀하게 포착할 수 있어, 기존 방법보다 높은 정확도와 재현율이 기대된다.

## II. 관련 연구

피싱 웹사이트 탐지를 위한 기존 연구들도 주로 URL의 구조적 특성과 도메인 정보를 활용한 탐지 기법을 중심으로 이루어졌다. 권현 외 2인은 뉴럴네트워크 이용하여 악성 URL 탐지를 위하여 URL의 길이, 연속된 단어의 수, 도메인 엔트로피, IP 주소, IP 지리적 위치, TTL(Time to Live) 평균 시간 및 표준편차, 도메인 유지 시간, 접속된 나라별 순위, 그리고 Passive DNS 변화 빈도수 등의 특성을 이용

하였다[1]. 오수민 외 1인은 머신러닝 기반으로 피싱 사이트 탐지를 위하여 URL, 도메인, 웹사이트 메타데이터와 관련된 특성을 사용하였다. 총 59개의 특성 중 상위 중요도 특성 몇 가지를 보면 웹사이트의 PageRank, 도메인 나이, 외부 하이퍼링크 비율, Google Index, Https 포함 여부 등을 사용하였다[2]. 이러한 기법들은 피싱 웹사이트와 정상 웹사이트를 구분하는 데 유용한 정보를 제공하지만, URL 분석을 위한 데이터 특성의 범위는 제한적일 수 있다.

본 연구에서는 기존 연구에서 사용된 URL 길이, URL 단축 서비스 사용 여부, ip 주소 포함 여부 등의 기본적인 특성과 함께 피싱 사이트 탐지 성능을 향상시키기 위해 다양한 URL 특징을 추가로 활용한다.

## III. 연구 방법

본 연구에서는 데이터 수집 후, 전처리를 통해 피싱 사이트 탐지 성능을 향상시킬 것으로 기대되는 다양한 데이터 특성을 여러 머신러닝 모델로 학습한다. 데이터 수집부터 모델 평가까지의 방법을 아래에 단계별로 설명한다.

### 3.1 데이터 수집

본 연구에서는 Kaggle[3]에서 제공된 데이터 셋을 사용하였다. URL과 웹사이트 상태를 나타내는 특성 type 두 가지로 이루어져 있으며, type은 'benign', 'phishing', 'malware', 'defacement'로 구성되어 있다. 데이터는 총 651,191개의 URL이 포함되어 있고, 정상 URL 428,000개와 피싱 URL 223,088개로 구성되어 있다.

### 3.2 데이터 전처리

데이터 전처리 과정에서는 데이터 셋에 제공된 특성 type에서 'benign'은 정상 사이트, 'phishing', 'malware', 'defacement' 3가지를 'phishing'으로 통합하고 이를 다시 정상 사이트는 '0', 피싱 사이트는 '1'로 인코딩하였다. URL을 분석하여 정상과 피싱의 탐지를 위한 특성을 가공하였고, 기존 피싱 사이트 연구에 사용되었던 특성도 사용하여 URL 포함

총 39개의 특성을 사용한다. Category를 target으로 설정하고, URL을 포함하여 문자형 특성 6개를 제외한 32개를 사용한다. 특성은 기존 연구에 사용되었던 url 길이, 각 특수문자 개수('?', '-', '=', ':', '%', '+', '/'), https 유무, url 숫자 개수, url 문자 개수, url 단축서비스 사용 여부, ip주소 포함 여부, long\_url\_phishing, hyphen\_in\_url 15개에 새로운 특성으로 url 전체 특수문자 개수, url 경로 길이, 쿼리 파라미터 수, 서브도메인 길이, 서브도메인 개수, url\_entropy, 특정 키워드 유무, 특수문자 비율, 파일 확장자 길이, 파일 확장자 특수 문자 개수, 파일 확장자 숫자 포함 여부, url 경로 길이, 최상위 도메인, 최상위 도메인이 여러개인 경우, 의심스러운 파일 확장자, 호스트/도메인 명에 숫자가 많은 경우, 'www.' 포함 여부 까지 17개를 추가하였다.

### 3.3 모델링

피싱 사이트 탐지 성능을 비교하기 위하여 다양한 머신러닝 모델을 사용하였다.

Logistic Regression, Random Forest, Gradient Boosting, Decision Tree, KNN, XGBoost, MLP 모델을 사용하였으며, 학습에는 전체 데이터 셋의 80%를 사용하였다.

Logistic Regression 모델은 데이터 표준화를 위해, KNN은 성능과 정확도를 높이기 위해 StandardScaler를 적용하였다. Random Forest, Gradient Boosting은 100개의 트리로 구성하였으며, Decision Tree를 포함한 트리 기반 모델은 모두 클래스 불균형 문제를 완화하기 위해 가중치를 자동으로 조절하는 방법을 적용하였다. XGBoost 모델은 이진 분류 예측 오류를 최소화하기 위하여 위해 'logloss'를 지표로 적용하였다.

### 3.4 모델 평가

모델의 성능은 재현율(Recall), F1-Score, 정밀도(Precision)의 지표를 사용하여 평가한다. 상위 3개의 모델은 ROC와 AUC까지 확인한다. 본 연구에서 새롭게 제안한 특성을 사용하여 학습했을 때 피싱 사이트 탐지 성능을 향상시키는지 확인하기 위하여 기존 연구에서 선정한 특성만 학습한 모델도 같은

지표로 평가하여 성능을 비교한다.

새롭게 사용된 특성이 모델 성능 향상에 도움이 되었는지 확인하기 위해 모델에 맞는 특성 중요도 분석도 진행하였다.

## IV. 실험 및 결과 분석

학습을 진행한 7가지 모델의 성능을 확인하기 위해 재현율(Recall), F1-Score, 정밀도(Precision)를 측정하여 평가하였고 이 결과를 표 1에 정리하였다.

표 1. 성능 평가 지표

Table 1. Performance Evaluation Metrics

모 델	Recall	F1 Score	Precision
Logistic Regression	89.49%	90.60%	92.12%
Random Forest	97.11%	97.34%	97.58%
Gradient Boosting	94.69%	95.20%	95.78%
Decision Tree	96.29%	96.29%	96.29%
KNN	96.55%	96.75%	96.97%
XGBoost	96.92%	97.28%	97.66%
MLP	96.59%	96.80%	97.01%

기존 연구에서 사용한 특성만으로 학습을 진행한 모델도 성능 비교를 위해 같은 방법으로 평가를 진행하였으며 이를 표 2에 정리하였다.

표 2. 기존 연구 특성 성능 평가 지표

Table 2. Performance Evaluation Metrics of Previous Study Characteristics

모 델	Recall	F1 Score	Precision
Logistic Regression	81.45%	83.57%	88.29%
Random Forest	93.01%	92.84%	92.68%
Gradient Boosting	86.36%	87.79%	90.01%
Decision Tree	96.25%	96.25%	96.25%
KNN	91.89%	92.40%	93.00%
XGBoost	90.10%	91.17%	92.59%
MLP	92.40%	92.66%	92.93%

각 실험에서 7개 모델 중 성능이 가장 높게 나타난 Random Forest 모델의 성능을 비교해 보면 재현율(Recall)은 97.11%와 93.01%, F1-Score는 97.34%와 92.84%, 정밀도(Precision)는 97.58%와 92.68%로

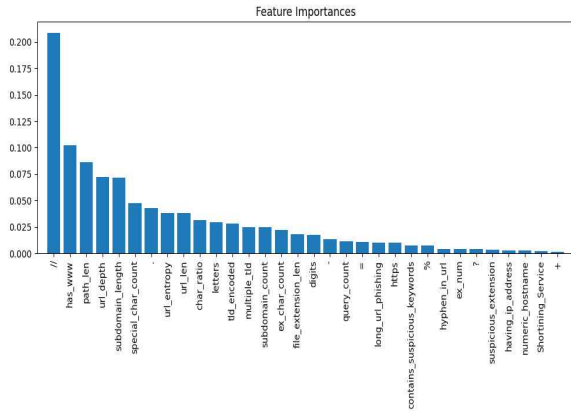


그림 1. 특성 중요도  
Fig. 1. Feature Importances

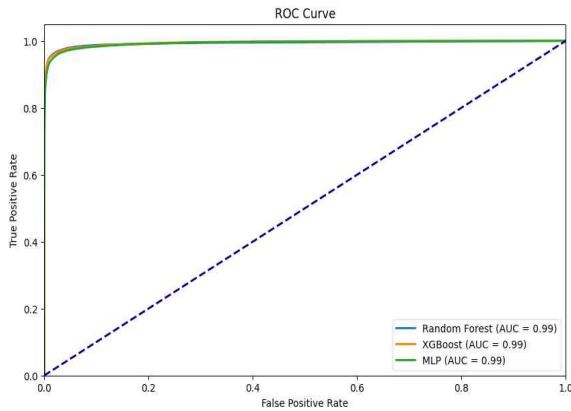


그림 2. 모델 분류 성능 평가  
Fig. 2. ROC Curve

세 가지 지표에서 모두 새롭게 제안한 특성을 함께 학습한 모델이 높은 성능을 나타냈다. 그림 1은 성능이 가장 높은 Random Forest 모델에 대하여 특성 중요도 분석을 진행한 결과이다.

특성 중요도 분석을 진행한 결과 ‘www.’ 포함 유무, URL 경로 길이, URL 경로 깊이, 서버도메인 길이 등 새롭게 제안한 특성이 정상 사이트와 피싱 사이트를 구분하는 데 도움이 되는 것을 확인할 수 있었다.

표 1의 성능 평가 지표에서 상위 3개 모델은 Random Forest, XGBoost, MLP이며, 이 세 가지 모델의 분류 성능 평가를 진행한 결과는 그림 2와 같다.

그림 2에 나타난 AUC 값을 보면 Random Forest와 XGBoost, MLP가 모두 0.99로 동일하게 나오지만, 성능 평가 지표와 ROC Curve까지 종합적으로

성능을 비교해 보면, Random Forest가 가장 우수한 모델임을 확인할 수 있었다.

## V. 결 론

본 연구는 URL만으로 피싱 사이트 여부를 탐지할 때 성능을 향상시킬 수 있는 새로운 특성을 제안하였고, 이러한 특성을 기존 연구에서 사용되던 특성과 함께 학습하였을 때 머신러닝 모델의 성능이 향상되는 것을 확인하였다. 특히 7가지의 머신러닝 모델 중 Random Forest 모델이 가장 우수한 성능을 보였고, 이 모델에서 특성 중요도를 분석하여 본 연구에서 제안한 새로운 특성들이 실제로 성능 향상에 기여했음을 보였다.

향후 연구에서는 URL만 사용한 피싱 탐지 기법과 다른 특성을 포함한 모델의 성능 비교를 진행하여 실시간성을 함께 고려한 향상 방법을 제안할 수 있을 것으로 기대한다.

## 참 고 문 헌

- [1] 권현, 박상준 and 김용철. "뉴럴네트워크 기반에 악성 URL 탐지방법 설계." 융합정보논문지, vol. 11, no. 5, pp. 30-37, 2021, <http://dx.doi.org/10.22156/CS4SMB.2021.11.05.030>
- [2] 오수민, 박민서, "머신러닝 기반 피싱 사이트 탐지 모델," 문화기술의 융합, Vol.10 No.4, 2024
- [3] Kaggle, <https://www.kaggle.com/datasets/sid321axn/malicious-urls-dataset> [accessde : Jun. 28, 2024]
- [4] 사준호 and 이상진. "피싱사이트 실시간 탐지 기법." 정보보호학회논문지, vol. 22, no. 4, pp. 819-825, 2012.
- [5] 이상윤, 윤현선, 호지수, 부석준, and 김혜정, "Zero-day 피싱 웹사이트 탐지를 위한 컨볼루션 오토인코더 기반 문자수준 URL 모형," in 한국정보과학회 학술발표논문집, 2019, pp. 788-790.
- [6] H.-S. Kim and M.-H. Kim, "Hybrid phishing site detection system with GRU-based shortened URL determination technique," Journal of IKEEE, vol. 27, no. 3, pp. 213-219, Sep. 2023.

# 내부자별 개별학습 기반 내부자 위협 탐지

김유준\*, 김영갑\*\*

## Insider Threat Detection Based on Insider-Specific Individualized Learning

Yujun Kim\*, and Young-Gab Kim\*\*

### 요 약

악의적인 내부자는 조직 내 보안 정책에 대한 지식과 자원에 대한 인가된 권한을 가지고 있어 탐지하기 어려울 뿐만 아니라 심각한 경제적 피해를 초래할 수 있다. 악의적인 내부자를 탐지하기 위해, 최근에는 실제 환경에서 수집하기 어려운 비정상 행동 데이터를 제외한 정상 행동 데이터만을 딥러닝 모델이 학습하는 재구성 기반 이상 탐지 방법이 활용되고 있다. 다만, 기존 재구성 기반 이상 탐지 방법은 내부자의 직위나 직무 등을 고려하지 않고, 모든 내부자의 정상 행동을 동일한 딥러닝 모델 파라미터로 학습하기 때문에, 특정 내부자의 비정상 행동이 정상 행동으로 추론될 수 있다. 이러한 문제를 해결하기 위해, 우리는 내부자마다 서로 다른 비정상 행동을 탐지할 수 있는 내부자별 개별학습 기반 내부자 위협 탐지 방법을 제안한다. 우리는 CERT r4.2 데이터셋을 사용해 제안한 방법의 성능을 측정하였으며, 오토인코더와 제안한 방법을 비교했을 때 제안한 방법이 오토인코더와 데이터당 동일한 실행 시간을 가지면서도 오토인코더보다 6.73%나 더 높은 93.68%의 AUC를 달성하였다.

### Abstract

Malicious insiders with knowledge of internal security policies and authorized access to organizational resources are challenging to detect and potentially cause significant economic damage. To detect these insiders, recent studies have utilized reconstruction-based anomaly detection methods, where deep learning models train only normal behavior data, excluding abnormal behavior data that is difficult to collect in real-world environments. However, existing reconstruction-based anomaly detection methods do not consider factors such as insiders' position or role, training all normal behaviors using the same model parameters. This can result in abnormal behaviors of specific insiders being inferred as normal. To address this, we propose an insider threat detection method based on insider-specific individualized learning that can detect distinct abnormal behaviors for each insider. We evaluated the proposed method's performance using the CERT r4.2 dataset. Compared to an autoencoder, our method achieved an AUC of 93.68%, which is 6.73% higher than that of the autoencoder, while maintaining the same running time per data.

### Key words

insider threat, anomaly detection, deep Learning

---

\* 세종대학교 정보보호학과, awakening95@sju.ac.kr

\*\* 세종대학교 정보보호학과, alwaysgabi@sejong.ac.kr(교신저자)

※ 이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구 결과임 (No.RS-2019-II190231, 공공 인프라 안전을 위한 인공지능 기반 영상보안 기술 및 시스템 개발)

## 1. 서 론

악의적인 내부자는 조직의 네트워크, 시스템, 데이터에 대한 접근 권한을 악용해 조직 내 정보에 대한 기밀성, 무결성, 가용성에 피해를 줄 수 있다. 또한, 악의적인 내부자는 외부자와 달리 조직의 보안 정책에 대해 잘 알고 있기 때문에 내부자 위협 탐지를 회피할 수 있다. Securonix의 2024 Insider Threat Report[1]에 따르면, 내부자 위협을 보고하는 조직이 2019년 66%에서 2024년 76%로 증가하였다. 또한, 66%의 조직이 내부자 위협에 대해 취약하다고 느끼고 있지만, 단 41%만이 내부자 위협 탐지 프로그램을 부분적으로 구현한 상태였다. 따라서, 내부자 위협을 효과적으로 탐지하기 위한 연구가 필요한 상황이다.

최근에는 내부자 위협을 탐지하기 위해 딥러닝 모델을 사용한 재구성 기반 이상 탐지 방법이 활용되고 있다. 해당 방법은 실제 환경에서 수집하기 어려운 비정상 행동 데이터를 제외한 정상 행동 데이터만을 딥러닝 모델이 재구성하도록 학습한다. 학습이 완료된 모델은 정상 행동 데이터를 비정상 행동 데이터보다 더 정확하게 재구성할 수 있게 된다. 따라서, 재구성 기반 이상 탐지 방법은 정상 행동 데이터와 비정상 행동 데이터 간의 재구성 차이를 바탕으로 내부자 위협을 탐지한다. 다만, 기존 재구성 기반 이상 탐지 방법은 모든 내부자의 정상 행동 데이터를 동일한 딥러닝 모델의 파라미터를 사용해 학습한다. 이로 인해, 내부자의 직위, 직무 등에 따라 정상 또는 비정상이 될 수 있는 행동을 딥러닝 모델

이 정상으로 학습할 경우, 해당 행동은 정상으로 추론될 가능성이 높다는 문제가 있다.

이러한 문제를 해결하기 위해, 본 논문은 내부자의 특성에 따라 다른 비정상 행동을 탐지하는 내부자별 개별학습 기반 내부자 위협 탐지 방법을 제안한다.

## II. 내부자별 개별학습 기반 내부자 위협 탐지

제안한 내부자별 개별학습 기반 내부자 위협 탐지 방법은 그림 1과 같이 먼저 이메일 로그, 장치 로그, 파일 로그 등 다양한 내부자 행동 데이터를 수집한다. 이후, 각 내부자의 로그인부터 로그아웃까지의 행동 데이터를 하나의 세션 데이터로 구성한다. 생성된 세션 데이터는 내부자별 개별학습 기반 내부자 위협 탐지 모델에 입력되며, 해당 모델은 사전 학습된 인코더와 각 내부자에게 할당된 디코더를 통해 세션 데이터를 재구성한다. 마지막으로, 식(1)과 같이 실제 세션 데이터  $x_i$ 와 재구성된 세션 데이터  $\hat{x}_i$  간의 평균 제곱 오차를 이상 점수로 사용하여 이상 여부를 추론한다.  $i$ 는 세션 데이터의  $i$ 번째 차원을 의미한다.

$$\text{이상 점수} = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (1)$$

제안한 내부자별 개별학습 기반 내부자 위협 탐지 모델은 세션 데이터를 사용하여 2단계로 구성된

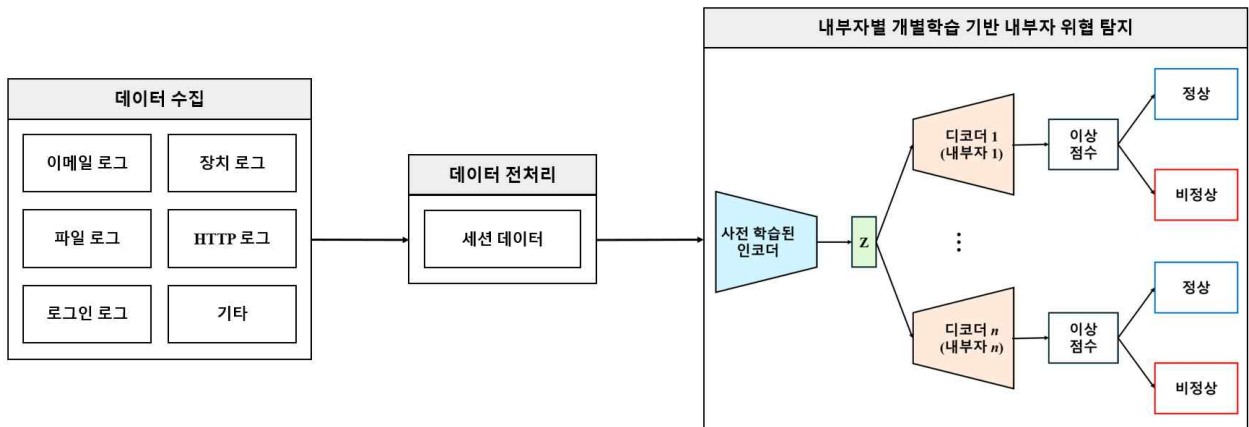


그림 1. 내부자별 개별학습 기반 내부자 위협 탐지 방법 개요

Fig. 1. Overview of the insider threat detection method based on insider-specific individualized learning



학습을 수행한다. 1단계 학습에서는 모든 내부자의 정상 데이터를 재구성하도록 오토인코더를 학습한 후, 내부자의 정상 데이터에 대한 잠재 표현을 출력하는 오토인코더의 인코더를 획득한다. 오토인코더는 식(1)과 같이 실제 세션 데이터와 재구성된 세션 데이터 간의 평균 제곱 오차를 손실 함수로 사용한다.

2단계 학습에서는 1단계 학습에서 사전 학습된 인코더를 각 사용자에게 할당된 디코더와 연결하고 각 디코더를 학습한다. 각 디코더의 구조는 1단계 학습에서 사용된 오토인코더의 디코더와 동일하다. 사전 학습된 인코더와 연결된 각 사용자의 디코더는 해당 사용자의 정상 데이터를 재구성하도록 학습된다. 이때 사용자별 디코더만 학습하고, 사전 학습된 인코더는 학습하지 않는다. 각 사용자의 디코더는 해당 사용자의 정상 데이터를 재구성하도록 학습되어, 해당 사용자의 비정상 데이터를 민감하게 탐지할 수 있게 된다. 2단계 학습 모델은 1단계 학습 모델과 동일한 평균 제곱 오차를 손실 함수로 사용한다.

본 논문에서 내부자마다 서로 다른 비정상 행동을 탐지하기 위해 내부자별로 오토인코더를 사용하지 않은 이유는 다음과 같다: 1) 학습 데이터가 부족한 내부자의 오토인코더는 학습이 어려워 비정상 행동 탐지가 어려워진다. 2) 각 오토인코더는 서로 다른 정상 데이터로 학습되므로, 각 사용자의 비정상 데이터에 대한 이상 점수 기준이 크게 달라질 수 있다. 예를 들어, 한 내부자의 비정상 데이터에 대한 이상 점수가 다른 사용자의 정상 데이터에 대한 이상 점수보다 낮아 이상 탐지를 실패할 수 있다. 결과적으로, 모든 내부자에 대한 학습을 가능하게 하고 내부자 간의 비정상 데이터에 대한 이상 점수 기준이 유사하도록 전체 내부자의 정상 데이터에 대해 사전 학습된 인코더를 사용한다.

### III. 실험 및 결과

#### 3.1 실험 데이터

본 논문에서 제안한 방법의 성능을 평가하기 위

해 기존 내부자 위협 탐지 연구에 많이 사용되는 CERT r4.2 데이터셋[2]을 사용한다. CERT r4.2는 약 17개월 동안 1,000명의 사용자로부터 수집된 로그를 포함하고 있으며, 405,380개의 장치 로그, 2,629,979개의 이메일 로그, 445,581개의 파일 로그, 28,434,423개의 HTTP 로그, 854,859개의 로그인 로그로 구성된 총 32,770,222개의 로그를 포함하고 있다.

CERT r4.2는 표 1과 같이 세 가지 내부 위협 시나리오를 포함하고 있으며, 시나리오 1은 30명의 사용자로부터 345개의 로그, 시나리오 2는 30명의 사용자로부터 6,765개의 로그, 시나리오 3은 10명의 사용자로부터 213개의 로그를 포함하고 있다.

표 1. CERT r4.2의 내부자 위협 시나리오  
Table 1. Insider threat scenarios of CERT r4.2

시나리오	내용
1	이동식 드라이브를 사용하거나 근무 시간 외에 일하지 않았던 사용자가 이동식 드라이브를 사용하고 근무 시간 외에 로그인하기 시작하며, 데이터를 wikileaks.org에 업로드한다. 그 후 얼마 지나지 않아 사용자는 조직을 떠난다.
2	사용자가 구직 사이트를 탐색하고 경쟁사에 고용을 요청하기 시작한다. 사용자는 조직을 떠나기 전에 이전보다 훨씬 빈번하게 이동식 드라이브를 사용하여 데이터를 탈취한다.
3	불만을 가진 시스템 관리자가 키로거를 다운로드하고 이동식 드라이브를 사용해 상사의 컴퓨터로 전송한다. 다음 날, 수집된 키로거를 사용하여 상사 계정으로 로그인하고 조직에 대량의 이메일을 발송하여 혼란을 유발한다. 이후 시스템 관리자는 즉시 조직을 떠난다.

#### 3.2 실험 방법

본 실험에서는 제안한 방법의 내부자 위협 탐지 정확도를 측정하기 위해 AUC(area under the curve)를 측정하였다. AUC는 ROC(receiver operating characteristic) 곡선의 아래 면적이며, 제안한 방법에서 ROC는 비정상 행동에 대한 이상 점수의 임계값을 변경함으로써 획득할 수 있다. 우리는 제안한 방법의 1단계 학습에 사용된 오토인코더와 2단계 학

습에 사용된 제안한 모델의 AUC와 CPU 기준 데이터당 실행 시간을 비교하였다. CPU 기준 데이터당 실행 시간은 AMD EPYC 7452 32-Core Processor를 사용하여 측정되었다.

본 실험에서는 CERT r4.2 데이터셋에서 사용자의 로그인부터 로그아웃까지 발생한 로그 데이터를 Le 외[3]가 제안한 데이터 전처리 방법을 통해 하나의 세션 데이터로 생성하였으며, 이를 6:2:2 비율로 나누어 훈련, 검증, 테스트 데이터로 사용하였다. 제안한 방법의 단계별 학습에는 동일한 훈련 및 검증 데이터를 사용하며, 단계마다 300 에포크로 훈련을 진행하였다. 본 실험에서는 1단계 학습에서 가장 낮은 검증 손실을 기록한 에포크의 인코더를 사용해 2단계 학습을 수행하였다. 각 단계에서 사용된 인코더와 디코더는 3개의 선형 계층으로 구성되어 있으며, 인코더는 세션 데이터의 차원에서 128, 64, 32차원으로 차례로 축소하고, 디코더는 이를 반대로 복원하여 원본 차원으로 복구한다.

### 3.3 실험 결과

표 2은 제안한 방법의 1단계 학습에 사용된 오토 인코더와 2단계 학습에서 사용된 제안한 모델이 가장 낮은 검증 손실을 달성했을 때의 AUC 값과 CPU 기준 데이터당 실행 시간을 보여준다. 실험 결과는 제안한 모델이 오토인코더와 비교해 동일한 데이터당 실행 시간을 가지면서도 6.73%나 더 높은 93.68%의 AUC를 달성하였다. 제안한 모델과 오토 인코더의 데이터당 실행 시간이 동일한 이유는 내부자의 세션 데이터에 대한 이상 점수를 측정할 때 사용되는 두 모델의 파라미터 수가 동일하기 때문이다.

표 2. 제안한 모델과 오토인코더의 AUC 및 CPU 기준 데이터당 실행 시간 비교

Table 2. The comparison of AUC and CPU-based running time per data between the proposed model and autoencoder

모델	AUC	Running Time (ms)
오토인코더	86.95	0.16
제안한 모델	93.68	0.16

그림 2는 단계별 학습에 사용된 모델의 에포크별 AUC 값을 보여준다. 그림 2에서 제안한 모델은 14 에포크부터 오토인코더가 224 에포크에서 달성한 가장 높은 AUC인 88.09%보다 높은 AUC를 달성하였다. 결과적으로, 실험 결과는 내부자마다 서로 다른 비정상 행동을 탐지하기 위해 제안한 방법이 데이터당 실행 시간 변화 없이 내부자 위협 탐지 정확도를 향상시켜준다는 것을 보여준다.

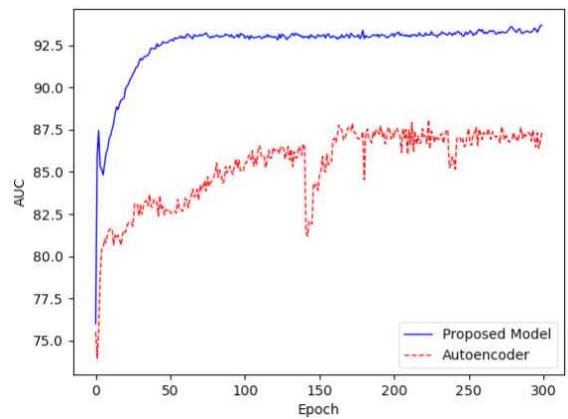


그림 2. 제안한 모델 및 오토인코더의 에포크별 AUC  
Fig. 2. AUC per epoch of the proposed model and autoencoder

## IV. 결 론

본 논문에서는 내부자의 직위와 직무 등에 따라 내부자별로 다른 비정상 행동을 탐지할 수 있는 내부자별 개별학습 기반 내부자 위협 탐지 방법을 제안하였다. 실험 결과, 제안한 방법은 오토인코더 기반 내부자 위협 탐지 방법과 동일한 데이터당 실행 시간을 유지하면서도 6.73% 더 높은 93.68%의 AUC를 달성하여 제안한 방법의 우수성을 보여준다. 향후 연구에서는 내부자별 비정상 행동을 효과적으로 탐지하기 위한 데이터 전처리 방법을 연구해 내부자 위협 탐지 성능을 향상시킬 것이다.

## 참 고 문 헌

[1] “2024 Insider Threat Report”, Securonix, Accessed: Oct. 28, 2024. [Online]. Available:

<https://www.securonix.com/resources/2024-insider-threat-report>

- [2] B. Lindauer, “Insider Threat Test Dataset”, Carnegie Mellon University, Accessed: Oct. 28, 2024. [Online]. Available: <https://doi.org/10.1184/R1/12841247.v1>
- [3] D. C. Le, N. Zincir-Heywood, and M. I. Heywood, “Analyzing data granularity levels for insider threat detection using machine learning”, IEEE Transactions on Network and Service Management, vol. 17, no. 1, pp. 30-44, January 2020.

# 형사사건 수사기록 비식별화를 위한 LLaMA3 기반 개인정보 태깅 및 비식별화 처리 연구

최주현\*, 배소은\*, 김령현\*, 김태연\*, 김경종\*\*

## A Study on LLaMA3-Based Personal Data Tagging and De-identification for Criminal Investigation Records

Juhyun Choi\*, Soeun Bae\*, Ryeonghyeon Kim\*, Taeyeon Kim\*, and Kyungjong Kim\*\*

### 요 약

본 연구에서는 기존 형사사법 분야 비식별화 기술의 수작업 의존성과 처리 속도 한계로 발생하는 문제를 해결하고자, 형사사건 수사기록 내 민감한 개인정보를 자동으로 식별하고 비식별화하는 프로세스를 제안한다. 파인튜닝된 LLaMA3 모델을 통해 이름, 주민등록번호, 주소 등의 개인정보를 정밀하게 태깅하고 비식별화 규칙에 따라 변환함으로써, 개인정보 보호의 높은 성능과 처리 효율을 달성하였다.

### Abstract

This study proposes an automated approach to identify and de-identify sensitive personal information in criminal investigation records, addressing the limitations of manual effort and processing speed in current de-identification methods within the criminal justice system. By fine-tuning the LLaMA3 model, the process accurately tags and de-identifies personal information such as names, resident registration numbers, and addresses according to established de-identification rules, enhancing both data protection and processing efficiency

### Key words

Criminal Justice, De-identification, Automatic Tagging, LLaMA3 Model, Fine-tuning

## 1. 서 론

데이터의 가치와 중요성이 증가함에 따라 데이터 공개에 대한 인식이 확산되고 있으며[1], 이러한 상황 속에서 형사사법을 포함한 다양한 분야에서 개인정보 보호를 위한 비식별화 기술의 필요성이 강

조되고 있다. 기존 연구에서는 형사사법 분야의 비식별화 기준을 재정립하고, 패턴 및 사전 기반의 알고리즘을 활용하여 수사기록 내 이름, 주민등록번호, 주소 등의 개인정보를 비식별화하는 기술을 개발하였다. 그러나, 이 알고리즘은 피고인 및 피해자와 같은 주요 인물의 이름을 자동으로 추출하지 못

\*경찰대학 치안데이터과학연구소, wisdompol@police.go.kr, \*\*치안데이터과학연구소, leeyeongul@police.go.kr(교신)

※ 본 연구는 2024년도 정부(경찰청)의 재원으로 과학치안진흥센터의 지원을 받아 수행된 연구임.

(No.PR10-02-000-24, AI 기반 범죄수사지원)

하여 수작업으로 알파벳을 대체해야 하고, 띄어쓰기 기반의 토큰화 및 대용량 사전 사용으로 인해 처리 속도가 지연되는 한계를 가지고 있다[2].

본 연구에서는 이러한 한계를 극복하고자 파인튜닝한 LLaMA3 모델을 이용하여 형사사건 수사기록 내 개인정보를 자동으로 태깅하고 비식별화하는 프로세스를 제안한다. 연구에서는 LLaMA3 모델의 학습 데이터 셋 구축 및 파인튜닝 과정에서 사용된 기술적 접근법을 상세히 설명하고, 다양한 개인정보 항목을 정확하게 식별하고 태깅하는 모델의 성능을 평가한다. 아울러, 제안된 비식별화 처리 방식이 기존 방식 대비 효율성 측면에서 얼마나 개선되었는지를 검토함으로써 형사사법 분야에서의 실질적인 적용 가능성을 논의하고자 한다.

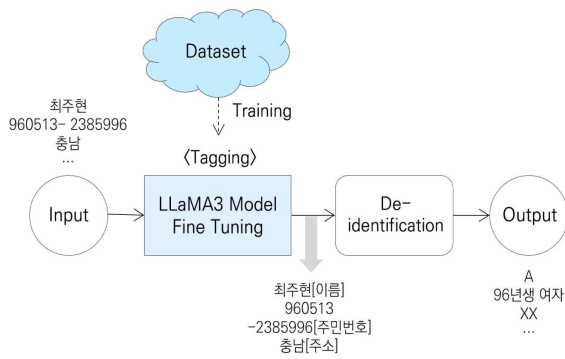


그림 1. 개인정보 태깅 및 비식별화 처리 프로세스  
Fig. 1. Personal Information Tagging and De-identification Processing Process

## II. 데이터 셋 구축

데이터셋은 연구자의 실무 경험을 바탕으로 수사 기록에서 발생할 수 있는 여러 시나리오를 반영한 가상의 데이터를 중심으로 구성되었으며, 실제 현장에서 발생할 수 있는 비식별화 상황과 표현을 보다 충실히 반영하기 위해 GPT 모델을 통해 다양한 형태의 텍스트 데이터를 생성하였다. 총 52,448개의 문장으로 구성된 데이터셋은 이름, 주민등록번호, 주소, 연락처, 건물명 등의 다양한 개인정보 유형을 포함하며, 실제 상황에 대한 적응력을 높이기 위해 비식별화가 필요하지 않은 문장도 일정 비율로 포함하였다.

표 1. Input 데이터 셋 예시  
Table 1. Input Dataset Example

구분	input 예시
개인정보 포함	저는 개인택시를 하고 있고 주소는 삼척시 근덕면 궁촌리 1914 소나무아파트 101호입니다.
개인정보 불포함	일을 마치고 퇴근을 했는데 같이 일하는 형이 제 휴대폰으로 전화해 만나자고 하였다.

또한 데이터는 입력-출력 쌍의 형식으로 구성되어 있으며, 입력에는 원문 텍스트가, 출력에는 비식별화가 필요한 항목에 대한 태깅 정보만을 포함하였다.

표 2. Output 데이터 예시  
Table 2. Output Dataset Example

유형	output 예시
이름	최주현[이름], 배소은[이름], 주현[이름], 경종[이름] 등
주민번호	960513-1234567[주민번호]
주소	충남[주소], 아산시[주소], 신창면[주소], 100-50[주소]
연락처	010-1234-5678[연락처]
건물·건물명	한국[건물명], 경찰서[건물], 부재[건물명], 부동산[건물],

## III. Llama3 모델 파인튜닝

본 연구의 목표는 한국어 형사사건 수사기록에서 발생하는 민감한 개인정보를 효과적으로 식별하고 태깅하는 것이다. 이를 위해, 오픈 소스 대형 언어 모델인 LLaMA3 시리즈 중 한국어에 특화된 beomi/Llama-3-Open-Ko-8B<sup>1)</sup> 모델을 선택하여 파인튜닝을 하였다.[3]

학습 환경으로는 별도 설치 없이 GPU와 Python을 바로 사용할 수 있는 Google Colab을 활용하였다. Google Colab은 고성능 GPU와의 호환성을 제공하며, 대규모 학습에 필요한 인프라를 간편하게 사용할 수 있어 본 연구 환경으로 선정하였다.

파인튜닝 과정에서는 파라미터 효율적 파인튜닝 (PEFT) 방식인 LoRA(Low-Rank Adaptation)를 적용하여, 모델의 전체 파라미터를 학습하지 않고 일부 주요 파라미터만 학습하도록 하였다. LoRA는 기존

1) <https://huggingface.co/beomi/Llama-3-Open-Ko-8B>

모델의 네트워크 구조를 유지하면서도 효율적인 파인튜닝을 가능하게 하며, 메모리와 학습 비용을 절감할 수 있는 기법이다.[4]

또한, 4비트 양자화를 통해 모델의 메모리 사용량을 줄여 대규모 데이터셋 학습을 보다 효율적으로 수행할 수 있도록 하였다. 이러한 최적화 기법을 통해 메모리와 연산 비용을 줄이면서도 개인정보 추출 성능을 극대화할 수 있었다. [그림 2]는 해당 방법을 사용하여 작성한 파인튜닝 모델 코드이다.

**Step 1: Initialize Trainer**

```
SFTTrainer(
    model = model,
    tokenizer = tokenizer,
    train_dataset = dataset,
    dataset_text_field = "text",
    max_seq_length = max_seq_length,
    dataset_num_proc = 2,
    packing = False
)
```

**Step 2: Set Training Arguments**

```
TrainingArguments(
    per_device_train_batch_size = 4,
    gradient_accumulation_steps = 8,
    warmup_steps = 10,
    num_train_epochs = 1,
    max_steps = 150,
    learning_rate = 2e-4,
    fp16 = not is_bfloat16_supported(),
    bf16 = is_bfloat16_supported(),
    logging_steps = 10,
    optim = "adamw_8bit",
    weight_decay = 0.1,
    lr_scheduler_type = "cosine",
    seed = 123,
    output_dir = "outputs"
)
```

**Step 3: Execute Training**

```
trainer.train()
```

그림 2. 파인튜닝 모델 학습 관련 pseudo 코드  
Fig. 2. Pseudocode for Fine-tuning Model Training

파인튜닝 전후의 LLaMA3 모델은 개인정보 추출과 태깅의 정확도에서 큰 차이를 보였다. 예를 들어, 입력 문장 "최주현(960513-2123454)은 010-1234-5678 번으로 대출광고 문자가 발송되어 배소는 상담사와 대출 상담을 하고, 전북 전주시 완산구 경찰택배에서 아산시 신창면 100-50 대학택배로 우리은행 체크카드 1매를 전달했다."를 테스트한 결과, 파인튜닝 이전의 PLM 모델은 문장을 일반 텍스트로 그대로 출력했으나, 파인튜닝된 PEFT 모델은 "최주현[이름],

960513-2123454[주민번호], 010-1234-5678[연락처], 배소은[이름], 전북[주소], 전주시[주소], 완산구[주소], 경찰택배[건물명], 아산시[주소], 신창면[주소], 100-50 [주소], 대학택배[건물명]"와 같이 민감 정보를 정확히 태깅하여 출력하는 것을 확인하였다. 이와 같이, 파인튜닝된 LLaMA3 모델이 개인정보 추출 성능을 크게 향상시키는 데 효과적임을 확인할 수 있었다.

**Step 1: Initialize Language Model for Inference**

```
Load model with FastLanguageModel.for_inference(model)
```

**Step 2: Prepare Input Data**

```
Format input text using alpaca_prompt with specified text
Tokenize formatted text using tokenizer
Set return type to tensors (pt) and move to CUDA (GPU)
```

**Step 3: Set Up Text Streamer**

```
Initialize TextStreamer with tokenizer for output streaming
```

**Step 4: Generate Output**

```
Use model.generate with parameters:
- tokenized inputs
- text streamer for real-time output display
- max_new_tokens set to 128
```

그림 3. 파인튜닝 모델 테스트 관련 pseudo 코드  
Fig. 3. Pseudocode for Fine-tuning Model Testing

**IV. 비식별화 처리 및 텍스트 생성**

본 연구에서는 파인튜닝된 LLaMA3 모델이 제공하는 태깅 결과와 python 코딩을 통해 개인정보가 비식별 처리된 텍스트를 생성하였다. 모델이 텍스트 내에서 이름, 주민번호, 주소, 연락처, 건물명으로 식별된 개인정보를 태깅하여 출력하면, 비식별 텍스트 생성 과정에서 이름의 경우, 태깅된 순서대로 알파벳이 자동으로 순차 부여된다. 주민번호는 ‘6자리 숫자-7자리 숫자’ 형식으로 구성되어 있을 시, 앞 6 자리에서는 처음 두 자리를 가져오고, 뒷 7자리에서는 첫 번째 숫자가 1, 3, 5인 경우 ‘남자’, 2, 4, 6인 경우 ‘여자’로 부여하여 ‘xx년생 남자 또는 여자’의 형태로 비식별화된다. 연락처의 경우, 비식별화 처리를 위해 새로운 연락처를 생성할 시 실제 연락처와 동일한 연락처가 생성될 가능성을 고려할 필요가 있다. 통신사별 원배정 국번에 따르면, 중간번호는 반드시 원배정 국번에 따라 2000부터 9999 사이의 숫자로 생성되어 랜덤하게 개통된다. 이에 통신사 미배정 국번인 010-0000-0000부터 010-1999-9999 범위 내에서 난수값을 생성하여 임의로 부여함으로써 실제 연락처와의 중복 가능성을 최소화하고자 한다.

이전 연구에서는 주민등록번호와 연락처를 일괄적으로 숫자를 '#'으로 대체하여 원본 데이터의 정보를 모두 소거하는 방식을 사용했으나, 본 연구에서는 데이터의 유용성을 유지하기 위해 일부 원본 형식을 남겨두는 방식으로 처리한다. 그 외 주소, 건물명의 경우, 원본 글자수만큼 'x'로 대체하여 처리한다.

```

foreach (value, tag) in matches do
  if tag == "Name" then
    if value not in name_mapping then
      Assign a new alphabet letter to name_mapping[value] based
      on name_count;
      Increment name_count;
    end
    Map anonymized_items[value] to name_mapping[value];
  end
  else if tag == "Address" then
    Replace anonymized_items[value] with "x" repeated for the
    length of value;
  end
  else if tag == "Contact" then
    Generate a random phone number in the format
    '010-xxxx-xxxx';
    Map anonymized_items[value] to this random number;
  end
  else if tag == "Resident Registration Number" then
    Split value into parts at '-';
    if format_matches then
      Extract year and gender digit;
      Determine gender based on gender digit;
      Map anonymized_items[value] to formatted string with birth
      year and gender;
    end
    else
      Replace anonymized_items[value] with
      "XXXXXX-XXXXXX";
    end
  end
  else if tag in ["Building", "Building Name"] then
    Replace anonymized_items[value] with "x" repeated for the
    length of value;
  end
end
    
```

그림 4. 비식별화 처리 함수 관련 pseudo 코드  
Fig. 4. Pseudocode for De-identification Processing Function

앞서 제시한 예시에서, 모델은 "최주현[이름], 960513-2123454[주민번호], 010-1234-5678[연락처], 배소은[이름], 전북[주소], 전주시[주소], 완산구[주소], 경찰택배[건물명], 아산시[주소], 신창면[주소], 100-50[주소], 대학택배[건물명]"과 같이 입력 문장

에 포함된 개인정보를 정확히 식별하고 태깅하는 성능을 보였다. 이를 바탕으로 비식별화 코드를 실행한 결과, 'A(96년생 여자)은 010-1952-2625번으로 대출광고 문자가 발송되어 B 상담사와 대출 상담을 하고, xx xxx xxx xxxx에서 xxx xxx xxxxxx xxxx로 우리은행 체크카드 1매를 전달했다.'와 같이 텍스트가 생성되어 다양한 형태의 개인정보가 정확하게 비식별화되었음을 확인할 수 있었다.

```

### 지시사항:
최주현(960513-2123454)은 010-1234-5678번으로 대출광고 문자가 발송되어
배소은 상담사와 대출 상담을 하고, 전북 전주시 완산구 경찰택배에서 아산시
신창면 100-50 대학택배로 우리은행 체크카드 1매를 전달했다.

### 응답:
최주현[이름], 960513-2123454[주민번호], 010-1234-5678[연락처], 배소은
[이름], 전북[주소], 전주시[주소], 완산구[주소], 경찰택배[주소], 아산시
[주소], 신창면[주소], 100-50[주소], 대학[건물명], 택배[건물]
비식별화된 텍스트: A(96년생 여자)은 010-1952-2625번으로 대출광고 문자가
발송되어 B 상담사와 대출 상담을 하고, xx xxx xxx xxxx에서 xxx xxx xxx
xx xxxx로 우리은행 체크카드 1매를 전달했다.
    
```

그림 5. 비식별화 처리 결과  
Fig. 5. De-identification Processing Result

## V. 프로세스 평가 및 한계점

본 연구에서 개발한 비식별화 태깅 모델은 형사 사건 수사기록 내 민감한 개인정보를 자동으로 식별하고 태깅하는 데 있어 높은 성능을 보였다. 모델은 이름, 주민등록번호, 연락처, 주소, 건물명 등 다양한 개인정보 항목을 정확히 태깅하여 비식별화된 텍스트를 생성하는 데 효과적이었다. 또한, 동일한 데이터 패턴이 반복될 경우 일관성 있게 비식별화 태깅을 수행하여 안정적이고 신뢰성 있는 결과를 제공하였다.

특히, 기존 연구에서는 사람 이름을 일일이 추출하여 알파벳으로 태깅하는 과정이 수작업으로 이루어졌으나, 본 연구에서는 파인튜닝된 모델을 통해 이름, 주민등록번호, 주소 등 주요 개인정보 항목을 자동으로 추출하고 알파벳으로 태깅할 수 있게 하였다. 이를 통해 수작업을 최소화하면서 개인정보 태깅 작업의 효율성과 정확성을 크게 향상시킬 수 있었다.

그러나 몇 가지 한계점도 발견되었다. 첫째, 비식별화할 항목의 표현이 모호하거나 맥락에 따라 다양한 해석이 가능한 경우, 모델이 정확한 태깅을 수

행하지 못할 가능성이 있었다. 예를 들어, "문경이  
가 문경에서 밥을 먹었다"와 같은 문장에서는 "문  
경"이 인명과 지명으로 모두 사용되어 혼동이 발생  
할 수 있으며, "사고 경위"와 "경위로 근무하는 경  
찰관"과 같이 동일한 단어가 서로 다른 의미로 쓰  
일 때 모델이 이를 구분하지 못해 태깅 오류가 발  
생할 가능성이 있었다. 이러한 다의어의 맥락적 처리  
문제는 모델의 비식별화 정확성에 영향을 미치는  
요인 중 하나로 작용하였다.

또한, 학습된 데이터셋의 제한으로 인해 일부 상  
황에서 사물이나 소프트웨어 명칭을 사람 이름으로  
오인할 가능성이 있었다. 예를 들어, "LLM"과 같은  
소프트웨어 명칭이 이름으로 잘못 인식되는 경우가  
발생할 수 있었다. 이는 학습 데이터에서 사물이 주  
어로 등장하는 경우를 충분히 반영하지 못한 데 기  
인한 한계로, 향후 이러한 사례를 반영한 추가 데이  
터 확보가 필요하다.

아울러, 본 연구에서 사용된 학습 데이터셋이 단  
문 위주로 구성되어 있어 여러 문장으로 이루어진  
긴 텍스트가 입력될 경우 모델의 정확도가 다소 떨  
어지는 문제가 있었다. 이는 복잡한 구조의 실제 수  
사기록을 처리할 때 태깅 정확성에 영향을 미칠 수  
있는 요소로 작용할 수 있으며, 추후 연구에서는 긴  
텍스트나 문단 단위의 데이터를 학습에 포함시켜  
이러한 문제를 해결할 필요가 있다.

향후 연구에서는 이러한 평가와 한계점을 바탕으  
로 실제 수사기록 데이터를 더욱 다양하게 반영하  
여 학습 데이터의 범위를 확장하고, 사물과 사람을  
보다 명확히 구분할 수 있는 알고리즘을 보완할 계  
획이다. 또한, 복잡한 고유명사와 다양한 맥락을 더  
정확하게 처리하기 위해 추가적인 개체명 인식 모  
델의 도입 및 개선을 검토할 예정이다. 이를 통해  
형사사건 수사기록의 비식별화 성능이 한층 향상되  
기를 기대한다.

### Acknowledgement

본 연구는 2024년 정부(경찰청)의 재원으로 과학  
치안지능센터의 지원을 받아 수행된 연구임.  
(No.PR10-02-000-24, AI 기반 범죄수사지원.)

### 참 고 문 헌

- [1] M. R. Kim, Y. J. Roh, and S. H. Kim, "A study on the use of criminal justice information big data in terms of the structuralization and categorization," *Journal of the Korean Society for Information Management*, vol. 36, no. 4, 2019.
- [2] K. J. Kim, S. E. Bae, B. K. Woo, and J. O. Kim, "A study on the de-identification of criminal case investigation records," *Journal of Digital Forensics*, vol. 16, no. 3, 2022.
- [3] J. Myung, J. Park, J. Son, K. Lee, and J. Han, "Efficient technical term translation: A knowledge distillation approach for parenthetical terminology translation," *arXiv preprint, arXiv:2410.00683*, 2024.
- [4] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, and W. Chen, "LoRA: Low-rank adaptation of large language models," *arXiv preprint, arXiv:2106.09685*, 2021.



# ResNet 기반 철도차량 주요 부품 분류 및 경량화 연구

이호정\*, 이수민\*, 곽대영\*, 서영주\*, 김동주\*\*

## Development and Compression of ResNet-based Classification Model for Key Components of Railway Vehicles

Ho Jung Lee\*, Soo Min Lee\*, Dae Young Kwak\*, Young Joo Suh\*, Dongju Kim\*\*

본 논문에서는 인공지능 융합 고속차량 특화 서비스를 제공하기 위해 철도차량의 대차 부품 및 경·중정비 주요 부품을 분류하는 모델을 개발하고 이를 경량화하는 연구를 소개한다. 전통적인 철도차량 정비 과정에서는 부품의 재고 확인 및 공급 문제로 인해 정비작업이 비효율적으로 운용되어왔다. 이러한 문제들을 해결하기 위해 최근 해외 고속차량 운영에서는 인공지능을 활용한 디지털 방식의 운영 시스템 도입이 활발히 이루어지고 있다. 예로, 독일에서는 AI 기반 실시간 운영 상황 분석을 통해 열차 지연을 최소화하고 있으며, 일본에서는 제철 작업에 필요한 유지보수 인력을 예측하는 데 AI를 활용하고 있다[1]. 본 연구에서는 철도차량 경·중정비동에서 직접 수집한 이미지를 데이터셋으로 활용하여 분류 모델을 개발했다. 대차 부품 데이터셋은 대차에 부착되어 육안으로 측면에서 구별 가능한 25종의 부품에 대한 14,919장, 경·중정비 부품 데이터셋은 269종의 핵심 부품에 대한 31,278장의 이미지로 구성되었다. 실험에서는 잔차 블록을 활용해 깊은 층에서의 정보 손실을 최소화하여 특징이 잘 유지되며 뛰어난 분류 성능을 보인 ResNet 구조를 Baseline 모델로 선정하였다[2]. 이를 기반으로 ResNet-18과 ResNet-50 구조를 적용해 대차 부품 및 경·중정비 주요 부품 분류 모델을 개발한 결과, ResNet-18은 각각 0.9870, 0.9966의 정확도를, ResNet-50은 상대적으로 낮은 0.9850, 0.9888의 정확도를 보였다. 이에 정확도와 경량성 면에서 우수한 성능을 보인 ResNet-18 모델을 대상으로 현업 활용도를 높이기 위한 경량화를 진행하였다. 가장 높은 성능을 보인 ResNet-18 모델에 대해 70%의 Structured Pruning 및 Fine-tuning을 적용한 결과, 대차 부품 분류 모델에서의 정확도 손실은 0.44%에 그쳤으나, 파라미터 개수는 약 11.1893M개에서 1.0019M개로 10.12배, FLOPs는 9.73배 감소하였다. 경·중정비 부품 분류 모델에서도 정확도 손실은 0.16%에 불과했으며, 파라미터 개수는 약 11.3145M개에서 1.0103M개로 11.20배, FLOPs는 9.73배 감소하였다. 이로써 가중치 삭제와 미세조정을 통해 최적화된 ResNet-18 모델이 크기 및 추론 속도 면에서 유의미하게 경량화되었음을 확인하였다. 본 연구를 통해 개발된 철도차량 주요 부품 분류 모델이 현업에서 실질적으로 활용되어 정비 효율이 향상되고 차량 유지보수 비용 절감에 기여하기를 기대한다.

### Keywords

Railway Vehicles, Artificial Intelligence, Image Classification, Model Compression, Structured Pruning

### 참고문헌

- [1] UIC, "The journey toward AI-enabled railway companies" International Union of Railways & McKinsey & Company. R&D Report, 2024
- [2] K.He, X.Zhang, S.Ren and J.Sun, "Deep Residual Learning for Image Recognition", Proceedings of the IEEE International Conference on CVPR, pp. 770-778, 2016.

\* 포항공과대학교 인공지능연구원, {ghwj0112, hibokchi, brise60, yjsuh}@postech.ac.kr,

\*\* 포항공과대학교 인공지능연구원(교신저자), kkb0320@postech.ac.kr

※ 본 논문은 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업이며 (No.2022R1A6A1A03052954), 정부(과학기술정보통신부)의 재원으로 한국지능정보사회진흥원의 지원을 받아 수행된 연구이며(2024년 지역특화형 초거대 AI 클라우드 팜 실증 및 확산 환경 조성), 「과학기술정보통신부 방송통신발전기금」을 지원받아 정보통신산업진흥원(DI014-24-1009, 철도차량 가용성 향상을 위한 AI 융합형 통합정비 공정 메타버스 플랫폼 개발)의 지원을 받아 수행된 연구이나 과학기술정보통신부의 공식의견과 다를 수 있음.

# 해양 선박구조물의 인식을 위한 의미론적 분할 기술

이지은<sup>1</sup>, 임재준<sup>2</sup>, 차상빈<sup>3</sup>, 김민<sup>4</sup>, 박상홍<sup>5</sup>, 최인오<sup>6</sup>

## Semantic segmentation for machinery recognition on marine vessels

Jieun Lee<sup>1</sup>, Jaejun Lim<sup>2</sup>, Sangbin Cha<sup>3</sup>, Min Kim<sup>4</sup>, Sanghong Park<sup>5</sup>, and Inoh Choi<sup>6</sup>

### 요약

해상 물류 및 어군 탐지 분야에서 UAV(무인 항공기)를 활용한 자율화가 활발하게 진행되고 있으나 선박 구조물 인식 기술의 한계로 현재까지 완전한 자율화에는 어려움이 존재한다. 자율주행 자동차와 달리 해상 선박 환경에서는 로프, 윈치 등과 같은 복잡하고 검출이 어려운 객체들로 인해 UAV 사고가 빈번하게 발생한다. 본 논문에서는 문제를 해결하기 위해 선박 위험 구조물 인식을 목표로 의미론적 분할 기술을 적용시켜 정확도를 향상시키고 UAV의 해상에서의 자율화에 기여하고자 한다.

### Abstract

In the field of maritime logistics and fishing group detection, the use of UAVs (unmanned aerial vehicles) is advancing rapidly. However, achieving complete autonomy remains challenging due to the limitations of ship structure recognition technology. Unlike autonomous vehicles, UAVs frequently encounter accidents in marine environments because of complex and difficult-to-detect objects, such as ropes and winches. To address this issue, this paper aims to enhance accuracy and contribute to the autonomy of UAVs at sea by applying semantic segmentation technology to recognize hazardous structures on ships.

### Key words

Semantic segmentation, UAV Autonomy, Ship Structure Recognition, Object Detection, Autonomous Navigation

## 1. 서론

의미론적 분할(Semantic segmentation)은 이미지의 모든 픽셀들을 특정 클래스에 할당하는 기술로, 해상 이미지에서 객체를 분류하고 경계를 확실히 표

현하기에 유용하다[1]. 본 연구에서는 U-Net 기반 딥러닝 모델을 사용하여 해상에서의 선박구조물을 인식하는 방법을 제안하며, U-Net의 encoder-decoder 구조는 다양한 해상도에서 객체의 특징을 추출, 복원하는 것에 효과적이기 때문에 이를 활용하였다

<sup>1</sup>국립부경대학교 스마트모빌리티공학과, ylducls@pukyong.ac.kr, <sup>2</sup>휴몬랩, jjl1003@pukyong.ac.kr,

<sup>3</sup>국립부경대학교 전자공학과, teddy@pukyong.ac.kr, <sup>4</sup>한국해양과학기술원, skymin@kiost.ac.kr,

<sup>5</sup>국립부경대학교 전자공학과, radar@pknu.ac.kr, <sup>6</sup>국립부경대학교 스마트모빌리티공학과, inoh@pkun.ac.kr(교신저자)

※ 이 논문은 2023년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2022R1G1A1004830)

[2]. 학습된 모델은 해상 환경에서 촬영한 여러 종류의 선박 이미지로 성능을 평가하며 정확도와 IoU 지표로 검증하였다[3]. 본 연구로 해상이라는 특수 환경에서의 UAV의 객체 인식 및 분할 성능을 높이고 해상 물류와 어군 탐지 등의 자율성의 향상을 기대한다.

표 1. 데이터 세트  
Table 1. Data sets

Data sets	Images	Labels	Ratio
Training Data	180	180	90%
Test Data	10	10	5%
Validation Data	10	10	5%

## II. 딥러닝 기반 의미론적 분할

본 연구에서는 해상에서의 선박구조물을 인식 및 분할을 위해 U-Net 모델을 선정하였다. U-Net은 encoder-decoder 구조로 되어있으며, 입력한 이미지의 경계를 확실하게 표현할 수 있다. 또한 다양한 해상도에서 feature를 효과적으로 추출 및 복원할 수 있는 기능이 뛰어나서 해상 환경에서의 복잡한 객체를 구분하기에 적합한 모델이다 [1].

모델의 input layer에 입력변수인 이미지가 들어오면 hidden layer는 convolution과 pooling 연산을 통해 입력데이터로부터 값을 예측하기 위한 중요한 특성을 추출한 후 feature map을 생성한다. 다음의 식은 선정한 모델의 출력값  $y$ 를 정의한다.

$$\hat{y}_{i,j} = \text{softmax}(W \cdot x_{i,j} + b) \quad (1)$$

여기서  $\hat{y}_{i,j}$ 는  $(i,j)$  위치의 예측 클래스 확률,  $W$ 는 가중치,  $b$ 는 편향,  $x_{i,j}$ 는 입력 이미지의 픽셀 값을 나타낸다. 모델의 학습 과정에서 손실 함수는 다음과 같이 정의한다.

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c}) \quad (2)$$

여기서  $N$ 은 batch size,  $C$ 는 클래스 수,  $y_{i,c}$ 는 실제 클래스 레이블이다. 손실 함수의 최소화를 위해

adam optimizer를 사용하여 가중치  $W$ 를 계속 업데이트한다. Accuracy와 Intersection over Union (IoU) 지표를 사용하여 모델 성능을 평가하며, 실험 결과 U-Net이 기반인 모델이 다른 모델보다 높은 성능을 나타내는 것을 확인하였다 (그림 1). IoU 지표로 평가한 모델의 성능은 선박구조물과 같은 복잡한 객체를 분할하는 데 있어서 높은 성능을 보였다(표 2). 따라서 UAV 기반 해상 물류와 어군 탐지 등의 높은 안전성과 효율성을 높일 수 있을 것이다.

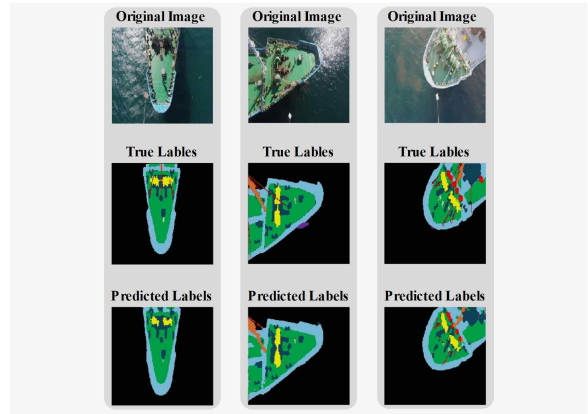


그림 1. 시멘틱 세그멘테이션 결과  
Fig. 1. Semantic segmentation result

표 2. 데이터셋 매트릭  
Table 2. Data sets Matrix

Data sets Matrix	results (%)
Global Accuracy	90.15
Mean Accuracy	59.06
mean IoU	50.35
Weighted IoU	83.63
Mean BF Score	67.58

## III. 결 론

본 논문에서는 선박구조물을 인식, 분할하고 UAV의 해상 자율운행을 위한 방법으로 U-Net 딥러닝 모델 기반의 의미론적 분할 기술을 제안하였다. U-Net 모델은 선박구조물과 같은 복잡한 객체의 경계를 확실히 분할할 수 있으며, 높은 정확도와 IoU 지표로 성능을 확인했다. 이는 UAV 자율주행 시스템의 해상에서의 드론 물류 이동 서비스, 선망 어선의 어군 탐지 등의 해상 시스템의 효율성과 안전성을 동시에 향상시킬 수 있을 것으로 예상된다.

참 고 문 헌

- [1] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. CVPR, 2015*, pp. 3431-3440.
- [2] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI, 2015*, pp. 234-241.
- [3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR, 2016*, pp. 770-778.

# 키포인트 기반 상호작용 자세 분석을 통한 폭력 상황 인식 및 실시간 감지 모델 개발

이수연\*, 김소희\*\*, 함종수\*\*\*

## Development of a Real-Time Violence Detection Model through Keypoint-Based Interaction Analysis

Suyeon Lee\*, Sohee Kim\*\*, and Jongsoo Ham\*\*\*

### 요 약

본 논문은 CCTV 영상에서 폭행을 효과적으로 감지하기 위해 신체 키포인트 분석과 LSTM 모델을 활용하는 연구를 다룬다. AlphaPose와 YOLO 모델을 통해 두 사람을 객체로 감지하고, 각 프레임에서 추출된 신체 키포인트를 분석하여 자세를 분류한다. 이후, 분류된 상호작용 자세를 LSTM 모델에 입력하고, 시간에 따른 자세 변화 데이터와 함께 학습하여 폭력적인 장면을 감지한다. 본 논문에서 상호작용을 분석한 데이터를 함께 학습한 모델과 그렇지 않은 모델의 성능 비교를 통해 폭행 장면을 탐지할 수 있는 효과적인 접근법을 제안한다.

### Abstract

This paper presents a study utilizing body keypoint analysis and an LSTM model to effectively detect assaults in CCTV footage. Using the AlphaPose and YOLO models, two individuals are detected as objects, and body keypoints are extracted from each frame to classify poses. Then, the LSTM model are trained with the classified interaction poses and extracted body keypoints to identify violent scenes. This paper proposes an effective approach for assault detection by comparing the performance of models that incorporate interaction analysis data with those that do not.

### Key words

Assault detection, Keypoint analysis, Pose classification, LSTM, Action recognition

---

\* 고려사이버대학교, syeun023@gmail.com

\*\* 한국방송통신대학교, somoon0422@gmail.com(공동 1저자)

\*\*\* 한양대학교 인공지능융합학과, hamjongsoo0724@gmail.com(교신저자)

## I. 서론

갑작스런 폭행이 발생할 경우 피해자가 즉각적으로 신고하기에는 어려움이 있다. 이와 같은 이유로 안전과 보안을 강화하기 위한 실시간 영상 감시 시스템은 현대 사회에서 점점 더 중요한 역할을 하고 있다. 이를 위해 본 연구에서는 시계열 모델인 Stacked LSTM과 신체 키포인트 추출 기술을 결합한 접근법을 활용하며, 두 사람의 키포인트를 토대로 상호작용을 분석하는 방법을 제안한다.

## II. 관련 연구

폭행과 같은 비정상적 행동을 감지하는 연구는 활발히 진행되고 있다. Sultani et al. (2018)은 비디오 클립을 기반으로 이상 행동을 탐지하는 기법을 제안했으며, 비디오 시퀀스를 클립 단위로 나누어 각 클립의 이상치 점수를 계산하여 비정상적 행동을 식별하고자 했다[1].

사람의 행동을 분석하기 위해 신체 키포인트 추출 기술이 널리 사용되고 있으며, 이는 비디오 시퀀스에서 시각적 정보를 추약하여 주요 특징을 추출하는 데 효과적이다. Fang et al. (2017)은 AlphaPose를 제안하여, 다중 사람 상황에서도 높은 정밀도로 키포인트를 추출할 수 있는 상향식 접근 방식을 설명하였다. 본 연구에서는 AlphaPose가 사람의 경계 상자를 기반으로 신체 자세를 추정하고, 복잡한 장면에서도 자세들 간의 관계를 파악하여 상호작용을 라벨링하는 방법을 차용하였다[2].

폭행 감지와 같은 비정상적 행동 검출에서는 재현율이 매우 중요한 지표로 여겨진다. 놓치는 폭행 상황이 발생할 경우 보안적 측면에서 큰 위험이 있기 때문에, 모든 폭행 상황을 빠짐없이 탐지하는 것이 필수적이다. 선행 연구에서도 재현율을 높이기 위해 다양한 딥러닝 기반 접근법과 위치 추정 전략을 사용하여 폭행과 같은 상황에서 높은 재현율을 달성하고자 했다(Mumtaz et al., 2022)[3]. 본 연구 역시 높은 재현율을 주요 평가 지표로 삼아 모델의 성능을 평가하고자 한다.

## III. 연구 문제 정의

본 연구의 목표는 실시간으로 두 사람 사이에서 발생하는 폭력적인 행동을 효과적으로 감지할 수 있는 모델을 개발하는 것이다. 기존 감시 시스템은 정적 규칙이나 단일 객체 기반 분석에 치우쳐 있어 다수의 객체가 포함된 복잡한 상호작용을 인식하는데 한계가 있었다. 본 연구에서는 AlphaPose를 통해 두 사람의 신체 키포인트를 추출하고 이를 토대로 두 사람의 상호작용을 분석함으로써 이러한 문제를 해결하고자 한다.

## IV. 연구 방법

본 연구의 목표는 실시간으로 폭력적인 행동이 일어나고 있는 상황을 정확하게 감지할 수 있는 모델을 개발하는 것이다. 이를 위해 AI 허브에서 제공하는 폭행 관련 이상행동 CCTV 영상 데이터를 활용하여[4], 데이터 전처리, 키포인트 추출 및 분석, 시계열 모델 학습의 단계를 거쳤다.

### 4.1 데이터 전처리

AI 허브에서 제공한 CCTV 영상을 사용하였으며, 각 비디오 파일에 폭력 상황이 발생한 시간 범위를 파악, 폭력 상황의 시작 및 종료 프레임을 추출하고, 이를 바탕으로 폭력(assault) 및 비폭력(normal) 프레임을 정의했다.

### 4.2 객체 탐지 및 키포인트 추출

YOLO 모델을 사용하여 각 프레임에서 사람을 탐지한 후, 사람의 경계 상자 추출 결과를 바탕으로 crop한 이미지를 AlphaPose 모델에 입력하여 신체의 주요 관절 키포인트를 추출한다. AlphaPose는 사람의 경계 상자를 기반으로 17개의 신체 관절(총 34개의 좌표)을 추출하며, 여러 사람이 있는 장면에서도 높은 정밀도를 제공한다.

AlphaPose는 주로 사람의 경계 상자를 기반으로 자세를 정제하는 상향식 접근 방식을 사용하여, 여

러 사람이 있는 복잡한 장면에서도 높은 정밀도를 제공한다. 또한, 비용 효율적이면서도 성능이 적절하여 본 연구에서 신체 키포인트를 추출할 모델로 선정하였다[5].

사람이 한 명도 추출되지 않은 경우에는 상호작용도 있을 수 없으므로 비폭력 상태로 판단하게끔 시스템을 구성하였다.

### 4.3 자세 분류 및 상호작용 정보 추출

AlphaPose로 추출된 키포인트를 사용하여 각 관절의 상대적 위치와 각도를 계산한 후 이를 토대로 각 프레임에서의 자세 및 상호작용을 분석하였다.

분석에는 관절 간의 각도를 계산하여 특징적인 현재 자세를 분류할 수 있도록 한다. 본 논문에서는 이러한 방법으로 *standing*, *walking*, *punching*, *fighting*, *falling* 자세 및 상호작용 정보를 추출한다.

$$|y_{left-wrist} - y_{nose}| < T_{punch} \quad (1)$$

$$|y_{right-wrist} - y_{nose}| < T_{punch} \quad (2)$$

예를 들어, 식 (1)과 (2)는 키포인트를 바탕으로 현재 상호작용 상태가 *punching*일 가능성이 있음을 판단하는 기준을 나타낸다. 왼손목 혹은 오른손목에 해당하는 키포인트가 코의 키포인트와 비슷한 위치(높이)에서 검출된다면, 이는 싸움 중 주먹을 휘두르는 자세일 가능성이 있다.

이러한 계산을 통해 신체 부위 간의 관계를 파악하고 특정한 행동 패턴을 분류할 수 있으나, 위에서 예로 든 *punching*으로 분류된 자세는 코와 손목의 거리가 가깝다라는 사실만을 나타내고 실제로는 얼굴을 만지는 등의 동작과 구별할 수 없기 때문에 주먹을 휘두르는 동작이라고 확정할 수 없다. 따라서, 폭력 상황 여부는 이 결과를 LSTM 모델에 입력하여 출력된 결과를 보고 판단하게 된다.

### 4.4 시계열 모델 학습

앞서 추출한 키포인트 정보와 사람의 자세 및 상호작용 정보를 특성으로 하는 시계열 데이터의 학

습을 위해 Stacked LSTM(Long Short-Term Memory) 모델을 사용하였다. Stacked LSTM 모델은 여러 LSTM 계층을 쌓아 시간에 따른 움직임 변화를 더 깊이 학습할 수 있다. 이러한 구조를 통해 폭행 상황에서의 복잡한 행동 패턴을 보다 정교하게 모델링하고, 시스템이 다양한 폭행 시나리오를 놓치지 않고 감지할 수 있도록 한다.

3개의 연속된 프레임을 하나의 시퀀스로 묶어 모델에 입력했으며, 각 시퀀스의 마지막 프레임에 해당하는 라벨을 타겟으로 설정하였다. 학습 과정에서는 조기 종료(early stopping) 기법을 사용하여 과적합을 방지했다.

### 4.5 시스템 평가

본 연구에서 제안한 시스템의 성능 평가는 테스트 데이터셋을 사용하였고, 정확도(accuracy), 재현율(recall), F1-스코어 등의 지표를 기반으로 평가했다. 특히, 폭력 상황의 재현율을 중점적으로 분석하여, 폭력적 행동을 놓치지 않고 감지할 수 있는 능력을 측정했다.

그림 1에서 Model1은 상호작용 정보 없이 학습한 모델을 나타내고, Model2는 상호작용 정보를 같이 학습한 모델을 나타낸다.

상호작용 정보를 포함하여 모델을 학습한 경우와 그렇지 않은 경우의 성능 차이는 다음과 같다.

상호작용 정보를 포함하지 않고 학습한 경우, 전체 정확도는 55.49%였으며, 폭력 상황의 재현율이 72.28%에 그쳤다.

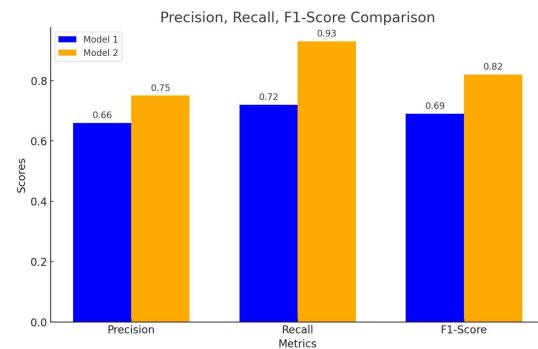


그림 1. 정밀도, 재현율, F1-스코어 비교  
Fig. 1. Precision, Recall, F1-Score Comparison

상호작용 정보를 포함하여 모델을 학습한 경우, 전체 정확도가 72.30%로 향상되었고, 폭력 상황의 재현율이 93.61%로 향상되었다.

## V. 결 론

본 연구는 두 사람의 신체 키포인트를 분석하여 상호작용 정보를 추출하고, 이를 시계열 모델 (LSTM)에 입력하여 학습하는 방법을 제안하였다. 그리고 상호작용 정보를 포함하여 학습한 결과, 폭력 감지의 정확도와 재현율이 향상되었음을 확인하였다. 따라서 이러한 시스템으로 기존의 정적 규칙 기반의 감시 시스템이 가진 한계를 극복하고, 더 복잡한 행동 패턴을 감지할 수 있다는 가능성을 제시하였다.

향후 연구에서는 객체 간의 상호작용 종류를 추가하여 정밀한 차이를 학습함으로써 모델의 성능을 한층 더 향상시킬 수 있을 것으로 기대된다.

또한, 이러한 연구 결과를 바탕으로 실시간 폭력 감지 시스템의 성능을 더욱 강화하여, 갑작스런 폭행이 발생할 경우에 피해자가 즉각적으로 신고하기 어려운 상황에서 빠르게 조치가 이루어질 수 있기를 기대한다.

## 참 고 문 헌

- [1] W. Sultani, C. Chen, and M. Shah, "Real-World Anomaly Detection in Surveillance Videos," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 6479-6488.
- [2] N. Mumtaz, N. Ejaz, S. Habib, S. M. Mohsin, P. Tiwari, and S. S. Band, "An Overview of Violence Detection Techniques: Current Challenges and Future Directions," Papers With Code, 2022.
- [3] M. Asif and T. S. Choi, "Shape from Focus Using Multilayer Feedforward Neural Network," in IEEE Transactions on Image Processing, vol. 10, no. 11, 2001, pp. 1670-1675.
- [4] AIHub, <https://www.aihub.or.kr/aihubdata/data/view.d>

- [5] H. Fang, S. Xie, Y. Tai, C. Lu, et al., "RMPE: Regional Multi-person Pose Estimation," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2334-2343.
- [6] S. Yan, Y. Xiong, and D. Lin, "Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition," Department of Information Engineering, The Chinese University of Hong Kong, 2018.
- [7] A. Toshev and C. Szegedy, "DeepPose: Human Pose Estimation via Deep Neural Networks," Google, 2014.
- [8] S.-H. Zhang, R. Li, X. Dong, P. Rosin, Z. Cai, X. Han, D. Yang, H. Huang, and S.-M. Hu, "Pose2Seg: Detection Free Human Instance Segmentation," Tsinghua University, BNRist, Cardiff University, Tencent AI Lab, 2019..



# 족적 검색 데이터 구축을 위한 다중라벨 분류 모델 개발

김호연\*, 권현섭\*, 서영주\*, 김동주\*\*

## Development of a Multi-label Classification Model for the Construction of a Shoeprint Retrieval Database

Hoyeon Kim\*, Hyunseob Kwon\*, Youngjoo Suh\* and Dongju Kim\*\*

### 요 약

범죄 현장에서 족적 정보로부터 범인을 추정하기 위해 사용되는 족적검색시스템은 감식관들이 직접 족적 이미지에 보이는 문양 정보를 수집하고 이를 기반으로 검색을 수행한다. 매년 수만 개의 신발을 감식관이 직접 검수하는 것은 한계가 있어 시스템의 확장성을 위해서는 문양 데이터를 빠르게 구축하는 방법이 필요한 상황이다. 본 논문에서는 다중라벨 분류 모델을 기반으로 문양 정보를 추출하는 모델을 통해 문양 데이터셋을 구축하는 방법을 제안한다. 다중라벨 분류 모델의 성능 비교를 수행하였을 때 EVA-V2 모델에 Asymmetric loss를 결합한 모델이 적합함을 확인했다. 또한, 실제 족적 문양 데이터셋과 오토라벨링한 데이터셋의 검색 성능을 비교하여 제안한 방법이 유효함을 확인하였다. 본 연구를 기반으로 족적 검색 시스템을 확장하는데 기여할 수 있을 것으로 기대한다.

### Abstract

The footwear search system, which is used to estimate the criminal from shoeprint information at a crime scene, is a system in which forensic investigators directly collect pattern information visible in shoeprint images and perform searches based on this pattern data. Since it is difficult for forensic investigators to directly inspect tens of thousands of shoes each year, a method to quickly build pattern data is needed to expand the system. In this paper, we propose a method for constructing a pattern dataset using a model that extracts pattern information based on a multi-label classification model. When comparing the performance of multi-label classification models, we confirmed that a model combining the EVA-V2 model with Asymmetric Loss was suitable. In addition, we confirmed the validity of the proposed method by comparing the search performance of the actual shoeprint pattern dataset and the auto-labeled dataset. We expect that this study will contribute to expanding the footwear search system.

### Key words

Multi-label classification, Shoeprint pattern, Shoeprint search system, Auto-labeling

---

\* 포항공과대학교 인공지능연구원, ghdus566@postech.ac.kr, kert06@postech.ac.kr, yjsuh@postech.ac.kr

\*\* 포항공과대학교 인공지능연구원, kkb0320@postech.ac.kr (교신저자)

※ 본 논문은 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업이며(No. 2022R1A6A1A0305295413), 정부(과학기술정보통신부)의 재원으로 한국지능정보사회진흥원의 지원을 받아 수행된 연구이며, (2024년 지역특화형 초거대 AI 클라우드 팜 실증 및 확산 환경 조성), 과학기술정보통신부·경찰청이 공동 지원한 '폴리스랩 2.0 사업'의(RS-2023-00281072) 지원을 받아 수행된 연구임.

## 1. 서 론

범죄 현장에서 발견할 수 있는 증거물은 생물학적 증거, 화학적 증거, 그리고 패턴 증거가 있다. 패턴 증거 중 족적은 용의자의 걸음걸이, 신장, 나이 및 성별 등 인적 사항을 추정하기 위한 주요한 단서로 활용되며, 현장의 지문, DNA 흔적이 없는 경우 초기 수사 방향을 좁히는 중요한 역할을 한다. 경찰청은 1997년 족·윤적감정시스템(Footwear and Tire Information System, FTIS)을 과학적범죄수사시스템(Scientific Crime Analysis System, SCAS)에 도입하였고, FTIS를 통한 족적 검색 처리 건수는 2018년 280건, 2021년 602건, 2023년 1,159건으로 최근 급격히 증가하는 추세이다.[1] 또한 2022년에는 족·윤적 감정서의 증거 능력을 높이기 위해 Korean Laboratory Accreditation Scheme (KOLAS)에서 인증한 족·윤적 감정관 자격제도가 도입되고 2023년 폴리스랩 2.0 과제를 수행하는 등 족·윤적 감식 및 검색 기술의 효율성 제고와 성능 고도화를 위해 노력하고 있다.

족적 검색 기술은 현장의 유류 족적 이미지와 DB에 등록된 신발 바닥 이미지를 전문적인 감식관이 직접 검색 및 대조하여 유력한 신발을 감정하여 단서로써 활용되며 주로 이미지 검색과 문양 키워드 검색 두 가지 방법으로 연구되고 있다.[2] 이미지 검색은 유류 이미지와 등록 이미지 사이의 유사 정도를 수치화하는 방식을 활용하여 과정이 직관적이지만 이미지의 도메인 요소(배경, 색깔 등)가 일치하지 않으면 검색 정확도와 신뢰성이 낮은 한계점이 있다.[3] 이러한 특성으로 인해 현재 주로 사용되는 족적 이미지 검색 방식은 유류 족적 이미지에 대해 문양 정보를 인식하고 인식된 문양의 유사 정도를 수치화하여 검색하는 방식을 활용한다. 문양 검색은 문양 정보만을 활용하여 검색하기에 이미지의 특성에 영향을 받지 않아 정확하면서도 효율적인 검색이 가능한 방법이다.[4] 하지만, 매년 수만 개의 새로운 신발들의 정답 문양을 감식관이 직접 라벨링 및 검수하기에 공수가 많이 들며, 전문적인 감식관이 아니라면 일관성이 있는 라벨링이 어려워 한정적인 등록 이미지만을 활용해야 한다는 한계점

이 있다.[5]

본 논문에서는 이러한 한계점을 해결하고자 다중 라벨 분류 모델을 통해 족적 이미지 내 문양을 추출하여 오토라벨링(Auto-labeling)을 수행하여 족적 데이터베이스 수량을 증가시켜 족적 검색 시스템의 확장성을 높이고자 한다. 모델의 성능에 주요한 영향을 끼치는 요소를 백본(Backbone) 모델과 손실 함수(Loss function)로 구분하여 비교한 결과 백본 모델에서는 EVA-V2가, 손실 함수에서는 ASL을 적용했을 때 Weighted F1-score 0.886으로 가장 높은 성능을 보임을 확인했다. 해당 모델의 문양 추출 결과를 문양 검색 시스템에 입력했을 때 실제 문양 검색 성능과 유사함을 확인할 수 있어 제안한 방법이 유효함을 확인하였다. 본 논문의 연구 결과를 기반으로 족적 문양을 일관성 있으면서도 정확한 문양 데이터를 구축할 수 있으며, 족적 이미지를 기존보다 높은 정확도로 검색할 수 있을 것으로 기대한다.

## II. 데이터셋

본 논문에서의 족적 데이터셋은 경찰청의 족적 검색 시스템에서 사용되는 등록 족적 데이터를 활용하였다. 족적 이미지 검색에 사용되는 문양 정보는 약 300개이며, 주로 신발 바닥에서 확인할 수 있는 다양한 평면도형(선, 사각형, 원 등)들의 세부적인 조합으로 구성되어 있다. 족적 이미지에서 보이는 문양 정보는 특정 문양들이 일정하게 구성되어 넓은 범위로 확인되는 전역적 문양과 세부적으로 좁은 범위에서 확인되는 지역적 문양이 있다. 한 이미지 내 문양 정보는 가시적으로 보이는 모든 문양을 포함하고 각 문양은 같은 중요도를 가지고 있어 지역적 문양과 전역적 문양을 모두 활용했을 때 족적 검색 정확도가 높게 측정된다. 현재 구축된 문양 데이터는 특정 클래스의 수가 많고, 대부분 클래스의 데이터 수가 적을 때 나타나는 long-tailed 분포 형태를 보인다. 일반적으로 많이 확인되는 지역적 패턴이 다수 분포 영역인 head에 속하며, 전역적 패턴들이 소수 분포 영역인 long-tailed 영역에 속해있으며, 이러한 불균형한 데이터 분포는 모델 학습 시 일반화된 예측 성능 저하의 주된 요인이다.

본 연구에서는 문양 데이터의 두 가지 특성을 고려하여 모델을 선정하고, 최종적으로 실제 검색 시스템의 족적 데이터와 성능 비교를 통해 선정한 모델을 통해 오토라벨링한 족적 데이터의 검색 성능을 비교하여 제안한 방법이 유효한지 확인하고자 한다.

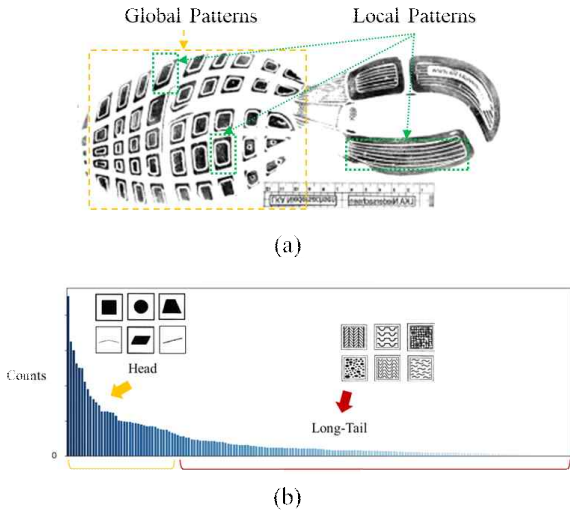


그림 1. 족적 문양 이미지의 특성, (a) 전역적/지역적 문양, (b) Long-tailed 분포

Fig. 1. Characteristics in shoeprint pattern images, (a)global/local patterns, (b)Long-tailed distribution

### III. 실험 및 결과

족적 이미지의 전역적, 국소적 요소에 나타나는 문양의 특징 추출 성능을 비교하기 위해 백본 모델의 경우 Convolutional neural network (CNN) 기반 모델과 Vision transformer (ViT) 기반 모델을 선정하여 성능을 비교하고자 한다. CNN 기반 모델은 ConvNeXt V2, ViT 기반 모델은 EVA-V2 모델을 선정하였으며 두 모델은 다중라벨 분류 분야에서 높은 성능을 보이는 모델이다.[6-7] 손실 함수는 다중라벨 분류에서 기본적으로 사용되는 Binary cross entropy (BCE), Label Smoothing 기법을 다중라벨 데이터에 적용한 Multi-label smoothing loss (MSL), 그리고 long-tailed 분포와 같이 불균형한 라벨 분포에 적합한 Asymmetric loss (ASL) 를 비교하여 성능을 비교하고자 한다.[8] 모델 학습에 사용되는 데이터 셋은 학습:검증:테스트를 3:1:1 비율로 랜덤샘플링

하였으며, backbone 모델 크기는 모두 base 모델을 활용했다. optimizer는 AdamW, 학습률은  $1e-4$ , 배치 크기는 8로 모두 같은 환경으로 설정하여 학습하였다. 다중라벨 분류 모델의 성능을 평가하기 위해 Overall F1-score (OF1), Average per-class F1-score(CF1)와 라벨 수의 영향을 고려한 Weighted F1-score(WF1) 지표를 선정하였으며, 모델의 예측 결과에 대한 성능 지표는 [표. 1]과 같다. 또한, 제안한 모델을 통해 구축한 데이터가 문양 검색 기반 족적 검색 시스템에 유효한지 검증하기 위해 실제 문양 정보 데이터로 구축한 데이터베이스와 모델 예측 결과 데이터만으로 구축한 데이터베이스의 검색 정확도를 비교한다. 검색 성능을 평가하기 위해 일반적으로 사용되는 Top 1% 및 5% 검색 정확도로 선정하였으며, 검색 성능 지표는 [표. 2]와 같다.

백본 모델에 따른 다중 분류 성능을 비교했을 때 전반적으로 EVA-V2의 성능이 높게 보여 족적 문양 추출에 더 적합한 모델임을 확인할 수 있었다. 족적 문양이 다양한 크기와 형태로 구성되어 있어 이러한 이미지의 특징 추출에 있어 높은 성능으로 보이는 것으로 분석된다. 세부적으로 손실 함수를 비교한 결과, ASL이 가장 높은 성능을 보임을 확인하여 학습 과정에서 클래스 불균형에 특화된 손실 함수를 사용하는 것이 유효함을 확인하였다. MSL 경우 가장 낮은 성능을 보임을 확인했는데 이는 단순히 정규화를 적용하는 것이 오히려 학습에 부정적인 영향을 보이는 것으로 분석된다. 최종적으로 EVA-V2 모델에 ASL loss를 사용해서 학습한 결과가 성능이 가장 높아 족적 문양 추출에 적합한 모델임을 확인하였다.

표 1. 다중라벨 분류 모델의 문양 추출 성능 지표  
Table 1. Results of multi-label classification pattern extraction models

백본 모델	손실 함수	OF1	CF1	WF1
ConvNeXt V2	BCE	0.728	0.831	0.829
	MSL	0.478	0.605	0.658
	ASL	0.779	0.848	0.846
EVA-V2	BCE	0.791	0.858	0.852
	MSL	0.539	0.620	0.678
	ASL	<b>0.836</b>	<b>0.888</b>	<b>0.886</b>

검색 성능을 비교했을 때 오토라벨링한 데이터가 Top-1% 정확도에서 5.082% 높아 제안한 방법으로 구축한 데이터가 기존보다 더 정확한 검색 결과를 제시하고 있는 것으로 확인하였다. 하지만 Top-5% 성능에서 2.37% 낮은 정확도를 보여 문양 예측이 어려운 항목에 대해서는 성능이 낮은 것으로 확인되나, 이는 다중라벨 분류에 적합한 분류기를 추가하고 최적화를 통해 개선할 수 있을 것으로 기대한다.

표 2. 문양 정보 기반 검색 성능 비교  
Table 2. Comparison of searching performance based on shoeprint patterns

문양 데이터	Top-1% 정확도(%)	Top-5% 정확도(%)
기존 문양 데이터	63.668	<b>90.189</b>
오토라벨링 데이터	<b>68.75</b>	87.819

#### IV. 결 론

본 논문에서는 족적 검색 데이터 구축을 위해 다중라벨 분류 모델 기반의 문양 추출 모델을 제안하고 실제 족적 검색 시스템에서의 성능을 확인하였다. 다중라벨 분류 모델의 성능에 영향을 미치는 요소를 백본 모델과 손실 함수로 구분하여 성능을 비교하였으며, EVA-V2와 ASL을 활용한 모델이 모든 성능 지표에서 우수함을 보임을 확인했다. 또한 제안한 모델을 활용하여 오토라벨링한 족적 문양 데이터셋을 구축하고 실제 문양 검색 시스템에 적용했을 때, 기존 시스템과 유사한 성능을 보여 제안한 기술이 유효함을 확인하였다. 최종적으로 본 연구를 바탕으로 문양 검색 데이터베이스를 빠르게 확장하여 족적 검색 시스템의 활용도를 높일 수 있기를 기대한다.

#### 참 고 문 헌

- [1] Korean National Police Agency, Open Data Portal, "National Police Agency\_footprint search processing status", <https://www.data.go.kr/en/data/15030190/fileData.do>, Aug. 2024.
- [2] Y. Wu, X. Dong, G. Shi, X. Zhang, and C. Chen, "Crime Scene Shoeprint Image Retrieval: A Review", *Electronics*, vol. 11, no. 16, p. 2487, Aug. 2022.
- [3] S. Alizadeh, H. B. Jond, V. V. Nabiyevev, and C. Kose, "Automatic Retrieval of Shoeprints Using Modified Multi-Block Local Binary Pattern", *Symmetry*, vol. 13, no. 2, p. 296, Feb 2021.
- [4] M. Budka, A. W. U. Ashraf, M. Bennett, S. Neville, and A. Mckrill, "Deep multilabel CNN for forensic footwear impression descriptor identification", *Applied Soft Computing.*, vol. 109, p. 107496, Sep 2021.
- [5] M. Asif and T. S. Choi, "Creating a Forensic Database of Shoeprints From Online Shoe-Tread Photos", *2023 IEEE/CVF Winter Conference on Applications of Computer Visioin*, pp. 858-868, 2023.
- [6] S. Woo, S. Debnath, R. Hu, X. Chen, I. S. Kweon, and S. Xie, "ConvNeXt V2: Co-Designing and Scaling ConvNets With Masked Autoencoders". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16133-16142, 2023.
- [7] Y. Fang, Q. Sun, X. Wang, T. Huang, X. Wang, and Y. Cao, "EVA-02: A Visual Representation for Neon Genesis". *arXiv preprint, arxiv.org/abs/2303.11331.*, 2023.
- [8] T. Ridnik, et al., "Asymmetric Loss for Multi-Label Classification", *Proceedings of the IEEE/CVF international conference on computer vision*, pp.82-91, 2021.

# 데이터 질에 따른 객체탐지 학습방법론에서 학습능력 변화

이 종 성\*

## Changes in Learning Capability in Object Detection Learning Methodologies Based on Data Quality

Jongsung Lee\*

### 요 약

인공 지능 및 기계 학습 방법론에서 데이터 도메인에 따른 학습 능력과 그 결과에 대한 부분은 많은 연구가 이루어져 있다. 일반적으로 학습에 있어서 질 좋은 데이터가 있을 때 학습이 우수하게 이루어지지만 항상 학습데이터가 좋게만 주어질 수는 없다. 가령 우천시나 미세먼지가 많은 상황에서의 CCTV 데이터는 청명한 날의 학습데이터보다 그 질이 떨어지게 된다. 이러한 경우 학습능력에 영향을 미칠 수밖에 없게 되는데 학습데이터의 질에 따른 학습 능력 변화에 대한 연구는 제한적이다. 따라서 본 연구에서는 객체 탐지 방법 중 YOLOv8s 방법론을 대상으로 학습데이터의 질을 흐려짐, 노이즈, 잘려짐 관점에서 변화 시켜보고 그 변화가 학습능력에 미치는 영향을 알아보았다. 그 결과 이상상태의 학습데이터를 포함한 경우가 보다 열악한 환경에서의 객체 탐지에 강건한(Robust) 결과를 보였다.

### Abstract

In artificial intelligence and machine learning methodologies, there has been extensive research on learning capabilities and outcomes based on data domains. Generally, when there is high-quality data for learning, the training is performed excellently; however, it is not always possible to have high-quality training data. For example, CCTV data collected during rainy weather or high levels of fine dust has a lower quality than training data collected on clear days. In such cases, the learning capability is inevitably affected, yet research on changes in learning capability based on the quality of training data is limited. Therefore, this study focuses on the YOLOv8s object detection method and investigates the impact of changing the quality of training data from the perspectives of blurriness, noise, and cropping. The results showed that including training data under abnormal conditions resulted in more robust object detection outcomes in harsher environments.

### Key words

Machine Learning, YOLO, Learning according to Data Quality, Data Augmentation

---

\* 국립한국교통대학교 산업경영공학과, Email: js.lee@ut.ac.kr

※ 2024년 한국교통대학교 지원을 받아 수행하였음

## 1. 서 론

인공 지능 및 기계 학습 방법론에서 데이터 도메인에 따른 학습 능력과 그 결과에 대한 부분은 많은 연구가 이루어져 있다. 일반적으로 학습에 있어서 좋은 데이터가 있을 때 학습이 우수하게 이루어지지만 항상 학습데이터가 좋게만 주어질 수는 없다. 가령 우천시나 미세먼지가 많은 상황에서의 CCTV 데이터는 청명한 날의 학습데이터보다 그 질이 떨어지게 된다. 이러한 경우 데이터의 질이 인공지능 모델의 학습능력에 영향을 미칠 수밖에 없게 되는데 학습 데이터의 질에 따른 인공지능 모델의 학습 능력 변화에 대한 연구는 제한적이다.

Prisyach 등과 Maharana 등, Akbiyik는 그들의 연구에서 데이터 증강 시 의도적인 노이즈 삽입이 인공지능 학습에서 데이터 과적합을 방지하여 학습 성능을 높인다는 것을 보였으며 Panda와 Roy의 연구와 Kim 등의 연구에서는 노이즈에 강건한 모델을 제안하였다. [1-5]

본 연구에서는 객체 탐지 모델인 YOLOv8s를 활용하여 현실에서 발생할 수 있는 이상데이터 요인인 흐려짐 (blur), 노이즈(noise), 잘려짐(crop or occlusion)에 대해서 각 요인의 수준을 변화시키고 그에 따라 만들어진 영상 데이터를 학습하였을 때 학습한 데이터에 따라 인공지능 모델이 얼마나 정확하게 정상 데이터와 이상 데이터에 대해서 비교하였다. 데이터 질에 따른 학습 능력 변화에 대한 기존 연구들은 노이즈와 같이 한가지 요인에 대한 학습능력 변화를 탐구하였으나 본 연구에서는 현실에서 발생할 수 있는 복합적인 이상 데이터에 대해서 각 요인별 학습의 영향 정도를 파악하고 요인들이 중첩되었을 때에 학습에 미치는 영향을 탐구하여 이상 데이터가 실제 학습에 미치는 현실적인 영향을 정량적으로 탐구하였다.

## II. 학습데이터 및 연구방법

ImageNet(<https://www.image-net.org>)에서 Sandglass, Coat, Tank, Mushroom, Plane, Train, Forklift, Butterfly, Monkey, Rabbit, Flog, Cat 12가지 종류의 이미지를 사용하여 모델 학습을 진행하였다. 각 이미지당 1,000개씩 총 12,000개의 데이터세트이고 Fig. 1과 같이 LabelImg를 이용해 Labeling하여 label

과 좌표를 Figure 5와 같이 class(label), x, y, width, height 5개의 지표로 구성하였다.

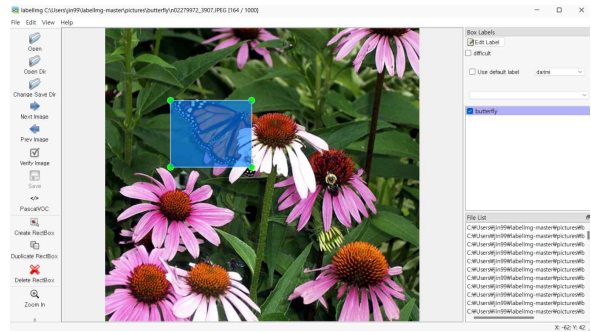


Fig. 1. Labeling Process

검증에 사용된 데이터는 Normal 데이터와 전체 데이터에서 Blur, Noise, Crop의 변형을 준 Abnormal 데이터로 설정하였다. Abnormal 데이터 생성은 blur와 noise, crop 이미지를 생성하였다. blur는 gaussian blure를 적용하였고 noise는 Salt&Pepper noise를 적용하였다. crop은 label된 데이터의 Bounding box를 기준으로 상하 좌우 모든 방향에 대해서 무작위로 정해진 수준의 그림을 삭제하는 방법을 활용하였다. Table 1은 이상데이터 생성 기법을 정리한 것이다.

Table 1. Abnormal data generation methods

blur	gaussian blur 적용 $G(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}}$
noise	Salt & Pepper Noise 적용 $f(x) = \begin{cases} P\alpha & x = \alpha \\ 1 - P\alpha, & \text{otherwise} \end{cases}$
crop	Labeling을 마친 이미지의 좌표 정보를 바탕으로 객체만 추출한 후에 총 4가지 방향(상하좌우)으로 랜덤하게 x% 삭제

Table 1의 이상데이터 생성 기법에 따라 생성된 이상데이터는 Fig. 2와 같다. Fig. 2의 (a)와 (b)는 blur 수준별 이미지 생성 결과이고 (c)와 (d)는 noise 수준 별 데이터 (e)와 (f), (g)는 crop 수준 별 이상 데이터 생성 결과이며 마지막으로 (h)는 3가지 요인을 모두 적용하였을 때 생성된 결과 이미지 예시이다.

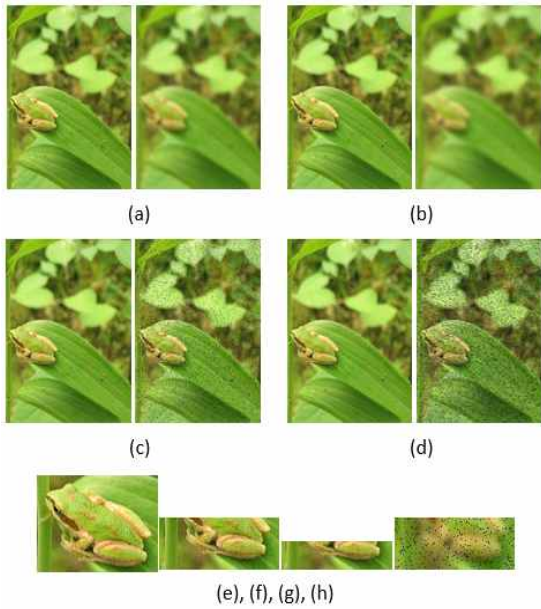


Fig. 2. Abnormal data generation results

정상데이터와 생성된 이상 데이터를 활용하여 YOLOv8s를 학습시켰다. Fig. 3은 본 연구의 전체적인 학습과 테스트 결과를 개념적으로 표현한 것이다. 모델 학습에 1) 정상 데이터를 학습하였을 경우, 2) blur 데이터를 학습하였을 경우, 3) noise 데이터를 학습하였을 경우, 4) crop 데이터를 학습하였을 경우, 5) blur와 noise를 중첩하여 적용한 데이터를 학습하였을 경우 6) blur와 noise를 중첩한 데이터와 정상 데이터를 혼합하여 학습하였을 경우 7) blur와 noise, crop, 정상데이터를 혼합한 데이터를 학습하였을 경우로 학습 모델을 나누고 각 학습 모델에 대해서 a) 정상데이터, b) blur 데이터, c) noise 데이터, d) crop 데이터, e) blur&noise 중첩데이터, f) blur&crop 중첩데이터, g) noise&crop 중첩데이터, h) blur&noise&crop 삼중첩 데이터에 대해서 mAP(mean Average Precision) 값을 비교하였다.

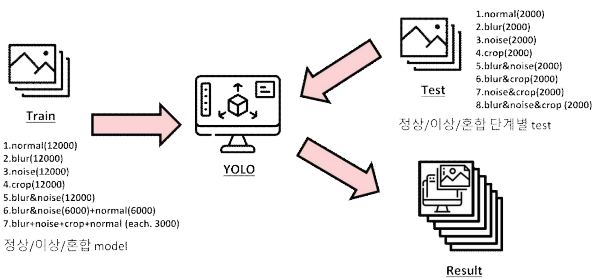


Fig. 3 Conceptual diagram of this study

### III. 연구 결과

Fig. 4는 정상 데이터만을 학습시켰을 때 정상데이터와 이상데이터에 대한 mAP를 막대그래프로 표현한 것이다.

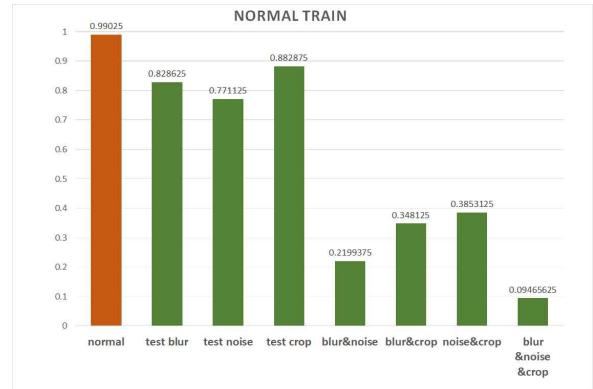


Fig. 4. mAP for the YOLO model based on normal data

정상 모델로 학습한 경우 정상 데이터에 대한 mAP가 0.99이상으로 매우 정확하게 객체를 탐지하는 것이 확인되었다. 그러나 blur와 noise 등 이상 데이터에 대해서는 객체 탐지 정확도가 감소하는 것을 확인할 수 있었다. 특히나 요인이 중첩되어 나타나는 이상 데이터의 경우에는 탐지 정확도가 급감하여 blur와 noise가 중첩 되어 나타난 이상 데이터의 경우 탐지 정확도가 0.21 수준이었고 blur와 noise, crop 이 동시에 적용된 경우 탐지 정확도가 0.09정도로 10%에도 미치지 못하는 것이 확인되었다. 정상 데이터 만을 가지고 학습을 한 경우 이상 데이터에 대해서 객체 탐지 정확도가 떨어짐을 알 수 있었고 특히나 두가지 이상의 요인이 중첩되어 데이터 질이 떨어졌을 경우 정상 데이터를 학습한 모델의 판정 정확도가 현저히 떨어지는 것을 확인할 수 있었다.

정상 데이터만을 학습하였을 때 이상데이터에 대한 객체 탐지 정확도가 현저히 떨어지는 것을 방지하기 위해서 이상데이터를 학습한 경우들에 대해서 mAP를 측정하였으며 T-Test를 이용하여 각 모델들이 정상 및 이상 데이터에 대해 객체 탐지를 할 때 통계적으로 유의미한 차이가 있는지 여부를 탐구하였다.

( ) = mAP

결과 모델 순위	Total mAP	test dataset 요인별 비교							
		blur	noise	crop	blur noise	blur crop	noise crop	blur noise crop	normal
1	combination (0.837), blur&noise( 0.817)	blur (0.984)	noise (0.984), combination (0.981)	crop (0.983)	combination (0.969)	blur (0.757)	noise (0.821)	blur&noise (0.68)	combination (0.986), noise (0.989), normal (0.99)
2	noise (0.786)	combination (0.956)	blur&noise (0.937)	normal (0.882)	blur&noise (0.934)	blur&noise (0.643)	combination (0.764)	combination (0.642)	blur (0.969)
3	blur (0.599), crop (0.448), normal (0.565)	blur&noise (0.933)	normal (0.771)	combination (0.871), noise (0.872)	noise (0.643)	crop (0.59)	blur&noise (0.716)	noise (0.295)	blur&noise (0.94)
4		noise (0.855)	blur (0.694)	blur&noise (0.755)	blur (0.342)	combination (0.526)	crop (0.678)	crop (0.197)	crop (0.362)
5		normal (0.828)	crop (0.316)	blur (0.713)	normal (0.219)	normal (0.348)	normal (0.385)	normal (0.094), blur (0.094)	
6		crop (0.339)			crop (0.122)	noise (0.34)	blur (0.237)		

Fig. 5. T-test results of learning models according to the test data sets

Fig. 5는 mAP를 통계적으로 비교하였을 때 가장 성능이 좋은 순서대로 나열한 결과이다. 같은 행에 위치한 모델은 통계적으로 서로 차이가 없는 모델을 의미한다. 정상 데이터 학습 모델(normal), blur 데이터 학습 모델(blur), noise 데이터 학습 모델(noise), crop 데이터 학습 모델(crop), blur&noise 중첩 데이터 학습 모델(blur&noise), blur&noise 중첩데이터와 정상 데이터 학습 모델(combination) 모델이 비교 되었다.

Fig. 5의 Total mAP는 각 모델의 전체 평균 mAP를 의미하며 이는 test data set인 blur, noise, crop, blur&noise, blur&crop, noise&crop, blur&noise&crop과 normal 데이터 전체에 대한 정확도 결과를 평균한 것이다. 나머지 열에 표현된 결과는 각 test data set에 대해서 각 모델의 mAP를 나타낸 것이다. 예를

들어 noise 열의 첫 번째 행에 있는 noise(0.984)와 combination(0.981)은 noise가 섞인 test 데이터에 대해서 noise가 있는 데이터로만 학습한 모델과 combination 모델이 통계적으로 유의미한 차이가 없이 일순위에 랭크되었으며 각각 noise 테스트 데이터에 대한 mAP가 0.984와 0.981임을 의미한다.

Total mAP로 비교할 때에 combination 모델이 blur&noise 모델과 함께 평균적인 정확도가 가장 높았고 이들 모델이 정상 데이터 뿐만아니라 이상데이터에 대해서도 높은 정확도를 보여주는 것을 확인할 수 있었다. 따라서 학습에 있어서 이상 데이터를 사용하여 학습 하는 것이 정상 데이터만을 가지고 학습하는 것에 비해서 보다 강건한 평가 성능을 보여주는 것을 확인할 수 있었다.



#### IV. 결 론

본 연구에서는 현실에서 발생할 수 있는 blur, noise, crop 등의 이상 데이터에 대해 정상 데이터만을 학습한 모델의 객체 탐지 성능이 이상 데이터에 대해서는 떨어지는 것을 확인하였으며 특히 두 개 이상의 요인이 중첩되어 이상 데이터가 발생하는 경우에는 현저하게 평가 성능이 떨어지는 것을 확인하였다. 이를 보완하게 위해서 모델 학습을 시킬 때 이상데이터를 포함한 모델 학습을 하는 것이 평균적으로 높은 mAP 값을 가지는 것을 확인하였으며 blur와 noise, crop 의 세가지 요인을 대상으로 할 때 blur&noise와 정상 데이터를 혼합하여 학습 시킨 combination 모델이 이상 데이터에 대해서 가장 강한 평가 성능 보여주는 것을 확인하였다. 따라서 모델이 이상데이터에 대해서 높은 객체 탐지 성능을 가지고 위해서는 이상 데이터를 학습하는 것이 좋으며 특히나 중첩된 이상 데이터가 있는 경우에는 중첩된 이상데이터를 학습하는 것이 전반적으로 우수한 성능을 나타내는 것을 확인하였다.

본 연구는 YOLO 객체 탐지 방법론에 대해서 학습데이터를 변경하였을 때 평가 성능의 변화를 탐구하였는데 추후에는 Retina-Net과 Fast/Faster-RNN 등의 다른 객체 탐지 방법론에서도 동일한 결과가 나오는지 확인하여 일반적인 인공지능 학습 방법론에서도 동일한 결과가 나오는지 확장할 예정이다.

#### 참 고 문 헌

- [1] T. Prisyach, V. Mendelev, and D. Ubskiy, "Data Augmentation for Training of Noise Robust Acoustic Models", Proc. Social Networks and Texts: 5<sup>th</sup> International conference, AIST 2016 Yekaterinburg, Russia, April 7-9, 2016.
- [2] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques", Proc. Global Transitions Proceedings, pp 91-99, 2022
- [3] M. E. Akbiyik, "Data augmentation in training CNNs: injecting noise to images." arXiv preprint

arXiv:2307.06855, 2023.

- [4] P. Panda, and K. Roy, "Implicit adversarial data augmentation and robustness with noise-based learning." Neural Networks, 141, pp120-132.
- [5] E. Kim, J. Kim, H. Lee, and S. Kim, "Adaptive data augmentation to achieve noise robustness and overcome data deficiency for deep learning." Applied Sciences, vol. 11 no. 12, 5586. 2021

# 협업필터링을 위한 레이블 스무딩 기반 변분 오토인코더

윤우성\*, 김영현\*, 정찬우\*, 최여준\*\*, 조윤식\*\*\*

## Variational Autoencoder with Label Smoothing for Collaborative Filtering

Wooseong Yun\*, Yeonghyeon Kim\*, Chanwoo Jeong\*, Yeojun Choi\*\*, and Yoonsik Cho\*\*\*

### 요 약

디지털 콘텐츠의 급증에 따라 효율적인 정보 탐색을 위한 추천 시스템의 필요성이 커지고 있다. 기존의 Mult-VAE와 같은 협업 필터링 모델은 아이템 순위를 반영하기 위해 Softmax 함수를 사용하지만, 이로 인해 확률 분포가 특정 아이템에 편향되어 일반화 성능이 저하될 수 있다. 이를 해결하기 위해 본 연구는 Label Smoothing 기법을 적용하여 특정 아이템에 대한 과도한 확률 부여를 억제하고, 아이템의 다양성을 증진시키는 새로운 방법을 제안한다. MovieLens20M 및 Netflix 벤치마크 실험을 통해 제안된 방법의 효과를 검증했으며, 하이퍼파라미터 실험을 통해 추천의 정확도와 다양성 간 균형을 유지하는 최적의 Label Smoothing 강도를 도출하였다.

### Abstract

With the rapid increase in digital content, the demand for recommendation systems to facilitate efficient information retrieval has grown. Traditional collaborative filtering models, such as Mult-VAE, commonly use the Softmax function to reflect item rankings. However, this approach can lead to skewed probability distributions that overfit to specific items, ultimately damaging generalization performance. To address this issue, we propose a novel method incorporating label smoothing to prevent the over-assignment of probability to specific items. Our experiments on MovieLens20M and Netflix datasets, we demonstrate the effectiveness of our proposed method. Additionally, hyperparameter experiment allows us to identify an optimal level of Label Smoothing that balances recommendation accuracy and diversity.

### Key words

Recommendation System, Collaborative Filtering, Variational Autoencoder, Label Smoothing,

---

\* 중앙대학교 AI학과 석사과정, {dntjd0804, hyuns6100, chwchong}@cau.ac.kr

\*\* 중앙대학교 AI학과 박사과정, tmscks94@cau.ac.kr

\*\*\* 중앙대학교 AI학과 교수, yoonsik@cau.ac.kr (교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과 (IITP-2024-RS-2024-00438056)와 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2024-00419201)

## I. 서론

온라인 플랫폼에서의 디지털 콘텐츠들이 증가함에 따라, 추천시스템은 정보 과부하 해결 및 사용자의 검색 경험 개선에 필수적인 존재가 되었다. 이들의 목표는 각 사용자의 선호도를 파악해 적절한 아이템을 추천해주는 것이다.

협업 필터링(Collaborative Filtering) [1, 2, 3, 4, 5, 6]은 가장 대중적인 방법으로서 사용자 간 또는 아이템 간의 유사성을 바탕으로 취향이 비슷한 사용자들이 선호하는 아이템을 추천하는 방법이다. 최근에는 딥 뉴럴 네트워크 (Deep Neural Network)의 성공에 힘입어, Variational Autoencoder(VAE) [7]와 같은 방법이 추천시스템에 도입되어왔다.

그러나 VAE에서 일반적으로 사용되는 BCE(Binary Cross-Entropy) 재구성 손실은 모든 아이템을 동일하게 취급하기 때문에 각 사용자에게 추천된 아이템의 순위를 고려하는 NDCG 평가지표와는 적합하지 않다. Mult-VAE [5]는 이를 해결하기 위해 Multinomial Likelihood와 이를 최적화하기 위해 Softmax 함수를 사용해 각 아이템의 상대적인 순위를 반영한다. 하지만 Softmax 함수는 입력값을 지수로서 계산하기에, 약간 큰 입력값에 대해 지나치게 높은 확률을 부여하는 경향이 있다. 이는 특정 클래스에 대한 확률을 급격히 증가시켜, 특정 아이템에 대한 과도한 확신으로 이어지고 결국 모델의 일반화 성능을 저하시킬 수 있다.

Image Classification 분야에서는 손실함수에 Label Smoothing을 적용하여 특정 클래스에 과도하게 집중되는 것을 억제한다 [8]. 이를 통해 모델은 다양한 클래스에 접근할 수 있으며, 일반화 성능에서도 강점을 보인다. 이를 바탕으로 본 논문에서는 Label Smoothing에서 영감을 받은 새로운 재구성 함수를 제안한다. 우리의 제안된 방법은 모델이 특정 아이템에 과도하게 높은 확률을 부여하는 것을 억제함으로써 다양한 아이템을 탐색할 수 있도록 유도한다. 이를 통해 모델은 아이템의 순위를 더 잘 고려할 수 있으며, 표준 벤치마크에서 상당한 성능을 개선한다.

## II. 본론

### 2.1 Preliminary

각 사용자  $u \in \{1, \dots, U\}$ 에 대해, 잠재 변수  $\mathbf{z}_u \in \mathbb{R}^k$ 는 다변량 가우시안 분포  $N(0, \mathbf{I}_k)$ 에서 샘플링된다. 샘플링된 잠재 변수  $\mathbf{z}_u$ 는 비선형 함수  $f_\theta(\cdot)$ 를 통해 변환되어,  $I$ 개의 아이템에 대한 확률 분포  $\pi(\mathbf{z}_u)$ 를 계산한다. 이 확률 분포  $\pi(\mathbf{z}_u)$ 는 일반적으로 다항 분포로서, 사용자의 아이템 클릭  $\mathbf{x}$ 를 샘플링하는 데 사용된다.

$$\begin{aligned} \mathbf{z}_u &\sim N(0, \mathbf{I}_k), \quad \pi(\mathbf{z}_u) \propto \exp\{f_\theta(\mathbf{z}_u)\}, \\ \mathbf{x}_u &\sim \text{Mult}(N_u, \pi(\mathbf{z}_u)), \end{aligned} \quad (1)$$

여기서  $N_u$ 는 사용자  $u$ 의 총 클릭 수이며, 사용자의 아이템 클릭  $\mathbf{x}_u$ 는  $N_u$ 회 시도 동안  $\pi(\mathbf{z}_u)$ 를 기반으로 생성된다. 이러한 생성 과정을 통해 VAE 프레임워크가 적용되며, 데이터의 우도  $p(\mathbf{x}) = \int p(\mathbf{x}|\mathbf{z})p(\mathbf{z})d\mathbf{z}$ 를 최대화하는 것을 목표로 한다. 그러나 비선형 매핑  $f_\theta(\cdot)$ 에 신경망을 사용하는 경우,  $p(\mathbf{x})$ 를 계산하는 것이 어려워진다. 이를 해결하기 위해 변분 추론이 적용되며, 주요 목표는 증거 하한(ELBO, Evidence Lower Bound)을 최대화하는 것이다.

$$\begin{aligned} \log p(\mathbf{x}_u; \theta) &\geq E_{q_\phi(\mathbf{z}_u|\mathbf{x}_u)}[\log p_\theta(\mathbf{x}_u|\mathbf{z}_u)] \\ &\quad - \text{KL}[q_\phi(\mathbf{z}_u|\mathbf{x}_u) \| p(\mathbf{z}_u)] := L(\mathbf{x}_u; \theta, \phi). \end{aligned} \quad (2)$$

여기서  $q_\phi(\mathbf{z}_u|\mathbf{x}_u)$ 는 계산 불가능한 사후 분포  $p(\mathbf{z}_u|\mathbf{x}_u)$ 를 최적으로 근사하도록 최적화된다. VAE 학습 방식에서는 디코더와 인코더의 매개변수  $\theta$ 와  $\phi$ 가 각각 업데이트되며, 이를 통해 모델을 학습시킨다.

### 2.2 Proposed Method

이번 장에서는 Label Smoothing 기법을 적용한 우리의 새로운 손실함수를 소개한다. 우리의 방법은

모델이 특정 클래스를 과신하는 것을 막음으로써 과적합을 방지할 수 있으며, 모델이 좀 더 다양한 아이템을 추천하도록 촉진한다. 우리의 최적화 함수는 식 (2)의  $\mathbf{x}_u$ 를  $\mathbf{x}'_u = \mathbf{x}_u(1-a) + a/2$ 로 대체하여 모델의 학습 과정을 개선한다.

$$\begin{aligned} \log p(\mathbf{x}'_u; \theta) &\geq E_{q_\phi(\mathbf{z}_u, \mathbf{x}'_u)} [\log p_\theta(\mathbf{x}'_u | \mathbf{z}_u)] \\ &- \text{KL}[q_\phi(\mathbf{z}_u | \mathbf{x}'_u) \| p(\mathbf{z}_u)] := L(\mathbf{x}'_u; \theta, \phi). \end{aligned} \quad (3)$$

### III. 실험

#### 3.1 실험 세팅

본 연구는 Mult-VAE의 모든 프로토콜을 따른다. Mult-VAE의 전처리 코드를 기반으로 훈련, 검증 및 테스트 데이터셋을 고정하고, 강한 일반화 설정 [9] 하에서 모델을 평가한다. 모델 학습에는 훈련 세트의 전체 클릭 기록을 사용하며, 검증 및 테스트 단계에서는 사용자 클릭 기록의 80%를 학습에 활용하고, 나머지 20%에 대해 예측 성능을 평가한다. 또한 Label Smoothing의 영향력을 조절하는 하이퍼파라미터  $\alpha = 0.05$ 로 설정하였다.

#### 3.4 주요 결과

표 1은 선형 및 비선형 협업 필터링 모델들과 비교하여 제안된 방법의 성능을 나타낸다. 본 연구의 모델은 대부분의 평가 지표에서 주요 베이스라인 모델들을 초과하는 성능을 보였다. 특히, Label Smoothing을 적용하여 순위 정보를 반영함으로써, 사용자 선호도 예측의 정확도를 향상시킬 수 있었

다. 다만, MovieLens20M 데이터셋에서 Mult-VAE와 비교했을 때, Recall@50에서 약간의 성능 저하가 관찰되었다. 이는 Label Smoothing이 순위에 따른 아이템 예측 정확도를 높이는 데 기여하지만, 순위가 낮은 아이템들의 예측 확률이 유사하게 형성됨에 따라 새로운 아이템이 추가될 경우 이들 간의 확률 차이가 축소되어 성능 저하로 이어질 수 있음을 시사한다.

#### 3.5 Hyperparameter Sensitivity

표 2는 미세한 스무딩부터 강한 스무딩까지 Label Smoothing의 강도 조절에 따른 추천 성능의 변화를 보여준다. 실험 결과,  $\alpha = 0.05$ 일 때 약간의 스무딩 효과로 최적의 추천 정확성을 달성할 수 있음을 확인하였다. 반면,  $\alpha$ 가 증가함에 따라 모델의 예측이 지나치게 평탄해져 추천 정확성이 저하되는 경향을 확인하였다. 이는 적절한 수준의 Label Smoothing이 추천의 정확도와 아이템의 다양성 간의 균형을 효과적으로 유지할 수 있음을 의미한다.

표 2. Label smoothing의  $\alpha$ 에 따른 모델의 성능 비교

$\alpha$	MovieLens20M		
	NDCG@100	Recall@50	Recall@20
0.01	0.429	0.533	0.397
0.05	<b>0.430</b>	<b>0.535</b>	<b>0.399</b>
0.1	0.429	0.530	0.399
0.2	0.426	0.528	0.397

$\alpha$	Netflix		
	NDCG@100	Recall@50	Recall@20
0.01	0.388	<b>0.444</b>	0.354
0.05	<b>0.388</b>	0.443	<b>0.355</b>
0.1	0.388	0.441	0.355
0.2	0.386	0.438	0.354

표 1. 표준 벤치마크에서의 성능 비교

Model	MovieLens20M			Netflix		
	NDCG@100	Recall@50	Recall@20	NDCG@100	Recall@50	Recall@20
WMF	0.386	0.498	0.360	0.351	0.404	0.316
SLIM	0.401	0.495	0.370	0.379	0.428	0.347
CDAE	0.418	0.523	0.391	0.376	0.428	0.343
Mult-VAE	0.426	<b>0.536</b>	0.395	0.386	0.440	0.348
Ours	<b>0.430</b>	<b>0.535</b>	<b>0.399</b>	<b>0.388</b>	<b>0.443</b>	<b>0.355</b>

#### IV. 결 론

본 연구에서는 VAE 기반의 추천시스템에 Label Smoothing 기법을 도입하여 모델이 특정 아이템에 과도한 확률을 부여하는 문제를 효과적으로 해결한다. 표준 베이스라인 모델과의 비교 실험에서 유의미한 성능 향상을 달성하며, Label Smoothing을 통해 확률 분포를 보다 균형 있게 형성하여 추천의 정확도를 향상시킬 수 있음을 확인하였다. 또한, 스무딩의 강도를 조절하는 하이퍼파라미터에 대한 다양한 실험을 통해 적절한 수준의 Label Smoothing이 추천 정확도와 아이템 다양성 간의 균형을 유지하는 데 중요한 역할을 한다는 것을 보였다.

#### 참 고 문 헌

- [1] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems", *Computer*, Vol. 42, No. 8, pp. 30-37, August 2009.
- [2] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets", *Proceedings of the Eighth IEEE International Conference on Data Mining*, pp. 263-272, December 2008.
- [3] X. Ning and G. Karypis, "Slim: Sparse linear methods for top-n recommender systems", *Proceedings of the 11th IEEE International Conference on Data Mining*, pp. 497-506, December 2011.
- [4] Y. Wu, C. DuBois, A. X. Zheng, and M. Ester, "Collaborative denoising auto-encoders for top-n recommender systems", *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pp. 153-162, February 2016.
- [5] D. Liang, R. G. Krishnan, M. D. Hoffman, and T. Jebara, "Variational autoencoders for collaborative filtering", *Proceedings of the 2018 World Wide Web Conference*, pp. 689-698, April 2018.
- [6] X. Yu, X. Zhang, Y. Cao, and M. Xia,

- "VAEGAN: A collaborative filtering framework based on adversarial variational autoencoders", *Proceedings of the 2019 International Joint Conference on Artificial Intelligence (IJCAI)*, Vol. 19, pp. 4206-4212, August 2019.
- [7] D. P. Kingma, "Auto-encoding variational bayes", *arXiv preprint arXiv:1312.6114*, 2013.
- [8] R. Müller, S. Kornblith, and G. E. Hinton, "When does label smoothing help?," in *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [9] B. MARLIN, "Collaborative Filtering: A Machine Learning Perspective", *University of Toronto*, pp. 2239-2239, 2004.

# Multi-head Cross Attention을 활용한 의도 계층화

김은선\*, 강유진\*\*, 조윤식\*\*\*

## Hierarchical Intent Representation Using Multi-head Cross Attention

EunSun Kim\*, Yujin Kang\*\*, and Yoon-Sik Cho\*\*\*

### 요 약

주관성이 많이 반영되는 의도 분류 태스크는 원본 데이터의 특성을 파악하는 것이 중요하다. 또한 50가지 이상의 많은 라벨들을 다루며, 일부 라벨 의미 중복 및 오류가 존재하는 추가적인 어려움이 있다. 본 논문에서는 텍스트와 라벨, 다른 라벨들과의 연관성을 고려한 Multi head Cross Attention 계층화 방법을 제안한다. 특히, 전체 데이터 샘플을 평균화한 방법보다 각 라벨에 대해 평균화한 방법이 더 우수한 군집화 현상을 보여주며, 라벨의 계층 의미를 더 우수하게 반영됨을 확인한다.

### Abstract

In intent classification tasks, where subjectivity is reflected in a significant role, understanding the characteristics of the original data is crucial. Additionally, handling over 50 labels presents further challenges due to semantic overlap and occasional errors in label meaning. This paper proposes a Multi-Head Cross Attention hierarchical method that considers the relationships between text, labels, and other labels. Notably, the approach of averaging per label, rather than averaging over the entire dataset samples, demonstrates a superior clustering effect, effectively capturing the hierarchical meaning of the labels.

### Key words

Intent detection, Cross Attention, Hierarchical Clustering, Large Language Model, Artificial Intelligence

### 1. 서 론

주어진 프롬프트를 참고하여 답변을 생성할 수

있는 초거대 언어 모델은 데이터 증강 기법[1]에 활발히 적용되고 있다. 몇 개의 적은 원본 데이터 샘플을 참고하여 많은 합성 데이터를 생성할 수 있어 불균형 태스크에 특히 활발히 활용되고 있다.

\* 중앙대학교 AI학과 석사과정, eunsun121@cau.ac.kr

\*\* 중앙대학교 AI학과 박사과정, zinjin32@cau.ac.kr

\*\*\* 중앙대학교 AI학과 교수, yoonsik@cau.ac.kr (교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과 (IITP-2024-RS-2024-00438056)와 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. RS-2024-00419201)

그러나 주관성이 심한 자연어 처리 분야의 경우, 생성된 합성 데이터를 적용 시 성능이 낮아지는 현상[2]이 있다.

특히, 의도 분류와 같이 주관성이 심한 태스크의 경우, 원본 데이터의 특성을 잘 반영한 고품질의 합성 데이터를 선택하는 방법이 매우 중요하다. 또한 의도 분류는 여러 가지의 라벨들이 있어, 해당 특성들을 고려하여 분류를 해야한다. 기존 연구들은 attention 메커니즘[3]이나 contrastive learning[4]을 사용해 의도 분류를 수행하였다. 그러나 의도 분류는 단순히 라벨을 예측하는 것을 넘어, 라벨의 의미가 유사하거나, 라벨링 오류가 존재하는 경우도 있기에 추가적인 어려움이 있다.

의도 분류 태스크가 가진 이러한 어려움들을 해결하기 위해, 본 연구에서는 원본 데이터의 특성을 파악하기 위한 계층화 방안을 제시한다. 또한 라벨의 의미에 따른 단순 계층화가 아닌, 라벨링된 텍스트와의 다른 라벨들 간의 관계를 고려한 더 정교화된 계층화를 제안한다.

## II. 본 론

### 2.1 데이터셋

본 논문에서는 Banking77[5] 데이터셋을 활용한다. 해당 데이터셋은 13,242개의 샘플 데이터와 77개의 라벨로 이루어졌고, 'card\_arrival', 'card\_linking', 'exchange\_rate'와 같은 은행 도메인의 데이터셋이다.

의도 분류 태스크에서 많이 활용되는 데이터셋이지만, 라벨링 의미 중복 및 라벨링 오류 문제[6]가 제기되었다. 때문에 본 논문에서는 의도 분류 태스크에서 빈번히 발생될 수 있는 라벨링 오류를 완화한 계층화 의도 분류 방법을 제안한다.

### 2.2 제안 방법

Multi-head Cross Attention을 활용하여 라벨링된 텍스트와 다른 라벨들간의 관계를 고려한다. 문장 임베딩을 Query로, 라벨 임베딩을 Key와 Value로 적용하여 Attention score를 산출한다. 이를 통해 해당 문장이 각 라벨들에 대해 얼마나 관련 있게 보는지를 attention score를 통해 계산한다. 이후 마지막

레이어의 multi-head attention 값을 평균화한 후, 라벨 간 코사인 유사도를 계산하여 라벨 간 유사성을 도출한다.

$$Attention(Q, K, V) = softmax\left(\frac{Q \cdot K^T}{\sqrt{d_{head}}}\right) \cdot V \quad (1)$$

$$MultiHead(Q, K, V) = [head_1; \dots; head_n] \cdot W^O \quad (2)$$

이후 계층적 군집화 알고리즘 중 하나인 Agglomerative Hierarchical Clustering을 통해 해당 라벨들에 대한 계층화를 수행한다. 해당 알고리즘은 여러 데이터들에서 가장 거리가 가까운 데이터를 찾은 뒤, 하나의 클러스터로 묶는다. 이렇게 모은 데이터가 하나의 전체 클러스터로 묶일 때까지 과정을 반복한다. 해당 알고리즘을 바탕으로 덴드로그램 시각화도 가능하며, 해당 알고리즘으로 적용하여 데이터의 특성을 살핀다. 위 흐름은 아래 그림1과 같다.

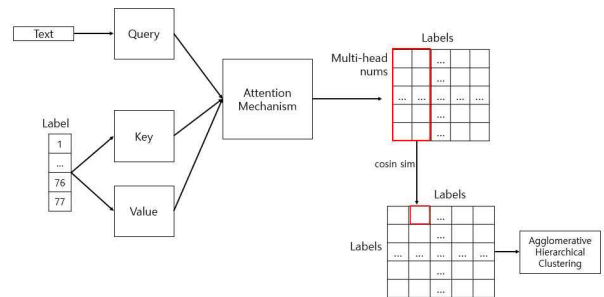


그림 1. Multi head Cross Attention Hierarchical Clustering 구조화

Fig. 1. Multi head Cross Attention Hierarchical Clustering

## III. 실험

### 3.1 실험 방법

실험에는 OPT-6B 모델[7]을 활용하여 텍스트와 라벨의 임베딩 값을 반영한 후, 마지막 레이어의 multi head attention score 값을 활용한다.

추출된 multi-head attention score값 평균화에는 모든 샘플들에 대해 평균화하는 방법과 각 라벨별로 평균화하는 두 가지 방법이 있다. 본 실험에서는 두 가지 방법 모두 적용하여 결과를 비교해본다.

### 3.2 실험 결과

그림2와 같이 모든 라벨에 대해 평균화한 값보다, 그림3과 같이 각 라벨별로 평균화한 방법이 더 우수한 군집화 현상을 보인다.

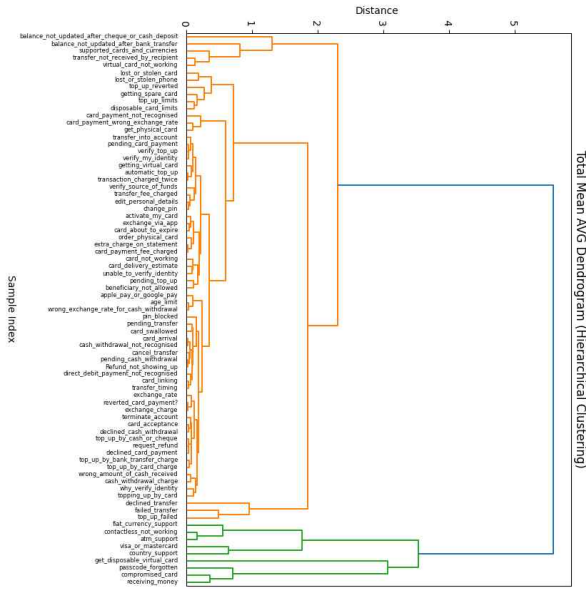


그림 2. 전체 multi-head attention score 평균 계층화  
Fig. 2. Total average of Multi head Cross Attention Hierarchical Clustering

반영하는 것을 알 수 있다.

'declined\_cash\_withdrawal','cash\_withdrawal\_not\_recognised'은 현금 인출과 관련하여 각각 거절, 인식 오류에 해당된다. 각 라벨 평균화 방법에서 더 가까이 군집화 됨을 확인할 수 있다. 또한 카드 결제 오류 관련 라벨인 'declined card payment', 'card\_payment\_not\_recognised' 등이 더 근접한 거리에서 군집화를 형성하고 있다.

#### IV. 결 론

본 논문에서는 원본 데이터의 특성을 파악해야하고, 라벨링 오류가 있는 의도 분류 태스크에서 multi-head cross attention을 활용한 계층화 방법을 제시하였다. 해당 방법론은 텍스트와 라벨 간, 다른 라벨들과의 관계를 더 잘 반영하여 계층성을 더욱 효과적으로 탐색한다. 향후 해당 방법을 활용하여 의도 분류에 적용할 수 있다.

#### Acknowledgement

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과(IITP-2024-RS-2024-00438056)와 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2024-00419201)

#### 참 고 문 헌

[1] Yang, S., Liu, X., Dong, X., & Fu, B. (2024, June). Mini-DA: Improving Your Model Performance through Minimal Data Augmentation using LLM. In Proceedings of the Fifth Workshop on Data Science with Human-in-the-Loop (DaSH 2024) (pp. 25-30).

[2] Li, Z., Zhu, H., Lu, Z., & Yin, M. (2023). Synthetic data generation with large language models for text classification: Potential and limitations. Proceedings of the EMNLP. 2023

[3] Yu, Shuai, et al. "ACJIS: A novel attentive cross

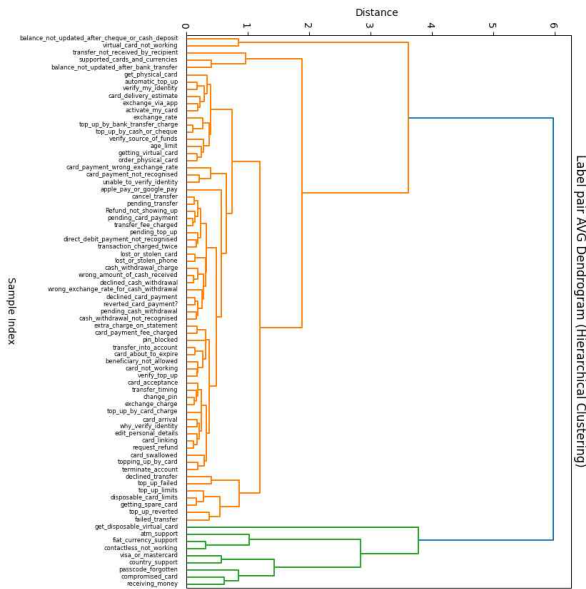


그림 3. 라벨별 multi-head attention score 평균 계층화  
Fig. 3. Label per average of Multi head Cross Attention Hierarchical Clustering

각 라벨별 평균화 계층 구조가 더 세밀하게 나뉘어져 있는 반면, 전체 평균화 계층 구조는 더 큰 범위로 그룹화 되어 있어, 대략적인 유사도만



- approach for joint intent detection and slotfilling." 2018 International Joint Conference on Neural Networks(IJCNN). IEEE, 2018.
- [4] Singhal, Bhavuk, Ashim Gupta, and AmrithKrishna. "IntenDD: A Unified Contrastive Learning Approach for IntentDetection and Discovery." Findings of the Association forComputational Linguistics: EMNLP. 2023.
- [5] Casanueva, I., Temčinas, T., Gerz, D., Henderson, M., & Vulić, I. Efficient intent detection with dual sentence encoders. Proceedings of the 2nd Workshop on Natural Language Processing for Conversational AI. 2020.
- [6] Ying, Cecilia, and Stephen Thomas."Label errors in BANKING77." Proceedings of the Third Workshop on Insights from Negative Results in NLP. 2022.
- [7] Zhang, Susan, et al. "Opt: Openpre-trained transformer language models." arXiv preprintarXiv:2205.01068. 2022.



# Unity-ROS를 활용한 가상 LIDAR 센서 개발

이세진\*, 유선진\*\*

## Development of Virtual LIDAR Sensors Using Unity-ROS

SeJin Lee\* and Sunjin Yu\*\*

### 요약

최근 무인 로봇의 수요가 증가함에 따라 모바일 로봇의 성능 평가와 테스트 효율성이 중요한 과제로 떠오르고 있다. 본 연구는 고가의 실제 센서 대신 Unity에서 가상 LIDAR 센서를 구현하여 테스트 비용을 줄이고 성능 평가를 효율적으로 수행하고자 한다. RPLIDAR의 S1 LIDAR 모델을 바탕으로 Unity와 ROS2를 연동하여 데이터를 검증했으며, 이를 통해 프로토타입 개발의 효율성을 높일 수 있었다.

### Abstract

As demand for autonomous robots increases, the need for efficient performance evaluation and testing of mobile robots has become crucial. This study aims to reduce testing costs and streamline performance assessment by implementing a virtual LIDAR sensor in Unity instead of using expensive physical sensors. Using the 2D LIDAR model of RPLIDAR S1, data was validated by integrating Unity with ROS2, enhancing prototype development efficiency.

### Key words

ROS, unity, Virtual Reality, LIDAR, sensor

## 1. 서론

최근 무인 로봇과 자율 시스템에 대한 관심이 급증하고 있으며, 이러한 기술은 물류, 자율 주행, 배송 등 다양한 산업 분야에서 활용되고 있다[1]. 무인 로봇은 공장 자동화부터 물류 창고 운영, 자율 배송 차량까지 폭넓게 사용되고 있다. 그러나 이러한 무인 로봇의 기능을 설계하고 성능을 평가하

는 과정에는 상당한 시간과 비용이 소요된다.

실제 장비와 센서는 고가이기 때문에, 반복적인 성능 검증이나 실험적 평가를 수행하는 것이 현실적으로 어려운 과제이다[2]. 무인 로봇의 IMU(Inertial Measurement Unit), LIDAR(Light Detection and Ranging), 그리고 카메라와 같은 센서들은 정밀한 주행에 필수적이지만, 초기 개발 단계에서 이러한 장비를 모두 갖추어 테스트하는 것은 비용 부담이 크다.

\* 국립창원대학교 문화융합기술협동과정 석사과정, 20237203@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다

따라서, 가상 환경에서 이러한 센서들을 구현하고 활용할 수 있는 방법에 대한 연구가 점점 중요해지고 있다.

## II. 데이터 비교를 통한 LIDAR 제작

본 연구에서는 Unity를 이용하여 가상의 센서 시스템을 개발하고, 이를 통해 실제 환경과 유사한 데이터를 얻을 수 있는 방법을 제안한다.

LIDAR는 2D LIDAR인 RPLIDAR사의 S1을 선정하였으며, 채택한 이유는 3D LIDAR에 비해 처리해야 할 정보량이 적어 성능 평가와 테스트 효율이 높기 때문이다. Unity Raycast 기능을 이용하여 S1 LIDAR의 Sample Rate, Range, Scan Rrate등 비슷한 스펙을 가진 가상 센서를 구현하였다.



그림 1. 가상현실 환경(위)과 실제 현실 환경(아래)  
Fig. 1. Virtual and Real-Life Environments

Unity에서 가상의 LIDAR 센서를 구현하고 ROS2와 연결하였다. 이를 위해 그림1에서 볼수 있듯이 실제와 유사한 환경을 Unity에서 구축하고, 가상 LIDAR 데이터를 수집하였다. 수집된 데이터는 RVIZ2에서 실제 S1 LIDAR의 포인트 클라우드와 비교하여 오차를 분석하고, 차이를 보이는 Raycast 값을 조정하여 정확도를 개선하였다. 이 과정에서 실물 LIDAR는 noise가 생성된다는 점을 확인하였다.

$$x_{\text{noisy}} = x_{\text{measured}} + \epsilon, \epsilon \sim N(0, \sigma^2) \quad (1)$$

이후, 가상 환경에서도 실제 환경과 유사한 센서 특성을 반영하기 위해 가우시안 노이즈(Gaussian Noise)를 추가하였다. 이러한 가우시안 노이즈 적용을 통해 현실적인 측정 오차를 가상 센서 데이터에 반영함으로써, 가상의 LIDAR 데이터를 실제 환경과 비슷하게 사용할 수 있다.

## III. 결 론

본 연구에서는 Unity를 활용하여 가상 LIDAR를 구현하고, 실제 환경에서의 데이터와 비교하였다. 가상 센서를 활용한 이 접근 방식은 개발 초기 단계의 비용 절감과 빠른 프로토타이핑을 가능하게 한다. 그러나 가상 환경에서는 실제 LIDAR의 반사율, 투과성, 간섭 등 물리적인 한계와 복잡성을 완벽하게 재현하지 못하는 단점이 있다. 이러한 한계를 극복하기 위해, IMU나 카메라와 같은 다른 가상 센서들도 구현하고 실적인 센서 모델링 기법을 도입하여 가상 센서의 정확도를 개선하는 방안을 모색할 예정이다.

## 참 고 문 헌

- [1] B. J. Yoon, and S.R. Yoo, "Mpleless Navigation Based on DQN Considering Moving Obstacles, and Training Time Reduction Algorithm" Journal of the Korea Institute of Information and Communication Engineering, 25.3: 377-383, 2021.
- [2] S. J. Jung and H. J. Hwang, "비대면 시대를 위한 물류로봇 기술 연구 동향". The Journal of The Korean Institute of Communication Sciences, 38(1), 21-26, 2020.

# RAG 시스템을 이용한 GPT 기반 무인 차량 VR 시뮬레이션

우동현\*, 유선진\*\*

## GPT-based Unmanned Vehicle VR Simulation using RAG System

Donghyun Woo\* and Sunjin Yu\*\*

### 요 약

최근 GPT와 같이 대규모 언어 모델의 활용이 VR 시뮬레이션에서의 사용자 경험 개선에 큰 잠재력을 보이고 있다. 그러나 LLM의 '환각' 현상으로 인해 잘못된 정보를 제공하여 시스템의 신뢰성을 저하시킨다. 특히, 군사 시뮬레이션과 같이 정확한 정보 전달이 중요한 분야에서는 환각 현상으로 문제를 발생시킬 수 있다. 본 연구에서는 RAG 시스템을 VR 시뮬레이션에 적용하여 LLM의 응답 정확성을 향상시키고자 하였다. 무인 차량 관련 데이터를 수집하고 전처리하여 데이터베이스를 구축한 후 Unity 엔진 기반의 VR 시뮬레이션에 통합하였다. 이를 통해 무인 차량의 상황 판단 및 응답 능력이 개선됨을 확인하였다.

### Abstract

The application of large language models, such as Generative Pre-trained Transformer(GPT), has shown significant potential for enhancing user experience in Virtual Reality(VR) simulations. However, the phenomenon of "hallucination" in Large Language Model(LLM) can lead to the provision of misinformation, undermining system reliability. This issue is particularly problematic in fields where accurate information delivery is critical, such as military simulations. In this study, we aimed to improve LLM response accuracy by applying a Retrieval-Augmented Generation(RAG) system to VR simulation. After collecting and preprocessing data related to unmanned vehicles to build a database, we integrated the database into a Unity engine-based VR simulation. This approach confirmed improvements in situation assessment and response capabilities of unmanned vehicles.

### Key words

unity, VR simulation, unmanned vehicle, RAG system,

## 1. 서 론

VR(Virtual Reality) 기술의 발전은 사용자에게 몰입감 높은 경험을 제공하며, 군사 시뮬레이션 분야

에서도 다양한 응용 가능성을 보여주고 있다. GPT(Generative Pre-trained Transformer)와 같은 대규모 언어 모델(Large Language Models, LLM)은 자연어 처리 분야에서 뛰어난 성능을 보인다. 그러나 LLM

\* 국립창원대학교 첨단방위공학과정, dhwoo1119@gs.cwnu.ac.kr,

\*\* 국립창원대학교 문화테크노학과, sjyu@changwon.ac.kr(교신저자)

※ 본 논문은 2023년도 국방기술품질원의 재원으로 방산혁신클러스터 지원(DCL2020L, 2020년 방산혁신클러스터 방산소재 부품연구실 사업)을 받아 수행된 연구의 일부임

은 '환각(Hallucination)' 현상으로 잘못된 정보를 제공할 수 있다. 이러한 현상은 군사 시뮬레이션과 같은 정보의 정확성이 필수적인 분야에서 큰 문제를 발생시킬 수 있다.

이러한 문제를 해결하기 위해, RAG(Retrieval-Augmented Generation) 시스템은 LLM의 응답 생성 과정에서 외부 데이터베이스를 참조하여 문제를 해결하는 방법으로 제시되고 있다. 본 논문에서는 RAG 시스템을 적용하여 무인 차량의 응답 정확성과 VR 시뮬레이션의 사용자 경험을 높이고자 한다.

## II. 시스템 설계 및 구현

무인 차량의 이동, 작전 수행, 상황 대응 등에 필요한 데이터 획득을 위해 학술 논문, 연구 보고서, 군사 시뮬레이션 등에서 자료를 수집하였다. 수집된 데이터는 한글 자연어 처리 라이브러리인 KoNLPy의 Okt(Open Korea Text) 모델을 활용하여 불필요한 조사와 어미를 제거하고 핵심 단어를 추출하였다. 전처리된 데이터는 인덱싱되어 LLM이 참조할 수 있는 데이터베이스로 구축되었다. 구축된 데이터베이스를 기반으로 RAG 시스템을 구현하여, 그림 1과 같이 입력된 질의에 대해 데이터베이스에서 관련 정보를 검색하고, 응답을 생성한다.

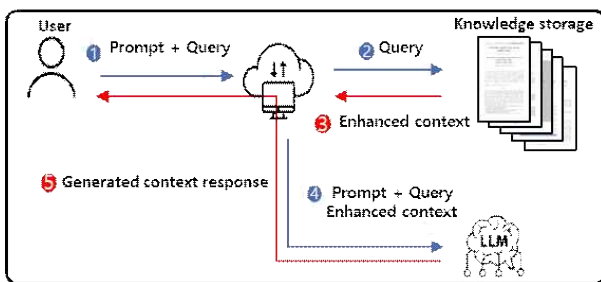


그림 1. LLM RAG 시스템 적용  
Fig. 1. Application of LLM RAG system

구현된 RAG 시스템은 Unity 엔진 기반의 VR 시뮬레이션에 통합되어 VR 환경에서 상호작용할 수 있도록 구현하였다. 무인 차량이 직면할 수 있는 상황을 반영하여 적의 발견 및 대응, 장애물 회피, 경로 재설정 등의 시나리오를 제작하였다. 그림 2에 보이는 것과 같이 제작된 시나리오는 무인 차량이

인식된 장애물을 회피하는 가상 상황을 포함하며, VR 환경에서 사용자 상호작용을 통해 시나리오의 변수를 조절할 수 있도록 구현되었다.



그림 2. 개발된 시스템을 적용한 무인 차량 시뮬레이션 구현  
Fig. 2. Implementation of an unmanned vehicle simulation applying the developed system

## III. 결 론

본 연구에서는 RAG 시스템을 VR 시뮬레이션에 적용하여 LLM의 응답 정확성을 향상시키는 방법을 제시하였다. 이를 통해 무인 차량의 상황 판단 및 응답 능력이 개선되었으며, 사용자 경험도 향상되었다. 향후 연구에서는 데이터 증강 기법을 적용하여 데이터베이스를 확장하고, 다양한 상황 데이터를 추가로 수집할 계획이다. 또한, 시스템의 성능 최적화를 통해 실시간 처리 능력을 향상시키고, 제스처 인식 및 음성 명령 등 다양한 상호작용 기술을 도입하여 사용자 몰입감을 더욱 높이고자 한다.

## 참 고 문 헌

- [1] D. M. Park and H. J. Lee. "Literature Review of AI Hallucination Research Since the Advent of ChatGPT: Focusing on Papers from arXiv." *Informatization Policy* 31.2 (2024): 3-38. <https://doi.org/10.22693/NIAIP.2024.31.2.003>
- [2] M. S. Jung, J. J. Song, S. H. Park and J. K. Moon, "A study on the military use of the new future battlefield environment metaverse." *The Journal of the Convergence on Culture Technology* 8.2 (2022): 179-185. <https://doi.org/10.17703/JCCT.2022.8.2.179>

# 전시상황을 고려한 VR 협업 민간인 대피 훈련 콘텐츠

강인주\*, 유선진\*\*

## VR Collaboration with Civilian Evacuation Training Contents Considering Exhibition Situation

Inju Gang\* and Sunjin Yu\*\*

### 요 약

최근 전쟁과 같은 비상 상황에서 공습으로 인한 민간인 피해가 증가하고 있다. 이러한 피해를 예방하기 위해 공습 시 효과적인 대피 방법을 이해하는 것이 중요하다. 본 논문은 VR을 활용한 민간인 협업 대피 콘텐츠를 제안하며, Unity 게임 엔진, XR Interaction Toolkit, PUN 2를 활용하여 멀티플레이 및 음성 채팅 기능을 설계하였다. 체험자들은 다양한 장애물을 협력하여 극복하며 대피소로 이동하고, 실시간 협력 임무를 수행한다. 향후 연구에서는 다양한 위기 상황을 반영한 시나리오 확장과 점수 평가 시스템을 개발할 계획이다.

### Abstract

In recent years, civilian damage caused by air strikes in emergency situations such as war is increasing. In order to prevent such damage, it is important to understand effective evacuation methods during air strikes. This paper proposes civilian collaborative evacuation content using VR, and designs multi-play and voice chat functions using the Unity game engine, XR Interaction Toolkit, and PUN 2. Experienced people overcome various obstacles by cooperating to move to the evacuation center and perform real-time cooperation missions. In future research, we plan to develop a scenario expansion and score evaluation system that reflects various crisis situations.

### Key words

Virtual Reality, XR Interaction Toolkit, multi-play

## 1. 서 론

최근 전쟁과 같은 비상 상황에서 공습으로 인한 민간인 피해가 증가하고 있다. 이러한 피해를 예방

하기 위해서는 공습 시 효과적인 대피 방법을 이해하는 것이 중요하다. VR(Virtual Reality)은 다양한 산업에서 교육과 훈련의 중요한 도구로 활용되고 있다. VR을 활용한 콘텐츠는 체험자가 독립적으로

\* 국립창원대학교 문화융합기술협동과정 석사과정, 20247306@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr(교신저자)

※ 본 논문은 2023년도 국방기술품질원의 재원으로 방산혁신클러스터 지원(DCL2020L, 2020년 방산혁신클러스터 방산 소재 부품연구실 사업)을 받아 수행된 연구의 일부임

행동할 수 있는 환경을 제공하여 흥미를 유발하고 학습 효과를 높인다[1]. 또한, 대피 과정에서 타인과의 협력이 필수적이다. 타인과 협력을 통해 자기 주도적 학습 능력뿐만 아니라 의사소통 능력까지 함께 발전시킬 수 있다[2]. 그래서 본 논문은 폭격 발생 시 VR을 활용한 민간인 협업 대피 훈련 콘텐츠를 제안하고자 한다.

## II. 시스템 구성 및 시나리오

### 2-1. 시스템 구성

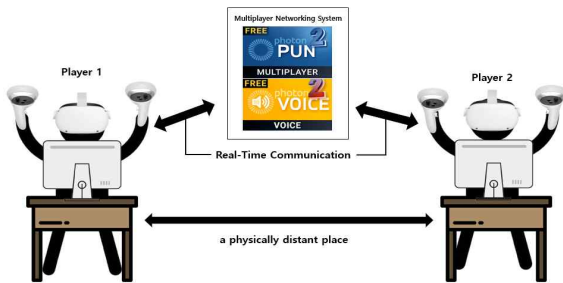


그림 1. 시스템 구성  
Fig. 1. System Configuration

본 콘텐츠는 그림 1과 같이 체험자가 물리적으로 떨어진 공간에 배치되고 체험자는 Meta의 Oculus Quest2 장비를 착용하여 체험한다. 또한 콘텐츠 환경은 Unity 게임 엔진과 Unity에서 제공하는 XR Interaction Toolkit을 활용하였다. 추가로 Unity 기반의 멀티플레이어 네트워킹 시스템인 PUN(Photon Unity Networking) 2를 사용하여 멀티플레이 및 보이스 채팅을 제공하도록 설계했다.

### 2-2. 시나리오

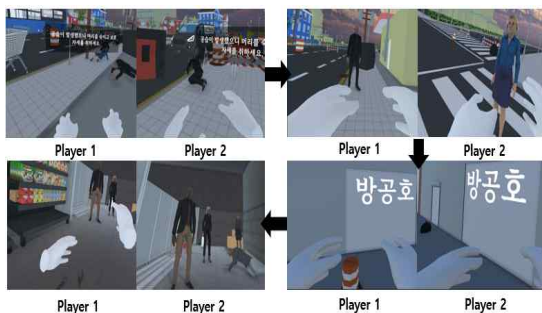


그림 2. 콘텐츠 시나리오 주요장면  
Fig. 2. Content Scenario

콘텐츠의 시나리오 장면은 그림 1에서 확인할 수 있다. Scene이 시작될 때, 도시에서 갑작스러운 공습경보가 울린다. 체험자는 머리를 숙이고 보호 자세를 취해야 하며 이때 각 팀원은 자신의 역할을 정하고 빠르게 대피 준비를 해야 한다. 대피소로 가는 경로는 여러 갈래로 나뉘어 있으며, 경로마다 붕괴한 구조물과 다양한 장애물이 있다. 이동 중에는 구조물을 들어 올리거나 부상자를 안전한 장소로 옮기는 등의 협력 임무가 주어진다. 대피소 근처에서 문이 잠겨 있거나 대피소에 접근할 수 없는 경우, 체험자들이 다 함께 문을 열거나 연기를 차단하는 등 협력 임무가 주어진다. 만약 모든 임무가 완료되면 체험자들은 대피소에 들어가야 한다. 만약 주어진 대피 시간이 초과하거나 협력이 실패하면 경고 메시지가 표시되고 성공적으로 대피하면 콘텐츠가 종료된다.

## III. 결론

본 논문은 폭격 발생 시 VR을 활용한 민간인 협업 대피 훈련 콘텐츠를 제안하였다. 이러한 콘텐츠는 협력을 통해 팀원 간의 상호작용 및 실제 상황에서 필요한 대처 능력을 효과적으로 훈련할 수 있다. 향후 연구에서는 다양한 위기 상황을 반영한 시나리오 확장과 점수 평가 시스템을 도입하여 발전시킬 예정이다.

## 참고 문헌

- [1] Yoo-Kyung Chung, Development of VR Fire-extinguishing Experience Education Contents Using UX Design Methodology, Journal of the Korean Content Society 17(3), pp. 222-230, 2017
- [2] Hye-Jin Park, A Case Study on the Effects of Cooperative learning by analysis of the reflection journals, Journal of Educational Innovation Research, 28(3), pp 225 - 250, 2018



# 메타버스 오피스에서의 행위 기반 사용자 식별 방법

강길욱\*, 구자훈\*, 김영갑\*\*

## Behavioral-Based User Identification Method in the Metaverse Office

Giluk Kang\*, Jahoon Koo\*, and Young-Gab Kim\*\*

### 요 약

메타버스(Metaverse)에 대한 관심이 증가함에 따라, 여러 메타버스 애플리케이션(application)이 게임을 넘어 재택근무, 보건의료, 국방 등 다양한 분야에 걸쳐 폭넓게 활용되고 있다. 특히, 코로나-19 팬데믹(COVID-19 pandemic) 이후 재택근무에 대한 근로자들의 선호도가 높아지면서 메타버스 오피스(office) 환경을 지원하는 애플리케이션에 대한 수요가 증가했다. 그러나 해당 애플리케이션은 시간과 장소에 구애받지 않고 업무를 수행할 수 있는 유연성을 제공하기 위해 많은 기밀 자산을 포함하고 있어 이러한 자산에 대해 적절한 사용자의 접근제어가 필요하다. 이에, 본 연구에서는 메타버스 오피스 환경에서 기밀 자산을 안전하게 보호하기 위한 사용자 인증 기술의 핵심 기술인 행위 기반 사용자 식별 방법을 제안한다. 제안 방법은 기계 학습을 기반으로 메타버스 오피스 환경에서 널리 사용되는 키보드 입력 시의 손 움직임 패턴을 분석하여 특정 사용자를 식별함으로써, 악의적 사용자의 비인가 접근을 탐지하고 차단하는 데 중점을 둔다.

### Abstract

With growing interest in Metaverse, various Metaverse applications have been utilized in many domains (e.g., game, remote work, healthcare, and military). In particular, the demand for applications that support Metaverse offices has increased after the COVID-19 pandemic because employees have increased their preference for remote work. However, this application requires access control to confidential assets, as the assets are included in the office to offer work flexibility without time and location constraints. Thus, this study proposes a novel behavioral-based user identification method that is a key technique for securely protecting confidential assets in the Metaverse office. The proposed method focuses on detecting and preventing unauthorized access by malicious users in the Metaverse office as it identifies the specific user using machine learning via hand movement patterns during keyboard input.

### Key words

user identification, Metaverse office, remote work, biometrics, machine learning

---

\* 세종대학교, giluk1027@sju.ac.kr, sigmao@sejong.ac.kr

\*\* 세종대학교, alwaysgabi@sejong.ac.kr(교신저자)

※ 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. RS-2024-00335672)

## 1. 서 론

메타버스(Metaverse)는 초월이라는 뜻을 가진 접두사인 메타(Meta)와 현실 세계를 의미하는 유니버스(universe)가 합쳐진 용어로, 여러 산업 전반에 혁신을 제공할 수 있는 잠재력을 인정받아 유망 기술로 주목받고 있다 [1]. 이러한 관심에 따라, 국외 빅테크(big-tech) 기업(예: Meta, Microsoft, Apple)도 메타버스를 위한 몰입형 기기와 애플리케이션을 개발하거나 관련 벤처기업에 투자를 꾸준히 진행하고 있다. 특히, 코로나-19 팬데믹(COVID-19 pandemic) 동안 국내외 기업이 재택근무를 장려함으로써 호라이즌 워크룸(Horizon workrooms)과 같은 많은 메타버스 오피스 애플리케이션(application)이 개발되고 활용되고 있다 [2]. 최근 팬데믹이 종료되었음에도 근로자들이 사무실 근무제를 거부함에 따라 기업들은 하이브리드(hybrid) 형태의 자율 근무의 형태로 여전히 메타버스 오피스 애플리케이션을 활용하고 있다.

그러나, 메타버스 오피스에는 시간과 공간의 제약 없이 업무 연속성을 보장하기 위해 다양한 기밀 문서(예: 지식재산권, 사업 전략 문서, 재무제표)가 저장되어 있어 산업스파이와 같은 악의적인 공격자가 접근할 시에 해당 문서가 유출되어 막대한 재산 피해를 야기된다 [3]. 또한, 공유 오피스의 개념 등과 함께 다수 기업이 업무 공간과 각종 기기를 공유함에 따라 특정 회사 근로자가 인증된 몰입형 기기를 잠시 두고 자리를 비운 사이에 다른 경쟁사의 직원이 몰래 기기를 착용하여 불법적으로 기밀 문서에 접근할 가능성이 있다. 이러한 기밀문서 유출을 방지하기 위해, 메타버스에서 사용자 인증 및 식별 기술을 제안하고 있지만 몇 가지 문제를 여전히 가지고 있다. 예를 들어, 지식기반 인증 방법은 대부분 비밀번호 기반 일회성 인증임에 따라 접속 세션이 상대적으로 긴 메타버스에는 적합하지 않으며 생체기반 인증도 값비싼 생체인식 센서가 요구되고 프라이버시(privacy) 우려도 꾸준히 제기되고 있다 [4,5]. 또한, 현재 제안된 행위 기반 인증 방법들은 게임 분야에 초점이 맞춰져 동적인 작업을 요구함에 따라 정적인 메타버스 오피스 환경에는 적합하지 않다. 이러한 이유로, 본 논문에서는 메타버

스 오피스에서의 행위 기반 사용자 식별 방법을 제안한다. 기존 제안된 연구들과 달리, 제안 방법은 값비싼 생체 센서가 필요하지 않으며 사용자 입력 행위를 기반으로 지속적으로 사용자의 행동 패턴을 분석하고 사용자를 식별함으로써 무자각 지속 인증으로 확장될 수 있다. 또한, 게임 분야를 넘어 메타버스 오피스를 채택하여 재택근무를 수행하는 다양한 분야에 사용할 수 있어 범용성이 높다는 장점이 있다. 본 논문은 다음과 같이 구성된다. 2장에서는 관련된 연구를 분석한다. 3장에서는 메타버스 오피스에서의 행위 기반 사용자 식별 방법을 제안한다. 4장에서는 제안 방법의 성능을 평가한다. 마지막으로, 5장에서는 본 연구의 결론에 관해 서술한다.

## II. 관련 연구

최근 메타버스의 관심에 꾸준히 증가함에 따라, 메타버스에서의 사용자 식별 및 인증 연구가 광범위하게 제안되었다. 대표적으로, Bu et al. [6]는 3D 원형 구체 기반 사용자 인증 방법을 제안했다. 본 인증 방법은 몰입형 기기의 컨트롤러(controller)를 이용하여 구체를 돌려가며 사용자 패스워드를 입력하는 방식으로 일회성 인증으로만 사용 가능하다는 단점이 있다. Li et al. [7]는 뇌파 기반 사용자 생체 인증 방법을 제안했다. 그들의 방법은 몰입형 기기에 뇌파 측정을 위한 추가적인 센서를 부착해야 하며, 인증을 위해 비디오 시청을 하는 동안 눈을 깜박이거나 머리 움직임이 제한되는 단점이 있다. LaRubio et al. [8]는 시선 행위 기반 사용자 인증 방법을 제안했다. 그들은 무작위로 위치되는 점 표시를 보도록 하거나 이미지를 보게 하는 등의 작업에서의 사용자 시선 움직임을 통해 사용자를 인증함으로써 일반적인 오피스 환경에 적합하지 않다. Li et al. [9]는 공 던지기 작업 기반 사용자 인증 방법이 제안했다. 본 방법은 공 던지는 초기 동작을 기반으로 향후 동작을 예측하여 사용자를 인증한다. 그러나, 이는 여전히 정적인 오피스 환경에 적합하지 않으며 공 던지는 초기 동작은 사용자끼리 유사할 가능성이 있으며 단순한 동작만을 사용함에 따라 모방 공격에 취약하다.

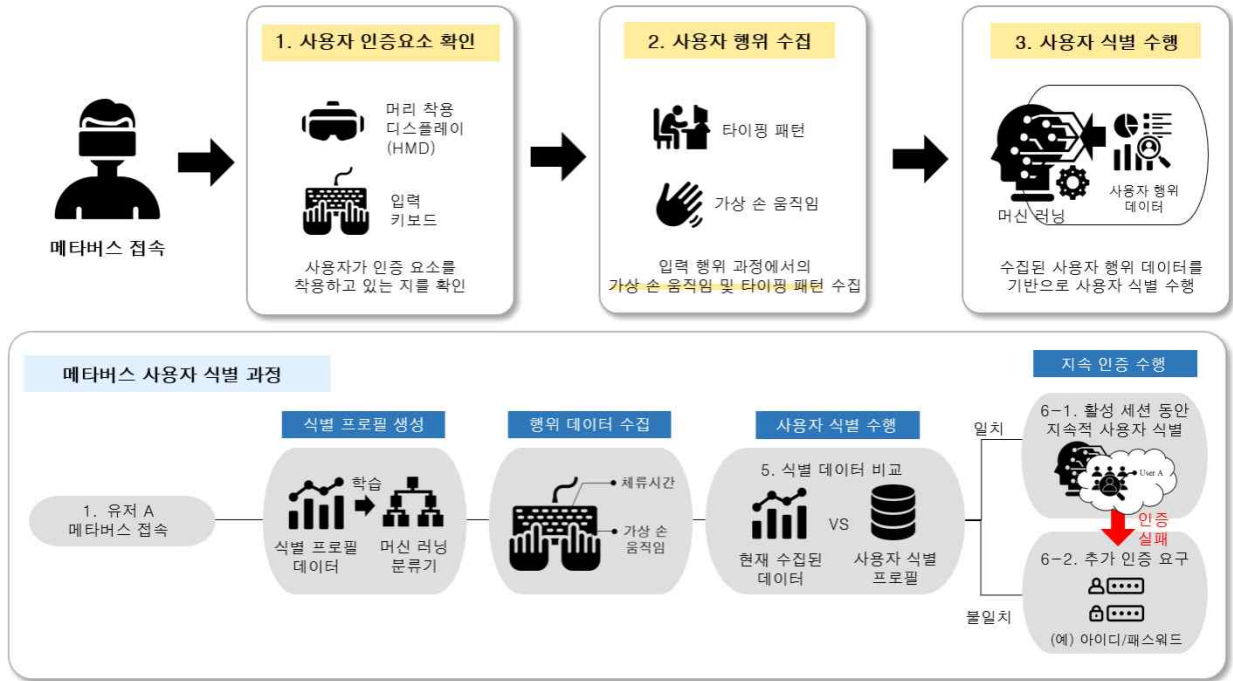


그림 1. 메타버스 오피스에서의 사용자 식별 방법  
 Fig. 1. Behavioral-based user identification method in the Metaverse office

### III. 메타버스 오피스에서의 행위 기반 사용자 식별 방법

본 논문에서 제안하는 방법은 국방, 보건의료 등과 같은 여러 분야에서 범용적으로 활용되는 메타버스 오피스에서 사용자 키보드 입력 시의 가상 손 움직임 행위를 기반으로 사용자를 식별한다. 그림1은 제안 방법의 전체 개요를 보여주며 자세한 동작 방식은 아래와 같다. 먼저, 유저 A가 메타버스에 접속하면 사용자 식별 프로필을 생성하기 위해 특정 텍스트를 2회 반복하여 입력한다. 특정 텍스트가 입력되는 과정에서 사용자의 가상 손 움직임 위치와 키 체류시간(dwell time)이 수집되고, 입력이 끝난 후에 기계 학습 분류기를 통해 해당 데이터가 학습되게 된다. 학습이 완료된 후, 유저 A는 회사 업무를 수행하기 위해 키보드 입력을 수행하는 과정에서 계속해서 가상 손 움직임과 키 체류시간이 수집된다. 그런 다음, 수집된 데이터가 학습된 분류기를 통해 기존 식별 프로필과 지속적으로 비교됨으로써 사용자 식별이 수행된다. 만약, 비교 과정에서 정확도가 관리자가 정의한 수준 이하로 떨어지게 되면 사용자 재인증이 요구된다.

### IV. 실험

본 논문에서 제안한 사용자 식별 방법의 유효성을 입증하기 위해, 15명의 참여자를 모집하여 데이터 수집을 진행하고, 이 중 무작위로 선정된 4명에 대해선 다른 날에도 데이터를 추가 수집하여 실험을 수행했다. 해당 실험은 세종대학교 윤리위원회를 통해 심의면제(심의면제번호:SUIRB-HR-E-2024-014) 되었으며, 실험에 앞서 모든 참가자는 실험 내용 및 과정에 대해 숙지했다.

본 실험은 스티브 잡스 연설문의 일부를 사용자 식별 프로필 생성과 실제 업무를 위한 텍스트로 채택했다. 즉, 학습을 위한 데이터와 검증을 위한 데이터를 동일하게 채택했다. 수집 데이터는 position, normal, direction, quaternion이 관절마다 수집되었으나 position만을 사용했으며, 수집 과정에서 오른손과 왼손이 같이 수집된 경우에만 사용했다. 기계 학습 분류기는 랜덤 포레스트(random forest)를 사용했다. 실험 결과에 따르면, 무작위로 선정된 참가자 5명은 그림2에서 보듯이 모두 80% 이상 정확도를 가졌으며 참가자 3명은 90% 이상의 정확도를 가짐

에 따라 제안 식별 방법이 메타버스 오피스에서 적절히 사용자를 식별할 수 있음을 보여준다.



그림 2. 특정 사용자별 식별 정확도  
Fig. 2. Identification accuracy for specific users

#### IV. 결 론

본 논문에서는 기존에 제안된 메타버스에서의 사용자 식별 및 인증 기술의 문제점을 해결하기 위한 메타버스 오피스에서의 행위 기반 사용자 식별 방법을 제안했다. 제안 방법은 메타버스 오피스에서 가장 흔히 사용되는 키보드 입력 행위를 활용하여 사용자를 식별함으로써 무작위 지속 인증으로 확장될 수 있으며, 무작위로 선정된 5명의 사용자에 대해 80% 이상 정확도를 보여줌에 따라 충분히 사용자 식별 방법으로써 활용이 가능할 것으로 판단된다. 향후 연구에서는 다양한 특징 세트를 구성하여 추가적인 실험을 진행하고자 한다.

#### 참 고 문 헌

[1] 이병권, "메타버스(Metavesre)세계와 우리의 미래", 한국콘텐츠학회지, 19권, 1호, pp. 13-17, 6월 2021년.  
[2] 류성한, 윤혜정, 박재현, 장영훈, "메타버스 개념 및 현황에 대한 논의와 향후 연구 방향 제안", 지식경영연구, 23권, 2호, pp. 1-13, 6월 2022년.  
[3] G. Kang, J. Koo, and Y. G. Kim, "Security and Privacy Requirements for the Metaverse: A Metaverse Applications Perspective", IEEE Commun. Mag. vol. 62, issue 1, pp. 148-154, 2024.

[4] H. Jang, S. Park, J. Woo, J. Ha, and L. Kim, "Authentication System Based on Event-related Potentials Using AR Glasses", Proc. BCI, pp. 1-4, 2023.  
[5] D. Progonov, H. Naumenko, O. Sokol, and V. Derkach, "User Authentication on Headset-Like Devices by Bioacoustic Signals", Proc. ETAA, pp. 31-47, 2023.  
[6] B. Zongyoug, Z. Haojun, X. Weiqiang, Z. Yi, L. Zitao, L. Weiqi, and G. BoYu, "Secure Authentication with 3D Manipulation in Dynamic Layout for Virtual Reality", Proc. VRW, pp. 955-956, 2023.  
[7] S. Li, S. Savaliya, L. Marino, A. M. Leider, and C. C. Tappert, "Brain signal authentication for human-computer interaction in virtual reality", Proc. CSE and EUC, pp. 115-120, 2019.  
[8] K. LaRubbio, J. Wright, B. David-John, A. Enqvist, and E. Jain, "Who do you look like? - Gaze-based authentication for workers in VR", Proc. VRW, pp. 744-745, 2022.  
[9] M. Li, N. K. Banerjee, aS. Banerjee "Using Motion Forecasting for Behavior-Based Virtual Reality (VR) Authentication", Proc. AlxVR, pp. 31-40, 2024.

# 모터 전류 신호를 활용한 양방향 LSTM 기반 이상 탐지 알고리즘

김도윤\*, 김해영\*, 반재필\*\*, 구교권\*\*\*

## Bidirectional LSTM Based Anomaly Detection Algorithm Using Motor Current Signal

Doyun Kim\*, Haeyoung Kim\*, Jaepil Ban\*\*, and Gyogwon Koo\*\*\*

### 요 약

대부분의 산업 기계는 유도 모터를 원동기로 사용하기 때문에 고장을 진단하는 것이 중요하다. 이전까지는 산업현장에서 여러 기계 학습 기법을 이상 탐지에 활용했다. 이 방법은 제한된 데이터 셋을 사용하고, 센서의 진동 데이터를 사용한다는 한계점이 있다. 본 논문에서는 이러한 한계를 극복하기 위한 딥러닝 모델로 양방향 LSTM을 사용했다. 먼저, 사용한 모델과 전류 데이터 셋에 대해 설명하고, 사용한 모델의 성능에 대한 실험 결과를 보여준다. 제안한 알고리즘은 AUC(Area Under the ROC Curve)가 0.83으로 이상 탐지를 수행하는 것을 보여준다.

### Abstract

Fault diagnosis is important because most industrial machines use induction motors as their driving force. Previously, various machine learning techniques were applied for anomaly detection in industrial settings. However, these methods have limitations, such as relying on limited data sets and using only vibration data from sensors. This paper addresses these limitations by employing a bidirectional LSTM model within a deep learning framework. First, it explains the utilized model and current dataset, followed by experimental results demonstrating the performance of the proposed model. The proposed algorithm demonstrates anomaly detection with an AUC (Area Under the ROC Curve) of 0.83

### Key words

deep learning, anomaly detection, induction motor, bidirectional LSTM, current dataset

### 1. 서 론

이상 탐지 및 고장 진단 방법이 많은 다양한 분야에서 점점 더 중요해지고 있다. 이는 특히 항공기, 기차, 자동차, 발전소 및 화학 공장과 같은 안전과 신뢰성, 안전성 및 효율성을 개선하기 위해 모터

\* 국립금공과대학교 전자공학과 석사과정, doo207@kumoh.ac.kr, sunh@kumoh.ac.kr

\*\* 국립금공과대학교 전자공학부 교수, banjp@kumoh.ac.kr

\*\*\* 대구경북과학기술원 지능형로봇연구부 선임연구원, gkoo@dgist.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부에서 지원하는 DGIST 기관고유사업에 의해 수행되었습니다 (24-IT-02)

밀접하게 관련된 분야에 해당한다[1]. 유도 모터는 모든 회전 기계의 핵심적인 구성 요소로, 산업 기계의 약 90%가 원동기로 사용할 만큼 매우 중요한 구동 장치이다. 따라서 고장과 생산 중단을 방지하기 위해 유도 모터의 상태를 예측하는 것이 필요하다. 유도 모터의 고장은 상당한 재정적 손실, 예기치 않은 정지, 관련 수리 비용으로 인해 큰 부담이 될 수 있다[2].

이전까지 산업 현장에서는 다양한 기계 학습 기법을 이상 탐지에 활용했다[3,4]. [3]은 kNN (Nearest Neighbor)과 SVM (Support Vector Machine) 알고리즘을 모터 진동 데이터에 적용해 고장 징후를 탐지한 경우가 있다. 이 모델들은 각각 95% 이상의 정확도로 높은 성능을 보여, 머신러닝을 통해 모터 진동 데이터를 효과적으로 분석해 고장 징후를 예측할 수 있음을 입증했다. 다만, 이 방식은 추가적인 진동 센서가 필요하다는 한계가 있다. [4]는 함수 기반 및 트리 기반 알고리즘을 사용해 철도차량 모터의 고장을 진단한 경우가 있다. 이 방법은 진폭, 주파수, RPM, 등의 센서 데이터를 활용해 기계 학습 모델로 고장을 분류하였으며, 특히 랜덤포레스트 알고리즘에서 가장 높은 정확도를 보였다. 이는 기계 학습이 모터 이상 탐지에 효과적임을 나타내지만, 첫 번째 사례와 같이 데이터 셋이 제한적이라는 단점이 존재한다. 요약하자면, 기존 기계 학습 연구는 공통적으로 제한적인 데이터 셋과 추가적인 센서의 진동 데이터만 사용한다는 한계가 있다.

이러한 한계를 극복하기 위해, 모터의 전류 데이터를 이용해 모터의 고장을 진단하는 딥러닝 기법을 제안한다. 모터의 진동 신호 계측할 때는 가속도 진동 센서가 필요하지만, 전류 데이터는 전류 입력 모듈만 있다면 계측이 가능하기 때문에 추가 센서 없이 데이터를 수집할 수 있어 진동 데이터보다 경제적이다[5]. 딥러닝은 기계 학습과 달리 인공 신경망을 기반으로 다양한 층을 통해 데이터를 자동으로 분석하고 처리할 수 있다. 따라서 본 논문에서 제안하는 모터 이상탐지 알고리즘은 딥러닝 모델 중 데이터에서 더 많은 정보를 추출 가능한 장점이 있는 양방향 LSTM(Long Short-Term Memory)을 사용하여 모터의 이상 탐지를 수행하고 사용한 모델

의 성능에 대한 실험 결과를 보여준다.

본 논문의 구성은 다음과 같다. 모터 이상 탐지를 위한 인공신경망을 보여주고, 제안한 인공신경망의 실험 결과를 보여준다. 마지막으로 결론을 말한다.

## II. 모터 이상 탐지를 위한 인공신경망

이 논문에서는 모터의 이상 탐지를 위한 딥러닝 모델로 양방향 LSTM을 사용했다. 양방향 LSTM은 기울기가 사라지는 문제를 극복하고 장기 의존성을 학습할 수 있는 구조이다. 순환신경망의 장기 의존 문제를 개선하기 위해 LSTM은 기억 셀을 추가하며, 이 기억 셀에는 입력, 망각, 출력 게이트를 포함한다. 이를 통해 내부의 상태가 여러 단계에 걸친 오류를 기억하는 데 도움을 준다.

LSTM 네트워크는 입력 시퀀스와 출력 시퀀스 간의 매핑을 계산한다[6]. 입력 시퀀스는  $X = (X_1, X_2, \dots, X_n)$ 이고, 출력 시퀀스는  $y = (y_1, y_2, \dots, y_n)$ 이다. 여기서  $n$ 는 데이터의 갯수를 의미한다. 이는 다음 방정식을 통해 계산된다.

$$O_F = \sigma(W_F X_t + W_h h_{t-1} + b_F) \quad (1)$$

$$O_I = \sigma(W_I X_t + W_h h_{t-1} + b_I) \quad (2)$$

$$O_O = \sigma(W_O X_t + W_h h_{t-1} + b_O) \quad (3)$$

$$C_t = C_{t-1} \otimes O_{F_t} + O_{I_t} \quad (4)$$

$$\otimes \tanh(W_C X_t + W_h h_{t-1} + b_C)$$

$$h_t = O_O \otimes \tanh C_{t-1} \quad (5)$$

식 (1-5)에서  $W_F, W_I, W_O, W_h, b_F, b_I, b_O, b_C$ 는 각각 입력, 망각, 출력 게이트와 기억 셀의 가중치와 바이어스를 나타낸다.  $h_{t-1}$ 은 이전 숨겨진 층의 유닛을 의미하며, 세 개의 게이트 가중치와 요소별로 더해진다. 식 (4)의  $C_t$ 는 현재 기억 셀 유닛이 된다. 식 (5)는 이전 숨겨진 유닛 출력과 이전 기억 셀 유닛의 요소별 곱셈을 나타낸다. 입력 시퀀스  $X$ 는  $h_{F,t} = (h_{F,1}, h_{F,2}, \dots, h_{F,n})$ 으로 순방향으로 계산되고,  $h_{B,t} = (h_{B,1}, h_{B,2}, \dots, h_{B,n})$ 으로 역방향으로 계산된다. 최종 출력 시퀀스  $y$ 는  $h_{F,t}$ 과  $h_{B,t}$ 에 의해 형성된 각 기억 셀의 최종 출력  $y_t$ 로 나타난

다. 세 개의 게이트에 대해 비선형성을 추가하는 활성화 함수로 출력 범위가 -1에서 1 사이로 확장된 시그모이드 함수인  $\tanh$ 와 시그모이드 곡선을 갖는  $\sigma$ 가 식 (1-5)에 표시되어 있다. 여기서,  $(t-1)$ 과  $t$ 는 각각 이전과 현재의 단계를 나타낸다. 이 모델은 이전 시점과 미래 시점의 정보를 모두 활용할 수 있어 전체 데이터로부터 효과적으로 학습할 수 있도록 도와준다. 따라서 본 논문에서는 순방향과 역방향의 정보를 추출해서 이용할 수 있는 이상 탐지 알고리즘으로 양방향 LSTM을 채택하여 제안한다.

### III. 실험

#### 3.1 데이터 셋

본 논문에서는 모터 전류 데이터만을 활용하여 이상을 탐지하기 위해 [7]의 데이터를 활용했다. 모터는 정격 속도의 50, 75, 100%로 계측했다. 실험에서는 모터의 고장자 단락을 이상 탐지에 사용했다.

학습 데이터 셋과 검증 데이터 셋을 8:2의 비율로 나누어서 학습을 진행했으며, 윈도우 슬라이딩 기법을 사용하여 윈도우 길이를 800으로 설정했다. 정규화 방법으로는 최소-최대 스케일러를 사용하여 값을 0과 1사이로 변환했다. 최종적으로, 학습 데이터 셋은 67,500개의 정상 데이터로 구성하였으며, 평가 데이터 셋은 39,375개의 정상 및 이상 데이터로 구성하였다.

#### 3.2 실험 설정

전류 데이터만을 활용하여 모터 이상 탐지 실험을 하였다. 실험을 진행한 환경은 jupyter notebook을 사용하였고, 하이퍼 파라미터는 손실함수는 MSE, 최적화 기법은 Adam을 사용하였다.

제안한 이상 탐지 알고리즘의 성능을 검증하기 위해 배치 크기와 학습률을 표 1, 2와 같이 변경하여 실험을 진행하였다. 임계값은 원본 데이터의 실험 데이터 셋과 모델이 예측한 실험 데이터 셋의 차이를 이용하여 설정하여 이상 탐지의 기준으로 사용하였다.

표 1. 학습률 실험

Table 1. Learning rate experiment

Learning rate	AUC
0.01	0.8281
0.001	0.8288
0.0001	0.8119
0.00001	0.6858
0.000001	0.4988

표 2. 배치 크기 실험

Table 2. Batch size experiment

Batch size	AUC
64	0.6459
128	0.7668
256	0.8288

#### 3.3 실험 결과

학습률에 대한 실험의 경우, 0.001일 때 AUC(Area Under the ROC Curve)가 0.8288로 가장 좋은 성능을 보였다. 학습률이 감소함에 따라 AUC가 낮아지는 경향을 볼 수 있었다. 다음, 배치 크기에 대한 실험은 256일 때 AUC가 0.8288로 가장 좋은 성능을 보였다. 각 실험에 대해 배치 사이즈가 증가함에 따라 AUC가 높아지는 경향을 볼 수 있었다. 따라서 제안한 알고리즘은 학습률 0.001과 배치 사이즈가 256일 때 가장 좋은 성능을 보여주는 것을 확인했다.

### IV. 결론

본 논문에서는 모터의 전류 신호만을 활용하여 이상 탐지를 하기 위해 딥러닝 모델 기반의 양방향 LSTM 모델을 제안했다. 이 모델은 입력, 망각, 출력 게이트를 포함하여 양방향 정보를 활용할 수 있다. 모터 전류 데이터만을 활용해 윈도우 슬라이딩 기법과 최소-최대 스케일러 기법으로 정규화한 데이터를 실험에 활용했다. 마지막으로 제안한 알고리즘은 실험을 통해 선정된 파라미터로 학습을 진행했을 때, 모터의 이상 탐지를 정상적으로 수행하는 것을 볼 수 있었다.

### 참고 문헌

- [1] Isermann, Folf. "Model-based fault-detection and diagnosis-status and applications.", Annual Reviews in control, vol. 29, pp. 71-85
- [2] Gundewar, Swapnil K., and Prasad V. Kane. "Condition monitoring and fault diagnosis of induction motor.", Journal of Vibration Engineering & Technologies, vol. 9, pp. 643-674
- [3] 변준호. 변시우, "기계 학습을 활용한 모터 진동의 고장 징후 분석", 한국산학기술학회, vol. 24, pp. 710-713
- [4] 정훈. 김주원, "기계적 모터 고장진단을 위한 머신러닝 기법", Journal of Korean Society of Industrial and Systems Engineering, vol. 40, pp. 57-64
- [5] J. -H. Han and S. -U. Park, S. -K. Hong, "A Study of Vibration and Current Data Characteristic Analysis for Motor Mechanical Fault Level Determination by Deep Learning", The Transactions of the Korean Institute of Electrical Engineers, vol. 70, pp. 1509-1516
- [6] Shahid, Farah, Aneela Zameer, and Muhammad Muneed. "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM", Chaos, Solitons & Fractals, vol. 140, 110212
- [7] Bruinsma, Sietze, et al. "Motor current and vibration monitoring dataset for various faults in an E-motor-driven centrifugal pump", Data in brief, vol. 52, 109987



# EfficientAD 기반 환편기 편물 검단 모델의 경량화 연구

황병일\*, 김다현\*, 여동훈\*, 최상민\*, 서영주\*, 김동주\*\*

## A Study on the Lightweighting of Anomaly Detection Model for EfficientAD-Based Circular Machine Fabric

Byeong Il Hwang\*, Da Hyun Kim\*, Dong Hun Yeo\*, Sang Min Choi\*, Young Joo Suh\*, and Dongju Kim\*\*

환편기는 회전하는 원형 배열의 바늘로 편물을 생산하는 위편기 기계로, 기름 누출이나 부품의 노후화와 같은 구조적 결함 외에도 먼지, 바람 등 외부 환경 요인으로 인해 편물에 결함이 발생할 수 있다[1]. 현재 재직 현장에서는 일정 길이의 편물이 만들어진 후 숙련된 작업자가 육안으로 결함을 검출하는 방식이 사용되고 있으나, 이러한 방법은 작업자의 피로도에 따라 결함을 놓치는 경우가 생기며, 특히 위·경사 결함이 일정 길이 이상 누적될 경우 생산된 편물을 모두 폐기해야 하는 경제적 손실을 초래할 수 있다. 이러한 문제를 해결하고자 본 연구에서는 자동화된 결함 검출을 통해 편물의 품질을 실시간으로 보장하고 생산 과정에서 발생하는 경제적 손실을 줄이는 것을 목표로 하였다.

본 연구에서는 EfficientAD 모델을 기반으로 환편기에서 생성된 원단을 실시간으로 촬영하고 분석하여 결함을 자동으로 탐지하는 모델을 개발하였다. EfficientAD는 높은 정확도와 빠른 추론 속도를 제공해 실시간 결함 탐지에 적합한 딥러닝 모델로, 초기 실험에서 AUROC 99.86%와 100 FPS의 추론 속도를 확보하였다. 그러나 환편기의 제한된 공간과 비용 문제로 인해 mini-PC 사양 또는 온-디바이스 환경에서 운영 가능하도록 모델의 경량화가 요구된다.

이를 위해 본 연구는 EfficientAD 모델의 경량화를 위한 가지치기와 양자화 방법을 적용하였다. 가지치기의 경우 중요도가 낮은 가중치를 제거하는 비구조적 가지치기를 사용하여 모델의 복잡도를 줄였다. 비구조적 가지치기 적용 후 AUROC는 99.24%, FPS는 82.2로 다소 성능 저하가 있었으나 모델 크기가 줄어드는 효과가 있었다. 또한, NVIDIA TensorRT 최적화 엔진을 활용한 FX 양자화를 통해 모델을 최적화한 결과, AUROC 99.42%, FPS 368.4를 기록하여 검출 성능은 소폭 감소했으나 추론 속도가 3.68배 향상되는 경량화 성능을 확인할 수 있었다. 이처럼 FX 양자화는 현장에 적용하기에 충분한 수준의 경량화와 실시간 성능을 제공함을 실험을 통해 입증하였다.

본 연구는 환편기에서의 편물 결함 자동 검출 모델을 경량화함으로써 기존 수작업 대비 결함 탐지율을 크게 향상시키고, 편물 생산 과정에서 품질 관리를 더욱 효과적으로 수행할 수 있는 가능성을 제시한다. 이를 통해 결함 탐지의 자동화가 생산 품질과 경제적 손실을 줄이는 데 기여할 수 있기를 기대한다. choi sang min

### Key words

Anomaly Detection, EfficientAD, Lightweight, Pruning, Quantization

### 참고문헌

- [1] 김다현, “자기 지도 학습 기반 환편기 원단 결함 실시간 고속 탐지 시스템 개발”, 2024년 한국통신학회 동계종합학술발표회
- [2] K.Batzner, L.Heckler and R.Konig, “EfficientAD: Accurate Visual Anomaly Detection at Millisecond-Level Latencies”, Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024, pp. 128-138

\* 포항공과대학교 인공지능연구원, {bihwang, kdhyun8011, tgc01098, bestmin, syjsuh}@postech.ac.kr,

\*\* 포항공과대학교 인공지능연구원(교신저자), kkb0320@postech.ac.kr

※ 본 논문은 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업이며 (No.2022R1A6A1A03052954), 정부(과학기술정보통신부)의 재원으로 한국지능정보사회진흥원의 지원을 받아 수행된 연구이며(2024년 지역특화형 초거대 AI 클라우드 팜 실증 및 확산 환경 조성), 「과학기술정보통신부 방송통신발전기금」을 지원받아 정보통신산업진흥원(DI014-24-1009, 철도차량 가용성 향상을 위한 AI 융합형 통합정비 공정 메타버스 플랫폼 개발)의 지원을 받아 수행된 연구이나 과학기술정보통신부의 공식의견과 다를 수 있음.



# 진화형 텍스트-음성 변환 : 데이터 축적 기반 음성 합성 모델 진화 학습 방법

김민재\*, 고병진\*\*, 윤종완\*\*, 박태준\*\*

## Evolving TTS : Evolutionary Training Text-to-Speech Model based on Cumulative Data

Minjae Kim\*, Byungjin Ko\*\*, Jongwan Yoon\*\*, and Taejoon Park\*\*

### 요 약

본 연구는 사용자와 인공지능의 상호작용을 통해 단계적으로 목표 화자의 음성을 합성하는 모델인 Evolving Text-to-Speech (Evolving TTS)를 제안한다. Evolving TTS는 수 초의 음성으로 학습 과정 없이 추론이 가능한 Zero-shot 학습 기법과, 전이 학습 기반의 Voice Cloning 기법을 병렬적으로 적용하여 짧은 시간 안에 결과를 확인할 수 있고, 이후 음성 인식을 통해 누적된 목표 화자의 음성을 이용한 단계적 학습으로 높은 품질의 음성을 합성할 수 있다. Evolving TTS는 새로운 학습 기법을 제시함으로써, 음성 합성 분야의 궁극적 목표인 출력 음성 품질의 향상과 학습 소요 시간 단축 사이의 조율을 통해 사용자에게 인공지능 TTS에 대한 접근성과 편의성을 제공함으로써 맞춤형 음성 비서의 상용화를 기대할 수 있다.

### Abstract

This study proposes Evolving Text-to-Speech (Evolving TTS), a model that synthesizes the target speaker's voice gradually through interaction between the user and artificial intelligence. Evolving TTS applies a zero-shot learning technique that enables inference without training using a few seconds of speech, and a transfer learning-based voice cloning technique in parallel, allowing quick results. Subsequently, stepwise learning using accumulated target speaker's speech through speech recognition enables the synthesis of high-quality voices. By introducing a new training approach, Evolving TTS aims to balance the enhancement of output speech quality and the reduction of training time, providing users with accessibility and convenience to AI-based TTS, ultimately paving the way for the commercialization of personalized voice assistants.

### Key words

evolving TTS, zero-shot learning, voice cloning, speech recognition

---

\* 한양대학교, flom97@hanyang.ac.kr,

\*\* 한양대학교에리카, byungjinko@hanyang.ac.kr(교신저자) jongwanyoon@hanyang.ac.kr(교신저자),  
taejoon@hanyang.ac.kr(교신저자)

※ 이 논문은 정부(산업통상자원부)의 재원으로 한국산업기술기획평가원(KEIT)의 지원을 받아 수행된 (1415187250, 100%) 연구임

## I. 서 론

음성 합성(Speech Synthesis)은 목표하는 화자의 음성을 기계를 통해 합성하는 기술로, 발화의 자동화를 통해 사용자에게 편의성을 제공해주는 기술이다. 인간의 음성을 인공적으로 생성하는 기술은 음성 비서, 내비게이션, 교육용 도구, 장애인을 위한 보조 기술 등 다양한 분야에서 응용 가능성이 높아 주목받고 있다.

딥러닝 기반의 음성 합성 기술은 인간의 목소리를 더 자연스럽고 유창하게 모방하는 것을 목표로 많은 연구가 이루어져 왔다. 수많은 연구 과정을 통해 현재는 인간의 목소리와 거의 흡사하며 매우 자연스러운 음성까지 생성이 가능하게 되었다. 그러나 목표로 하는 화자의 자연스러운 음성을 생성하기 위해서는 텍스트-음성 쌍으로 이루어진 데이터가 많이 필요하다.

음성 데이터의 복잡한 전처리 과정과 학습에 소요되는 시간을 줄이기 위한 학습 기법 중 하나인 Zero-shot 학습 기법은, 별도의 학습 과정 없이 화자 인코더를 통해 목표 화자의 음성을 분석하고 이를 추론 과정만으로 목표 화자의 음성을 합성하는 기법이다. 이러한 방법은 전사된 데이터 없이 음성만으로 빠른 시간 안에 비슷한 목소리를 합성할 수 있으나 학습 과정을 거치지 않았기 때문에 결과 음성의 품질이 다소 떨어지며, 사전학습 모델의 성능에 따라 결과가 차이가 심하다는 단점이 있다.

다른 방법으로 Voice Cloning 또한 연구되었는데, 이는 특정 인물의 음성을 학습한 모델을 사용하여 목표하는 화자의 음성으로 복제하는 전이 학습 기법을 사용한 기술이다. Zero-shot Learning과는 다르게 목표 화자의 음성에 대한 학습 과정을 필요로 하며, 음성 품질이 목표 화자의 데이터양에 따라 달라진다는 차이가 있다. Voice Cloning은 음성 데이터가 학습 과정을 거치기 때문에 합성된 음성의 품질 향상을 기대할 수 있으나, 기존 사전학습된 화자의 음성, 억양 등이 남아있기 때문에 온전히 새로운 화자의 음성을 합성하기 위해서는 학습 데이터의 양과 품질이 보장되어야 한다는 단점이 있다.

본 연구에서는 학습 시간을 단축함과 동시에, 합

성된 음성의 품질을 점진적으로 높일 수 있도록 하며, 수집된 데이터를 기반으로 모델이 학습됨에 따라 중간 결과를 사용자로 하여금 확인할 수 있게 하여 사용자와 모델이 상호작용하는 새로운 학습 기법을 제안한다.

초기 단계의 모델은 적은 양의 데이터를 Zero-shot Learning을 통해 빠르게 학습된 모델을 사용자에게 제공한다. 이후, 음성 인식을 통해 사용자의 데이터를 누적해가며, 이를 본 모델의 학습을 통해, 다소 품질이 떨어지는 음성을 합성하는 모델이 개선되어가는 과정을 사용자가 확인할 수 있다.

이후 누적된 데이터 양에 따라 학습된 모델은 파라미터가 지속적으로 변경되며, 점차 목표하는 음성 품질에 도달하게 된다. 결국 학습과정 없이 빠른 시간 내에 음성 합성 모델을 사용자에게 제공함과 동시에, 사용자로부터 데이터를 제공받고 이를 이용한 학습을 통해 다시 사용자에게 개선된 품질의 음성을 제공하는 상호작용을 통한 인공지능 음성 합성 모델의 진화를 목표로 한다.

본 논문의 구성은 다음과 같다. 2장은 본 논문에서 제안하는 Evolving TTS 모델에 대해 자세하게 기술한다. 3장은 본 모델의 학습에 대한 실험 환경과 합성된 음성의 품질 평가를 포함한 실험 결과를 기술한다. 마지막으로 4장은 결론에 대해 기술하고 본 연구의 한계점 및 향후 연구에 대해 기술한다.

## II. Evolving TTS

Evolving TTS는 초기 단계에서 Zero-shot Learning을 통해 추론과정만으로 빠른 시간 안에 목표 화자의 음성을 합성하고, 이후 지속적인 음성 인식을 통해 데이터를 점차 누적하여 본 모델의 학습을 통해 합성된 음성의 품질을 높이는 제안 기법이다. 기존 TTS의 다량의 데이터셋이 필요함과 동시에 학습 속도의 느림으로 인한 결과 확인이 느리다는 단점과, Zero-shot Learning의 음성 품질의 저하를 동시에 극복하는 새로운 모델을 제안한다. 본 연구에서는 목표하고자 하는 화자의 음성과 최대한 유사하게 음성을 합성하는 것을 목표로 하며, 학습된 음성을 지속적으로 확인할 수 있고, 누적된 음성 데이터를

바탕으로 모델의 파라미터를 점진적으로 수정하여 초기 다소 낮은 품질의 음성을 점진적으로 높은 품질의 음성으로 진화시킬 수 있다.

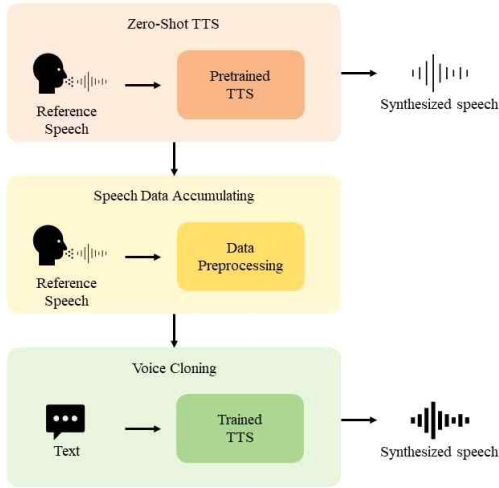


그림 1. Evolving TTS의 시스템 개요  
Fig. 1. System Overview of Evolving TTS

그림 1은 본 논문에서 제안하는 Evolving TTS 시스템의 전체구조를 나타낸다. Evolving TTS의 초기 파이프라인은 Zero-shot TTS 모듈이다. 짧은 시간 내에 목표 화자의 음성을 합성하기 위해 Style Encoder를 사용한다. Style Encoder는 복합 컨볼루션(Convolution) 레이어와 멀티 헤드 어텐션(Multi Head Attention) 레이어를 통해 참조 음성의 특징 벡터를 빠르게 추출한다. 이후 사용자가 입력한 문장과 함께 Stochastic Duration Predictor를 통해 음소 길이를 계산한다. 계산된 음소 길이 벡터와 가중치들은 Residual Coupling Block을 통과한 이후 Wav2vec Decoder를 통해 음성 벡터로 출력되며, 최종적으로 합성기를 통해 사람이 들을 수 있는 파형(음성)을 합성하게 된다.

Evolving TTS의 중간 파이프라인은 음성 데이터 누적 모듈로, 이전 Zero-shot TTS 진행 과정에서 사용된 참조 음성을 전처리 과정을 통해 본 모델의 학습을 준비한다. 음성 인식 모델은 Self-Supervised Learning 기법으로 좋은 성능을 보이는 Facebook의 Wav2vec 2.0[2] 모델을 사용하였으며, 참조 음성을

문자로 전사시켜 참조 음성과 문자 데이터를 이용해 Text-to-Vector 모듈을 학습시킨다.

Evolving TTS의 마지막 파이프라인은 Voice Cloning 모듈이다. 초기 단계의 Zero-shot TTS와 같은 구조를 가지며, 누적된 데이터를 통해 학습된 모델과 합성기를 통해 목표 화자의 음성을 출력한다. 이때, Zero-shot TTS에 사용했던 참조 음성의 특징을 추출하는 모듈인 Style Encoder는 선택에 따라 사용 가능하며, 이후 또 다른 화자의 목소리로 변경할 수 있도록 한다.

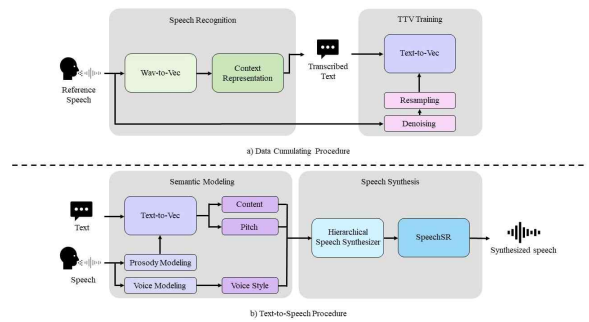


그림 2. Evolving TTS의 데이터 누적 과정과 음성 합성 시스템 구조

Fig. 2. The architecture of data accumulation process and the speech synthesis system in Evolving TTS

그림 2는 그림 1의 시스템 개요를 포함한 데이터 누적 과정과 음성 합성 시스템의 구조이다. TTS의 기본적인 구조는 현재 Zero-shot TTS에서 좋은 성능을 보여주는 HierSpeech++[3] 모델을 수정하여 사용하였다. Zero-shot TTS와 Voice Cloning은 b)를 따르며, 데이터 누적 과정은 a)를 따른다. 각 모듈은 학습과 추론이 동시다발적으로 이루어지며, 사용자와의 상호작용을 통해 최종적으로 음성 합성의 고질적 문제점인 음성 품질과 학습 시간 사이의 타협점을 제시한다. 모델 학습이 진행됨에 따라 사용자는 간헐적으로 음성의 품질을 중간 출력으로 확인할 수 있으며, 원하는 유사도까지 학습 척도를 직접 조절하는 것이 가능하다.

### III. 실험 환경 및 결과

본 모델의 학습은 워크스테이션에서 진행하였으며, 하드웨어 스펙은 다음과 같다. CPU는 Intel

Xeon Silver 4210R, RAM은 128GB, GPU는 RTX 3090 24GB 2way, 운영체제는 Ubuntu 20.04를 사용했다. 학습 시 사용한 하이퍼파라미터는 아래의 표와 같다.

표 1. 학습 하이퍼파라미터

Table 1. Training Hyperparameters

하이퍼파라미터	설정 값
에포크	10000
평가 간격	100
배치 크기	4
학습률	2e-4 (감쇠율 적용)
감가율 ( $\lambda$ )	0.999
최적화 함수	AdamW

학습에 사용한 데이터셋은 총 3가지로, 사전학습을 위한 LibriVox[4] 데이터셋과 이후 학습을 위한 영어 데이터셋 LJSpeech[5], 한국어 데이터셋 KSS Dataset[6](Korean Single Speaker Dataset) 을 이용하였다.

진화형 학습 검증을 위해 사용된 데이터셋의 길이에 따라 결과를 나누어 MOS(Mean Opinion Score) 평가를 진행하였다. 초기 단계의 Zero-shot TTS부터 학습이 충분히 완료되었다고 판단되는 지점까지 여러 구간으로 나누어 동일 인원에게 청취 테스트를 진행했다. MOS는 1~5점의 점수로 구분되며, 1점은 노이즈가 매우 심해 의사소통이 불가능한 품질, 2점은 노이즈가 있어 알아듣기 어려운 품질, 3점은 어느정도 알아들을 수 있는 품질, 4점은 사람의 목소리와 비슷한 품질, 5점은 실제 사람이 말하는 것처럼 들리는 품질로 점수를 부여하여 평가한다.

표 2. LJSpeech 데이터셋의 길이별 학습에 대한 MOS 평가  
Table 2. MOS Evaluation of Length-dependent Training on the LJSpeech Dataset

참조 음성 길이	MOS
5초 (Zero-shot)	4.20
30초 (Zero-shot)	4.39
2분	4.45
5분	4.60
10분	4.62

표 3. KSS 데이터셋의 길이별 학습에 대한 MOS 평가

Table 3. MOS Evaluation of Length-dependent Training on

the KSS Dataset

참조 음성 길이	MOS
5초 (Zero-Shot)	1.14
30초 (Zero-Shot)	1.57
10분	2.86
30분	3.57
60분	4.57

영어 데이터셋의 평가 결과는 사전학습에 사용한 데이터셋과 동일한 언어로 초기 Zero-shot 결과부터 높은 MOS를 보였으며, 이후 데이터가 누적됨에 따라 점진적인 성능 향상을 보였다. 사전학습에 사용되지 않은 언어인 한국어 데이터셋을 적용했을 때, 초기 결과는 비교적 낮은 MOS를 보였으며, 데이터 누적 시간에 비례한 학습이 진행됨에 따라 높은 성능 향상을 확인할 수 있었다.

## V. 결론 및 향후 과제

본 논문에서는 개인화된 인공지능 비서에 대한 접근성과 편의성을 높이기 위한 새로운 학습 기법을 제안하였다. 기존 음성 합성의 문제점인 학습 소요 시간과 음성의 품질을 극복하기 위한 학습 기법인, Zero-shot TTS와 Voice Cloning을 병렬적으로 적용함으로써, 음성의 품질이 점차 높아지는 과정을 사용자가 중간에 확인할 수 있도록 Evolving TTS라는 새로운 시스템을 구현하였다. 제안 시스템의 성능 평가를 위해 참조 음성의 길이별로 합성 결과를 MOS를 통해 측정해 봄으로써 제안 시스템이 점차 음성의 품질을 높이고, 진화해가는 음성을 직접 확인할 수 있음을 보였다.

본 논문에서는 이전에 시도된 바 없던 사용자와의 상호작용을 통한 진화형 인공지능 TTS의 중간 결과 출력과 음질의 단계적 향상을 보인다는 점에서 연구에 의의를 찾을 수 있다. 그러나 제안 시스템은, 상용화를 위해 언어별 사전학습 모델 준비가 필수적이며, 음성 인식의 정확도가 떨어진다는 점 등에서 본 연구의 한계를 찾을 수 있다. 따라서 상기 한계점을 극복하기 위한 연구를 본 논문의 향후 과제로 한다.

## 참 고 문 헌

- [1] A. Oord et al., "WaveNet: A Generative Model for Raw Audio," in arXiv, 2016.
- [2] A. Baevski, Y. Zhou, A. Mohamed, and M. Auli. wav2vec 2.0: A 2022. framework for self-supervised learning of speech representations. Proc. Adv. Neural Inf. Process. Syst., 33:12449-12460, 2020.
- [3] S. Lee, H. Choi, S. Kim, and S. Lee. "HierSpeech ++: Bridging the Gap between Semantic and Acoustic Representation of Speech by Hierarchical Variational Inference for Zero-shot Speech Synthesis," arXiv:2311.12454, 2023.
- [4] "LibriVox - Free public domain audiobooks," <https://librivox.org/>.
- [5] K. Ito and L. Johnson, "The LJ Speech Dataset," <https://keithito.com/LJ-Speech-Dataset/>, 2017.
- [6] K. Park, "KSS dataset: Korean single speaker speech dataset," <https://kaggle.com/bryanpark/korean-single-speaker-speech-dataset>, 2018.

# 게이트 순환 유닛을 활용한 유도가열롤 온도 예측 모델

이강현\*, 박상백\*, 고병진\*\*, 윤종완\*\*, 박태준\*\*

## Temperature Prediction Model for Induction Heating Roll Using Gated Recurrent Unit

Gang Hyeon Lee\*, Sang Baek Park\*, Ko Byungjin\*\*, Jong-wan Yoon\*\*, and Taejoon Park\*\*

### 요 약

본 논문은 유도가열롤의 온도를 정밀한 예측하기 위한 딥 러닝의 종류인 시계열 알고리즘을 사용한 기술 기반의 방법을 제시한다. 게이트 순환 유닛을 기반으로 한 온도 예측 모델을 사용하여, 윈도우 사이즈가 10인 시계열 데이터를 통해 온도 예측의 평균 제곱 오차가  $0.0033^{\circ}\text{C}$ 로 측정되었다. 이는 윈도우 사이즈 7의 예측 결과보다 성능이 개선되었으며, 제안된 접근법으로 최적의 윈도우 사이즈를 통해 정확한 온도 예측을 보장한다. 그 결과 섬유 산업의 품질의 일관성을 기대할 수 있다.

### Abstract

This paper presents a novel, time-series algorithm-based method, a type of deep learning, for accurately predicting the temperature of induction heating rolls. By utilizing a temperature prediction model based on a gated recurrent unit, the mean squared error for temperature predictions using a window size of 10 was measured at  $0.0033^{\circ}\text{C}$ . This represents an improvement in performance compared to the predictions obtained with a window size of 7, ensuring accurate temperature predictions through the proposed approach with an optimal window size. As a result, consistency in quality within the textile industry can be expected.

### Key words

induction heating roll, gated recurrent unit, temperature prediction model

## I. 서 론

최근 섬유 산업은 고부가가치 기능성 섬유의 품질 향상을 위해 카렌딩 공정에 집중하고 있다. 해당

공정은 서로 반대로 회전하는 유도가열롤을 활용하여 섬유를 연속적으로 제조하는 과정이다. 기존의 유도가열롤의 온도를 예측 및 제어 연구는 수식 기반 모델에 의존해왔다[1]. 그러나 이러한 모델들은

\* 한양대학교 인공지능융합학과, rndns147@hanyang.ac.kr, sbp0783@hanyang.ac.kr

\*\* 한양대학교 스마트융합공학부, byungjinko@hanyang.ac.kr(공동 교신저자)

\*\* 한양대학교 지능형로봇학과, jongwanyoon@hanyang.ac.kr(공동 교신저자)

\*\* 한양대학교 로봇공학과, taejoon@hanyang.ac.kr(공동 교신저자)

※ 이 논문은 정부(산업통상지원부)의 재원으로 한국산업기술기술평가원(KEIT)의 지원을 받아 수행된 사업(1415187250, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 융합보안핵심인재양성사업(IITP-2024-RS-2024-00423071, 50%)연구임



유도가열 공정의 복잡한 비선형성을 충분히 반영하지 못한다. 예로 들어, 유도가열롤에서 발생하는 전자기장과 물질의 상호작용, 열전달 과정 및 재료 특성 변화 등은 매우 복잡하고 동적인 요소들로[1], 단순히 현재 상태만을 고려한 제어는 불안전하고 근시안적인 측면이 존재한다. 이러한 문제를 해결하기 위해 최근에는 데이터에서 직접 패턴을 학습하고 비선형성을 예측할 수 있는 딥러닝이 활용되고 있다.[2] 특히, 순환 알고리즘 모델은 시계열 데이터에서의 예측이 효과적이기 때문에 온도 예측에 적합하게 적용되고 있다[2].

그렇지만, 순환 알고리즘 모델의 성능은 입력 데이터의 형태에 크게 의존한다. 입력 데이터의 길이와 패턴은 모델이 학습하는 정보의 양과 품질에 직접적인 영향을 미치기 때문이다[3]. 이에 따라 예측 성능이 달라질 수 있으며, 최적의 입력 데이터 길이를 설정하는 것은 예측 성능을 높이기 위한 필수적인 과정이다. 본 논문에서는 최적의 윈도우 크기(Window Size)를 고려하여 순환 알고리즘의 대표적인 모델 중 하나인 게이트 순환 유닛(Gated Recurrent Unit, GRU)을 활용한 정합도 약 99.73% 성능의 유도가열롤 온도 예측 모델을 제안한다.

## II. 방법론

### 2.1 학습 데이터 구성

윈도우는 시계열 데이터 처리와 자연어 처리에서 활용되는 개념으로, 데이터를 처리할 때 일정한 크기의 데이터 범위를 설정하고 그 범위를 이동시키며 처리하는 방식이다. 이를 통해 전체 시계열 데이터를 여러 윈도우로 나누어 분석함으로써, 연속적인 정보나 패턴을 파악할 수 있다.

그러나 윈도우 크기가 작으면 중요한 패턴을 놓칠 수 있으며, 반대의 경우 노이즈가 증가하거나 과적합이 발생할 위험이 있다[4]. 따라서 문제의 시간적 패턴을 효과적으로 포착할 수 있는 적절한 윈도우 크기를 선택하는 것이 중요하다.

$$X_t^{(i)} = [x_t, x_{t+1}, \dots, x_{t+w-1}] \quad (1)$$

$$X_{t+1}^{(i+1)} = [x_{t+1}, x_{t+2}, \dots, x_{t+w}] \quad (2)$$

본 논문에서는 수식 1과 수식 2같이 슬라이딩 윈도우를 활용하여, 데이터를 균일하게 분할하였다. 데이터 총 크기는 유도가열롤 작동한 시점에서부터 설정 온도에 도달한 뒤 다시 초기 온도에 도달할 때까지의 시간이다. 수식 1과 2는 윈도우 크기가  $w$  이  $w$ 개의 데이터가 있으며  $t$ 는 데이터 순서이다.  $i$ 는 윈도우의 순서를 의미한다.

### 2.2 게이트 순환 유닛을 활용한 온도 예측 모델

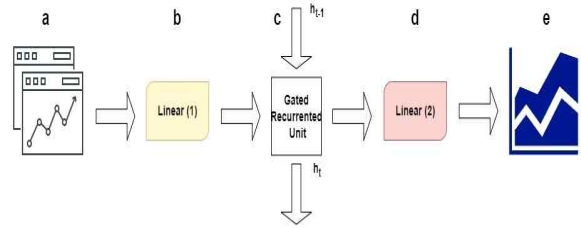


그림 1. 유도가열롤 온도 예측 모델 구조  
Fig. 1. Structure of the Heating Roll Temperature Prediction Model

게이트 순환 유닛은 다른 순환 신경망 모델에 비해 상대적으로 구조가 간단하여 학습 속도와 메모리 효율성이 높다. 그래서 해당 모델을 온도 예측 모델로 설정하였다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Target - Predict)^2 \quad (3)$$

그림 1은 전체적인 유도가열롤 온도 예측 프레임워크를 설명하는 그림이다. a)는 시계열 데이터를 슬라이딩 윈도우를 통해 일정한 크기의 데이터로 나눈다. b) a에서 취득한 데이터를 선형 레이어(1)와 비선형 함수인 ReLU를 활용하여 온도 데이터의 특징을 추출한다. c) 게이트 순환 유닛 레이어에 있는 업데이트 게이트와 리셋 게이트의 조합을 활용하여, 추출된 특징들의 시간적 관계성을 효과적으로 학습한다. d) 출력 레이어(2)에서는 다시 선형 레이어를 통해 최종 온도 예측값을 생성한다. e) 마지막으로

모델의 온도 예측값과 실제 온도 값을 수식 3과 같이 평균제곱오차(MSE) 손실기반으로 역전파하여 모델을 학습시켰다.

### III. 실험

#### 3.1 유도가열롤의 구성

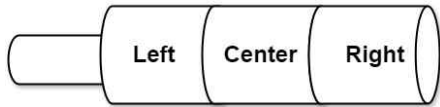


그림 2. 유도가열롤 구성

Fig. 2. Induction heating roll configuration

본 논문에서는 유도가열롤 표면을 그림 2와 같이 왼쪽, 중앙, 오른쪽으로 나누어 분석하였다. 이는 각 부분에 균일하지 않는 열전달을 고려하여 보다 정교한 온도 예측을 가능하게 하기 위함이다.

#### 3.2 실험 설정

수집된 유도가열롤 데이터는 상부 롤 속도(mpm), 중앙온도(°C), 유도롤 전류(A), 유도롤 전압(V), 설정 목표 온도(°C), 롤 왼쪽 압력(kg/cm<sup>2</sup>) 및 온도(°C), 롤 오른쪽 압력(kg/cm<sup>2</sup>) 및 온도(°C), 롤 심부 온도(°C)로 구성되어 있다. 데이터의 진행했으며 지표는 수식(1)에서처럼 평균 시간 크기는 0초부터 17,487초(sec)이다. 데이터는 하루에 1개를 수집했으며 데이터 내용은 초기 온도에서부터 설정 온도까지 도달한 다음, 다시 초기 온도에 도달할 때까지이다. 학습 데이터는 15일, 테스트 세트는 4일로 구성되어 있다. 목표 온도를 각각 120°C, 140°C, 160°C, 180°C로 설정했다.

$$\text{평균 } MSE = \frac{1}{4} \sum_{i=0}^{N=3} (\text{Target} - \text{Predict})^2 \quad (4)$$

이후 성능 지표로 수식 4와 같이 테스트 결과 4개를 평균 제곱 오차(Mean Square Error; MSE)를 사용하였다.

$$\text{정합도} = \frac{1}{4} \sum_{t=0}^3 \sum_{t=0}^T \left[ 1 - \frac{\text{실제값}_t - \text{예측값}_t}{\text{실제값}_t} \times 100 \right] \quad (5)$$

또한, 수식 5에 나온 것처럼 정합도를 사용하여 유도가열롤 온도 예측 모델의 정확도를 나타냈다.

모델의 파라미터를 최적화하기 위해, Adam을 사용했다[ref]. 학습률은 2e-4, 배치 사이즈는 256 그리고 epoch은 100으로 설정했다. 환경은 pytorch를 사용했으며 cpu는 Intel(R) Core(TM) i7-9700f cpu @ 3.00GHz, gpu는 NVIDIA GeForce RTX 2080 Ti이다.

#### 3.3 실험 결과

표 1. 윈도우 크기에 따른 각 유도롤 부위 온도 예측 평균 제곱 오차

Table 1. Mean Squared Error (MSE) of Temperature Predictions for Each Induction Roll Section by Window Size

window size	MSE Left(°C)	MSE Center(°C)	MSE Right(°C)
4	0.0181	0.0179	0.0180
5	0.0176	0.0178	0.0175
6	0.0049	0.0042	0.0041
7	0.0859	0.0874	0.0864
8	0.0168	0.0170	0.0163
9	0.0058	0.0054	0.0043
10	0.0035	0.0033	0.0018

표 1를 바탕으로 윈도우 사이즈가 10일 때 가장 우수한 성능을 보이는 것으로 확인되었다. 이는 성능이 가장 안 좋았던 윈도우 사이즈 7에 비해서 왼쪽은 0.0824, 중앙값은 0.084, 오른쪽은 0.0846 성능이 올라갔다. 이를 통해 윈도우 사이즈 10이 최적의 크기인 것을 확인할 수 있었다.

표 2. 윈도우 크기에 따른 각 유도롤 부위 온도 예측 정합도

Table 2. Goodness of Fit for Temperature Prediction of Each Induction Roll Section by Window Size

Window Size	정합도 Left (%)	정합도 Center (%)	정합도 Right (%)
4	97.825	97.784	97.791
5	97.760	97.722	97.795
6	99.463	99.392	99.426
7	89.818	89.614	89.685
8	97.521	97.629	97.677
9	99.089	99.305	99.394
10	99.728	99.529	99.530

표 2에서도 윈도우 사이즈가 10일 경우 왼쪽, 가운데, 오른쪽 평균 99.595%의 정합도를 보이며 가장 성능이 좋은 것으로 나타났다. 반대로 윈도우 사이즈가 7일 경우 마찬가지로 평균을 났을 때, 89.705%의 정합도를 보였다. 이를 통해 평균제곱오차뿐만 아니라 정합도에서도 가장 좋지 못한 성능을 보였다.

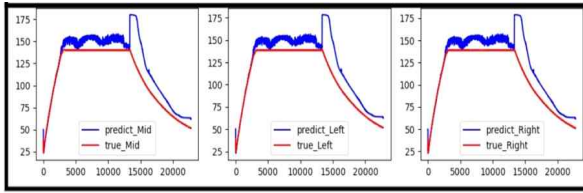


그림 3. 윈도우 사이즈 7일 때  
Fig. 3. Window Size of 7

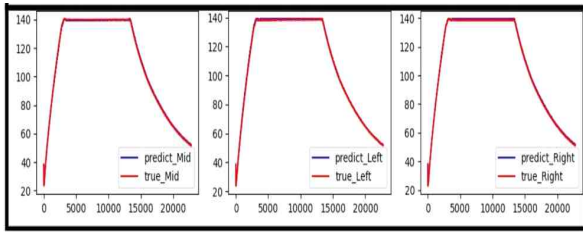


그림 4. 윈도우 사이즈 10일 때  
Fig. 4. Window Size of 10

그림 3은 윈도우 사이즈가 7이고, 설정 온도를 175°C일 때, 첫 번째부터 중앙, 왼쪽, 오른쪽의 실험값이다. x축은 시간(sec), y축은 온도(°C)를 나타낸다. 파란색은 예측 온도 값이며, 빨간색은 실제 온도 값이다. 설정 온도와 상관없이 예측 초반부터 끝날 때까지 정확한 온도를 예측하지 못하는 것을 알 수 있다.

그림 4는 윈도우 사이즈가 10이고, 설정 온도를 140°C일 때 실험값이며, 그림 3과 마찬가지로 x축은 시간(sec), y축은 온도(°C)를 나타낸다. 그림 3의 윈도우 사이즈가 7일 때 보다 예측값이 실제값과 유사한 모습이다.

#### IV. 결 론

본 연구에서는 게이트 순환 유닛을 기반으로 한 시계열 예측 모델을 통해 유도가열롤의 온도 예측

을 실험했다. 특히, 윈도우 사이즈가 10일 때, 예측 정확도가 높았다. 또한, 게이트 순환 유닛은 비선형적 시계열 데이터의 패턴을 효과적으로 학습하여 정확한 온도 예측이 가능함을 확인하였다. 향후 연구에서는 다양한 재질의 섬유에 대해서 유도가열롤을 이용한 작업에서 온도 예측을 정확하게 하는 모델을 개발할 것이다.

#### 참 고 문 헌

- [1] Rötzer, Florian, et al. "Temperature control for induction heating of thin strips." IFAC-PapersOnLine 53.2 (2020): 11968-11973.
- [2] Guo, Hao, et al. "Process modelling of induction welding for thermoplastic composite materials by neural networks." SAMPE 2019-Charlotte, NC, May 2019 (2019).
- [3] Tran, Trang Thi Kieu, et al. "A review of neural networks for air temperature forecasting." Water 13.9 (2021): 1294.
- [4] Dong, Limei, et al. "Prediction of streamflow based on dynamic sliding window LSTM." Water 12.11 (2020): 3032.
- [5] Vaitheeswaran, Sowmya Sevoor, and Varun Raj Ventrpragada. "Wind Power Pattern Prediction in time series measurement data for wind energy prediction modelling using LSTM-GA networks." 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2019.

# 언어 모델을 활용한 질문 난이도 자체 측정 방법론

이주은\*, 김립\*, 남궁혜림\*, 김남규\*\*

## Self-Evaluation of Question Difficulty Using Language Models

Joeun Lee\*, Lib Kim\*, Namgung Haelim\*, Namgyu Kim\*\*

### 요약

최근 프롬프트는 언어 모델 성능 향상의 중요한 요소로 자리 잡고 있으며, 이에 따라 다양한 프롬프트 설계 기법이 활발히 연구되고 있다. 그중에서도 중간 사고 과정을 보여주는 CoT(Chain-of-Thought) 프롬프트는 복잡한 추론 문제에서도 높은 성능을 보여주어 주목받고 있지만, 질문에 대해 인간이 수작업으로 주석을 달아주는 과정이 필요하다는 측면에서 많은 시간과 비용이 요구된다. 이러한 문제를 해결하기 위해 주석을 달 가치가 있는 질문만을 선별하는 연구가 꾸준히 진행되고 있으며, 본 연구는 이러한 배경에서 언어 모델을 활용하여 주석을 달기에 적합한 질문 세트를 선별하고자 한다. 본 연구에서는 라벨이 없는 질문 쌍을 대상으로 토너먼트 방식을 적용해 상대적으로 난도가 더 높은 질문을 선택하는 과정을 반복함으로써, 모델이 자체적으로 질문의 난이도를 평가할 수 있도록 하였다. 제안 방법론은 언어 모델 기반 자체 난이도 측정을 통해 주석을 달 가치가 있는 질문 세트를 효율적으로 제공하며, 궁극적으로 주석이 필요한 질문의 수를 제한함으로써 주석 작업에 소요되는 시간과 비용을 절감할 수 있을 것으로 기대한다.

### Abstract

Recently, prompts have become key to improving language model performance, sparking research into various design techniques. Among these, Chain-of-Thought (CoT) prompts have shown strong results in complex reasoning tasks, but they require costly manual annotation. To address this, research focuses on identifying questions worth annotating. Our study proposes using a language model to select such questions through a tournament-based method, where the model iteratively chooses the more difficult question. This self-evaluation approach helps efficiently determine which questions merit annotation, reducing both time and costs in model inference.

### Key words

prompt engineering, chain-of-thought prompts, question selection, self-evaluation of difficulty, annotation efficiency

---

\* 국민대학교 비즈니스IT전문대학원 석사과정, Email: {wndms2047 fpdkfflq9231, nghaelim98}@kookmin.ac.kr

\*\* 국민대학교 비즈니스IT전문대학원 교수(교신저자), Email: ngkim@kookmin.ac.kr

## I. 서론

최근 딥러닝 기술의 발전으로 비정형 데이터를 효과적으로 분석할 수 있는 기반이 마련되었다. 특히 자연어 처리 분야에서 대규모 언어 모델(Large Language Model, LLM)은 방대한 양의 데이터를 학습하는 과정을 통해 다양한 질문에 대한 답변을 생성하는 능력을 갖추게 되었다. 언어 모델에 특정 태스크를 수행하도록 지시하는 입력 텍스트인 프롬프트는 사용자가 원하는 결과를 효과적으로 도출하며, 모델의 성능을 극대화하기 위해서는 적절한 프롬프트 설계의 중요성이 강조되고 있다. 이에 따라 프롬프트에 관한 연구가 꾸준히 진행되고 있으며, 그중 모델이 문제를 해결하는 과정에서 중간 단계의 추론을 포함하도록 유도하는 CoT 프롬프트는 복잡한 QA(Question & Answering) 태스크에 있어 언어 모델의 성능을 크게 향상시켰다[1].

그러나 현재의 CoT 프롬프트 기법은 사람이 질문에 수작업으로 주석을 달아야 하는 비효율성이 존재한다. 이러한 문제를 해결하기 위해, 합리적인 비용으로 주석을 달 수 있도록 불확실성 측정 지표를 도입하여 질문 풀에서 주석을 달기 가장 유용한 질문을 선택하는 것을 목표로 하는 Active Prompting 기법이 제안되었다[2]. 하지만 해당 방식 역시 각 질문에 대해 여러 번 답변을 요청해야 하기에 비용과 시간 측면에서 여전히 효율적이지 못하다. 본 연구에서는 이러한 한계를 해결하기 위해 불확실성 지표를 언어 모델이 스스로 난이도를 평가하는 방식으로 대체하였다. 난이도는 외부 지표가 아닌 모델 스스로가 가장 잘 알고 있는 정보 중 하나이므로, 모델이 질문의 난이도를 자체 평가할 수 있다면 더 적절한 질문을 선택할 수 있다.

본 연구는 기존 불확실성 측정 지표를 사용하는 대신 언어 모델이 스스로 난이도를 평가하도록 요청하는 새로운 접근법을 제안한다. 모델에 프롬프트를 제시하는 것만으로 질문의 난이도를 판단함으로써 더 효과적인 질문 선택이 가능할 것으로 기대한다. 이를 통해 주석 처리의 효율성을 높이고, 보다 직관적인 난이도 측정을 통해 언어 모델의 성능을 개선하는 방안을 모색한다.

## II. 언어 모델을 활용한 자체 난이도 평가

본 장에서는 언어 모델을 활용한 자체 난이도 측정 방법론을 통해 가장 난도가 있는 질문을 선정하는 단계별 프로세스를 설명한다(그림 1).

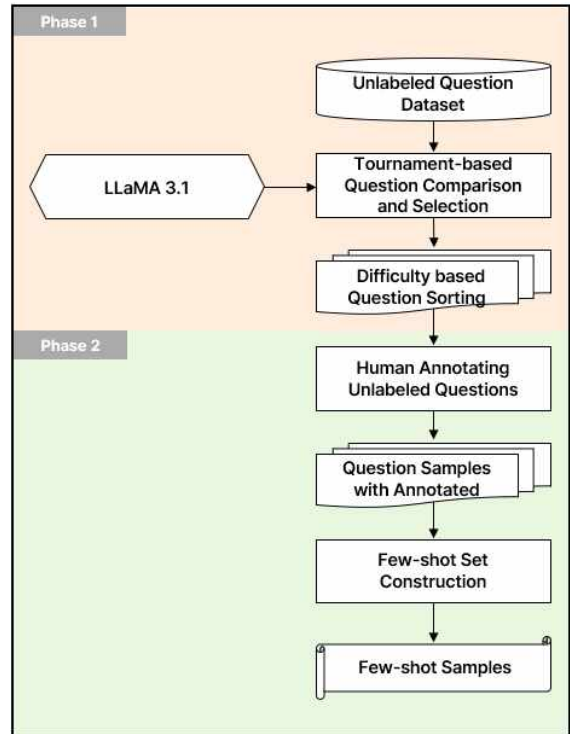


그림 1. 제안 방법론 개요

Fig. 1. Overall Research Process

먼저 Phase 1에서는 다양한 난이도로 구성된 라벨 없는 질문 풀에서 두 개의 질문을 토너먼트 방식으로 비교하여, 더 난도가 높은 질문을 선택할 수 있도록 프롬프트 기반의 자체평가 시스템을 구축하였다. 질문 A와 질문 B는 질문 풀 내에서 무작위로 선택되며, 모델은 두 질문의 난이도를 평가한 후 난도가 더 높다고 판단되는 질문의 인덱스를 출력한다. 본 과정은 총 8개의 질문이 선정될 때까지 반복되며, 선정된 질문은 모델이 자체 평가한 난이도에 따라 내림차순으로 정렬된다. 해당 과정을 통해 주석을 달아야 할 질문을 난이도에 기반하여 효율적으로 선별할 수 있다.

Phase 2에서는 선정된 8개의 질문에 기반해 인간 주석자가 직접 주석을 달고, 각 질문에 대한 답변과 사고 과정에 관련된 정보를 추가한다[3]. 주석이

달린 8개의 질문을 기반으로 Few-shot 세트를 구성하며, 해당 세트는 모델이 특정 태스크를 수행하는데 필요한 예시를 제공하여 추론 성능을 극대화할 수 있도록 설계된다. 구성된 Few-shot 세트는 추론 시 테스트 질문 앞에 추가되는 방식으로 활용되며, 모델이 질문을 처리할 때 다양한 예시를 참조하게 하여 성능을 향상한다. 본 과정을 통해 최종적으로 모델은 주어진 태스크에 대해 보다 정확하고 신뢰할 수 있는 답변을 생성할 수 있게 된다.

### III. 실험

본 연구에서는 제안 방법론의 성능을 검증하기 위해 다음과 같은 실험을 설계하였다(표 1).

표 1. 제안 방법론 실험 설계

Method	Dataset	Model	Evaluation Metric
Active-Prompt(base) Proposed	GSM8K	LLaMA3.1 70B	Accuracy

먼저 난이도가 미리 제시된 대학 수학 능력 평가 시험 문서를 사용하여 파일럿 실험을 진행하고자 한다. 구체적으로 2점부터 4점까지의 난이도가 제시된 문제들로 구성된 문서를 활용하였고, 제안 방법론을 통해 모델이 자체적으로 난이도를 판단할 수 있는지에 대한 측정을 수행한다. 언어 모델로는 LLaMA 3.1 70B 모델을 사용하여 추론을 진행할 예정으로, 모델 스스로가 질문의 난이도를 정확히 평가할 수 있는지에 대한 검증을 목표로 한다. 이후 기준값으로 설정된 Active Prompting 기법과의 성능 비교를 진행하여 불확실성 지표가 언어 모델을 사용한 난이도 측정 방식으로 대체될 수 있음을 확인하고자 한다.

### IV. 결론

본 연구에서는 최근 자연어 처리 분야에서 효과적으로 사용되는 언어 모델을 활용하여 난이도를 측정하는 방법론을 제안하였다. 제안 방법론의 성능을 확인하기 위해 라벨이 달리지 않은 QA 문서를

사용해 실험을 설계하였고, 별도의 지표 설정 없이도 모델이 스스로 난이도를 파악할 수 있음을 확인하고자 한다. 제안 방법론을 통해 주식 작업에서의 비용과 시간을 절감할 수 있으며, 궁극적으로 모델의 추론 성능을 향상할 수 있을 것으로 기대한다.

### 참고 문헌

- [1] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). "Chain-of-thought Prompting Elicits Reasoning in Large Language Models," *Advances in neural information processing systems*, 35, 24824-24837.
- [2] Diao, S., Wang, P., Lin, Y., & Zhang, T. (2023). "Active Prompting with Chain-of-thought for Large Language Models," *arXiv preprint arXiv:2302.12246*.
- [3] Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). "Active Learning with Statistical Models," *Journal of artificial intelligence research*, 4, 129-145.

# CMAPSS 데이터의 잔여 유효 수명 예측을 위한 병렬 연결 구성의 딥러닝 기반 모델

김해영\*, 김도윤\*, 반재필\*\*, 구교권\*\*\*

## CMAPSS Data Remaining Validity Prediction Using Deep Learning-Based Models with Parallel Connectivity Configuration

Haeyoung Kim\*, Doyun Kim\*, Jaepil Ban\*\*, and Gyogwon Koo\*\*\*

### 요 약

터보팬 엔진의 RUL(Remaining useful life) 예측은 엔진이 고장 나기 전 예방 정비를 통해 엔진의 성능을 보장하고, 고장 전 엔진의 상태를 평가할 수 있다. 최근 많은 딥러닝 모델들이 RUL 예측을 시행하고 있으며 RUL 예측에 대한 좋은 성능을 보여주고 있다. 하지만 다소 복잡한 전처리 방법을 사용하여 RUL 예측을 위해 모델을 학습하는 경우가 많다. 이 방법은 비효율적인 노력이 소요될 수 있고, 모델의 복잡함에 의한 과적합을 초래할 수도 있다. 따라서 본 논문에서는 이러한 문제를 해결하기 위해 간단한 전처리 과정과 2차원 합성곱층(2D convolutional layer)을 병렬로 연결한 새로운 구조를 통해 문제를 해결할 수 있는 딥러닝 기반 방법을 제안한다. 먼저, 제안한 알고리즘의 구조에 대해서 설명하고 학습 데이터의 RUL 데이터 구성, 제안한 알고리즘에 대한 다양한 실험들과 RMSE 9.99로 기존 알고리즘들 성능 대비 0.5% 상승했음을 증명했다.

### Abstract

The prediction of the Remaining Useful Life (RUL) of a turbofan engine allows for evaluating the engine's condition before failure and ensuring performance through preventive maintenance. Recently, many deep learning models have been applied to RUL prediction, demonstrating good performance. However, many of these models rely on complex preprocessing methods for training, which can require inefficient effort and lead to overfitting due to model complexity. Therefore, this paper proposes a deep learning-based method to address these issues by using a simple preprocessing process and a novel structure that connects two-dimensional convolutional layers in parallel. Firstly, the structure of the proposed algorithm is explained, along with the configuration of the RUL data in the training dataset. Through various experiments conducted on the proposed algorithm, it demonstrated a 0.5% improvement over existing algorithms with an RMSE of 9.99.

### Key words

deep learning, time series, RUL predict, turbofan dataset, Neural network

---

\* 국립금공과대학교 전자공학과 석사과정, sunh@kumoh.ac.kr ,do0207@kumoh.ac.kr

\*\* 국립금공과대학교 전자공학부 교수, banjp@kumoh.ac.kr

\*\*\* 대구경북과학기술원 지능형로봇연구부 선임연구원, gkoo@dgist.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부에서 지원하는 DGIST 기관고유사업에 의해 수행되었습니다 (24-IT-02)

## I. 서론

터보팬 엔진은 가스터빈 엔진의 일종으로 항공기에 추진력을 공급하는 동력원으로 사용된다. 터보팬 엔진의 고장은 비행 중 손쓸 수 없는 사고를 유발하기 때문에 심각한 인명 피해와 큰 재산 손실이 발생할 수 있다[1]. 그러므로 터보팬 엔진의 정비는 필수적으로 시행해야 한다.

엔진의 정비로 상태 기반 정비를 많이 시행하고 있다[2,3]. 상태 기반 정비이란 실시간 모니터링에서 얻은 센서 출력값을 이용해 엔진의 결함을 정비하는 전략을 의미한다. 이전에 엔진의 상태를 확인하기 위해 엔진의 작동을 정지하고 분해, 수리하던 예방 정비와는 다르게 실시간으로 엔진의 상태를 파악할 수 있고, 그에 따라 높은 노동력과 비용을 절감할 수 있었다. 그렇기에 요즘은 상태 기반 정비 전략을 이용해 엔진이 정상적으로 동작할 수 있는 시간을 의미하는 잔여 유효 수명(Remaining Useful Life, RUL)을 예측하는 알고리즘들이 많이 개발되는 추세이다[3,4].

잔여 유효 수명을 예측하는 알고리즘으로 다양한 딥러닝 모델들이 개발되고 있다. 이전 연구들 중 RNN(Recurrent Neural Network)을 이용해 엔진의 데이터를 이용해 RUL로 고장을 진단하는 알고리즘이 많이 연구되었다[5,6]. [5]는 1차원 합성곱 층(1D convolution layer)만을, [6]은 LSTM(Long Short-Term Memory)만을 사용한 알고리즘이다. 많은 RNN 계열 알고리즘은 시계열 데이터를 시간의 측면에서 데이터의 변화량을 학습하기 때문에 시계열 학습에 용이하지만, 다변량 시계열 데이터의 변량 간의 관계를 고려할 수 없다는 문제점이 있다. 이 점을 보완하고자 최근 CNN(Convolution Neural Network)과 RNN을 결합한 알고리즘이 많이 제안되었다[7-9]. [7]은 2차원 합성곱 층(2D convolution layer)과 LSTM, [8]은 2차원 합성곱 층과 Bi-LSTM(Bidirectional Long Short-Term Memory), [9]는 2차원 합성곱 층, Bi-LSTM, LSTM을 이용해 모델을 구성했다. CNN 계열을 이용해 변량 간의 관계를 고려할 수 있지만 RNN 계열은 이전 시간 단계의 계산들이 모두 수행되어야 현재 시간 단계의 계산이 수행되기 때문에

학습이 매우 느려진다는 문제점[10]을 가지고 있고 앞서 설명한 모델 모두 해결하지 못한 문제점이다.

본 논문에서는 CMAPSS 데이터의 잔여 유효 수명 예측을 위한 병렬연결 구성의 딥러닝 기반 모델을 제안한다. 논문에서 제안하는 알고리즘은 복잡한 전처리를 사용하지 않고 학습을 진행하기 때문에 다른 데이터 적용에 용이하고, 다양한 커널 크기를 사용해 센서 간의 다양한 특징을 추출하여 학습할 수 있다. 이 점은 한 번의 여러 특징을 고려할 수 있기 때문에 RUL 예측 성능 향상에 기여할 수 있다.

## II. RUL 추정 알고리즘

### 2.1 제안하는 모델

본 논문에서 제안하는 딥러닝 모델의 구조는 Fig 1과 같다. 모델에서 사용한 층은 2차원 합성곱 층, 완전 연결 층(Fully connected layer), 배치 정규화(Batch Normalization), ReLU, 드롭아웃(Drop out)을 사용하였다. 모든 2차원 합성곱 층 이후에는 배치 정규화, ReLU, 드롭아웃을 연결해 사용하여 학습의 성능을 높였다. 2차원 합성곱 층은 Fig 2에서 Conv로 2차원 합성곱 층의 커널 정보를 ( $H \times W, C$ )로 나타냈다.  $H$ 는 커널의 높이,  $W$ 는 커널의 넓이,  $C$ 는 커널의 개수,  $\oplus$ 은 채널 방향으로 concatenation 합을 의미한다. 본 논문에서 제안하는 모델은 시간에 흐름에 따라 측정된 센서의 특징으로 RUL을 추정하는 이전 방식과 다르게 30 cycle에 걸쳐 측정된 데이터의 특징으로 RUL을 추정한다. 모델의 구조는 2차원 합성곱 층을 이용해 센서들의 특징을 추출하고 조합하여 완전 연결 층을 이용해 회귀를 진행한다. 2차원 합성곱 층의 커널 정보를 모두 다르게 설정하여 입력되는 다변량 시계열 데이터에서 이웃되는 센서들의 특징을 다양하게 추출할 수 있도록 설정했다. 이 기법으로 다양한 센서의 특징 표현을 추출하고 취합할 수 있게 되어 성능을 높일 수 있었다.



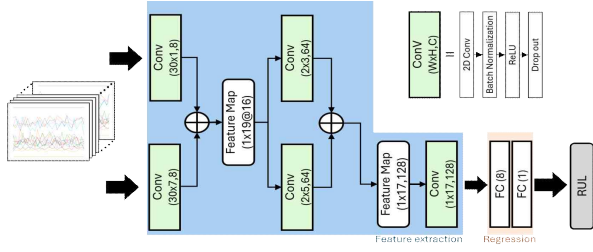


그림 1. 모델의 구조  
Fig. 1. model architecture

### III. 실험

#### 3.1 실험 설정

본 논문에서 사용한 데이터는 C-MAPSS 데이터셋 [11]을 사용했다. C-MAPSS 데이터셋은 21개의 센서 데이터를 각 cycle에 맞게 측정해 엔진의 상태를 센서값으로 출력한 센서 데이터이다. C-MAPSS 데이터셋은 학습 데이터의 RUL이 제공되지 않기 때문에 [12,13]에서 제안한 RUL 추정 모델로 선형 열화 모델을 사용하였다. [12]에서 제안한대로 본 논문에서도 최대 RUL을 125로 지정하였다. 학습 데이터 셋과 검증 데이터 셋은 8:2 비율로 나누었고 윈도우 길이는 30으로 설정해 윈도우 슬라이딩 기법을 사용했다. 학습 데이터 셋은 (16504, 30, 21), 검증 데이터 셋은 (4127, 30, 21) 로 정의했다.

실험을 실행한 환경은 Anaconda의 jupyter notebook에서 진행되었고, 실험에 사용된 하이퍼 파라미터로 옵티마이저는 Adam, 배치 크기는 500, 손실함수로 MSE를 사용했다. MSE의 수식은 (1)과 같다.

$$MSE = \frac{1}{K} \sum_{k=1}^K (y_k - y_{k,true})^2 \quad (1)$$

$K$ 는 한 데이터셋에서 가지는 엔진의 개수를 의미한다. 실험에서 사용한 데이터셋에서는  $K = 100$ 이다.  $y_k, y_{k,true}$ 는 각각  $k$ 번째 엔진의 데이터를 이용해 모델이 예측한 RUL 값,  $k$ 번째 엔진의 실제 RUL값을 의미한다. 논문의 평가지표로 RMSE를 사용했다. 수식은 (2)와 같다.

$$RMSE = \sqrt{\frac{1}{K} \sum_{k=1}^K (y_k - y_{k,true})^2} \quad (2)$$

표 1. 알고리즘의 구조 실험

Table 1. Structural experiments of algorithms

Algorithms Structural	RMSE	
	Avg	Std
related work[9]	11.02	0.43
only conv layer[9]	11.07	0.37
<b>Parallel Conv</b>	<b>10.34</b>	<b>0.50</b>

수식에서 사용된 값들은 모두 MSE와 같다. RMSE를 사용해 평가한 최근 모델들과 결과를 비교할 것이다. 또 regularization으로 L2 regularization을 사용하였다. 모든 실험은 실험의 신뢰도를 높이기 위해서 같은 실험을 10번 실행하여 평균한 결과를 나열했다.

#### 3.2 제안한 알고리즘의 최적화 실험

이전 논문[9]의 알고리즘을 좀 더 개발하여 논문에서 제안하는 알고리즘을 만들게 되었다. 먼저 [9]의 알고리즘에서 RNN 계열의 층들을 모두 제외하고 학습한 모델, 병렬 2차원 합성곱 층만을 병렬로 쌓은 모델을 학습해 실험을 진행했다. 결과는 Table 1과 같다. RNN 계열을 모델에서 제거했을 때의 결과와 이전 논문의 결과가 큰 차이가 없음을 확인했고, RNN 계열이 큰 학습효과를 주지 않는 것을 증명했다. 제안하는 알고리즘의 논문 구조로 변경해 학습을 진행했을 때 RMSE가 10점대로 가장 좋은 RMSE 값을 가지는 것을 확인할 수 있었다.

Table 2는 Table 1에서 가장 결과가 좋았던 구조로 학습을 최적화 실험을 진행한 결과이다. 기존의 학습률은 0.01 이었고, 0.001, 0.0001을 학습률로 두어 추가 실험을 진행했다. 결과, 0.001이 가장 좋은 결과를 얻었고 다음 실험은 Table 2의 가장 좋은 결과인 0.001을 학습률로 두어 추가 실험을 진행했다.

Table 3는 드롭아웃 비율에 대해 최적화 실험을 진행한 결과를 나타냈다. 기존의 드롭아웃 비율은 0.2였고, 0.5, 0.4, 0.3, 0.1, 드롭아웃 층 제거를 추가로 실험을 진행했다. 결과, 0.1이 가장 좋은 결과를 얻어 모든 실험 중 가장 좋은 성능을 보여주는 것을 확인했다. 따라서 학습률 0.001과 드롭아웃 비율 0.1로 최종 모델을 선정했다.

표 2. 알고리즘의 학습률 실험

Table 2. Algorithm learning rate experiment

Hyper Parameter	RMSE	
	Avg	Std
learning rate	0.01	10.34
	<b>0.001</b>	<b>10.05</b>
	0.0001	10.48

표 3. 알고리즘의 드롭아웃 비율 실험

Table 3. Algorithm Drop out rate experiment

Hyper Parameter	RMSE	
	Avg	Std
Dropout rate	0.5	11.31
	0.4	11.01
	0.3	10.43
	0.2	10.05
	<b>0.1</b>	<b>9.99</b>
	0.0	10.26

### 3.3 최신 알고리즘들과 결과 비교

논문에서 제안한 알고리즘과 최근 C-MAPSS 데이터셋을 사용한 RUL 추정 알고리즘들을 비교한 결과를 Table 4에 나타냈다. 여러 계열의 층을 쌓아 구성한 알고리즘들과 비교했을 때 합성곱 층만을 사용한 논문에서 제안한 알고리즘이 훨씬 더 좋은 성능을 보여주는 것을 입증했다. 이를 통해 논문에서 제안한 알고리즘은 병렬로 쌓은 2차원 합성곱 층 구조를 이용해 다양한 센서의 특징 관계를 학습하는 것이 성능 증진에 기여했음을 증명할 수 있었다.

표 4. 최근 알고리즘들과 예측 결과 비교

Table 4. Comparison of prediction results with recent algorithms

Method	RMSE
2D-CNN + LSTM [7]	16.13
LSTM + FC [6]	14.08
1D CNN [5]	13.32
2D-CNN + Bi-LSTM [8]	13.22
2D-CNN + Bi-LSTM + LSTM [9]	10.04
<b>Proposed Method</b>	<b>9.99</b>

## IV. 결 론

본 논문에서는 터보팬 엔진 데이터를 이용해 엔진의 잔여 유효 수명(Remaining Useful Life, RUL)을 예측하는 딥러닝 모델을 제안했다. 제안한 모델은 2차원 합성곱 층을 병렬로 연결해 다양한 센서 관계를 특징으로 추출해 학습의 성능을 높였다. 또, 간단한 전처리를 사용해도 학습에 좋은 성능을 보이기 때문에 같은 모델에 다른 데이터를 적용할 때의 큰 용이성을 보인다. 마지막으로 제안한 모델은 여러 최적화 실험을 통해 정해진 파라미터로 학습 결과를 내놓았을 때, 최신 알고리즘들과 비교하여 가장 좋은 결과를 보여주었기 때문에 C-MPASS 데이터셋에 매우 알맞게 최적화된 모델임을 입증할 수 있었다.

## 참 고 문 헌

- [1] <https://www.bbc.com/korean/articles/cy4ln3ve4m9o>
- [2] H. N. Nair and C. V. Stewart, "Robust focus ranging", Proc. CVPR, pp. 309-314, 1992.
- [3] 최석주. "신뢰도 기반 정비체계 적용을 위한 연구." 국내석사학위논문 공주대학교 일반대학원, 2021. 충청남도
- [4] T. S. Choi and J. Yun, "Three-dimensional shape recovery from focused image surface". Opt. Eng., vol. 39, May 2000.
- [5] Li, X., Ding, Q., Sun, J.-Q., "Remaining useful life estimation in prognostics using deep convolution neural networks". Reliability Engineering and System Safety, vol. 172, pp. 1-11, Dec. 2018.
- [6] Lin, R., Yu, Y., Wang, H., Che, C., Ni, X., "Remaining useful life prediction in prognostics using multi-scale sequence and Long Short-Term Memory network". Journal of Computational Science, Vol. 57, Nov. 2021.
- [7] Kong, Z., Cui, Y., Xia, Z., Lv, H., "Convolution and Long Short-Term Memory Hybrid Deep Neural Networks for Remaining Useful Life Prognostics". Applied Sciences, Oct. 2019.
- [8] Wang, L., Zhu, Z., Zhao, X., "Dynamic predictive

- maintenance strategy for system remaining useful life prediction via deep learning ensemble method”.Reliability Engineering and System Safety, vol. 245, Feb 2024.
- [9] Back, M., Kim, H., Ban, J., Koo, g., “터보팬 엔진의 잔여 유효 수명 예측을 위한 2D CNN 기반 모델”. 대한 전자공학회 추계학술대회 논문집, pp. 261-263, 2023.
- [10] Shiri, F, M., Perumal, T., Mustapha, N., Mohamed, R,“A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU”arXiv preprint arXiv:2305.17473, Jun. 2023.
- [11] Frederick, D, K., DeCastro, J., Litt, J.S,“User’s guide for the commercial modular aero-propulsion system simulation (C-MAPSS), 2007.
- [11] Heimes, F,“Recurrent Neural Networks for Remaining Useful Life Estimation”. International Conference on Prognostics and Health Management, 2008.
- [12] Babu, G, S., Zhao, P., Li, X, L,“Deep convolutional neural network based regression approach for estimation of remaining useful life”. Database Systems for Advanced Applications, pp. 214-228, Mar 2016.

# ISA 기반 March 패턴 생성기의 FPGA 구현 및 검증

정인수\*, 강주완\*, 김태준\*, 조승준\*, 문병인\*\*

## FPGA Implementation and Verification of an ISA-Based March Pattern Generator

Insu Jeong\*, Joowan Kang\*, Taejun Kim\*, Seungjun Jo\*, and Byungin Moon\*\*

최근 메모리 수요의 증가와 관련 기술의 급속한 발전으로 메모리 성능이 빠르게 발전하고 있다. 이에 고속으로 고용량 메모리의 테스트를 수행하는 방법에 대한 연구가 활발히 진행되고 있다. March 알고리즘은 메모리 테스트 알고리즘 중 테스트 방식이 단순하여 빠르게 테스트 패턴을 생성할 수 있다. 하지만 단순한 March 알고리즘을 사용하더라도 고속으로 동작하는 메모리의 경우 테스트 장비의 패턴 생성 속도와 메모리의 동작 속도 차이로 인해 병목 현상이 발생한다. 이에 본 논문에서는 테스트 패턴 생성 시간을 줄인 패턴 생성기의 하드웨어 구조[1]를 활용하여, 명령어 집합 구조(Instruction Set Architecture, ISA) 기반의 1-way 패턴 생성기를 FPGA로 설계하고, 이를 검증한 결과를 제시한다. 본 논문은 Xilinx사의 xc7z020clg484-1 FPGA에 제안한 ISA 기반 1-way 패턴 생성기를 구현하였으며, Instruction Memory에 저장된 March C-, March B 알고리즘의 instruction에 맞게 테스트 패턴을 생성하는 것을 확인했다. 본 논문에서 사용한 패턴 생성기의 하드웨어 구조와 시뮬레이션 결과 파형은 그림 1, 2와 같다.

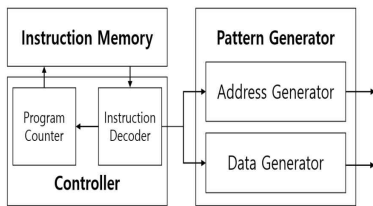


그림 1. 패턴 생성기 하드웨어 구조

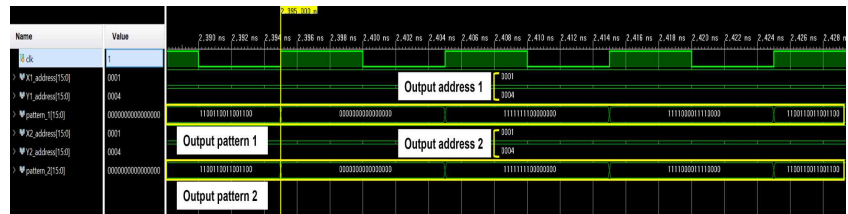


그림 115. 패턴 생성기 시뮬레이션 결과 파형

### Key words

march test, high speed memory, pattern generator, hardware structure, FPGA verification

### 참고문헌

- [1] 김정환, 송하윤, 문병인, “초고속 메모리 테스트를 위한 패턴 생성기”, 한국반도체테스트학회 학술발표대회 논문집, 제 24권, pp. 154-156, 2023.

\* 경북대학교 대학원 전자전기공학부, {insu8944, jwkang1231, ktj016669, jo73012}@knu.ac.kr

\*\* 경북대학교 대학원 전자전기공학부, 경북대학교 IT대학 전자공학부, bihmoon@knu.ac.kr, 교신저자

※ 본 연구는 산업통상자원부(20019363)와 KSRC 지원 사업인 미래반도체소자 원천기술개발사업의 연구결과로 수행되었습니다.

※ 본 연구는 IDEC에서 EDA Tool을 지원받아 수행하였습니다.

# 미래 인지전 수행을 위한 정보 획득 방안 연구

이종화\*, 전재현\*\*

## A Study on Information Acquisition Method for Cognitive Warfare

Jonghwa Lee\*, and Jaehyeon Jun\*\*

본 논문은 인지전 개념의 등장 배경과 특징을 살펴보고, 미래 우리 군의 인지전 수행을 위한 개념 도출과 인지전의 가장 큰 특징인 실시간 전쟁 상황에 대한 신뢰성 있는 정보획득 방안을 제시하는데 목적을 두고 있다. 이를 위해 본 연구는 인지라는 개념이 사람들의 관심을 끌기 위한 배경으로 소셜 네트워크 서비스(SNS)의 등장과 발전 그리고 이에 대한 영향을 연구하는 과정에서 뇌인지와 뇌신경과학의 발달과 함께 인지전이라는 용어가 군사적으로 등장하였다고 판단하였다. 소셜 네트워크 서비스가 군사적으로 활용된 사례를 비교 분석하기 위해 2022년 러시아-우크라이나 전쟁을 분석의 대상으로 삼고, 러시아와 우크라이나에서 각각 진행된 인지전의 수행 양상을 7개 전쟁 단계를 분류하고 일자별로 분석하였다. 본 연구에서는 분석한 인지전 수행 양상에 따라 인지전을 수행하기 위한 특징을 디지털 플랫폼 운용, 신뢰성 있는 정보의 획득과 유통, 소셜 네트워크 서비스(SNS)를 통한 뇌인지의 즉각적인 반응으로 도출하고, 이에 맞는 인지전의 개념을 정립하였다. 본 연구에서 제시한 인지전의 개념을 기반으로 우리 군의 인지전 수행을 위한 군사적 측면에서의 전장 정보 획득 방안은 첫 번째, 우리 군·사단에서 창설되는 공보작전의 기본 대대인 정훈대에서 SNS를 활용할 수 있는 기획 및 계획의 발전과 함께 두 번째로 미군의 실시간 디지털 유통 정보 시스템인 DVIDS 운영 플랫폼을 살펴보고 이를 통한 우리 군의 정보 획득 방안을 제시하였다.

### Key words

Cognitive Warfare, Russia-Ukrainian War, Information Warfare, Social Network Service, DVIDS

### 참고문헌

- [1] 강신욱, “인지전 개념과 한국국방에대한 함의: 러시아-우크라이나 전쟁을 중심으로,” 『국방정책연구』, 통권제139호, pp. 183-186
- [2] 김상현, “인지전의 공격 양상과 대응에 관한 연구: 2014년 크림반도 합병과 2022년 우크라이나 전쟁을 중심으로,” 『21세기정치학회보』 제32집 4호, pp. 21-49.
- [3] 김유향, “소셜미디어와 인터넷공간에서의 정치적 소통,” 『평화연구』 2011년 가을호, pp. 199-218
- [4] 김진호, 최영찬. “미래전에 대비한 한국군 인지전 발전 방향,” 『국방정책연구』, 통권제141호, pp. 93-132.
- [5] 송태은, 북한의 사이버 위협 실태와 우리의 대응, IFANS 주요국제문제 분석 2023-09
- [6] 송태은, 이스라엘-하마스 전쟁의 사이버 인지전: 전개양상과 함의, IFANS 주요국제문제 분석 2024-11.
- [7] 조상근 외, “2021년 이스라엘-팔레스타인 분쟁에서의 인지전 사례 연구,” The Journal of the Convergence on Culture Technology, Volume 8 Issue 6(2022), pp. 537-542.

\*한국국방연구원, jonghwa31@kida.re.kr, \*\*한국국방연구원, theejay@kida.re.kr.

# 인공지능기반 배병해충 분류 웨어러블 장치 연구

당탄부\*, 김재석\*, 김순천\*, 장진호\*, 유광현\*\*

## A Study on wearable device for classification of pear pests based on artificial intelligence

Dang Thanh Vu\*, Kim Jae-seog\*, Kim Sun-cheon\*, Jang Jin-ho\*, and Yu Gwang-hyun\*\*

최근 몇 년간 컴퓨터 비전 분야에서 인공지능 모델은 뛰어난 성과를 보여왔다. 농업 분야에서 인공지능은 농장의 해충, 식물 질병 및 영양 부족을 감지하는데 도움이 되고, 이를 기반으로 하는 정밀 농업은 수확의 전반적인 품질과 정확성을 향상시키는데 도움이 된다. 본 연구에서는 정밀 농업 이후 수확한 배를 선과장의 선별 과정에서 정상 배와 질병 및 해충이 있는 배로 구분하는 인공지능기반 배병해충 분류 웨어러블 장치를 연구한다. 인공지능 모델은 YOLO(You Only Look Once) 모델의 최신 버전인 YOLOv5와 YOLOv8로 배 병해충 분류 모델을 구성하여 성능을 비교한다. 데이터셋은 공개데이터셋인 AI-HUB 과수화상병 촬영이미지를 사용하여 학습하고, 자체적으로 선과장에서 수집한 배 병해충 데이터를 통해 테스트하였다. YOLOv5와 YOLOv8에서 얻어진 객체 탐지 결과 중 분류 결과만을 사용해서 분류 정확도를 계산하였다. 실험 결과 자체 수집한 배병해충 분류 결과는 YOLOv5기반 모델은 81%, YOLOv8기반 모델은 77%를 달성하였다. 전체 배병해충 분류 웨어러블 장치는 사용자가 카메라 및 통신 장치가 달린 밴더와 인공지능 서버로 구성된다. 사용자가 웨어러블 장치를 구동하면 촬영된 이미지가 무선으로 서버로 전송되고 서버에서 배병해충을 분류하고 결과를 LED를 통해 알려준다. 인공지능 서버는 실험결과 우수한 성능을 보인 YOLOv8기반 배병해충 분류 모델이 탑재되고, 사용된 이미지는 저장된다. 이를 통해 인공지능기반 배병해충 분류 웨어러블 장치가 선과장에 도착한 배를 선별하는 선별자들을 도와 더 효과적으로 배병해충 선별작업을 도울 수 있을 것으로 기대된다.

### Key words

artificial intelligence, pear pest, classification, YOLOv5, YOLOv8

### 참고문헌

- [1] Jeong-Tae, K., Hui-Yeon, J., and Eun-Jeong C., "딥러닝 기법을 이용한 머신 비전 기술 최근 응용 동향", The Magazine of the IEIE, Vol. 43-11, pp.18-26, Nov, 2016.
- [2] Redmon, J., "You only look once: Unified, real-time object detection", In Proceedings of the IEEE conference on computer vision and pattern recognition, pp.779-788, July, 2016.
- [3] Terven, J., Córdova-Esparza, D. M., and Romero-González, J. A., "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas", Machine Learning and Knowledge Extraction, Vol.5-4, pp.1680-1716, Nov, 2023.

\* (주)에이아이시드, dtvu@aiseed.kr, bossrobot@aiseed.kr, ai\_seed@aiseed.kr, verycosy@kakao.com

\*\* 전남대학교산학협력단, sayney1004@gmail.com(교신저자)

※ 이 논문은 과학기술정보통신부와 정보통신산업진흥원으로부터 '2024년 지역디지털기초체력지원(광주\_SW서비스)(S-2001-24-1005)'과 과학기술정보통신부·광주광역시 공동 지원한 '인공지능 중심 산업융합 집적단지 조성사업'으로 지원을 받아 수행된 연구 결과입니다.

# 인공지능을 위한 배병해충 이미지 데이터셋 연구

르황안\*, 이진\*, 이주환\*, 박수혁\*\*, 김민경\*\*\*, 김진영\*

## A Study on image dataset of pear pests for artificial intelligence

Le Hoang Anh\*, Lee Jin\*, Lee Ju-hwan\*, Park Su-hyeok\*\*, Kim Min-kyung\*\*\*, and Kim Jin-young\*

최근 몇 년간 컴퓨터 비전 분야에서 인공지능 모델은 뛰어난 성과를 보여왔다. 그러한 배경에는 CIFAR10, ImageNet, Pascal VOC 등과 같은 대규모 데이터셋이 있었으며, 이는 이미지 분류(Classification), 객체 탐지(Object Detection), 분할(Segmentation) 분야에서 인공지능 모델이 뛰어난 성능을 보여주는 기반이 되었다. 인공지능이 일반적인 컴퓨터 비전 분야를 넘어 다양한 분야로 확장되면서 도메인에 맞는 데이터셋이 중요하게 되었다. 본 연구에서는 농장에서 수확한 배를 선과장 및 선별장으로 보내 배를 선별하는 과정에서 정상 배와 병해충이 있는 배로 구분하는 도메인에 필요한 배병해충 이미지 데이터셋 및 인공지능 모델을 통한 데이터셋 검증을 수행하고자 한다. 배병해충 이미지 데이터셋을 구축하기 위하여 전남지역 선과장으로 부터 비상품성 배를 구매하였다. 구매한 비상품성 배는 배병해충 전문가를 통해 일일이 검수하여 7개의 클래스별로 분류하였다. 분류된 배병해충은 일반적으로 상품 사진을 촬영하는 포토박스를 만들어 정면, 측면, 상부 3면에서 스마트폰 고해상도로 촬영하였다. 촬영된 고해상도 이미지를 클래스 폴더별로 정리하고, 이를 다시 한 번 배병해충 전문가를 통해 검수하였다. 최종적으로 촬영된 전체 배병해충 이미지는 10,390장이며 클래스별로는 흑성병 1,733장, 과피열룩 36장, 복숭아순나방 131장, 복숭아식식나방 72장, 배각지벌레 579장, 잎말이나방류 100장, 기타 7,739장이다. 인공지능용 배병해충 이미지 데이터셋 검증을 위해 YOLO(You Only Look Once) 모델의 최신 버전인 YOLOv5와 YOLOv8로 배병해충 탐지 및 분류 모델을 구성하였다. YOLOv5와 YOLOv8에서 얻어진 객체 탐지 결과와 바운딩박스의 클래스 결과만을 사용한 분류 정확도를 평가지표로 삼았다. 실험 결과 YOLOv5 모델의 mAP50은 93.57%, F1-Confidence는 88%, 분류 정확도는 96.88%이고, YOLOv8 모델의 mAP50은 90.83%, F1-Confidence는 86%, 분류 정확도는 93.75%를 달성하여 배병해충 이미지 데이터셋이 인공지능 모델에 효과적이라는 것을 확인하였다. 이를 통해 인공지능을 위한 배병해충 이미지 데이터셋이 배병해충과 관련된 인공지능 모델 및 서비스를 연구 및 개발할 때 도움을 줄 수 있을 것으로 기대된다.

### Key words

artificial intelligence, pear pest, dataset, YOLOv8

### 참고문헌

- [1] Redmon, J., "You only look once: Unified, real-time object detection", In Proceedings of the IEEE conference on computer vision and pattern recognition, pp.779-788, July, 2016.

\* 전남대학교 지능전자컴퓨터공학과, andrewlee1807@gmail.com, lijn96@naver.com, qualify12@gmail.com, beyondi@jnu.ac.kr(교신저자)

\*\* 전남대학교 전자공학과, psh011215@naver.com

\*\*\* 전남대학교산학협력단 지능전자연구실, getlucky@chonnam.ac.kr

※ 이 논문은 과학기술정보통신부와 정보통신산업진흥원으로부터 '2024년 지역디지털기초체력지원(전남\_SW서비스) (S-2001-24-1008)'과 과학기술정보통신부·광주광역시가 공동 지원한 '인공지능 중심 산업융합 집적단지 조성사업'으로 지원을 받아 수행된 연구 결과입니다.





# 협동지능 기반 자가-적응형 재난 대응 프레임워크

박지훈\*, 김영갑\*\*

## A Collaborative Intelligence-Based Self-Adaptive Disaster Response Framework

Jihoon Park\*, and Young-Gab Kim\*\*

### 요약

현재 각종 재난 및 범죄 예방 등을 목적으로 다양한 지자체에 CCTV 통합 관제센터가 설치되고 있지만, 사전 대응 불가, 육안 관제 의존, 규칙 기반 이상 탐지의 높은 오탐율 등의 한계로 효과적인 재난 예측 및 대응이 어렵다. 또한, 기존에 발생하지 않던 신종 재난의 발생으로 인한 인적·물적 피해가 발생함에 따라 다양한 재난을 신속하게 예측하고 판단하여 대응할 수 있는 시스템 개발의 필요성이 대두되었다. 이 논문에서는 다양한 재난을 예측하고 사전에 대응하며, 만약 재난이 발생하였더라도 자율적으로 재난에 대응할 수 있는 협동지능 기반의 자가-적응형 재난 대응 프레임워크를 제안한다.

### Abstract

Various municipalities have established integrated CCTV monitoring centers to prevent disasters and crimes. However, these centers face limitations, such as the inability to implement proactive measures, reliance on manual monitoring, and a high rate of false positives in rule-based anomaly detection systems, making it difficult to predict and respond to disasters effectively. Furthermore, new types of disasters, which were not previously encountered, have resulted in significant property and human losses. As a result, there is an increasing need for a system that can swiftly predict, assess, and respond to various disasters. We propose a collaborative intelligence-based self-adaptive disaster response framework that predicts and preempts disasters and autonomously responds when a disaster occurs.

### Key words

self-adaptive system, collaborative intelligence, disaster response automation, anomaly detection

### 1. 서론

최근 각종 재난 및 안전 사고의 발생 빈도가 증가함에 따라, 각종 재난·사고 및 범죄 예방 등의 목

적으로 다수의 지자체와 공공기관에 CCTV의 설치가 꾸준히 증가하고 있다. e나라지표에 따르면, 공공기관에 설치된 CCTV의 수는 지난 2020년 약 133만대에서 2024년 약 176만대로 약 32% 가량 증가

\* 세종대학교 정보보호학과, syr03099@sju.ac.kr

\*\* 세종대학교 정보보호학과, alwaysgabi@sejong.ac.kr(교신저자)

※ 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. RS-2024-00419201).

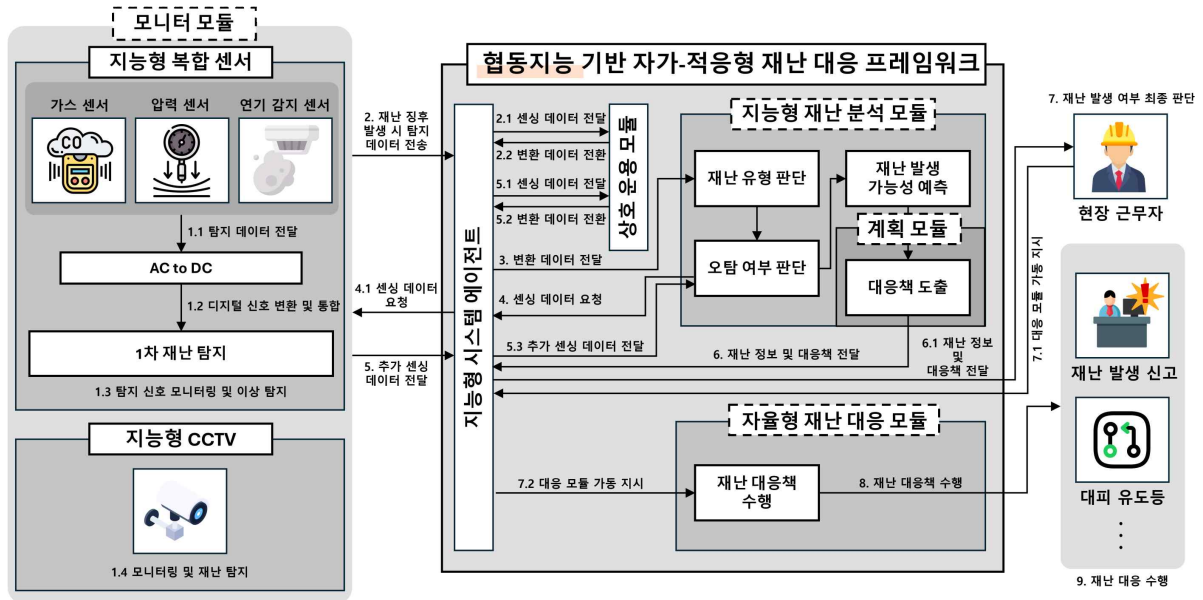


그림 1. 협동지능 기반 자가-적응형 재난 대응 프레임워크 개요

하였으며, 행정안전부는 모든 CCTV를 2027년까지 모두 지능형 CCTV로 대체할 계획을 수립하는 등 정책적 지원도 함께 이루어지고 있다[1,2]. 또한, 설치되는 CCTV가 증가함에 따라, 많은 지자체에서 CCTV 통합 관제센터를 운영하여 재난 및 범죄를 탐지하고자 노력하고 있다. 이러한 국가적 차원에서의 노력에도 불구하고, 기존 대부분의 CCTV 기반 재난 탐지 시스템은 발생한 재난 및 범죄 탐지에 중점을 두고 있으며, 이러한 사후 대응 방식은 재난이 발생하기 전에 선제적으로 대응하기에 한계가 있고, 그로 인해 인적·물적 피해를 효과적으로 최소화하지 못하고 있다[3]. 특히, 재난 및 범죄 발생 시 특정 장소에서의 실시간 모니터링은 보안관제요원의 육안 관제에 대한 의존도가 높고, CCTV의 설치가 증가하더라도 사각지대가 여전히 존재함에 따라 예상치 못한 사고가 발생하더라도 탐지하지 못할 확률이 높다[4]. 이와 더불어, 현재 설치된 지능형 CCTV의 규칙 기반 이상 탐지 기술의 경우 환경 변화나 복잡한 상황에서 오탐율이 높아 실제 상황에서의 적용이 어려우며, 최근에는 기존에 발생하지 않던 새로운 유형의 재난 (예: 군중 밀집 사고 등)의 발생으로 인해 대규모 인명과 재산 피해가 발생하는 사례 또한 증가하고 있다[5].

따라서, 다양한 재난을 신속하게 예측하고 선제

적으로 대응할 수 있으며, 재난이 발생하였을 경우 자율적으로 대응할 수 있는 시스템의 필요성이 더욱 강조되고 있다. 이러한 필요성을 바탕으로, 본 논문에서는 협동지능 기반의 자가-적응형 재난 대응 프레임워크를 제안한다. 제안된 프레임워크는 Monitor, Analyze, Plan, Execute over a Knowledge (MAPE-K)를 기반으로 여러 CCTV 및 센서의 데이터를 통합하여 재난 발생 가능성을 분석하고, 발생한 재난에 대해 자율적으로 대응책을 제시하여 재난 발생으로 인한 피해를 최소화할 수 있도록 설계되었다.

## II. 협동지능 기반 자가-적응형 재난 대응 프레임워크

본 논문에서 제안하는 협동지능 기반 자가-적응형 재난 대응 프레임워크는 지능형 CCTV 및 센서에서 수집한 데이터를 기반으로 다양한 재난 발생 징후를 식별하고, 재난 발생 가능성을 예측하여 자율적인 재난 대응을 수행하는 것을 목표로 한다. 그림 1은 본 논문에서 제안하는 프레임워크의 개요로써, 프레임워크를 구성하는 구성 요소와, 각 구성 요소의 역할을 나타낸다. 제안된 프레임워크는 모니터 모듈, 지능형 재난 분석 모듈, 계획 모듈, 지능형

재난 대응 모듈로 구성되어 있으며, 각 과정에서 생성된 데이터는 Knowledge로써 더 효과적인 대응책을 제시하기 위해 서로 공유된다.

모니터 모듈은 다양한 재난을 탐지하기 위해 여러 센서를 통합한 지능형 복합 센서와 지능형 CCTV로 구성된다. 지능형 복합 센서에서 수집된 센싱 데이터는 AC to DC 변환기를 통해 디지털 신호로 통합되며, 지능형 CCTV는 범죄 및 재난 상황 탐지를 위한 이상 객체 또는 이상 현상 (예: 연기 발생 등)을 식별하고, 각 수집된 데이터를 1차적인 재난 발생 여부 탐지에 사용한다. 모니터 모듈은 수집한 데이터를 기반으로 재난이 발생할 가능성이 높은 재난 징후를 탐지하여 지능형 시스템 에이전트를 통해 센싱 데이터 및 영상 데이터를 전달하고, 상호운용 모듈을 통해 이기종 기기로부터 수집된 데이터를 공통된 형태로 통합하여 지능형 재난 분석 모듈에 제공한다.

지능형 재난 분석 모듈은 제공받은 데이터를 기반으로 발생한 재난의 유형 (예: 화재 또는 군중 밀집 사고 등) 및 1차 재난 탐지의 오탐 여부를 판단하고자 시도한다. 이때, 지능형 재난 분석 모듈이 재난을 판단하기 위한 데이터가 부족한 경우, 지능형 재난 분석 모듈은 지능형 시스템 에이전트를 통해 센서 및 CCTV로부터 추가 데이터를 제공받을 수 있다. 마지막으로 분석 모듈은 데이터를 기반으로 재난의 유형을 특정하고, 재난이 발생할 가능성을 예측한다.

재난이 특정되고 발생 가능성이 예측되었다면, 계획 모듈은 특정된 재난에 특화된 대응책을 식별하여 지능형 시스템 에이전트에 전달하며, 이때 식별되는 대응책은 일반 대중의 안전 및 대피를 최우선으로 한다. 식별된 재난 대응책은 현장 근무자에게 전달되어 현장 근무자가 재난 발생 여부에 대한 최종 결정을 수행한다. 현장 근무자가 재난 발생 정보 및 대응책을 최종 검토하고, 재난 발생 여부와 대응책이 유효하다 판단한 경우, 현장 근무자의 대응 모듈 가동 지시에 따라 자율형 재난 대응 모듈이 작동되어 식별된 재난 대응책이 수행된다.

결과적으로 본 프레임워크는 각종 센서, 지능형 CCTV 및 MAPE-K를 구성하는 각 요소들이 서로

상호작용하며, 협동지능을 기반으로 재난을 예측 및 판단하고 최적의 대응책을 식별하여 자율적으로 재난에 대응한다.

### III. 결론

본 논문에서는 다양한 재난을 신속하게 예측하고 자율적으로 대응할 수 있는 협동지능 기반 자가-적응형 재난 대응 프레임워크를 제안하였다. 제안된 프레임워크는 MAPE-K를 기반으로 여러 센서 및 지능형 CCTV 데이터를 통합 분석하여, 재난 발생 가능성을 실시간으로 예측하고, 자율적인 대응책을 제공함으로써 재난으로 인한 피해를 최소화할 수 있도록 설계되었다. 본 프레임워크는 기존 CCTV 기반 시스템의 한계를 극복하고, 다양한 신종 재난 상황에 보다 효과적으로 대응할 수 있는 가능성을 제시하였다.

### 참 고 문 헌

- [1] e-나라지표, "공공기관 CCTV 설치 및 운영", [https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx\\_cd=2855](https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=2855).
- [2] 행정안전부, "지자체 CCTV 지능형 관제체제로 전환, 현장에서 답을 구한다", May 2024.
- [3] F. Carta, C. Zidda, M. Putzu, D. Loru, M. Anedda, and D. Giusto, "Advancements in Forest Fire Prevention: A Comprehensive Survey", *Sens.*, vol. 23, no. 14, pp. 1-26, July 2023.
- [4] J. Kim, M. -G. Kim, and S. Pan, "A study on implementation of real-time intelligent video surveillance system based on embedded module", *JVIP*, vol. 35, no. 2021, pp. 1-22, November 2021.
- [5] A. Elmetwally, R. Eldeeb, and S. Elmougy, "Deep learning based anomaly detection in real-time video", *Multimed. Tools Appl.*, pp. 1-17, May 2024.

# 공중 무인체계 상호운용성 평가 개념 확대

홍 정 희\*

## Expanding the Concept of Interoperability Evaluation for Unmanned Aerial System

Hong, JeongHee\*

공중 무인체계(UAS)는 크게 무인기(UAV)와 통제장비(GCS)로 구성되어 있고, 무인체계마다 무인기와 통제장비 간 고유한 프로토콜의 연통형(stove-pipe) 방식으로 운용되어 현재는 무인체계 간 상호연동이 불가하다. 점차 다영역작전 등을 대비한 공중 무인체계의 운용개념과 요구사항은 다기종 무인체 운용으로 변화되어, 하나의 통제장비로 타 무인체계의 무인기를 통제하고 그 무인기에서 획득한 영상/이미지 데이터를 원 통제장비에서 수신 가능하도록 요구하고 있다.[1] 이를 위해서는 통제장비 개발 시 공통된 표준 아키텍처와 표준 연동 프로토콜 적용이 필요하며[1], 이러한 표준의 적용을 확인하고 상호운용을 보장하기 위해서는 무인기와 통제장비 간에도 상호운용성 평가가 필요하다. 그러나 '각각 다른 운용 목적을 가진 2개 이상의 체계 간', '서로 다른 체계 간'의 정보교환과 운용 능력의 기존 국방 상호운용성 개념[2]하에서는 무인체계를 하나의 체계로 보고, 해당 무인체계와 연동되는 타 지휘통제체계 등 외부 연동체계 간에만 상호운용성 평가를 수행하여 왔다. 따라서 우선 공중 무인체계 상호운용성 개념을 무인체계 내부를 포함하여 확대시킬 필요가 있다. 확대된 공중 무인체계 상호운용성 개념을 기반으로 상호운용성 평가를 수행하기 위해서는 해당 평가항목을 식별하여 도출해야 하며, 이때 평가항목 적용여부를 결정하는 사업특성분류 유형이 필요하다. 현재는 체계별 사업특성에 따라 4개의 사업특성분류 유형 [2]이 있다.

본 논문에서는 중대형급 연구개발 국방 공중 무인체계 관련하여 다기종 무인체 운용으로 확대되는 상호운용성 요구에 따라 기존 상호운용성 개념을 보완하여 공중 무인체계 상호운용성 개념을 정의하였다. 이를 바탕으로 상호운용성 평가를 위한 범위를 공중 무인체계를 위한 표준 연동 아키텍처와 인터페이스 등을 제공하는 NATO의 STANAG-4586[1]을 기반으로 분석 식별하여 제안하였다. 또한 상호운용성 평가에 필요한 사업특성분류 유형에 대하여 기존 유형에 적절하지 않음을 보이고 공중 무인체계의 운용개념과 특성에 기반하여 사업특성분류 유형 신설을 제안하였다.

### Key words

Defense Interoperability test, Interoperability, Unmanned Aerial System, Unmanned Aerial Vehicle, UAS, UAV

### 참고문헌

- [1] NATO Standardization Office(NSO), "Standard Interfaces Of Unmanned Aircraft(UA) Control System(UCS) For NATO UA Interoperability interface Control Document", AEP-84 Vol. II, Edition A, Version 1, 2017.
- [2] 국방부, "국방상호운용성 관리지시", 2023.

---

\*한국국방연구원, jhhong@kida.re.kr

# 위성 DEM을 활용한 경사와 거리에 따른 경로 제안

변공규\*, 유선진\*\*

## Proposing the Path by Slope and Distance using Satellite DEM

Gongkyu Byeon\* and Sunjin Yu\*\*

### 요 약

본 연구에서는 위성 DEM(Digital Elevation Model) 데이터를 이용하여 경사도와 거리에 가중치를 적용하는 A\* 알고리즘 기반 경로 최적화 시스템을 개발한다. 이 VR 시스템은 경사도와 거리를 동시에 고려하여 최적의 이동 경로를 제공하며, 경사도 가중치와 거리 가중치를 통해 경사가 완만하고 거리가 짧은 경로를 우선적으로 선택하도록 설계되었다. 실험 결과 경사도가 낮고 거리가 적절한 경로를 최적 경로로 선정하여 병력 피로도를 낮추고 빠른 이동에 기여했다.

### Abstract

In this study, we develop an A\* algorithm-based path optimization system that applies weights to slope and distance using satellite DEM(Digital Elevation Model) data. This VR system provides an optimal travel route by considering both slope and distance at the same time, and is designed to preferentially select a route with a gentle slope and a short distance through slope and distance weights. As a result of the experiment, a route with a low slope and an appropriate distance was selected as the optimal route, thereby lowering troop fatigue and contributing to fast movement.

### Key words

virtual reality, DEM data, slope weight, distance weight, path simulation

## I. 서 론

군사 작전에서 지상군의 이동 경로를 최적화하는 것은 작전 성공과 병력의 안전을 위해 매우 중요하다[1]. 지형의 경사, 거리, 잠재적 위협 요소를 고려하여 최적의 경로를 선택하는 것은 전투 효율성을 높이고 병력의 피로를 줄이는 데 중요한 역할을 한

다. 본 연구에서는 유니티 엔진을 활용하여 A\* 알고리즘 기반 VR 경로 최적화 시스템을 개발한다[2]. 이 시스템은 경사와 거리에 가중치를 적용하여 지상군에게 가장 적합한 경로를 제공한다. 산지에 특화된 지형 특성을 분석하고 거리를 고려함으로써, 전시 상황에서 작전 효율성을 극대화하는 것을 목표로 한다.

\* 국립창원대학교 문화융합기술협동과정 박사과정, bgk9517@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr (교신저자)

※ 본 논문은 2023년도 국방기술품질원의 재원으로 방산혁신클러스터 지원(DCL2020L, 2020년 방산혁신클러스터 방산소재 부품연구실 사업)을 받아 수행된 연구의 일부임

## II. 경사와 거리 가중치를 이용한 경로 제안

지역은 대한민국의 경상남도를 연구 대상 지역으로 선정하여, 해당 지역의 DEM(Digital Elevation Model) 데이터를 활용해 지형을 모델링한다. 유니티 엔진에서 Terrain 오브젝트를 생성하고, 고도 데이터를 기반으로 3D 지형과 VR 환경을 구현한다. 본 연구에서는 경사 가중치( $\alpha$ )와 거리 가중치( $\beta$ )를 도입하여 경로의 효율성을 평가하는 점수 계산 방법을 제안한다. 경로의 총 경사 비용은 각 구간의 경사를 계산하고 경사 가중치  $\alpha$ 를 곱하여 계산한다.

$$\text{SlopeCost} = \sum_{i=1}^{n-1} \left( \frac{|h(i+1) - h(i)|}{d(i, i+1)} \times \alpha \right) \quad (1)$$

시스템의 최종 점수는 각 가중치를 적용한 총 거리와 총 경사 비용의 합의 역수를 취하여 계산한다. 거리가 짧고 경사가 완만할수록 더 높은 점수를 부여할 수 있다.

$$\text{Score} = \frac{1}{(\text{Distance} \times \beta) + \text{SlopeCost}} \quad (2)$$

예를 들어, 경사 가중치  $\alpha$ 를 5, 거리 가중치  $\beta$ 를 1.2로 가정하고, 경로의 총 거리와 총 경사 비용이 각각 3500m와 450일 경우, 최종 점수는 다음과 같이 계산한다:

$$\text{Score} = \frac{1}{(3500 \times 1.2) + 450} \approx 0.000215 \quad (3)$$



그림 1. 유니티 엔진으로 구현된 경로 제안 시스템  
Fig. 1. Proposition System Implemented by Unity Engine

탐색된 경로는 유니티 환경에서 실시간으로 시각화되며, 각 경로의 총 거리, 경사 비용, 총 비용이 분석되어 제공된다. 그림 1은 시작 지점부터 끝 지점까지 선택된 경로에 따라 각 가중치에 따른 세 개의 라인을 나타낸다. 색상은 빨강, 초록, 파랑에 맞춰 각각 A, B, C 경로를 지정하였다.

표 1. 세 가지 주요 경로에 대한 기울기 및 거리 평균 점수  
Table 1. Slope and Distance Average Score for Three Main Paths

Paths	Average Score		
	Slope Score	Gradient Score	Total Score
A	634.7341	33.08394	0.0001567294
B	711.6739	795.0484	0.0001263937
C	730.4618	1523.865	0.0001132697

표 1은 세 가지 경로에 대한 기울기 및 거리 평균 점수이다. 경사와 거리를 모두 고려한 결과, 상대적으로 낮은 경사와 적절한 거리를 가진 경로 A가 최적의 경로로 선택되었다.

## III. 결 론

본 연구에서 제안한 경사와 거리 가중치 기반의 A\* 알고리즘은 지상군의 이동 경로 최적화에 실질적인 도움을 줄 수 있다. 특히, 경사와 거리를 동시에 고려함으로써 보다 현실적인 경로를 VR 환경 안에서 제시할 수 있다. 향후 연구에서는 다양한 지형 조건과 날씨, 적의 위치와 같은 실시간 변수 등 동적인 요소를 추가하여 알고리즘의 현실성을 더욱 높일 예정이다.

## 참 고 문 헌

- [1] R. Qiu, Z. Ju, X. Gong, X. Zhang, G. Tao and J. Gong, "Real-Time Terrain-Aware Path Optimization for Off-Road Autonomous Vehicles." In 2024 IEEE Intelligent Vehicles Symposium (IV), pp. 2078-2084, 2024.
- [2] H. Ge, Z. Ying, Z. Chen, W. Zu, C. Liu and Y. Jin, "Improved A\* algorithm for path planning of spherical robot considering energy consumption," Sensors, vol.23, no.16, 7115, 2023.

# 실시간 기상 상태와 도로 기하구조 데이터를 활용한 도로 위험도 예측 기법 연구

박종우\*, 우리나라\*, 서영균\*\*, 이우진\*\*\*

## Study on road severity prediction method using real-time weather condition and road geometry data

Jong Woo Park\*, Rinara Woo, Young-Kyoon Suh\*\* and Woo Jin Lee\*\*\*

### 요 약

본 논문은 도로 위험도 예측 방법을 탐색한다. 기존 방법은 주로 정적 데이터, 즉 교통사고 발생 빈도와 교통사고 데이터의 회귀분석에 의해 도출된 영향계수를 기반으로 도로 위험성을 산출하였다. 그러한 방법의 문제점 중 하나는 신규 도로와 같이 사고는 발생하지 않았으나 환경적인 조건에 의해 위험할 수 있는 상황을 도로 위험도에 반영하지 못한다는 것이다. 이러한 문제를 해결하기 위해, 본 논문은 실시간 기상 상태와 도로 기하구조 데이터를 활용하여 도로 위험도를 예측할 수 있는 새로운 기법을 제안한다. 실험 결과, 제안한 방법이 실시간 기상 데이터 및 도로 기하구조를 반영하여 도로 위험도를 효과적으로 예측할 수 있음을 확인하였다.

### Abstract

This paper explores a road severity prediction method. Existing methods mainly calculate road risk based on static data, such as traffic accident occurrence frequency and coefficients derived from regression analysis of traffic accident data. One of the problems with the existing methods is that they do not reflect situations that may be dangerous due to environmental conditions, such as new roads, even though no accidents have occurred in the road severity. To address this concern, this paper proposes a new method to predict road severity utilizing real-time weather conditions and road geometry data. Our experimental results show that the proposed method can effectively predict road severity by reflecting real-time weather data and geometry.

### Key words

Road Severity, Traffic Accident, Road Geometry, Weather Condition

---

\* 경북대학교 임베디드소프트웨어연구센터, {jwpark, skymin16}@cest.re.kr

\*\* 경북대학교 IT대학 컴퓨터학부, yksuh@knu.ac.kr

\*\*\* 경북대학교 IT대학 컴퓨터학부, woojin@knu.ac.kr(교신저자)

※ 본 연구는 산업통상자원부와 한국산업기술진흥원의 “지역혁신클러스터육성(R&D, P0025885)”사업의 지원을 받아 수행된 연구결과임.

## 1. 서 론

최근 교통사고를 줄이기 위해서 도로의 위험도 정보를 운전자에게 미리 제공하여 운전자가 사전에 주의 운전하여 사고를 예방하도록 하는 서비스가 제공되고 있으며, 도로 위험도 산정 방법에 관한 다양한 연구 결과가 발표되었다[1,2].

우리나라에서는 대표적으로 그림 1과 같이 한국도로교통공단의 ‘위험도로예보시스템’[3]에서 도로 링크별 위험등급 정보를 제공한다. 이 시스템은 도로 링크의 교통사고 건수, 사상자 수, 사상자 환산계수 등을 도로 위험지수의 기본값으로 채택하여 교통사고 다발 구역이 사고 위험도가 높다는 점을 반영하고 있다. 그리고 기상 상태, 소통 상황, 주야 시간대를 영향 변수로 설정하여 환경적 위험성을 반영함으로써 최종 도로 위험도 지수를 결정한다.

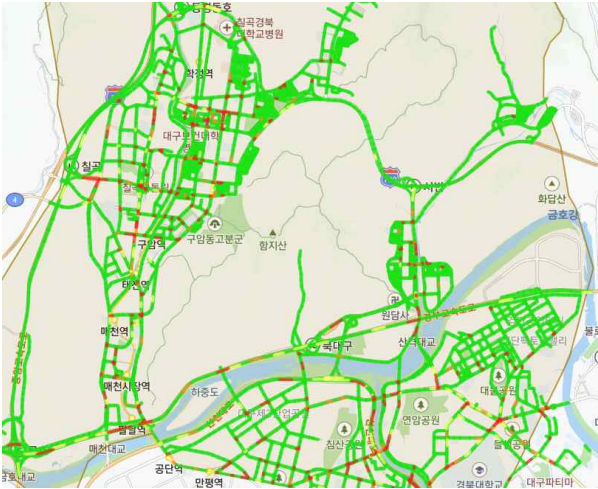


그림 1. 도로교통공단 위험도로예보시스템  
Fig. 1. Korea Road Traffic Authority TAAS Caution Forecast System

그러나 영향 변수의 계수를 도출할 때 교통사고 DB를 다중 회귀 분석하는 방법을 이용하면서 기상 상태가 맑을 때보다 강우, 강설, 안개일 때 영향력이 더 작은 때도 있으며, 주야 시간대도 주간과 야간의 영향력이 같게 도출되는 경우가 있어 운전자가 주의해야 하는 상황과 반대로 표현되는 경우가 보인다.

표 1, 2는 한국도로교통공단에서 제공하는 2020년에서 2023년까지 최근 3년간의 교통사고 DB의

raw data[4]를 수집하여 기상 상태별 사고 건수당 사상자 비율과 주야간 구분별 사고 건수당 사상자 비율을 분석한 자료이다.

표 1. 기상 상태별 사고 건수당 사상자 비율  
Table 1. Casualty rate per accident by weather condition

	맑음	비	흐림	눈	안개
사망자수	1.3%	1.9%	2.8%	1.7%	7.8%
중상자수	27.4%	28.8%	29.7%	24%	44.7%
경상자수	107.5%	108.9%	103.5%	130.8%	94.2%
부상신고자수	9.2%	8.8%	9%	10.5%	7.3%

표 2. 주간/야간 구분별 사고 건수당 사상자 비율  
Table 1. Casualty rate per accident by day/night

	주간	야간
사망자수	1.2%	1.7%
중상자수	27.5%	27.6%
경상자수	108.2%	106.4%
부상신고자수	9.1%	9.3%

분석 결과에 따르면 맑은 날씨보다 비가 오는 날씨에 약 1.4배, 흐린 날씨에 약 2.1배, 눈이 오는 날씨에 약 1.3배, 안개가 낀 날씨에 약 5.9배 더 많은 사망자가 발생했다는 점을 확인할 수 있다. 또한 주간보다 야간에 약 1.4배 더 많은 사망자가 발생했다는 점을 확인할 수 있다.

따라서 기상 상태와 시간대에 따른 영향계수에 대한 조정이 필요하다. 또한 도로 기하구조와 곡선 반경에 따른 안전속도 등 다양한 조건이 반영된 도로 위험도 예측 기법에 대한 연구가 필요하다.

본 논문은 실시간 기상 상태와 주야간 시간대, 노면 상태, 곡선반경, 곡선반경별 안전속도를 영향 변수로 설정하고 변수별 영향계수를 설정하여 도로 위험도를 실시간으로 예측할 수 있는 새로운 기법을 제안한다. 그리고 제안한 방법이 실시간 기상 데이터 및 도로 기하구조를 반영하여 도로 위험도가 효과적으로 예측할 수 있음을 확인하였다.

## II. 도로 위험도 예측 기법

기상 상태와 주야간 시간대의 영향계수는 표 1과 2에 따라 설정하며, 기상 상태는 맑은 날씨, 시간대는 주간을 기준으로 각각 1로 설정하고, 다른 조건



들은 사망자 수 비율에 맞춰 조정한다.

노면 상태, 곡선반경, 제한속도에 따른 영향계수는 기존 연구에서 제시한 안전속도 산정에 대한 가중치를 참조한다[5,6].

노면 상태에 따른 영향계수는 건조, 젖음, 비, 적설, 결빙으로 구분된다. 노면 상태의 실시간 정보 수집이 어렵기 때문에, 도로에서 가장 가까운 기상 관측소의 실시간 기상 정보를 활용하여 노면 상태를 설정한다. 맑은 날씨에는 ‘건조’, 강우 시에는 ‘비’, 강설 시에는 ‘적설’, 강우 또는 강설 시에 기온이 영하 10도 이하일 경우에는 ‘결빙’으로 반영한다. 젖음 상태는 기상 상태 정보만으로 판별하기 어렵기 때문에 제외한다.

곡선반경에 대한 영향계수는 460m 이하, 460m 초과 710m 이하, 710m 초과 2500m 이하를 기준으로 사고 발생 비율로 결정한다.

곡선반경별 안전속도는 곡선반경에 따라 물리적으로 안전한 최대 속도를 의미한다. 각 도로 링크의 제한속도가 안전속도 이하이면 영향계수는 1이며, 안전속도를 초과하면 제한속도를 안전속도로 나누 값이 곡선반경별 안전속도에 대한 영향계수로 결정된다. 이 방식에 의하여 안전속도와 제한속도 간 차이가 클수록 영향계수가 증가한다.

변수별 영향계수는 일반적인 조건을 기준으로 한 사고율에 대한 수치이므로 식 1과 같이 모든 변수의 영향계수를 곱하여 최종 도로 위험지수를 산출한다.

$$s = x_{weather} \times x_{time} \times x_{roadcond} \times x_{curve} \times x_{speed} \quad (1)$$

각 영향계수의 최솟값은 1이며, 기상 상태는 5.915, 주야간 시간대는 1.383, 노면 상태는 1.603, 곡선반경은 14, 곡선반경별 안전속도는 4.04의 최댓값을 가질 수 있다. 따라서 최종 도로 위험지수는 1 ~ 741.685 범위의 값으로 산출된다.

위험등급은 도로 위험지수의 구간에 따라 0 ~ 7 까지 총 8단계로 구분한다. 위험지수가 50 이하일 때는 0단계이며, 그 이상은 10단위로 한 단계씩 증가한다. 그리고 110 이상일 때는 7단계이다.

아래 그림 2, 3은 대구지역을 대상으로 흐린 날

씨와 맑은 날씨에서 링크의 도로 위험등급을 표시한 지도이다.

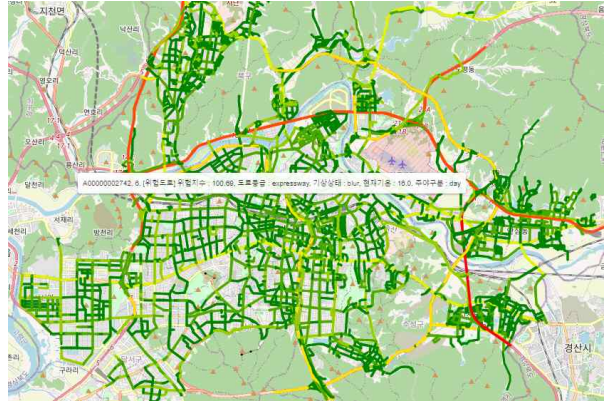


그림 2. 대구지역 도로 위험등급 지도(흐림, 주간)  
Fig. 2. Road severity level map in Dae-gu(blur, day)

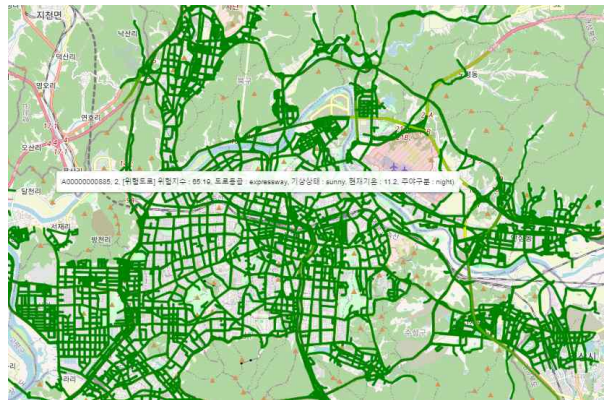


그림 3. 대구지역 도로 위험등급 지도(맑음, 주간)  
Fig. 3. Road severity level map in Dae-gu(sunny, day)

지도에 표시된 위험등급을 확인하면 흐린 날씨에는 대체로 고속도로와 도시고속도로의 위험등급이 일반 도로보다 높게 산출되는 경향이 있다. 맑은 날씨에는 흐린 날씨에 위험등급이 높았던 고속도로와 도시고속도로의 위험등급이 줄어든다. 이는 영향 변수 중 높은 최댓값을 가지는 기상 상태가 도로 위험지수에 크게 영향을 미치기 때문이며, 곡선반경별 안전속도도 최댓값이 비교적 높지만, 정적 데이터 변수이므로 실시간 데이터에 의한 도로 위험지수는 기상 상태에 따라 크게 변화된다.

### III. 결 론

본 논문에서는 기상 상태와 도로 기하구조를 반

영한 도로 위험도를 예측하는 기법을 제안하였다. 제안된 방법은 기상 상태와 도로 기하구조에 따른 교통사고율을 기반으로 결정된 영향 계수값을 반영하여 도로 위험지수를 산출한다.

다만 영향계수를 결정할 때 도로의 형태에 따라 분류하지 않았고, 위험등급을 결정하는 구간을 실험적으로 정의하였기 때문에 이를 객관적으로 검증하는 방안에 관한 연구를 진행할 예정이다. 또한 도로 위험도를 판별할 수 있는 추가적인 영향 변수를 분석하여 도로 위험지수 산출식에 추가할 예정이다.

### 참 고 문 헌

- [1] 권건안, 박상민, 정하림, 권철우, 윤일수, “교통 빅데이터를 이용한 전국 도로 안전성능합수 및 안전등급 개발 연구”, 한국ITS학회논문지, 제18권 제5호, 2019. 10.
- [2] 양충현, 김진국, “동적·정적 자료 기반 도로위험도 산정 알고리즘 개발”, 한국ITS학회논문지, 제19권 제4호, 2020. 8.
- [3] 한국도로교통공단 위험도로예보시스템: <https://taas.koroad.or.kr/TCFS/>
- [4] 한국도로교통공단 교통사고분석시스템: [https://taas.koroad.or.kr/gis/mcm/mcl/initMap.do?menuId=GIS\\_GMP\\_STS\\_RSN](https://taas.koroad.or.kr/gis/mcm/mcl/initMap.do?menuId=GIS_GMP_STS_RSN)
- [5] 강승림, 박창호 “고속도로 선형조건과 GIS 기반 교통사고 위험도지수 분석”, 대한교통학회지, 제21권 제1호, 2003. 2.
- [6] 김근혁, 이준범, 이수범, 강동수, 홍지연, “도로환경특성을 고려한 안전속도 산정에 관한 연구”, 한국도로학회 논문집, 제14권 제4호, 2012. 8.

# 기계학습법을 이용한 국내 중기 전력 수요 예측

이근철\*

## Domestic Monthly Load Forecasting using Machine Learnings

Geun-Cheol Lee\*

본 연구는 한국의 월별 전력 수요를 예측하기 위해 회귀 기반의 예측 모델을 제안한다. 제안된 모델은 기온, 태양 복사량 등의 기상 변수, 캘린더 데이터, 제조업 활동 지표 등 다양한 외부변수를 포함하여 국내 중기 전력 수요에 영향을 미치는 주요 요인을 고려하였다. 기존 전력 수요 연구에서 사용된 요인뿐만 아니라, 국내 전력 특성을 반영한 요인을 고려함으로써 예측 정확도를 높이고자 하였다. 본 연구에서는 이러한 예측 변수를 데이터 분석을 통해 선정하고, 기존의 단변량 시계열 모델과 최신 머신러닝 기법과의 비교 실험을 수행하였다. 비교 대상 모델로는 Holt-Winters, SARIMA, Prophet, XGBoost, Random Forest, RNN, LSTM 등이 포함되며, 각 모델의 예측 성능을 MAE, RMSE, MAPE와 같은 평가 지표를 통해 분석하였다. 실험 결과, 제안된 회귀 모델은 모든 평가 지표에서 기존 방법들을 상회하는 성능을 보여주었으며, 특히 기존 방법들을 혼합한 앙상블 방법보다도 높은 예측 정확도를 기록하였다. 본 연구는 중기 전력 수요 예측을 위한 회귀 기반 접근법이 효과적임을 입증하며, 국내 전력 시스템 관리와 자원 배분에 기여할 수 있을 것으로 기대된다.

### Key words

mid-term load forecasting, time series analysis, regression, machine learning, ensemble methods

---

\*건국대학교 경영대학, gclee@konkuk.ac.kr

※ 이 논문은 2024년도 건국대학교 KU학술연구비 지원에 의한 논문임

# 물리학 정보 기반 도메인 적응: 낙동강 하구의 염분 침입 예측

유창연\*, 강지훈\*\*

## Physics-Informed Domain Adaptation: Predicting Salinity Intrusion in the Nakdong River Estuary

Changyeon Yu\*, and Jihoon Kang\*\*

본 논문은 낙동강 하구의 염분 침입 예측을 위한 물리 기반 도메인 적응 프레임워크를 제시하며, 물리적 원리를 적용해 머신러닝 예측 성능을 향상시키고자 한다. 염분 유입 역학은 강과 해양 조건의 영향을 받아 복잡하며, 수문 정책의 변화에 따라 그 양상이 달라진다. 본 연구는 2022년 2월부터 2023년 6월까지 낙동강 하류에서 10개 모니터링 관측소를 통해 수집된 염분, 수온, 용존 산소 및 pH 데이터 등을 사용한다. 2022년과 2023년의 수문 운영 정책이 상이하여 염분 수준에 영향을 미쳤으며, 특히 2022년 데이터는 수문의 개폐 여부에 큰 영향을 받았을 것으로 판단된다[1]. 도메인 차이에 따른 모델의 일반화 성능 저하 문제를 해결하기 위해, 먼저 유전 알고리즘을 통해 2023년 데이터를 기반으로 최적화된 1차원 물리 모델[2],[3]의 계수를 추정하였다. 추정된 계수를 바탕으로 2022년 데이터의 이론적 염분 값을 계산하고, 실제 관측 값과의 불일치 정도에 반비례하는 중요도 가중치를 설정하였다. 중요도 가중치를 부여하는 방식은 데이터가 2023년 수문 상시 개방 조건의 염분 패턴과 일치하는 정도를 반영하여 관련성이 높은 데이터를 강조한다. 설정된 중요도 가중치는 확장된 선형 회귀 모델의 손실 함수에 통합되어 정확도를 높인다. 본 연구는 도메인 적응 프레임워크 내에서 물리 기반 가중치 전략이 복잡한 환경 시스템 모델의 신뢰성을 높여, 다양한 수문학적 응용에 적용 가능한 구조화된 방법론을 제공할 수 있음을 보여준다.

### Key words

Domain Adaptation, Nakdong River Estuary, Physics-informed, Salinity Intrusion

### 참고문헌

- [1] Woo, J. W., Kim, Y. J., and Yoon, J. S., "Prediction of salinity of Nakdong River estuary using deep learning algorithm (LSTM) for time series analysis," J. Korean Soc. Coast. Ocean Eng., vol. 34, no. 4, pp. 128-134, August 2022.
- [2] Nguyen, D., Umeyama, M., Shintani, T., and Nakaza, C., "Effect of topography on salinity distribution: A case study in a high-tidal range estuary," J. Japan Soc. Civ. Eng., Ser. B1 (Hydraulic Engineering), vol. 68, pp. I\_271-I\_276, January 2012.
- [3] Kim, D., Park, H., and Park, S., "The investigation of sea water intrusion length on opening of Nakdong River estuary barrage using numerical simulation model," J. Korean Soc. Hazard Mitig., vol. 16, no. 5, pp. 299-309, October 2016.

---

\* 한국공학대학교 컴퓨터공학과, kite99@tukorea.ac.kr,

\*\* 한국공학대학교 경영학부, jhkang82@tukorea.ac.kr(교신저자표시)

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (RS-2024-00415938, 2024년 산업혁신인재성장지원사업)

# 대기 순환 고해상도 이미지 복원 모델의 견고한 구축을 위한 물리학 기반 초해상도 GAN

김재만\*, 강지훈\*\*, 양정아\*\*\*

## Physics-Informed Super-Resolution GAN for the Robust Construction of an Atmospheric Circulation High-Resolution Image Restoration Model

Jaeman Kim\*, Jihoon Kang\*\*, Junga Yang\*\*\*

대기의 고해상도 영상은 정확한 기상 분석 및 예측을 위해 필수적이지만, 기후예측자료의 낮은 공간 분해능으로 인해 객체 탐지, 분류, 변화 탐지 등의 분석의 정확도를 떨어뜨리며, 효과적인 대기 및 기후 모델링에 있어서 세부적인 공간 분석의 어려움은 야기한다. 그러나 고해상도 영상의 재구성을 위한 전통적인 방법은 대기의 복잡한 역학을 보존하는 데 어려움을 겪는 경우가 많으며, 이로 인해 대기 구성 요소의 필수 기능이 손실되고 수치해석을 위한 복원과 재현을 보장하지 못한다. 이러한 기존의 이미지 분야에서, 저해상도 이미지를 고해상도로 재구성한 연구는 데이터 보간에서 유래된다[1]. 이후 Neural Network의 등장과 함께, CNN(Convolutional Neural Network) 모델 기반 저해상도에서 고해상도 이미지로의 매핑을 학습하도록 설계된 일련의 컨볼루션 레이어 구성 아키텍처가 활용되었지만, High texture에서 스무딩이 발생하거나 Detail이 떨어지며 지각적인 표현성이 부족하다는 평가가 제기되었다[2]. 2014년에는 이미지를 생성에만 초점을 맞춘 GAN(Generative Adversarial Network) 모델이 등장하였으나, 픽셀 단위의 손실 함수를 최소화하는 데 중점을 두기 때문에, 이 과정에서 데이터의 특성을 잘 반영하지 못할 경우 비현실적이거나 물리적으로 불가능한 이미지가 생성되는 경우들이 관찰되었다[3]. 이러한 한계를 해결하기 위해 본 연구는 AOGCM(대기-해양 대순환 모델)의 예측된 기상장 데이터를 재구성하여 ERA5 재분석 자료의 해상도와 근사한 고해상도 성능을 가진 영상 이미지를 복원하는 것을 목표로 하는 물리 도메인 기반 초해상도 생성 적대 신경망(SR-GAN)을 제시한다. 제안된 SR-GAN은 대류-확산 방정식을 통합함으로써 대기의 이동 및 확산 과정을 제어하는 물리적 제약을 포함하여 생성된 고해상도 이미지가 현실적인 대기 역학을 보존하도록 보장한다.

### Key words

Domain Adaptation, Physics-Informed, Super-Resolution, Neural Network, Generative Adversarial Network

### 참고문헌

- [1] A. Ruiz-Arias, J. Tovar-Pescador, et al., "Intercomparison of Interpolation Techniques for the Estimation of Solar Radiation in North America", *Solar Energy*, vol. 129, pp. 100-112, 2017.
- [2] J. Z. Schwarz, J. S. Stein, et al., "Fast and Accurate Image Super Resolution by Deep CNNs", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 3, pp. 679-693, 2018.
- [3] X. Wang, K. Yu, et al., "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks", *Proc. Eur. Conf. Comput. Vision Workshops*, pp. 63-79, 2018.

\* 한국공학대학교 컴퓨터공학과, 2017312004@tukorea.ac.kr,

\*\* 한국공학대학교 경영학부, jhkang82@tukorea.ac.kr,

\*\*\* 건국대학교 사회환경공학과, yangja.1985@gmail.com (,교신저자)

이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (RS-2024-00415938, 2024년 산업혁신인재성장지원사업)



# 다중 해상도 트랜스포머 기반 자기 지도 사전학습 모델을 활용한 화자 검증 시스템

이재환\*, 서영주\*\*

## Leveraging Multi-resolution Transformer based Self-supervised Pre-trained Model for Speaker Verification

Jaehwan Lee\*, and Young-Joo Suh\*\*

### 요약

본 연구는 다중 해상도 트랜스포머 기반 음성 표현 학습 모델인 MR-HuBERT를 활용하여 자기 지도 학습 기반 화자 검증(Speaker Verification, SV) 성능을 향상시키는 방법을 제안한다. 또한 다중 해상도 특징을 효과적으로 하위 네트워크에 전달하기 위한 feature fusion 모듈을 제안한다. 실험 결과, 다중 해상도 특징과 효과적인 feature fusion 기법을 사용한 SV 모델이 기존 모델보다 향상된 성능을 보였다.

### Abstract

This paper proposes a method to enhance the performance of self-supervised speaker verification (SV) by utilizing MR-HuBERT, a multi-resolution transformer-based speech representation learning model. Additionally, an improved feature fusion module is proposed to effectively convey multi-resolution features to the downstream network. Experimental results demonstrate that the SV model, employing multi-resolution features and an improved feature fusion technique, achieved enhanced performance compared to existing model.

### Key words

speaker verification, multi-resolution, self-supervised, feature fusion

## 1. 서론

화자 검증(Speaker Verification, SV)은 등록 발화와 테스트 발화의 유사도를 계산하여 두 발화의 화

자가 일치하는 지 여부를 판별하는 태스크이다.

딥러닝 기반 SV 연구는 많은 수의 화자를 포함하는 데이터셋을 이용하여 closed-set 화자 분류 태스크로 지도 학습을 진행한 뒤, 학습된 모델을 이용

\* 포항공과대학교 인공지능대학원, jhwan98@postech.ac.kr

\*\* 포항공과대학교 인공지능대학원, yjsuh@postech.ac.kr (교신저자)

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2019-II191906)과 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원(No.2022R1A6A1A03052954)과 2024년도 정부(과학기술정보통신부)의 재원으로 한국지능정보사회진흥원의 지원(2024년 지역특화형 초거대 AI 클라우드 팜 실증 및 확산 환경 조성)을 받아 수행된 연구이며, 「과학기술정보통신부 방송통신발전기금」을 지원받아 정보통신산업진흥원(D1014-24-1009, 철도차량 가용성 향상을 위한 AI 융합형 통합정비 공정 메타버스 플랫폼 개발)의 지원을 받아 수행된 연구이나 과학기술정보통신부의 공식의견과 다를 수 있음.

하여 발화 쌍에 대한 검증 점수를 계산하여 테스트 하는 방식이 주류를 이루었다[1].

하지만 많은 라벨링된 데이터를 구하는 것의 한 계로 인해, 최근에는 자기지도학습(Self-Supervised Learning, SSL)을 활용한 연구가 많이 이루어졌다[2]. SSL 기반 화자 검증 연구는 사전학습 모델 (Pre-trained model, PTM)이 많은 양의 음성 데이터로 학습되어 음성의 일반적인 특징을 학습할 수 있고, 적은 양의 라벨링 데이터로 하위 네트워크를 파인튜닝하여 좋은 성능의 SV 모델을 구축할 수 있다는 장점이 있다. 하위 네트워크를 훈련시킬 때는 PTM의 여러 계층의 출력을 가중합(weighted-sum)하여 하위 태스크 네트워크의 입력으로 전달하는데, 그 이유는 PTM은 사전 학습 과정에서 여러 계층에 걸쳐 특징이 분산되어 저장되기 때문이다[4].

최근 제시된 많은 음성 분야 PTM은 여러 계층의 트랜스포머 구조를 활용하지만 이러한 PTM은 하나의 해상도(resolution) 상에서 음성 특징을 학습한다는 단점이 있다. 현존하는 많은 음성 SSL 모델들은 20ms의 고정 해상도의 음성 유닛을 기반으로 학습되지만, 더 낮은 해상도의 정보 또한 화자 인식과 같은 태스크에 도움이 될 수 있는 정보를 추출하는데 도움을 줄 수 있다.

위 문제점에 기반하여 최근 다중 해상도 상에서 음성의 특징을 학습한 PTM인 Multi-Resolution HuBERT(MR-HuBERT)가 제시되었다[5]. 하지만 이 모델의 SV 하위 태스크 성능에 관한 연구는 충분

히 이루어지지 않았다. 따라서 본 연구에서는 이 모델을 활용하여 SV 하위 태스크 성능을 기존 단일 해상도 모델과 비교하여 다중 해상도 특징의 효용성을 확인하고자 한다. 또한, PTM의 특징 정보를 효율적으로 취합(feature fusion)하여 하위 태스크 네트워크에 전달하기 위해 [3]에서 소개한 Adaptive Multi-scale Feature Fusion(AMFF) 모듈을 개선하여 SV 성능을 높이고자 한다.

위의 방법을 바탕으로 Voxceleb2 데이터셋을 이용하여 파인튜닝을 진행한 결과, 기존 단일 해상도 모델보다 SV 성능이 향상됨을 확인하였다.

## II. 다중 해상도 특징을 활용한 SV 모델

### 2.1 MR-HuBERT

그림 1은 본 연구에서 제시한 SV 모델의 전체 아키텍처를 보여준다.

MR-HuBERT는 두 가지 해상도를 가지는 3개의 트랜스포머 인코더로 구성된다. 입력 음성 신호는 특징 추출 모듈(Conv Feature Extractor)을 거쳐 음성 유닛 시퀀스로 변환된 뒤 마스킹이 적용된다. 이후 High Resolution Encoder를 통과한 뒤, 저해상도로 다운 샘플링된다. 이는 Low Resolution Encoder를 거친 후 다시 원래 해상도로 업 샘플링되며, 이 출력은 첫 High Resolution Encoder의 출력과 더해져 마지막 인코더를 통과한다.

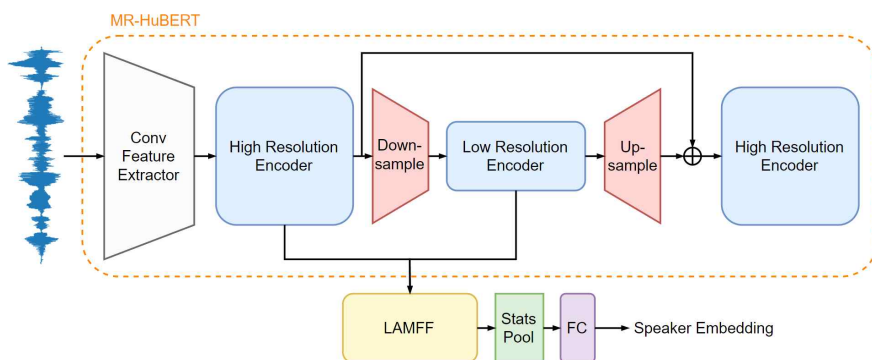


그림 1. MR-HuBERT 기반 Speaker Verification 모델  
Fig. 1. MR-HuBERT based Speaker Verification model



## 2.2 Low-layer Adaptive Multi-scale Feature Fusion(LAMFF)

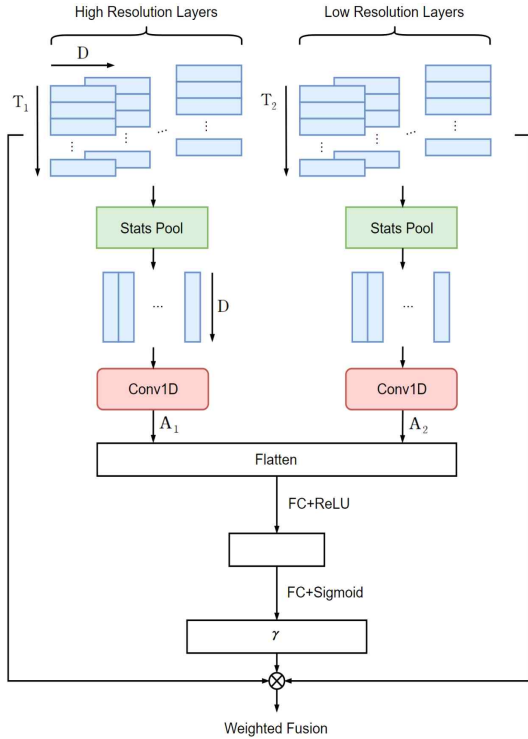


그림 2. Low-layer Adaptive Multi-scale Feature Fusion 모듈

Fig. 2. Low-layer Adaptive Multi-scale Feature Fusion module

그림 2는 본 연구에서 제시한 feature fusion 모듈을 보여준다.

이 모듈은 AMFF를 기반으로 한 모듈이며, 하위 계층의 representation을 중점적으로 활용한다는 특징이 있다. Masked Unit Prediction 태스크로 훈련된 PTM은 주로 하위 계층에 화자와 관련된 특징 정보가 분포하기 때문에[7, 8], MR-HuBERT의 모든 인코더를 사용하는 것이 아닌, 첫 High Resolution Encoder와 Low Resolution Encoder의 출력을 feature fusion 모듈의 입력으로 사용한다. 또한 squeeze 과정에서 단순 Average Pooling으로 인한 정보 손실을 최소화하기 위해 각 해상도별로 Statistic Pooling, 커널 크기 1의 1D 컨볼루션 연산을 적용한 뒤, 모두 flatten시켜 excitation 과정의 입력으로 사용한다. 또한, 해상도에 상관없이 이를 단일 해상도 모델의 하위 계층에도 적용하여 성능을 비교한다. 이를

Low-layer Adaptive Multi-scale Feature Fusion(LAMFF)이라 한다.

## III. 실험 결과

Fine-tuning 데이터셋은 5,994명의 발화 데이터를 가지는 VoxCeleb2의 dev 파트를 사용하며, 추가적인 데이터 증강 기법은 사용하지 않는다.

하위 네트워크는 PTM의 representation 능력을 중점적으로 평가하기 위해 단순 fully-connected 네트워크를 사용한다. 비교 모델은 HuBERT[6]와 AMFF 기반 SV 모델로 설정하며, 학습은 [3]에서 사용한 2-step 학습 방식을 따른다. 본 연구에서는 LayerDrop을 사용하지 않기 때문에 baseline 모델 결과는 직접 진행한 실험 결과를 사용한다. MR-HuBERT, HuBERT 모델은 S3PRL<sup>1</sup>의 체크포인트를 사용하였으며, 모두 base 타입이다.

평가를 위해 VoxCeleb1 test set(Vox1-O)을 사용하며, trial pair 점수는 코사인 유사도로 계산한다. SV 성능평가는 Equal Error Rate(EER)을 사용한다.

표 1. 실험 결과

Table 1. Experiment result

모델	EER(%)	Params.
(1) HuBERT+AMFF+linear(baseline)	1.87	96.14M
(2) HuBERT+LAMFF+linear	1.85	96.15M
(3) MR-HuBERT+AMFF+linear	1.87	97.91M
(4) MR-HuBERT+LAMFF+linear	1.82	97.92M

표 1. 은 실험 결과를 보여준다. (1), (3) 결과를 봤을 때, 기존 AMFF 모듈을 사용할 경우 MR-HuBERT는 HuBERT와 성능 차이가 크게 없지만, LAMFF를 사용했을 때 약 0.01%의 파라미터 수 증가로 더 향상된 성능을 보였다. 단, (3)이 (1)보다 더 큰 모델임에도 불구하고 성능에 큰 변화가 없는 것으로 봤을 때, 단순 파라미터 수 증가에 따른 성능 향상이라고 보긴 어렵다. 이는 다중 해상도 특징 정보를 활용하는 것 뿐만 아니라 효과적인 feature fusion 기법의 영향도 중요하다고 볼 수 있다.

<sup>1</sup><https://github.com/s3prl/s3prl>

#### IV. 결론

본 논문에서는 다중 해상도 PTM과 효과적인 feature fusion 모듈을 이용하여 SV 하위 태스크에 적용했을 때 단일 해상도 모델보다 성능이 향상됨을 확인하였다.

#### 참 고 문 헌

- [1] Variani, Ehsan, et al. "Deep neural networks for small footprint text-dependent speaker verification." in *Proc. ICASSP*. 2014.
- [2] Chen, Sanyuan, et al. "Unispeech-sat: Universal speech representation learning with speaker aware pre-training." in *Proc. ICASSP*. 2022.
- [3] Li, Yishuang, et al. "SR-HuBERT: An Efficient Pre-Trained Model for Speaker Verification." in *Proc. ICASSP*. 2024.
- [4] Yang, Shu-wen, et al. "Superb: Speech processing universal performance benchmark." in *Proc. Interspeech*. 2021.
- [5] Shi, Jiatong, et al. "Multi-resolution HuBERT: Multi-resolution speech self-supervised learning with masked unit prediction." ICLR 2024, The Twelfth International Conference on Learning Representations. ICLR. 2024.
- [6] Hsu, Wei-Ning, et al. "Hubert: Self-supervised speech representation learning by masked prediction of hidden units." *IEEE/ACM transactions on audio, speech, and language processing* 29 (2021): 3451-3460.
- [7] Pasad, Ankita, Bowen Shi, and Karen Livescu. "Comparative layer-wise analysis of self-supervised speech models." in *Proc. ICASSP*. 2023.
- [8] Chen, Sanyuan, et al. "Why does self-supervised learning for speech recognition benefit speaker recognition?." arXiv preprint arXiv:2204.12765 (2022).

# ResNet 기반의 DELG 모델을 활용한 범죄 현장 족윤적 이미지 검색 성능 고도화

류지연\*, 이창엽\*, 서영주\*, 박진\*\*

## Enhanced Crime Scene Footprint Image Retrieval Using an Optimized DELG Model with ResNet

Ji-yeon Ryu\*, Chang-yeop Lee\*, Young-joo Suh\* and Jin Park\*\*

### 요 약

본 연구는 범죄 현장에서 수집된 족윤적 이미지의 검색 성능을 향상시키기 위한 방법을 제안한다. 본 방법에서는 빛, 바닥 재질, 혈흔 등 환경적 요인에 의해 품질이 저하된 족윤적 이미지를 Super Resolution과 Cropping 기법을 통해 정제하였다. 또한, 글로벌 및 로컬 특징을 동시에 추출할 수 있는 Deep Local and Global Features(DELG) 모델을 도입하여 족윤적 도메인 특성을 반영한 변형 구조를 설계하고, 백본으로 ResNet의 4, 5번째의 레이어를 활용함으로써 검색 정확도와 성능을 기존 모델 대비 17.33% 개선하였다. 검증을 위해 경찰청에서 제공한 족윤적 이미지 455장과 직접 수집한 이미지 55장을 활용하였으며, 다양한 거리 측정 방법을 통해 평가의 신뢰성을 높였다.

### Abstract

This study proposes a method to enhance the retrieval performance of footprint images collected from crime scenes. The proposed method refines degraded footprint images affected by environmental factors such as lighting, blood and floor material using Super Resolution and Cropping techniques. Additionally, a customized structure of the Deep Local and Global Features (DELG) model was designed to reflect the specific characteristics of the footprint domain by extracting both global and local features simultaneously, utilizing the 4th and 5th layers of ResNet as the backbone. This approach achieved a 17.33% improvement in retrieval accuracy and performance. For validation, 455 footprint images provided by the National Police Agency and 55 images collected independently were used. Various distance metrics were applied to ensure the reliability of the evaluation.

### Key words

footprint image retrieval, crime scene analysis, DELG model, super resolution, feature extraction, distance metrics

---

\* 포항공과대학교 인공지능연구원, yoen114@postech.ac.kr, lcy8417@postech.ac.kr, yjsuh@postech.ac.kr

\*\* 포항공과대학교 인공지능연구원, spider@postech.ac.kr

※ 본 논문은 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업이며(No. 2022R1A6A1A0305295413), 정부(과학기술정보통신부)의 재원으로 한국지능정보사회진흥원의 지원을 받아 수행된 연구이며, (2024년 지역특화형 초거대 AI 클라우드 팜 실증 및 확산 환경 조성), 과학기술정보통신부·경찰청이 공동 지원한 '폴리스랩 2.0 사업'의(RS-2023-00281072) 지원을 받아 수행된 연구임.

## I. 서론

죽운적 이미지를 이용한 범죄 현장 분석은 사건 수사의 중요한 단서가 될 수 있다. 특히, 죽운적 이미지를 이용한 이미지 검색 기술은 범죄자 추적 및 증거 수집에 매우 유용하다. 그러나 죽운적 이미지는 촬영 환경, 빛, 바닥 재질 등의 환경적 요인에 따라 품질이 크게 달라질 수 있으며, 이는 이미지 간 시각적 유사도를 낮추어 검색 모델의 정확도를 저하하는 주요 원인이 된다. 이러한 특성으로 인해 기존의 이미지 검색 모델을 죽운적 이미지에 그대로 적용하기에는 한계가 존재한다[1]. 또한, 각 데이터 도메인의 특성 때문에 모델을 그대로 적용하기에는 어려움이 있어, 이를 해결하기 위한 맞춤형 접근 방식이 필요하다.

죽운적 이미지의 한계를 해결하기 위해, 본 논문에서는 죽운적의 품질을 개선하기 위한 다양한 전처리 기법의 파이프라인과 도메인 특성을 고려한 특징 추출 모델의 변형을 제안한다. 이 연구에서 선정한 Deep Local and Global Features (DELG) 모델은 이미지의 글로벌 및 로컬 특징을 동시에 추출할 수 있는 강점이 있어, 죽운적 이미지의 복잡한 패턴과 미세한 차이를 효과적으로 분석할 수 있다. 특히, DELG 모델은 글로벌 특징을 통해 전체적인 맥락을 이해하고, 로컬 특징을 통해 이미지의 세부적인 정보를 캡처할 수 있어 죽운적 이미지와 같이 복잡한 패턴을 가진 데이터에 적합하다[2].

본 연구에서는 DELG 모델의 백본으로 ResNet을 활용하되, 특히 Stage 4와 Stage 5 레이어만을 사용하도록 커스텀 구조를 도입하였다. ResNet의 Stage 4는 로컬 특징을 추출하는 역할을 하며, 이는 죽운적 이미지의 작은 패턴이나 세부적인 모양을 강조하는 데 적합하다. Stage 5는 글로벌 특징을 추출하여 이미지의 전반적인 형태와 윤곽을 파악하는 데 활용된다. 이와 같이 ResNet의 4, 5번째 레이어만을 사용하여 모델의 복잡도를 줄이고, 죽운적 이미지의 주요 특징을 효과적으로 추출할 수 있도록 설계하였다[3]. 이를 통해 범죄 현장에서의 죽운적 이미지 검색 성능을 고도화하는 것을 목표로 하였으며, 다양한 전처리 기법과 모델 학습 방법을 실험하고 그 성능을 분석하였다.

## II. 연구 방법

이미지 품질 향상을 위해 데이터 전처리 과정에서 Cropping, Super resolution, Inpainting 기법을 적용하였다[4]. 먼저 Cropping은 관심 영역을 분리하여 가늌자와 같은 불필요한 배경 정보를 제거하는데 사용된다. 이어서 Super resolution 과정을 통해 크기와 해상도가 작아진 죽운적 이미지를 고해상도로 변환하여 세부적인 정보를 복원함으로써 분석의 정밀성을 개선하였다. 마지막으로 Inpainting을 통해 죽운적 이미지의 접합이나 손상된 부분을 복구하여 전체적인 이미지의 품질을 향상시켰다. 이러한 전처리 과정을 통해 모델이 죽운적 이미지의 본질적인 특징을 더욱 정확하게 추출하고, 더 높은 성능을 도출할 수 있도록 하였다.

다음으로 DELG 모델의 변형을 제안하였다. 기존 DELG 모델은 부분적인 특징과 전체적인 특징을 모두 포괄하는 방식으로 작동하여, 죽운적 이미지에서 불필요한 정보까지 포함하게 되어 처리 효율성을 떨어뜨리는 한계가 존재했다[1]. 이러한 한계를 극복하기 위해 본 논문에서는 ResNet의 4, 5번째 레이어를 각각 로컬, 글로벌 추출기로 활용하고, 죽운적 이미지의 글로벌 및 로컬 특징의 가중치를 달리해 중요한 부분을 강조하여 특징을 추출하였다[5].

그림 1에서는 ResNet의 Stage 4 레이어에서 추출된 로컬 특징 맵을 시각화한 것이다. 채널이 진행됨에 따라 죽운적 이미지의 세부 문양이 강조되며, 중요한 패턴이 더욱 뚜렷하게 나타나는 것을 확인할 수 있다. 그림 2는 Stage 5 레이어에서 추출된 글로벌 특징 맵을 보여준다. 채널이 진행됨에 따라 죽운적 이미지의 전반적인 윤곽과 구조가 더욱 강조되며, 전체 형태를 잘 포착하는 특징이 나타난다. 이를 통해 죽운적 이미지의 전체적인 형태를 효과적으로 파악할 수 있는 정보를 제공한다.

본 연구에서 사용된 DELG 모델의 변형은 ResNet의 Stage 4와 Stage 5 레이어만을 사용하여 커스텀화된 백본 구조를 가지고 있으며, 이를 'ResNet\_STAGE45'로 명명하였다. ResNet\_STAGE45에서 추출된 특징을 보다 효율적으로 변환하고 결합하기 위해 Generalized Mean Pooling(GeM)을 도입하였다. GeM은 각 특징의 중요도에 따라 가중치

를 조정하며 풀링을 수행하여, 전통적인 평균 풀링이나 최대 풀링 방식보다 더 세밀한 정보를 보존할 수 있게 한다[1]. 특히, GeM에서 p 값을 4로 설정하여 미세한 차이를 더 효과적으로 반영할 수 있도록 하였으며, 이는 죽운적 이미지와 같이 디테일한 특징을 필요로 하는 작업에서 성능을 극대화하는 데 기여하였다. 이를 통해 모델은 이미지의 중요한 특징을 보다 잘 포착하고, 이를 기반으로 더 정확한 이미지 분석과 검색을 수행할 수 있다. GeM을 도입함으로써, 다양한 환경에서도 검색의 정확도를 크게 향상시키며, 이로 인해 여러 환경에서의 적용 가능성이 증가하였다.

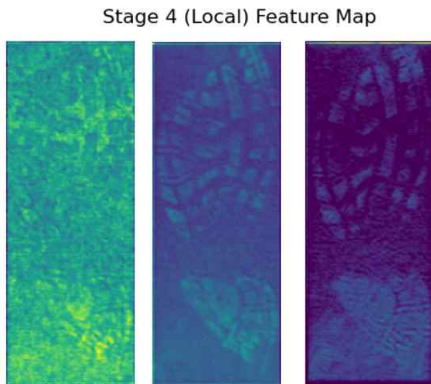


그림 1. Stage 4 로컬 특징 맵  
Fig 1. Stage 4 Local Feature Map

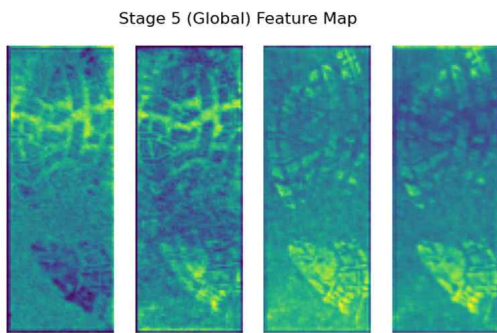


그림 2. Stage 5 글로벌 특징 맵  
Fig 2. Stage 5 Global Feature Map

모델 학습 과정에서는 ArcFace Loss를 활용하여 클래스 간의 차이를 더 명확하게 구분하였다. ArcFace는 죽운적 이미지처럼 미세한 차이가 있는 데이터를 처리하는 데 매우 효과적이며, 이를 통해 학습된 모델은 높은 분별력을 갖게 된다[5]. 또한, Spatial Attention 모듈을 적용하여 이미지의 중요한

지역적 특징에 가중치를 부여하였다. 이 모듈은 지역적 특징을 더욱 강조하여 글로벌 및 로컬 특징을 동시에 학습할 수 있도록 하였다.

결과적으로, ResNet\_STAGE45 구조에서 추출된 글로벌 및 로컬 특징은 각각 GeM Pooling과 Spatial Attention 모듈을 거쳐 ArcFace Loss를 적용받아 학습되었으며, 이를 통해 죽운적 이미지 검색 성능을 최적화하였다.

### III. 실험 및 결과

본 연구에서는 경찰청에서 제공한 죽운적 데이터를 활용하여 죽운적 이미지 검색 모델의 성능을 평가하였다. 사용된 데이터는 쿼리 이미지와 참조 이미지로 나뉘며, 쿼리 이미지는 경찰청에서 제공된 545장의 죽운적 이미지와 직접 수집한 55장의 죽운적 이미지로 구성되었으며 학습에 활용되는 데이터이다. 참조 이미지는 정답 이미지를 ShoeRinsics 모델로 전처리하여 사용하였다[7]. 또한 100만장의 ImageNet 데이터로 학습된 R101-SOLAR 모델의 가중치를 활용하여 전이 학습한 모델로 평가를 진행하였다[6]. 마지막으로 Distance metrics 간의 비교 실험을 통해 최적의 거리 측정 방법을 선정하고, 이미지 검색 분야에서 널리 사용되는 기존 모델들과의 성능을 비교하였다.

실험 결과, 표 1과 같이 기존 학습된 DELG 모델의 성능이 42.67%였던 반면, 직접 커스텀하여 학습시킨 DELG 모델은 Top-50 Accuracy 60.00%로 더 우수한 성능을 보였다. 또한, Super Resolution과 Inpainting을 통해 이미지 해상도와 복원을 개선한 전처리 기법을 적용하여 검색 성능을 고도화하였고, 이를 통해 약 0.18%의 성능 향상을 확인할 수 있었다.

또한, 표 2와 같이 Distance metrics을 변경하여 Euclidean, Cosine, Minkowski, Chebyshev 등의 다양한 거리 계산 방법을 사용해 검색 성능을 평가한 결과, Euclidean distance가 가장 좋은 성능을 나타냈다. 이는 죽운적 이미지의 미세한 차이를 효과적으로 반영하는 Euclidean distance가 죽운적 이미지 검색에 가장 적합함을 시사한다. 반면, Chebyshev distance는 Top-50 Accuracy 29.50%로 가장 낮은

성능을 보였는데, 이는 이 방법이 미세한 차이를 반영하지 않고 두 이미지 간 최대 차이만을 측정하기 때문이다. 이러한 결과는 죽운적 이미지의 특성에 맞춘 전처리와 거리 계산 방법의 선택이 검색 성능에 중요한 영향을 미친다는 점을 보여준다.

표 1. 기존 대비 커스텀 DELG 모델 성능 비교

Table 1. Performance Comparison of Distance Metrics

Model	Top-50 Accuracy (%)
<b>Customized DELG + Preprocessing</b>	<b>60.18</b>
<b>Customized DELG</b>	<b>60.00</b>
DELG	42.67
SOLAR	48.00
Token	48.84
NetVLAD	12.50

표 2. Distance Metrics 성능 비교

Table 2. Performance Comparison of Distance Metrics

Distance Metrics	Top-50 Accuracy (%)
Euclidean	<b>60.00</b>
Manhattan	57.67
Minkowski(p=2.2)	58.84
Chebyshev	29.50
Cosine	59.50
Pearson	59.67

#### IV. 결 론

본 연구에서는 범죄 현장에서 수집된 죽운적 이미지를 효과적으로 검색하기 위해 기존 DELG 모델을 커스텀하여 성능을 고도화하였다. 커스텀 DELG 모델은 ResNet의 Stage 4와 Stage 5 레이어만을 사용하고, 글로벌 및 로컬 특징에 가중치를 달리하여 죽운적 이미지의 특성을 더욱 세밀하게 반영하였다. 이를 통해 기존 DELG 모델 대비 Top-50 Accuracy에서 약 17.33%의 성능 향상을 달성하였으며, Super resolution 및 Inpainting 전처리 기법을 추가하여 죽운적 이미지의 해상도와 품질을 개선함으로써 더욱 높은 검색 정확도를 얻을 수 있었다.

추가적으로, 다양한 Distance metrics을 적용한 실험에서 Euclidean distance가 가장 높은 성능을 나타내어 죽운적 이미지의 미세한 차이를 구별하는데에 최적임을 확인하였다. 이 연구의 결과는 죽운적 이미지의 특성에 맞춘 모델 구조 및 전처리 기

법의 선택이 검색 성능에 큰 영향을 미친다는 점을 시사하며, 이를 통해 실제 범죄 수사에서 효과적인 죽운적 이미지 검색 솔루션을 제공할 수 있는 가능성을 제시하였다.

#### 참 고 문 헌

- [1] Y. Cao, A. Araujo, J. Sim, and T. Weyand, "Unifying Deep Local and Global Features for Efficient Image Search," IEEE Int. Conf. on Computer Vision (ICCV), pp. 4590-4600, October 2021.
- [2] H. Jégou, M. Douze, C. Schmid, and P. Pérez, "Aggregating Local Descriptors into a Compact Image Representation," IEEE Trans. Pattern Anal. Machine Intell., vol. 34, no. 9, pp. 1704-1716, September 2010.
- [3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, June 2016.
- [4] X. Wang, K. Yu, S. Wu, J. Gu, C. Dong, Y. Qiao, and C. C. Loy, "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," European Conf. on Computer Vision (ECCV), pp. 63-79, September 2018.
- [5] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 4690-4699, June 2019.
- [6] O. Russakovsky, et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. of Computer Vision, vol. 115, no. 3, pp. 211-252, January 2015.
- [7] S. Shafique, B. Kong, S. Kong, and C. C. Fowlkes, "Creating a Forensic Database of Shoeprints from Online Shoe Tread Photos," Proc. Winter Conf. on Applications of Computer Vision (WACV), pp. 858-868, January 2023.

# 딥러닝 기반 얼굴 탐지 모델 성능 연구

당탄부\*, 김재석\*, 최지현\*, 장진호\*, 유광현\*\*

## A Study on deep learning based face detection model

Dang Thanh Vu\*, Kim Jae-seog\*, Choi Ji-hyun\*, Jang Jin-ho\*, and Yu Gwang-hyun\*\*

본 논문에서는 딥러닝 기반의 얼굴 탐지 모델에 대한 성능 연구를 제시한다. 최근 몇 년간 컴퓨터 비전 및 인공지능 분야에서 딥러닝 모델은 뛰어난 성과를 보여왔다. 얼굴 영역 탐지는 보안 및 감시시스템, 얼굴 인식 시스템, 그리고 사용자 인터페이스와 같은 분야에서 필수적인 기술이다. 전통적인 컴퓨터 비전 방법들은 특징 추출 알고리즘을 기반으로 얼굴 영역 탐지를 수행해왔으나, 복잡한 환경과 다양한 얼굴 형태에 대하여 성능 한계가 있었다. 딥러닝의 발전으로 인하여 이러한 한계는 극복되고 있으며, 특히 합성곱 신경망(CNN, Convolutional Neural Network)기반 모델들이 얼굴 탐지에서 탁월한 성능을 발휘하고 있다. 본 연구에서는 객체 탐지 분야에서 널리 사용되는 합성곱 신경망 계열 중 하나인 YOLO(You Only Look Once)를 기반으로 얼굴 탐지 모델을 구성하여 성능을 비교한다. 데이터셋은 공개데이터셋인 SCUT-FBP5500을 전체 4,405장 이미지에서 얼굴 마스크는 111,251개를 사용하였고, 얼굴 탐지 모델로는 YOLOv5와 YOLOv8을 과인 튜닝하였다. 실험 결과 YOLOv5기반 얼굴 탐지 모델의 Precision-Recall Curve는 91.7%, F1-Confidence Curve는 91%이고, YOLOv8기반 얼굴 탐지 모델의 Precision-Recall Curve는 93.5%, F1-Confidence Curve는 91%로 YOLOv8기반 얼굴 탐지 모델 성능이 더 높은 것으로 확인되었다. 이를 통해 딥러닝 기반 얼굴 탐지 모델에서 YOLO 계열을 사용할 때 기준을 제시할 수 있을 것으로 기대되고, 추후에는 다양한 데이터셋을 통해 응용 연구를 수행할 계획이다.

### Key words

deep learning, face detection, YOLO, YOLOv5, YOLOv8

### 참고문헌

- [1] Jeong-Tae, K., Hui-Yeon, J., and Eun-Jeong C., "딥러닝 기법을 이용한 머신 비전 기술 최근 응용 동향", The Magazine of the IEIE, Vol. 43-11, pp.18-26, Nov, 2016.
- [2] 문형진, and 김계희., "사용자 인증을 위한 딥러닝 기반 얼굴인식 기술 동향", 대한산업경영학회, Vol.17-3, pp.23-29, Sep, 2019.
- [3] Redmon, J., "You only look once: Unified, real-time object detection", In Proceedings of the IEEE conference on computer vision and pattern recognition, pp.779-788, July, 2016.
- [4] Terven, J., Córdova-Esparza, D. M., and Romero-González, J. A., "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas", Machine Learning and Knowledge Extraction, Vol.5-4, pp.1680-1716, Nov, 2023.

\* (주)에이아이시드, dtvu@aiseed.kr, bossrobot@aiseed.kr, ai\_seed@aiseed.kr, verycosy@kakao.com

\*\* 전남대학교산학협력단, sayney1004@gmail.com(교신저자)

※ 이 논문은 과학기술정보통신부와 정보통신산업진흥원으로부터 '2024년 지역디지털기초체력지원(광주\_SW서비스)(S-2001-24-1005)'과 과학기술정보통신부·광주광역시 공동 지원한 '인공지능 중심 산업융합 집적단지 조성사업'으로 지원을 받아 수행된 연구 결과입니다.

# 음악 소스 분리 모델을 활용한 Wave-U-Net 기반 자동 음악 믹싱 시스템

윤진성\*, 심영준\*, 서영주\*\*

## Automatic Music Mixing Using Music Source Separation Models and Wave-U-Net

Jinsung Yoon\*, Youngjun Sim\* and Young-Joo Suh\*\*

### 요 약

음악 믹싱은 개별 트랙 형태로 악기를 녹음하고, 이를 전문가의 지식을 이용하여 적절히 다듬어 섞는 과정을 뜻한다. 최근 몇 년 동안 딥러닝을 활용한 자동 음악 믹싱이 연구되었지만, 개별 트랙 녹음과 믹싱이 되지 않은 드라이한 녹음 데이터의 부족은 이러한 모델의 성능을 제한시켰다. 본 연구에서는 음악 소스 분리 모델을 이용하여 개별 트랙 녹음을 추출하고 무작위로 볼륨이 조절된 녹음을 생성하여 데이터 부족 문제를 해결한다. 이후 Wave-U-Net 기반의 자동 음악 믹싱 모델을 학습하여 자동 음악 믹싱 시스템을 설계한다.

### Abstract

Music mixing refers to the process of recording individual tracks of instruments and refining them using expert knowledge to blend them appropriately. In recent years, there has been research into automated music mixing utilizing deep learning; however, the lack of dry recording data with individual track recordings and mixing has limited the performance of such models. In this study, we address the data scarcity issue by employing music source separation models to extract individual track recordings and generate recordings with randomly adjusted volumes. Subsequently, we train a Wave-U-Net-based automatic music mixing model to design an automated music mixing system.

### Key words

automatic music mixing, music source separation, Wave-U-Net, acoustic AI, deep learning

---

\* 포항공과대학교 인공지능대학원, truestar2001@postech.ac.kr, youngjunsim@postech.ac.kr

\*\* 포항공과대학교 인공지능대학원, yjsuh@postech.ac.kr(교신저자)

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2019-II191906)과 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원(No.2022R1A6A1A03052954)과 2024년도 정부(과학기술정보통신부)의 재원으로 한국지능정보사회진흥원의 지원(2024년 지역특화형 초거대 AI 클라우드 팜 실증 및 확산 환경 조성)을 받아 수행된 연구이며, 「과학기술정보통신부 방송통신발전기금」을 지원받아 정보통신산업진흥원(D1014-24-1009, 철도차량 가용성 향상을 위한 AI 융합형 통합정비 공정 메타버스 플랫폼 개발)의 지원을 받아 수행된 연구이나 과학기술정보통신부의 공식의견과 다를 수 있음.



## 1. 서론

음악 믹싱은 여러 개별 악기 트랙을 적절히 조정하여 하나의 곡으로 완성하는 과정으로, 볼륨, 주파수, 공간감 등을 조절해 청각적으로 균형 잡힌 사운드를 구현하는 작업이다. 특히 볼륨 조절은 트랙 간 볼륨의 상대적 크기를 조정해 특정 악기를 강조하거나 최적의 밸런스를 맞추는 믹싱의 핵심 작업이다.

최근에는 딥러닝 기술의 발달로 자동 음악 믹싱이 활발히 연구되고 있으며, 특히 볼륨 조절 작업에서 Wave-U-Net[1] 기반 모델들의 가능성이 입증되었다. 그러나 믹싱된 음악에서 분리된 개별 악기 트랙 데이터의 부족과, 믹싱 되지 않은 개별 트랙 악기 데이터의 부재는 자동 음악 믹싱 모델의 학습을 능력을 저하시키고 성능의 제약을 야기한다.

본 연구에서는 음악 소스 분리 모델 Hybrid Transformer Demucs(HT Demucs)[2]을 활용하여 믹싱된 음악에서 분리된 개별 트랙을 추출하고, 각 분리된 트랙의 볼륨을 무작위로 조정하여 자동 음악 믹싱을 위한 데이터셋을 생성한다. 이를 통해 데이터 부족 문제를 극복하고, Wave-U-Net 구조를 활용하여 자동으로 볼륨 밸런스를 조절하는 음악 믹싱 시스템을 제안한다.

## II. 자동 음악 믹싱 시스템

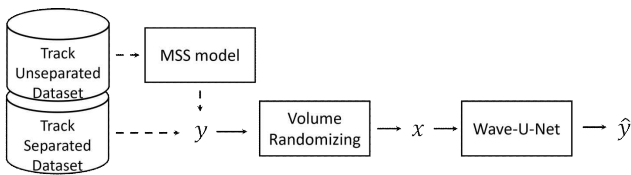


그림 1. 자동 음악 믹싱 시스템 모식도

Fig. 1. Automatic music mixing system architecture

제안된 자동 음악 믹싱 시스템의 전체 구조는 그림 1과 같다. HT Demucs와 무작위 볼륨 조절을 통해 딥러닝 기반 음악 믹싱 모델을 위한 훈련-타겟 데이터 쌍을 생성한다. 이후 Wave-U-Net 구조의 음악 믹싱 모델을 학습하여 믹싱 전의 개별 트랙들의 볼륨을 조정해 최적의 믹싱된 음악으로 변환한다.

### 2.1 타겟 데이터셋 생성

개별 악기 트랙 녹음이 존재하지 않는 무작위 음악 데이터를 HT Demucs에 통과시켜 분리된 개별 악기 트랙을 추출한다. HT Demucs는 Wave-U-Net 구조에서 가장 깊은 층을 트랜스포머 층으로 교체하고 트랙 혼합 녹음 입력을 높은 품질의 4개의 개별 악기 트랙 녹음(베이스, 드럼, 보컬, 그 외)으로 분리하여 출력하는 딥러닝 기반 음악 소스 분리 모델이다. 개별 악기 트랙이 없는 음악 데이터셋 GTZ AN[3]의 음악 300곡을 HT Demucs를 통해 개별 악기 트랙을 생성하고, 개별 악기 트랙이 있는 MUSDB18[4] 데이터셋의 음악 100곡과 병합하여 총 400곡의 타겟 데이터셋을 구축한다.

### 2.2 훈련 데이터셋 생성

$$scaling = \begin{cases} 1/(4-6x) & 0 \leq x < 0.5 \\ 1+6(x-0.5) & 0.5 \leq x < 1 \end{cases} \quad (1)$$

타겟 데이터 셋의 개별 악기 트랙의 볼륨을 인위적으로 조정하여 믹싱 되지 않은 훈련 데이터셋을 생성한다. 식 (1)은 균등분포에서 샘플링된  $x$ 를 통해 0.25~4 사이의 무작위 볼륨 조절 인자를 생성하는 식을 나타낸다. 0.25~4 사이의 무작위 값을 볼륨 조절 인자로 사용하면 1 이상의 인자가 추출될 확률이 더 높아지고, 결과적으로 볼륨을 낮출 확률보다 키울 확률이 높아지게 된다. 식 (1)을 통해 0.25~1과 1~4의 볼륨 조절 인자가 추출될 확률이 같아지도록 하였다. 각 트랙에 해당 조절 요소를 곱하여 볼륨이 조절된 새로운 악기 트랙을 생성한다. 이를 통해 자동 음악 믹싱 모델을 위한 훈련-타겟 데이터 쌍을 구축한다.

### 2.3 음악 믹싱

구축한 데이터셋을 이용해 믹싱 되지 않은 개별 악기 트랙을 믹싱된 개별 악기 트랙으로 변환하는 모델을 학습시킨다. 음악 믹싱 모델은 U-Net 구조를 음성 데이터에 적용될 수 있도록 변형한 Wave-U-Net을 사용한다.

일반적인 U-Net과 달리 1차원 합성곱 신경망(CNN)으로 이루어진 인코더를 사용하여 특징을 추출하고, 선형 보간법과 1차원 합성곱 신경망으로 이루어진 디코더를 사용하여 원본 녹음을 복원한다. 이 과정에서 U-Net과 마찬가지로 인코더와 디코더 사이에는 스킵 연결이 추가된다.

$$L_{total} = L_{spec} + \alpha L_{energy} \quad (2)$$

식 (2)는 모델의 손실 함수를 나타낸다.  $L_{spec}$ 은 스펙트로그램으로 변환된  $\hat{y}$ 과  $y$ 의  $L_1$ 거리이며,  $L_{energy}$ 는  $\hat{y}$ 과  $y$  각각의 제곱의  $L_1$ 거리이다. 손실 함수의 가중치인  $\alpha$ 는 1을 사용했다.

### III. 실험

표 1은 모델에 따른 볼륨 조절 성능을 평가한 것이다. 비교 모델은 Auto-Encoder를 사용하였다. Randomized는 무작위로 볼륨을 조절한 훈련 데이터셋이며, Wave-U-Net + MSS는 음악 소스 분리 모델을 이용하여 데이터셋을 증강한 후 Wave-U-Net으로 자동 믹싱하는 시스템이다. 각 모델은 4의 배치 크기로 50000 스텝 동안 훈련되었다. 검증 데이터로는 MUSDB18 데이터셋 중 훈련 때 사용되지 않은 50 곡을 사용하였다. 평가 지표로는 평균 음량을 나타내는 LUFS(Loudness Units Full Scale)의 차이를 실제값과 예측값 사이의 오차를 백분율로 나타내주는 지표인 MAPE(Mean Absolute Percentage Error)로 계산한 값을 사용하였다. 또한, 원본 신호와 복원된 신호 간의 왜곡 수준을 나타내는 SDR(Signal-to-Distortion Ratio)을 사용하였다.

Wave-U-Net 기반 모델은 Auto-Encoder에 비하여

표 1. 모델 별 LUFS, SDR 점수

Table 1. LUFS, SDR for each model

모델	베이스		드럼		보컬		그 외	
	LUFS ↓	SDR ↑	LUFS ↓	SDR ↑	LUFS ↓	SDR ↑	LUFS ↓	SDR ↑
Radomized	0.265	3.810	0.305	2.076	0.334	1.203	0.261	4.162
Auto-Encoder[5]	0.115	-6.647	<b>0.097</b>	-5.328	0.081	<b>-5.561</b>	<b>0.073</b>	-6.048
Wave-U-Net[1]	0.115	<u>-5.900</u>	0.137	<u>-4.789</u>	<b>0.072</b>	<u>-6.114</u>	0.078	<b>8.112</b>
Wave-U-Net + MSS	<b>0.106</b>	<b>7.480</b>	<u>0.110</u>	<b>6.135</b>	<u>0.076</u>	-6.489	0.080	<u>8.016</u>

모든 악기 트랙에서 전반적으로 높은 점수를 기록하였으며, 음원 소스 분리 모델을 추가하였을 때 베이스와 드럼 트랙에서 큰 폭의 성능 향상을 보였다.

표 2는 손실 함수 설정에 따른 볼륨 조절 성능을 LUFS로 평가한 것이다.  $L_{time}$ 은  $\hat{y}$ 과  $y$ 를 스펙트로그램으로 변환하지 않고  $L_1$ 거리를 구한 값이고,  $L_{spec}$ 은  $\hat{y}$ 과  $y$ 를 스펙트로그램으로 변환 후  $L_1$ 거리를 구한 값이다. 실험 결과,  $L_{time}$ 과  $L_{energy}$ 를 함께 사용하는 것이 ‘베이스’에서,  $L_{spec}$ 만 사용하는 것이 ‘드럼’과 ‘보컬’에서,  $L_{spec}$ 과  $L_{energy}$ 를 함께 사용하는 것이 ‘그 외’에서 가장 좋은 성능을 나타냈다.

표 2. 손실 함수 설정 별 LUFS 점수

Table 2. LUFS for each loss function setting

손실 함수	베이스	드럼	보컬	그 외
$L_{time}$	0.116	0.133	0.066	0.170
$L_{spec}$	0.110	<b>0.116</b>	<b>0.063</b>	0.081
$L_{time} + L_{energy}$	<b>0.108</b>	0.175	0.066	0.079
$L_{spec} + L_{energy}$	0.115	0.137	0.072	<b>0.078</b>

### IV. 결론

본 논문에서는 음악 소스 분리 모델과 Wave-U-Net을 이용하여 자동 음악 믹싱 모델을 제안하였다. 제안된 방법은 개별 악기 트랙이 없는 녹음을 음악 소스 분리 모델을 활용해 개별 악기 트랙 녹음을 생성하여 데이터셋 부족 문제를 해결할 수 있음을 확인한다. 또한 무작위 볼륨 조절을 통해 훈련 데이터셋을 생성하고 이를 이용해 Wave-U-Net 기반 자동 음악 믹싱 시스템을 설계할 수 있음을 보여준다.

참 고 문 헌

- [1] Daniel Stoller, Sebastian Ewert, and Simon Dixon. Wave-u-net: A multi-scale neural network for end-to-end audio source separation, 2018.
- [2] Simon Rouard, Francisco Massa, and Alexandre D'efossez. Hybrid transformers for music source separation. In ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1-5, 2023
- [3] Tzanetakis, G., Cook, P.: Musical genre classification of audio signals. IEEE Trans. Speech Audio Process. 10(5), 293-302 (2002)
- [4] Z. Rafii, A. Liutkus, F.-R. Stöter, S. I. Mimitakis, and R. Bittner, "The MUSDB18 corpus for music separation," Dec. 2017. [Online]. Available: <https://doi.org/10.5281/zenodo.1117372>
- [5] G. E. Hinton, R. R. Salakhutdinov ,Reducing the Dimensionality of Data with Neural Networks. Science 313, 504-507 (2006). DOI: 10. 1126/ science. 1127647

# Attention Module 기반의 CAM 확장 방법을 이용한 약 지도 영상 분할

주선기\*, 유종빈\*\*

## Weakly supervised image segmentation using CAM extension method based on Attention Module

Sunki Joo\*, and Jongbin Ryu\*\*

### 요 약

Weakly supervised semantic segmentation은 픽셀 단위 라벨 데이터를 확보하기 어려운 상황에서 적은 양의 라벨 데이터를 활용해 픽셀 단위 라벨 데이터를 확보하는 연구다. Weakly supervised semantic segmentation에서 주로 사용되는 CAM은 모델이 이미지를 분석할 때 집중하는 영역을 시각화하며, 이를 픽셀 단위 라벨 데이터로 사용한다. 그러나 CAM은 전경과 객체 구분 문제 및 지역적 한계를 가진다. 이를 개선하기 위해 본 연구는 CAM의 영역을 확장하는 attention module을 제안한다.

### Abstract

Weakly supervised semantic segmentation is a study that utilizes a small amount of labeled data to obtain pixel-level label data in situations where acquiring pixel-level labels is difficult. In weakly supervised semantic segmentation, the commonly used CAM visualizes the regions the model focuses on when analyzing an image, and this is used as pixel-level label data. However, CAM has limitations such as difficulty in distinguishing foreground objects and issues with localized attention. To solve this, we propose an attention module that expands the regions identified by CAM.

### Key words

Weakly supervised semantic segmentation, Attention module, Class activation map

### 1. 서 론

컴퓨터 비전 분야에서 인공지능의 연구는 다양한 task에서 진행되어 왔다. 대표적으로 classification, object detection, segmentation이 있다. 그리고 이러한 연구들은 실생활에도 빠르게 적용되고 있다.

하지만 성공적인 연구 결과들을 그대로 실생활에 적용하는 데에는 많은 제약이 존재 한다. 그중 하나가 segmentation task의 데이터 확보 문제다. segmentation task는 모델을 학습하기 위하여 이미지의 픽셀 단위 라벨 데이터가 필요하다. 픽셀 단위 라벨 데이터의 경우 라벨링 과정에서 많은 시

---

\*아주대학교 인공지능학과, dollyin@ajou.ac.kr, \*\*아주대학교 인공지능학과, jongbinryu@ajou.ac.kr (교신저자)

간과 노동력이 필요하다. 또한 의료 분야와 같은 특정 도메인에서는 높은 수준의 전문성 또한 요구된다.

그래서 등장한 연구 분야가 Weakly supervised semantic segmentation(WSSS)다. WSSS는 더 적은 양의 정보를 가지고 있는 라벨 데이터를 활용하여 픽셀 단위 라벨 데이터를 생성하여 segmentation 모델 학습에 사용하는 방법을 연구한다. WSSS 연구에서 사용하는 대표적인 방법 중 하나가 class activation map(CAM) [1]을 사용하는 것이다. 이미지 단위 라벨 데이터가 존재 할 때, 이미지 단위 라벨링이 되어 있는 데이터로 학습된 classification 모델을 사용하여 CAM을 그려낼 수 있다. CAM은 인공지능 모델이 이미지 추론을 하였을 때, 추론 결과를 도출하기 위하여 이미지의 어느 영역에 집중을 하고 있는지 보여주는 데이터다. 그리고 WSSS 연구에서는 CAM이 집중하고 있는 영역을 픽셀 단위 라벨 데이터로 사용한다.

이러한 CAM을 사용한 픽셀 단위 라벨 데이터 생성은 몇가지 도전 과제에 직면해 있다. CAM이 완벽하지 않기 때문이다. CAM은 객체의 일부만을 집중을 하거나 전경과 객체를 구분하는데 많은 어려움을 겪고 있다. 그렇기 때문에 WSSS 연구는 이러한 CAM의 불완전성을 극복하는데 초점이 맞춰져 있다.

본 연구에서는 CAM이 더 넓은 영역을 표현하고 중요하지 않은 영역을 배제 하도록 하는 attention

module을 제안한다. attention module은 refine CAM을 만들어내고, refine CAM을 original CAM과 결합해 final CAM을 만들어내는 방식이다.

본 연구에서는 attention module의 성능을 검증하기 위하여 WSSS 분야에서 검증된 모델들인 DSRG [2], PuzzleCAM [3] 모델에 attention module을 결합하여 CAM IoU의 변화를 관찰 하였다. 연구에 사용된 데이터셋은 PascalVOC2012 [4]이다. PascalVOC2012 데이터셋은 segmentation task를 위하여 만들어진 benchmark 데이터셋 중 하나로 WSSS 연구에서도 주로 사용되고 있는 데이터 셋이다.

DSRG와 attention module을 결합한 실험에서 segmentation 결과의 mIoU가 57.6에서 59.2로 향상되었다. PuzzleCAM과 attention module을 결합한 실험에서는 CAM mIoU가 55.4에서 56.3으로 향상 됐다.

## II. 방법론

본 연구에서 제안한 attention module을 적용할 때, attention module을 적용한 이후에 CAM을 구성하고 있는 각각의 class 채널의 값을 조정하는 class adjustment 과정을 거친다. 이 항목에서는 attention module과 class adjustment을 결합한 방법론과 실험에 대해 서술 한다.

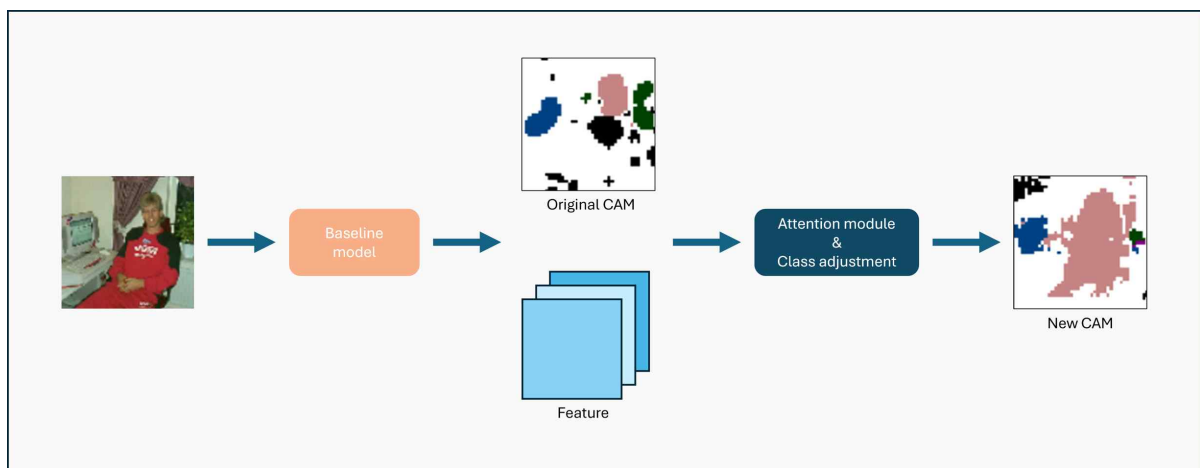


그림 1. Attention module 적용 구조.  
Fig. 1. Attention module application structure.

### 2-1. Attention module, class adjustment

본 연구에서 제안한 attention module의 적용은 두가지 주요 컴포넌트인 original CAM과 refine CAM이 결합하는 과정으로 진행된다. 먼저 refine CAM을 생성하기 위하여 feature(f)를 Affinity Matrix(유사도 행렬) 계산 (1) 및 정규화 (2)를 통해 affinity 행렬 A를 구한다. 그리고 A를 original CAM과 곱하여 refine CAM을 생성한다. (3) A를 구하는 방식인 Affinity Matrix 계산 및 정규화가 self-attention 메커니즘의 핵심 아이디어와 유사하기 때문에 attention module이라고 명칭 하였다.

$$A = ReLU(f_{norm}^T \cdot f_{norm}) \quad (1)$$

$$A_{norm} = \frac{A}{\sum_j A_{ij} + \epsilon} \quad (2)$$

$$CAM_{refine} = CAM_{original} \cdot A_{norm} \quad (3)$$

위와 같은 과정을 거쳐 생성한 refine CAM은 class adjustment 과정을 거쳐 조정된다. class adjustment 과정에서는 refine CAM의 class 정보를 담고 있는 각 채널의 값들을 해당 채널의 최대 값이 0이상일 때 최대값으로 나누어 준다. 이렇게 조정된 refine CAM을 다시한번 original CAM과 더 해주어 final CAM을 생성 한다.

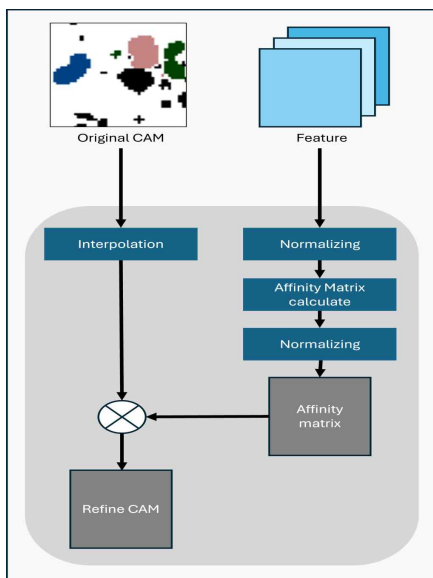


그림 2. Attention module 구조  
Fig. 2. Attention module structure

위와 같은 attention module 적용 과정을 거쳐 final CAM은 original CAM에 비해 더 넓은 영역의 중요 정보에 집중할 수 있게 되고 중요도가 떨어지는 정보를 배제할 수 있게 된다.

### III. 실험 및 결과

본 논문에서는 PascalVOC2012 데이터셋을 사용했다. PascalVOC2012 데이터셋은 컴퓨터 비전 연구에서 널리 사용되는 이미지 분류, 객체 탐지, 및 세그멘테이션을 위한 데이터셋이다. 이 데이터셋은 다양한 task에서 다루며, 특히 Object Detection과 Segmentation에서 중요한 벤치마크 데이터셋으로 활용되고 있다. 20개의 클래스로 구성되어 있으며, train set 5717개, valid set 5823개로 총 11,540개의 이미지로 구성되어 있다.

실험에 사용된 base model은 DSRG와 PuzzleCAM이다. DSRG는 seed cam을 사용한 초기 라벨 데이터로 segmentation 학습을 하며 seed cam을 지속적으로 보강하는 방식을 사용한 모델로서 WSSS 연구의 초석을 마련한 모델이다. PuzzleCAM은 이미지를 4등분하여 각각의 이미지에서 서로다른 feature를 생성해서 CAM을 구하는 방식을 사용한 모델이다.

실험의 결과는 표 1과 같이 나왔다. 평가 지표로 mIoU를 사용하였고, DSRG와 PuzzleCAM 두 개의 모델에서 성능 향상을 관찰하였다.

표 1. PascalVOC2012 데이터셋을 사용한 DSRG와 PuzzleCAM의 실험 결과.

Table 1. Experimental Results of DSRG and PuzzleCAM Using the PascalVOC2012 Dataset.

	baseline mIoU	baseline+ours mIoU
DSRG [2]	57.6	59.2
PuzzleCAM [3]	55.4	56.3

### III. 결 론

본 연구는 CAM을 사용한 WSSS 연구에서 CAM의 확장을 위하여 attention module을 제안 하였다.

실험을 통하여 기존 검증된 모델에 attention module의 적용 하였을 때 성능이 향상 되는 것을 확인 할 수 있었다. 학습된 모델이 생성한 CAM에 attention module을 사용하여 중요한 픽셀에 가중치를 부여하는 방식의 효과를 입증한 것이다.

본 연구를 토대로 attention module과 그 적용을 정교하게 발전시켜 나가고, 다른 방법론들과 결합하여 성능 향상을 이끌어 내는 것이 앞으로의 연구 과제다.

## 참 고 문 헌

- [1] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba, "Learning deep features for discriminative localization," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2921-2929, 2016.
- [2] Zilong Huang, Xinggang Wang, Jiasi Wang, Wenyu Liu, Jingdong Wang, "Weakly-supervised semantic segmentation network with deep seeded region growing," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7014-7023,
- [3] Sanghyun Jo, In-Jae Yu, "Puzzle-cam: Improved localization via matching partial and full features," 2021 IEEE International Conference on Image Processing (ICIP), 639-643, 2021.
- [4] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, A. Zisserman, "The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results," <http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html>.

# 화자 얼굴 사진을 이용한 음성 합성 시스템

정우열\*, 서영주\*\*

## Speech Synthesis System Based on Speaker's Face Images

Wooyeol Jeong\* and Young-Joo Suh\*\*

### 요약

대부분의 Text-To-Speech (TTS), Voice Conversion (VC) 시스템은 특정 화자의 녹음된 음성을 바탕으로 발화를 합성한다. 학습 시에 보지 못한 화자의 음성을 합성하려면, 일반적으로 고품질의 녹음된 발화가 필요하다. 본 논문에서는 저화질의 얼굴 이미지를 이용해 화자의 목소리를 추정하고, 이를 바탕으로 자연스러운 음성을 합성하는 시스템을 제안한다. 실험을 통해, 제안된 시스템은 원본 화자의 목소리처럼 자연스럽게 얼굴과 목소리 사이 높은 일관성을 보이는 음성을 합성할 수 있음을 확인하였다.

### Abstract

Most text-to-speech (TTS) and voice conversion (VC) systems synthesize utterance based on the voice of a specific speaker. To synthesize the voice of an unseen speaker during training, high-quality recordings are generally required. This paper proposes a system that estimates a speaker's voice using low-resolution facial images and synthesizes natural-sounding speech based on this estimation. Experimental results demonstrate that the proposed system can synthesize speech with a naturalness close to that of the original speaker's voice, while maintaining high consistency between the face and the voice.

### Key words

text-to-speech, voice conversion, speech synthesis, face-to-speech, artificial neural network

## 1. 서론

Text-to-Speech (TTS)는 텍스트와 화자의 목소리를 입력으로 받아, 해당 화자의 목소리로 텍스트를 발

\* 포항공과대학교 인공지능대학원, jungwy0106@postech.ac.kr

\*\* 포항공과대학교 인공지능대학원, yjsuh@postech.ac.kr (교신저자)

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2019-II191906)과 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원(No.2022R1A6A1A03052954)과 2024년도 정부(과학기술정보통신부)의 재원으로 한국지능정보사회진흥원의 지원(2024년 지역특화형 초거대 AI 클라우드 팜 실증 및 확산 환경 조성)을 받아 수행된 연구이며, 「과학기술정보통신부 방송통신발전기금」을 지원받아 정보통신산업진흥원(D1014-24-1009, 철도차량 가용성 향상을 위한 AI 융합형 통합정비 공정 메타버스 플랫폼 개발)의 지원을 받아 수행된 연구이나 과학기술정보통신부의 공식의견과 다를 수 있음.



화하는 음성을 합성하는 분야이다. 음성 변환(Voice Conversion, VC)은 입력된 발화의 내용을 유지하면서, 이를 원하는 화자의 목소리로 변환하는 분야이다. 딥러닝의 발전에 따라 TTS와 음성 변환 분야는 빠르게 성장하고 있다. YourTTS[3]의 경우, multi-lingual zero-shot multi-speaker TTS와 음성 변환을 수행할 수 있는 모델로, VITS[5]를 기반으로 하고 있다. YourTTS는 VCTK dataset에서 zero-shot multi speaker TTS 분야의 State-of-the-Art (SOTA)를 달성했으며, zero-shot voice conversion 분야에서도 SOTA에 근접한 성능을 보였다.

생물학 및 신경학 분야의 연구들에 따르면, 사람의 얼굴은 그 사람의 목소리와 연관되어 있다는 것이 증명되었다[1,2]. 이를 통해 사람들은 얼굴을 보고 성별, 나이, 일부 지역적 특색과 같은 정보를 얻어 음성 특성과 특징을 유추할 수 있다.

Vision-Transformer (ViT)[4]는 자연어 처리 분야에서 주로 사용되던 Transformer 구조를 컴퓨터 비전 분야에 적용한 연구로, 기존 연구보다 더 적은 연산량으로 높은 성능을 보이며 많은 연구들에서 활용되고 있다.

본 논문에서는 화자의 얼굴을 바탕으로 목소리의 특성과 특징을 추정하여 화자 임베딩(speaker embedding)으로 사용하는 음성 합성 시스템을 제안한다. 현재 대부분의 음성 합성 연구들은 새로운 화자의 목소리로 발화를 합성하기 위해 고품질의 녹음된 목소리를 필요로 하지만, 제안된 시스템은 고품질의 녹음된 목소리가 없는 경우에도 저화질의 얼굴 사진만으로 자연스러운 음성을 합성할 수 있도록 한다.

## II. 얼굴 사진을 활용한 음성 합성 시스템

본 논문에서 제안하는 시스템의 전반적인 구조는 그림 1과 같다. 얼굴 사진을 활용한 음성 합성을 위해, 사전 학습(pre-trained)된 YourTTS[3]와 ViT[4]를 사용한다. 화자의 얼굴 이미지를 ViT에 입력하여 얼굴의 특성을 나타내는 얼굴 임베딩(facial embedding)을 추출한 뒤, 이를 화자의 음성 특징을 나타내는 화자 임베딩으로 변환하는 mapping model

에 넣어 화자 임베딩을 예측한다. ViT의 CLS 토큰을 얼굴 임베딩으로 사용하며, mapping model은 768차원의 얼굴 임베딩을 512차원의 화자 임베딩으로 변환한다. Mapping model은 fully-connected layer, ReLU activation function, Batch normalization으로 구성되며, L2 normalization을 적용하여 화자 임베딩을 예측한다.

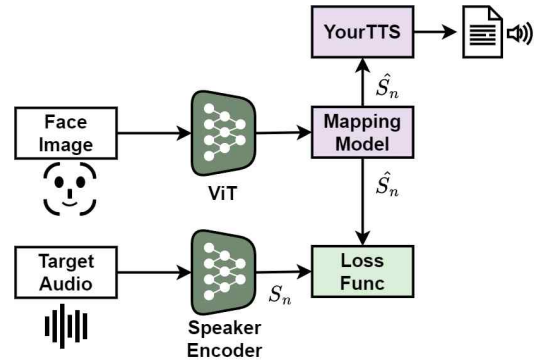


그림 1. 모델 구조  
Fig. 1. Model architecture

화자의 목소리를 기반으로 사전 학습된 화자 인코더(speaker encoder)[6]를 사용해 ground-truth 화자 임베딩을 얻고, 예측된 화자 임베딩과 ground-truth 화자 임베딩 사이의 코사인 유사도와 평균 제곱 오차(mean square error)를 최소화하도록 mapping model을 학습한다. 이때 손실 함수는 다음과 같이 정의된다.

$$L = \frac{1}{N} \sum_{n=1}^N \alpha \cos(\hat{S}_n, S_n) + \beta mse(\hat{S}_n, S_n) \quad (1)$$

위 식에서  $\hat{S}_n$ 은 얼굴 임베딩을 이용해 예측한 화자 임베딩을 의미하고,  $S_n$ 은 사전 학습된 화자 인코더로 얻은 ground-truth 화자 임베딩을 의미한다.  $\cos()$ 는 코사인 유사도를 나타내며,  $mse()$ 는 평균 제곱 오차를 의미한다.  $\alpha$ 와  $\beta$ 는 hyper-parameter로, 값의 크기를 고려하여 코사인 유사도를 강조하도록 설정하였다.

화자의 얼굴 사진을 통해 예측한 화자 임베딩을 사용하여 음성을 합성한 결과를 평가하기 위해 15명의 평가자와 함께 정성평가를 진행했다. 얼굴과

목소리의 일치 정도를 측정하기 위해 1-5 scale mean opinion scores in terms of face-voice consistent degree (MOS-FVC)을 실시하고 자연스러움을 측정하기 위해 1-5 scale naturalness mean opinion score (NMOS)를 실시했으며, 결과는 표 1과 같다.

‘Ground-truth’는 화자의 원본 음성을 의미하고, ‘Using  $\hat{S}_n$ ’은 화자의 얼굴 사진을 통해 예측한 화자 임베딩을 사용하여 합성한 음성을 의미하며, ‘Using  $S_n$ ’은 화자의 목소리를 통해 얻은 화자 임베딩을 활용하여 합성한 음성을 의미한다. 실험을 통해 저화질의 얼굴 사진으로 자연스러운 음성을 합성할 수 있음을 확인하였다.

표 1. 합성된 음성 및 원본 음성에 대한 1-5 scale Mean Opinion Score (MOS)

Table 1. 1-5 scale Mean Opinion Score (MOS) of synthesized speech and original speech

	Ground-truth	Using $\hat{S}_n$	Using $S_n$
MOS-FVC (1-5 scale)	4.400	3.467	2.233
NMOS (1-5 scale)	4.767	4.133	3.267

### III. 결 론

본 논문에서는 얼굴 사진을 활용해 음성을 합성하는 인공지능 기반 시스템을 제안한다. 제안된 시스템은 학습 시에 보지 못한 새로운 화자의 저화질 얼굴 사진을 통해서도 목소리 특성을 예측하여 자연스러운 음성을 합성할 수 있음을 보여준다.

### 참 고 문 헌

[1] Lauren Mavica and Elan Barenholtz, “Matching voice and face identity from static images”, Journal of experimental psychology. Human perception and performance, Vol. 39, No. 2, pp. 307-312, 12, 2013.

[2] Smith, H. M. J., Dunn, A. K., Baguley, T., and Stacey, P. C., “Concordant Cues in Face s and Voices: Testing the Backup Signal Hypothesis”,

Evolutionary Psychology, Vol. 14, 02, 2016

[3] Edresson Casanova, Julian Weber, Christopher Shulby, Arnaldo Candido Junior, Eren Gölge and Moacir Antonelli Ponti, “YourTTS: Towards Zero-Shot Mutli-Speaker TTS and Zero-Shot Voice Conversion for everyone“, Proceedings of the 39<sup>th</sup> International Conference on Machine Learning, PMLR 162:2709-2720, 2022.

[4] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterhiner, Mostafa Dehghani, Matthias Minderer Georg Heigold, Sylvain Gelly, Jakob Uszkoreit and Neil Houlsby, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale“, International Conference on Learning Representation, 2021.

[5] Jaehyeon Kim, Hungil Kong and Juhee Son, “Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech”, Proceedings of the 38<sup>th</sup> International Conference on Machine Learning, PMLR 139, 2021

[6] Hee Soo Heo, Bong-Jin Lee, Haesung Huh and Joon Son Chung, “Clova Baseline System for the VoxCeleb Speaker Recognition Challenge 2020”, arXiv preprint arXiv:2009.14153, 2020.

# 신뢰성모델 기반 대학생 중도탈락 분석 프로세스 개발

장성호\*, 유승훈\*\*, 강지훈\*\*\*

## Development of a Reliability-Based Model for Analyzing University Student Dropout Rates

Seongho Jang\*, Seunghun You\*\*, and Jihoon Kang\*\*\*

본 논문은 대학생의 중도탈락 가능성을 시간 경과에 따라 모니터링할 수 있는 신뢰성 모델 기반 분석 프로세스를 개발하는 것을 목적으로 한다. 기존의 대학생 중도탈락 연구는 학습 참여, 교육 만족도, 진로 역량 등의 변수를 활용한 상관 분석 및 다중 회귀 분석[1], 수업 외 활동이 학업 중단에 미치는 영향을 구조방정식모형을 통해 분석하는 연구[2] 등 다양한 관점에서 중도탈락 요인을 다루고 있다. 또한, AI 발전에 따라 중도탈락과 졸업률 이진 분류로 예측하는 연구도 활발히 진행 중이다[3]. 그러나 이러한 연구들은 시간 흐름에 따른 중도탈락의 영향을 분석하지 못하며, 대부분 일시적 스냅샷에 기반한 분석이다. 이에 따라 본 논문에서는 2013년부터 2024년까지 입학한 약 4,500명의 학생 데이터를 활용하여 신뢰성 모델 기반의 Cox 비례위험 모형을 적용한 시간에 따른 중도탈락 분석을 수행하였다. 독립 변수로는 입학 정보, 대학 성적, 장학금 수혜 여부, 비교과 활동, 대학생 활 적용 검사 결과 등을, 종속 변수로는 최종 학적 변동(졸업, 중도탈락, 재학)을 설정하였다. 모델링은 학과, 입학 전형, 학점 그룹(상위, 하위) 등의 세부 요인에 따라 개별적으로 수행하였다. 분석 결과, 대학생활 적응 검사의 다양한 지표들이 모델에 유의미한 영향을 미쳤으며, 특히 학자금 지원, 교수와의 관계 만족도, 진로 준비 필요성 인식 등이 주요 예측 요인으로 나타났다. 본 연구는 시간 경과에 따른 중도탈락 가능성의 변화를 심도 있게 분석할 수 있는 방법론을 제시함으로써, 대학생 중도탈락 방지 및 지원 프로그램 개발에 기여할 수 있을 것으로 기대된다.

### Key words

big data, statistics model, college dropout, college students

### 참고문헌

- [1] S. Joo and Y. J. Jeong, "The Influence of College Freshmen's Institutional Commitment and Student Engagement on Satisfaction with General Education," *Journal of Education & Culture (JOEC)*, vol. 28, no. 5, pp. 301-322, 2022.
- [2] C. Y. Back and B.S. Lee, "Examining the Relationship between College Students' Out-of-Class Activities and Institutional Departure: Focusing on the Mediating Effect of Institutional Commitment," *Journal of Higher Educational Research*, vol. 5, no. 1, pp. 1-25, 2022.
- [3] Matti, "Predicting student dropouts with machine learning: An empirical study in Finnish higher education," *Technology in Society*, vol. 76, 2024.

\* 한국공학대학교 스마트팩토리융합학과, jang0212@tukorea.ac.kr,

\*\* 한국공학대학교 IT경영학과, tmdgns3480@tukorea.ac.kr,

\*\*\* 한국공학대학교 경영학부, jhkang82@tukorea.ac.kr (교신저자표시)



# 반주시과제 눈 움직임을 활용한 노인 우울증 평가

김가혜\*, 김종일\*

## Assessment of Late-Life Depression Using Eye Movements in Anti-Saccade Tasks

Kahye Kim\* and Joong Il Kim\*

노인 우울증(Late-life depression, LLD)은 60세 이상의 사람들에게 발생하는 가장 흔한 정신 질환 중 하나이다. 실행 기능의 결핍은 우울한 노인의 주요 신경인지 결손 중 하나로 강조되었으며, 실행 기능의 감소는 개인의 독립적 삶을 방해하여 우울 증상을 지속시키고 동시에 자살과 같은 더 나쁜 결과를 초래할 위험성이 있다 [1]. 실행 기능은 주요 반응 억제, 작업 기억 정보모니터링, 작업 간 전환과 관련된 능력으로 개념화할 수 있으며 주로 신경 심리학적 테스트를 이용하여 측정되고 있으나 오랜 측정 시간으로 인한 참가자의 심리적 고통이 호소되며 운동 및 언어 능력에 따른 결과 차이가 발생할 수 있다. 따라서 실행 기능을 평가하는 다양한 신경생리학적 지표가 제안되고 있으며 눈 움직임은 비침습적으로 뇌의 기능을 간접적으로 측정할 수 있는 통찰력을 제공한다. 눈 움직임 평가는 특정 과제를 통해서 진행할 수 있으며, 그 중 정/반주시과제(pro/anti saccade task)는 시야에서 주변에 나타나는 표적 방향으로 눈의 자동적인 방향을 주시/억제하는 능력을 탐색하여 실행기능의 주요 구성 요소인 인지 억제를 조사하는 눈 추적 패러다임이다. 특히 반주시과제의 성공적인 수행을 위해서는 고차원적인 인지 제어를 필요로 하므로 고위험 우울 환자를 식별하는 도구로 활용가능성이 제안되었다[2]. 본 연구에서는 정/반주시과제를 수행하는 동안의 눈의 움직임을 기록하여 LLD 위험성을 가진 노인의 특징을 확인하고, 눈의 움직임이 LLD와 건강한 노인의 특성을 확인하는데 유용한지를 탐색하였다. 지역 코호트를 통해 수집된 60세 이상 인지적으로 건강한 노인 225명을 대상으로 우울평가척도를 통해 건강대조군(n=188)과 LLD 위험군(n=37)으로 분류한 후 신경심리검사를 평가한 결과, 주의집중력, 언어기능, 시공간기능, 기억력 점수에서는 군 간 유의한 차이를 보이지 않았으나 LLD 위험군의 전두엽 집행기능의 점수는 대조군 대비 유의하게 낮았다. 비디오 시선 추적 기술을 사용하여 정/반주시과제를 수행하는 동안의 눈 움직임을 조사한 결과에서는 정주시과제에서 두 그룹 간 차이가 나타나지 않았으나 반주시과제 시 LLD 위험군에서 낮은 정답율 및 오류율 증가를 보였으며 수정된 억제 오류의 잠복기 및 반응시간에서 건강한 대비 유의한 차이가 확인되었다. 이러한 눈 움직임은 LLD에서 나타날 수 있는 실행 기능 구성 요소의 낮은 효율성을 평가하는데 활용될 수 있음을 나타내며, 인지 억제와 관련된 눈 움직임에 대한 평가가 우울한 노인 환자를 식별하는 첫 번째 단계가 될 수 있음을 시사한다.

### Key words

late-life depression, anti saccadic eye movements, executive function

### 참고문헌

- [1] Zhangying Wu et al., Longitudinal Association Between Cognition and Depression in Patients With Late-Life Depression: A Cross-Lagged Design Study, *Front Psychiatry*, 2021 Oct 22;12:577058.
- [2] Negin Chehrehnegar et al., Executive function deficits in mild cognitive impairment: evidence from saccade tasks, *Aging Ment Health*, 2022 May;26(5):1001-1009

\*한국한의학연구원 디지털임상연구부, kkh2@kiom.re.kr(교신저자)

※ 본 연구는 한국한의학연구원 기본사업 과제의 지원을 받아 수행된 연구임(No. KSN2312022).

# 암 치료에 활용하기 위한 나노입자 광열 특성 분석

김 정 대\*

## Nanoparticles photothermal properties analysis for cancer therapy

Jung-Dae Kim\*

나노입자는 다양한 고유의 특성을 보유하므로, 여러 과학 분야에서 중요한 역할을 한다[1,2]. 일반적으로 바이오 응용 분야에서 표적 약물 전달(Targeted drug delivery)[3], 바이오센싱(Biosensing)[4], 광열 암 진단 및 치료(Photothermal cancer diagnosis and therapy)[5] 연구에 그 활용도가 높다. 특히, 광열 치료는 크기와 모양에 따라 표면 플라즈몬 공명(Surface plasmon resonance; SPR) 조절이 가능한 나노입자를 활용하여, 특정 파장의 빛을 강하게 흡수한 후 국부적인 열로 변환하여 선택적으로 암세포를 괴사시킨다. 이러한 나노입자 중에서 비구형 구조인 금 나노로드를 활용하는 세 가지 장점이 있다. 첫째, 종횡비에 따라 SPR 파장의 조절이 가능하여 인체 내 깊게 위치한 암세포를 목표로 할 수 있는 근적외선 영역대의 빛을 사용할 수 있다. 둘째, 일반적인 구형 나노입자보다 높은 종횡비를 구성할 수 있어 높은 광열 변환 효율을 가진다. 마지막으로 표면을 다양하게 기능화시킬 수 있어 표적에 알맞은 결합체를 합성하여 효율적인 국부 치료가 가능하다. 낮은 강도의 빛으로 고효율을 내기 위해서는 개별 금 나노로드의 광열 특성 분석이 중요하지만, 현재까지 상용화된 장비는 미비하다. 본 논문에서는 일반적인 현미경에 몇 가지 기능만 추가하여 개별 금 나노입자의 광열 특성을 분석하는 기술에 대해 제시한다.

### Key words

photothermal cancer therapy, nanoparticles, photothermal properties, surface plasmon resonance

### 참고문헌

- [1] J. Gao, H. Gu and B. Xu, "Multifunctional magnetic nanoparticles: design, synthesis, and biomedical applications", *Accounts Chem. Res.*, vol. 42, pp. 1097-1107, August 2009.
- [2] C. Burda, X. Chen, R. Narayanan, and M. A. El-Sayed, "Chemistry and properties of nanocrystals of different shapes", *Chem. Rev.*, vol. 105, pp. 1025-1102, March 2005.
- [3] F. Zobi, "Diatom biosilica in targeted drug delivery and biosensing applications: Recent Studies", *Micro*, vol. 2, pp. 342-360, June 2022.
- [4] I. Khan, K. Saeed and I. Khan, "Nanoparticles: properties, applications and toxicities", *Arab. J. Chem.*, vol. 12, pp. 908-931, November 2019.
- [5] L. Zhao, X. Zhang, X. Wang, X. Guan, W. Zhang and J. Ma, "Recent advances in selective photothermal therapy of tumor", *J. Nanobiotechnol.*, vol. 19, pp. 335, October 2021.

---

\* 한국한의학연구원 디지털임상연구부, kjd@kiom.re.kr

※ 본 연구는 2024년도 한국한의학연구원 “한의이론 기반 스마트 건강노화 관리기술 개발 (KSN231202)”과제의 지원을 받아 수행된 연구 결과임

# 헬스케어의 진보: 전신 경락 자극을 위한 침대형 음파 장치 개발

김 정 대\*

## Advancements in healthcare: development of a bed-Type sound wave device for whole-body meridian stimulation

Jung-Dae Kim\*

현대사회에서 경락 자극 치료법은 인체 내 에너지 흐름을 조절하여 건강을 증진하는 중요한 역할을 담당하고 있다[1]. 그러나 기존 방식은 시술에 전문가가 필요하고, 통증 및 화상 같은 부작용 위험이 있어, 보다 편리하고 안전한 경락 자극 기법 개발이 요구된다[2]. 본 연구는 비침습적인 음파 자극을 이용한 침대형 전신 자극 장치를 개발하여 경락을 효과적으로 자극하는 새로운 기술을 제안한다. 음파 진동은 기존의 모터 방식에 비해 소음, 내구성 문제를 개선하며, 진동수(Hz)와 세기(Intensity)를 정밀하게 조절할 수 있어 사용자의 신체 특성에 맞춘 맞춤형 자극이 가능하다[3]. 전신 진동 운동이 근육 기능 강화, 유연성 향상, 골밀도 증가, 혈류량 증가 등 다양한 신체기능 개선에 효과적인 것을 입증된 바와 같이[4], 음파 자극 장치는 심층 조직을 통과하여 전신의 경락과 신경계를 자극할 수 있어 더욱 광범위한 치료 효과가 기대된다. 특히, 경락을 통해 자극된 신경계 및 연조직이 체내 에너지 대사와 혈류 개선에 긍정적인 영향을 미치며, 이를 통해 전통적 한의학 치료법과 현대 음파 기술을 융합한 혁신적 치료 방안을 제시할 수 있다[5]. 본 연구는 안전성과 효율성을 동시에 갖춘 새로운 경락 자극 장치 개발을 목표로 하며, 미래의 차세대 헬스케어 기술로서 신체 기능 강화와 건강 증진에 기여할 것으로 기대된다.

### Key words

sound wave stimulation, meridian stimulation, healthcare, non-invasive stimulation

### 참고문헌

- [1] S. Chang, "The meridian system and mechanism of acupuncture-a comparative review. Part 1: the meridian system", Taiwan. J. Obstet. Gynecol., vol. 51, pp. 506-514, December 2012.
- [2] C. D. Myers, B. A. White and M. W. Heft, "A review of complementary and alternative medicine use for treating chronic facial pain", J. Am. Dent. Assoc., vol. 133, pp. 1189-1196, September 2002.
- [3] S. Cerciello, S. Rossi, E. Visonà, K. Corona and F. Oliva, "Clinical applications of vibration therapy in orthopaedic practice", MLTJ-Muscles Ligaments Tendons J., vol. 6, pp. 147-156, May 2016.
- [4] D. J. Cochrane, "Vibration exercise: the potential benefits", Int. J. Sports Med., vol. 32, pp. 75-99, February 2011.
- [5] J. Fei, L. Gao, H.-H. Li, Q.-L. Yuan and L.-J. Li, "Electroacupuncture promotes peripheral nerve regeneration after facial nerve crush injury and upregulates the expression of glial cell-derived neurotrophic factor", Neural Regen. Res., vol. 14, pp. 673-682, April 2019.

---

\* 한국한의학연구원 디지털임상연구부, kjd@kiom.re.kr

※ 본 연구는 2024년도 한국한의학연구원 "전신 경락 자극을 위한 침대형 음파 발생장치 개발 (PSN2312010)"과제의 지원을 받아 수행된 연구 결과임

# 시선추적기술을 이용한 경도인지장애의 신속보기 눈운동과 인지기능 평가

김 중 일\*

## Assessment of Saccadic Eye Movements and Cognitive Function in Mild Cognitive Impairment Using Eye Tracking Technology

Joong Il Kim\*

경도인지장애(MCI)는 치매와 정상 노화의 중간 단계의 인지기능이 저하된 상태이다[1]. 기억 장애 유무에 따라 기억상실형(amnestic MCI, aMCI)과 비기억상실형(non-amnestic MCI, naMCI)으로 분류되며, 치매로 진행할 가능성이 높은 질환으로 알려져 있다. 특히, 경도인지장애의 주의집중력과 억제기능 저하는 신속보기(saccade) 눈운동에도 영향을 미칠 수 있어, 신속보기 눈운동의 변화는 경도인지장애를 평가하는 중요한 지표가 될 수 있다[2,3]. 본 연구는 2019년부터 2022년까지 지역 코호트에서 모집된 건강한 노인 221명과 경도인지장애 환자 140명(aMCI: 96명, naMCI: 44명)의 시선추적 데이터를 비교 분석하였다. 시선추적 기술을 이용한 눈운동은 정방향(prosaccade)과 역방향(antisaccade) 신속보기를 통해 주의집중력을 평가하고, 신속보기를 수행(go)하다가도 시각적 단서(cue)에 따라 신속보기를 수행하지 않도록(no-go) 하는 과제(task)를 통해 억제기능을 평가하도록 설계하였다[3]. 주요 매개변수로는 시선의 이동 경로, 눈운동 잠복기, 정반응 비율 등을 활용하였다. 연구 결과, 기억상실형 경도인지장애 환자군은 건강한 노인군에 비해 정방향 신속보기에서 눈운동 잠복기가 유의하게 길었다. 역방향 신속보기에서는, 기억상실형 경도인지장애 환자군이 눈운동 잠복기가 짧아 표적에 대한 반응 속도는 빨랐으나, 정반응 비율이 유의하게 낮아 수행 능력은 떨어진 것으로 나타났다. 또한, 경도인지장애 환자군 모두에서 정방향과 역방향 표적 자극에 대해 반응하지 않은 비율이 건강한 노인군에 비해 높았다. 신속보기를 멈추고 시선을 고정해야 하는 no-go 조건에서 기억상실형 경도인지장애군의 시선 고정 시간이 유의하게 짧았다. 본 연구에서는 시선추적기술을 활용하여 주의집중력과 억제기능 평가 방법을 제시하고, 신속보기 눈운동의 특성이 경도인지장애의 조기 진단 및 질환 진행 예측을 위한 디지털 바이오마커로 활용될 가능성이 있음을 제안한다.

### Key words

mild cognitive impairment, eye-tracking, saccadic eye movements, attention, inhibitory control

### 참고문헌

- [1] E. Oh and A.Y. Lee, "Mild Cognitive Impairment", J Korean Neurol Assoc, vol. 34, no. 3, pp. 167-175, August 2016.
- [2] S. Kim, "Ocular Motor Abnormalities in Alzheimer's Dementia", Clin Neuroophthalmol, vol. 7, no. 2, pp. 31-36, December 2017.
- [3] J. Opwonya, et al. "Eye movement changes as an indicator of mild cognitive impairment", Front Neurosci, vol. 17, pp. 1-10, June 2023

\*한국한의학연구원 디지털임상연구부, jikim@kiom.re.kr(교신저자)

※ 본 연구는 한국한의학연구원 기본사업 과제의 지원을 받아 수행된 연구임(No. KSN2312022).



## 고혈압과 수면시간의 연관성 연구

이 범 주\*

### A study of association between hypertension and sleep duration

Bum Ju Lee\*

고혈압은 인간의 많은 장기들에 중대한 손상을 발생시키며 심혈관질환관련 사망원인의 요인들중 하나이다 [1]. 최근에 이르러, 수면시간과 고혈압의 연관성에 대한 연구들이 수행되었는데, 전세계적으로 짧은 수면시간은 고혈압 발달의 위험요인이 될수 있다는 연구결과들이 보고되었다 [1-6]. 따라서, 본 연구는 한국인 성인을 대상으로 고혈압과 수면시간사이의 연관성을 도출하는 것이다. 질본관리청 국민건강영양조사 2022년 데이터((IRB no. 2018-01-03-4C-A)를 이용하여 총 4041명(남성 1742명, 여성 2299명)의 데이터를 이진 회귀분석을 이용하여 odds ratio (OR)를 분석하였다. 고혈압의 기준은  $140\text{mmHg} \leq$  수축기혈압 또는  $90\text{mmHg} \leq$  이완기혈압 또는 고혈압약 복용을 고려하였다. 수면시간은 주중 (근무일) 하루 평균 수면시간을 기준으로 시간(hour) 단위로 기술하였다. 남성에서 crude 분석에서는 수면시간과 고혈압사이에서 통계적 유의성이 나타났으나 (OR=1.103 [1.003-1.213],  $p=0.044$ ), 나이 및 체질량지수를 보정한 model 1과 나이, 가족수입, 교육수준, 직업, 가족구성원수, 결혼여부, 복부둘레, 체질량지수를 보정한 model 2에서는 연관성이 사라졌다. 여성에서는 모든 모델들에서 통계적 유의성이 나타나지 않았다. 이러한 연구결과는 향후 고혈압 관련 위험요인 분석에 활용 가능하다.

#### Key words

hypertension, sleep duration, association

#### 참고문헌

- [1] Q. Wang et al., "Short sleep duration is associated with hypertension risk among adults: a systematic review and meta-analysis", *Hypertens Res.*, vol. 35, pp. 1012-1018, Oct 2012.
- [2] J. M. Bock et al., "Sleep Duration and Hypertension: Epidemiological Evidence and Underlying Mechanisms", *Am J Hypertens.* vol. 35, pp. 3-11, Jan 2022.
- [3] J. E. Gangwisch et al., "Short sleep duration as a risk factor for hypertension: analyses of the first National Health and Nutrition Examination Survey", *Hypertension*, vol. 47, pp. 833-839, May 2006.
- [4] H. B. Deng et al., "Short Sleep Duration Increases Metabolic Impact in Healthy Adults: A Population-Based Cohort Study", *Sleep*. vol. 40, pp. 10, Oct 2017.
- [5] P. Bansil et al., "Associations Between Sleep Disorders, Sleep Duration, Quality of Sleep, and Hypertension: Results From the National Health and Nutrition Examination Survey, 2005 to 2008", *J Clin Hypertens (Greenwich)*, vol. 13, pp. 739-743, Oct 2011.
- [6] J. E. Gangwisch, "A Review of Evidence for the Link Between Sleep Duration and Hypertension", *Am J Hypertens.*, vol. 27, pp. 1235-1242, Oct 2014.

---

\* 한국한의학연구원, bjlee@kiom.re.kr

※ This study was supported by the Korea Institute of Oriental Medicine (KIOM; Grant nos. KSN1824130, KSN1923111, and KSN2312022) funded by the Korean government.

# 악력과 대사성질환과의 연관성 연구

이 범 주\*

## An association between metabolic disorders and grip strength

Bum Ju Lee\*

최근에 이르러 약한 악력은 많은 대사성 및 심혈관 질환들과 연관성이 높은 것으로 보고되고 있고, 이러한 질환들은 비만지표들과도 매우 높은 연관성이 있다고 보고되고 있다 [1-4]. 따라서, 본 연구에서는 선행연구들과 현 진행중인 연구들을 기반으로 한국인 성인에서 악력 및 비만지표와 대사성질환들과의 연관성을 파악한다. 이러한 연관성 분석을 위해 질병관리청에서 제공하는 공개데이터[5]를 기반으로 complex samples binary logistic regression을 이용하여 odds ratio를 도출하였다. 심근경색/협심증과 악력과의 연관성 연구에서는 비만지표보다도 악력지표들이 본 질환과 연관성이 더 높았으며, 특히 상대악력이 절대악력보다 연관성이 더 높은 경향을 나타냈다 [1]. 대사증후군과 악력의 연구에서는 대사증후군이 절대 및 상대악력과 연관성이 있는 것으로 나타났으나, 실제로 악력보다 비만지표(허리와 키의 비율 지표)가 대사증후군과 더 높은 연관성을 나타냈다 [2]. 또한, 고혈압/당뇨와 악력과의 연관성에서는 두 질병 모두 악력과 연관성이 있는 것으로 나타났으나, 비만지표들이 악력지표 보다 두 질병과 연관성이 더 높은 경향을 보였다. 따라서, 이러한 질병들 모두 악력 및 비만지표와 연관성이 존재하는 것으로 나타났으나, 각 질병들마다 더 높은 연관성을 보이는 지표는 다를 수 있다는 사실을 밝혀냈다. 이러한 연구결과는 향후 정보기술을 기반으로 대중보건 분야에서 대사성질환에 대한 구체적이고 다양한 위험요인 도출에 활용이 가능할 것으로 기대된다.

### Key words

grip strength, metabolic disorders, association

### 참고문헌

- [1] B. J. Lee et al., "Association of relative hand grip strength with myocardial infarction and angina pectoris in the Korean population: a large-scale cross-sectional study", *BMC Public Health*, vol. 24, pp. 941, Apr 2024.
- [2] B. J. Lee, "Grip Strength, Anthropometric Indices, and Their Combination in Screening for Metabolic Syndrome in the Korean Population", *J Clin Med*, vol. 13, pp. 5988, Oct 2024.
- [3] J. P. Lopez-Lopez et al., "The prediction of metabolic syndrome alterations is improved by combining waist circumference and handgrip strength measurements compared to either alone", *Cardiovasc Diabetol.*, vol. 20, pp. 68, Mar 2021.
- [4] C. Shen et al., "Association between handgrip strength and the risk of new-onset metabolic syndrome: A population-based cohort study", *BMJ Open*, vol. 10, pp. e041384, Oct 2020.
- [5] S. Kweon et al., "Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES)", *Int J Epidemiol.*, vol. 43, pp. 69-77, Feb 2014.

---

\*한국한의학연구원, bjlee@kiom.re.kr

※ This study was supported by the Korea Institute of Oriental Medicine (KIOM; Grant nos. KSN1824130, KSN1923111, and KSN2312022) funded by the Korean government.

# 임상 데이터 기반 안면신경마비 치료를 위한 의사결정 지원 모델 개발

임미홍\*

## Development of a model for making clinical decisions for treatment of facial palsy based on clinical data

Mi Hong Yim\*

안면신경마비는 흔히 볼 수 있는 신경병증 질환 중 하나로 얼굴근육 마비로 눈과 입이 틀어지고 심할 경우 안구 건조, 침 분비 감소, 눈물 분비 장애, 청각 장애 등을 초래하는 질환이다[1]. 안면신경마비는 한방치료 선호도가 높은 질환으로 병원과 한방병원 및 자가관리의 다양한 차원의 관리가 필요한 질환이다. 따라서 본 연구에서는 임상 데이터를 기반으로 안면신경마비 치료를 위한 한의사의 의사결정을 지원하는 모델을 개발하고 그 성능을 평가하였다. 2021년부터 2024년까지 대한민국의 2개 한방병원에서 모집된 19세 이상의 안면신경마비 환자 166명의 임상 연구 자료를 분석하였다. 안면신경마비 환자를 대상으로 환자의 상태에 대한 문진, 안면신경마비 평가, 활력징후, 신체계측값 등을 바탕으로 전자침 치료, 한약 치료, 건식 부항 치료 여부를 분류하는 모델을 구축하였다. 모델은 엘라스틱 넷(E-net, Elastic Net), 익스트림 그래디언트 부스팅(XGBoost, Extreme Gradient Boost), 랜덤 포레스트(RF, Random forest), 뉴럴 네트워크(NN, Neural network) 알고리즘을 사용하여 개발되었다. 모델은 전체 데이터의 70%인 train set와 validation set를 사용하여 생성되었고 성능은 전체 데이터의 30%인 test set를 사용하여 평가되었다. 모델 생성을 위한 초모수(hyperparameters)는 5-fold 교차 검정을 사용하여 조정되었다[2]. 모델에서 선택된 환자의 상태에 대한 문진, 활력징후, 신체계측값 변수들에 대한 기여도를 결정하기 위해 상대 변수 중요도를 계산하였다. 모델의 성능은 임계값에 영향을 받지 않는 AUROC (Areas Under the Receiver Operating characteristic Curve)값을 기준으로 평가하였고 임계값에 따라 confusion matrix를 통해 얻어지는 F1 점수, Kappa, 민감도, 특이도 등을 계산하였다. 모델의 성능 비교 결과, 전자침 치료 모델에서는 XGBoost 알고리즘, 한약 치료 모델에서는 E-net 알고리즘, 건식 부항 치료 모델에서는 RF를 적용한 모델이 가장 높은 AUROC 값을 나타냈고 F1 점수와 Kappa에서도 비교적 높은 값을 나타냈다(전자침, AUROC = 0.88 [95% CI, 0.731, 0.983]; 한약, 0.842[0.703, 0.953]; 건식 부항, 0.886[0.769, 0.968]). 전자침 치료 모델에서는 표면 근전도 평가, 신경계 병력 여부, 삶의 질 설문, 총 수면시간 등이, 한약 치료 모델에서는 표면 근전도 평가, 후유증 증상 여부, 신경계 병력 여부 등이, 건식 부항 치료 모델에서는 표면 근전도 평가, 삶의 질 설문, 안면신경마비 육안 평가 등이 각 모델에 강한 영향을 주는 변수로 선택되었다. 임상 데이터 기반 기계학습 알고리즘을 사용한 안면신경마비 치료 모델은 한의사가 안면신경마비 치료를 위한 의사결정을 하는데 도움을 줄 것으로 예상된다.

Key words

clinical data, facial palsy, machine learning, treatment

### 참고문헌

- [1] Finsterer J. Management of peripheral facial nerve palsy. *Eur Arch Otorhinolaryngol.* 2008 ; 265(7) : 743-52.
- [2] LeDell, Erin, Maya Petersen, and Mark van der Laan. "Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates." *Electronic journal of statistics* 9.1 (2015): 1583.

\*한국한의학연구원 디지털임상연구부, mh.yim@kiom.re.kr

※ This work was supported by the Korea Institute of Oriental Medicine (KSN2312022, KSN1824130, KSN1923111).

# 인구·사회·경제학적 변수 및 건강 관련 변수 기반 노인 우울증 분류 모델 비교

임미홍\*

## Comparison of models for classification of depression in the elderly based on socio-economic, demographic variables and health-related variables

Mi Hong Yim\*

우울증은 노인들 사이에서 가장 흔한 정신 건강 문제로, 전 세계적으로 13.3%에서 28.4%의 유병률이 추정된다[1-2]. 우울증은 삶의 질과 신체적 및 정신적 건강에 파괴적 영향을 미치며, 건강한 노화를 방해하는 주된 요인으로 알려져 있다. 따라서 본 연구에서는 인구·사회·경제학적 변수 및 건강 관련 변수를 기반으로 노인 우울증 분류 모델을 개발하고 그 성능을 비교하였다. 2010년부터 2019년까지의 국민건강영양조사 자료를 바탕으로 65세 이상의 성인 13,622명을 대상으로 분석하였다. 우울증 분류 모델을 구축하기 위해 3개의 기계학습 알고리즘(Elastic Net, Extreme Gradient Boost, Neural network)이 사용되었고 인구·사회·경제학적 변수 기반 우울증 분류 모델, 건강 관련 변수 기반 우울증 분류 모델, 인구·사회·경제학적 변수 및 건강 관련 변수 기반 우울증 분류 모델을 생성하였다. 모델 생성을 위해 전체 데이터의 70%인 train set와 validation set이 사용되었고 모델 성능 평가를 위해 전체 데이터의 30%인 test set가 사용되었다. 모델의 성능 평가는 임계값에 영향을 받지 않는 AUROC (Areas Under the Receiver Operating characteristic Curve)값을 기준으로 하였다. 노인 우울증 모델의 성능 평가 결과, 익스트림 그래디언트 부스팅 알고리즘을 사용한 모델이 높은 AUROC 값을 나타냈다. 입력 변수를 바탕으로 모델 성능을 평가해 보면, 건강 관련 변수 기반 우울증 분류 모델(AUROC = 0.725 [95% CI, 0.682, 0.767]), 인구·사회·경제학적 변수 및 건강 관련 변수 기반 우울증 분류 모델(0.707 [0.666, 0.744]), 인구·사회·경제학적 변수 기반 우울증 분류 모델(0.655 [0.613, 0.694]) 순으로 AUROC 값이 높게 나타났다. 건강 관련 변수에서는 만성질환 개수, 이상지질혈증 현재 유병 여부, 1년간 음주 빈도 등이, 인구·사회·경제학적 변수에서는 필요 의료서비스 미충족 여부, 연령, 결혼상태 등이 모델에 강한 영향을 주는 변수로 선택되었다. 기계학습 알고리즘을 사용한 노인 우울증 분류 모델은 임상 및 연구자와 보건 정책담당자의 의사결정에 도움이 될 것으로 예상된다.

### Key words

depression, socio-economic variables, demographic variables, health-related variables, machine learning

### 참고문헌

- [1] Abdoli, N.; Salari, N.; Darvishi, N.; Jafarpour, S.; Solaymani, M.; Mohammadi, M.; Shohaimi, S. The Global Prevalence of Major Depressive Disorder (MDD) among the Elderly: A Systematic Review and Meta-Analysis. *Neurosci. Biobehav. Rev.* 2022, 132, 1067-1073.
- [2] Hu, T.; Zhao, X.; Wu, M.; Li, Z.; Luo, L.; Yang, C.; Yang, F. Prevalence of Depression in Older Adults: A Systematic Review and Meta-Analysis. *Psychiatry Res.* 2022, 311, 114511.

\*한국한의학연구원 디지털임상연구부, mh.yim@kiom.re.kr

※ This work was supported by the Korea Institute of Oriental Medicine (KSN2312022).

# AI 기반 혀 이미지 분석을 통한 한의학 설진의 정량화 연구

최우수\*

## Quantitative Study of Tongue Diagnosis in Traditional Korean Medicine Using AI-based Image Analysis

Woosu Choi\*

한의학에서 설진은 망진의 하나로, 비침습적으로 혀의 색상, 형태, 모양 등을 종합적으로 관찰하여 환자의 건강 상태를 평가하는 중요한 과정이다. 그동안 다양한 혀 이미지를 촬영하고 이를 분석함으로써 설진을 객관화하려는 연구가 활발히 진행되어왔다.[1,2] 최근에는 기술 발전과 함께 AI 기술의 활용이 더욱 용이해지면서 여러 분야에서 AI 기술이 접목된 연구가 진행되고 있으며, 특히 영상처리 분야에서는 비약적인 발전이 이루어지면서 설 이미지 분석 분야에도 다양하게 활용되고 있다.[3] 본 논문에서는 기존의 혀 이미지 분석 기술에 AI 기술을 결합하여 분석을 위한 혀 영역을 보다 정확하게 추출하는 방법을 소개한다. 또한, 혀 이미지 분류를 통해 진단을 수행하는 기법을 제시하며, 기존의 전통적인 기계학습 기법과 AI 기술이 접목된 경우의 결과를 비교 분석한다. 이러한 연구를 통해 얻어진 결과물의 활용 방법도 논의하며, 향후 한의학 진단의 정확성을 높이는 데 기여할 수 있는 가능성을 제시한다.

### Key words

Tongue Diagnosis, Artificial Intelligence, Image Segmentation, Image Classification

### 참고문헌

- [1] C. J. Jung, K. H. Kim, J. S. Jang, & Y. J. Jeon, "Development and evaluation of an indirect illumination for tongue image acquisition", Journal of The Institute of Electronics and Information Engineers, vol. 51 no. 11, pp. 221-228, 2014
- [2] W. Choi, J. Kim, & K. H. Kim, "Tongue Indices with Upper Respiratory Tract Infection for Application in Diagnostic Systems Without Face-to-face Visits". Journal of Biomedical Engineering Research, vol. 44 no. 4, pp. 242-254, 2023
- [3] X. Wang, J. Liu, C. Wu, J. Liu, Q. Li, Y. Chen, & J. Chen, "Artificial intelligence in tongue diagnosis: Using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark" Computational and structural biotechnology journal, vol. 18, pp. 973-980 2020

---

\*한국한의학연구원, harrius@kiom.re.kr

※ 본 연구는 한국한의학연구원 기관고유 AI 한의사 개발을 위한 임상 빅데이터 수집 및 서비스 플랫폼 구축 (KSN1923111) 과제의 지원을 받아 수행되었음.

# 자가조립형 섬유를 활용한 피부 부착형 유기센서 개발

최 승 환\*

## Development of Skin-Attachable Sensor using Self-Assembled Fibers

Seunghwan Choy\*

자가조립형(self-assembly) 섬유 소재를 활용한 피부 접촉형 고감도 센서를 개발하고 활용 가능성에 대해 필요성이 대두되고 있다[1]. 최근 자가조립형 섬유는 뛰어난 기계적 특성과 생체 적합성, 그리고 유기 소재와의 뛰어난 협응성을 보여주며 다양한 센서 응용 분야에서 주목받고 있다[2]. 특히, 유기 트랜지스터(organic transistor) 기술을 접목함으로써, 높은 민감도와 유연성을 갖춘 센서를 구현할 수 있다[3].

본 연구에서는 나노 사이즈의 자가조립형 섬유를 합성하고 자체의 고유한 특성을 평가하였다. 자가 조립된 형태의 섬유는 뛰어난 결정성을 가지며 넓은 영역에 걸쳐 자가조립이 성공적으로 일어난 것을 확인할 수 있었다. 또한, 피부 부착능, 유연성, 수중 안정성, 낮은 피부저항 특성을 보이며 센서 소재로의 가능성을 보여주었다. 또한, 유기 트랜지스터 채널에 자가조립형 소재를 도입하였을 때, 높은 트랜스컨덕턴스(transconductance)와 output 특성을 보여 유기 센서 복합소재로써 가능성을 확인하였다.

### Key words

self-assembly, fiber, organic sensor

### 참고문헌

- [1] P. Wang, M. Wang, J. Zhu, Y. Wang, J. Gao, C. Gao, and Q. Gao, "Surface engineering via self-assembly on PEDOT:PSS fiber: Biomimetic fluff-like morphology and sensing application", Chem. Eng. J., vol. 425, pp. 131551, December 2021.
- [2] J. Kim, J. H. Kang, W. H. Kwon, I. Lee, S. J. Park, C. -H. Kim, W. Jeong, J. S. Choi, and K. Kim, "Self-assembling biomolecules for biosensor applications", Biomater. Res., vol. 2, pp. 127, December 2023.
- [3] J. H. Schon, H. Meng, and Z. Bao, "Self-assembled monolayer organic field-effect transistors", Nature, vol. 413, pp. 713-716, October 2001.

---

\*한국한의학연구원, schoy@kiom.re.kr

※ 본 연구는 한국한의학연구원 '전신 경락 자극을 위한 침대형 음파 발생장치 개발 (PSN2312010)' 과제의 지원을 받아 수행되었음.

# 자산관리셸 표준 API기반 데이터 시각화 웹 애플리케이션 개발

정은아\*, 송병훈\*\*

## Development of a Web Application with Digital Twin Server Based on AAS Standard API

Euna Jeong\* and Byunghun Song\*\*

### 요 약

자율제조 환경에서 물리적 자산의 디지털 정보를 표준화하여 관리하고 공유할 수 있는 자산관리셸(Asset Administration Shell, 이하 AAS) 시각화 프레임워크가 필요하다. 현재 AAS 모델링 시각화 툴인 AASX Package Explorer 는 윈도우 운영체제에만 동작하고 시각화 기능이 부족해 자산 상태를 직관적으로 확인하기 어렵다. 본 논문에서는 디지털 트윈 서버와 연결해 AAS 데이터를 사용자에게 직관적으로 시각화 할 수 있는 웹 애플리케이션을 제안한다. 이 애플리케이션은 AAS 표준 API를 준수하여 오픈소스 디지털 트윈 서버와 호환되며, 자율제조 환경에서 자산관리와 데이터 활용의 효율성을 증대시킬 것으로 기대된다.

### Abstract

The fourth industrial revolution necessitates frameworks for autonomous manufacturing that standardize and share digital information of physical assets in real-time. The Asset Administration Shell (AAS) is essential for efficient asset management and interoperability. However, the primary tool, AASX Package Explorer, is limited to Windows and lacks effective visualization features, hindering easy asset status monitoring. This study proposes a web application that connects with digital twin servers in real-time, visualizing AAS data for intuitive asset monitoring. By adhering to AAS standard APIs, the application ensures compatibility with major open-source digital twin servers, enhancing asset management and data utilization in autonomous manufacturing environments

### Key words

Asset Administration Shell, Digital Twin, Web Application, Autonomous Manufacturing

### 1. 서 론

자율제조[1]의 주요 목적은 디지털 기술을 기반으로 인간의 개입을 최소화한 상태에서 제품의 기

획, 설계부터 생산 및 공급까지 가능하게 하는 시스템을 구축하는 것이다. 이와 관련하여 물리적 자산의 디지털 정보를 체계적으로 관리하고 공유하기 위한 자산관리셸(Asset Administration Shell, 이하

\*한국전자기술연구원, dorocy.jung@keti.re.kr, \*\*한국전자기술연구원, bhsong@keti.re.kr(교신저자)

※ 이 논문은 2024년 정부(산업통상자원부)의 재원으로 한국산업기술기획평가원의 지원을 받아 수행된 연구임 (RS-2024-00442448, 전기차 부품 제조 AI 공통 플랫폼 기술개발 및 실증)

AAS)[2]이 중요해졌다. 현재 다양한 디지털 트윈 서버가 활발하게 개발되고 있지만, 이들 서버와 연결하여 데이터를 사용자에게 시각화해주는 애플리케이션이 부족한 실정이다. 특히, AAS 데이터를 편집하기 위한 도구인 AASX Package Explorer[3]는 Windows 운영체제에서만 설치할 수 있으며, 시각화 기능이 부족해 자산의 상태를 직관적으로 파악하기 어려운 문제가 있다. 이를 극복하기 위해 본 연구에서는 다양한 디지털 트윈 서버와 실시간으로 연결하여 자산 데이터를 시각화하는 AAS 기반 애플리케이션을 설계 및 구현하였다. 구현된 애플리케이션은 AAS 표준 API[1]를 준수하며, Basyx[4], Fa3st[5] 등 다양한 오픈소스 디지털 트윈 서버와 호환될 수 있다. 이는 자율제조 환경에서 자산관리와 데이터 활용을 보다 효율적으로 지원하는데 기여할 것이다.

## II. 설계

본 애플리케이션은 다양한 디지털 트윈 서버와 연결하여 자산 데이터를 시각화하고, 사용자에게 직관적인 정보를 제공하는 것이다. 이를 위해 AAS 표준 API를 활용하여 자산 정보를 효과적으로 조회하고 표시할 수 있는 기능을 구현한다.

애플리케이션 기능은 크게 3가지 부분으로 구성되며 디지털 트윈 서버 정보 입력부, API 호출부, 시각화 데이터 맵핑부이다.

### 1) 디지털 트윈 서버 정보 입력부

해당 부분은 연결하고자 하는 디지털 트윈 서버의 도메인 주소를 입력하는 부분이다. 사용자는 수동으로 대상 서버 주소를 입력하며, 서비스에서는 입력된 주소를 API 주소 체계에 매칭시켜 사용자가 원하는 디지털 트윈 서버와 연결한다.

### 2) API 호출부

표준 문서에서 제안하는 API 체계를 준수하여 디지털 트윈 서버와 데이터를 주고 받기 위한 REST API 호출 부분이다. 사용자가 입력한 디지털 트윈 서버를 기반으로 AAS, Submodel 등을 호출하여 본 애플리케이션에서 사용할 자산관리셀 모델의 정보를 읽어온다.

### 3) 데이터 시각화부

사용자에게 실제 자산의 디지털 정보를 직관적으로 시각화하는 부분이다. API 호출부를 통해 전달된 데이터는 자산관리셀 정보, 자산의 정보, 자산 세부 기능 정보를 통해 사용자는 직관적으로 데이터를 확인할 수 있다.

[그림 1]은 본 서비스의 시퀀스 다이어그램이다. 사용자가 입력한 서버 정보를 기반으로 디지털 트윈 서버와 연결하여 API를 호출하고 데이터를 받아오는 과정을 순서대로 나타낸다. 제공되는 API 중 전체 AAS 정보와 Submodel 모델정보를 불러오는 과정을 대표하여 나타냈으며 하위 정보 또한 표준 API를 기반으로 호출이 가능하다.

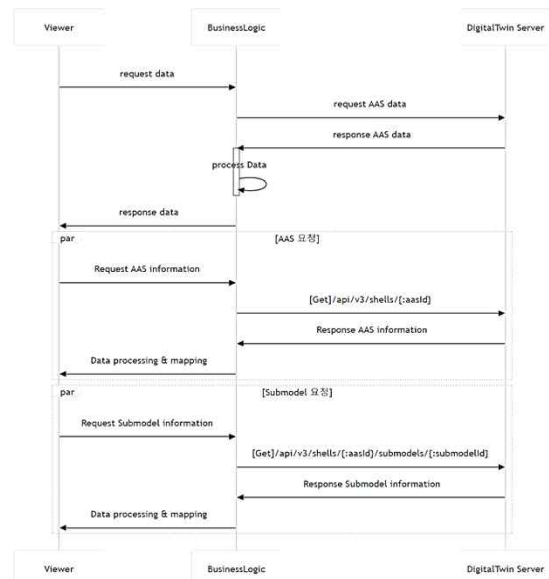


그림 1. 웹 서비스 시퀀스 다이어그램  
Fig. 1. Sequence diagram of web service

## III. 구현 및 검증

구현환경은 Mac 운영체제에서 Typescript 언어를 기반으로 개발되었으며, 연동하기 위한 디지털 트윈 서버는 Fa3st를 사용하였다. 디지털 트윈 서버와 연결이 성공적으로 완료되면 표준 API 주소 체계에 맞춰 데이터를 호출하며 그에 따른 AAS 모델 정보를 확인할 수 있었다. 구현된 웹 서비스의 결과를 검증하기 위해 델타로봇 생산을 위한 공급망 기반 AAS 사용 시나리오에 적용하였다. 사용자는 구현된 웹 서비스를 통해 구입하기 위한 델타로봇의 정보



를 확인한다고 가정한다. 실제 최종 제품인 델타로봇의 AAS 파일은 디지털 트윈 서버인 Fa3st를 통해 데이터를 디지털화 한다. 구현된 웹 서비스는 관리자를 통해 디지털 트윈 서버의 정보를 입력받아 연결을 수행하고, 성공적으로 연결되면 API를 호출하여 모델의 데이터를 불러오는 것을 확인할 수 있었다. 성공적으로 불러온 데이터는 추가적인 로직을 통해 프로세싱과 매핑과정을 거쳐 UI에 적용되어 사용자는 델타로봇 정보를 보다 직관적으로 확인할 수 있었다.

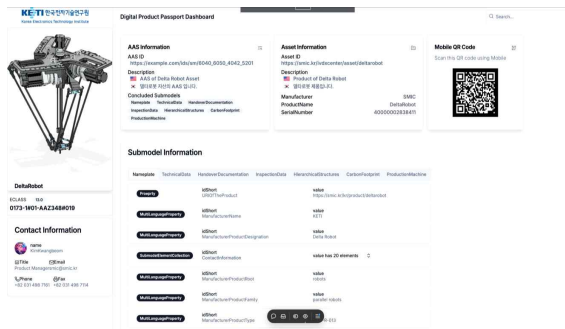


그림 2. AAS Get API 호출 결과  
Fig. 2. Result of AAS GET API

Submodel Information						
Nameplate	TechnicalData	HandoverDocumentation	InspectionData	HierarchicalStructures	CarbonFootprint	ProductionMachine
Property	idShort	URIOfTheProduct	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	ManufacturerName	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	ManufacturerProductDesignation	value	URIOfTheProduct		
			value	URIOfTheProduct		
SubmodelElementCollection	idShort	ContactInformation	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	ManufacturerProductRoot	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	ManufacturerProductFamily	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	ManufacturerProductType	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	OrderCodeOfManufacturer	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	ProductArticleNumberOfManufacturer	value	URIOfTheProduct		
			value	URIOfTheProduct		
Property	idShort	SerialNumber	value	URIOfTheProduct		
			value	URIOfTheProduct		
Property	idShort	YearOfConstruction	value	URIOfTheProduct		
			value	URIOfTheProduct		
Property	idShort	DateOfManufacture	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	HardwareVersion	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	FirmwareVersion	value	URIOfTheProduct		
			value	URIOfTheProduct		
MultiLanguageProperty	idShort	SoftwareVersion	value	URIOfTheProduct		
			value	URIOfTheProduct		
Property	idShort	CoutryOfOrigin	value	URIOfTheProduct		
			value	URIOfTheProduct		
File	idShort	CompanyLogo	value	URIOfTheProduct		
			value	URIOfTheProduct		
SubmodelElementCollection	idShort	Markings	value	URIOfTheProduct		
			value	URIOfTheProduct		

그림 3. Submodel Get API 호출 결과  
Fig. 3. Result of Submodel GET API

### III. 결 론

본 논문에서는 디지털 트윈 서버 정보를 통해 서버와 연결하여 AAS 데이터를 보다 직관적으로 확인할 수 있는 웹 서비스를 제안하였다. AAS 표준 API를 통해 디지털 트윈 서버와 연동하여 데이터를 받아오는 것을 확인할 수 있을 뿐만 아니라 추가적인 처리를 통해 사용자가 보다 직관적으로 데이터를 확인할 수 있다. 또한, QR코드를 통해 모바일로도 접근이 가능한 것을 확인할 수 있었다. 현재 디지털 트윈 서버 한 개만 연결이 가능하며 여러개의 디지털 트윈 서버를 실행할 경우, 각 서버에 개별적으로 연결하여 데이터를 확인할 수 있기 때문에 효율성이 떨어지는 한계가 존재한다. 추후에는 여러개의 디지털 트윈 서버 정보를 입력하여 각 서버와의 연결을 관리하고 사용자의 선택에 따라 포괄적으로 관리할 수 있는 기능을 추가할 예정이다.

### 참 고 문 헌

- [1] D. Y. Lee. Autonomous Manufacturing: Complete Intelligentization beyond Complete Automation. 한국정밀공학회 학술발표대회 논문집, 강원. 2023
- [2] Specification of the Asset Administration Shell Part 1: Metamodel IDTA Number: 01001-3-0. June 2023.
- [3] IDTA-01005-3-0\_SpecificationAssetAdministrationShell\_Part5\_AASXPackageFileFormat.pdf, April 2023.
- [4] Kuhn, T., Antonio, P. O., & Schnicke, F. "Eclipse BaSyx-Industrie 4.0" einfach machen Das Open-Source-Betriebssystem für die Industrie 4.0. atp magazin, 63(10), 54-61, 2021
- [5] Michael Jacoby. "FAST\_Service - An Open Source Implementation of the Reactive Asset Administration Shell", IEEE Transaction on Image Processing, vol. 10, no. 11, pp. 1670-1675, November 2001.

# 사용자 친화적인 UI 기반 학습 모델 서빙 및 관리에 관한 연구

박종원\*, 신준호\*\*

## Research on serving and managing trained models with user-friendly UI

Jong-Won Park\*, and Jun-Ho Shin\*\*

### 요 약

본 논문에서는 학습 모델 배포와 관련된 복잡성을 최소화하면서 사전 학습된 모델을 서빙하기 위해 설계된 사용자 친화적인 UI 기반 모델 서빙 및 관리 애플리케이션을 소개한다. 광범위한 파이프라인 설정과 구성이 필요한 기존 모델 서빙 솔루션과 달리 이 애플리케이션은 프로세스를 간소화하여 사용자가 손쉽게 모델을 서빙할 수 있도록 지원한다.

### Abstract

In this paper, we present a user-friendly, UI-based model serving and management application designed to serve pre-trained models while minimizing the complexity associated with deploying trained models. Unlike traditional model serving solutions that require extensive pipeline setup and configuration, this application simplifies the process and make it easy for users to serve models.

### Key words

user friendly UI, model serving, pre-trained models, deploying trained models

## 1. 서 론

학습 모델 서빙이 비즈니스와 연구에서 중요한 단계로 자리 잡으면서, 이를 효율적으로 관리 및 배포하는 다양한 플랫폼들이 개발되고 있다. 그러나 모델 서빙은 MLOps 파이프라인[1] 중 한 단계로서 학습 파이프라인과 깊이 연관되어 있어 솔루션을 활용하기 위해서는 복잡한 설정과 이해가 필요하다. 이러한 복잡성은 모델 배포에 소요되는 시간과 비

용을 증가시키고, 특히 모델 학습 이후 사전 학습된 모델을 배포하려는 사용자들에게는 높은 진입 장벽으로 작용할 수 있다.

이에 본문에서는 사용자 친화적인 UI 기반의 간편한 방식으로 사전 학습된 모델을 서빙할 수 있는 애플리케이션을 제안한다. 본 애플리케이션은 사용자 친화적인 UI를 통해 모델 등록, 입/출력 유형 정의, 도커 이미지 생성[2] 등의 기능을 통해 효율적인 모델 서빙이 가능하다. 특히, 복잡한 파이프라인

\*한국전자기술연구원, jongunee@keti.re.kr, \*\*한국전자기술연구원, jhshin@keti.re.kr(교신저자)

구축 과정 없이 간단한 버튼 클릭만으로 해당 모델을 기반으로 한 추론 서버를 띄울 수 있어 모델 배포 단계에서 생산성을 향상할 수 있다.

본 논문에서는 이러한 애플리케이션의 기능과 장점을 소개함으로써 모델 배포 효율성 및 생산성을 높이는 방법을 제시하고자 한다.

## II. 애플리케이션 주요 기능 및 아키텍처

본 논문에서 제안하는 애플리케이션의 아키텍처는 다음 그림 1과 같으며 각 다음과 같은 역할을 한다.

- 사용자 인터페이스: 사용자가 접근할 화면으로 코드 없이 원하는 동작이 가능함
- 저장소: 사전 학습된 모델과 모델 서빙의 필요한 파일들, 도커 이미지 등을 저장 및 관리
- 데이터베이스: 모델을 관리를 위한 정보를 포함
- 추론 서버: 학습 모델을 기반으로 한 예측 가능 서버

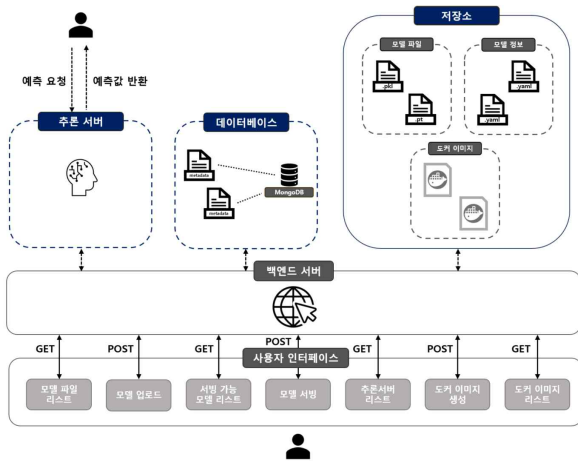


그림 1. 애플리케이션 아키텍처  
Fig. 1. Application architecture

## III. 모델 정의

학습된 모델을 업로드하는 과정에서 모델에 대한 정보를 추가해 저장한다. 필요한 정보는 프레임워크명, 확장자명, 입/출력 유형 정보, API 정보 등이 있다.

- 프레임워크/확장자명: 학습에 사용된 프레임워크와 저장된 확장자명을 의미

- 입/출력 유형: 추론 서버를 사용해 예측할 때 필요한 입/출력 정보
- API 정보: 추론 서버가 제공하는 입력 정보 예시로 즉각적인 테스트에 사용됨

모델을 업로드하며 추가 정보를 저장함으로써 프레임워크에 따른 처리가 가능하고 후에 추론 서버에서 예측할 때 활용하여 손쉬운 서빙이 가능해진다.



그림 2. 모델 정의 UI  
Fig. 2. Model definition UI

```

_id: ObjectId('659b9aad46f28b9020946c6e')
model_name: "wine_quality"
tag: "4so4ef5n6goakxea"
framework: "sklearn"
input_type: "JSON"
output_type: "JSON"
api_data: Object
  fixed acidity:float: "7.4"
  volatile acidity:float: "0.7"
  free sulfur dioxide:float: "11"
  total sulfur dioxide:float: "34"
  density:float: "0.9978"
  pH:float: "3.51"
  sulphates:float: "0.56"
  alcohol:float: "9.4"
  chlorides:float: "1.9"
  citric acid:float: "0"
  residual sugar:float: "1.9"
    
```

그림 3. 데이터베이스에 저장된 모델 정보  
Fig. 3. Model information stored in database

## IV. 모델 서빙

모델 서빙이란 머신러닝으로 학습된 모델을 실제 애플리케이션 환경에서 사용자가 접근하고 활용할 수 있도록 배포하고 운영하는 과정을 의미한다.

본 논문에서는 파이프라인 구성없이 UI 기반으로 버튼만으로 모델 서빙이 가능한 방식을 제안한다. 아래 그림 4의 Run 버튼을 클릭하면 백엔드 서버는 저장소의 모델 파일과 데이터베이스에 저장된 모델 정보를 입력받아 서버를 띄워준다. 이때 정의된 입/출력 유형에 맞게 변환을 해주는 작업을 해주어 사용자는 해당 유형의 데이터 타입의 입력값을 주어 추론 서버로부터 예측값을 반환받을 수가 있다.

## Saved Model List

[Saved Model File List](#) [Saved Model List](#) [Activated Server List](#) [Docker Image List](#)

Supported types

- Framework: Scikit-learn, Keras, Pytorch, Tensorflow
- Input: Numpy/Ndarray, PandasDataFrame, PandasSeries, JSON
- Output: NumpyNdarray, PandasDataFrame, PandasSeries, JSON

No.	Model Name	Tag Name	Framework	Input Type	Output Type	Action
1	wine_quality	4s04ef5n6goakxea	sklearn	JSON	JSON	<a href="#">Run</a>   <a href="#">Create Image</a>

그림 4. 서빙 가능한 모델 리스트 UI  
Fig. 4. Servable model list UI

서빙을 통해 띄운 추론 서버는 아래 그림5와 같은 API 명세서를 제공하며 모델 정의 단계에서 입력한 API 정보를 확인할 수 있다. 예시 API 정보를 통해 사용자는 어떤 입력을 주어야 모델의 예측값을 받을 수 있는지 확인할 수 있고 추론 서버의 반환값의 유형을 확인해 즉시 활용할 수 있다.

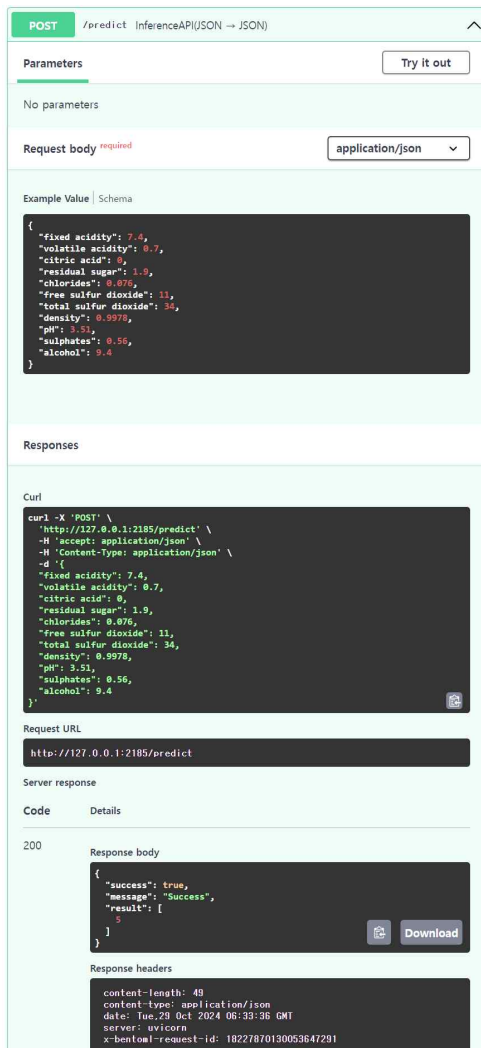


그림 5. 모델 예측 테스트 화면  
Fig. 5. Model prediction test screen

## V. 확장성

추가로 애플리케이션에서는 학습 모델에 모델 서빙 코드를 담은 도커 이미지 파일을 생성해 주는 기능을 포함한다. 이미지화를 통해 쿠버네티스와 같은 컨테이너 오케스트레이션 시스템과 쉽게 통합할 수 있다. 또한, 현재는 JSON 및 배열 형태의 기본적인 입/출력 데이터 유형을 지원하고 있지만, 입력, 언어 프롬프트 등 다양한 데이터 유형까지 확장하여 모델의 활용성을 높일 수 있다.

## VI. 결론

MLOps 생애 주기에서는 무중단 학습 및 배포를 할 수 있어야하므로 일반적으로 학습 모델 운영을 위한 파이프라인 구축이 필요하다. 이에 대한 지식이 없거나 모델 서빙만을 필요로 하면 본 논문에서 제안하는 방식은 높은 활용도를 보여줄 것으로 예상된다. 또한, 챗봇과 같은 언어 모델, 객체 인식 문제 등 범용적인 환경과 데이터만이 필요한 문제에서는 파인튜닝 없이 사전 학습된 모델이 바로 사용될 수 있기 때문에 활용성을 기대할 수 있다.

## 사사

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-00187, 제조 장비·로봇의 지능화를 위한 엣지 브레인 프레임워크 개발)

## 참고 문헌

- [1] Google, <https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning>. August 2024.
- [2] Docker, <https://docs.bentoml.com/en/latest/guides/iotypes.html>, October 2024.
- [3] BentoML, <https://docs.bentoml.com/en/latest/guides/iotypes.html>, October 2024.

# Unity 엔진을 활용한 AAS 3D 데이터 연동 및 시각화 방법

이지현\*, 이재현\*\*

## Methods for Integrating and Visualizing AAS 3D Data Using the Unity Engine

Jihyun Lee, Jaehyun Lee

### 요 약

스마트 제조와 산업 자동화의 발전에 따라 다양한 시스템 간의 자산 데이터를 실시간으로 통합하고 시각화하는 방법이 중요한 과제로 떠오르고 있다. 이에 본 논문에서는 디지털 트윈 (Digital Twin) 구현의 핵심 구성요소인 Asset Administration Shell (AAS)를 활용하여 자산의 3D 정보를 체계적으로 관리하고, Unity 엔진과의 실시간 데이터 통신을 통해 3D 모델을 시각화하는 방안을 제안한다. 제안하는 방법은 자산의 3D 모델 정보에 대한 표준화된 AAS 메타데이터 구조와 API를 기반으로 AAS와 Unity 간의 상호운용성을 향상시킬 수 있으며, 향후 다양한 제조 산업 환경에서 3D 자산 관리와 시각화 업무의 효율성을 개선할 수 있을 것이다.

### Abstract

The advancement of smart manufacturing and industrial automation has highlighted the need for real-time integration and visualization of asset data across diverse systems. This paper proposes a method to systematically manage asset information by utilizing the Asset Administration Shell (AAS), a core component in implementing Digital Twins. By establishing real-time data communication with Unity engine, this approach enables the visualization of 3D models of its asset. The proposed method, based on a standardized AAS metadata structure and API, aims to enhance interoperability between AAS and Unity. This approach is expected to improve the efficiency of 3D asset management and visualization tasks in various manufacturing industry environments in the future.

### Key words

Unity, AAS, smart manufacturing, 3D model visualization, interoperability, digital twin

## 1. 서 론

최근 스마트 팩토리, 사물인터넷 (IoT), 인터넷 보안 등 다양한 산업 분야에서 Asset Administration Shell (AAS)에 대한 관심이 급격히 증가하고 있다.

이는 스마트 제조와 산업 자동화에서 디지털 트윈 (Digital Twin)이 점차 중요해지고 있기 때문이다. 디지털 트윈은 물리적 자산의 가상 복제본으로, 자산을 모니터링하고 제어할 수 있게 한다. 이러한 실시간 운영을 위해서는 표준화된 데이터 교환과 상

\* 한국전자기술연구원, jihyunlee97@keti.re.kr, \*\* 한국전자기술연구원, jh.lee@keti.re.kr

※ 이 논문은 2024년 정부(산업통상자원부)의 재원으로 한국산업기술기술평가원의 지원을 받아 수행된 연구임 (RS-2024-00442448, 전기차 부품 제조 AI 공통 플랫폼 기술개발 및 실증)

호 운용성이 필수적이며, AAS는 이를 위한 핵심 프레임워크로 자리 잡고 있다 [1].

AAS는 다양한 시스템 간의 데이터 통신을 표준화된 방식으로 제공한다. 이를 통해 AAS는 산업 전반에서 데이터의 통합과 상호 운용성을 촉진하고, 다양한 제조환경에서 자산의 효율적 운영을 가능하게 한다. 특히 오픈 소스 소프트웨어의 발전은 AAS 기반의 디지털 트윈 구현을 더욱 용이하게 만들어, 여러 시스템 간의 데이터 교환과 협업을 지원하는 주요 역할을 하고 있다 [2].

본 연구는 AAS와 Unity 간의 3D 데이터 실시간 상호연동 방법을 제안하고, 이를 통해 자산의 3D 모델을 보다 효율적으로 시각화할 수 있게 하는 시스템을 구축하는 것을 목표로 한다. AAS와 Unity의 통합 가능성을 탐구함으로써, 스마트 제조와 디지털 트윈의 실시간 모니터링 및 시각화에 대한 새로운 접근법을 제시하고자 한다.

## II. 본 론

### 1. 3D 데이터의 상호운용을 위한 서브 모델 정의

본 연구에서는 AAS와 Unity 엔진 간에 자산의 3D 데이터를 상호 연동하기 위하여 자산의 3D 모델 파일 관련 메타 정보를 AAS의 표준 규격 기반으로 IDTMetaverse Submodel로 정의하였으며, 다음의 Submodel Elements를 포함하도록 구성하였다.

- UsageProgram: 해당 모델을 사용하는 프로그램 (예: Unity)
- RequiredProgramModule: 모델 실행에 필요한 프로그램 모듈 (예: obj\_importer)
- FileFormat: 파일 형식 (예: obj)
- File: 실제 3D 모델 파일

다만, 자산의 3D 파일과 관련하여 현재 Unity의 무료 버전은 STEP 파일 형식을 지원하지 않기 때문에, 본 연구에서는 대체 포맷으로 OBJ 파일을 채택하여 연구를 진행하였으나 시스템이 지원하는 다양한 포맷의 3D 모델 파일도 IDTMetaverse Submodel에 적용 가능하다.

그림 1의 예시로 구현된 IDTMetaverse Submodel을 보면, AAS와 Unity 간 데이터 통신에서 OBJ 파

일 포맷을 지원하고, AAS 표준 규격 기반으로 3D 데이터 호환성을 확장하는 데 있어 중요한 역할을 한다.

```
SM "IDTMetaverse" [https://example.com/ids/sm/6492_7032_7042_0669]
Prop "UsageProgram" = Unity
Prop "RequiredProgramModule" = obj_importer
Prop "FileFormat" = obj
File "File" => /aasx/files/starterKit.obj
```

그림 1. IDTMetaverse Submodel의 구성

Fig. 1. Composition of the IDTMetaverse Submodel

### 2. 3D 데이터 연동을 위한 AAS 서버 구성

본 연구에서 AAS 모델과 Unity 간의 데이터 상호연동을 위해 Docker 환경 기반의 Eclipse BaSyx 서버를 구성하였다. BaSyx[3]는 오픈 소스 플랫폼으로, 산업 4.0에서 AAS의 다양한 구현을 지원하며, AAS 모델을 효율적으로 관리 및 배포할 수 있는 기능을 제공한다. 이를 통해 다양한 애플리케이션 간의 상호 운용성을 보장할 수 있으며, AAS 기반의 디지털 자산 통합을 위한 핵심 역할을 수행한다.

#### 2.1 Docker와 Docker Compose를 활용한 서버 구성

Docker는 애플리케이션의 배포와 관리를 간소화하는 컨테이너 기반의 가상화 도구이다. 이를 통해 BaSyx 서버 환경을 쉽게 설정하고 운영할 수 있으며, 독립된 실행 환경에서 애플리케이션을 구동할 수 있는 장점이 있다. 특히, Docker Compose는 여러 개의 Docker 컨테이너를 정의하고, 일괄적으로 실행할 수 있는 도구로, 본 연구에서는 Docker Compose를 통해 BaSyx 서버를 구성하였다. BaSyx 플랫폼은 서버가 가동되면 AAS 모델을 저장 및 관리할 수 있는 환경을 제공하며, 사용자 정의한 AAS 모델을 플랫폼에 업로드하고 필요 시 수정 및 조회할 수 있도록 한다. 이와 같이 서버가 성공적으로 구동되면 BaSyx 플랫폼 상에서 업로드된 AAS 모델들이 활성화되어, API 호출을 통해 해당 모델의 데이터를 외부 애플리케이션에서 접근할 수 있다.

#### 2.2 Unity와 AAS API를 통한 데이터 통신

Unity 내에서의 데이터 접근을 위해, 그림 2와 같이 BaSyx 서버에 저장된 AAS 모델의 API 엔드포인트

트를 활용하였다. BaSyx는 AAS 모델과 외부 애플리케이션이 상호작용할 수 있는 RESTful API를 제공한다. 이 API는 표준화된 HTTP 요청을 통해 AAS의 특정 모델 또는 속성에 접근할 수 있게 해주며, 이를 통해 Unity는 필요한 정보를 실시간으로 불러올 수 있다.

본 연구에서는 Unity 프로젝트 내에서 API 주소를 사전 정의하여, AAS 내의 모델 및 그 속성 정보를 필요한 시점에 호출할 수 있도록 설정하였다. 이를 통해 Unity는 BaSyx 플랫폼에서 제공하는 모든 데이터, 즉 AAS 모델의 메타데이터 및 3D 모델 파일 정보를 동적으로 가져와 사용할 수 있다. 예를 들어, AAS 모델에 포함된 3D 객체와 관련 메타데이터를 Unity의 런타임 환경에서 실시간으로 로드하고, 사용자가 원하는 방식으로 시각화할 수 있다.

[1] M. Jacoby, M. Baumann, T. Bischoff, H. Mees, J. Müller, L. Stojanovic, and F. Volz, "Open-source implementations of the reactive Asset Administration Shell: A survey," *Sensors*, vol. 23, no. 5229, pp. 1-22, May 2023.

[2] E. Figueroa, W. Pavon, H. Calvopiña, D. Mideros, and F. C. De la Cruz, "Integration of the Asset Administration Shell (AAS) for Smart Manufacturing: State-of-Art and Future Opportunities," *Preprints*, vol. 2024, no. 06.1271, pp. 1-16, June 2024.

[3] <https://eclipse.dev/basyx/>

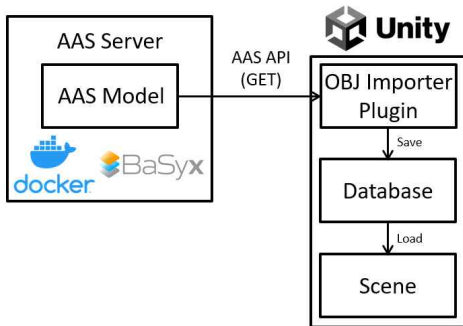


그림 2. Unity와 AAS의 3D 데이터 연동 과정  
Fig. 2. 3D Data Integration Process between Unity & AAS

### III. 결 론

본 연구에서는 AAS와 Unity 엔진 간 실시간 데이터 연동 방법을 제안하였으며, 자산의 3D 모델의 시각화 및 데이터 통합 가능성을 확인하였다. Unity의 OBJ 파일 활용과 런타임 데이터 로딩 설정을 통해 AAS 데이터를 시각화하고, BaSyx 및 Docker 기반의 API 구조를 통해 데이터 교환을 구현하였다. 이를 통해 AAS와 Unity 간의 3D 데이터 상호 운용성을 향상시킬 수 있었으며, 향후 다양한 파일 형식을 지원할 경우 스마트 제조 시스템에서의 활용 가능성을 더욱 더 높일 수 있을 것으로 예상된다.

### 참 고 문 헌

# 자산관리셸 모델을 활용한 로봇의 작업계획 프로세스 설계

류승민, 이하림

## Robotic Task Planning Process with Asset Administration Shell Model

SeungMin Ryu, Harim Lee

### 요 약

최근 유연한 생산 시스템에서 로봇을 활용하는 다양한 방법에 대해 많은 연구가 진행되고 있다. 그중 로봇이 자신이 해야 할 일을 스스로 인지하고 수행하기 위한 기술인 작업계획(Task Planning) 능력은 로봇을 활용한 유연 생산 시스템에서 필수적인 능력이다. 이에 본 논문에서는 디지털 트윈 표현 모델인 자산관리셸(Asset Administration Shell)을 활용하여 생산 제품의 모델에 로봇이 수행해야 할 작업 내용을 모델링하고 이를 활용하여 로봇의 작업을 계획하는 프로세스에 대해 설계한다.

### Abstract

Recently, various methods of utilizing robots in flexible manufacturing systems have been actively studied. Among these, task planning, a technology that enables robots to autonomously recognize and perform their tasks, is an essential capability in flexible manufacturing systems using robots. In this paper, we design a process that models the tasks a robot needs to perform using the Asset Administration Shell (AAS), a digital twin representation model, for the production product. We then utilize this model to plan the tasks for the robot in the manufacturing process.

### Key words

Task Planning, Asset Administration Shell, Flexible Manufacturing System

## 1. 서 론

최근 소비자들의 니즈가 다양해지고 제품의 수명 주기가 짧아짐에 따라 생산라인의 변경에 유연한 대처가 가능한 유연한 생산 시스템의 필요성이 커지고 있다. 하지만 기존의 로봇을 활용하는 공정에서는 정해진 환경에서 일정한 작업만을 반복적으로 수행하기 때문에 환경 변화나 작업 내용의 변경에

유연하게 대처하지 못한다는 단점이 존재한다.[1] 예를들어, 생산부품을 옮기는 픽애플레이스 작업에서 생산부품이 변경되어 플레이스의 위치가 변경되거나 작업 내용이 변경되는 상황에서 공정을 멈추고 사람이 다시 교시해야 하는 상황이 발생한다. 이렇기 때문에 로봇이 스스로 변화된 작업 내용을 인지하여 작업을 수행한다면, 사람의 개입없이 다양한 공정에서 로봇의 활용이 수월해질 것이다. 따라서

\* 한국전자기술연구원, tmd2938@keti.re.kr,



본 연구에서는 물리적 자산의 디지털 트윈을 표현하고 관리하는 프레임워크인 자산관리셸(Asset Administration Shell, AAS)을 활용하여 최근 로봇의 유연성추가해주기 위한 장치인 틸체인저를 활용하기 위해 생산 제품을 잡기위해 필요한 엔드이펙터의 정보, 생산제품이 결합 또는 플레이스 되어야할 위치정보 등을 모델링해주고, 상위 시스템에서 자산관리셸 정보를 교환하고 로봇 제어에 활용하여 변경된 작업 내용에 유연하게 대처할 수 있는 프로세스를 설계한다.

## II. 작업계획 프로세스

제품의 자산관리셸 모델을 활용하여 로봇의 작업을 계획하기 위해서는 자산관리셸 모델에 로봇이 수행해야 할 작업 내용이 들어가야 한다. 그림 1은 픽애플레이스 작업이 필요한 생산 제품의 자산관리셸 모델을 AASX Package Explorer[2]에서 본 예시이다. 자산관리셸은 여러개의 서브모델로 구성되어있으며 작업계획에 활용하기 위해 하나의 사용자 정의 서브모델인 OperationalData 서브모델을 구축하였다. OperationalData 서브모델에는 작업 내용에 맞게 사용해야 하는 엔드이펙터의 정보를 명시하고 본 생산 제품이 로봇을 활용한 공정에서 어떤 작업이 필요한지에 대한 작업 내용을 추가해주었다. 픽애플레이스 작업이 필요한 생산제품의 경우 서브모델에는 픽애플레이스 작업을 명시해주고 사용해야할 엔드 이펙터의 모델과 플레이스 해야 할 위치정보가 들어있는 6자유도 정보를 작성해 주었다.

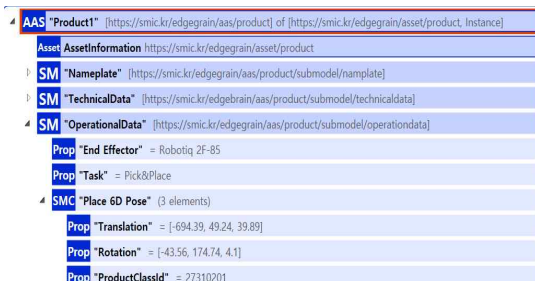


그림 1. 픽애플레이스 해야 할 생산제품의 자산관리셸 예시

Fig. 1. Example of an Asset Administration Shell for a Pick-and-Place Production Product

그림 2은 로봇이 작업을 수행하기까지의 프로세스를 순서도로 나타낸 그림이다. 비전 시스템에서 생산 제품을 촬영하고 현재의 생산 제품에 맞는 자산관리셸을 AAS Registry에 접속하여 접속 정보를 받아오게 된다. 이때 모델링한 자산관리셸의 정보를 상위 시스템에서 활용하기 위해 다양한 소프트웨어에서 활용이 가능하도록 API(HTTP, MQTT)를 구축해야한다. 이러한 Reactive AAS를 구현하기 위한 다양한 오픈 소스들이 존재하는데 본 연구에서는 독일의 Fraunhofer IOSB연구소에서 구현한 Fa3st Service[]를 활용하여 Reactive AAS를 구현하였으며, HTTP API를 통해 상위 시스템에서 자산관리셸의 정보를 교환한다.

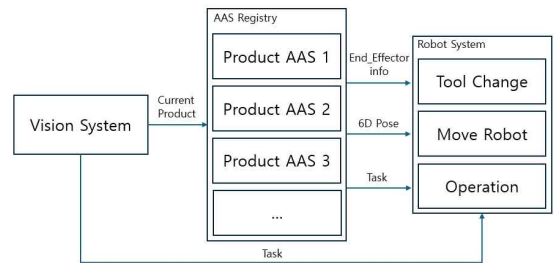


그림 2. 로봇 작업의 전체 구성도  
Fig. 2. Overall Structure of Robot Tasks

HTTP API를 통해 받은 생산제품의 End Effector 정보, 수행해야 할 Task 정보, 작업 수행을 위한 6D Pose 정보 등을 기반으로 틸체이저를 활용해 상황에 맞는 로봇의 End-Effector로 변경하고 로봇의 작업을 수행할 수 있다. 이를 위해 HTTP API를 ROS2 기반의 Topic 형태로 변경하여 로봇 시스템에 전달해주어야 한다. 로봇 제어에는 오픈소스 기반의 로봇 제어 프레임워크인 MoveIt2[4]를 기반으로 시뮬레이션 환경인 디지털 트윈을 구축하여 경로를 계획하고 로봇을 안전하게 제어할 수 있도록 하였다. ROS2 Topic으로 변경된 생산제품의 6D Pose 정보를 MoveIt API를 활용하여 로봇을 프로그래밍 없이 쉽게 제어할 수 있도록 로봇 시스템을 구축하였다.

## III. 결 론

본 연구에서는 다관절 로봇을 대상으로 작업을 계획할 때 생산제품의 자산관리셸 모델을 활용하여

로봇이 수행해야 할 작업 내용이 변경되더라도 사람의 개입없이 자산관리셀 정보를 기반으로 로봇의 작업을 변경하는 프로세스를 제안하였다. 본 연구에서 제안하는 프로세스를 통해 가변적인 환경에서 빠르고 유연하게 대처 가능한 로봇 프로그램 개발이 가능해질 것으로 기대한다. 또한 산업 자동화 및 제조 분야에서 디지털 트윈 구현과 상호운용성을 증진시키기 위한 자산관리셀 기술을 활용하여 다양한 제조사의 장비나 시스템 간에도 일관된 데이터 이해를 통해 데이터 상호운용이 가능할 것으로 기대된다.

### 참 고 문 헌

- [1] 백승혁, 신성호, 강래영, 이성주, 이주순, 육태경, 전창현, 김종원, 유연국, 박성호, 윤준호, 이준석, 이규빈. (2021). 물품조립 AI - 조립설명서 기반 로봇의 작업계획 생성 시스템. 로봇과 인간, 18(1), 3-15.
- [2] AASX Package Explorer.  
<https://github.com/eclipse-aaspe/package-explorer>
- [3] FraunhoferIOSB Fa3st-Service.  
<https://github.com/FraunhoferIOSB/FAAAST-Service>
- [4] MoveIt2.  
<https://github.com/moveit/moveit2>

# 이기종 설비의 실시간 데이터 연동을 위한 디지털 트윈 기반 시뮬레이션 방안에 관한 연구

유현수\*, 이진성\*\*, 윤성준\*\*\*

## Research of simulation based on digital twin for real-time data linkage of heterogeneous equipment

Yu Hyeonsoo\*, and Lee Jinseong\*\*, and Yoon seongjun\*\*\*

### 요 약

본 논문에서는 FMS 라인에서 모니터링 및 프로세스 최적화를 위해 설계된 디지털 트윈 기반 시뮬레이션 방안을 제시한다. Visual Components의 3D 모델링을 사용하여 OPC UA를 통해 실시간 데이터 연동 환경을 구축하여 70개의 신호를 맵핑하고 4가지 시나리오를 시뮬레이션한다. 이를 통해 제조 환경에서 생산성 향상의 방향성을 제안하고자 한다.

### Abstract

We propose a digital twin-based simulation method designed for monitoring and process optimization in FMS lines. Using 3D modeling of Visual Components, we build a real-time data linkage environment via OPC UA, map 70 signals, and simulate four scenarios. Through this, we aim to suggest a direction for productivity improvement in the manufacturing environment.

### Key words

digital twin, simulation, opc ua, visual components

## 1. 서 론

디지털 트윈은 물리적 개체, 시스템 또는 프로세스의 가상 표현으로 물리적 개체의 실시간 데이터를 지속적으로 업데이트한다. 2000년대 초반에 처음 구상된 이 개념은 이후 Industry 4.0의 핵심 구성 요

소로 발전하여 제조, 의료 및 기타 부문의 운영 효율성을 향상하고 가동률을 높여 정확한 미래 예측이 가능하다.[1] 디지털 복제본을 생성함으로써 물리적 시스템에 변경 사항을 구현하기 전에 가상 시나리오를 수행하고 위험을 평가하여 통제된 디지털 환경에서 프로세스를 최적화할 수 있다. 제조 분야

\* 한국전자기술연구원, gustn1847@keti.re.kr, \*\* 한국전자기술연구원, jinsl127@keti.re.kr(교신저자),  
\*\*\* 한국전자기술연구원, sjyoon@keti.re.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-00866, 대규모 확장성 및 고신뢰 분산 시뮬레이션을 지원하는 제조 디지털 트윈 프레임워크 기술 개발)

에서 생산 라인을 시뮬레이션하여 병목 현상을 식별하고 낭비를 줄일 수 있다.[2] 마켓앤마켓은 디지털 트윈 기술은 급속한 성장을 이어가며, 2023년 101억 달러에서 2028년 1,101억 달러로 61.3%의 초고속 성장을 유지할 것으로 전망한다.[3] 그러나 디지털 트윈의 구현은 이기종 데이터 소스의 통합, 실시간 데이터 처리, 시뮬레이션 복잡성 관리에서 어려움을 겪고 있다.[4]

본 논문에서는 디지털 트윈의 구현을 위한 시뮬레이션에 중점을 두고 이기종 설비의 실시간 데이터 연동 및 공정 최적화 방안을 제안한다.

## II. 모델링 및 데이터 연동

모니터링 및 시뮬레이션 디지털 트윈 구현을 위해 Visual Components 프로그램을 통해 3D 모델링을 진행한다. 그림 1.과 같이 FMS 라인을 설계하는데 사용된 장비는 1) Robot, 2) Machine, 3) Conveyor, 4) Track이고, 추가로 5) Fence, 6) Table을 통해 생산 라인의 레이아웃을 표현했다.

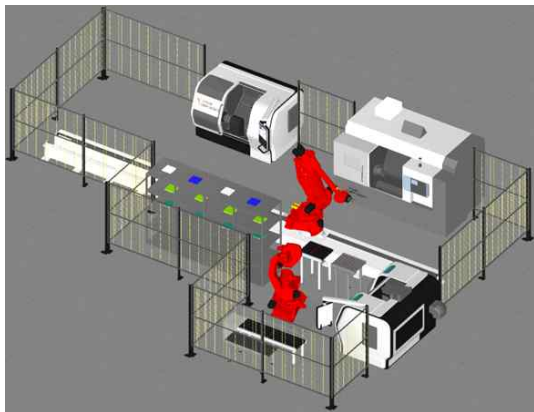


그림 1. FMS 라인 3D 모델  
Fig. 1. 3D Model of FMS Line

OPC UA 기반으로 구성된 서버 및 데이터 통신을 통해 설비에서 발생하는 데이터를 실시간 연동하는 환경을 구축한다. 그림 2.는 로봇과 기계 각각의 OPC UA 서버를 클라이언트 역할을 하는 Visual Components에서 3D 모델과 실시간 데이터를 연동하는 구조를 보여준다.

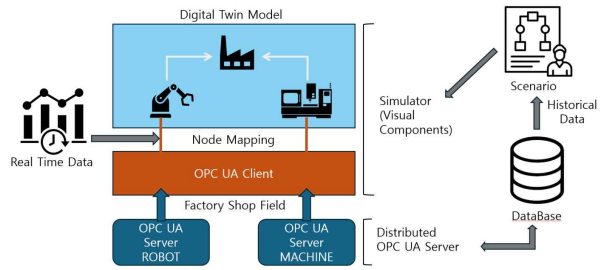


그림 2. OPC UA 서버 연동 및 시뮬레이션 아키텍처  
Fig. 2. Architecture of OPC UA server linkage and simulation

로봇의 Axis, Gripper 값과 가공기의 Call Status, Product Information 값 등 70개의 신호를 대응하는 노드에 맵핑하고 50ms의 주기로 업데이트한다. 그림 3.과 같이 실시간 데이터(UaExpert)와 3D 모델 노드에 맵핑된 연동 데이터(Visual Components)를 확인한다.

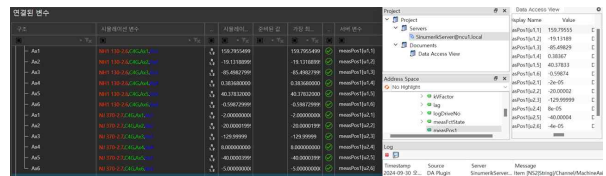


그림 3. 실시간 데이터 노드 맵핑 화면(Visual Components / UaExpert)  
Fig. 3. Real-time data node mapping screen(Visual Components / UaExpert)

## III. 시뮬레이션

실제 FMS 라인에서 생산되는 공정이 다른 제품 3가지 데이터를 DataBase에 저장하고 Historical data를 토대로 시나리오를 만들어 시뮬레이션한다. 그림 4.의 4가지 시나리오는 하나의 제품이 생산되는 경우 3가지와 3개의 제품이 동시에 생산되는 경우 1가지이다. 표 1.에서 볼 때 Case 2가 가장 생산성이 높았으며 Case 1은 1개의 제품을 생산했음에도 불구하고 Case 4와 시간이 크게 차이 나지 않았다. 이로 인해 생산하는 제품의 종류보다 사용한 가공기의 수와 우선순위가 생산량과 사이클 타임에 영향을 미치는 것을 알 수 있고 특히, OP20에서 병목 현상이 발생하는 문제점을 파악할 수 있다.

그림 4. 시뮬레이션 시나리오  
Fig. 4. Simulation scenario

표 1. 시나리오별 시뮬레이션 시간

Table 1. Scenario simulation time

Case	사용한 가공기	Simulation time
1	OP10, 20, 30	08:38
2	OP10, 20	06:06
3	OP20, 30	08:13
4	OP10, 20, 30	08:59

#### IV. 결 론

본 논문에서는 3D 모델링과 OPC UA 기반 데이터 연동을 통해 로봇, 가공기 등 구성 요소의 실시간 데이터를 시뮬레이션 환경에서 확인하고 시나리오별 싸이클 타임을 확인하여 디지털 트윈 기반 시뮬레이션 방안을 제안했다. 향후 맵핑 데이터 포인트를 늘리고 다양한 시나리오를 통해 공정 구성 및 제품 유형에 따른 생산 효율성을 높이는 연구를 확장 진행할 계획이다.

#### 참 고 문 헌

[1] Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. Digital twin in industry: State-of-the-art. \*IEEE Transactions on Industrial Informatics\*, vol. 15, no. 4, pp. 2405-2415, April 2019.

[2] Fuller, A., Fan, Z., Day, C., & Barlow, C. Digital twin: Enabling technologies, challenges and open research. \*IEEE Access\*, vol. 8, pp. 108952-108971, May 2020.

[3] Digital Twin Matket - Global Forecast to 2028.

[4] Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. Characterising the digital twin: A systematic literature review. \*CIRP Journal of Manufacturing Science and Technology\*, vol. 29, pp. 36-52, Jun 2020.

# 대규모 언어 모델을 활용한 협동 로봇 제어 시스템 개념 설계

이정현\*, 김윤한\*\*

## Conceptual Design for Control System for Collaborative Robots Using Large Language Models (LLM)

Jeonghyun Lee\*, and Yunhan Kim\*\*

### 요 약

본 연구에서는 대규모 언어 모델을 활용한 협동 로봇의 제어 시스템을 위한 새로운 개념 설계를 제안한다. 기존에 협동 로봇을 제어하기 위해서는 로봇 프로그래밍에 대한 이해를 바탕으로 작업자에 의해 동작이 수동으로 프로그래밍 되었다. 이러한 로봇 프로그래밍 과정을 보다 직관적이고 효율적으로 개선하기 위해, 본 연구에서는 인간의 언어를 이해하고 추론할 수 있는 대규모 언어 모델을 활용하고자 한다. 구체적으로, 대규모 언어 모델 기반의 협동 로봇 제어 시스템은 인간의 자연어로 이루어진 명령을 입력받아 로봇의 작업 계획을 자동화한다. 이러한 개념 설계를 검증하기 위해, 본 연구에서는 실제 협동 로봇을 활용하여 로봇의 동작을 계획하고 가시화할 수 있는 시스템을 구축하였다. 해당 시스템을 통해, 대규모 언어 모델 기반의 제어 시스템의 개념 설계를 검증하고 이를 고도화하여 실제 제조 현장에서 인간-로봇 상호작용을 직관적이고 효율적으로 개선할 수 있다.

### Abstract

In this study, we propose a new conceptual design for a collaborative robot control system that utilizes large-scale language models. To control existing collaborative robots, movements were manually programmed by the operator based on an understanding of robot programming. To make this robot programming process more intuitive and efficient, this research seeks to leverage large language models that can understand and reason about human language. Specifically, a large language model-based collaborative robot control system accepts instructions consisting of human natural language and automates the robot's work planning. In order to verify this conceptual design, this research utilized an actual collaborative robot to build a system that can plan and visualize robot movements. This system allows you to validate and advance conceptual designs of large language model-based control systems to intuitively and efficiently improve human-robot interaction on real-world manufacturing sites.

### Key words

Large Language Model, Deep Learning, Collaborative Robot

---

\*한국전자기술연구원, jeonghyunlee@keti.re.kr, \*\*한국전자기술연구원, yunhan.kim@keti.re.kr(교신저자표시), ...

## 1. 서 론

협동 로봇(Collaborative Robot)은 사람과 같은 공간에서 사람과 물리적으로 상호작용하며 사람의 보조 역할로 작업하는 로봇을 말한다. 제조 기업들은 기존의 산업용 로봇에 비해 협동 로봇이 다양한 공정 및 작업에 유연하게 활용된다는 점에서 높은 잠재력을 보유한 것으로 여기며 최근 협동 로봇을 적극적으로 도입하고 있다.[1]

기존에 협동 로봇은 로봇 프로그래밍 언어를 활용하여 정해진 환경에서 동작을 수행하기 위한 과정을 사전에 입력하여, 정해진 동작을 반복적으로 수행한다. 그러나 기존의 로봇 동작 제어 과정은 모두 규칙 기반 작업이기 때문에 협동 로봇은 다양하고 복잡한 명령의 숨겨진 패턴과 추세를 발견하지 못한다. 또한 규칙 기반이다보니 특정된 업무 영역에서 제한된 사람들만 사용할 수 있다.

본 논문에서는 상기한 한계를 개선하기 위하여 대규모 언어 모델을 사용할 것을 제안한다. 대규모 언어 모델(Large Language Model, LLM)은 대규모 텍스트 데이터를 학습하여 단어의 앞 뒤뿐만 아니라 전체 문맥을 바탕으로 주어진 텍스트에 맞는 텍스트를 생성하는 모델이다. OpenAI는 GPT 시리즈를 공개하였는데, 해당 모델들은 다양한 소스의 텍스트 데이터를 혼합하여 사전 학습함으로써 지식을 습득하고, 광범위한 도메인에서 일관되고, 맥락적으로 관련성 있는 텍스트를 생성한다. GPT-4.0이 각종 시험에서 상위 10%에 해당하는 성적을 거둔 결과로 봤을 때 LLM 모델이 자연어 명령을 이해하고 그에 맞는 결과를 생성한다는 것을 알 수 있다[2].

인간의 언어를 이해하고 추론할 수 있는 LLM을 활용하면, 협동 로봇이 실시간으로 명령의 패턴과 추세를 발견하여 자연어를 이해하고 사용자의 명령에 따라 특정 동작이 실행 되도록 할 수 있다. 이때 사용자와의 상호작용을 통해 지속적인 학습이 이루어지도록 한다.

이러한 접근 방식은 제조업의 다양한 작업 환경에서 Pick and place, 기계 조작과 조립 등을 처리하는 데 있어 로봇이 감지, 적응, 재프로그래밍할 수 있는데 도움이 될 수 있다. 또한 자연어를 데이터로 다루기 때문에 사용법이 쉬워 기존 로봇 프로그래밍 방법보다 사용자의 범위를 확대할 수 있다.

따라서 LLM을 사용해 사용자의 명령에 따른 제조업의 협동 로봇 자동화 실현을 위한 본 연구를 제안한다.

## II. 제안하는 대규모 언어 모델 기반 협동 로봇 제어 시스템

### 2-1. 개념 설계 검증을 위한 협동 로봇 테스트 베드 구축

본 논문에서 다루고 있는 개발 환경은 다음과 같다. 제어 프로그램은 협동 로봇 모델의 동작을 제어하기 위해 주어지는 소프트웨어이다.

표 1. 개발 환경

Table 1. Development environment

협동 로봇 모델	Techman Robot TM5-700
제어 프로그램	TMflow v.1.86
제어 환경	Ubuntu 20.04 / ROS2 Foxy
개발 언어	Python

다음은 로봇 컨트롤러 이미지이다. HDMI, Gripper USB, LAN 포트에 케이블을 연결했다.



그림 1. 로봇 컨트롤러  
Fig. 1. robot controller

## 2-2. LLM을 활용한 로봇 제어 시스템 개념 설계

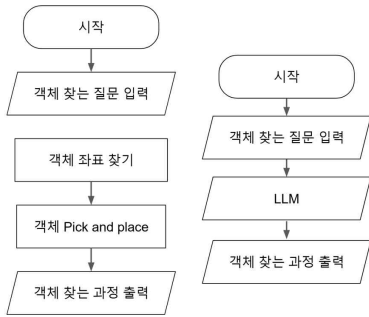


그림 2. LLM 사용하지 않은 동작 흐름도(왼), LLM 사용한 동작 흐름도(오)  
Fig. 2. LLM unused operation flowchart (left), Operation flowchart using LLM (right)

LLM을 사용하지 않은 동작 흐름도는 프론트엔드단에서 들어온 자연어를 LLM이 아닌 언어 모델의 입력으로 두고 그에 맞는 로봇의 동작 수행 과정을 결과값으로 출력한다.

Techman Robot 컨트롤러와 통신하기 위해 TM ROS2 드라이버를 설치한 후, 로봇을 제어할 수 있는 프로그램인 TMflow에 연결해 로봇 이동 등을 제어할 수 있게 한다.

TM Robot의 TMvision(내장비전시스템)으로 HTTP 메소드를 사용하여 원격 ROS 컴퓨터에서 이미지 데이터를 받을 수 있게 한다.

저장된 이미지 데이터를 객체 인식 인공지능 모델에 입력해 객체의 좌표를 구한다.

앞서 구한 좌표로 로봇을 이동시켜 Pick and place를 진행한다. 이 때 해당 좌표에 위치한 객체가 하나만 있다면 그 물체를 바로 Pick and place를 진행한다. 하지만 두 가지 이상이라면 z좌표를 비교해 찾고자 하는 객체보다 z좌표가 더 큰 값이 있다면 그 물체를 다른 방향으로 random 값을 추가해 위치를 옮긴다.

LLM을 사용한 동작 흐름도는 입력으로 텍스트 데이터가 입력되면 LLM 모델을 통해 최종 결과를 알아내고 로봇 스스로 목적 달성을 위한 최적화된 동작 순서를 생성해 과정을 출력한다.

## 2-3. 협동 로봇의 동작 시각화

또한 개발 환경에서 MoveIt2 패키지를 설치해 RViz GUI를 사용할 수 있다. 이는 충돌 없는 동작 수행을 계획하고 실행하는 방법을 사용자에게 보여준다.

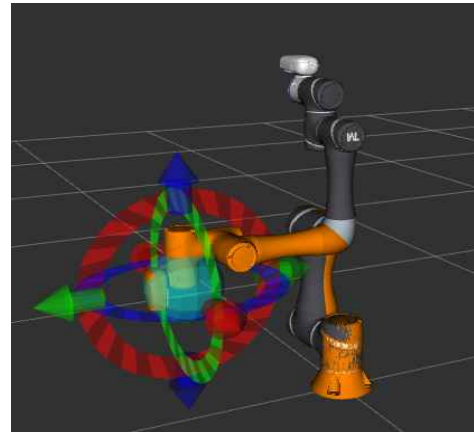


그림 3. RViz에서 로봇을 시각화한 모습  
Fig. 3. Visualizing robots in RViz

여기서 주황색으로 표현된 부분은 목표 지점이고 회색으로 표현된 부분은 현재 지점이다.



그림 4 로봇의 기존 위치(왼), 목표 지점으로 이동한 로봇(오)  
Fig. 4.(left) Existing position of the robot (left), Robot moved to target point (right)



회색으로 표현된 위치(원)에서 주황색으로 표현된(오) 목표 지점으로 로봇을 동작 시킨 이미지이다

### III. 결 론

그동안 제조업에서 협동 로봇을 활용한 동작 제어 과정은 규칙 기반 학습 방식으로 진행되어 왔다. 그러나 제조업에서 발생하는 다양한 데이터와 복잡한 작업 환경에서 높은 정확도로 실시간 결과를 출력하기에는 한계가 있었다. 특히 새로운 데이터가 입력될 때 즉각적으로 정확한 답을 출력하지 못하는 문제점이 있었다. 이를 개선하기 위해 본 논문에서는 LLM의 도입을 제안하였다. LLM은 높은 정확도로 결과를 출력할 뿐만 아니라, 다양한 작업과 환경에 유연하게 대응할 수 있는 장점을 가지고 있어 기존 문제를 효과적으로 해결할 수 있을 것으로 기대된다.

또한 로봇 트래킹 기술을 활용하여 동작을 제어하고 명령 수행 과정을 출력하는 방안을 검토하였다. 이 부분은 향후 연구 과제로 남겨두어 추가적인 연구가 이루어져야 할 것이다.

### 참 고 문 헌

- [1] Deloitte, "Deloitte Survey on AI Adoption in Manufacturing", 2020.
- [2] OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, and Florencia Leoni Aleman et al, "GPT-4 Technical Report", arXiv, pp. 6, 2024

# 제조 환경 모니터링을 위한 OPC UA 발행/구독 모델 연구

이원희\*

## A Research on OPC UA PubSub Model for Manufacturing Environment Monitoring

Wonhee Lee\*

### 요 약

제조 환경의 공정은 많은 수의 로봇, 디바이스, 센서 등이 연결되어 동작하고, 다수의 현장 작업자가 공정을 운영한다. 특히 제조현장을 안정적으로 관리하기 위한 운영 애플리케이션이 지능화되고 필요한 데이터가 다양해지면서, 데이터 연계 인터페이스의 통합 표준 기술의 요구가 있어왔다. 이러한 표준 기술은 벤더 종속적이지 않도록 상호운용성을 제공함과 동시에 수 많은 장치 연결을 위한 확장성을 제공하는 형태로 기술이 발전해 왔다. 특히 제조분야에서의 데이터 연동 표준인 OPC UA는 오랜시간 숙성된 기술로 제조 벤더사가 필수적으로 장치에 빌트-인 하여 시장출시되어 왔다. 다만, OPC UA의 기본 구조상 많은 수의 장치 연결을 담보하지 못하는 구조적 한계로 발행/구독에 관한 표준연구가 진행되어야 왔으며, 적용 방법에 관한 기술개발이 필수적으로 요구된다. 이에 본 논문에서는 제조 환경을 모니터링할 수 있는 디바이스 데이터를 수집하여 OPC UA 발행/구독 서비스 가능한 구조를 설계하고, 테스트베드 환경 구성을 통해 기능 검증한 결과이다.

### Abstract

The manufacturing environment process is operated by connecting a large number of robots, devices, and sensors, and a large number of field workers operate the process. In particular, as the operational applications for stable management of the manufacturing site become intelligent and the required data diversifies, there has been a demand for integrated standard technology for data linkage interfaces. This standard technology has developed in a way that provides interoperability without vendor dependency and scalability for connecting a large number of devices. In particular, OPC UA, a data linkage standard in the manufacturing field, is a technology that has been matured for a long time and has been released to the market by manufacturing vendors by building it into their devices. However, due to the structural limitations of the basic structure of OPC UA that cannot guarantee a large number of device connections, standard research on publication/subscription has been conducted, and technology development on application methods is essential. Accordingly, this paper collects device data that can monitor the manufacturing environment, designs a structure that enables OPC UA publication/subscription service, and verifies its function through a testbed environment configuration.

### Key words

OPC UA PubSub, Published/Subscribed, Open Platform Communication Unified Architecture

---

\*한국전자기술연구원 자율제조연구센터, zsed50@keti.re.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-00962, 제조 현장 맞춤형 Industrial XR 프레임워크 개발)

## 1. 서 론

최근 AI, 빅데이터 등 데이터를 중심으로 한 제조공정혁신 개편으로 자율제조가 전세계적 흐름이다. 이에 따라 제조공정을 구성하는 다양한 형태의 데이터를 활용한 지능화 서비스가 연구개발되어 제조현장에 적용되고 있으며, 특히 제조산업 전반에 데이터 수집과 통합 관점에서의 표준 적용에 관한 정책적 요구가 있다. 관련 표준으로 OPC UA, AAS 등의 인터페이스 표준이 필수 적용 기술로써 지속적인 연구개발이 진행되어 왔고, 이에 따라 관련 장치를 보급하고 있는 벤더사에서 해당 기술을 적용한 장치를 시장에 선보이고 있다.[1][2] 이와같이 제조현장에 적용되고 있는 표준 기반 장치는 기본적으로 서버-클라이언트 구조 모델을 수용하고 있는 상황으로 제조현장의 다양한 장치, 센서 데이터를 연동하기에는 구조적 한계가 있다. 이와같은 문제를 해결하기 위한 방안으로 MQTT, AMQP와 같은 발행/구독 모델을 적용한 OPC UA PubSub[3] 기술이 연구소와 대학 중심으로 선행 연구되어 왔고, 최근에는 PubSub 모델이 적용된 형태의 상용 디바이스가 시장에 출시되고 있다.

이에 본 논문에서는 제조현장의 OPC UA 표준을 지원하는 않는 레거시 통신 장치의 데이터를 수집하고 OPC UA 발행/구독 모델로 변환하여 확장성을 제공할 수 있는 테스트베드를 구성하고, OPC UA PubSub 구조모델의 특징을 분석 검증하였다.

## II. OPC UA 발행/구독 테스트베드 설계 및 구현

### 1) OPC UA 발행/구독 테스트베드 구조

OPC UA 발행/구독 테스트베드는 OPC UA Publisher와 OPC UA Subscriber로 구성되며, 기능적으로 OPC UA Publisher는 레거시 장치 연결을 위한 Modbus/TCP Function과 데이터 발행을 위한 OPC UA Publisher MetaDataSet Function의 구조를 갖는다. Modbus/TCP 통신 연결을 통해 레거시 장치 데이터를 수집하고, 이를 OPC UA Publisher MetaDataSet을 구성한다. MetaDataSet은 OPC UA Publisher가 발행하는 데이터 송신 프로토콜을 구조화하는 작업으로 OPC UA 오픈소스를 활용하였다.[4] 이와같이 변환

된 MetaDataSet을 기반으로 OPC UA 데이터를 주기적으로 발행한다. 그림 1은 OPC UA 발행/구독 테스트베드의 구조를 도시화한 모습이다.

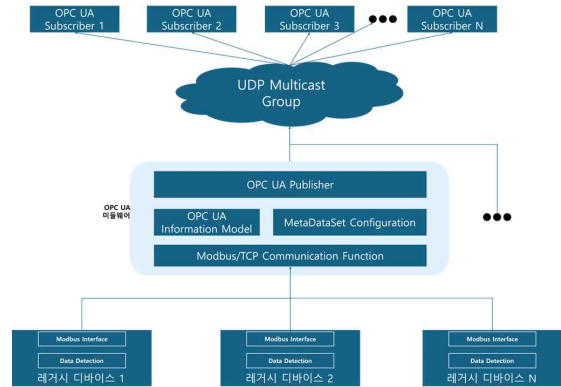


그림 1. OPC UA 발행/구독 테스트베드 설계 구조  
Fig. 1. OPC UA PubSub Testbed Design Structure

레거시 디바이스는 제조현장의 환경 모니터링을 위한 장치로 분산된 구조로 제조현장 환경을 모니터링 후 온도 데이터를 다양한 형태로 발생시킨다. OPC UA 미들웨어는 레거시 디바이스 데이터를 OPC UA Publisher로 변환해주는 역할을 수행하며 Modbus/TCP를 통해 데이터를 수집하고 OPC UA Publisher MetaDataSet을 구성한다. 이와같이 구성된 MetaDataSet은 UDP Multicast Group으로 OPC UA 구조로 데이터를 주기적 발행한다. OPC UA Subscriber는 OPC UA Publisher가 발행하는 동일한 UDP Multicast Group 내에 있는 다수의 장치로써 데이터를 주기적으로 구독하는 역할을 수행한다.

### 2) OPC UA 발행/구독 테스트베드 구축



그림 2. OPC UA 발행/구독 테스트베드 구성  
Fig. 2. OPC UA PubSub Testbed Configuration

그림 2는 OPC UA 발행/구독 기능 시험을 위한 테스트베드 구성의 모습으로 레거시 디바이스인 열화상 카메라가 3개 설치되어 특정 포인트의 최대, 최소, 평균 3개의 온도 데이터를 발생시킨다. OPC UA 미들웨어는 C언어 기반의 OPC UA 오픈소스인 open62541을 활용해서 개발하였으며 레거시 디바이스인 열화상카메라 대상 Modbus/TCP 연결을 통해 데이터를 수집한다. 수집 데이터를 활용하여 OPC UA Publisher MetaDataSet을 구성후 OPC UA Publisher 셋업을 하였다. 특히 MetaDataSet 구성은 발행 데이터를 구조화 및 체계화 하는 작업으로써 OPC UA 정보모델링 된 데이터 식별자 정보인 NodeId와 데이터 타입 정보를 활용하여 구성하였다. 구성된 MetaDataSet은 NetworkMessage의 Payload로 활용되어 주기적으로 UDP Multicast Group에 발행한다. 테스트베드는 OPC UA Publisher 15개로 구성하였다.

OPC UA Subscriber는 OPC UA Publisher가 발행하는 UDP Multicast Group과 동일한 구성을 하였으며, 주기적으로 발행한 NetworkMessage를 구독하여 필요한 데이터를 상위 애플리케이션에서 활용할 수 있도록 개발하였고, 테스트베드는 OPC UA Subscriber 15개로 구성하였다.

### 3) OPC UA 발행/구독 테스트베드 기능검증 결과

OPC UA 발행/구독 테스트베드에서의 OPC UA 미들웨어는 OPC UA PubSub을 지원하지 않는 제조 디바이스 대상 OPC UA PubSub 환경을 제공할 수 있음을 확인할 수 있다. 이를위해 OPC UA PubSub 기능을 결과는 다음과 같다.



그림 3. OPC UA Publisher 발행 결과  
Fig. 3. OPC UA Publisher result

그림 3은 레거시 디바이스인 열화상카메라의 데이터를 OPC UA 미들웨어가 수집 후 OPC UA Publisher를 구성한 결과이다. 열화상카메라가 발생하는 최대, 최소, 평균 온도 데이터를 이용하여 OPC UA 정보모델링 매핑 후 데이터 변화를 감지하기 위해 MonitoredItem 및 Subscription Function을 이용해 데이터를 콜백하였다. OPC UA 정보모델은 Folder Type의 FLIR Object 노드를 구성하여 데이터를 구조화하였으며, FLIR Object 노드 내에 열화상카메라 데이터를 모델링하여 데이터를 체계화하였다. 이와같이 OPC UA 정보모델링 된 결과는 그림 4와 같다.

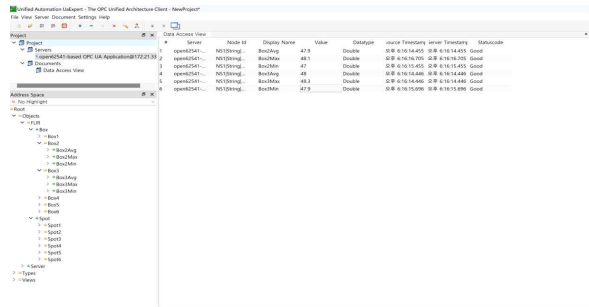


그림 4. OPC UA Publisher 정보모델  
Fig. 4. OPC UA Publisher Information Model

그림 4와 같이 정보모델링 된 OPC UA Publisher는 UDP Multicast Group에 주기적으로 데이터를 발행하고, 발행된 데이터는 동일한 UDP Multicast Group 내 있는 OPC UA Subscriber가 구독한다. 이와같이 OPC UA Subscriber가 구독한 데이터는 상위 애플리케이션 및 OPC UA Subscriber 간 활용 가능한 환경을 제공한다. 그림 5는 OPC UA Subscriber가 구독한 결과이다.

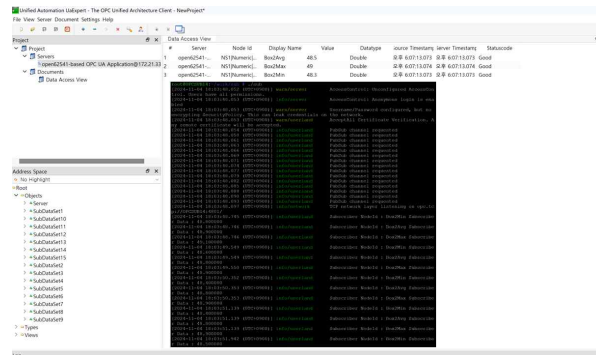


그림 5. OPC UA Subscriber 구독 결과  
Fig. 5. OPC UA Subscriber result

### III. 결 론

본 연구는 제조현장의 디바이스 to 디바이스, 디바이스 to 애플리케이션 간의 상호운용성, 확장성을 제공할 수 있는 표준 연구에서 시작했다. 특히 제조현장은 수많은 디바이스가 네트워크 통신으로 연동되고 있고, 제조데이터 연동에 필수적으로 요구되는 표준 기술로써 OPC UA가 현장적용 되는 상황에서 기존 제조 벤더사의 OPC UA 서버 구조가 갖는 잠재적 문제를 해결할 수 있는 방안을 제시하고자 하였다. 이에 개발된 테스트베드를 통해서 OPC UA PubSub에 대한 주요한 특징을 직접 확인할 수 있었으며, 비연결성 구조모델을 통한 제조 디바이스의 확장성 제공이 가능함을 검증하였다. 특히 IT 분야에서 다루던 PubSub 구조모델을 OT 분야에 접목시켜 많은 수의 제조 디바이스를 확장시킬 수 있는 방안을 도출해 낼 수 있었다는 점에서 매우 의미있는 연구였다.

향후 기 구현된 결과물을 실제조공정 설비를 연동시킬 수 있는 테스트베드로 확장해 나갈 계획이며, OPC UA Specification 전체를 포함하는 상용 SDK를 활용하여 고도화 시킬 예정이다. 또한 상위 애플리케이션으로 XR, VR 등의 디지털 콘텐츠 서비스를 위한 데이터 연동 기술로 활용하여 다양한 사례를 확보하고자 한다.

### 참 고 문 헌

- [1] OPC Foundation, <https://opcfoundation.org>
- [2] IDTA, <https://industrialdigitaltwin.org>
- [3] A. Hoday, C. Chrysoulas, T. Watt, M. Wollschlaege, M. de Sousa and L. Maglaras, "Broker-Less OPC UA PubSub Communication Model Performance Analysis," 2024 IEEE 30th International Conference on Telecommunications (ICT), pp. 1-5
- [4] open62541 Open Project, <https://open62541.org>

# 공정 최적화 모델 도출을 위한 프로세스 마이닝 시각화 도구 개발 및 검증 연구

이용호\*, 이원희\*\*

## Development and Validation of a Process Mining Visualization Tool for Deriving Process Optimization Models

Yongho Lee\*, and Wonhee Lee\*\*

### 요 약

본 연구에서는 공정 최적화를 위해 프로세스 마이닝 기법을 활용하여 도출된 모델과 데이터 분석 결과를 효과적으로 시각화할 수 있는 통합 도구를 개발하고 이를 실제 산업 데이터를 적용한 사례로 검증한 결과를 제시한다. 먼저, 이벤트 로그 데이터를 기반으로 한 프로세스 마이닝 모델의 자동으로 생성하는 기능을 구현하고, 생성된 모델의 이해도를 높이기 위해 다양한 시각화 기법을 적용하였다. 데이터 분석과 시각화를 위한 사용자 인터페이스를 개발하여 사용자가 복잡한 프로세스를 쉽게 탐색하고 분석할 수 있도록 도구를 개발하였다. 또한, 본 연구는 개발된 도구의 성능과 유효성을 평가하기 위해 실제 산업 데이터를 이용하여 검증하였다. 이를 통해, 공정 개선, 병목현상 식별, 성과 분석 등 분석 도구의 유용성과 실효성을 확인하였으며 그 결과로 기업 또는 조직내 프로세스 이해도 향상과 데이터 기반 의사결정의 고도화에 기여할 수 있을 것으로 기대된다.

### Abstract

In this study, we develop an integrated tool to effectively visualize the derived models and data analysis results using process mining techniques for process optimization, and present the results of validating the tool by applying real industrial data. First, we implemented the automatic generation of process mining models based on event log data and applied various visualization techniques to enhance the comprehensibility of the generated models. A user interface for data analysis and visualization was developed to enable users to easily explore and analyze complex processes. Furthermore, the study was validated using real industrial data to evaluate the performance and validity of the developed tool. Through this study, we confirmed the usefulness and effectiveness of the analysis tool for process improvement, bottleneck identification, and performance analysis, and as a result, it is expected to contribute to the improvement of process understanding and data-based decision-making in companies or organizations.

### Key words

Process mining, Data-driven, Manufacturing, Visualization, Implementation

\* 한국전자기술연구원 자율제조연구센터, yongho@keti.re.kr

\*\* 한국전자기술연구원 자율제조연구센터, zsed50@keti.re.kr (교신저자)

※ 이 논문은 2024년 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-01024, 대규모 CPPS 최적 운용을 위한 국제표준 기반 제조 디지털 트윈 관리 기술 개발)

## 1. 서 론

산업 현장에서의 공정 최적화는 기업 경쟁력 강화를 위한 핵심 과제 중 하나로, 특히 제조업 분야에서는 복잡한 프로세스의 효율성을 높이기 위해 데이터 기반의 접근법이 필수적이다[1]. 이를 위해 데이터 마이닝의 한 기법인 프로세스 마이닝이 활용되는데 이벤트 로그에서 활동 간의 적절한 관계를 찾아내기 위해 데이터 패턴을 구조적 규칙이나 알고리즘을 통해 프로세스 흐름을 자동으로 도출하는 기법이다[2]. 하지만 기존의 프로세스 마이닝 도구는 제조업의 복잡한 데이터 구조와 공정 특성을 충분히 반영하지 못하는 경우가 많다[3]. 이러한 한계를 해결하기 위해 본 연구에서는 제조업 특성에 맞춘 통합 도구를 개발하고자 하였다. 개발된 도구는 복잡한 프로세스를 쉽게 이해하고 분석할 수 있도록 프로세스 마이닝의 핵심기법인 프로세스 발견(Discovery)기법을 적용하였으며[4], 사용자 인터페이스를 통해 복잡한 공정을 직관적으로 탐색하고 성과 지표를 평가할 수 있도록 시각화 기능을 제공한다. 본 연구는 개발된 도구의 성능과 유효성을 검증하기 위해 실제 산업 데이터를 적용하여 기능을 검증하였다. 본 논문에서는 개발된 도구의 설계와 구현 과정을 설명하고 사례 연구를 통해 도구의 효과를 입증하며 이를 바탕으로 제조업 내 프로세스 최적화의 실질적 적용 가능성을 제시한다.

## II. 프로세스 마이닝 시각화 도구 설계 및 구조

본 논문에서 제안하는 프로세스 마이닝 시각화 도구는 웹 기반으로 클라이언트가 API 서버에 HTTP 요청을 통해 데이터를 업로드하고 프로세스 마이닝 분석을 통해 분석 결과를 조회할 수 있도록 한다. 웹 표준인 HTTP 프로토콜을 기반으로 설계된 API 방식인 Restful API로 클라이언트가 다양한 엔드포인트를 통해 기능에 접근할 수 있다. 이를 통해 클라이언트는 HTTP 메서드를 사용해 리소스를 생성, 조회, 수정, 삭제 하는 등의 작업을 수행한다. API 서버의 Node.js는 HTTP 요청 처리, 비즈니스 로직 실행, RServe 및 데이터베이스와의 통신을 담당한다. PM2

는 Node.js 애플리케이션의 안정적인 운영을 지원하는 프로세스 관리 도구로, 자동 재시작, 로드 밸런싱, 로그 관리, 실시간 모니터링 등의 기능을 제공한다. 이를 통해 TCP 프로토콜을 사용해 데이터 분석 서버(RServe)와 통신하여 프로세스 마이닝 알고리즘을 실행하고, 데이터베이스(MariaDB)와 통신하여 분석 결과를 저장하거나 조회한다. 그 구조는 <그림 1>과 같다.

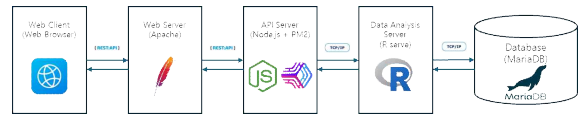


그림 1. 웹 서비스 구조

Fig. 1. Web service architecture

## III. 프로세스 마이닝 시각화 도구 구현 및 검증

### 1. 기능 구성

제조업 특화 도구 개발을 위해 프로세스 마이닝 기반 제조 공정 분석 프레임워크를 활용하여 기능을 개발하였다[5]. 도구의 기능은 크게 데이터 импорт(Data Import)와 프로세스 분석 및 데이터 시각화(Process Discover & Data Visualization) 두 가지 주요 영역으로 나뉘며, 각각의 영역에는 <그림 2>와 같이 세부 기능이 포함되어 있다.

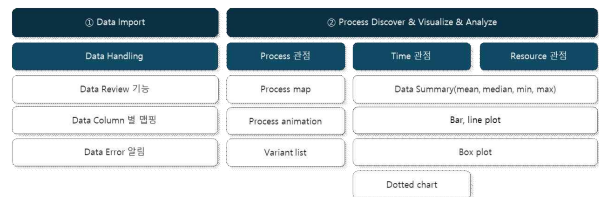


그림 2. 프로세스 마이닝 시각화 도구 기능

Fig. 2. Process mining visualization tool function

### 2. 기능 구현 및 검증

해당 도구의 데이터 импорт 단계는 프로세스 분석을 위한 데이터를 준비하고 품질을 확보하는 역할을 한다. Data Review는 데이터를 검토하고, 프로세스 마이닝 알고리즘에 적합한 데이터 형식과 조

건을 점검하여 분석에 적합한지 확인한다[6]. Data Column은 컬럼별로 데이터를 표시하여 필요한 전처리나 변환이 필요한 컬럼을 식별할 수 있게 한다. Data Error는 데이터에서 잘못된 형식이나 누락된 값을 찾아 수정하여 데이터 품질을 높인다. 프로세스 발견과 시각화 단계에서는 준비된 데이터를 바탕으로 프로세스 마이닝을 통한 프로세스 효율성 향상을 위해 <그림 2>에 따라 프로세스(Process) 분석, 시간(Time) 분석과 리소스(Resource) 세 가지 분석 축에 따라 Process Map, Process Animation, Variant List, Data Summary, Bar 및 Line Plot, Box Plot, Dotted Chart 기능을 사용하여 다양한 시각화와 분석을 제공한다. 검증에 사용된 데이터는 한국전자기술연구원의 스마트제조공정혁신센터의 가공·물류 공정의 Log Data를 활용하였다.

프로세스 분석에서는 Process Map을 사용하여 전체 프로세스의 단계와 활동 간의 연결을 시각적으로 표현함으로써 공정의 각 단계에서 소요되는 시간과 빈도를 한눈에 파악할 수 있다. <그림 3>에서와 같이, Process Map을 통해 프로세스의 구조와 흐름을 명확하게 이해할 수 있으며, 이를 통해 병목 현상이 발생하는 구간을 식별하고, 효율성을 높일 수 있는 지점을 찾아낼 수 있다.

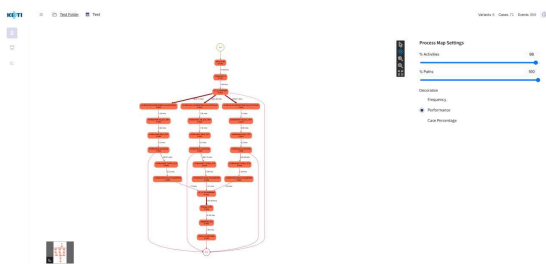


그림 3. 프로세스 맵 도출 기능 구현

Fig. 3. Process map discover function implementation

본 연구에서는 프로세스 Discovery를 위해 Inductive Miner 알고리즘을 사용하여 프로세스 맵을 생성하였다[7]. Inductive Miner는 트리 구조를 기반으로 계층적 프로세스 모델을 생성하여 병렬성과 루프 구조를 효과적으로 반영하는 데 강점이 있다. 이 알고리즘은 다른 알고리즘에 비해 계산 복잡도가 높아 모델을 구현하는 데 시간이 오래 걸릴 수 있으나,

구조화된 프로세스 처리가 가능하고 안정적이며 실용적인 모델을 도출할 수 있다는 장점을 가진다. 이를 기반으로 <그림 4>와 같이 Process Animation 기능을 구현하여, 데이터의 각 Case ID가 시간 흐름에 따라 프로세스 모델 내에서 이동하는 방식을 애니메이션으로 시각화하였다. 이 시각화는 프로세스의 각 단계 간 연결을 명확하게 보여주어 프로세스의 동적인 흐름을 이해하는 데 도움을 준다.

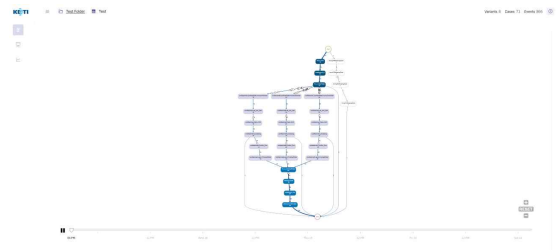


그림 4. 프로세스 맵 애니메이션 기능 구현

Fig. 4. Process map animation function implementation

또한, Variant List 기능인 <그림 5>는 프로세스 패턴을 파악하여 업무 효율화를 달성하고 특히 변형을 탐지하는 데 목적이 있다. Variant List는 다양한 프로세스 변형을 리스트 형태로 제공하여 동일한 프로세스 내에서 발생하는 변형을 비교하고 분석할 수 있게 한다. 이를 통해 상위 프로세스 패턴을 식별하고, 비정상적인 변형이나 예외 상황을 파악하여 프로세스의 표준화 및 최적화를 할 수 있다.

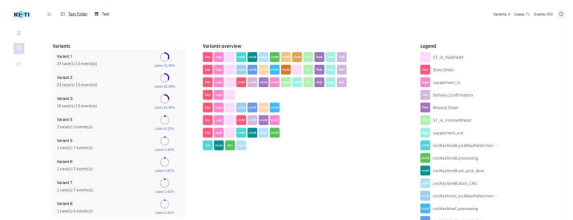


그림 5. 프로세스 변형 패턴 리스트 구현

Fig. 5. Process variant pattern list implementation

시간 분석은 작업 단계별 소요 시간과 대기 시간을 기반으로 프로세스의 비효율적인 요소를 식별하는 데 중점을 두며, 리소스 분석은 자원의 활용 빈도와 작업 시간의 변동성을 통해 리소스 배치의 최적화를 추구한다. 우선, Data Summary는 각 이벤트의 소요 시간과 대기 시간의 평균, 중앙값, 최솟값,



최대값을 제시함으로써 <그림 6>과 같이 전체적인 시간 통계를 제공한다. 이를 통해 작업 표준 시간과 비교하여 프로세스 시간을 줄일 수 있는 요소를 식별할 수 있다. 더 나아가, 시간이 긴 구간을 추적하여 원인을 파악하고, 리소스 조합에 따른 평균 리드 타임과 최적 리드 타임을 계산함으로써 공정 성능 향상에 기여할 수 있다. 리소스 분석에서는 동일한 방식으로 각 리소스의 평균 소요 시간과 대기 시간을 파악하고, 이 정보를 바탕으로 자원의 효율성을 검토한다.



그림 6. 데이터 요약 기능 구현  
Fig. 6. Data summary function implementation

Bar 및 Line Plot은 각 활동 단계에서 발생하는 이벤트의 횟수와 소요 시간을 시각화함으로써 특정 활동에서의 처리 빈도와 소요 시간의 분포를 <그림 7>과 같이 한눈에 파악할 수 있게 한다. 이를 통해 작업의 효율성을 비교 분석할 수 있다. 리소스 분석에서도 시간에 따른 리소스 사용 빈도를 Bar Plot으로, 소요 시간을 Line Plot으로 표현하여 각 리소스의 작업 효율성과 변동성을 파악할 수 있다. 특히, 이 분석은 리소스 간 작업 시간의 차이를 확인하고, 지연이 발생하는 패턴을 추적함으로써 프로세스 리드 타임을 감소시킬 수 있는 전략적 개선 방안을 제시하는 데 유용하다.

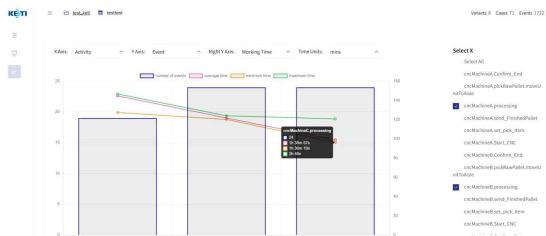


그림 7. Bar, Line plot chart 구현  
Fig. 7. Bar, Line plot chart implementation

<그림 8>은 이벤트 로그에서 이상치를 탐지하고 시간 및 리소스의 분포를 분석하는 데 사용되는 Box Plot으로 각 활동의 시간 분포와 이상치를 박스 플롯으로 나타냄으로써 이상 변화가 발생하는 지점을 식별할 수 있다. 리소스 분석에서는 리소스 사용량의 분포와 변동성을 시각화하여, 프로세스 애니메이션과 함께 활용하여 특정 리소스가 과부하되거나 병목 현상이 발생하는 구간을 탐지할 수 있다. 이로써 리소스 배치의 문제점을 파악하고, 기계적 결함이나 제조상의 문제를 확인하여 개선할 수 있는 근거를 제공한다.

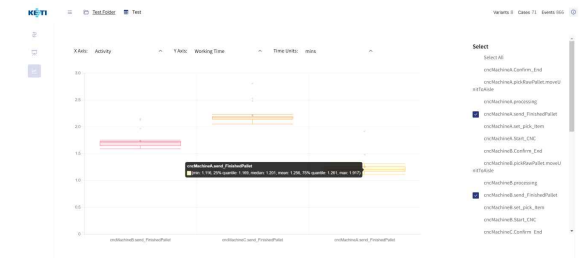


그림 8. Box plot chart 구현  
Fig. 8. Box plot chart implementation

마지막으로, <그림 9> Dotted Chart는 케이스별 이벤트의 시간 분포를 점 차트로 시각화하여 프로세스의 전체 흐름을 분석한다. 시간에 따른 이벤트 발생을 시각적으로 표현함으로써 프로세스 진행 경향을 예측할 수 있으며, 이를 통해 미래의 프로세스 흐름을 파악할 수 있다.

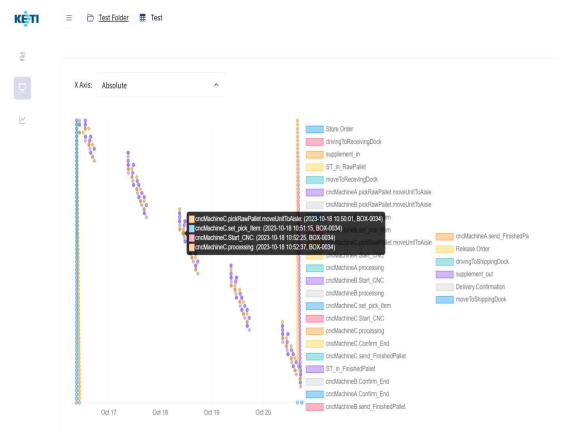


그림 9. Dotted chart 구현  
Fig. 9. Dotted chart implementation

#### IV. 결 론

본 연구는 제조업의 공정 최적화를 위한 데이터 기반 접근법으로서 프로세스 마이닝을 활용한 시각화 도구를 개발하고, 이를 통해 산업 현장에서의 적용을 위한 도구의 기능을 검증했다. 제조업의 특성을 반영한 제조 공정 분석 프레임워크를 기반으로 맞춤형 프로세스 마이닝 도구를 제안하였으며, 개발된 도구가 복잡한 공정을 이해하고 분석할 수 있도록 다양한 시각화 기능을 제공했다. 결론적으로, 본 연구는 프로세스 마이닝을 통한 제조업 공정 최적화를 실질적인 가능성을 확인했다. 본 도구가 제조업의 기업 내 프로세스 이해도 향상과 데이터 기반 의사결정 고도화에 기여할 수 있을 것으로 기대된다. 향후 연구에서는 기존 연구했던 프로세스 마이닝 기반 제조 공정 분석 프레임워크의 기능들을 모두 구현하고 각 분석방법을 이용하여 KPI를 도출해주는 기능까지 적용하고자 한다. 또한 실시간 분석 기능을 추가해 더욱 고도화된 분석 지원을 할 수 있을 것이다. 이를 통해 목적에 따른 분석기법을 사용하여 모델 생성 및 시각화, 프로세스 효율화, 요소 간 상관관계 분석 등 제조공정 개선이 가능할 것이다.

#### 참 고 문 헌

- [1] Lee, J., Lapira, E., Bagheri, B., & Kao, H-A, "Recent advances and trends in predictive manufacturing systems in big data environment.", *Manufacturing Letters*, 1(1), pp. 38-41, October, 2013
- [2] Van der Aalst, W. M. P., & Weijters, A. J. M. M. "Process mining: a research agenda", *Computers in Industry*, 53(3), pp. 231-244, April, 2004.
- [3] Reinkemeyer, L. "People: The Human Factor. In: Reinkemeyer, L.", (eds) *Process Mining in Action*, Springer, pp. 27-30, March, 2020.
- [4] Rudnitckaia, J., Venkatachalam, H. S., Essmann, R., Hruška, T., & Colombo, A. W., "Screening process mining and value stream techniques on

- industrial manufacturing processes: Process modelling and bottleneck analysis.", *IEEE Access*, 10, pp. 24203-24214, March, 2022.
- [5] Lee, YH., Lee, WH., "Analysis of research trends related to process mining analysis techniques.", *대한산업공학회 춘계공동학술대회 논문집*, pp. 600-602, May, 2023.
- [6] Van der Aalst, W. M. P., "Process Mining: Discovery, Conformance and Enhancement of Business Processes.", Springer, April, 2011.
- [7] Lugaresi, G., Ciappina, A. D., Rossi, M., & Matta, A., Exploiting a combined process mining approach to enhance the discovery and analysis of support processes in manufacturing. *International Journal of Computer Integrated Manufacturing*, 36(2), pp. 169-189. June, 2022.

# 정밀금속 가공 공구마모 측정시 작업자 시선 추적 데이터 분석

장성수\*, 송병훈\*\*

## A study on measuring tool wear analyzing with eye-tracking data in precision machining

Seongsu Jhang\*, and Byounghoon song\*\*

### 요 약

정밀금속 가공은 정밀성, 반복성이 요구되며 이는 품질, 생산성과 직결되며, 동일한 반복가공(양산)에서의 동일 품질 유지가 필수적이다. 이때, 반복 가공 공정에서 발생하는 공구마모 정도는 제품 품질에 영향을 주며 공구마모도 측정, 보정 작업을 통해 공구마모에 따른 절삭 품질 유지 가능하다. 공구마모 측정 시 측정 작업자의 측정 스킬(노하우)에서 정밀한 측정 공구의 제어와 제품의 측정 포지션(휴먼에러, 작업자 컨디션) 및 제품 표면 오염도 등에 따라서 측정치 보정(오프셋 보정)에서의 오차, 품질 저하의 주요 원인으로 본 논문에서는 아이트래커를 활용하여 숙련공의 노하우를 취득, 분석하였다.

### Abstract

The precision and repeatability required in precision machining are critical to product quality and productivity. Maintaining consistent quality in repetitive machining processes is essential. Tool wear is a significant factor affecting product quality in such processes. While measuring and compensating for tool wear can mitigate these effects, factors such as operator skill, measurement tool precision, and product positioning can introduce errors. This research aims to acquire and analyze the expertise of skilled workers using eye-tracking technology to identify and address these challenges, thereby improving the accuracy and consistency of tool wear measurement.

### Key words

eye-tracker, tacit knowledge, precision machining, tool wear measuring, offset calibration

## 1. 서 론

고정밀도가 필요한 제조공정인 자동차 부품 제조 공정에서 절삭가공은 주요한 공정 중 하나이다. 이 공정에서는 가공시 품질을 유지하기 위해 공구마모

관리가 필수적이며 양산 과정에서 공구 마모에 대한 보정, 가공 설비에서의 보상이 필요하다. 이를 측정하는 작업자의 숙련도, 컨디션에 따라 측정오차에 의한 오프셋 보정 오차 발생 및 휴먼에러로 인한 품질 저하가 발생하기도 하며 대량생산이 이뤄

\*한국전자기술연구원 ICT융합연구센터, jss0221@keti.re.kr, \*\*bhsong@keti.re.kr(교신저자표시)

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (No.P141200001, 빅데이터 활용 마이스터 로봇화 기반구축)

지는 동시에 동일 조건에서의 제품 측정, 측정 장치 사용 시 숙련공의 측정 스킬(노하우) 유지, 오프셋 보정 주기 관리가 품질 유지에 중요한 요소 중 하나이다[1, 2] 본 논문에서는 작업자 노하우 요소를 측정, 정량화하기 위해 아이트래커(착용자 시선 패턴 및 추적 데이터 취득 장치)를 사용하여 데이터를 취득, 분석하였다.

## II. 아이트래커 데이터 취득 및 분석

이 논문에서는 숙련공(공정 노하우 취득 작업자)의 시선분포 및 추적 데이터를 취득하기 위해 아이트래커를 사용하였다. 아이트래커는 작업자가 바라보는 시선 및 안구운동 정보를 취득할 수 있는 설비로 암묵지(노하우)를 취득하는데 사용되기도 한다 [3-5]. 하지만 정밀금속가공 업종에서 공구마모 측정을 위한 측정장비 사용, 오프셋 보정 등의 세부 작업에서의 활용 사례는 거의 없어 데이터 취득 및 분석에 대한 충분한 참고 자료가 없기 때문에 본 연구에서는 표준화된 작업 프로세서를 기반으로 작업 지시서를 작성, 절차에 따른 공정 데이터를 수집하였다. 1시간 단위로 가공 된 제품의 중요치수(전장)를 측정, 측정된 치수를 근거로 마모도를 판단, 공구마모보정을 통해 Offset 작업을 시행하는 절차에서 작업자에게 아이트래커 장비를 착용시켜 데이터를 취득, 수집했다.

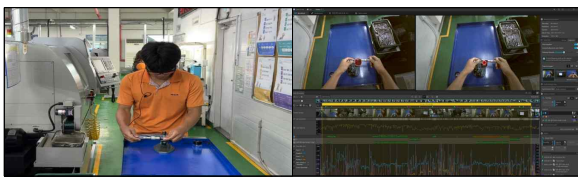


그림 1. 작업 및 아이트래커 분석 화면  
Fig. 1. Eye-tracker data acquisition

수집되는 데이터의 종류는 동공속도, 크기, 도약 안구 운동(Gyroscope, Accelerometer X,Y,Z Value), 시선의 위치 및 순서(Point X, Point Y, Point Z), AOI(Area Of Interest) 작업영역에서 작업자의 시선 분포도 등이 있으며 전처리를 통해 AOI를 기준으로 데이터의 패턴 분석을 수행하였다.

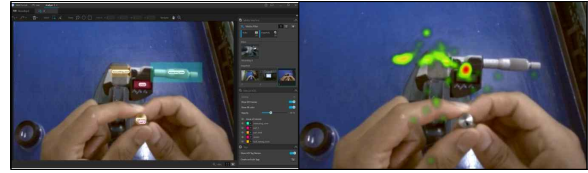


그림 2. AOI 히트맵 분석 결과  
Fig. 2. Eye-tracker data analysis

AOI는 크게 측정공구의 수동 조작부, 게이지, 소재로 구분하였으며 AOI별 total, average duration of fixation, pupil diameter을 포함한 20여 가지의 변수 데이터를 확인하였다. 시선분포별 상관관계를 파악하기 위해 피어슨 상관 계수를 적용하여 AOI 내 단속운동 변수, 시선 고정 시간 및 횡수 등에 대한 상관관계를 분석하여 공구마모 측정 공정에서의 암묵지 판단요소를 도출했다. 그림 3과 같이 시선 고정 시간, 횡수, 동공의 변화 세 변수에 대한 데이터의 상관계수가 높았으며 특히 Tool setting zone으로 설정한 AOI에서 빈도수가 두드러지게 나타났다. 전체 작업 프로세스(시계열) 중 핵심 공정 영역별 시선고정 시간을 기준으로 주요 관심영역별 공정 핵심요소를 파악할 수 있었으며, 분석 결과 공구마모 측정 작업시(Tool setting zone) 공구마모 측정 장비 조작부에 시선이 오래 분포했음을 확인했고 이는 메인 조작부와 미세 조절부의 정밀한 조정을 통해 숙련요소를 적용, 종속변수로는 측정물을 손으로 접촉하여 발생하는 열팽창 시간(약 20sec), 영점 조정, 작업자별로 발생할 수 있는 측정력이 있는 것으로 인터뷰 및 작업 데이터 분석을 통해 확인하였다. 대체적으로 주요 작업부위에서의 시선 고정 기간, 횡수는 작업 중요도와 비례할 것으로 추정되며 기타 신체적인 조건, 반사행동 등은 관련이 없음이 상관관계분석으로도 드러났다.

AOI 내 단속운동, 시선 고정 및 AOI 진입 횡수를 도출했으며 식(1)의 피어슨 상관계수를 적용했다.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 + (y_i - \bar{y})^2}} \quad (1)$$

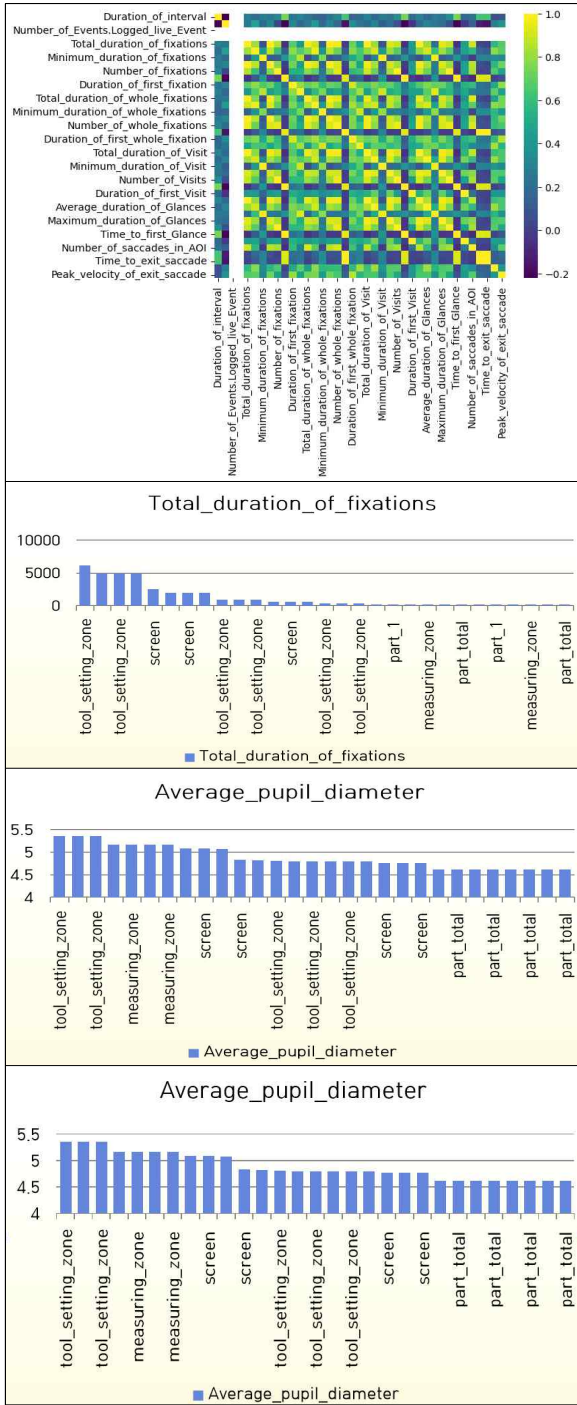


그림 3. 변수별 상관관계 분석 결과 및 AOI별 특성  
Fig. 3. Correlation of variables

### III. 결 론

본 논문에서는 숙련된 작업자의 정밀금속가공 공구마모 측정 및 오프셋 보정 공정에 대한 숙련요소 취득. 분석을 위해 아이트래커 장비를 사용하여 데이터를 취득했으며 작업 패턴별 작업자의 노하우를

디지털화하기 위한 변수 추출방법을 연구했다. 작업 시 주요 관심영역인 AOI에서 Tool setting zone(공구 미세조정 구역)에 대한 시선분포도가 가장 높았으며 이에 대한 데이터의 패턴에서 시선 고정 시간, 횡수, 동공 상태 세 가지 변수에 대한 상관관계수가 가장 높게 나타났다. 이를 통해 숙련도 판단 시 해당 변수를 활용할 수 있으며 공구마모 측정 장비에 대한 시선분포 영역 설정 시 소재 고정부와 인접한 수동 조절부에 도달하는 아이트래킹 데이터를 활용할 수 있음을 확인했다.

### 참 고 문 헌

- [1] 송철원, et al. 곡면가공에 있어서 공구마모예측 CAM 시스템 개발. 한국정밀공학회 학술발표대회 논문집, 2008, 37-38.
- [2] 황주효; 진교홍. 공구 교체 시점 기반 라벨링을 활용한 공구 수명 예측 모델. Journal of the Korea Institute of Information & Communication Engineering, 2024, 28.5.
- [3] WATANABE, Yuta, et al. Study on classification of the tacit knowledge using an eye-tracking interface: experiment of observation pictures and assessment of reproduction drawing. In: Knowledge-Based and Intelligent Information and Engineering Systems: 14th International Conference, KES 2010, Cardiff, UK, September 8-10, 2010, Proceedings, Part IV 14. Springer Berlin Heidelberg, 2010. p. 475-484.
- [4] 이학준, et al. 아이트래커 기반 절삭공정 숙련도 관련 데이터 표준화. In: Proceedings of KIIT Conference. 2022. p. 66-69.vol. 39, May 2000.
- [5] 장성수; 이진성. 아이트래커와 XAI 를 활용한 숙련공 암묵지 분석에 관한 연구. 대한산업공학회 춘계공동학술대회 논문집, 2023, 613-617.

# Text 기반 조색사 암묵지 학습을 위한 경량화 모델 연구

장성수\*, 유동휘\*, 권재영\*\*

## A Study on Models for Tacit Knowledge Learning of Colorists Based on Text Data

Seongsu Jhang\*, Donghui Yoo\*, and Jaeyoung Kwon\*\*

### 요 약

인공지능, 심층 신경망(Deep Neural Network) 기반의 OpenAI의 Chat-GPT, 구글의 BARD 등 거대언어모델(LLM) 성능이 검증되며 다양한 분야에서 서비스, 비즈니스 모델로 확산되고 있어 유스케이스와 함께 오픈소스가 공개되고 있다. 하지만 정형화된 데이터를 다루지 않는 사람의 감성, 감각에 의존하는 암묵지 요소가 필요한 공정에 범용적인 알고리즘, 모델들을 적용하기가 어렵다. 대부분의 중소, 중견기업은 작업자의 노하우, 감각에 의존하며 고객사에서도 표준화되지 않은 요구사항을 제시하는 경우가 많다. 본 논문에서는 비정형 데이터 및 감성 정보를 기반으로 조색, 염료 업종에서 정보를 추론하는 모델을 제안, 자연어처리에 주로 사용되는 LSTM과 비교하였다.

### Abstract

Based on AI service like OpenAI's ChatGPT, Google's Bard, and Microsoft's Copilot, large language models (LLMs) have been considered important in field of various industries and are being integrated into various services and business models. The transformation of natural language into machine-readable tokens and the processing of unstructured data have become increasingly common, with numerous open-source tools and use cases becoming available. However, applying generic algorithms and models to processes that rely heavily on human intuition and tacit knowledge, remains challenging. This paper presents a novel model that infers color formulation information from unstructured, subjective data. We conducted a study to validate, compare models using RGB benchmark dataset.

### Key words

tacit knowledge, deep learning, artificial neural network, tokenization, lstm, tenn

### 1. 서 론

최근, 대규모 데이터를 활용한 거대언어모델

(LLM) 기반의 서비스, 비즈니스 모델 등이 선보여지며 자연어처리 및 딥러닝 기반의 다양한 어플리케이션에 대한 여러 산업에 걸친 적용사례가 증가

\*한국전자기술연구원 ICT융합연구센터, jss0221@keti.re.kr, elight@keti.re.kr, \*\*jykwon@keti.re.kr(교신저자표시)

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (No.P0027635, 디지털전환 확산지원체계 구축 사업, 내역사업 : 산업 DX 지원 플랫폼)

하고 있으며 자연어 전처리 방법 중 하나인 토큰화를 통한 기계어로의 변환, 비정형 데이터의 처리가 증가하고 있다. 하지만 정형화된 데이터를 위주로 다루고 있는 산업계에서는 비정형 데이터를 활용한 사람의 감성, 감각에 의존하는 암묵지를 포함한 공정에 적용한 사례가 거의 없다. 특히, 도색, 사출 등 색상 정보가 필수적으로 반영되는 조색 업종, 공정에서 대기업을 제외한 대부분의 중소기업은 특정 색상의 필렛, 염료 등을 배합하는 과정에서 조색사의 노하우, 감각에 의존하며 고객사에서 표준화되지 않은 요구사항으로 비정형 색상, 감성정보를 지닌 발주를 하는 경우가 많다. 본 논문에서는 비정형 및 감성 정보를 기반으로 배합정보를 추론하는 모델을 제안하여 이를 검증하였다.

국내에서는 해외 제조사의(수입) 색차계와 안료에 대한 의존도가 높아 국산 원료에 대한 표준화, 데이터화의 진행이 미진한 실정이다. 이 때문에 색상 정보를 지닌 도료 등을 고객사 납품시 표준화된 고객의 요구 필렛(원료)의 정보가 부족하여 명도, 채도, 색상의 비정형적인 데이터(감성적인 말뭉치에 기반한 정보)가 제공되어, 이러한 추론 모델에 있어서 그 성능, 활용도가 점차 증가하고 있다.

언어모델의 성능과 학습 데이터를 고려한 색상별 수치, 안료 등의 배합 정보와 정형화된 데이터(본 논문에서는 RGB color code chart 정보 활용)를 학습시켜 유의미한 추론 모델을 개발할 수 있다.

## II. 학습 데이터셋 선정 및 선행연구

학습을 위한 데이터의 경우, 선행 연구에서는 - CIEL\*a\*b\* 표색계, CIELAB 색공간(표준화) 정보 기반의 기업 정보를 활용[1-3]하였으나 본 연구에서는 모델의 범용적 특성을 파악하기 위해 RGB chart와 해당 데이터에서 보유한 감성 정보, 비정형 색상 이름표를 활용하였다. 또한, 기업별 보유 중인 안료, 도료 등의 배합 정보 등의 비정형 데이터의 규모, 분포가 다르며 기업마다 그 편차가 크기 때문에 RGB Color Codes Chart를 색상에 대한 배합 정보의 샘플 데이터셋 400여 종으로 활용하였다. RGB 색상 정보에도 light, dark, pale, deep, medium 등 주관

적인 요소가 포함된 여러 색상과 매칭될 수 있는 감성적인 말뭉치의 학습 가능 여부 등을 확인할 수 있도록 데이터셋을 구성했다. 색상정보별 RGB 코드 데이터(3dimension의 0~255 사이 숫자로 구성)를 활용하였다. 학습 결과는 CIELAB 색공간(표준화) 정보 기반의 라벨링과 추론 결과를 비교하여 모델의 정확도를 도출하여 비교가 가능하며 LCH 색공간 데이터를 사용(LCH Format) 과 LAB format 데이터를 사용하여 성능 분석이 가능하다. 인공지능경망을 활용한 레시피 추론 모델에서는 가장 높은 성능의 데이터 포맷에서도 값이 1보다 작은 결과치가 60%에 불과했다[4].

## III. 학습 모델

본 논문에서 학습모델은 감성적인 데이터를 포함한 것을 상정하여 Text 형태의 비정형 데이터를 적용할 수 있도록 LSTM 기반 모델과 토큰화 전처리 기능을 가진 심층 신경망 기반 모델 두 가지를 적용, 검증하였다. 두 모델 모두 인공지능경망 기반으로 데이터의 수가 많지 않아 적은 수의 파라미터를 가진 소규모 모델을 선정하였으며 해당 모델에서도 일정 수준의 성능을 확인했다[5]. 제안한 모델의 구조는 그림 1. 과 같으며 각각 모델의 파라미터 수는 LSTM은 1,459개, 제안한 TeNN(Tokenizing encoder Neural Network)는 923개로 구성하여 학습 및 비교를 수행하였다.

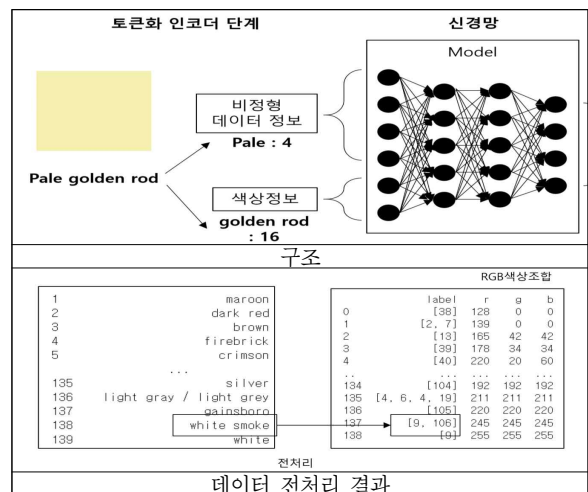


그림 1. 토큰화 기반 신경망 모델 구조  
Fig. 1. Model structure based on tokenizaing layer

학습결과는 그림 2. 의 loss, accuracy를 비교하였을 때 LSTM에 비하여 TeNN에서 더 좋은 성능이 확인되었다.

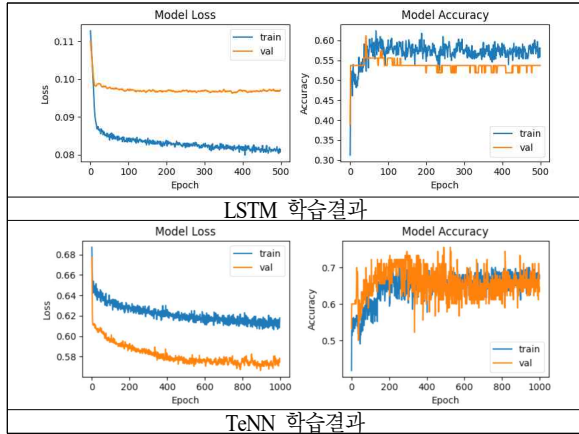


그림 2. 모델별 학습 결과 비교  
Fig. 2. Learning results

추론 결과, CIELAB 기반 Delta E 값을 산출하여 원색상과의 차이를 비교했다. 30개 샘플데이터 기준, 약 73.8의 Delta E 값이 산출되었다.

$$\Delta E = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2} \quad (1)$$



그림 3. 샘플데이터 추론 결과  
Fig. 3. Sample data inference results

#### IV. 결 론

본 논문에서는 조색 공정에서 숙련된 조색사의 노하우를 학습시키기 위한 비정형 데이터 기반 조색 배합 비율 추론 모델을 제안하였고 RGB color chart 데이터를 활용하여 비정형 감성 언어를 학습시킬 수 있는 LSTM, TeNN 두 모델을 비교했다. 학습 성능은 TeNN이 우수하였지만 데이터셋, 모델 규모 등을 고려하였을 때 개선이 필요하다. 향후, 실 데이터를 기반으로 모델 개선을 수행할 예정이다.

#### 참 고 문 헌

- [1] 황규석; 박창원. 목표색상 재현을 위한 페인트 안료 배합비율의 예측. 한국유화학회지, 2008, 25.4: 438-445.
- [2] SPEHL, J., et al. Application of backpropagation nets for color recipe prediction as a nonlinear approximation problem. In: Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94). IEEE, 1994. p. 3336-3341.
- [3] BISHOP, J. M., et al. Genetic optimisation of neural network architectures for colour recipe prediction. In: Artificial Neural Nets and Genetic Algorithms: Proceedings of the International Conference in Innsbruck, Austria, 1993. Springer Vienna, 1993. p. 719-725.
- [4] BISHOP, J. M.; BUSHNELL, M. J.; WESTLAND, Stephen. Application of neural networks to computer recipe prediction. Color Research & Application, 1991, 16.1: 3-9.
- [5] JHANG, Seongsu; YOO, Donghwi; KOWN, Jaeyong. Research on color recipe recommendation based on unstructured data using TENN. arXiv preprint arXiv:2408.09094, 2024.



# 머신러닝 기반 절삭공구 마모량 추정

유동휘\*, 남종필\*, 이진성\*\*

## Machine learning-based estimation of cutting tool wear

D.H Yoo\*, J.P Nam\*, and J.S. Lee\*\*

### 요약

제조 현장에서 절삭공구의 길이를 측정하기 위해서 가공 공정이 완료된 후 별도의 측정 장비를 통해 마모량을 측정하는 과정이 필수적이다. 이는 공정 시간을 증가시키고 생산성이 저하된다는 문제가 있다. 최근 설비 모니터링 기술의 발전으로 실시간 센서를 활용하여 전류, 가속도 등의 데이터 취득을 통해 설비의 상태를 감시하는 연구가 이루어지고 있다. 본 연구에서는 실시간 설비 데이터를 활용하여 절삭공구의 마모량을 실시간으로 추정하여 기존의 측정 과정을 생략함으로써 생산성을 향상시킨다.

### Abstract

In manufacturing environments, the measurement of cutting tool wear requires the use of dedicated measurement equipment after the machining process is complete. This post-process measurement increases cycle time and negatively impacts productivity. Recent advancements in condition monitoring system have enabled real-time condition monitoring of machinery by gathering data such as current and acceleration through sensors. In this study, we propose a cutting tool wear estimating method in real-time by utilizing machine data, thereby eliminating the need for conventional post-process measurement and significantly enhancing productivity.

### Key words

Machine tools, milling, wear monitoring artificial intelligence, big data

## 1. 서론

최근 빅데이터 및 인공지능 기술이 발전함에 따라 실시간 데이터를 활용하여 설비의 이상을 감지하거나 상태를 진단하는 기술에 대한 관심이 높아지고 있다. 특히, 가공 공정에서 절삭공구의 상태를

파악하는 것은 제품의 품질과 생산성을 유지하는데 중요하다. 그러나 기존의 공구 마모 측정 방법은 주로 공정 후 별도의 측정 장비를 사용해야 하는 번거로움이 있으며, 이는 공정 시간을 증가시키고 불필요한 비용을 초래한다. 이러한 한계를 극복하기 위해 최근에는 측정 장비 없이 실시간 설비 데이터

\* 한국전자기술연구원, elight@keti.re.kr

\*\* 한국전자기술연구원, jinsl127@keti.re.kr(교신저자)

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (RS-2023-00259643, 전자(가전) 산업의 벨류체인 고도화를 위한 지능형 협업공장 구축 기술 개발)

를 활용해 공구의 마모를 추정하는 연구들이 활발히 진행되고 있다[1,2]. 가공 중 발생하는 전류, 진동 등 다양한 센서 데이터를 바탕으로 공구의 마모를 실시간으로 추정하는 기술은 측정 과정을 생략할 수 있어 공정 시간을 단축하고, 교체 시기를 예측하여 생산성을 향상시킬 수 있다[3].

## II. 머신러닝 기반 절삭공구 마모량 추정

본 연구에서는 엔드밀 절삭공구의 마모량을 추정하기 위해 반복 가공 실험을 진행하였다. 가공 공정 중 설비 데이터, 센서 데이터 및 절삭공구 마모량을 취득하여 인공지능 모델 학습 데이터셋을 구축하였고 머신러닝을 활용하여 절삭공구의 마모량을 추정하는 모델을 학습하였다. 추출한 설비 데이터 중 분석에 활용한 스핀들 축 진동, 스핀들 서보 드라이브 전류 데이터는 Fig. 1과 같다. Mean, variance, skewness, kurtosis 등 11개의 통계적 특성인자를 활용하여 시계열 데이터의 가공 정보를 특성인자로 요약하였다. 추출한 특성인자 값과 절삭공구의 길이, 직경 값을 Randomforest regressor, XGBoost regressor, Bayesian ridge regressor로 학습하였다. 각 모델의 결과를 앙상블하여 절삭공구의 길이, 직경 값을 예측하였다. 학습 결과, Fig. 2와 같이 실시간 설비 데이터를 활용하여 측정 장비의 개입 없이 절삭공구의 마모량을 추정할 수 있다.

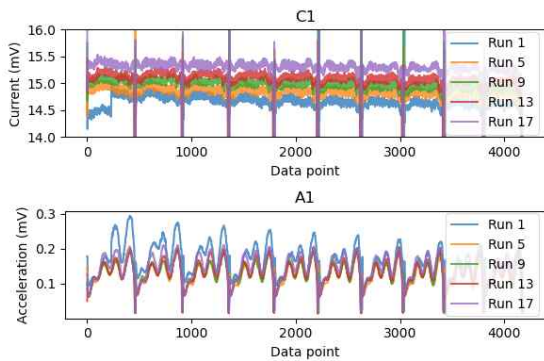


Fig. 1. (Top) Current data; (bottom) acceleration data during the machining process

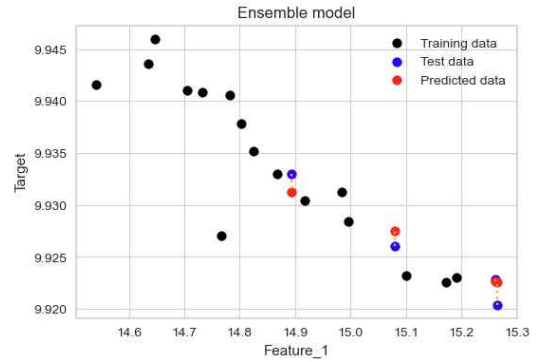


Fig. 2. Prediction results

## III. 결 론

본 연구에서는 머신러닝을 활용하여 절삭공구의 마모량을 실시간으로 추정하는 모델을 개발하였다. 시편의 반복 가공 실험을 통해 취득한 가공 중 설비의 진동, 전류 데이터를 활용하여 머신러닝 기반 인공지능 모델을 학습하였다. 실험 결과, 인공지능 모델을 통해 실시간 설비 데이터를 활용하여 절삭공구의 마모량을 추정할 수 있음을 보여준다. 추후 다양한 가공 조건 및 형상이 주어졌을 때 절삭공구의 마모량을 추정하기 위한 실험이 진행될 필요가 있다.

## 참 고 문 헌

- [1] M. Wang, J. Zhou, J. Gao, Z. Li and E. Li, "Milling tool wear prediction method based on deep learning under variable working conditions," IEEE Access, vol. 8, pp. 140726-140735, 2020
- [2] Y. Chen, Y. Jin and G. Jiri, "Predicting tool wear with multi-sensor data using deep belief networks", Int J. Adv. Manuf. Techno. vol. 99, pp.1917-1926, August 2018.
- [3] X. Xu, J. Wang, B. Zhong, W. Ming and M. Chen, "Deep learning-based tool wear prediction and its application for machining process using multi-scale feature fusion and channel attention mechanism", Measurement, vol. 177, p.109254, June 2021.

# 가공 데이터를 활용한 최적 절삭 조건 추정 시스템

유동휘\*, 권재영\*\*

## Data-driven estimation system for optimal cutting condition in machining processes

D.H. Yoo\*, and J.Y. Kwon\*\*

### 요약

전통적인 가공 공정에서는 사용자 경험과 실험적 방법에 의존해 피드 속도, 스핀들 속도 등의 절삭 조건을 설정하였다. 본 연구에서는 실시간 장비 데이터를 활용하여 최적의 절삭 조건을 자동으로 도출하기 위한 데이터 기반 최적 절삭 조건 추천 시스템을 제안한다.

### Abstract

In traditional machining processes, cutting conditions such as feed rate and spindle speed were typically set based on user experience and experimental methods. In this study, we propose a data-driven system that leverages real-time machine data to automatically determine the optimal cutting conditions. This system aims to provide optimized cutting recommendations to improve the efficiency and precision of machining processes.

### Key words

Data-driven machining, optimal cutting condition, sensor fusion

## 1. 서론

전통적인 가공 공정에서는 절삭 조건(피드 속도, 스핀들 속도, 절삭 깊이 등)을 주로 작업자의 경험과 실험적 방법에 의존하여 설정해 왔다. 이러한 접근은 오랜 시간 축적된 경험적 지식과 가공 핸드북 정보를 바탕으로 일정한 수준의 품질과 신뢰성을 제공하였으나 가공 환경, 공작물 재질, 공구 상태 등의 변수에 따라 공정의 성능이 달라질 수 있어 일관성을 유지하기 어렵다는 한계를 가지고 있다.

비효율적으로 설정된 절삭 조건은 가공 공정에서 생산성과 품질을 저하시키며, 공구 마모의 가속화, 공정 시간 증가, 에너지 소비 증가 등 다양한 문제를 초래할 수 있다. 이러한 문제를 해결하기 위해 최근에는 데이터 기반의 절삭 조건 최적화 연구가 활발히 진행되고 있다[1,2]. 데이터 기반 접근은 다양한 설비데이터 및 측정 장치를 통해 공정에서 취득되는 대규모 데이터를 바탕으로 최적의 절삭 조건을 추정할 수 있다. 이러한 접근은 기존의 경험적 방법에서 발생하는 주관적 판단을 최소화하여 공정

\*한국전자기술연구원, elight@keti.re.kr

\*\*한국전자기술연구원, jykwon@keti.re.kr(교신저자)

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (No.P0027635, 디지털전환 확산지원체계 구축 사업, 내역사업 : 산업 DX 지원 플랫폼)

의 신뢰도를 향상시키고, 절삭 속도, 피드 속도 등 다수의 공정 요인들이 서로 상호작용하며 발생하는 복잡하고 비선형적인 관계를 학습할 수 있다.

## II. 최적 절삭 조건 추정 시스템

본 연구에서는 최적 절삭 조건을 추정하기 위한 시스템을 제안한다. 최적 절삭 조건을 탐색하기 위한 시스템 순서도는 Fig. 1과 같다. 우선 유저가 원하는 작동 조건의 범위와 가공 시험 시편을 결정한다. 그 후 머시닝 센터에서 시편 가공을 수행한 후 데이터를 취득한다. 취득 데이터는 크게 네 가지 항목으로 구성된다. 작동 조건은 절입량, 피드 속도, 스핀들 속도 등 가공 조건에 관한 데이터이다. 머시닝 데이터는 가공 중 발생하는 데이터에 대한 정보로 가공 시간, NC (Numerical control) 코드, 스핀들 위치, 톨 카운트 등 설비에서 발생하는 데이터이다. 센서 데이터는 추가적으로 부착한 전류, 진동 센서를 통해 취득되는 공정 데이터이다. 품질 데이터는 가공이 끝난 후 표면 거칠기, 공차, 편차 등 결과물의 측정값에 관한 데이터이다.

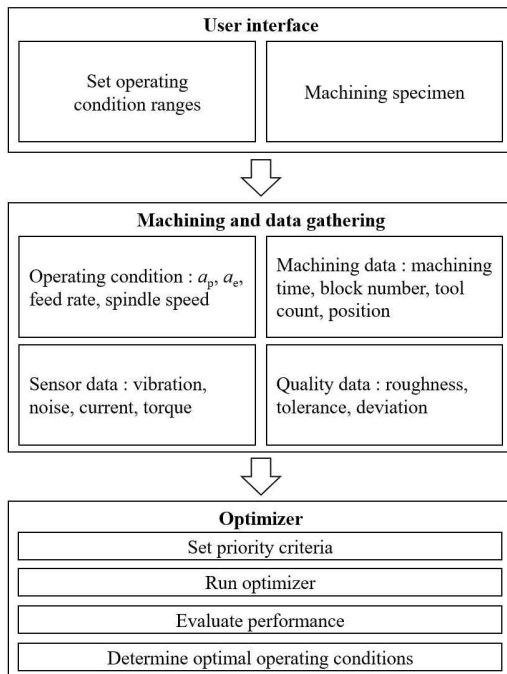


Fig. 1. Data-driven estimation process for optimal cutting condition

측정된 결과는 저장된 후 최적화 알고리즘에 전달한다. 계산에 앞서 공정의 우선순위 중요도를 설정한다. 가공 상황에 맞도록 가공시간, 정밀도 등 여러 항목에 대한 우선순위 및 최소 조건을 설정한다. 데이터의 특성을 고려하여 다항회귀, 머신러닝, 딥러닝 등의 알고리즘을 선택한 후 계산을 수행한다. 계산한 결과를 바탕으로 각 가공 조건에서의 성능을 평가하고 최적의 가공 조건을 결정하여 출력하게 된다. 다항회귀를 활용하여 공구 마모도에 대한 스핀들 속도, 피드 속도의 최적화 결과 예시는 Fig. 2와 같다.

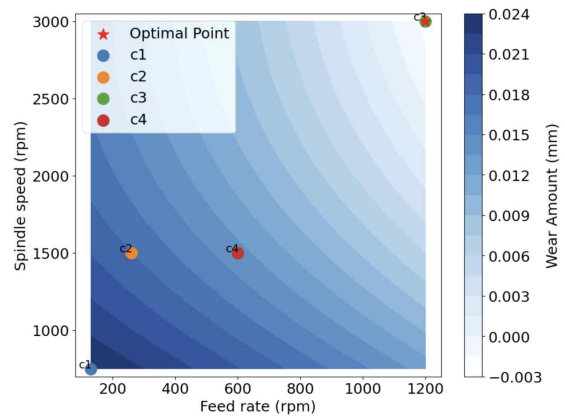


Fig. 2. Optimization results for feed rate and spindle speed on tool wear

## III. 결 론

본 연구는 공정 데이터를 기반으로 절삭 공정의 최적 절삭 조건을 추정하는 시스템을 제안하였다. 제안한 방법은 기존의 작업자 경험 의존 방식에 비해 데이터를 기반으로 객관적이고 재현 가능한 절삭 조건을 추정할 수 있다. 향후 연구에서는 최적 공정을 추정하기 위한 다양한 재질 및 형상의 시편을 체계화하고 비선형성을 예측하기 위한 딥러닝 알고리즘을 적용할 필요가 있다.

## 참 고 문 헌

- [1] I.N. Tansel, B. Ozcelik, W.Y. Bao, P. Chen, D. Rincon, S.Y. Yang, A. Yenilmez, "Selection of optimal cutting conditions by using GONNS",

International Journal of Machine Tools and  
Manufacture, Volume 46, Issue 1, pp. 26-35,  
2006.

- [2] C. Prakasvudhisarn, S. Kunnapapdeelert & P.  
Yenradee, "Optimal cutting condition determination  
for desired surface roughness in end milling",  
International Journal of Advanced Manufacturing  
Technology Volume 41, pp. 440-451, 2009.

# 정밀 금속가공 CNC의 지능형 애플리케이션 연계를 위한 통합서버 시스템 연구

권재영\*, 이원희\*\*

## A Research on Integrated Server System for Intelligent Application Linkage of Precision Metal Processing CNC

Jaeyoung Kwon\* and Wonhee Lee\*\*

### 요 약

현대 금속 가공 산업의 생산성 향상을 위한 노력은 하드웨어와 소프트웨어가 결합된 시스템 패키지의 도입 등의 형태로 가속화되고 있다. 이는 기존의 장비 중심적인 공급 방식에서 벗어나, 지능형 및 자율형 애플리케이션을 통해 생산 효율성을 극대화하려는 움직임으로 해석된다. 이러한 변화에 따라, 정보통신 기술(ICT) 기반의 장비 운영 최적화 소프트웨어를 탑재한 시스템 패키지 시장이 급격히 성장하고 있다. 또한, 제조 환경의 변화로 인해 다품종 소량 생산 시스템이 확산되면서, 다양한 장비의 유연한 운영 및 관리를 위한 통합 서버 솔루션의 중요성이 더욱 부각되고 있다. 본 연구에서는 이러한 산업 동향을 반영하여, 서로 다른 프로토콜을 사용하는 이기종 공작기계 CNC 및 외부 센서 데이터를 수집하고 제어하며, 애플리케이션과 제조 데이터를 연동할 수 있는 통합서버 시스템을 설계하였다.

### Abstract

Efforts to enhance productivity in the modern metalworking industry are accelerating in the form of system packages that integrate hardware and software. This shift marks a departure from traditional equipment-centered supply methods, aiming to maximize production efficiency through intelligent and autonomous applications. Consequently, the market for system packages equipped with ICT-based equipment optimization software is rapidly expanding. Additionally, as multi-product, low-volume production systems become more prevalent in changing manufacturing environments, the importance of integrated server solutions for the flexible operation and management of diverse equipment has become more pronounced. This study reflects these industry trends by designing an integrated server system capable of collecting and controlling data from heterogeneous CNC machine tools and external sensors using various protocols, enabling connectivity between applications and manufacturing data.

### Key words

heterogeneous CNC machine tools, Various industrial protocols, intelligent application, integrated server

---

\*한국전자기술연구원 ICT융합연구센터, jykwon@keti.re.kr, \*\*자율제조연구센터, zsed50@keti.re.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-00962, 제조 현장 맞춤형 Industrial XR 프레임 워크 개발)

## 1. 서론

다양한 제조사의 CNC 시스템을 사용하는 공작기계 환경에서의 통합서버 설계는 몇 가지 중요한 고려 사항을 수반한다. 첫째, CNC 제어 시스템의 표준화된 통신 프로토콜 적용이 중요하다. 과거에는 각 CNC 제조사별 고유 프로토콜에 의존하는 경향이 강하여 데이터 공유 및 통합에 제약이 많았다. 둘째, 통합 서버 시스템은 다양한 시스템에 대한 호환성이 확보되어야 한다. 시장 점유율이 높은 FANUC, SIEMENS, HEIDENHAIN 등 주요 CNC 제조사와의 호환성을 보장하고, 산업계에서 가장 보편적으로 활용되는 전류 센서 및 진동 센서 등 외부 센서 데이터를 통합적으로 처리할 수 있는 로직을 설계해야 한다. 셋째, 애플리케이션과 연동성을 고려해야 한다. IT 시스템과 OT 시스템 간 개발 환경 및 개발 도구의 차이로 인해 시스템 연계 및 추가 개발에 어려움을 겪는 경우가 많다. 본 연구에서는 이러한 문제점을 해결하기 위해, 서로 다른 CNC 시스템을 사용하는 절삭가공 환경에서 표준 프로토콜 기반의 통합서버 시스템을 설계한다.

## 11. 선행연구 및 상용 소프트웨어 분석

다양한 프로토콜을 기반으로 공작기계 CNC의 실시간 데이터를 수집하기 위한 레거시 서버 인터페이스에 대한 연구가 진행되어 왔으나[1], 주로 Windows 운영체제와 네트워크에 제약이 있는 OPC Classic에 중점을 두고 있다. 또한, 표준 통신 인터페이스 중 하나인 MTConnect 기반의 데이터 수집 시스템을 구현하는 연구가 있었지만[2-3], 장비 제어 및 피드백 제어에 한계가 존재한다.

모든 공작기계 CNC 제어기가 국제 표준 IEC 62541 OPC UA 통신 프로토콜을 지원하는 것은 아니기 때문에, 산업 현장에서 이기종 CNC 시스템 간의 연동에 현실적인 제약이 발생하고 있다. 따라서 CNC에 종속적인 데이터 수집 시스템 구축의 비효율성을 개선하여 일관된 방식으로 데이터 통신을 하고자 한다.

또한, 산업계에서 보편적으로 사용되는 상용 소프트웨어와 표준 정보모델에 대해 분석을 수행하였다.

### 1) MT-LINKi

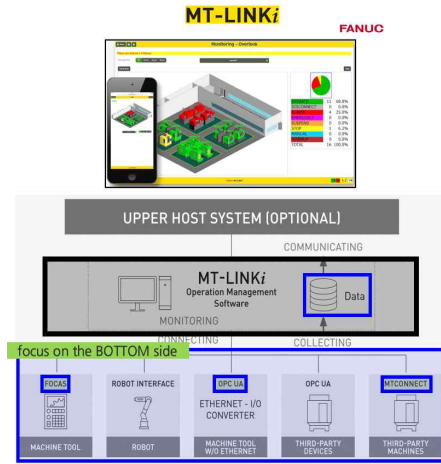


그림 1. MT-LINK i 개념도  
Fig. 1. MT-LINKi Conceptual Diagram

MT-LINK i는 FANUC에서 개발한 상용 소프트웨어로 FOCAS, OPC UA, MTConnect 등 다양한 프로토콜을 통해 CNC 및 로봇으로부터 제조 데이터를 수집하고 시각화하는 기능을 제공한다. 하지만 OT 영역에 집중되어 있어 이기종 기계 설비 데이터 수집에 주로 활용된다. 애플리케이션과의 데이터 교환에는 MongoDB를 사용하기 때문에 실시간 데이터통신이 중요한 서비스 애플리케이션과의 연동 시에는 성능 제약이 발생할 수 있다.

### 2) UMATI

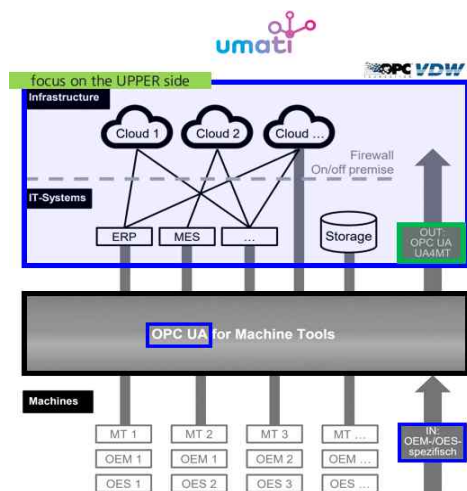


그림 2. UMATI 개념도  
Fig. 2. UMATI Conceptual Diagram

OPC Foundation과 VDW에서 제시하는 표준 정보 모델은 산업용 장비 간의 상호운용성을 확보하기 위한 중요한 기반을 마련한다. 이 모델은 장비마다 상이한 노드 주소 및 값에 대한 표준화된 정보 교환을 가능하게 하여 데이터 통합 및 시스템 연동에 기여한다. 그러나 Shop Floor Level의 이기종 설비 연결성을 고려하지 않아, 실제 산업 현장에서는 Kepware 또는 BFC Gateway와 같은 추가적인 변환 미들웨어가 필요하다는 단점이 있다. 이는 시스템 구축 비용 및 복잡성을 증가시키는 요인으로 작용할 수 있다.

### 3) 통합서버 설계컨셉

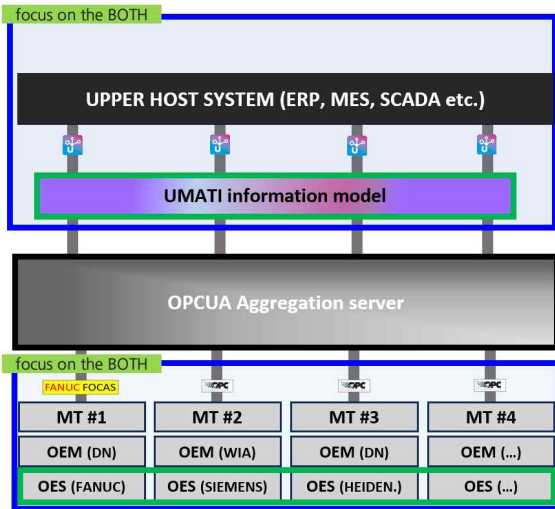


그림 3. 통합서버 개발컨셉도  
Fig. 3. Aggregation Server Architecture Diagram

본 통합서버 설계는 기존 상용 소프트웨어와 정보 모델의 한계를 보완하고 OT와 IT 측면을 모두 고려하여 설계하였다. 특히, 서로 다른 통신 프로토콜을 사용하는 이기종 CNC 환경을 통합적으로 관리하기 위해 OPC UA를 지원하지 않는 FANUC CNC를 포함하여 SIEMENS, HEIDENHAIN 등 다양한 제조사의 CNC 장비를 위한 OPC UA 통합 클라이언트를 탑재하였다. 이를 통해 각 CNC 시스템으로부터 OPC UA 기반으로 데이터를 일괄 수집하고, 표준화된 포맷으로 변환하여 제공함으로써 상호운용성을 확보하였다. 뿐만 아니라, VR, AR, XR 등 미래 지향적인 애플리케이션과의 연동을 고려하여 JSON 형식의 데이터 제공 기능을 통한 확장성을 고려하였다.

### III. 이기종 CNC 및 멀티센서 통합서버의 설계

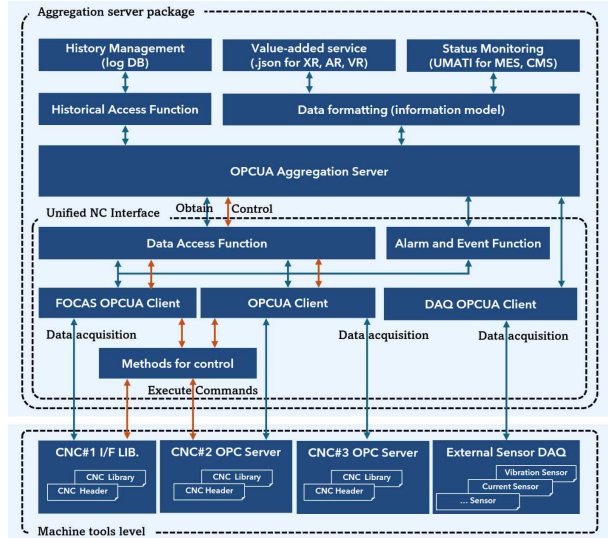


그림 4. 통합서버 시스템 구성도  
Fig. 4. Aggregation Server Configuration Diagram

본 연구에서 제안하는 Aggregation Server 시스템의 구성은 다음과 같다. 첫째, 기계설비들의 휘발성 실시간 데이터 이력관리를 위한 DB기반 History management. 둘째, 제조 IT를 넘어서는 AR, VR, XR 등 Head Mount Display를 통한 다양한 Add-On 서비스와 연동을 위한 json formatting. 셋째, MES, SCADA 등의 표준 정보모델 기반의 연결성 확보를 위한 UMATI formatting. 넷째, FANUC CNC데이터를 수집하기 위한 NC communication API인 FOCAS OPCUA Client. 다섯째, OPCUA 서버가 있는 CNC 데이터 수집을 위한 OPCUA Client. 마지막으로, 진동센서 및 전류센서 등 별치형 센서 데이터 수집을 위한 DAQ OPCUA Client로 설계하였다.

그림5의 UML 다이어그램은 FANUC, SIEMENS, HEIDENHAIN 세 가지 CNC 시스템과 별치형 센서를 대상으로 하는 통합 서버의 설계안을 보여준다. 통합서버는 각 CNC 시스템에 대한 상태 모니터링 및 애플리케이션과의 표준포맷 형태의 데이터 통신을 목적으로 하며, 각기 다른 통신 프로토콜과 데이터 형식을 통합하여 중앙에서 관리할 수 있도록 설계되었다.

다이어그램 상단의 Aggregation Server는 전체 시스템의 중앙 관리 역할을 하며, CNC 시스템별로 OPC UA 클라이언트를 통해 통신한다. 또한, 각기



다른 CNC 시스템의 프로토콜 지원 여부를 확인하고, 이를 기반으로 연결 상태를 관리하는 방식을 보여 준다. 특히, 다양한 상황에서의 통신 흐름을 세부적으로 보여주기 위해 Alternative 블록을 활용하여 처리 과정을 구체화하였다. 예를 들어, 각 CNC 시스템에 OPC UA 연결을 시도할 때, 연결 가능 여부에 따라 다른 시나리오가 실행된다. 연결이 성공적으로 이루어지면 데이터 전송을 위한 설정이 진행되고, 만약 OPC UA 프로토콜이 지원되지 않는 경우에는 대체 방법을 통해 데이터를 수집할 수 있도록 다. 이처럼 다양한 상황에 유연하게 대응할 수 있도록 설계함으로써, 다양한 제조 시스템을 중앙에서 효율적으로 모니터링하고 관리하여 생산 환경의 통합 운영을 지원할 수 있게 된다.

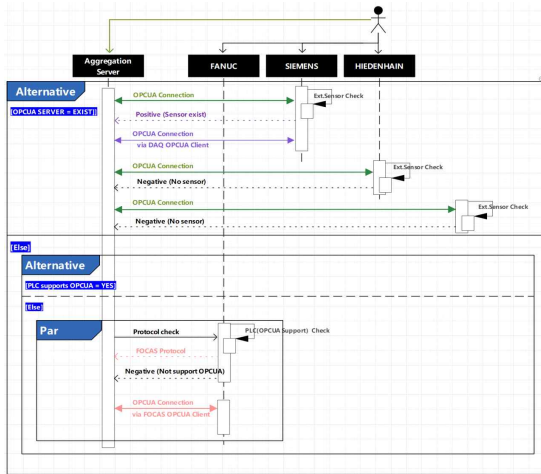


그림 5. 통합서버 시스템 다이어그램  
Fig. 5. Aggregation Server System Diagram

각 CNC 제어기에 따른 상세 시스템 다이어그램을 설계했으며, 그림 6에서는 FANUC CNC 컨트롤러의 데이터 수집 및 제어를 위한 UML 다이어그램을 예시로 제시한다. FANUC CNC의 데이터 수집에는 FOCAS (FANUC Open CNC API Specifications) Library를 사용한다. FOCAS 라이브러리는 이더넷 또는 HSSB (High Speed Serial Bus)를 통해 CNC 및 PMC (Programmable Machine Control) 컨트롤러에 접근하여 데이터를 수집할 수 있도록 지원한다. 이 시퀀스 다이어그램은 FANUC CNC와 FOCAS OPCUA Client 그리고 통합서버 간 데이터 통신 프로세스를 단계적으로 설명한다. 각 개체 간의 상호작용을 통해 제조 데이터 수집과 연계를 위한 절차를 각 스텝별로 상세히

설명하고 있다. 주요 기능으로는 개체들간의 연결 (Connection), 알람 및 이벤트(Alarm & Event), 데이터 취득 (Data Access) 세 가지로 구성되어있다. 시스템 흐름은 FOCAS OPCUA Client에서 MachineToolconnectRequest() 메서드를 통해 FANUC.CNC.dll 모듈로 초기연결 설정을 요청한다. 그리고 FANUC\_FOCASLibraryconnectRequest() 메서드를 호출하여 FANUC CNC에 연결요청을 보내고, 연결 오류가 없는 경우 ConnectionSuccess()가 호출되어 ConnectSuccess() 메서드와 ConnectOKCallback()메서드를 통해 FOCAS OPCUA Client와 통합서버 간 연결을 맺는다. 만약 연결이 실패할 경우 connectionNC\_Callback()이 호출되어 connectRetry()메서드를 통해 재시도 Loop가 이어지도록 설계하였다. 두 번째 기능인 Alarm & Event 기능에서는 FANUC CNC.dll모듈에서 AlarmStatusRequest()를 호출하여 FANUC의 알람상태를 주기적으로 요청한다. 이때, 알람 상태 체크는 1초간격(check.Trigger.1s.)으로 수행되며, 가변타이머를 통해 상황에 맞는 주기설정과 이벤트 변경 감지가 가능하도록 구현하였다. 알람 이벤트 리스트에 변동이 있는 경우 UpdateNewA/E()를 통해 새로운 알람 상태를 업데이트하고 서버로 전송하고, 변동이 없는 경우 ReleaseA/E\_update()로 반환한다.

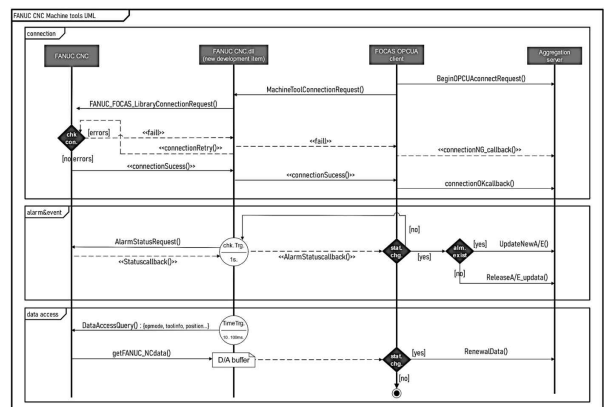


그림 6. FOCAS OPCUA Client 시퀀스 다이어그램  
Fig. 6. Sequence Diagram for FOCAS OPCUA Client

세 번째, Data Access Function에서 FANUC CNC.dll 모듈은 DataAccessQuery() 메서드를 통해 FANUC CNC의 제조 데이터를 요청하며, 여기에는 작업 모드(Operation mode), 축 위치(Axis position), 공구 정보(Tool information) 등의 주요 정보가 포함 된다. 데이터는 10ms에서 100ms 간격으로 수집되어 D/A 버퍼에 저장되며, 이는 실시간 데이터 처리를

가능하게 한다. Aggregation 서버는 RenewalData() 메서드를 통해 수집된 데이터를 최신 상태로 갱신한다.

이 다이어그램은 FANUC CNC와 Aggregation 서버 간의 통신 흐름을 구체적으로 보여주며, 초기 연결 설정, 알람 상태 모니터링, 주기적 데이터 수집 및 실시간 상태 업데이트 과정이 포함된다. 또한, CNC, OPC UA 클라이언트, 통합 서버 간의 통신 과정과 각 단계의 보안 설정 및 데이터 모니터링 흐름은 그림 7.에서 상세히 설명된다.

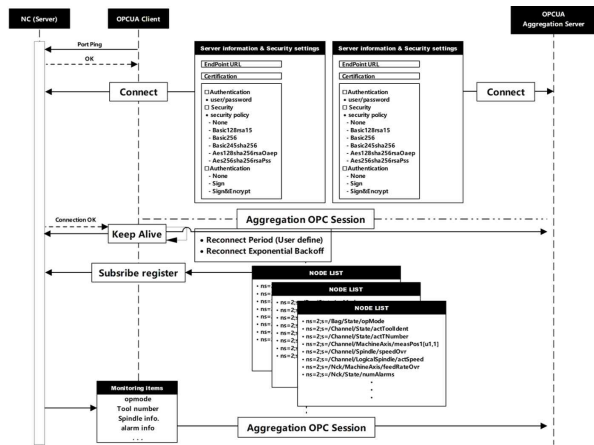


그림 7. 통합서버 시퀀스 다이어그램  
Fig. 7. Sequence Diagram for Aggregation Server

본 연구에서는 FANUC CNC에서 지원하지 않는 보안(PKI) 서비스를 추가하여, 민감한 제조 데이터를 보호하기 위해 Sign 또는 Sign & Encrypt 옵션을 설정할 수 있도록 구현하였다. OPC UA 서버를 통해 엔드포인트 URL과 사용자 인증 정보를 설정하고, Basic128rsa15, Basic256, Basic256sha256 등의 보안 정책을 선택할 수 있다. 연결이 완료되면 통합 서버와 OPC 세션이 시작되며, 세션 유지를 위해 Keep Alive 메커니즘이 작동하고, 네트워크 불안정 시 Exponential Backoff를 통해 안정성을 유지한다. Aggregation 서버는 OPC UA 클라이언트에 데이터 항목 구독 요청을 보내며, 구독 항목에는 작동 모드, 공구 번호, 스핀들 정보, 알람 정보 등이 포함된다. Aggregation 서버는 주기적으로 각 노드에서 데이터를 수집하고 최신 상태로 갱신하여, 제조 환경에서 여러 CNC 장비의 데이터를 통합 모니터링하고 애플리케이션과 쉽게 연동되도록 하며, 실시간 데이터 접근성과 보안성을 강화한다.

#### IV. 결 론

금속 가공 산업의 가치사슬 내에서 생성되는 산업 데이터는 암묵지적 성격이 강해, 그간 작업자의 경험과 지식에 의존해 전수 및 보존이 어려웠다. 하지만 디지털 전환을 통한 체계적인 표준화가 요구되는 대표적 산업 분야이다. 특히, FANUC, SIEMENS, HEIDENHAIN의 이기종 CNC와 멀티 센서의 OT 데이터를 통합 관리하고, 이를 IT 레벨의 지능형 애플리케이션과 표준화된 방식으로 연계할 필요성이 높아지고 있다. 본 연구를 통해 CNC 시스템의 상태와 가공 데이터를 통합 서버에서 일관되게 수집하고 모니터링함으로써, 제조 환경에서의 상호운용성과 실시간 데이터 접근성이 향상될 것으로 기대된다. 또한, 다양한 프로토콜을 지원하는 CNC에 대한 기술 표준을 제시함으로써 금속 가공 산업의 디지털 전환 모델에 신뢰성을 더하고, 국내외 시장 경쟁력 향상에 기여할 수 있을 것이다.

#### 참 고 문 헌

- [1] W. Kim and M. Sung, "Standalone OPC UA Wrapper for Industrial Monitoring and Control Systems," IEEE Access, vol. 6. Institute of Electrical and Electronics Engineers (IEEE), pp. 36557 - 36570, 2018.
- [2] C. Liu, X. Xu, Q. Peng, and Z. Zhou, "MTConnect-based Cyber-Physical Machine Tool: a case study," Procedia CIRP, vol. 72. Elsevier BV, pp. 492 - 497, 2018.
- [3] R. L. Pushkov, L. I. Martinova, and S. V. Evstafieva, "Extending functionality of the control system using MTConnectdata collection standard," IOP Conference Series: Materials Science and Engineering, vol. 971. IOP Publishing, p. 042080, 01-Dec-2020.
- [9] R. Pushkov, L. Martinova. and S. Evstafieva, "Extending functionality of the control system using MTConnect data collection standard," ICMTMTE 2020, IOP Conf. Series: Materials Science and Engineering vol. 971, p. 042080, December 2020.

# 제조 기계에 관한 OPC Companion Specification의 객체 분석 및 활용사례 구현

남승욱\*, 오세명\*\*

## Object Analysis and Use Case Implementation of OPC Companion Specification for machines in manufacturing

Seungwook Nam\*, and Saemyung Oh\*\*

### 요 약

기계공학 및 산업에서 디지털화라는 변화를 맞이함에 따라 자동화 기술, 가상 제품개발 및 운전 등 다양한 융합기술로의 확대가 이루어지고 있다. 이러한 맥락에서 상호운용 표준 인터페이스 기술을 정의하는 OPC Foundation에서는 산업계의 다양한 이해당사자와 함께, 산업별 전반적으로 적용할 수 있는 기능 및 사례, 정보 모델을 개발하고 있다. OPC 40001 UA for Machinery 또한 그러한 맥락에서 개발되었으며 특정 기계 타입으로 확장되기 위한 기본 정보모델을 제공한다. 본 논문에서는 OPC 40001에서 제공하는 주요 Object Type에 대해 분석하고 실 제조 환경에 적용하기 위한 방법론을 제시한다.

### Abstract

As mechanical engineering and industry face the change of digitalization, various convergence technologies such as automation technology, virtual product development, and operation are being expanded. In this context, the OPC Foundation, which defines interoperability standard interface technology, is developing functions, cases, and information models that can be applied across industries with various stakeholders in the industry. OPC 40001 UA for Machinery was also developed in this context and provides a basic information model for expansion to specific machine types. In this paper, we analyze the main Object Types provided by OPC 40001 and present a methodology for applying them to actual manufacturing environments.

### Key words

OPC UA, Information Model, Companion Specification, Machinery, Semantic Interoperability

## 1. 서 론

산업 기계에 관한 공통 기능 및 사례, 정보모델을 제시하는 OPC 40001 UA for machinery는 현재

---

\*한국전자기술연구원, swnam@keti.re.kr, \*\*한국전자기술연구원, semoh@keti.re.kr

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술평가관리원에서 지원받아 수행한 연구임 (No.20022994, 표준기반 제어기 고도화 및 디지털트윈 기반 제어기 성능검증 기술개발)

Part1~3이 출시되어있다.

각 Part는 기계의 식별정보에 관한 정보모델 (Part1:Basic Building Blocks)[1], 센서 또는 액추에이터의 정보처리에 관한 객체와 정보모델의 정의(Part 2: Process Values)[2], 기계에 관한 작업 관리 정보 모델(Part 3: Job Management)[3]에 관한 것이다. 제조사, 시리얼 번호 등을 포함한 기본적인 기계 식별 정보에서 시작하여 운영 데이터에 관한 처리 및 작업지시에 이르기까지 객체 정의가 구체화 되어 가는 추세이다. 또한 OPC40001은 그림 1과 같이 사출기, 로봇, 머신툴 등 세부 기계에 관한 정보모델에서 공통으로 적용하기 위한 객체를 제공한다.

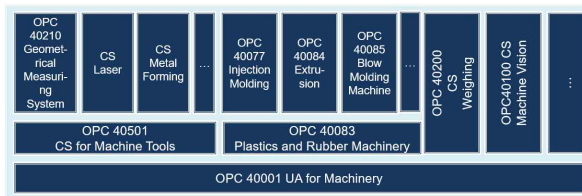


그림 1. OPC 40001을 기반으로 하는 Companion Specification 위계

Fig. 1. Companion Specification Hierarchy Based on OPC 40001

이와 같은 정보모델을 활용하여 제조사의 세부 기종별 표준 정보모델을 구현할 수 있으며 이를 활용하여 인터페이스 구현 시 다음과 같은 장점을 제공한다.

- OPC UA 기반의 상호운용 인터페이스 제공을 통한 연결성 제공
- 데이터에 관한 의미론적 상호운용성을 통해 비전문가의 쉬운 기계 데이터 이해
- OPC UA의 Security concept를 활용한 사이버 보안 및 유지관리에 관한 장점 등

## II. OPC 40001의 주요 객체 분석

Part1, Machine Identification and Nameplate 빌딩 블록은 기계를 전역적으로 고유하게 식별하고 기계에 대한 공급업체 정의 정보에 액세스하고 기계 식별을 위한 사용자별 정보를 관리하는 기능을 제공한다. 인터페이스 IMachineryItemVendorNameplate

Type 및 ObjectType MachineryItemIdentificationType은 다른 객체에서 참조하여 사용하는 일반적인 ObjectType이다. "2:Identification"(OPC 10000-100에 정의)을 가진 AddIn MachineIdentificationType은 MachineryItemIdentificationType에서 파생되므로 2:FunctionalGroupType에서 간접적으로 파생되고 인터페이스 IMachineVendorNameplateType 및 IMachineTagNameplateType을 구현한다. IMachineVendorNameplateType은 IMachineryItemVendorNameplateType의 하위 유형이므로 OPC 10000-100에 정의된 2:IVendorNameplateType에서 간접적으로 파생된다. IMachineryItemVendorNameplateType은 2:IVendorNameplateType에 정의된 속성을 변경하고 추가 Property를 정의한다. IMachineVendorNameplateType은 이러한 정의를 사용하여 별도의 필수 Property를 정의한다. IMachineTagNameplateType은 OPC 10000-100에 정의된 2:ITagNameplateType의 하위 유형이다.

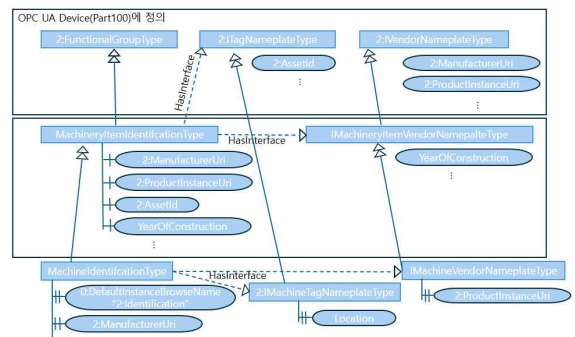


그림 2. Machine Identification and Nameplate Building Block의 구조

Fig. 2. Architecture of Machine Identification and Nameplate Building Block

Part 2: Process Value 정보모델은 센서 또는 액추에이터에서 제공하는 프로세스 값에 대한 정보를 제공한다. 또한 장치 식별에 대한 정보와 상태 정보, 오류 처리 방법 등이 포함된다. 특히 제조현장에서 자주 사용되는 기능인 알람 처리, Calibration 등 추가 메커니즘에 관한 객체도 포함된다.

해당 파트에서 제공하는 6개의 객체 정의 중, 본 논문에서는 대표적인 객체 유형인 ProcessValue - SetpointVariableType을 설명한다. VariableType은 0: AnalogUnitRangeType의 하위 유형이다. ProcessValue

-Type에서 3:AnalogSignal의 원하는 값을 정의하는데 사용되며 DataType, ValueRank 및 ArrayDimensions, 0:EngineeringUnits는 해당 변수와 동일해야 한다. 해당 VariableType의 인스턴스는 0:InstrumentRange를 제공하지 않으며 Process Value Setpoint 변수는 서버가 Process Value(Process Value Setpoint가 속한 변수)를 Process Value Setpoint로 가져오기 위해 내부 연산을 실행할 필요가 없다.

선택 사항으로 SubstituteValue 변수는 Process Value Setpoint를 더 이상 제어할 수 없을 때(통신 두절 등) 사용해야 하는 값을 제공한다. SubstituteValue는 Process Value Setpoint와 동일한 DataType, ValueRank 및 ArrayDimensions를 가져야 하며 Process Value Setpoint에 대해 정의된 것과 동일한 메타 데이터가 적용된다. LowLowDeviation, LowDeviation, HighDeviation 및 HighHighDeviation의 네 가지 선택적 편차 변수가 있으며 이들은 편차에 대한 한계 값을 정의한다. 편차는 백분을 또는 절대 값으로 정의할 수 있다. 편차 변수를 제공하는 ProcessValueSetpointVariableType의 모든 인스턴스는 스칼라 값을 가져야 한다. 또한 Subtype은 배열을 사용하여 생성하고 편차 변수에 대한 예상 동작을 정의할 수 있다. 그림 3은 편차 변수, Process-ValueSetpoint 및 범위의 관계를 보여준다.

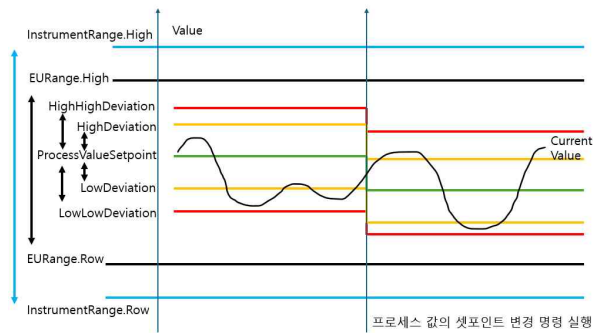


그림 3. 편차 변수, ProcessValueSetpoint 및 범위의 관계  
Fig. 3. The relation of the deviation Variables, the ProcessValueSetpoint, and the ranges

편차 한계에 도달하면 서버에서 0:DeviationAlarms을 생성할 수 있으며 편차 변수는 이 알람을 구성하는 데 사용된다.

선택 사항인 AutoDeviationAdjustment 변수는 구성 이 변경된 경우(0:EURange 또는 DeviationSensitivity)

DeviationSensitivity(제공된 경우) 설정에 따라 편차 변수가 자동으로 조정 되는지 여부를 정의한다. TRUE로 설정하면 편차 변수가 자동으로 조정되고 편차 변수 쓰기가 제한된다.

Part 3: Job Management는 ISA-95표준에 근거한 작업 관리 정보 모델에 대한 개요를 제공하며 작업 주문을 관리하기 위한 ObjectTypes에 대해 설명한다. 또한 해당 객체를 MachineryItem에 통합하는 방법, 작업 주문 실행 순서를 정의하는 메커니즘, 작업 주문 설명을 위한 사전 정의된 매개변수 개념이 정의되어있다. 그림 4는 JobManagementType에 대한 개요를 보여준다. 2:ISA95JobReceiverObjectType의 JobOrderControl 객체는 MachineryItem에 작업 주문을 추가하고 제어하며 또한 MachineryItem에서 현재 관리하는 모든 작업 주문(2:JobOrderList)에 대한 정보도 제공한다. JobOrderControl에서 관리하는 모든 작업 주문은 동일한 StateMachine을 지원하며 2:ISA95JobReceiverObjectType의 세부 정보는 OPC 10031-4에 정의되어 있다.

MachineryItem에서 작업 주문이 실행되면 JobOrderResults 개체에서 작업 주문 실행의 중간 결과 또는 최종 결과를 제공한다. 이 단계에서는 JobOrderControl 객체에서 작업 주문을 제어할 수 있다. JobOrderResults는 2:ISAJobResponseProviderObject-Type를 참조하며 결과를 노출하기 위한 다양한 메커니즘(메서드, 이벤트 또는 변수)을 제공한다. 클라이언트가 최종 결과를 받으면 MachineryItem에서 결과 값과 JobOrderControl에서 관리하는 작업 주문의 메타 데이터를 제거할 수 있다.

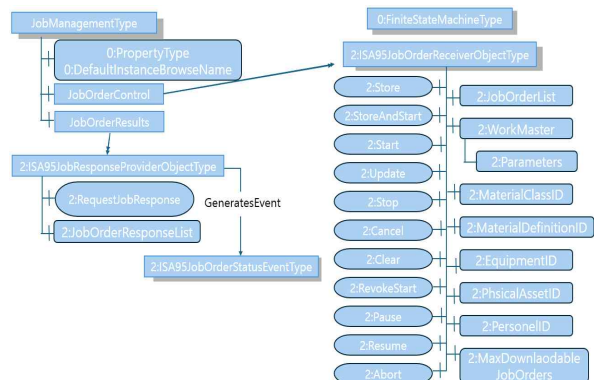


그림 4. JobManagementType의 개요  
Fig. 4. JobManagementType Overview

### III. Information Model의 적용 방안

OPC UA는 사용자측면에서의 상호운용성 외에도 개발자측면의 상호운용성을 제공한다. 즉 PC, 싱글보드 컴퓨터, 임베디드 등 다양한 플랫폼에 적용할 수 있다는 장점이 있다. 따라서 제조 현장의 기계 데이터를 받아 가공된 정보모델 형태로 데이터를 개방하는 게이트웨이의 역할은 제어기(PLC, CNC 등)에 직접 내장하거나 Edge 형태의 경량형 PC에 적용하여 단위 기계 별 게이트웨이를 구축할 수 있다. 또한 여러 기계의 데이터를 취합하여 포괄적 정보모델을 제공하는 방법도 가능하며 이와 같은 경우, 워크스테이션, 서버에서 Aggregator로써 구현될 수 있다.

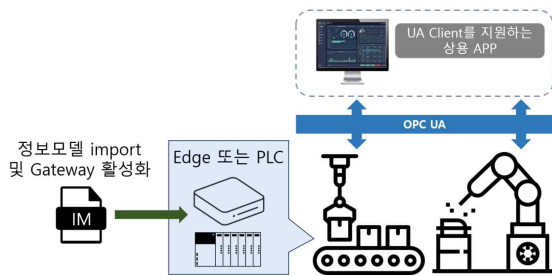


그림 5. Information Model 개념  
Fig. 5. Implementation concept of Information Model

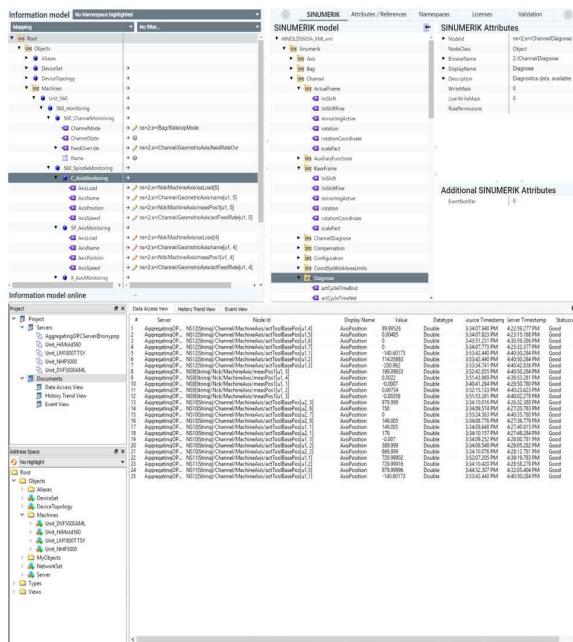


그림 6. Information Model 적용 예시  
Fig. 6. Implementation of Information Model

그림 6과 같이 상용 OPC UA 모델링 툴을 활용하여 가공기(Machine Tool)의 정보모델을 OPC40001과 40501을 참조로 생성한 후, PLC에 적용하여 사용자 정의의 데이터 규격에 따라 OPC UA 서버를 구현한 후 상용 OPC UA Client를 통해 접속하여 데이터가 정상 수신됨을 확인하였다.

### IV. 결 론

일반적으로 OPC UA는 단순 프로토콜로써 오해하는 경우가 많으며, 따라서 OPC UA를 적용했다는 것은 3rd Party 프로토콜을 OPC UA로 변환했다는 의미로 받아들여지기 쉽다. 하지만 OPC UA는 단순한 통신 프로토콜을 넘어 다양한 인터페이스 표준과 정보모델, 보안 등을 포괄적으로 지원하는 플랫폼에 가깝다. 특히 OPC UA Companion Specification을 기반으로 기계의 정보모델을 구현하여 적용한다는 것은 OPC UA 인터페이스를 활용할 수 있다는 것을 넘어 기계의 데이터를 체계적이고 이해하기 쉽게 가공할 수 있다는 것이 핵심이다. 이기종간 데이터를 공유하고 이해하기 위한 Description을 표준화된 형태로 제공하는 것은 장치산업에서는 일반적인 것이며(예: Profi-net의 GSDML) 이러한 경향이 OPC UA를 통해 기계 또는 공정 단위까지 확장되었다고 할 수 있다. OPC UA의 의미론적 상호운용성은 기존 공장 자동화 영역을 넘어 다양한 IT 솔루션과의 연계를 요구하는 시장의 요구사항을 만족할 수 있을 것으로 평가한다.

### 참 고 문 헌

- [1] OPC UA for Machinery Part 1: Basic Building Blocks (Release 1.01.0, Apr 2021)
- [2] OPC UA for Machinery Part 2: Process Values (Release 1.00, May 2023)
- [3] OPC UA for Machinery Part 3: Job Management (Release Candidate 1.0, Jul 2023)

# 적응 가능한 제조시스템 구현을 위한 필수 구성요소에 관한 연구

오세명\*, 남승욱\*\*

## Research on essential components for implementing adaptive manufacturing systems

Sae Myung Oh\* and Seung Wook Nam\*\*

### 요 약

기존 자동화시스템은 많은 인력이 투입된다는 점과 고정된 환경으로 인해 제품에 따라 쉽게 변화를 적용하기 어렵다는 점이 현시대의 시장 특성과 맞지 않다. 이를 개선하고자 최근 자율제조 시스템 구축에 대한 관심이 높아졌고, 이는 AI기술이 제조 적용됨에 따라 더욱 탄력을 받았다. 본 논문에서는 자율제조 시스템의 더욱 고도화된 적응 가능한 제조시스템을 논하고, 이를 구현하기 위한 필수 구성요소들을 열거하여 적응 가능한 제조시스템 구축에 기여하고자 한다.

### Abstract

Existing automated systems do not fit the characteristics of the current market because they require a lot of manpower and have a fixed environment, making it difficult to easily apply changes depending on the product. To improve this, interest in building autonomous manufacturing systems has recently increased, and this has gained more momentum as AI technology is applied to manufacturing. In this paper, we discuss a more advanced adaptive manufacturing system of autonomous manufacturing systems and list the essential components for implementing it, thereby contributing to the construction of an adaptive manufacturing system.

### Key words

autonomous manufacturing, optimization, OPCUA, interoperability, information model

### 1. 서 론

OT와 IT기술의 융합이 가능해지고, 그 경계가 얇아지면서 제조 혁신을 위한 다양한 플랫폼, 어플리

케이션이 나오고 있다. 이러한 제조 혁신의 초기에는 필드 영역의 데이터를 상위 IT단으로 효율적으로 전달할 수 있는 것에 집중하면서 IoT기술이 주요 이슈였다. 최근에는 AI기술의 재등장과 그 효과

\*한국전자기술연구원 자율제조연구센터, semoh@keti.re.kr, \*\*한국전자기술연구원 자율제조연구센터, swnam@keti.re.kr

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술기술평가원에서 지원받아 수행한 연구임 (No.20022994, 표준기반 제어기 고도화 및 디지털트윈 기반 제어기 성능검증 기술개발)

및 성능이 높아짐에 따라 AI 기반의 제조 혁신이 박차를 가하고 있는게 현실이다.

자율제조(Autonomous Manufacturing)시스템은 인간의 개입이 줄고, 기계와 로봇, 디바이스가 스스로 생산을 운영하는 시스템을 말한다. 하지만 생산 현장에는 항상 새로운 이벤트가 빈번히 발생하고, 기존과는 다른 또다른 결정을 해야할 경우가 많다. 또한, 소비자 욕구의 변화는 생산 시스템에 곧장 반영이 되어야 하는 만큼 인간의 개입이 없는 완전한 자율제조는 사실상 어려울 수 있다.

자율제조보다 더욱 고도화된 시스템인 적응 가능한 제조시스템(Adaptive Manufacturing System)은 제조 혁신을 일으킨 IoT, 5G, AI 등을 활용하여 구현 가능하다. 본 논문에서는 적응 가능한 제조시스템을 위해서 갖추어야 하는 특성 및 기술요소를 나열하고, 이를 설명한다.

## II. 지능적 중앙 제어

적응 가능한 제조 시스템에서 가장 중요한 역할을 하는 것이 “지능적 중앙 제어”를 담당하는 중앙 제어 모듈(Central Control Module, CCM)이다. 매우 복합적인 기술이 포함되어 있는 게이트웨이 성격의 시스템으로써 IoT, 디지털트윈, AI, 네트워크 기술을 기반으로 한 대규모 시스템이다.

적응성을 달성하기 위해서는 일반적으로 장비를 제어하는 PLC가 신속하게 재구성이 가능해야[1] 하기 때문에 CCM에서 통합 및 지시/명령이 가능해야 한다. 적응 가능한 제조 시스템은 이벤트에 반응하기 위해 실시간으로 데이터 취득이 필요하며, 이를 기반으로 가장 최적의 판단을 항상 변경하고, 생산 시스템에 반영할 수 있어야 한다. 또한, 오류가 없는지 확인 테스트를 위해 반영하기 이전에는 시뮬레이션이 이루어져야 하기 때문에[1] 이러한 부분들이 자동화 된다면 인간의 최소한의 노력으로 이를 자율화 할 수 있다.

AI기법으로는 지속적으로 학습이 가능한 신경망 모델 혹은 조합 최적화 알고리즘이 활용될 수 있다.

이러한 기능을 하는 개별의 모듈들이 일련의 반복적인 프로세스를 통해서 피드백을 통해 스스로

진화하고, 새로운 변화에 적응할 수 있다.

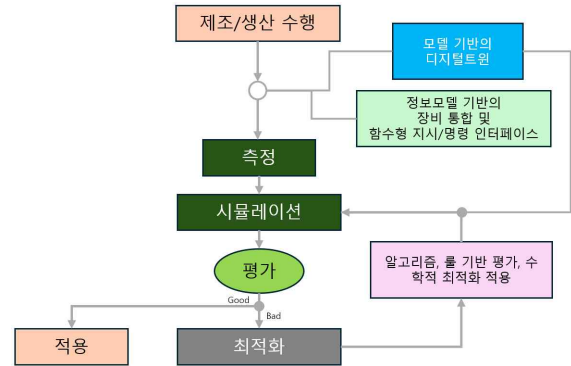


그림 1. 중앙 제어 모듈 운영 프로세스  
Fig. 1. Operation Process of CCM

## III. 객체 인식

생산 시스템은 일련의 작업을 담당하는 여러 대의 설비 그리고 이를 연결하고, 작업을 지원하는 로봇과 그 외 인프라 시스템이 하나의 라인을 구성하여 제품을 생산한다. 이는 생산에 개입하는 장비 및 시스템이 서로 다른 특성과 역할로 존재한다는 의미이며, 이를 중앙 제어 하는 게이트웨이 입장에서는 식별할 수 있는 기능이 필요하다.

OPCUA(OPC Unified Architecture)는 설비를 제어하는 각각의 제어기의 벤더사와 상관없이 통일된 통신 인터페이스를 제공하는 동시에 정보모델이라는 데이터 구조 공간이 존재하며, 이는 파일(xml)형태로 관리가 가능하다. 이러한 정보모델 기반으로 설비 및 시스템을 동일한 통신 방식으로 통합해제가 가능한 게이트웨이 기술이 필요하다.

## IV. 상호운용성

앞서 언급한 OPCUA 기술은 제조 산업에서의 이기종 간 상호운용성을 보장하도록 지원한다. 그 범위는 통신 인터페이스 뿐만 아니라 정보모델 및 동반사양(Companion Specification)을 통해 데이터에서의 의미론적 상호운용까지 광범위하다. 표 1은 OPCUA 동반사양 기반의 로봇과 머시비전 시스템의 상호운용에서 활용될 데이터 노드를 작성한 것이다.[2,3] 디바이스와 로봇 그리고 장비들은 제작사



가 서로 다르기 때문에 공통 데이터 규칙을 정하는 것이 쉽지 않다. 하지만 적용 가능한 생산 시스템에서 모든 시스템이 자율적으로 상호운용되기 위해서 표준이 필요하다.

표 1. 로봇과 머신비전 시스템 간 데이터 상호운용

Table 1. Data Interaction between robot and MV

	Robotics	MachineVision
1. Condition Diagnosis	OnPath, InControl(Bool), SpeedOverride(Double)	VisionStateMachineType>Preoperational
2. Current OpMode	OperationalMode(Bool)	
3. Safety Emergency	EmergencyStop, ProtectiveStop(Bool)	VisionSafetyTriggered(Bool)
4. Program Selection - program start - move - position arrival	TaskProgramName(String) TaskProgramLoaded(Bool) ExecutionMode(Enum) (0:Single, 1:AutoRepeat, 2:Step)	-
5. Camera On/Off	-	VisionStateMachineType>Halt/Operational
6. Result Management	-	GetResultComponentsById(Method)
7. Error or not	-	VisionStateMachineType>Error
8. Cycle Finish(Robot at Origin Position)	ActualPosition(Double)	-

## V. 자율 로봇

적용 가능한 제조시스템에서 로봇의 역할이 가장 중요하다. 사실상 현 생산시스템 컨셉에서 인간의 역할을 대신할 대상이기도 하다. 물리적인 움직임의 자유로움도 갖추어야 하지만 적응형 지능 또한 중요하다. 지능형 중앙 제어장치와 모든 로봇은 연동되며, 중앙 제어장치에 대해 지시(request), 질의(ask),

응답(response)를 수행하고, 피드백을 전달해야 한다. 자율 로봇은 새로운 지식을 습득하고, 변화에 대응할 수 있는 진화적 알고리즘(evolutionary algorithm)을 탑재하여야 하며, 상호운용이 가능한 통신 인터페이스로 연동되어야 한다.

## VI. 결 론

본 논문을 통해 적용 가능한 생산시스템의 구현에 필수적인 구성요소를 나열하고, 설명하였다. AI 기술의 발달과 정확도 그리고 연산 속도가 점점 개선이 된다면 적용 가능한 생산시스템의 그 모습은 더욱 뚜렷해질 것이다. 하지만 우려되는 점은 수많은 AI 관련 어플리케이션만 존재한다고 자율제조를 완성할 수 있는 것이 아니다. AI기술을 적용하기 이전인 OT단과 데이터 가공 역할을 하는 게이트웨이단 또한 많은 노력과 기술개발이 필요하다는 점을 배제하면 안된다.

OPCUA, 제어기 고도화 기술에 대한 많은 연구 및 개발 내용들이 늘어나고, 결과물이 다양해질수록 AI 성능 또한 올라갈 것이다.

## 참 고 문 헌

- [1] Fan Mo, Miriam Ugarte Querejeta, Joseph Hellewell, Hamood Ur Rehman, Miren Illarramendi Rezabal, Jack C. Chaplin, David Sanderson, Svetan Ratchev, "PLC orchestration automation to enhance human-machine integration in adaptive manufacturing systems", Journal of Manufacturing System, Vol 71, pp. 172-187, December 2023.
- [2] OPC Foundation, OPC 40100-1 OPCUA for Machine Vision, Part1: Control, configuration management, recipe management, result management, Release 1.0, August 2019
- [3] OPC Foundation, OPC 40010-1 OPCUA for Robotics, Part1: Vertical Integration Release 1.0, July 2019

# 연구데이터 관리 및 활용을 위한 교육과 가이드라인의 필요성

황바른\*, 이미경\*, 임형준\*

## The Need for Education and Guidelines to Enhance Research Data Management and Utilization

Barun Hwang\*, Mikyoung Lee\*, Hyung-Jun Yim\*

국내에는 정부출연연구기관과 대학을 비롯한 다양한 기관에서 데이터 기반 연구가 활발히 이루어지고 있으며, 이를 통해 다량의 연구데이터가 생산되고 있다. 최근 오픈 사이언스 확산과 더불어 연구데이터의 관리 및 공유를 통해 재사용을 촉진하는 것이 중요해지고 있으며, 이에 연구데이터 관리와 공유를 위한 플랫폼과 리포지터리도 증가하고 있다. 그러나 연구자들은 체계적으로 연구데이터를 관리하고 활용하는 데 기술 부족, 정책 및 규정의 미비 등으로 인해 어려움을 겪고 있다[1]. 따라서 연구데이터 리포지터리와 관련 기관에서는 연구자의 데이터 관리와 활용을 지원하기 위해 체계적인 교육과 구체적인 가이드라인 개발이 필요하다.

최근 연구데이터뿐만 아니라 연구 전 과정에 필요한 개발 및 분석 도구와 소프트웨어를 공유하고 활용하기 위한 '연구데이터커먼즈(Research Data Commons)' 개념이 주목받고 있다[2]. 연구데이터 관리와 활용을 활성화하기 위해서는 데이터를 효과적으로 분석하고 활용할 수 있는 기술 역량이 필요하다. 이를 위해 연구자들은 Python, R과 같은 프로그래밍 언어와 AI, 머신러닝을 활용한 다양한 데이터 분석 및 개발 도구를 활용할 수 있어야 한다. 연구데이터 관리의 주요 기능에는 DMP(Data Management Plan) 작성 지원, RDM(Research Data Management), 데이터 저장, 공유 및 접근, 데이터 보안, 윤리 및 저작권 등이 포함되며[3], 데이터 기반 연구를 수행하는 연구자들은 이러한 지식과 기술을 습득하고 이를 연구에 적용할 수 있어야 한다.

연구데이터 관리 및 활용 관련 교육과 가이드라인 제공을 통해 연구자들이 데이터를 보다 심도 있게 분석하고, 새로운 연구 주제를 발굴하며, 데이터를 재사용함으로써 연구의 가치를 극대화할 수 있다. 본 연구에서는 선행연구를 바탕으로 연구데이터 리포지터리에서 데이터 관리와 활용을 위해 필요한 데이터 수집, 관리, 인용, 공유, 윤리, 보안뿐만 아니라 데이터 분석 및 활용에 필요한 기술 역량 등의 사항에 대한 교육과 가이드라인을 구축하는 것을 목표로 한다. 이를 통해 연구자들에게 연구데이터 관리와 활용에 필요한 구체적인 지침을 제시하고, 연구데이터를 체계적으로 관리하고 활용할 수 있도록 지원하고자 한다.

### Key words

Research Data Management, Open Science, Data Service, Data Repository

### 참고문헌

- [1] 채현수 외 3인. (2021). 대학구성원의 연구데이터 관리 인식 및 경험 연구. 정보관리학회지, 38(4), 173-198.
- [2] 박성은 외 5인. (2024). 국가연구데이터커먼즈 체계 수립을 위한 연구데이터 관리자들의 인식에 관한 연구. 정보관리학회지, 41(1), 465-486.
- [3] 김주섭 외 2인. (2019). 연구 데이터 관리 및 서비스를 위한 핵심요소의 기능적 요건. 한국문헌정보학회지 (Journal of the Korean Society for Library and Information Science), 53(3), 317-344.

\*한국과학기술정보연구원, {barun, jerryis, hjyim}@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24LIMC3-01)

# KRDC 프레임워크를 활용한 지질과학 연구 데이터 기반 인공지능 모델 재현 및 활용

이원희\*, 조민희\*, 송사광\*, 이미경\*, 임형준\*

## Reproduction and Application of AI Models Based on Geoscience Research Data Using the KRDC Framework

Won-hee Lee\*, Minhee Cho\*, Sa-Kwang Song\*, Mikyoung Lee,\* and Hyung-Jun Yim\*

최근 인공지능 기술의 발전으로 인해 자연과학 분야에서도 인공지능을 결합한 연구가 지속적으로 증가하고 있다. 이러한 연구에서는 인공지능 모델이 예측이나 분석에 중요한 역할을 하기 때문에, 연구의 재현 가능성을 확보하는 것이 결과의 신뢰성을 높이는 데 필수적이다. 그러나 동일한 데이터와 인공지능 기술을 사용하더라도 연구자들이 보유한 컴퓨팅 환경과 리소스들의 차이로 인하여, 신뢰할 수 있는 연구 결과를 재현하는 데 어려움이 있는 상황이다. 이를 해결하기 위해 동일한 컴퓨팅 환경에서 다양한 연구 데이터를 활용하여 신뢰 가능한 연구 결과를 재현할 수 있는 국가연구데이터커먼즈(KRDC; Korea Research Data Commons) 프레임워크[1]를 활용한다. 본 논문에서는 KRDC 프레임워크를 활용하여 지질과학 분야 연구의 신뢰성과 향상된 모델 개발 가능성을 제시한다. 해당 프레임워크는 연구자가 필요한 태스크를 수행할 수 있는 애플리케이션들을 연결해 워크플로우를 구성하며, 이를 통해 지질과학 분야 연구 데이터의 활용을 극대화하고, 기존 연구의 재현과 모델 개선을 지원하는 방안을 제안한다.

본 논문에서는 지질과학 분야에서 활용되는 연구 중 하나인 탄성과 반사계수 예측을 위한 워크플로우를 개발하고 그 재현성을 검증한다. 워크플로우는 두 가지 주요 기능을 가진 앱들로 구성된다. 첫째, 데이터를 전처리하여 학습이 가능한 데이터셋으로 분리한다. 둘째, 사용자가 원하는 인공지능 모델을 선택하여 학습시킴으로써 탄성과 반사계수를 예측한다. 제안하는 워크플로우를 통해 사용자는 데이터셋을 전처리하여 학습용 데이터와 테스트용 데이터로 분리하고, 선택한 인공지능 모델을 이용해 학습과 예측을 수행할 수 있다. 위와 같이 작성된 워크플로우는 한국지질자원연구원 클러스터에 설치된 KRDC 프레임워크에 탑재되어 그 실효성을 증명하였다. 이를 통해 다양한 분야의 연구자들이 인공지능 학습 모델을 보다 쉽게 사용할 수 있는 컴퓨팅 환경을 제공하고, 연구 결과의 신뢰성을 향상시킬 수 있음을 확인하였다. 또한, 이 워크플로우는 지질과학 분야를 포함한 다양한 도메인에서도 활용될 수 있음을 확인하였다.

### Key words

Korea Research Data Commons, Research data, Reproducibility, Geological science

### 참고문헌

- [1] D. Kim, et al., "Development of Korea Research Data Commons Framework for Resource Federation Between Multi-clusters", *Database Research*, Vol. 40, No. 2, pp. 50-65, Aug. 2024.

\*한국과학기술정보연구원 연구데이터공유센터, {wonhee, mini, esmallj, jerryis, hjyim}@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24L1M03-01)

# 리소스 허브 중심의 연구데이터 통합플랫폼 설계

이미경\*, 조민희\*, 김다솔\*, 엄정호\*, 임형준\*

## Design of Research Data Platform as a Resource Hub

Mikyoung Lee\*, Minhee Cho\*, Dasol Kim\*, Jungho Um\* and Hyung-Jun Yim\*

본 논문에서는 리소스 허브 기능을 중심으로 하는 연구데이터 통합플랫폼에 대해 설명한다. 현재 한국과학기술정보연구원(KISTI)에서는 연구데이터를 체계적으로 공유·관리하고 검색 및 분석·활용을 지원하는 국가연구데이터플랫폼(DataON)을 개발·운영하고 있다[1]. 연구자들간 연구데이터를 공유·활용함으로써 연구 생산성을 향상시키고자하는 목적으로 개발된 DataON은 개인 연구자와 연구기관에서 수집한 연구데이터와 활용 측면에서 데이터 분석에 필요한 인프라와 분석 도구를 서비스하고 있다. 세계적인 추세인 오픈사이언스와 FAIR(Findable, Accessible, Interoperable, Reusable)원칙과 함께 데이터 공유는 데이터뿐만 아니라 리소스도 함께 공유·활용하는 데이터커먼즈 체계로 발전하고 있다. 유럽에서는 EOSC(European Open Science Cloud) EU Node[2]를 통해 연구자들에게 논문, 데이터, 소프트웨어, 서비스, 가이드라인, 도구 등을 제공하고 있다. KISTI에서는 연구데이터커먼즈를 위한 연구소프트웨어 공유·활용 체계 연구를 수행 중이다[3]. DataON이 연구데이터 통합플랫폼으로 발전하기 위해서는 출연(연), 데이터 전문센터 등에서 공유되는 연구데이터의 관리·통합뿐만 아니라 데이터 기반 연구에 필요한 모든 리소스(논문, 소프트웨어, 서비스, 인프라 등)의 관리 및 공유로 범위를 확대하여 지원해줄 필요가 있다. 우리는 연구데이터, 분석 소프트웨어, 모델, 소스 코드, API, 서비스, 인프라 등을 수집·연계하고 관리·검색하기 위한 카탈로그 기반 리소스 허브 서비스를 제공하고자 한다. 리소스 허브에서 제공되는 소프트웨어와 모델 등의 리소스는 각 기관의 데이터플랫폼에서 필요에 따라 개발된 소프트웨어와 Github와 같은 오픈 소스 저장소를 통해 공유된 소스 코드를 말하며, 연구자는 소프트웨어/소스의 실행 환경이 포함된 분석 도구 및 인프라를 사용하여 연구의 재현성을 높일 수 있다. 연구자는 연구데이터 검색을 통해 여러 곳에 산재되어 있는 연구데이터에 접근하며, 데이터 분석에 필요한 리소스도 사용할 수 있다. 연구데이터 통합플랫폼은 리소스 허브와 함께 분석도구, 주피터 노트북, 실행 환경기반 이미지, 분석 인프라 등을 서비스로 제공하여 연구자들에게 데이터 확보부터 분석까지 원스톱 연구를 수행할 수 있게 지원한다. 향후 연구데이터 통합플랫폼은 전 분야의 리소스를 공유·활용하는 데이터커먼즈 체계로 발전하고자 한다.

### Key words

Research Data, Research Resource, Research Data Platform, Resource Hub

### 참고문헌

- [1] 황미영 외 4인, “연구데이터 공유·관리·분석을 위한 국가연구데이터플랫폼(DataON) 서비스 운영”, 한국콘텐츠학회 논문지, vol.23, no.1, pp.36-47, 2023.
- [2] “European Open Science Cloud - EU Node”, <https://open-science-cloud.ec.europa.eu/>
- [3] 김다솔 외 4인, “멀티 클러스터 간 리소스 연계를 위한 국가연구데이터커먼즈 프레임워크 개발”, 데이터베이스연구, vol.40, no.2, pp.50-65, 2024.

---

\*한국과학기술정보연구원, {jerryis, mini, dskim, jhum, hjyim}@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 자체사업 창의도전연구과제로 수행된 연구입니다.(과제번호: (KISTI)24JR042-244)

# 데이터 증강을 활용한 소량 생산 제조 공정 불량 검출 기술 연구

변상봉\*, 이창용\*\*, 권소영\*\*, 한동철\*, 이성훈\*, 정용안\*, 조수현\*

## A Study on Defect Detection Technology for Low-Volume Manufacturing Processes Using Data Augmentation

Sang-Bong Byun\*, Chang-Yong Lee\*\*, So-Young Kwon\*\*, Dong Cheul Han\*, Sung hun Lee\*, Yong An Jung\*, and Soo-Hyun Cho\*

본 논문은 YOLOv5[1] 알고리즘을 이용해 소량 생산 제조 공정에서 발생하는 불량을 검출하는 기술에 대해 다룬다. YOLOv5를 활용한 불량 검출은 실시간 객체 탐지 능력을 통해 제조 공정에서 발생하는 불량을 신속하고 정확하게 식별하는 데 효과적이며, 특히 경량화된 모델로, 빠른 처리 속도를 제공하여 제조 공정에서 연속적인 이미지 스트림을 처리하고 실시간 검출이 필요한 환경에 적합하다. 본 연구에서는 제조 공정에서 수집된 불량 이미지에 대해 YOLOv5 모델을 학습시켜 불량 위치와 유형을 빠르게 감지할 수 있도록 했습니다. 하지만 소량 생산 제조공정은 인공지능 학습을 위한 대량의 이미지 데이터가 부족한 문제가 있다. 이를 해결하기 위해 데이터 증강(Data Augmentation)[2] 기법을 적용하여 소량의 불량 데이터를 다양한 형태로 증강, 학습 데이터의 양을 보완함으로써 불량 검출의 정확도를 높였다. 데이터 증강 기법으로는 회전, 확대, 색상 변환 등 다양한 변형을 통해 기존 불량 이미지의 변형 데이터를 생성하였으며, 이를 통해 모델이 다양한 불량 패턴을 학습하도록 했습니다. 이로 인해 소량 데이터 환경에서도 신뢰성 높은 검출 성능을 확보할 수 있다. 또한, 제조 공정 중 확보되는 불량 검출 이미지를 주기적으로 추가 학습시켜 알고리즘을 지속적으로 개선하고, 변화하는 공정 조건에 대응할 수 있도록 했다. 구동 디바이스는 NVIDIA Jetson AGX Orin이며, 해상도와 프레임 레이트는 해당 공정 불량 특성에 따라 가변적으로 활용한 계획이다. 이러한 기술은 다양한 제조 공정에 쉽게 적용 가능하며, 고정밀 불량 검출이 필요한 산업 전반에서 유용하게 활용될 수 있다.

### Key words

data augmentation, defect detection, artificial intelligence, edge device

### 참고문헌

- [1] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788, June 2016.
- [2] R. Guerrero, L. Carvalho, T. Bocklitz, J. Popp, J. L. Oliveira, "A Data Augmentation Methodology to Reduce the Class Imbalance in Histopathology Images", Journal of Imaging Informatics in Medicine 2024, vol. 37, pp. 1767 - 1782, Apr. 2023.

\* 구미전자정보기술원 지능형디바이스연구센터, email : sbbyun@geri.re.kr(교신저자), cataegu07@geri.re.kr, leesh@geri.re.kr, yajung@geri.re.kr, shcho@geri.re.kr

\*\* 국립금오공과대학교 일반대학원 전자공학과, email : lcy5200@kumoh.ac.kr, papaya4040@kumoh.ac.kr

※ 이 연구는 2024년도 산업통상자원부 및 한국산업기술기획평가원 연구비 지원에 의한 연구임(유도탄 제조 공정의 디지털전환을 위한 AI기반 고신뢰성 로봇-장비 및 위험관리 통합운영 기술 개발, RS-2024-00445778)

## HSI와 식생지수를 활용한 수목 군집 탐지

권산\*, 정재경\*, 아이다 아야노\*\*, 박찬\*\*, 어수행\*\*\*, 이재용\*\*\*\*, 김충식\*\*\*\*, 최호식\*

### Detecting tree clusters using hsi and vegetation indices

San Gwon\*, Jaekyeong Jung\*, Ayano Aida\*\*, Chan Park\*\*, Abel Eo\*\*\*, Jaeyong Lee\*\*\*\*, Chungsik Kim\*\*\*\*, and Hosik Choi\*

아까시나무는 문화재 관리에 있어 위해식물로서, 그 확산을 조기에 탐지하고 제어하는 것이 중요하다. 하지만 대면적 문화재의 경우, 지형상 제약과 넓은 범위로 인한 인력 문제 등으로 육안 탐지가 힘들다. 최근 이를 해결 하려는 방안으로 드론을 이용하여 측정 가능한 HSI 기술이 주목받고 있다. HSI(hyperspectral image)는 일반적인 이미지와 달리 가시광선, 근적외선을 모두 포함한 정보를 가지고 있는 이미지이다. 이를 위해, 연구는 HSI 데이터를 바탕으로 다양한 식생지수를 계산하여 수목의 건강 상태 및 유형을 구분하고, 머신러닝과 클러스터링 기법을 적용해 아까시나무 군집을 탐지하였다. 본 연구는 HSI와 식생지수를 결합하여 대면적 문화재 구역 내에서 아까시나무의 군집을 효율적으로 탐지하는 방법을 제시한다.

탐지 모델 로는 XGBoost와 DBSCAN 알고리즘을 사용하였다. XGBoost는 HSI 데이터의 복잡한 패턴을 학습 하여 아까시나무를 다른 식물 종과 구분하는 분류 모델을 구축하는 데 활용되었다. HSI는 촬영 환경 및 보정 방법에 따라 노이즈가 내재되어 있어, 탐지 결과에 영향을 미친다. 이를 보완하기 위해, 본 연구는 그룹화된 대역대의 특징변수와 인접 픽셀의 스펙트럼으로부터 추출한 정보를 활용한다. 또한 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)은 밀도 기반 클러스터링 기법으로, 아까시나무가 밀집된 지역을 군집화하여 그 분포를 명확히 시각화하고 탐지하는 데 사용되었다.

이 연구는 HSI 데이터와 식생지수를 활용한 머신러닝 기법이 대규모 문화재 구역 내에서 위해수종을 신속하고 정확하게 탐지할 수 있는 가능성을 제시하며, 향후 생태계 보존 및 관리 방안에 중요한 기여를 할 수 있을 것으로 기대된다.

#### Key words

hyperspectral image, xgboost, artificial intelligence, dbscan,vegetation indices

---

\* 서울시립대학교 도시빅데이터융합학과, jmt30269@uos.ac.kr, choi.hosik@uos.ac.kr(교신저자)

\*\* 서울시립대학교 조경학과

\*\*\* (주)크립토크

\*\*\*\* 전통문화대학교 전통조경학과

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NO.2022M3J6A1084845)

# 임의 부분 노이즈제거를 지원하는 윤곽선 매칭의 효율적인 색인 구축 방법

김범수\*

## An Efficient Index-Building Method for Boundary Matching that Supports Arbitrary Partial Denoising

Bum-Soo Kim\*

최근 AI 기술이 빠르게 발전하면서 양질의 원본 데이터를 확보하고 처리하는 것은 중요한 문제이다[1]. 본 논문에서는 대용량 윤곽선 이미지 데이터베이스에서 임의 부분에 있는 윤곽선 노이즈를 고려하여 윤곽선 이미지 매칭이 가능하도록 효율적인 색인 구축 방법을 제안한다. 임의 부분에 있는 윤곽선 노이즈를 고려하여 색인을 구축하는 것은 가능한 모든 노이즈제거 매개 변수에 따른 부분 노이즈제거를 고려해야하기 때문에 어려운 문제이다[2]. 기존 연구에서는 다차원 색인[3]을 활용하여 모든 가능한 노이즈제거 매개 변수에 의한 최소 경계 영역(MBR)을 구성하여 효율적인 단일 색인 생성 알고리즘을 제안하였다[4]. 그런데 이 알고리즘에서는 MBR를 구성하는데 많은 계산량이 발생한다. 본 연구에서는 MBR 구성 시 계산 과정을 최적화하여 빠른 색인을 구축한다. 즉, MBR 구성을 위해 필요한 최하점과 최상점을 계산할 때 반복적으로 계산하는 부분 노이즈제거 부분 값을 저장하여 이를 활용한다. 실험 결과, 기존 방법에 비해 1.6배에서 14.7배까지 구축 시간을 단축하는 것으로 나타났다. 향후 연구로는 임의 부분 노이즈를 고려한 윤곽선 매칭을 위해 AI 기술을 활용할 예정이다.

Key words

time-series matching, partial denoising, indexing, optimization

### 참고문헌

- [1] Y. Gong, G. Liu, Y. Xue, R. Li, and L. Meng, "A survey on dataset quality in machine learning," *Information and Software Technology*, Vol. 162, pp. 1-12, Oct. 2023.
- [2] B.-S. Kim, Y.-S. Moon, and J.-G. Lee, "Boundary Image Matching Supporting Partial Denoising Using Time-Series Matching Techniques," *Multimedia Tools and Applications*, Vol. 76, No. 6, pp. 8471-8496, Mar. 2017.
- [3] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, "The R\*-tree: An Efficient and Robust Access Method for Points and Rectangles," In *Proc. of the ACM SIGMOD Int'l Conf. on Management of Data*, Atlantic City, New Jersey, pp. 322-331, May 1990.
- [4] B.-S. Kim, "An Index-Building Method for Boundary Matching that Supports Arbitrary Partial Denoising," *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 23, No. 11, pp. 1343-1350, Nov. 2019.

---

\*한국건설기술연구원 미래스마트건설연구본부, bsk@kict.re.kr(교신저자)

※ 본 연구는 국토교통부 건설사업정보화(CALS) 사업으로 수행되었습니다(20240001-001, 24 건설사업정보화 사업기획 및 관리).

# Pointwise Convolutional Layer용 하드웨어 가속기 설계

이종윤\*, 권소영\*\*, 이창용\*\*, 변상봉\*\*\*, 이용환\*\*\*\*

## Design of Hardware Accelerator for Pointwise Convolutional Layer

Jong-Youn Lee\*, So-Young Kwon\*\*, Chang-Yong Lee\*\*, Sang-Bong Byun\*\*\*,  
and Yong-Hwan Lee\*\*\*\*

본 논문에서는 하드웨어 경량화 기법의 하나인 Pointwise Convolution 연산을 효과적으로 수행하기 위한 하드웨어 설계 방안을 제안한다. Pointwise Convolutional Layer는 Convolutional Neural Network (CNN)[1]에서 필터 크기가  $1 \times 1$ 인 특수한 형태의 Convolution 연산을 수행하는 계층으로, 필터 개수를 조절해 차원을 축소함으로써 네트워크의 계산 복잡도를 낮추고 모델 크기와 연산 효율성을 개선할 수 있다는 장점을 가진다[2]. 이러한 연산을 하드웨어에서 효과적으로 구현하기 위해, 본 논문은 주요 기능을 수행하는 모듈들을 기반으로 한 하드웨어 아키텍처를 설계하였다. 첫 번째 모듈은 입력되는 특징 맵을 메모리에 저장하고, 두 번째 모듈은 메모리에서 채널 단위로 값을 읽어와 채널 기반의 연산을 가능하게 한다. 마지막으로, 세 번째 모듈은 읽어들인 특징 맵 값과 필터값을 곱하여 연산을 수행하는 역할을 한다. 이와 같은 하드웨어 구현은 소프트웨어의 직렬 처리 방식과 달리 병렬 처리의 이점[3]을 최대한 활용할 수 있어, 실시간 데이터 처리 시스템에서도 높은 처리 성능을 제공할 수 있다.

### Key words

CNN, pointwise convolution, artificial intelligence, pipelining

### 참고문헌

- [1] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition", IEEE, vol. 86, pp. 2278-2324, 1998.
- [2] Andrew, G., and Zhu Menglong. "Efficient convolutional neural networks for mobile vision applications." Mobilenets 10 (2017): 151.
- [3] H. Zhang, L. N. Zhao, and X. F. Dai, "Hardware Design of a Pipeline Inspection Robot," Applied Mechanics and Materials, vol. 151, pp. 116-120, 2012.

---

\* 금오공과대학교 일반대학원 반도체시스템공학과 석사과정, ljjun0807@naver.com

\*\* 금오공과대학교 일반대학원 전자공학과 박사과정, papaya4040@kumoh.ac.kr, lcy5200@kumoh.ac.kr

\*\*\* 구미전자정보기술원 지능형디바이스연구센터, sbbyun@geri.re.kr

\*\*\*\* 금오공과대학교 전자공학부, 교신저자, yhlee@kumoh.ac.kr

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0017011, 2024년 산업혁신인재성장지원사업)



# 언어경관 이미지의 어노테이션 오류 유형화 및 요인 탐색

홍성연\*, 김여림\*\*

## Exploring annotation errors and influencing factors in linguistic landscape images

Seong-Yun Hong\* and Yeorim Kim\*\*

지난 10여년 간 빠르게 발전한 컴퓨터 비전 기술은 도시의 여러 분야를 탐색하고 연구하는데 폭넓게 활용되고 있다. 간판이나 공공 표지판 등 특정 장소에서 보여지는 언어적 요소에 관한 언어경관 연구에서도, 거리 시점 영상에 컴퓨터 비전 기술을 적용하여 연구의 공간적, 시간적 범위를 확장하고자 하는 시도가 이어지고 있다 [1-3]. 이러한 기술을 활용하여 신뢰할 수 있는 연구 결과를 얻기 위해서는 양질의 학습 데이터를 충분히 확보하는 것이 중요하다. 그러나 거리 시점 영상에 담긴 수많은 객체 중 언어경관 연구에 필요한 객체를 찾아 주석을 달아주는 어노테이션(annotation) 과정에서는 다양한 오류가 발생할 수 있다. 어노테이션 과정에서 발생하는 오류는 학습 데이터의 품질을 낮추고, 이는 다시 연구 결과의 정확성에 직접적인 영향을 주게 된다. 따라서 본 논문에서는 어노테이션 과정에서 나타나는 오류를 체계적으로 유형화하고, 오류 발생에 영향을 미친 주요 요인들을 탐구하고자 한다. 본 논문은 서울시를 대상으로 수집한 12,000여 장의 거리 시점 이미지를 직접 어노테이션하는 과정에서 발생한 오류를 분석 대상으로 한다. 분석 결과, 이미지의 해상도나 언어적, 시각적 복잡성도 오류 발생에 영향을 미치지만, 어노테이션 작업을 수행한 사람에 따른 차이도 확인할 수 있었다. 본 연구 결과는 어노테이션 과정의 오류 발생을 줄이고, 학습 데이터의 품질을 높이기 위한 기초 자료를 제공한다는 점에서 의미가 있다. 이를 토대로 양질의 언어경관 학습 데이터가 누적되고 이를 통해 컴퓨터 비전 모델의 정확성이 개선된다면, 향후 관련 연구의 활성화에 기여할 수 있을 것이다.

### Key words

annotation errors, computer vision, training data, linguistic landscape

### 참고문헌

- [1] Hong, S.-Y. (2020). Linguistic landscapes on street-level images. *ISPRS International Journal of Geo-Information*, 9(1), 57. <https://doi.org/10.3390/ijgi9010057>
- [2] Hong, S.-Y. (2022). Exploring ethnic segregation in linguistic landscapes. Presented at *12th Asian Conference in Regional Science*, Ulsan, Republic of Korea.
- [3] Hong, S.-Y., Kim, Y., & Lee, Y. (2023). Ethnic segregation on linguistic landscapes. *Environment and Planning B: Urban Analytics and City Science*, 50(7), 1916-1931. <https://doi.org/10.1177/23998083221150240>

---

\* 경희대학교 이과대학 지리학과, syhong@khu.ac.kr (교신저자), \*\*경희대학교 이과대학 지리학과, erin9906@khu.ac.kr

※ This work was supported by a grant from Kyung Hee University in 2023 (KHU-20230916).

# 자율주행차량의 객체 인식 기술 동향

박병주\*, 김요셉\*, 문일영\*\*

## Trends in Object Recognition Technology for Autonomous Vehicles

Byeong Ju Park\*, Yosep Kim\*, and Il-Young Moon\*\*

### 요약

자율주행차의 객체 인식 기술은 자율주행 시스템의 필수적 요소로 자리 잡고 있으며, 특히 카메라 기반 객체 인식은 풍부한 시각 정보를 제공함으로써 차량, 보행자, 도로 표지판 등을 정확하게 인식하는 데 중요한 역할을 한다. 본 논문에서는 자율주행차의 카메라 객체 인식 기술의 최근 동향과 데이터 증강, 전이 학습, 다중 센서 융합 등의 기법을 통해 성능을 향상시키는 방법들을 살펴본다. 또한, 다양한 환경 조건에서의 안정적인 인식을 위한 기술적 접근을 논의하며, 앞으로의 연구 방향에 대해 제시한다.

### Abstract

Object recognition technology for autonomous vehicles is an essential component of autonomous driving systems, particularly camera-based object recognition, which plays a crucial role in accurately detecting vehicles, pedestrians, road signs, and other elements by providing rich visual information. This paper analyzes recent trends in camera-based object recognition technology for autonomous vehicles, addressing various challenges and exploring methods to enhance performance through data augmentation, transfer learning, and multi-sensor fusion. Additionally, it discusses technical approaches to achieve stable recognition under diverse environmental conditions and presents future research directions.

### Key words

Autonomus veicles, Object recognition, Camera-based detection, Deep learning, Data augmentation, Multi-sensor fusion

### 1. 서론

자율주행 기술은 미래 교통의 패러다임을 변화시키고 있으며, 차량이 인간의 개입 없이 스스로 주행

할 수 있도록 하는 다양한 기술적 혁신이 요구된다. 이러한 자율주행 기술의 핵심 요소 중 하나는 차량 주변 환경을 정확하게 이해하고 대응하는 능력으로, 이를 위해 센서와 인공지능 기술의 융합이 필수적

\* 한국기술교육대학교 컴퓨터공학과, bj4906@koreatech.ac.kr, kys1479@koreatech.ac.kr

\*\* 한국기술교육대학교 컴퓨터공학과, iymoon@koreatech.ac.kr

※ 본 결과물은 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다.(2021RIS-004)

이다. 자율주행차는 주행 중 도로의 상황을 인식하고, 이를 바탕으로 경로 계획 및 충돌 회피와 같은 다양한 결정을 내려야 한다. 이를 위해 여러 종류의 센서가 사용되며, 그중에서도 카메라를 활용한 객체 인식 기술은 높은 해상도의 시각 정보를 제공함으로써 차량, 보행자, 도로 표지판, 교통 신호 등을 정확하게 인식하는 데 중요한 역할을 한다 [1, 2, 3, 6].

카메라는 상대적으로 저렴하고 많은 양의 정보를 제공할 수 있다는 이점 덕분에 자율주행 시스템에서 중요한 센서로 자리잡고 있다. 예를 들어, 라이다(LiDAR)나 레이더와 비교했을 때 카메라는 시각적으로 해상도가 높아 도로 표지판이나 교통 신호의 색상, 텍스트 등을 명확히 인식할 수 있다. 이러한 이유로 카메라 기반 객체 인식 기술은 자율주행 차량에서 필수적인 요소로 여겨지며, 특히 딥러닝 기술과의 결합을 통해 그 정확성과 효율성이 빠르게 발전하고 있다 [1, 3].

최근 몇 년간 딥러닝을 이용한 객체 인식 기술의 발전은 자율주행차의 카메라 인식 능력을 크게 향상시켰다. 합성곱 신경망(CNN, Convolutional Neural Network)의 발전과 더불어, 객체 인식 모델인 YOLO(You Only Look Once) [1], Faster R-CNN [2], SSD(Single Shot MultiBox Detector) [3] 등이 실시간 인식 성능과 높은 정확도를 달성하며 주요한 기술로 자리 잡았다. 하지만 자율주행 차량은 단순한 도로 환경뿐만 아니라 복잡하고 역동적인 상황에서도 안정적으로 작동해야 하며, 다양한 기후나 조명 조건에서 일관된 성능을 유지해야 한다. 이러한 이유로 데이터 증강(Data Augmentation), 전이 학습(Transfer Learning), 강화 학습(Reinforcement Learning) 등 다양한 접근법이 도입되고 있으며, 생성적 적대 신경망(GAN, Generative Adversarial Network)을 활용한 데이터 증강 또한 악조건에서의 성능 개선을 위해 주목받고 있다 [4].

본 논문에서는 자율주행차의 카메라 객체 인식 기술의 현재 동향과 이러한 기술들이 해결해야 할 도전 과제들을 살펴보고, 앞으로의 연구 방향에 대해 논의하고자 한다. 특히 카메라 객체 인식 기술의 발전과 함께 다중 센서 융합(Multi-sensor Fusion) 기

술이 자율주행 시스템의 안전성과 신뢰성을 높이는 데 어떤 역할을 하고 있는지에 대해서도 집중적으로 다룰 것이다 [5].

## II. 본론

카메라를 활용한 객체 인식 기술은 크게 두 가지로 분류된다. 전통적인 컴퓨터 비전 접근법과 딥러닝을 기반으로 하는 접근법이다. 초기의 객체 인식은 에지 검출, 색상 분할 등의 전통적인 방법을 사용하여 개발되었지만, 이는 복잡한 도로 상황에서의 인식 성능에 한계가 있었다. 최근에는 딥러닝 기술, 특히 합성곱 신경망(CNN, Convolutional Neural Network)이 도입되면서 객체 인식의 정확성과 안정성이 크게 향상되었다 [2]. 딥러닝 모델 중에서는 YOLO(You Only Look Once) [1], Faster R-CNN [2], SSD(Single Shot MultiBox Detector) [3]와 같은 모델들이 주로 사용되며, 이들은 실시간 성능과 높은 인식률을 동시에 달성하는 데 중점을 둔다 [6].

최근의 연구들은 다양한 환경 조건에서 안정적인 객체 인식을 위해 데이터 증강(Data Augmentation), 전이 학습(Transfer Learning), 그리고 강화 학습(Reinforcement Learning)을 이용한 기술들을 탐구하고 있다. 특히, 날씨나 조명 변화와 같은 악조건에서도 인식 성능을 유지하기 위해 생성적 적대 신경망(GAN, Generative Adversarial Network)을 사용한 데이터 증강이 주목받고 있다 [4]. 이 외에도, 라이다(LiDAR)와 레이더와 같은 센서 데이터를 카메라 영상과 융합하여 객체 인식의 신뢰성을 높이는 다중 센서 융합(Multi-sensor Fusion) 기술도 활발히 연구되고 있다 [5].

자율주행차 객체 인식의 주요 도전 과제는 다양한 도로 환경에서의 안정적인 인식과 실시간 처리 요구를 충족시키는 것이다. 다양한 기후 조건, 복잡한 도시 환경, 그리고 예측하기 어려운 도로 상황은 객체 인식 모델의 성능에 큰 영향을 미친다. 또한, 자율주행차는 안전성 확보를 위해 실시간으로 수많은 객체를 정확히 인식해야 하므로 모델의 경량화와 최적화 또한 중요한 과제로 떠오르고 있다.

### III. 결 론

카메라를 이용한 객체 인식 기술은 자율주행차의 핵심 요소로, 딥러닝의 발전과 함께 그 중요성이 더욱 커지고 있다. 다양한 환경에서의 인식 성능 향상과 실시간 처리의 효율성을 높이기 위해 많은 연구가 진행되고 있으며, GAN을 통한 데이터 증강이나 다중 센서 융합과 같은 기술이 그 해결책으로 제시되고 있다. 앞으로는 더욱 효율적이고 안정적인 인식 모델 개발을 위해 지속적인 연구와 기술 발전이 필요하다. 또한, 다양한 환경 조건에 적응할 수 있는 모델의 개발이 자율주행 기술의 상용화에 중요한 역할을 할 것이다.

위에서 언급된 도표와 데이터를 통해 자율주행차 객체 인식 기술의 현황을 보다 시각적으로 이해할 수 있으며, 이를 바탕으로 향후 연구 방향을 명확히 할 수 있을 것이다.

### 참 고 문 헌

- [1] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
- [2] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems*, 28, 91-99.
- [3] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In *European Conference on Computer Vision* (pp. 21-37).
- [4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, Y. Bengio et al., "Generative Adversarial Networks," arXiv preprint arXiv:1406.2661, 2014..
- [5] Zhao, H., Qi, X., Shen, X., Shi, J., & Jia, J. (2017). ICNet for Real-Time Semantic Segmentation on High-Resolution Images. In *Proceedings of the European Conference on*

*Computer Vision (ECCV)*, 405-420.

- [6] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834-848.

# 오픈 데이터 수집 및 활용을 위한 KRDC 프레임워크 기반 OAI-PMH 워크플로우 개발

김다솔\*, 조민희\*, 송사광\*, 이미경\*, 임형준\*

## Development of OAI-PMH Workflow Based on KRDC Framework for Open Data Collection and Utilization

Dasol Kim\*, Minhee Cho\*, Sa-Kwang Song\*, Mikyong Lee,\* and Hyung-Jun Yim\*

빅데이터, AI 기술의 발전으로 인해 데이터 중심(data-intensive) 연구개발로 환경이 변화하면서 연구에 필요한 데이터 확보의 중요성이 더욱 커지고 있다. 하지만, 다량의 실제 데이터를 확보하는 일은 여전히 어려운 과제이며 데이터 확보 방법 또한 제한적이다. 본 논문에서는 오픈 데이터를 연구에 활용하기 위해 OAI-PMH 프로토콜 기반의 데이터 하베스터를 개발하고 실제 오픈 데이터를 대상으로 수집, 연계, 활용을 위한 시나리오를 설계한다. 제안 시나리오를 연구소프트웨어 상호운용 플랫폼인 KRDC(Korea Research Data Commons) 프레임워크에 패키징된 앱으로 개발한다. KRDC 프레임워크는 연구소프트웨어와 분석 환경을 함께 제공하는 플랫폼으로 소프트웨어를 도커이미지인 앱으로 제작하고, 앱을 조합한 워크플로우로 동작시킨다[1]. 소프트웨어를 앱 형태로 프레임워크에 배포하여 공유, 활용이 가능하며 이를 실행단위인 워크플로우로 구성하여 다양하게 응용이 가능하다. 데이터 수집에 활용한 OAI-PMH(Open Archives Initiative Protocol for Metadata Harvesting)는 데이터 리파지토리 간 데이터 교환을 위한 표준 프로토콜로 본 논문에서는 이를 활용하여 데이터 하베스터를 개발한다.

본 연구에서는 실제 오픈 데이터를 수집, 활용하기 위한 워크플로우 구성을 설계하고 총 세 가지 앱을 구현하여 KRDC 프레임워크에 배포한다. 설계한 워크플로우 구성은 다음과 같다. (1) OAI-PMH harvester 앱을 활용하여 카탈로그를 수집하며, OAI-PMH 프로토콜 endpoint에서 데이터셋을 100개 단위로 XML 파일에 저장한다. (2) Catalog parser 앱을 연결하여 필수 메타데이터만 전처리하며, 해당 앱에서는 XML 파일에서 지정된 메타데이터를 파싱하고 이를 JSON 리스트로 변환한다. (3) Dataset downloader 앱은 JSON 파일에서 실제 데이터를 다운받아 사용자가 지정한 워크스페이스 출력경로에 저장한다. 사용자는 OAI-PMH 워크플로우를 통해 자신의 디렉토리에 다운받은 데이터를 사용하여 데이터 시각화, AI 모델 학습과 같이 다양하게 활용할 수 있다. 구현한 워크플로우의 동작과 유효성을 검증하기 위해, 유럽의 EGI DataHub와 Datasite를 대상으로 평가하며 정상적으로 카탈로그 수집, 전처리, 데이터셋 다운로드가 가능함을 확인하였다. 본 논문에서는 연구소프트웨어 공유 환경을 제공하는 KRDC 프레임워크를 기반으로 실제 데이터 수집, 활용을 위한 워크플로우를 개발하였다. 제안한 결과물을 활용하면 연구에 필요한 오픈 데이터를 확보할 수 있으며, KRDC 프레임워크 상에서 다양한 워크플로우 구성에 활용 가능하다. 향후 연구로는 구현한 워크플로우를 활용하여 다양한 응용 사례를 개발할 예정이다.

### Key words

Data Harvesting, Open Data, Research Software, Korea Research Data Commons

### 참고문헌

- [1] D. Kim, et al., "Development of Korea Research Data Commons Framework for Resource Federation Between Multi-clusters", *Database Research*, Vol. 40, No. 2, pp. 50-65, Aug. 2024.

\*한국과학기술정보연구원 연구데이터공유센터, {dskim, mini, esmallj, jerryis, hjyim}@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24L1M0C3-01)

# 메시지 큐 기반의 AMI Framework 설계

김장현\*, 김준하\*\*, 김광영\*\*\*

## Message Queue-Based AMI Framework Design

Kim JangHyun\*, Kim JunHa\*\*, and Kim GwangYeong\*\*\*

### 요약

AMI 기기의 S/W는 사업차수별, 제조사별 제품이 다양하여 효율적인 운영/유지 관리에 어려움이 있다. 이 논문에서는 여러 AMI 설비에 공통적으로 적용할 수 있는 서비스 확장형의 AMI Framework 도입을 제안한다. AMI Framework는 오픈소스 메시지 큐를 이용하여 데이터 전달의 신뢰성 확보와 비동기 메시지 처리가 가능하다. 메시지 큐 기반의 AMI Framework 적용을 통해 대규모 AMI 시스템의 안정적 운영과 유지관리가 가능하다.

### Abstract

The S/W of AMI devices has difficulty in efficient operation/maintenance as products vary by business order and manufacturer. In this paper, we propose the introduction of an AMI framework that can be applied to several AMI facilities in common. AMI framework can secure the reliability of data delivery and process asynchronous messages by using an open-source message queue. The stable operation and maintenance of large-scale AMI systems are possible through the application of the message queue-based AMI framework.

### Key words

AMI, framework, message queue, maintenance

### 1. 서론

한전은 정부의 AMI 지능형 전력망 구축 기본계획[1]에 발맞추어 2024년까지 전체 2,250만 호 모든 고객에 대하여 AMI 통신망 구축을 완료한다는 계획하에 연차 사업을 추진 중에 있다. 그러나 수년에 걸쳐 사업을 추진하며 AMI 기기는 연차별 도입 제품의 규격이 상이

하고 다수 업체·다종의 제품을 도입한 결과 효율적인 현장 설비(전력량계/모뎀/DCU) 운영 및 유지관리에 어려움을 겪고 있다. 본 논문에서는 다양한 AMI 기기에 공통적으로 적용할 수 있고 유연한 확장이 가능한 메시지 큐 기반의 표준 AMI Framework 설계에 대한 내용을 제안한다.

\*한전KDN, jhkim\_5049@kdn.com, \*\*한전KDN, junhakim.17@kdn.com, \*\*\*한전KDN, kimholy.0706@kdn.com

## II. AMI Framework 설계

AMI Framework란 AMI S/W 개발에 공통적으로 사용되는 구성요소와 아키텍처를 일반화하여 손쉽게 상위 어플리케이션을 구현할 수 있도록 여러 가지 기능을 제공해 주는 반제품 형태의 소프트웨어 시스템이다.

전력량계나 통신모뎀의 데이터를 수집하여 상위 AMI 서버로 전송하는 역할을 하는 DCU 장치는 제조사별/통신방식별로 다양하게 운영되어야 하므로 OS와 HW에 비종속적이고 개방적이며 AMI 이외의 다양한 신규 어플리케이션(DER 감시제어, 상판별 등) 연동이 가능한 서비스 확장형의 표준 AMI Framework를 도입하였다.

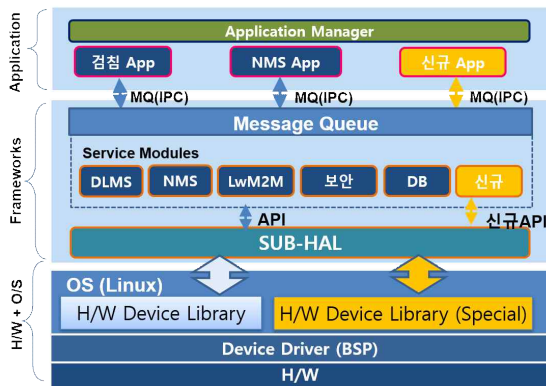


그림 1. AMI Framework 구성도  
Fig. 1. AMI Framework Diagram

AMI Framework는 다종의 OS와 HW에 적용할 수 있는 하드웨어 추상화 계층 (SUB-HAL), AMI 기능별 요소를 각각 서비스 단위로 구현한 서비스 모듈, 모듈 간 통신을 위한 메시지 큐(MQ), 그리고 모듈을 관리하는 Manager로 이루어져 있으며 상위의 어플리케이션은 AMI Framework를 활용하여 AMI 비즈니스 로직을 구현한다.

DLMS 모듈은 전력량계의 통신 프로토콜 패킷 처리 및 분석 서비스를 제공하고, NMS 모듈은 AMI 사업 차수별로 서로 다른 모뎀과의 통신 인터페이스 대응 서비스를 제공하고, LwM2M[2,3] 모듈은 상위 AMI 서버와 통신

LwM2M 프로토콜 패킷 처리 및 분석 서비스를 제공하고, Security 모듈은 KCMVP 기반의 DCU 기기 인증, 인증서 관리 및 모뎀 인증서 검증 서비스를 제공하고, DB 모듈은 검침 데이터 및 NMS 데이터를 저장/관리하는 서비스를 제공한다.

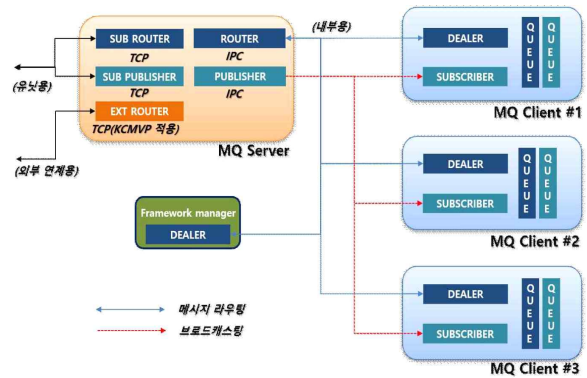


그림 2. Message Queue 설계도  
Fig. 2. Message Queue Design

AMI Framework는 기능별로 서비스 모듈을 분화하고 모듈 간 통신은 메시지 큐를 이용하여 표준화된 API를 정의함으로써 신규 모듈에 대한 확장성과 유연성을 제공한다. 메시지 큐는 MQ Server와 MQ Client로 나누어 기능과 역할을 아래와 같이 설계했다.

MQ Server는 MQ Client간 데이터 전송을 위해 프로세스 상태로 존재하며, MQ Client가 생성한 데이터의 중계역할을 수행한다. 데이터 유실방지와 원활한 데이터 처리를 위해 내부 큐를 사용하며 프로세스간 빠른 데이터 전송을 위해 IPC 방식의 인터페이스를 사용한다. 단, 외부(K-DCU 유닛 등) 데이터 연계를 위해 TCP로 동작할 수 있다.

MQ Client는 데이터를 생성 및 패키징하여 MQ Server로 전송하는 역할을 수행하고, 제어 명령 수신 시 이를 처리하는 일을 수행한다. 또한, 내부 큐를 사용하여 수신 메시지를 순차적으로 처리할 수 있으며 비동기 메시지 처리가 가능하다.

MQ Server의 Router는 목적지 프로세스의 주소를 확인하여 수신된 메시지를 UniCast로

전달하는 역할을 수행하며 Publisher는 수신된 메시지를 등록된 전체의 MQ Client에게 BroadCast로 전달하는 역할을 수행한다. Framework manager는 MQ Server를 이용하여 전체 모듈을 관리 및 제어하는 메시지를 보낼 수 있다.

### III. 결 론

AMI 기기 운영 S/W 표준화로 운영관리 및 유지보수 비용을 절감할 수 있으며 정부의 계 시별 요금제(TOU) 보급 확대 등의 정책에 대비해 고신뢰 AMI 인프라를 구축할 수 있다. 또한 AMI 표준 S/W 아키텍처 및 S/W 표준화 응용 기술을 내재화하여 제조사 의존성을 탈 피하고 자체 운영/유지보수 역량을 확보할 수 있다. 향후 AMI 기기 S/W 원격 업데이트를 통해 2,250만호 대규모 AMI 시스템의 안정적 운영과 관리 효율성을 제고할 수 있을 것이다.

### 참 고 문 헌

- [1] “제3차 지능형 전력망 기본계획 (2023~2027)”, 산업통상자원부
- [2] OMA-TS-LightwightM2M\_Core-V1\_1-20190617-A
- [3] OMA-TS-LightwightM2M\_Transport-V1\_1-20180710

-A



# 상호운용성 검증을 위한 DataHub와 KRDC 연계 방안

송사광\*, 김다솔\*, 조민희\*, 이미경\*, 임형준\*

## Interlinking Methodology between DataHub and KRDC for Interoperability

Sa-kwang Song, Dasol Kim, Minhee Cho, Mikyoung Lee, and Hyoungjun Yim

빅데이터와 인공지능이 전 세계적으로 연구와 산업 분야에 보편화되면서, 오픈 사이언스에 대한 관심과 지원이 증가하고 있다. 특히 연구데이터와 연구 자원의 공유와 활용을 촉진하는 여러 커먼즈 시스템이 개발 및 운영되고 있다. 그러나, 기관 및 분야를 넘어서는 연구데이터 커먼즈 시스템 간의 공유/활용 사례 개발은 여전히 도전적인 주제이기 때문에, EOSC, RDA, CODATA, WDS 등의 글로벌 기구 및 조직 등에서 기술적/의미적/조직적/법적 상호운용성(Interoperability)을 주제로 많은 연구와 논의가 진행 중에 있다.

본 연구에서는 유럽의 EGI(European Grid Initiative)와 KISTI(Korea Institute of Science Technology Information)의 커먼즈 시스템 간의 상호운용성 검증을 위한 연계 시나리오를 마련하고 이를 검증할 수 있는 연계 사례를 개발함으로써 기술적 관점에서의 상호운용성 검증을 수행하였다. DataHub는 EGI에서 개발한 연구데이터 허브 플랫폼으로 연구데이터의 공유활용을 용이하게 하기 위한 통합 인증체계, 통합 검색, 파일 브라우징, 웹기반 다양한 사용자 편의성 등을 제공하고 있다. KISTI는 다양한 연구데이터 리소스(데이터, SW, 서비스, 인프라 등)의 공유활용을 지원하기 위한 플랫폼으로 2021년부터 KRDC(Korea Research Data Commons)를 개발해 오고 있다. 상호운용성을 확보하기 위해서는 기술적/의미적/조직적/법적 계층에서의 각 방법론이 정립되어야 하나, 본 논문에서는 기술적인 관점에서의 연계 방법론을 소개한다. 특히, 통합 인증 및 데이터와 소프트웨어 연계 시나리오에 집중하여 소개한다. 연계 시나리오는 두 커먼즈 간의 통합인증 기능, DataHub 공유데이터의 KRDC 상 검색 기능 제공, DataHub의 연구데이터 실시간 다운로드 기능, 연구데이터의 실시간 분석/가시화 기능, Jupyter/VM/워크플로우 환경 지원 기능 등을 포함한다. 연구데이터 커먼즈 시스템간의 연계 활용을 위해서 가장 기본적으로 해결되어야 할 문제는 인증 체계(Authentication and Authorization Infrastructure)의 연동인데, 이는 한 시스템의 사용자가 다른 시스템 사용을 위한 인증(Authentication) 및 인가(Authorization)를 지원하는 기능이다. 이를 위해, KRDC와 DataHub는 SAML, OAuth, OpenID Connect 등의 대표적인 표준 프로토콜을 지원함으로써 연계를 지원하였다. DataHub의 공유데이터에 대한 정보(메타데이터) 수집을 위해 OAI-PMH 표준 프로토콜을 이용하였고, 수집된 정보의 통합 검색(Findable)을 지원하였다. 사용자가 원하는 연구데이터의 다운로드(Accessable)와 다운로드한 데이터의 분석/가시화를 위해서는 KRDC 앱을 개발/공개(Reusable)하고, 앱들을 일련의 프로세스로 연결하여 워크플로우로 구성하여 실행 및 결과를 확인할 수 있도록 하였다. 이를 통해 DataHub와 KRDC가 상호운용(Interoperable) 가능한 환경에서 사용자의 태스크를 수행할 수 있는 환경을 제공할 수 있었다. 이로써, FAIR(Findable, Accessable, Interoperable, Reusable) 원칙을 지원하는 KRDC 체계를 구현하였다.

향후 연구로는, 기술적 계층을 넘어, 의미적/조직적 계층의 상호 운용성 확보를 위한 KRDC 체계의 구축 방법론 개발과, 그 방법론의 검증을 위한 실제 운용 사례를 구축할 예정이다.

### Key words

Interoperability, Reseach Data Commons, Korea Research Data Commons, KRDC

### 참고문헌

- [1] 조민희외 3명, “연구소프트웨어 공유 활용을 위한 학술 인프라 설계”, 한국경영과학회 학술대회논문집, 2023.

\* 한국과학기술정보연구원 {esmallj, dskim, mini, jerryis, hjvim}@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24L1M0C3-01)

# 가상 소방 훈련에서의 다중 인원 동시 훈련에 따른 안전 모니터링 기술에 관한 연구

김치중\*, 장민혁\*

## A Study on the safety monitoring technology of multi-human simultaneous training in virtual fire training

Chijung Kim\*, Minhyuk Chang\*

최근 소방 훈련 분야에 화재 예방, 진압, 대피 훈련 등의 다양한 상황을 재현한 가상 교육/훈련 콘텐츠가 활용되고 있다. 이러한 가상훈련 콘텐츠 중 다중 훈련의 경우 VR HMD를 착용한 후에 제한된 공간 내에서 다수 인원이 동시에 훈련을 수행하기 때문에, 사용자 간의 충돌이나 훈련 공간 영역을 벗어나는 등의 안전사고 이슈가 존재한다. 본 연구는 다중 사용자의 가상 소방 훈련에서 실시간 위치 추적 및 모션캡처 기술을 활용하여 충돌 및 이탈 방지 안전 모니터링 기술에 관한 연구이다. 안전 모니터링은 광학식 위치 추적 카메라에서 센싱된 사용자의 위치 좌표와 전신 모션캡처를 이용한 관절 데이터를 실시간으로 사용한다. 첫 번째 충돌 모니터링은 관절 좌표를 활용하여 사용자-사용자 간의 신체 충돌 상황 중 팔을 휘젓는 상황과 이동 반경에 대하여 '한국인 표준 인체 치수'의 팔 길이와 보폭 등을 적용하여 안전 감지 거리 기준을 적용하였다. 또한 공간 내에 배치된 실제 물리 구조물-사용자 간의 충돌 상황도 감지된다. 두 번째 이탈 모니터링 기술은 사용자 위치 데이터를 기반으로 체험 공간을 벗어나지 않도록 안전 영역을 설정하고 이탈 여부를 감지한다. 세 번째, 모니터링된 값들은 사용자가 체험 중인 VR 화면에 경고 알람 프로토콜이 적용되어 사용자의 충돌/이탈이 감지될 경우, 위험 경고 화면 및 사운드가 송출된다. 경고 알람은 충돌/이탈 상황별 문구, 충돌 예상 방향, 충돌체와의 현재 거리에 대한 그래픽 경고와 감지된 거리 및 간섭 단계별 경고 사운드를 볼륨 변화로 제공한다. 또한 실제 공간을 가상공간으로 구성할 수 있도록 저작도구를 개발하여 맵 저작, 거리/영역 등 설정, 3D 데이터 연동 등의 저작기능을 제공한다. 본 기술은 Unreal 그래픽 엔진 에셋으로 제공되어 타 콘텐츠에서도 활용이 가능하며, 다중 인원(최대 10명)에 대한 동시 위치/모션 데이터 기반의 충돌 및 이탈 상황 감지를 통한 안전 모니터링 시스템 활용에 대한 효과성을 소방 훈련 대상으로 검증할 예정이다.

### Key words

Virtual fire training, Safety monitoring, Collision/Departure Alert Notification

### 참고문헌

- [1] Min-Hye Lee, et al. "Reminder module design to prevent collision accidents while wearing HMD", Journal of the Korea Institute of Information & Communication Engineering 26.11 (2022).
- [2] HanWolKim, et al. "Collision Avoidance Application for VR User.", Proceedings of KIIT Conference. 2022.

---

\*한국전자기술연구원, cjkim@keti.re.kr, minhyuk@keti.re.kr(교신저자)

※ 본 논문은 소방청의 실감기반 첨단소방훈련체계연구사업(200008389)의 지원을 받아 작성함

# 사용자 제어 의도(SSVEP) 검증을 위한 VR 콘텐츠 개발 연구

김서영\*, 김치중\*, 장민혁\*

## A Study on VR Contents Development for User Control Intention Verification(SSVEP)

Seoyoung Kim\*, Chijung Kim\*, Minhyuk Chang\*

본 연구는 사용자 제어 의도에 관한 SSVEP(Steady State Visually Evoked Potentials) 분석 연구에 활용할 수 있는 VR 콘텐츠 개발에 관한 것이다. SSVEP는 특정 주파수로 깜박이는 시각 자극을 사용자가 응시할 때, 후두엽에서 동일한 주파수의 뇌파 성분이 검출되는 것으로서 사용자 제어 의도 파악에 많이 사용된다. 본 연구의 VR 콘텐츠는 SSVEP 분석에 필요한 다양한 시각 자극을 사용자가 손쉽게 편집하고, 검출된 SSVEP에 따라 가상 캐릭터를 제어할 수 있는 VR 콘텐츠이다.

VR 콘텐츠 내 시각 자극은 Checkerboard(Contrast) 형태로 제시되며, SSVEP 분석 연구 특성에 맞게 사용자가 손쉽게 편집할 수 있도록 제공한다. 시각 자극 편집은 시각 자극의 개수(최대 6개), 주파수(VR HMD의 최대 주사율 90Hz 이내 소수점까지 입력), 모양(원형, 삼각형, 사각형, 오각형 등), 투명도(0~100%)를 편집할 수 있도록 제공한다. 특정 주파수가 입력된 시각 자극은 VR 내 캐릭터의 제어와 관련된 6종의 인터페이스(앞/뒤/좌/우/시점 변경/나가기)로 활용되며, 미로 찾기 VR 콘텐츠에 적용하여 사용자 제어 의도 검증에 활용될 수 있도록 개발하였다. 미로 찾기 VR 콘텐츠는 시점 변경을 통하여 VR 화면에 캐릭터 시점을 1인치 또는 3인칭으로 변경할 수 있으며, 미니맵을 통하여 캐릭터의 현재 위치와 방향을 확인하여 목적지 도달 미션을 제공한다.

본 연구로 개발 중인 시각 자극이 적용된 VR 콘텐츠를 활용하여 피험자들을 대상으로 제어 의도 분류 알고리즘의 정확도 평가 및 SSVEP 검증 연구를 진행할 예정이다.

### Key words

SSVEP, VR Interface, Control Intention, EEG frequency

### 참고문헌

- [1] Koo, Bonkon, et al. "Immersive BCI with SSVEP in VR head-mounted display." 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, 2015.
- [2] Hsu, Hao-Teng, et al. "Phase-approaching stimulation sequence for SSVEP-based BCI: a practical use in VR/AR HMD." IEEE Transactions on Neural Systems and Rehabilitation Engineering 29 (2021): 2754-2764.

---

\* 한국전자기술연구원, syoung@keti.re.kr, cjkim@keti.re.kr, minhyuk@keti.re.kr(교신저자)

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.RS-2023-00229074, 멀티모달 기반 Arousal-Valence 감정 전달 인터페이스 기술 개발)

# GPS 음영지역을 위한 고정밀 측위용 UWB 시스템 설계

김동언\*, 변상봉\*, 이성훈\*, 한동철\*, 조수현\*

## Design of a High-Precision Positioning UWB System for GPS Shadowed Areas

Dong-Eon Kim\*, Sang-Bong Byun\*, Sung hun Lee\*, Dong Cheul Han\* and Soo-Hyun Cho\*

본 논문에서는 GPS 음영지역에서 정밀 측위를 적용하기 위한 UWB 시스템을 설계한다. 고가도로의 밑이나 터널 또는 실내 환경은 GPS가 적용되지 않은 음영지역이다. GPS 음영지역에서도 작업량 관리 및 자동화 시스템 구축을 위해서 이동 로봇이나 작업자 등의 위치 정보가 필요한 경우가 많다. GNSS 시스템을 대체하여 GPS 음영지역에서 측위를 위한 시스템이 필요하다. 하지만, 라이다 등의 고가의 장비를 적용하기에는 고비용의 문제가 발생하고, 공사 및 건설 현장에 적용 시에는 작업의 진척에 따라서 지도를 재구축해야 하는 문제도 발생한다. 이러한 문제를 해결하기 위하여, UWB 시스템 구축을 통한 고정밀 측위 시스템을 제안한다[1][2][3]. UWB 시스템에서는 앵커와 태그 간의 메시지 충돌 및 간섭을 방지하기 위해서 액세스 관리가 필요하다. 따라서 본 논문에서는 턴 기반의 중앙 집중식 UWB 모듈을 활용한 양방향 거리 측정(TWR)의 메시지 시퀀스 시스템을 제안한다. 제안된 시스템은 앵커 마스터가 포함된 4개의 앵커와 태그로 구성된다. 이를 위해 딜레이의 제어와 구조화된 턴 테이킹을 도입하여 한 번에 하나의 장치만을 명확하게 정의하여, 설계된 순서에 따라 통신하도록 하여 문제를 해결한다. 태그는 앵커 마스터가 제공하는 Start turn 메시지로 시작점을 표시하여 거리 측정 시퀀스를 시작한다. 태그는 방송형 POLL 메시지를 보내고, 각 앵커는 메시지 충돌을 방지하기 위해 특정 전송 메시지의 딜레이를 지정 후 RESPONSE 메시지로 응답한다. 이러한 딜레이는 충돌을 피하고 앵커 간의 순차적 전송을 보장한다. 모든 응답이 수신되거나 타임아웃이 발생하면 태그는 앵커 마스터에게 END TRUN 메시지를 보내면서 거리 측정에 관한 턴을 종료한다. 이러한 거리 측정 시퀀스의 처리 후 새로운 태그 장치로 추가적인 거리 측정 주기를 위한 턴을 실행한다. 제안된 방법은 GNSS 정보의 획득이 어려운 GPS 음영지역에서 측위를 통한 이동로봇 및 작업자의 위치 인식이 X, Y축으로 구성된 평면상에서 15cm 이내의 정밀한 측위가 가능하다.

### Key words

UWB, TWR, Anchor master, message sequence

### 참고문헌

- [1] 최제익, 이재욱, 박소윤, 임지호, 신동원. (2024). UWB 기반 이동 앵커를 이용한 사람 추종 로봇 시스템. 한국지능시스템학회 논문지, 34(1), 74-80, 10.5391/JKIIS.2024.34.1.74.
- [2] 정인택, 장봉주. (2023). UWB를 이용한 실시간 실내 객체 측위 및 정량적 위치 정확도 검증 방법에 관한 연구. 한국산학기술학회 논문지, 24(12), 686-693, 10.5762/KAIS.2023.24.12.686.
- [3] 장병준. (2022). UWB 측위기술의 원리 및 동향. 한국전자과학회논문지, 33(1), 1-11.

\* 구미전자정보기술원 ICT디바이스연구센터, email : dekim@geri.re.kr(교신저자), sbbyun@geri.re.kr, leesh@geri.re.kr, cataegu07@geri.re.kr, shcho@geri.re.kr

※ 이 연구는 2024년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임(20017474)

# ScienceON의 과학기술 뉴스 수집 및 서비스 방안 연구

김진영<sup>†</sup>, 박정훈, 유수현

## A Study on ScienceON Science & Technology News Collection and Service Method

Jinyoung Kim<sup>†</sup>, Jeonghun Park, Suhyeon Yoo

ScienceON은 주로 과학기술 관련 연구·개발을 수행하는 연구자, 학생, 공무원, 기업 담당자 등이 사용한다. 주로 과학기술정보(논문, 특허, 보고서 등)를 열람하고 관련 서비스를 사용하기 위해 ScienceON에 접근한다. 그러나, 과학기술 분야 트렌드 파악을 위해 관련 분야의 핵심 뉴스를 빠르게 습득하고자 하는 요구를 반영하여 ScienceON에서는 최신 과학기술 뉴스를 수집하여 분야 트렌드 파악에 도움이 될 수 있는 뉴스들을 선별하여 서비스하고 있다.

ScienceON에서는 12대 국가전략기술(반도체·디스플레이, 이차전지, 첨단 모빌리티, 차세대 원자력, 수소 등)에 해당하는 뉴스들을 8개 국가(대한민국, 미국, 중국, 일본 등)의 51개 매체에서 수집하여 주 2회 서비스한다. 현행 서비스는 원내 담당자가 최근 트렌드, 조회 수, 기술 수준 등을 고려하여 대상 매체들을 검색하여 매 회 10건을 선별하여 제공했다.

이러한 방법은 타겟 뉴스 출처와 검색 키워드가 많아 뉴스 검색(탐색)에 많은 시간이 소요되며, 선정된 기사에 대한 요약 시 정확하지 않을 수 있다는 문제점이 있다. 그러므로 본 연구에서는 아래 그림 과 같이 해당 서비스 제공의 효율화와 정확도 향상을 위한 자동화 방안을 제안한다.

키워드 목록 작성 → 뉴스 자동 수집 → 키워드 및 매체 별 Top-K 결과 도출 → 자동번역/요약생성 → 10건 선택 → 최종 수정(요약 등) → 업로드

키워드 목록 작성 시 12대 국가전략기술 별 최근 트렌드에 맞는 키워드들로 작성하며, 뉴스 자동 수집은 키워드 목록에 있는 키워드를 대상 국가 언어로 번역하여 Google News RSS 피드를 활용하여 수집한다. 키워드/매체/국가 별 수집된 뉴스 결과에 대해 자동번역, 요약을 생성하여 담당자가 10건을 선택하여 최종 수정을 통해 업로드를 수행한다. 번역과 요약 생성은 ChatGPT API와 DeepL을 활용하고 있으나 추후 KISTI의 과학기술 특화 생성형 언어 모델인 KONI를 활용할 예정이다.

이를 통해 과학기술뉴스 제공을 위한 시간 절감과 정확도 향상을 얻을 수 있었으며, 업로드된 뉴스는 데이터 베이스에 저장되며 추후 ScienceON에서 제공되는 과학기술정보의 한 분류로 활용할 것이다.

추후 트렌드에 맞는 키워드 자동 선정, 핵심 뉴스 자동 선별에 대한 세부 연구를 추진할 예정이며 전 과정의 자동화를 통한 효율화 증대와 실질적인 사용자 요구에 맞는 뉴스를 선별할 수 있도록 정확도 향상에 대한 연구를 진행할 것이다.

Key words

ScienceON News, News Collection

\* 한국과학기술정보연구원 국가과학기술데이터본부 융합서비스센터, {jykim, jhpark<sup>†</sup>, yoosu}@kisti.re.kr

※ 본 연구는 한국과학기술정보연구원에서 수행하는 주요사업 '지능형 과학기술지식인프라 통합서비스 플랫폼 구축(K24L1M3C4-01)'의 지원을 받아 수행된 연구임.

# ScienceON의 서비스 및 기능 인큐베이팅 방안 연구

김진영, 박정훈<sup>†</sup>, 유수현

## A Study on the Service and Function Incubation Method of ScienceON

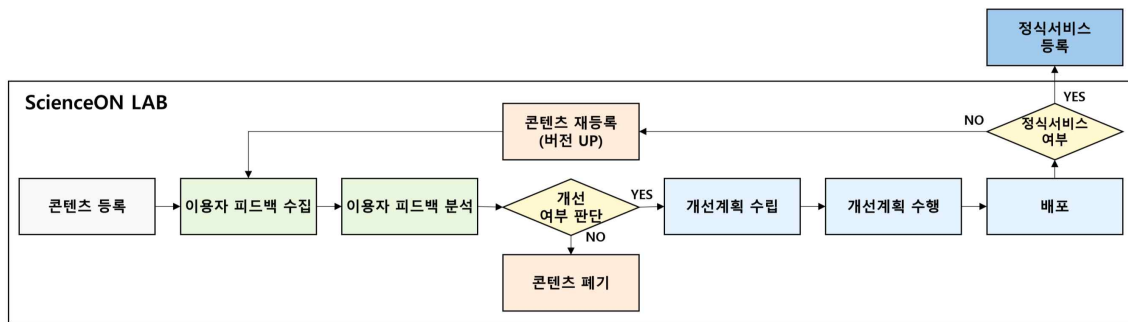
Jinyoung Kim, Jeonghun Park<sup>†</sup>, Suhyeon Yoo

ScienceON은 과학기술정보(논문, 특허, 보고서 등)와 관련된 다양한 서비스와 기능을 제공한다. 인공지능과 같은 다양한 기술을 발전과 연구 전주기에 활용할 수 있는 각종 서비스와 기능에 대한 요구가 발생하고 있으며, ScienceON에서는 이러한 요구에 맞는 다양한 서비스와 기능을 개발하여 제공한다.

서비스 운영 차원에서 다양한 창구를 통해 VOC(Voice of Customer)를 수렴하는데 주로 FGI(Focus Group Interview), NPS(Net Promoter Score), Q&A 등을 활용한다[1]. 이 방법들은 ScienceON 전체적인 의견을 수렴하는데는 크게 도움이 되나, 각 서비스와 기능의 세부 의견을 수렴하고 분석하기는 어렵다.

그러므로, ScienceON에서는 신규로 개발된 각 서비스와 기능에 대해 사용자가 직접 사용 후 의견을 직접 제시할 수 있는 창구인 ScienceON LAB을 운영하고 있다. 사용자 피드백은 편의성, 적절성, 활용성, 차별성, 종합 만족도에 대해 5점 척도로 점수를 부여할 수 있으며, 추가적으로 의견을 남길 수 있도록 설계했다. ScienceON LAB에서 다루는 콘텐츠는 서비스, 기능, 데이터셋으로 분류하고 있으며, 각 콘텐츠 분류에 맞게 다른 의미를 가질 수 있도록 항목을 정의했다.

수집된 사용자 피드백은 다음 그림에서와 같이 콘텐츠의 등록 시부터 수집/분석되며, 일정 기간 콘텐츠를 운영한 후 개선 여부에 대한 의사결정을 수행한 후 폐기/재등록/정식오픈으로 진행된다. 이는 ScienceON의 디지털 전환과 애자일개발을 통한 개발/운영의 효율화에 기여할 수 있다.



Key words

ScienceON, ScienceON LAB, Service Incuvation

[1] 이해진, 신재민, 배수영, 최희석, 이석형, “과학기술 지식인프라 통합서비스의 품질평가모델 연구,” Journal of Digital Contents Society, Vol. 22, No. 5, pp. 867-876, 2021.

\* 한국과학기술정보연구원 국가과학기술데이터본부 융합서비스센터, {jykim, jhpark<sup>†</sup>, yoosu}@kisti.re.kr

※ 본 연구는 한국과학기술정보연구원에서 수행하는 주요사업 ‘지능형 과학기술지식인프라 통합서비스 플랫폼 구축 (K24L1M3C4-01)’의 지원을 받아 수행된 연구임.

# 과학기술정보서비스 ScienceON 홈페이지의 사용성 평가 및 분석 연구

김진영, 박정훈, 유수현<sup>†</sup>

## A Study on the Usability Evaluation and Analysis of the ScienceON as a Science & Technology Information Service

Jinyoung Kim, Jeonghun Park, Suhyeon Yoo<sup>†</sup>

과학기술정보서비스인 ScienceON은 국내외 과학기술정보, 연구데이터, 각종 지식인프라를 통합하여 다양한 서비스를 제공하고 있으며 2019년에 오픈하여 현재까지 서비스되고 있다. ScienceON에서는 지속적인 사용자 의견수렴을 통해 개선을 추진하고 있으며 2019년부터 2022년까지의 VOC(Voice of Customer), NPS(Net Promoter Score)를 분석한 결과 '메인페이지가 복잡', 'UI/UX 불편', '용어 난해' 등의 의견이 수렴됐고, 불편사항들을 반영하여 2023년 11월에 개편된 홈페이지를 오픈했다.

본 연구에서는 개편된 홈페이지에 대한 사용성 평가 분석 결과와 도출된 개선 방안을 제시한다.

개편된 ScienceON 홈페이지에 대한 사용성 평가는 사용자 시나리오 중심 사용성 진단 및 분석을 통해 UI/UX의 개선점을 도출하고 적용함으로써 이용자들의 편의성, 접근성, 활용성 등을 강화하고 서비스의 질적 향상을 도모하기 위해 수행됐다.

시나리오는 개편된 ScienceON의 주요 서비스 및 기능 중 가장 많이 활용되는 '통합검색'과 타 학술정보서비스와 차별되는 '지식인프라'와 'MyON' 서비스를 대상으로 총 7개 Task(필수 4개, 선택: 4개)로 설계됐다. 조사대상은 ScienceON을 사용한 경험이 있는 공학 계열 연구자, Google Scholar, RISS, DBPia 사용 경험이 있는 사람, 기타 유사 지식인프라서비스를 사용한 경험이 있는 사람을 위주로 총 5명이 선정됐다. 조사방법은 사용성 평가 장비가 구비된 사용성 연구실에서 데스크톱 장비를 이용해 대상자가 직접 ScienceON을 이용하여 준비된 시나리오를 수행하고 수행한 결과를 연구원이 분석하는 방법을 활용했다. 분석기준은 시나리오 Task들을 수행하면서 발생하는 인터페이스 구성요소(용어와 심볼, 네비게이션, 레이아웃, 피드백)의 용의성 분석, 사용성 구성요소(효율성, 정확성, 일관성, 유연성, 의미성) 분석, UX 10 Usability Heuristics[1] 등을 적용했다.

사용성 평가 결과 Google Scholar의 만족도와 사용 빈도가 높게 나타났으며, 전반적으로 ScienceON이 제공하는 개념은 궁극적이지만 통합검색의 인터페이스와 검색 결과의 양과 질적 측면에서 다소 낮은 점수를 보였다. ScienceON은 다양한 연구지원 편의기능을 제공하고 있어 국내 연구자들의 연구활동 지원에 적합하지만, 모호한 용어 사용, 검색 결과 노출 등 다양한 부분에서 개선이 필요하다는 결론을 내렸다.

이 결과를 바탕으로 ScienceON의 통합검색 부분의 UI/UX에 대한 개선을 추진 중이며, 검색 결과의 질적 측면의 개선을 위한 연구를 수행하고 있다. 또한, 연구자들의 연구 전주기를 지원할 수 있으며 타 과학기술정보서비스와 차별화될 수 있는 MyON, AIHelper 등의 서비스를 고도화하고 있다.

Key words

Usability Evaluation and Analysis

[1] <https://www.nngroup.com/articles/ten-usability-heuristics/>

\* 한국과학기술정보연구원 국가과학기술데이터본부 융합서비스센터, {jykim, jhpark, yoosu<sup>†</sup>}@kisti.re.kr

※ 본 연구는 한국과학기술정보연구원에서 수행하는 주요사업 '지능형 과학기술지식인프라 통합서비스 플랫폼 구축(K24L1M3C4-01)'의 지원을 받아 수행된 연구임.

# 연구소소프트웨어 상호운용을 위한 KRDC 프레임워크 구현

조민희\*, 김다솔\*, 송사광\*, 이미경\*, 임형준\*

## Implementation of the KRDC Framework for Research SW Interoperability

Minhee Cho, Dasol Kim, Sa-kwang Song, Mikyoung Lee, Hyung-Jun Yim

본 논문에서는 연구소프트웨어의 상호운용성을 향상시키기 위한 KRDC(Korea Research Data Commons) 프레임워크를 제시한다. KRDC는 오픈소스를 기반으로 개발된 통합 도구로, 연구소프트웨어의 관리 및 활용을 간소화하여 다양한 연구 환경에서 효과적으로 소프트웨어를 관리, 공유 및 활용할 수 있도록 지원한다.

KRDC 프레임워크는 4가지 특징을 가지고 있다. 첫째 사용자 중심의 연구소프트웨어 관리 기능을 지원한다. 연구소프트웨어 코드를 개발하고 공유하는 과정에서 복잡한 기술적 설정 없이 직관적인 인터페이스를 제공하여 쉽게 작업할 수 있도록 돕는다. 또한 연구소프트웨어의 자동화된 배포 및 다양한 플랫폼과의 통합을 간편하게 설정할 수 있는 기능을 제공한다. 둘째 연구소프트웨어의 상호운용성을 촉진한다. KRDC는 유럽의 EGI(European Grid Initiative)와의 공동연구를 통해 DataHub 서비스와의 상호인증을 통한 데이터 연계를 검증하며, 상호운용성 강화를 지원하고 있다. 또한 외부 서버 및 클라우드에 KRDC프레임워크를 설치하여 KRDC 허브 서비스를 통해 연합함으로써 다양한 플랫폼 간 호환성을 검증하였다. 셋째 연구소프트웨어 상호운용을 위한 워크플로우 기능을 제공한다. KRDC는 단일 소프트웨어 및 다중 소프트웨어 실행을 지원하는 워크플로우 기능을 통해 연구소프트웨어 간 데이터의 입출력을 표준화하고, 효율성과 재사용성을 높이며 연구 과정에서 발생할 수 있는 비호환성 문제를 최소화하였다. 또한 다른 서버의 프레임워크에서 관리되고 있는 연구소프트웨어는 프록시 API를 통해 실행할 수 있어, 연구자들은 서로 다른 시스템에서 소프트웨어를 직접 실행하고 호출하지 않고도 통합적으로 관리하고 활용할 수 있다. 넷째, KRDC는 통합된 데이터 저장소를 제공하여, 여러 프레임워크에서 생성된 데이터를 중앙에서 관리하고 교환할 수 있도록 지원한다. 이를 통해 연구자들은 서로 다른 플랫폼의 데이터를 손쉽게 공유하고 협업할 수 있다. 마지막으로 KRDC는 쿠버네티스 기반의 유연한 인프라를 제공한다. 쿠버네티스에서 실행되도록 설계된 KRDC는 연구소프트웨어를 컨테이너화된 환경에서 실행하고 관리할 수 있도록 하여 연구자들이 연구소프트웨어의 개발 및 배포를 용이하게 하고, 확장 가능한 인프라를 활용할 수 있게 한다. 운영자는 쿠버네티스의 모놀식 구조를 활용하여, 다양한 오픈소스 도구를 설치하고 통합할 수 있어 필요에 따라 기능을 확장할 수 있다.

이러한 특징을 통해 KRDC는 연구소프트웨어의 상호운용성과 효율적인 관리, 그리고 글로벌 협업을 촉진하는 강력한 도구로, 앞으로 더 많은 연구 플랫폼과 통합되어 다양한 연구 분야에서 활용될 수 있을 것으로 기대한다. 향후 연구에서는 KRDC의 적용 범위를 더욱 확대하여, 국내외 연구자들이 사용할 수 있는 범용 연구 데이터 허브로 발전시키는 데 중점을 둘 계획이다.

### Key words

Research Software, Korea Research Data Commons, Interoperability

### 참고문헌

- [1] 조민희외 3명, “연구소소프트웨어 공유 활용을 위한 학술 인프라 설계”, 한국경영과학회 학술대회논문집, 2023.

\*한국과학기술정보연구원 (mini, dskim, esmallj, jerryis, hjyim)@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: (KISTI)K24L1M2C3-01)



# 다중 에이전트 환경에서의 최적 경로 탐색 기술

이종해\*, 김대영\*, 안병권\*\*

## Optimal Path Finding Technology in Multi-Agent Environment

JongHae Lee\*, DaeYoung Kim\*, and ByungKuon Ahn\*\*

본 논문에서는 다중 에이전트(Multi-Agent) 운용 환경에서 최단 경로 탐색 기술을 제안한다. 다수의 에이전트가 고정된 공간 내에서 동시에 최적 경로를 찾고, 충돌 없이 효율적으로 이동할 수 있도록 하기 위해 다양한 경로 탐색 알고리즘이 적용되었다[1]-[3]. 본 연구에서 사용된 주요 알고리즘은 Cooperative A-Star 알고리즘으로 다수의 에이전트가 실시간으로 협력하며 경로를 탐색한다. 각각의 에이전트가 독립적으로 경로를 계산하되, 다른 에이전트와의 충돌 가능성을 실시간으로 예측하고 조정한다. 충돌이 예상되는 경우 미리 정의된 우선순위에 따라 낮은 우선순위의 에이전트가 속도를 줄이거나 경로를 수정하여 충돌을 방지함으로써 전체 시스템의 효율성을 높일 수 있다. 중앙 집중형 관제 시스템(Multi-Agent Control System, MACS)은 각 에이전트가 목적지까지 도달하기 위한 최초 경로 탐색을 수행하여 최단 경로를 광역(Global) 지정 경로로 설정하여 주행하도록 한다. MACS는 실시간으로 충돌을 예측하고 에이전트들 간의 충돌 가능성이 발생하는 경우 우선순위에 따라 각 에이전트의 경로를 조정한다. 이 때 지역(Local) 주행 경로는 각 에이전트가 현재 시점에서 점유하고 있는 경로이며 MACS는 지역 주행 경로를 배제하여 우회 경로를 탐색한다. 특정 경로에서 충돌 가능성이 감지되면, 우선순위가 높은 에이전트가 기존 경로를 유지하고 우선순위가 낮은 에이전트는 대기하거나 새로운 우회 경로로 주행한다. 이와 같은 실시간 조정을 통해 에이전트들은 충돌이나 불필요한 정지 없이 원활하게 이동할 수 있다. 본 연구에서 제안한 다중 에이전트 운용 환경의 최단 경로 탐색 기술은 물류 및 생산 환경에서의 무인 자동화 시스템의 효율을 크게 향상시킬 수 있다. 제안하는 알고리즘을 활용한 충돌 방지와 각 에이전트의 실시간 경로 업데이트 기능은 다중 에이전트가 복잡한 환경에서도 원활하게 작동할 수 있도록 지원하며, 향후 다양한 자동화 시스템에서의 응용 가능성을 제시한다.

### Key words

autonomous robot, path finding, multi-agent, multi-agent control system

### 참고문헌

- [1] 김종선, 김세진, 주영훈, "AGV의 최적 경로 탐색 알고리즘," 2007년도 정보 및 제어 학술대회, 대한전기학회 2007년도 Oct. 2007.
- [2] 주영훈, 김종선, "경로 점유권을 이용한 AGV의 경로 제어 알고리즘," 한국지능시스템학회 논문지, vol. 18, no. 5, pp. 592-598, Aug. 2008.
- [3] 최미형, 서민호, 우제승, 홍순기, "다기준의사결정기법과 수정 A-STAR 알고리즘을 이용한 목적지 최적경로 탐색 기법 개발," Journal of the Korean Society of Industry Convergence, vol. 24, no. 6\_2, pp. 891-897, Dec. 2021.

---

\*(주)러셀로보틱스, jhlee@erussell.co.kr, dykim@erussell.co.kr, \*\*승실대학교, abg6312@ssu.ac.kr(교신저자)

# 마이크로컨트롤러와 지능이론을 이용한 유도전동기시스템의 고장진단

이종현\*, 장재표\*\*, 류민영\*\*\*, 오창욱\*\*\*, 이인수\*\*\*\*

## Fault diagnosis of Induction Motor System using Micro-controller and Intelligent Theory

Jong Hyun Lee\*, Jae Pyo Jang\*\*, Min Young Ryu\*\*\*, Chang Wook Oh\*\*\*, and In Soo Lee\*\*\*\*

### 요 약

본 논문에서는 라즈베리파이와 CNN을 이용한 유도전동기 고장진단시스템을 제안하였다. 제안한 방법에서는 유도전동기 시뮬레이터를 통해 획득한 진동 데이터를 CNN 신경회로망 모델로 학습하였고 라즈베리파이를 사용하여 데이터 수집 및 진단 시스템 동작을 하나의 컨트롤러로 수행하였다. 또한 진단 결과를 터치스크린 과 GUI를 통해 사용자가 직접 확인할 수 있다. 시스템의 성능은 유도전동기의 시뮬레이터에서 획득한 데이터를 사용하여 진단 실험 결과를 통해 시스템의 성능을 확인하였으며, SVM 및 MNN과의 고장진단 성능을 비교하였다.

### Abstract

In this paper, we propose a fault diagnosis system for induction motors using Raspberry Pi and CNN. The proposed method involves learning a CNN neural network model with vibration data obtained from an induction motor simulator. The Raspberry Pi is utilized to acquire data and operate the diagnosis system as a single controller. Additionally, users can directly check the diagnosis results through a touchscreen and GUI. The performance of the system is verified through diagnostic experiments using data obtained from the induction motor simulator and compared with the diagnosis results by SVM and MNN.

### Key words

Induction motor, convolution neural network, SVM, MNN, fault diagnosis system, raspberry pi

## I. 서 론

유도 전동기는 교류 전동기의 대표적인 예로, 브

러시가 없는 브러시리스(brushless) 형태로 직류전동기와 비슷한 출력 특성을 가지는 작동 형태를 가진

---

\* (주) 아진산업

\*\* 경북대학교 전자전기공학부

\*\*\* (주) 케이와이

\*\*\*\* 경북대학교 전자전기공학부 교수, 교신저자

다. 이에 따라 유도 전동기는 구조가 간단하고 유지보수가 용이하여 현재 현대 산업 응용 분야에서 기계 및 산업 설비를 구동하는 가장 중요한 장비 중 하나이다. 그러나 유도 전동기의 예기치 못한 고장은 생산 공장의 가동 중단을 초래하고, 상당한 경제적 손실을 유발할 수 있다. 따라서 유도 전동기의 상태를 모니터링하고 고장을 조기에 감지할 수 있는 고장진단 시스템의 개발이 필요하다.[1-2]

고장진단은 시스템 내에서 발생하는 고장을 탐지하고 이를 분류하는 과정으로, 크게 모델에 근거한 방법과 모델 근거하지 않는 방법으로 분류할 수 있다.[3] 모델에 근거한 방법은 시스템의 동작을 수학적으로 모델링한 모델을 근거로 하여 고장을 진단하지만, 시스템의 비선형성으로 인해 때문에 정확한 모델을 얻기가 쉽지 않다. 반면, 모델에 기반하지 않는 방법은 측정 데이터, 경험, 물리적 검토 등을 바탕으로 하며, 한계치 검사, 전문가 시스템 기법, 신경망(Neural network) 등의 방법이 포함된다. 특히 신경망은 병렬 처리 능력과 비선형함수 표현 능력, 학습 기능 등의 장점 덕분에 고장진단에 많이 활용되고 있다. 특히 이러한 방법은 기계 학습 알고리즘을 활용하여 모터 데이터의 복잡한 패턴을 분석하여 높은 정확도로 고장을 나타내는 미묘한 이상을 판단할 수 있다. 또한 신경망 기반 고장진단 기술은 대규모 모터 데이터 세트에 적합하여 모터를 지속해서 모니터링하고 향후 고장으로 이어질 수 있는 추세를 감지할 수 있다는 장점이 있다[3].

본 논문에서는 신호의 패턴인식에 강한 CNN 모델을 고장진단에 사용하였고, 유도 전동기 시스템에서 얻은 정상, 회전자 고장, 베어링 고장의 3가지 상태의 진동 신호를 사용하여 본 논문에서 사용한 CNN 모델의 고장 판단의 정확성을 검증하였다[4].

또한 라즈베리파이를 사용하여 고장진단 시스템을 구현하였다. 라즈베리파이를 통해 진동데이터의 수집과 CNN 모델을 이용한 고장진단을 컨트롤러 하나로 수행하였다. 그리고 터치스크린과 GUI(Graphical user interface)를 통해 사용자에게 실시간으로 모니터링 할 수 있게 하였고 SVM(Support Vector Machine)과 MNN(Multilayer Neural Network)을 이용한 고장 분류결과와 비교하였다.

## II. 유도전동기 고장진단 실험환경

유도 전동기의 고장진단 실험을 위해 그림 1과 같이 정상, 회전자 고장, 베어링 고장을 유발시킨 3개의 3상 유도전동기로서 시뮬레이터를 구성하였다. 중앙에 위치한 유도전동기는 정상상태, 왼쪽과 오른쪽에 위치한 유도전동기는 각각 회전자고장 상태와 베어링고장 상태이다. 진동 데이터 수집을 위하여 전동센서는 Renke사의 RS-WZ1A-N01-1-CX을 사용하였고, 고장진단을 위한 마이크로컨트롤러는 라즈베리 파이 제단의 Raspberry Pi 5를 사용하였다.

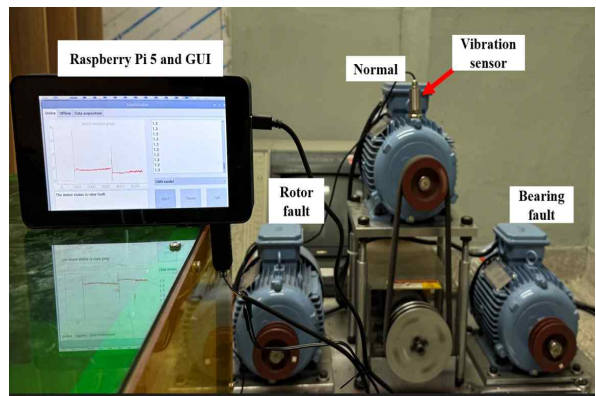


그림 1. 유도전동기 실험장치  
Fig. 1. Induction motor experimental unit

유도전동기에서 발생하는 고장의 종류는 회전자 고장과 베어링 고장 두 가지를 고려하였다. 고장진단 실험을 위해 회전자고장은 유도전동기의 회전자 bar에 드릴로 구멍을 내어 고장을 발생시켰으며, 베어링고장의 경우는 베어링에 쇳가루를 넣어 베어링의 마모를 발생시켰다.

## III. 제안한 CNN 기반 유도전동기 고장진단

논문에서는 컨볼루션신경회로망 (CNN), 서포트 벡터 머신(SVM)과 다층신경망(MNN)을 이용한 유도전동기 고장진단을 수행한다. 고장진단을 위해 본 논문에서 제안하는 고장진단 알고리즘은 그림 2와 같다. 유도전동기에서 진동센서를 사용하여 진동 데이터를 획득한 후 컨트롤러에 데이터를 전달하며 컨트롤러에서는 신경망 모델을 통해 고장을 진단한다. 그리고 고장진단 알고리즘의 순서도는 그림 3과 같다.

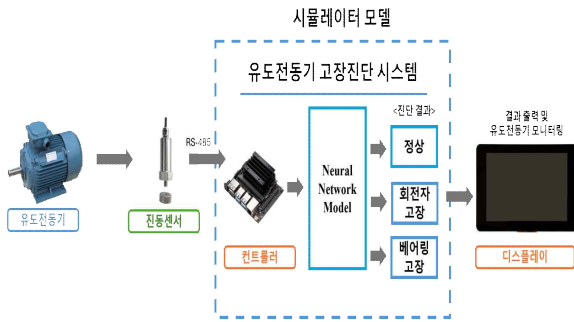


그림 2. 유도전동기 고장진단 시스템 구성도  
Fig. 2. Structure of the induction motor fault diagnosis system

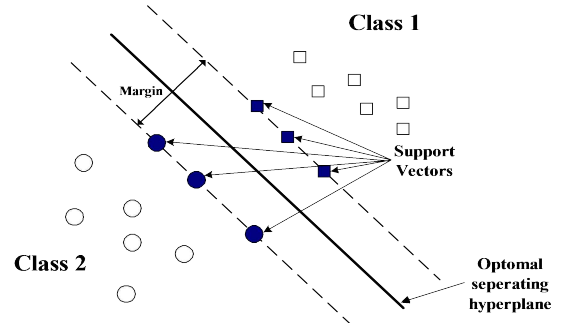


그림 4. SVM에 의한 클래스분류  
Fig. 4. Separation of two classes by SVM

### 3.1 Convolutional Neural Network

CNN은 기존의 패턴 인식 방식에서 분리되어 있던 특징 추출과 분류가 통합한 모델이다. CNN은 기본적으로 다수의 컨볼루션층(Convolutional layer)과 서브샘플링층(Subsampling layer)으로 구성되어 있다.[4] 유도 전동기 고장시스템에 사용된 CNN의 구조는 그림 3과 같이 같이 입력 레이어와 컨볼루션 레이어 2개, 맥스 풀링 레이어 2개 그리고 Fully connected Layer와 출력 레이어로 이루어져 있다.

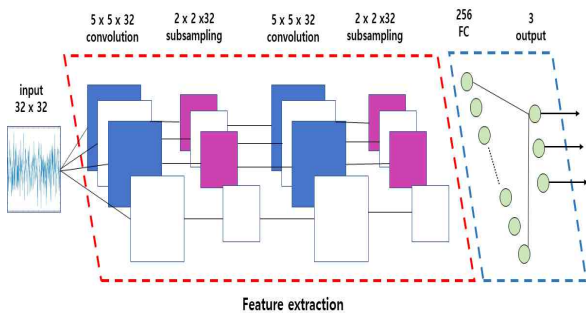


그림 3. CNN 모델의 구조  
Fig. 3. Structure of the CNN model

### 3.2 Support Vector Machine

SVM은 Vapnik [7]에 의해 개발된 학습알고리즘으로서 원래 이진분류기(binary classifier)였다. 즉, 그림 4와 같이 특징공간에서 초평면(hyperplane)과 가장 인접한 지지벡터(support vectors)들과의 거리인 여유(margin)를 최대로 하는 최대여유초평면(maximum margin hyperplane)을 찾아서 이진분류를 수행한다.[5-6]

그러나 모든 내적(inner product)을 비선형 커널함수(kernel function)로 교체하므로써 최대여유초평면이 원래의 입력공간보다는 훨씬 더 큰 차원을 가지는 비선형 사상(mapping) 특징공간에 적합하게 되었으며 이로 인해 비선형 분류가 가능한 분류기를 설계할 수 있다.

### 3.3 Multilayer Neural Network

다층신경망(MNN)은 선형함수만으로 구성되어 비선형 분류가 불가능한 단층 퍼셉트론의 한계를 보완한 신경망으로 구조는 그림 5와 같다[13]. 입력층과 출력층 사이에는 1개 이상의 은닉층이 존재하고 모든 계층의 노드들은 인접한 계층의 모든 뉴런과 서로 연결되어 있다[7].

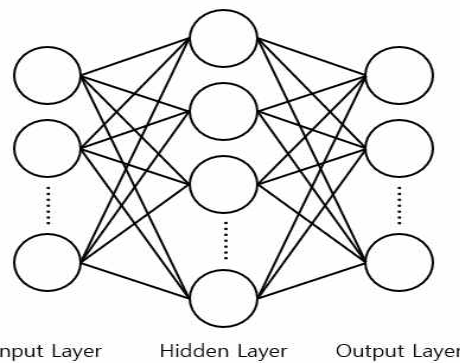


그림 5. 다층신경망의 구조  
Fig. 5. Architecture of Multi-layer neural network

다층신경망은 순전파와 역전파를 통해 학습을 수행한다. 순전파는 입력층에서 출력층 방향으로 가중치와 활성화 함수를 사용하여 출력값을 계산하는 방법이고 역전파는 순전파를 통해 계산된 신경망의

출력값과 실제값으로 손실함수의 값을 계산하고, 학습 알고리즘을 통해 손실함수의 기울기가 감소하는 방향으로 가중치를 갱신하는 방법이다.

#### IV. 실험결과 및 고찰

시스템 테스트를 위해 유도전동기 시뮬레이터를 사용하여 테스트를 진행하였다. 유도전동기 시뮬레이터는 각각 정상, 회전자 고장 그리고 베어링 고장 상태의 유도전동기 3개로 구성되어 있으며 실험을 각 시뮬레이터에서 데이터를 얻고 이를 세 가지 모델 (CNM, SVM, MNN)을 통해 학습하였다.

시뮬레이터의 진동 데이터는 100개의 데이터를 한 쌍으로 정의하여 각 상태(정상, 회전자 고장, 베어링 고장)마다 100쌍의 데이터 세트를 테스트로 사용하여 총 300쌍의 테스트 데이터 세트를 사용하여 모델의 성능을 테스트하였다.

실험에 사용된 CNN은 100개의 노드를 가진 입력층과 첫 번째 컨볼루션 층(32개 노드) 과 두 번째 컨볼루션 층(32개 노드), 16개의 노드를 Fully connected Layer, 3개의 노드를 가진 출력층으로 구성된다. 두 개의 컨볼루션층과 Fully connected Layer 은 활성화 함수인 ReLU가 사용되었다.

그림 6은 CNN 모델에서의 유도전동기 시뮬레이터의 고장진단 결과를 confusion matrix로 나타낸 것이다. Confusion matrix의 Target 라벨은 실제 유도전동기 상태를 의미하며 Diagnosis는 CNN 모델이 진단한 유도전동기의 상태를 의미한다. Confusion matrix에서 상태마다 CNN 모델의 진단결과가 Target 과 동일하게 진단하여 모든 상태에서 100쌍의 데이터를 전부 알맞게 진단한 것을 확인할 수 있다. 따라서 CNN 모델의 진단 정확도가 100%임을 확인할 수 있다,

그림 7은 SVM 모델에서의 유도전동기 시뮬레이터의 고장진단 결과를 confusion matrix로 나타낸 것이다. 고장진단에 사용된 SVM 모델은 C값을 1로  $\gamma$ 값을 0.01로 설정하였고 RBF 커널을 사용하였다. SVM 모델은 정상과 베어링 고장상태에서는 100쌍의 데이터를 전부 정확하게 분류하여 100%의 진단 정확도를 보였으나, 회전자 고장 부분에서 회전자

고장으로 89쌍, 베어링 고장으로 11쌍으로 진단하여 89%의 정확도를 나타내었다. 이는 추후 파라미터 튜닝을 통해 성능 향상을 기대할 수 있다.

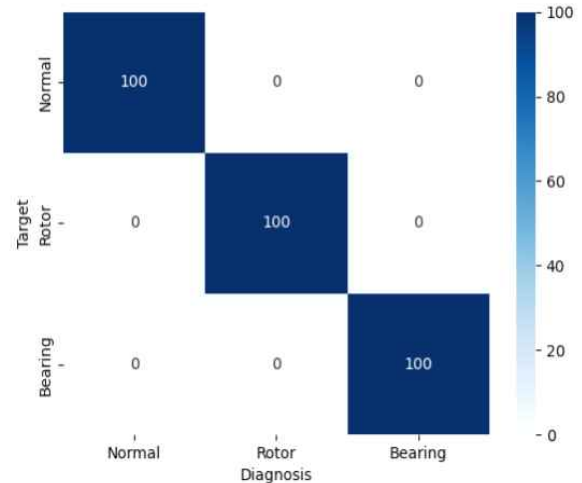


그림 6. CNN 모델의 고장진단 결과  
Fig. 6. Fault diagnosis results by CNN

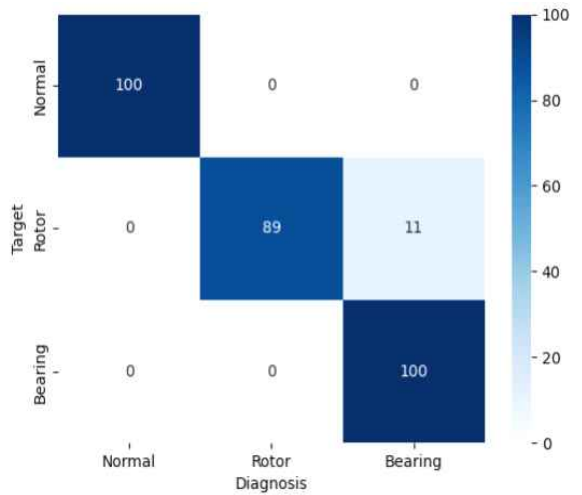


그림 7. SVM 모델의 고장진단 결과  
Fig. 7. Fault diagnosis results by SVM

그림 8은 MNN 모델에서의 유도전동기 시뮬레이터의 고장진단 결과를 confusion matrix로 나타낸 것이다. MNN은 100개의 노드로 구성된 1개의 입력층, 3개의 노드로 구성된 1개의 출력층, 2개의 은닉층으로 구성되었다. 첫 번째 은닉층의 노드 수는 32개, 두 번째 은닉층의 노드 수는 16개를 사용하였고, 은닉층의 활성화 함수로 ReLU 함수를 사용하였다. 학습 알고리즘으로 아담을 사용하였고, epoch

수는 3,000으로 학습을 진행하였다. MNN 모델은 CNN 모델과 동일하게 모든 상태에서 100쌍의 데이터를 정확하게 진단하여 100%의 정확도를 보여주었다.

고장진단을 위해서 제작한 GUI는 그림 9와 같다. GUI에는 온라인 실시간 모드, 오프라인 모드 및 데이터 수집 모드와 같은 세가지 동작 모드가 있다.

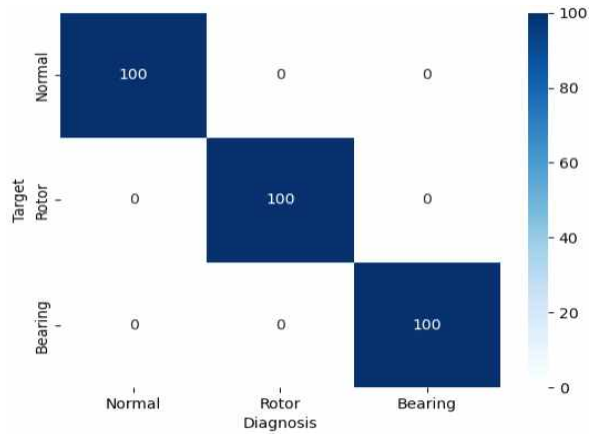


그림 8. MNN 모델의 고장진단 결과  
Fig. 8. Fault diagnosis results by MNN

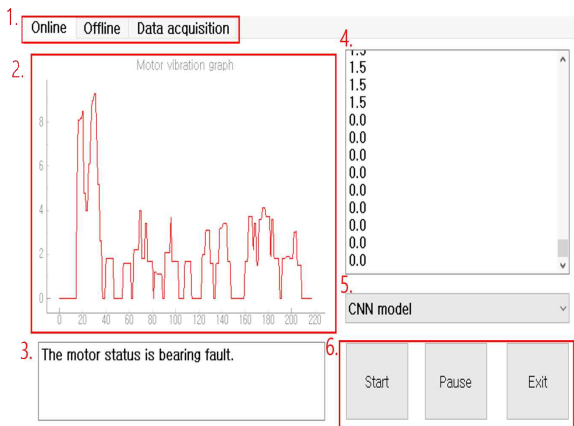
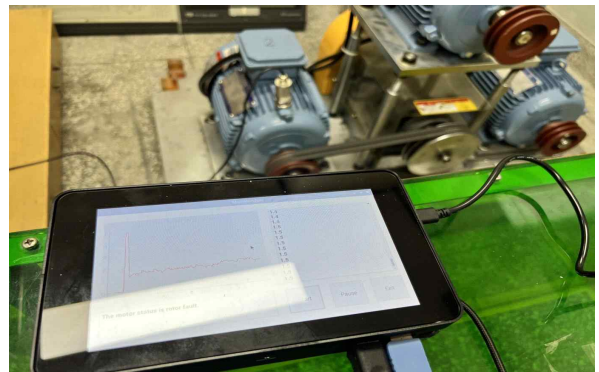


그림 9. 고장진단을 위한 GUI  
Fig. 9. GUI for fault diagnosis

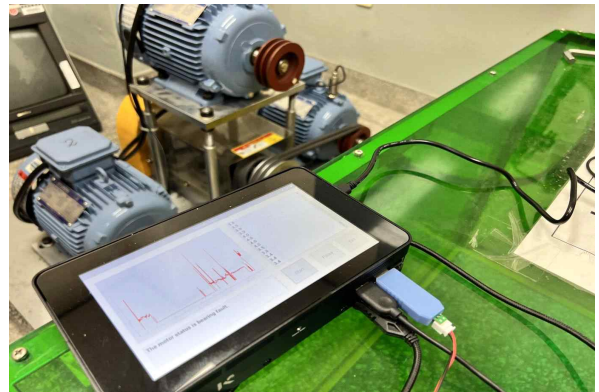
그림 10(a), (b) 및 (c)는 각각 정상상태, 회전자고장상태 및 베어링고장상태의 데이터로 학습한 모델을 탑재한 시스템을 사용하여 시뮬레이터의 모터 상태를 실시간으로 진단한 결과이다. 테스트는 CNN 모델을 사용하여 진행하였다. 시스템 테스트 결과 정상, 회전자 고장 그리고 베어링 고장상태에서 정확하게 진단하는 것을 확인할 수 있다.



(a) 정상상태  
(a) Normal



(b) 회전자고장  
(b) Rotor fault



(c) 베어링 고장  
(c) Bearing faultt

그림 10. 유도전동기 고장진단결과  
Fig. 10. Induction motor fault diagnosis results

## V. 결 론

본 논문에서는 라즈베리파이 5와 지능이론을 이용한 유도전동기 고장진단을 수행하였다. 실험을 위해 정상 상태, 회전자 고장 상태, 베어링 고장 상태 유도전동기가 있는 유도전동기 시뮬레이터를 제작하였다. 시뮬레이터를 통해 진동 데이터를 수집하였

고 CNN, SVM 및 다층신경망을 사용하여 고장진단을 수행하였고 성능을 비교하였다. 또한 터치스크린과 GUI를 통해 사용자에게 실시간으로 모니터링할 수 있도록 기능을 구현하였다. 실험결과 제안한 고장진단시스템이 정확하게 정상, 회전자고장 및 베어링고장을 잘 분류함을 알 수 있었다.

## 사 사

본 논문은 2024년도 과학기술정보통신부의 재원으로 연구개발특구진흥재단의 지원을 받아 수행된 연구임 (2024-DG-RD-0018)

## 참 고 문 헌

- [1] Wenjun Sun, Rui Zhao, Ruqiang Yan, Siyu Shao, Xuefeng Chen "Convolutional Discriminative Feature Learning for Induction Motor Fault Diagnosis", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 13 2017.6, 1350-1359 (10 pages).
- [2] In Soo Lee, "Fault Diagnosis System Development of Induction Motors Using Discrete Wavelet Transform and Neural Network", Journal of Korean Institute of Information Technology 9(8), 2011.8, 53-61 (9 pages).
- [3] Ron J. Patton, "Robust Model-Based Fault Diagnosis: The State of the ART", IFAC Proceedings Volumes, Vol. 27, no. 5, pp. 1-24, June 1994.
- [4] Rui Zhao, Ruqiang Yan, Zhenghua Chen, Kezhi Mao, Peng Wang, Robert X. Gao, "Deep learning and its applications to machine health monitoring", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 13 Mechanical Systems and Signal Processing 115 (2019) 213-237(26 pages).
- [5] Rakesh Kumar Jha, Preety D. Swami, "Fault diagnosis and severity analysis of rolling bearings using vibration image texture enhancement and

multiclass support vector machines", Applied Acoustics, Vol. 182, November 2021.

- [6] P.K. Kankar, Satish C. Sharma, S.P. Harsha, "Fault diagnosis of ball bearings using machine learning methods", Expert Systems with Applications, vol. 38, no. 3, pp. 1876-1886, 2011.
- [7] B.T. Pham, M.D. Nguyen, K.-T.T. Bui, I. Prakash, K. Chapi, D.T. Bui, "A novel artificial intelligence approach based on multi-layer perceptron neural network and biogeography-based optimization for predicting coefficient of consolidation of soil", Mechanical Systems and Signal Processing, vol. 21, no. 6, pp. 2560-2574, August 2007.

# DNN을 활용한 4축 DC 모터 기반 자세 제어 스테이지의 정밀도 향상 연구

변상봉\*, 이성호\*\*, 한동철\*, 이성훈\*, 정용안\*, 조수현\*

## A Study on Precision Enhancement of a 4-Axis DC Motor-Based Attitude Control Stage Using DNN

Sang-Bong Byun\*, Sung Ho Lee\*\*, Dong Cheul Han\*, Sung hun Lee\*, Yong An Jung\* and  
Soo-Hyun Cho\*

본 논문은 4축 DC 모터를 활용한 자세 제어 스테이지의 정밀도 향상을 위해 딥 러닝(DNN[1]-[2])을 적용한 연구이다. 제어 스테이지는 메타버스 기반 가상 환경에서 가정용 자전거의 기울기를 다양한 지형에 맞춰 실시간으로 조정하여 사용자가 오르막, 내리막, 커브 등을 직관적으로 경험할 수 있도록 한다. 이를 통해 몰입감이 높아져 사용자는 운동에 집중하고 동기부여가 높아져 지속적으로 운동을 이어가게 한다. 하지만 각 모터 사이의 미세한 구동 오차로 인해 동일한 입력값에서도 출력 차이가 발생하여 스테이지의 제어에 오차가 발생하여 몰입감을 저하시킬 수 있다. 이를 해결하기 위해, 각 모터의 입력값 대비 실제 출력값을 DNN으로 학습하여 보정값을 생성하고, 이를 실제 모터에 적용함으로써 입력과 출력이 정확히 일치하도록 한다. 이 접근법을 통해 15~20%의 제어 정밀도와 속도 향상을 기대할 수 있으며, 사용자가 원하는 만큼 모터가 정확하게 움직여 현실감을 높일 수 있다. 향후 연구에서는 모터 제어의 추가적 성능 향상을 위해 R-CNN, ResNet-50[3] 등의 모델을 도입하여 다양한 제어 방식의 적용 가능성을 연구하고, 이를 여러 분야로 확장할 계획입니다.

### Key words

deep learning, metaverse, artificial intelligence, motor

### 참고문헌

- [1] A. Fatemimoghadam, L. V. Iyer, N. C. Kar, "Real-Time Validation of Enhanced Permanent Magnet Synchronous Motor Drive Using Dense-Neural-Network-Based Control", IEEE Access., vol. 12, pp. 73323-73339, May 2024.
- [2] A. Tragoudaras, C. Antoniadis, Y. Massoud, "Enhancing DNN Models for EEG/ECOG BCI With a Novel Data-Driven Offline Optimization Method", IEEE Access., vol. 11, pp. 35888 - 35900, Apr. 2023.
- [3] V. Jain, S. Mohamed, D. Goswami, S. Stuijk, "DNN-based visual perception for high-precision motion control", 2024 ECC., pp. 2003 - 2009, Jun. 2024.

---

\* 구미전자정보기술원 지능형디바이스연구센터, email : sbbyun@geri.re.kr(교신저자), cataegu07@geri.re.kr, leesh@geri.re.kr, yajung@geri.re.kr, shcho@geri.re.kr

\*\* 아이엔티, email : k2sungho@hanmail.net

※ 이 연구는 2024년도 산업통상자원부 및 중소벤처기업진흥공단 연구비 지원에 의한 연구임(딥러닝을 활용한 메타버스 기반의 홈 사이클링, 2024-1-030)



# 시계열 진동 데이터 기반의 하이브리드 딥러닝을 이용한 작업파트너 로봇의 이상 검출 연구

최 승 환\*

## A Study on Anomaly Detection in Collaborative Robots Using Hybrid Deep Learning Based on Time-Series Vibration Data

Seung-Hwan Choi\*

본 논문에서는 작업파트너 로봇의 진동 데이터를 활용하여 고장 발생 전 이상 상태를 조기에 검출할 수 있는 연구를 수행하였다. 최근 제조 산업에서는 협동 로봇과 물류 로봇을 포함한 작업파트너 로봇의 사용이 증가하면서 생산성과 효율성이 크게 향상되고 있다[1]. 그러나 이러한 로봇은 기계적 한계로 인해 고장과 오작동이 발생할 수 있으며, 이는 생산성 저하뿐만 아니라 작업 현장에서의 안전사고로 이어질 수 있다[2]. 따라서, 산업 현장에서는 작업파트너 로봇의 상태를 실시간으로 모니터링하고, 고장이 발생하기 전에 이상 상태를 조기에 감지할 수 있는 기술이 필수적이다[3]. 본 연구에서는 작업파트너 로봇의 구동 모듈 내구성 시험을 통해 수집된 진동 데이터를 활용하여 하이브리드 딥러닝 모델 기반의 이상 상태 검출 방법을 제안한다. 제안된 모델은 시계열 데이터에 시간에 따른 가중치를 적용하여 고장 직전의 이상 신호를 효과적으로 검출할 수 있도록 설계하였다. 모델의 성능을 평가하기 위해 하이브리드 딥러닝 모델과 기존 단일 모델의 이상 검출 성능을 비교하였으며, 실험 결과 제안된 모델이 고장 직전의 이상 상태를 기존 모델보다 더 정확하게 검출하였음을 확인하였다.

### Key words

working partner robot, time-series vibration data, anomaly detection, hybrid deep learning

### 참고문헌

- [1] K. K. Ramachandran, "Intelligent Manufacturing: Leveraging Autonomous Robotics and AI", Computer Science and Information Technology Research (IJCSITR), Vol. 5, No. 2, pp. 20-32, Dec. 2024.
- [2] R. Kermenov, G. Nabissi, S. Longhi and A. Bonci, "Anomaly Detection and Concept Drift Adaptation for Dynamic Systems: A General Method with Practical Implementation Using an Industrial Collaborative Robot", Sensors, Vol. 23, No. 6, pp. 3260, Mar. 2023.
- [3] P. Kumar, S. Khalid and H. S. Kim, "Prognostics and Health Management of Rotating Machinery of Industrial Robot with Deep Learning Applications—A Review", Mathematics, Vol. 11, No. 13, pp. 3008, Jul. 2023.

---

\* 한국생산기술연구원, csw1496@kitech.re.kr

※ 본 논문은 한국생산기술연구원 기관주요사업 "제품 제조현장 작업파트너 로봇 기술 개발(5/6) (kitech EO-24-0007)"의 지원으로 수행한 연구입니다.

## micro-ROS가 적용된 VRS-RTK GNSS 모듈 설계

김동언\*, 변상봉\*, 이성훈\*, 한동철\*, 조수현\*

### Design of a VRS-RTK GNSS Module with micro-ROS Application

Dong-Eon Kim\*, Sang-Bong Byun\*, Sung hun Lee\*, Dong Cheul Han\* and Soo-Hyun Cho\*

본 논문에서는 고정밀 측위 성능을 만족하면서 비용 효율적인 시스템을 구축하기 위한 임베디드 기반의 VRS-RTK GNSS 모듈을 설계한다. 다양한 분야에서 자동화 시스템 적용에 대한 수요가 존재하고, 대부분의 시스템에서는 측위에 대한 성능이 요구되므로 cm급 정밀 측위를 위한 GNSS 시스템이 적용되는 경우가 많다 [1][2][3]. 하지만 고정밀 측위 장치의 경우에 RTK 리플레이 및 수신기 등이 함께 구성되어 구축 비용이 비싸고, 현장에 적용하기 위해서는 하드웨어 최적화 등의 추가 작업이 필요하다. 이를 해결하기 위하여, micro-ROS가 적용된 임베디드 기반의 VRS-RTK GNSS 시스템을 제안한다. 본 논문에서 제안된 시스템은 고정밀 측위 정보의 수신을 위해 ZED-F9K 모듈을 사용하여 I2C 버스를 통해 MCU인 ESP32와 통신한다. 설계된 디바이스는 GNSS 위치 정보 데이터를 전송하고 RTK 목적으로 RTCM(실시간 보정 메시지) 데이터를 수신한다. 이때, micro-ROS가 적용되어 MCU가 ROS2 프로토콜을 사용하여 호스트 머신과 통신할 수 있게 된다. 호스트 머신에는 우분투가 적용되어 ROS2 플랫폼에서 2개의 노드가 실행된다. micro-ROS 에이전트 노드는 MCU인 ESP32와 호스트 머신의 ROS2 시스템 간의 브릿지 역할을 한다. 여기서 GNSS 데이터를 /fix 토픽에 게시하고, RTCM 보정 메시지를 ESP32로 전달하기 위해 /rtcm 토픽을 구독한다. NTRIP 클라이언트 노드는 GNSS 데이터를 기반으로 /rtcm 토픽에 RTCM 메시지를 게시하여 실시간 보정에 사용한다. 본 논문에서는 임베디드 시스템에서 구동되는 micro-ROS가 적용된 VRS-RTK GNSS 시스템을 제안하였다. 제안된 방법은 라즈베리파이 혹은 PC 등 운영체제를 가진 기기에서 동작하는 방법에 비하여 실시간성이 개선된다.

#### Key words

uROS, VRS-RTK GNSS, RTCM, NTRIP

#### 참고문헌

- [1] Akhshirsh, G. S., Al-Salihi, N. K., & Hamid, O. H. (2021). A cost-effective GPS-aided autonomous guided vehicle for global path planning. *Bulletin of Electrical Engineering and Informatics*, 10(2), 650-657.
- [2] H. J. Yoon, and B. W. Park. "Assessment of Position Degradation Due to Intermittent Broadcast of RTK MSM Correction Under Various Conditions." *Journal of Positioning, Navigation, and Timing* 9.3 (2020): 237-248.
- [3] Y. K. Park, W. P. Kang, J. E. Choi, and B. J. Kim. "A study on the Evaluation of Real-Time Map Update Technology for Automated Driving." *Journal of the Korean Association of Geographic Information Studies*. 22.3 (2019): 146-154.

---

\* 구미전자정보기술원 ICT디바이스연구센터, email : dekim@geri.re.kr(교신저자), sbbyun@geri.re.kr, leesh@geri.re.kr, cataegu07@geri.re.kr, shcho@geri.re.kr

※ 이 연구는 2024년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임(20017474)

# 무인 지상로봇의 지상관제 시스템을 위한 모듈형 변환기의 표준 프로토콜 적용 연구

최지훈\*, 하옥균\*\*

## A Study on Applying a Standard Protocol to Modular Converters for the Ground Control System of Unmanned Ground Robots

Ji-Hoon Choi\* and Ok-Kyoon Ha\*\*

전 세계 병역 자원의 감소로 인명 피해를 최소화하고 전투 효율을 극대화하기 위해 무인 항공기(UAV)의 첨단화가 급속히 진행되고 있으며, UAV의 경우 NATO 회원국 간 STANAG-4586 프로토콜 표준 협정이 체결되어 군용 UAV의 기술 개발이 가속화되고 있다. 하지만 UAV 외에도 유-무인 복합 운영체제 (Manned-Unmanned Teaming, MUM-T) 구축에 필수적인 UGV(Unmanned Ground Vehicle)는 공통 표준 프로토콜이 부족하여 시스템 간 상호 운영성을 확보하기 어려운 상황이다. 따라서 본 논문에서는 무인 지상 이동체의 각기 다른 프로토콜을 해석하여 MQTT, ROS2, MAVLink 등의 무인 이동체를 위한 공통 프로토콜로 통합 및 변환하기 위한 VSM(Vehicle Specific Module)을 개발하기 위해 UAV의 STANAG-4586 공통 프로토콜을 적용하는 개념과 연구를 다룬다. 모듈형 VSM의 경우 고성능 PC를 이용한 CPU 내 스레드를 분할하여 구현할 수 있지만 이는 다른 프로토콜의 추가 및 삭제, 지상관제 시스템의 업그레이드 등의 상황에서 전체 프로그램 수정해야 하는 번거로움이 발생하게 된다. 이에 본 논문에서는 Slot Board 형태의 모듈화된 각각의 프로토콜에 해당하는 모듈형 VSM를 구현하여 교체 장착 및 시스템 변경이 용이하고 시스템 동작의 속도를 향상할 수 있도록 개발한다. 특히, 각기 다른 프로토콜로 운용되는 무인 지상 로봇에 대한 시스템의 상호 호환성을 확보하기 위해서는 개별의 프로토콜을 해석하여 STANAG-4586에 기반하는 표준 프로토콜로 변환하는 모듈형 VSM를 포함하는 지상관제 시스템에 대한 시스템 아키텍처를 제시하며, 지상관제 시스템에 내장되는 싱글 임베디드 보드에 VSM 기능을 구현하지 않고 각 프로토콜에 해당하는 모듈형 VSM를 제시함으로써 사용하지 않는 프로토콜 또는 새로운 프로토콜의 추가 시 지상관제 시스템 전체의 프로그램에 대한 수정 없이 모듈형 VSM의 교체 및 장착을 통해 시스템 변경을 편리하게 하는 지상관제 시스템 아키텍처를 제시한다.

### Key words

Unmanned ground robots, ground control system, network protocol, modular converter, STANAG-4586

### 참고문헌

- [1] Gyeongrae Nam, Jeonghwan Go, Cheolhee Kwon, Soyoung Jeong, "A Study on the Analysis and Improvement of STANAG 4586 / MAVLink Protocol for Interoperability Improvement of UAS," *Journal of the KIMST*, Vol. 23, No. 6, pp. 618-638, 2020.

\*아토즈, abcd7535@naver.com, \*\*경운대학교, okha@ikw.ac.kr

※ 본 연구는 방산혁신클러스터사업의 재원으로 국방기술진흥연구소, 경북·구미방산혁신클러스터사업단의 지원을 받아 수행되었음(No.DC2023SD).

# 정형외과 수술용 포지셔닝 로봇에서 일정한 힘과 거리로 골편을 견인하기 위한 제어 알고리즘 개발

방현희\*, 이현우\*, 이현주\*\*, 정상현\*\*\*, 박철우\*

## Development of a control algorithm for traction of bone fragments with constant force and distance in a positioning robot for orthopedic surgery

Hyun-Hee Bang\*, Hyun-Woo Lee\*, Hyun-Joo Lee\*\*, Sanghyun Joung\*\*\*, and Chul-Woo Park\*

### 요 약

본 논문은 골절 정복 수술의 효율성과 안전성을 향상하기 위한 골편 견인 제어 알고리즘을 제안한다. 골편 견인 제어 알고리즘은 견인 거리와 힘을 정밀하게 제어할 수 있는 두 가지 방식으로 구성된다. 제안된 알고리즘을 활용한 정형외과 수술용 포지셔닝 로봇 시스템은 의료진의 신체적 부담을 줄이고, 과도한 견인으로 인한 조직 손상을 방지할 수 있다.

### Abstract

This paper proposes a bone fragment traction control algorithm to improve the efficiency and safety of fracture reduction surgery. The control algorithm consists of two methods that can precisely control the traction distance and force of the bone fragment. The orthopedic surgical positioning robot system using the proposed algorithm can reduce the physical burden of medical staff in fracture surgery and prevent tissue damage caused by excessive traction.

### Key words

orthopedic surgery positioning robot, fracture reduction surgery, bone fragments traction

## 1. 서 론

골절 정복 수술은 부러진 뼈를 정상 위치로 복원하고 고정하는 절차를 의미한다[1]. 이 수술은 견인, 정복, 고정의 세 단계로 구성된다. 특히 대퇴골 골

절의 경우 근육의 긴장으로 인해 뼈가 어긋나기 때문에 이를 견인한 뒤 정확하게 정복하고 고정하는 과정이 필수적이다. 이 중 견인 작업은 최대 300N의 힘이 필요하며, 견인력과 견인 거리의 정확한 유지가 수술 결과에 중요한 영향을 미친다[2,3].

\* 경북대학교 융합의료기기로봇연구소, hyuntomb@gmail.com, hwlee.mdrip@gmail.com, pcw.mdrip@gmail.com

\*\* 경북대학교 의과대학 정형외과학교실, lidmania@daum.net

\*\*\* 주식회사 에어스, shjoung@airsurgical.net

※ 본 논문은 2024년도 과학기술정보통신부의 재원으로 연구개발특구진흥재단의 지원을 받아 수행된 연구임 (2024-DG-RD-0012)

현재 골절 수술에서 사용되는 견인 방식으로는 의료진이 직접 견인하는 방법과 수동 견인기를 갖춘 수술대를 사용하는 방법이 있다. 직접 견인은 의료진에게 과도한 신체적 부담을 주며, 수술 보조 인력이 추가로 필요하다. 수동 견인기는 조작이 비교적 편리하지만, 수술 중 환자의 움직임에 따라 견인력을 안정적으로 유지하기 어렵고, 견인력과 거리 측정이 불가능하여 과도한 견인으로 인해 환부가 손상될 위험이 존재한다.

이러한 문제를 해결하기 위해 정형외과 수술에서 의료진과 협동하여 환부의 포지셔닝을 지원하는 로봇을 개발하고 있고, 본 논문에서는 로봇 시스템을 이용하여 골편을 일정한 힘과 거리로 견인하기 위한 제어 알고리즘을 제안한다. 제안하는 알고리즘은 정확한 견인력과 거리를 유지하여 의료진의 부담을 줄이고, 수술의 효율성 및 환자의 안전성을 향상시킬 수 있다.

## II. 골편 견인 제어 알고리즘

본 연구에서 제안하는 정형외과 수술용 포지셔닝 로봇 제어 알고리즘은 견인 거리 제어 모드와 견인력 제어 모드의 두 가지 방식을 지원한다. 그림 1에 제어 알고리즘의 블록도를 제시하였다.

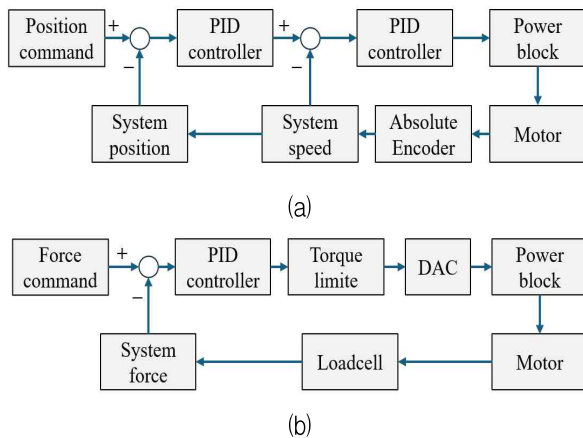


그림 1. 정형외과 수술용 포지셔닝 로봇 제어, (a) 견인거리 제어 알고리즘, (b) 견인력 제어 알고리즘  
Fig. 1. Control for orthopedic surgical positioning robot, (a) traction distance control algorithm, (b) traction force control algorithm, (c) control program

그림 1(a)의 거리 제어 모드에서는 절대치 엔코더를 통해 현재 위치를 검출하고, 사용자가 입력한 목표 위치와 비교하여 위치 추종 제어를 수행한다. 이때 모터의 구동 속도와 토크는 환자의 안전을 위해 미리 설정된 한계치를 초과하지 않도록 제어한다. 속도는 엔코더에서 산출된 값을 기준으로 제한되며, 토크가 한계치에 도달하면 시스템은 자동으로 견인을 중지한다. 이러한 제어 방식을 통해 견인 과정에서 과도한 힘이 가해지는 것을 방지하여 환자의 안전성을 확보할 수 있다.

그림 1(b)의 견인력 제어 모드에서는 로드셀을 통해 골편에 가해지는 힘을 측정하고, 이를 사용자가 설정한 목표 견인력과 비교하여 제어한다. DSP 기반 제어부는 실시간으로 견인력을 정확하게 조절하여 골편에 적절한 힘을 가할 수 있다.



그림 2. 정형외과 수술용 포지셔닝 로봇 시스템의 제어 프로그램

Fig. 2. Control program for positioning robot system for orthopedic surgery

그림 2에는 제안하는 알고리즘을 적용한 포지셔닝 로봇 제어 시스템의 프로그램 화면을 나타내었다. 수술 과정에서 수집되는 견인력, 거리, 속도 등의 정보가 사용자 인터페이스를 통해 수치와 그래프로 시각화되어 제공된다. 이를 통해 의료진은 수술 중 견인 상태를 직관적으로 파악하고 제어할 수 있다. 견인 초기에는 견인 거리 제어 모드로 골편을 적절한 위치로 이동시키고, 이후 견인력 제어 모드

로 전환하여 수술 중 환자의 움직임에도 일정한 견인을 유지할 수 있다. 또한 IMU 센서를 이용한 환자의 자세와 가속도 데이터도 측정되어 제공된다. 이러한 구성은 골절 수술의 정밀도를 높여 수술 효율성을 극대화하고, 과도한 견인으로 인한 조직 손상을 방지할 수 있다.

### III. 결 론

본 논문에서는 골절 정복 수술의 효율성과 안전성을 높이기 위해 견인 거리와 견인력을 정밀하게 제어할 수 있는 알고리즘을 제안하였다. 제안된 정형외과 수술용 포지셔닝 로봇 시스템은 거리 제어 모드와 힘 제어 모드의 두 가지 기능을 통해 의사의 부담을 경감하고, 과도한 견인으로 인한 조직 손상을 방지한다. 이를 통해 골절 수술을 더욱 정교하게 수행할 수 있으며, 환자에게 최적의 치료 결과를 제공할 수 있다.

### 참 고 문 헌

- [1] Y. H. Choi, S. J. Lee, C. W. Oh, H. H. Bang, H. W. Lee, & C. W. Park, "Locking Screw Insertion Path Guidance System in Minimally Invasive Fracture Reduction Surgery," *Journal of Institute of Control, Robotics and Systems*, vol. 29, no. 5, pp. 458-465, 2023
- [2] M. S. Rashid, S. Aziz, S. Haydar, S. S. Fleming, and A. Datta, "Intra-operative fluoroscopic radiation exposure in orthopaedic trauma theatre," *Eur. J. Orthop. Surg. Traumatol*, vol. 28, no. 1, pp. 9-14, Aug. 2018.
- [3] J. W. Kim and J. J. Kim, "Radiation Exposure to the Orthopaedic Surgeon during Fracture Surgery," *Journal of the Korean Orthopaedic Association*, vol. 45, no. 2, pp. 107-113, 2010.

# 무인 로봇 가감속 제어를 통한 정지 최소화 충돌 방지 기술

이종해\*, 김대영\*, 안병권\*\*

## Collision Avoidance Technology with Minimum Stopping Through Acceleration and Deceleration Control of Autonomous Robots

JongHae Lee\*, DaeYoung Kim\*, and ByungKuon Ahn\*\*

본 논문에서는 다수의 무인 로봇(Autonomous Robot)들이 제한된 공간에서 충돌 없이 효율적으로 운용될 수 있도록, 가감속 제어를 통한 정지 최소화 충돌 방지 기술을 제안한다. 다수의 무인 로봇이 동시에 운용되는 환경에서 충돌 방지는 중요한 문제로, 이를 위해 일반적으로 무인 로봇의 정지와 대기 시간이 발생한다[1]-[3]. 빈번한 정지는 전체 시스템의 물동량 저하로 이어지므로 정지 시간을 최소화하면서도 충돌을 방지할 수 있는 기술이 필요하다. 본 연구에서는 이 과정에서 두 대 이상의 무인 로봇이 합류하는 구역에서 감속 구간과 정지 구간을 정의한다. 우선 진입한 로봇이 감속 없이 합류 구역을 지나가도록 하고, 뒤따르는 로봇이 감속 구간에서 속도를 줄이며 주행한다. 교차 구역 내에서는 한 번에 한 대의 로봇만 진입할 수 있도록 제한하여 불필요한 정지를 최소화하고 각 로봇이 적절한 속도로 접근하도록 하여 원활한 흐름을 유지한다. 이러한 통제는 각 로봇의 주행 속도와 제동 거리, 폭, 그리고 이동 경로 등을 기준으로 실시간 조정된다. 본 연구에서 제안한 정지 최소화 충돌 방지 기술은 다수의 무인 로봇이 운영되는 복잡한 환경에서도 안정적인 충돌 방지와 효율적인 교통 흐름을 가능하게 한다. 이를 통해 로봇의 이동 시간을 단축하고, 운영의 생산성을 높일 수 있다. 본 연구는 향후 더욱 복잡한 물류 환경에서의 확장 가능성을 제시하며, 다양한 산업 분야에서 무인 로봇 운영 효율성을 증대할 수 있는 기반 기술이 될 것이다.

### Key words

autonomous robot, collision avoidance, multi-agent

### 참고문헌

- [1] A. Felner, R. S. Shimony, E. Boyarski, M. Goldenberg, G. Sharon, N. Strurtevant, G. Wagner, and P. Surynek, "Search-Based Optimal Solvers for the Multi-Agent Pathfinding Problem: Summary and Challenges," International Symposium on Combinatorial Search, Pittsburgh, USA, Jun. 2017.
- [2] R. Stern, N. Sturtevant, A. Felner, S. Koenig, H. Ma, T. Walker, J. Li, D. Atzmon, L. Cohen, T. K. Kumar, R. Bartak, E. Boyarski, "Multi-Agent Pathfinding: Definitions, Variants, and Benchmarks," International Symposium on Combinatorial Search, Napa, USA, Jul. 2019.
- [3] 정의황, 조재훈, 김용태, "다중 물류로봇의 협업을 위한 우선순위 기반 충돌회피 알고리즘," 한국지능시스템학회 논문지, vol. 29, no. 5, pp. 359-364, Oct. 2019.

---

\*(주)러셀로보틱스, jhlee@erussell.co.kr, dykim@erussell.co.kr, \*\*승실대학교, abg6312@ssu.ac.kr(교신저자)

# 다중 센서를 활용한 제조 공정 장비의 예지보전 기술 연구

변상봉\*, 이종윤\*\*, 한동철\*, 이성훈\*, 정용안\*, 조수현\*

## A Study on Predictive Maintenance Technology for Manufacturing Equipment Using Multiple Sensors

Sang-Bong Byun\*, Jong-Youn Lee\*\*, Dong Cheul Han\*, Sung hun Lee\*, Yong An Jung\* and Soo-Hyun Cho\*

본 논문은 제조 공정 장비의 고장을 다양한 센서를 통해 사전에 감지하는 기술에 대한 연구를 다룬다. 장비의 전기적 상태를 모니터링하기 위해 전류 및 전압 센서를 사용하여 전기적 이상 신호를 탐지하고, 진동 센서를 통해 장비 이상 시 발생하는 비정상적인 진동 패턴을 감지한다[1]-[2]. 또한, 초음파 수신기를 통해 이상 발생 시 발생하는 특이 소음을 초음파[3] 신호로 탐지하는 시스템을 연구한다. 이 시스템을 구축하기 위해 각 센서가 감지하는 정상 및 비정상 데이터를 연구하고, 이를 기반으로 이상 상황을 구분할 수 있는 기준을 설정하는 것이 필수적이다. 예를 들어, 전류 및 전압 센서는 불량 발생 시 특정한 주파수 변화나 전압 강하 등의 특징적인 신호를 나타내고, 진동 센서는 불량 상태에서 비정상적인 진동 세기나 주파수 변화를 보여준다. 초음파 수신기 또한 평소와 다른 주파수 범위의 소음을 감지함으로써 고장의 전조를 파악할 수 있다. 이 기술은 장비의 이상을 조기에 탐지함으로써, 고장이 발생하기 전에 수리나 유지 보수를 수행할 수 있게 하여 불필요한 장비 가동 중단을 방지하고 생산성을 유지하는 데 기여한다. 이를 통해 장비의 비정상적인 가동으로 인한 금전적 손실을 최소화하며, 예측 정비 및 예방 정비의 효율성을 높일 수 있다. 이러한 고장 감지 시스템은 장비의 운영 신뢰도를 향상시키며, 제조 공정의 지속적인 최적화를 가능하게 한다. 향후 다양한 공정 장비에 대응하기 위한 데이터 종류 추가에 대한 연구를 진행할 예정이다.

### Key words

predictive maintenance, ultrasonic wave, artificial intelligence, sensor

### 참고문헌

- [1] A. Abid, M. T. Khan, J. Iqbal, "A review on fault detection and diagnosis techniques: basics and beyond", *Artificial Intelligence Review*, vol. 57, pp. 3639-3664, Nov. 2020.
- [2] W. Xie, S. Gao, X. He, X. Zhu, J. Huang, W. Chen, L. Ma, H. Shu, W. Cao, L. Jiang, Z. Shu, "Vibration Data Analysis for Fault Detection in Manufacturing Systems", *ICIEA 2022*, pp. 851-857, Dec. 2022.
- [3] A. Moshrefi, F. Nabki, "Advanced Industrial Fault Detection: A Comparative Analysis of Ultrasonic Signal Processing and Ensemble Machine Learning Techniques", *Applied Sciences* 2024, vol. 14, pp. 6397, Jul. 2024.

---

\* 구미전자정보기술원 지능형디바이스연구센터, email : sbbyun@geri.re.kr(교신저자), cataegu07@geri.re.kr, leesh@geri.re.kr, yajung@geri.re.kr, shcho@geri.re.kr

\*\* 금오공과대학교 일반대학원 반도체시스템공학과, email : ljyun0807@naver.com

※ 이 연구는 2024년도 산업통상자원부 및 한국산업기술기획평가원 연구비 지원에 의한 연구임(유도탄 제조 공정의 디지털전환을 위한 AI기반 고신뢰성 로봇-장비 및 위험관리 통합운영 기술 개발, RS-2024-00445778)



# 서버간 상태 체크를 통한 검침데이터 유실방지 방법에 관한 연구

양아름\*, 유해근\*\*, 전지예\*\*\*

## Research on methods to prevent loss of meter reading data through status checking between servers

Yang Areum\*, Yoo Haegeun\*\*, and Jeon Jiye\*\*\*

### 요 약

AMI 상위서버시스템은 각 단위시스템의 고유기능에 집중하고 데이터 처리가 필요시에는 네트워크를 통해 시스템간 송·수신하여 처리하는 구조로 되어있다. 이러한 구조는 데이터 전송시 네트워크의 의존도는 증가하게 되어 네트워크 장애 등의 문제로 데이터의 유실될 상황이 발생하게 된다. 본 논문에서는 Heartbeat Check를 활용하여 검침데이터의 유실 가능성을 최소화한다.

### Abstract

The AMI upper server system focuses on the unique functions of each unit system and, when data processing is necessary, is structured to process data by sending and receiving between systems through the network. This structure increases the dependency on the network when transmitting data, resulting in data loss due to problems such as network failure. In this paper, Heartbeat Check is used to minimize the possibility of loss of meter reading data.

### Key words

meter reading data, heartbeat, status check

## I. 서 론

현대 시스템에서 실시간 데이터 처리의 중요성을 증가되어지고 있다. 사용자들은 정보를 실시간으로 보길 원하고, 운영자들은 데이터 유무 또는 출력되는 데이터의 형태에 따라 장애 등 문제에 대한 즉각적인 대응을 할 수 있기 때문이다. 전력산업에서

핵심요소 중 하나인 AMI도 실시간으로 데이터를 운영자 및 사용자에게 제공한다.

AMI(Advanced Metering Infrastructure)는 에너지를 효율적으로 관리할 수 있는 시스템 및 서비스로 소비자에게 실시간으로 에너지 사용량에 대한 정보가 제공되어 소비자 혹은 자동화된 기기 제어를 통해 에너지 사용을 제어하는 인프라를 뜻한다.

AMI시스템에서의 검침데이터는 하위설비인 전력

\*한전KDN(주), areum\_0603@kdn.com, \*\*한전KDN(주), yoohaegeun.59@kdn.com, \*\*\*한전KDN(주), jiye\_0524@kdn.com,

량계, SMGW(Smart Meter GateWay), 데이터집중장치(DCU)를 거쳐 상위서버시스템으로 도달하며 검침데이터는 설비데이터와 계량데이터로 이루어져 있다. 수집된 검침데이터를 처리하는 AMI 상위서버시스템은 크고 작은 단위시스템들이 유기적으로 연계되어 있고, 각각의 단위시스템들은 시스템별 역할(기능)을 분산하여 결합도가 높아 서버간 데이터를 송수신시 네트워크에 대한 의존도가 증가하게 되어 네트워크 장애 등의 문제로 데이터의 유실될 상황이 발생하게 된다.

이에 본 논문에서는 네트워크 상태와 데이터를 수신하는 대상 단위시스템의 상태를 파악하고, 상태 이상시 데이터를 임시저장소에 저장하고, 상태가 복구되면 저장된 데이터를 재송신하여 검침데이터의 유실가능성을 최소화하는 방안을 제시하고 있다.

## II. 검침데이터 유실 최소화 방법

본 논문에서는 검침데이터의 유실을 최소화하기 위해 상위 서버시스템 내 단위시스템별(서버) 데이터 송신부(이하 송신부)와 데이터 수신부(이하 수신부)로 나뉜다. 그리고 모든 데이터 송·수신에 대한 관장을 하는 데이터 관리부(이하 관리부)로 구성된다.

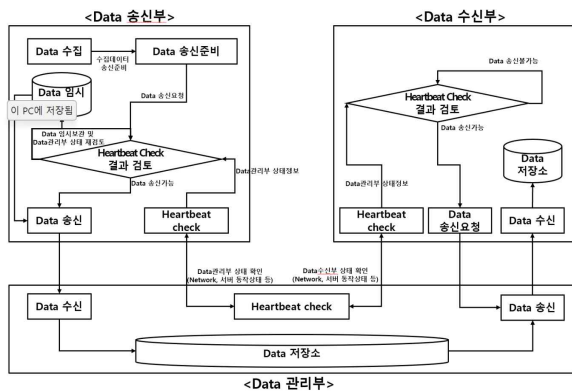


그림 1. 데이터 수집안정성 강화방법  
Fig. 1. How to enhance data collection stability

기본적인 동작은 DCU 등 하위설비로부터 실시간으로 받은 검침데이터는 송신부에서 관리부로 송신되어지고, 관리부에서 수신부로 송신됨으로써 안정적으로 데이터가 이동하게 된다. 이때, 송신/수신/관

리부에서는 Heartbeat Check 모듈을 통해 상태체크를 한다.

송신부는 하위설비로부터 받은 검침데이터는 관리부로 송신한다. 또한 송·수신부의 Heartbeat Check 모듈은 데이터 관리부의 Heartbeat Check 모듈을 통해 시스템 및 네트워크 상태를 수시로 확인하면서 데이터 송·수신 가능여부를 실시간으로 체크한다.

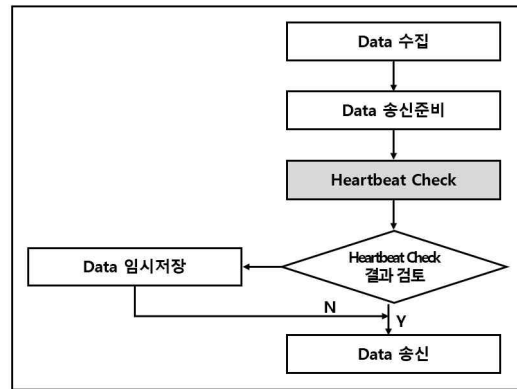


그림 2. 데이터 송신부 Flow  
Fig. 2. Data Transmission Department Flow

송신부는 관리부가 데이터 송신이 가능한 경우에만 데이터를 송신한다. 만약 시스템 및 네트워크 상태에 문제가 발견된다면, 데이터 송신을 정지하고 시스템 내부의 데이터 임시 저장소에 데이터를 임시 보관한다.

또한 송신부는 데이터 송신 가능 시점이 되면, 임시 저장소에 저장된 데이터를 우선적으로 송신한 후 실시간으로 수집된 데이터를 송신 처리가 이루어진다.

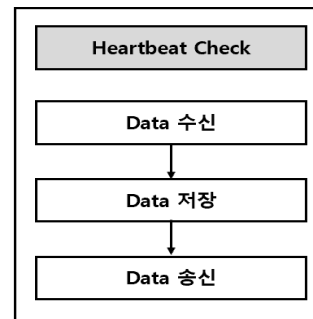


그림 3. 데이터 관리부 Flow  
Fig. 3. Data Management Flow

관리부는 송신부에서 수집한 데이터를 데이터 저장소에 실시간으로 저장되며 관리부의 Heartbeat Check 모듈은 송신부에서 요청한 서버 동작상태를 실시간으로 응답해주며, 자신의 데이터 수신 가능 여부를 제공한다. 또한, 관리부는 수신부가 데이터 송신이 가능한 경우에만 데이터를 송신하며, 시스템 및 네트워크 상태에 문제 발견시 데이터 송신을 정지한다. 데이터 송신 가능시점이 되면 저장된 데이터를 송신한다.

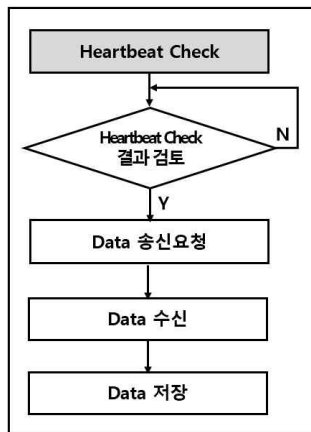


그림 4. 데이터 수신부 Flow  
Fig. 4. Data Reception Flow

수신부는 관리부로부터 송신한 데이터를 실시간으로 수신받아 처리하며, Heartbeat Check 모듈은 관리부에서 요청한 서버 동작상태를 실시간 응답을 해주며, 자신의 데이터 수신 가능여부를 제공한다.

### III. 결 론

본 논문에서는 검침데이터의 유실 가능성을 최소화하는 방안을 제안하였다. 제안된 방법은 데이터 관리부, 송신부, 수신부 각각 Heartbeat Check 모듈을 통해 서버 및 네트워크 상태를 실시간으로 파악하고, 상태 이상 시 송신할 데이터를 임시 저장소에 저장하고, 상태가 복구되었을 때 저장된 데이터를 재송신하는 구조이다.

기존 시스템의 경우 서버간 직접적으로 통신하거나 중간매개체를 활용하여 통신한다. 이 경우 수신

받을 시스템의 상태이상으로 데이터를 받지 못하는 상황이 발생하더라도 송신하는 시스템에서는 데이터 송신시도를 하게 될 것이다. 데이터가 송신되어도 수신 받을 시스템이 상태이상으로 받을 수 없기 때문에 이는 곧 데이터 유실로 직결된다.

AMI 상위서버시스템의 경우 사용자에게 보여지는 대부분이 검침데이터이며 이 데이터가 요금정산에 기반 데이터가 되기 때문에, 유실이 된다면 운영자 뿐만 아니라 사용자에게도 즉각적인 정보를 제공할 수 없는 어려움이 생긴다.

앞서 제안된 방법을 토대로 구성한다면 실시간으로 Heartbeat Check 모듈을 통해 상태이상을 파악하여 데이터 송수신을 진행하기 때문에 데이터 유실을 최소화할 수 있는 강력한 방안으로 보여진다.

### 참 고 문 헌

- [1] Do Hwan Oh and Jae Hyung Park, " Analysis of Data Processing Efficiency using Duplicated Data Removal in AMI", Smart Media Journal, Vol10, No. 2, pp.9-15, June, 2021
- [2] Seung Jae Oh, "Comparative Analysis of Asynchronous Communication Technology in Web Data Communication", Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, Vol7, No. 8, pp.419-426, August, 2017
- [3] Jea Suk Song and Sun Young Bae, "Open Source based Smart Meter Real-time Data Acquisition and Management using MQTT Protocol", Journal of Software Assessment and Valuation, Vol 20, No. 3, pp.75-87, September, 2024

# 마사지와 온열 요법이 결합된 지압베드의 통증완화 및 체성분 변화에 관한 임상연구

김재훈\*, 안도현\*\*, 김해용\*\*\*, 김수민\*, 장호경\*, 허성필\*\*\*\*

## A Clinical Study on Pain Relief and Body Composition Changes in Acupressure Bed Combining Massage and Thermal Therapy

Jaehun Kim\*, and Dohyun Ahn\*\*, Haeyong Kim\*\*\*, Soomin Kim\*, Hogyong Chang\*, and Sungphil Heo\*\*\*\*

### 요 약

비만의 치료법으로는 약물요법, 수술요법, 운동요법, 식이요법 등이 이용되고 있으나 해당 요법들은 부작용 및 위험성 등이 있으며, 장기간에 걸쳐 시행하여야 한다는 단점이 있다. 이러한 치료법 이외에 보조적인 방안으로는 온열요법과 마사지 요법이 있다. 온열요법의 경우 목, 등의 근막 이완 및 통증 완화에 영향을 주며, 마사지의 경우 통증을 감소시키고, 근육을 자극하여 혈액순환에 도움을 준다. 그리고 지방 연소가 활발해져 체중 감소에 효과가 있다고 보고 되고 있다. 따라서 본 연구에서는 통증 완화 및 체지방 감소를 위하여 온열과 기계적 마사지가 조합된 지압 침대를 사용하였을 때 사용자의 통증 척도 및 체성분의 변화를 확인하기 위해 임상적인 실험을 하였다. 실험은 지압 침대를 사용하기 전과 후로 나누어 진행하였다. 실험 결과 통증 척도와 체지방량, 체지방률에 대해서는 유의미한 감소가 보였지만 BMI에 대해서는 유의미한 감소를 확인하지 못했다. 이는 온열요법과 기계적 마사지가 체지방량과 체지방률 감소에 도움을 줌으로써 비만 예방과 관리, 다이어트에 도움이 되고, 통증 척도의 감소를 통해 스포츠 재활에 도움이 될 수 있음을 시사하며, 향후 정량화를 위해서는 추가적인 연구가 요구된다.

### Abstract

Treatments for obesity include pharmacotherapy, surgical interventions, exercise therapy, and dietary management, but these methods have associated side effects, risks, and the drawback of requiring long-term implementation. In addition to these treatment methods, thermotherapy and massage therapy are available as supplementary approaches. Thermotherapy aids in relaxing the fascia in areas such as the neck and back and alleviates pain, while massage therapy reduces pain, stimulates muscles, and improves blood circulation. It is also reported that these methods can enhance fat burning, contributing to weight loss. Therefore, this study conducted a clinical experiment to observe changes in users' pain scale and body composition by using a bed that combines thermotherapy and mechanical massage to manage pain and body fat increase. The experiment was divided into two phases: before and after using the bed. The results showed significant reductions in the pain scale, body fat mass, and body fat percentage, although there was no significant reduction in BMI. These findings suggest that thermotherapy and mechanical massage may aid in the prevention, management, and reduction of obesity by decreasing body fat mass and percentage, as well as assist in sports rehabilitation by reducing the pain scale. Further research is required to quantify these effects.

### Key words

massage, heat, acupressure bed, body, pain, composition

\*대구한의대학교 재활의료공학과, jaehun3817@naver.com, \*\*\*\*강릉원주대학교 정보통신공학과, spheo@gwnu.ac.kr  
\*\* 경북대학교 의용생체공학과, tigerace5012@nate.com, \*\*\*경북대학교 융합소프트웨어학과, twt@naver.com

※ 본 연구는 보건복지부·대구광역시·경상북도의 재원으로 한국한의학진흥원 한의약 산업 전주기 지원체계 구축사업의 지원을 받아 수행한 사업임

## 1. 서 론

비만은 전 세계적으로 매년 늘고 있다. 세계보건기구(WHO)에 따르면 비만 인구는 10억 명 이상으로 1990년과 비교해보았을 때, 2배 넘게 늘었고 그 중 어린이, 청소년의 비만 인구는 4배가량 증가한 것으로 조사되었다. 또한, 2022년 성인의 경우 43%가 과체중인 것으로 파악되었다. 대한민국도 빠르게 비만 인구가 늘고 있다. 대한민국 공식 전자 정보 누리집인 지표 누리에 따르면, 한국의 평균 비만율은 2015년 33.2%에서 2022년 37.2%로 약 4% 증가하였다. 이 통계는 비만 문제가 단순한 개인의 건강 문제를 넘어서 사회적 국가적 문제로 확산되고 있음을 시사한다.

비만은 체지방의 과도한 축적으로 체중 증가 및 통증이 유발되며, 이로 인해 심혈관 질환, 당뇨병, 고혈압, 소화기계, 근골격계 등의 건강 문제를 유발한다[1].

비만으로 인한 통증의 종류는 디스크·척추관협착증에 따른 허리통증, 고관절 충격·손상으로 발생하는 고관절통, 관절염에 따른 무릎 통증 등이 있으며, 이를 치료하지 않으면 더욱 심각한 합병증으로 이어질 수 있다.

비만의 치료법으로는 일반적으로 약물요법, 수술요법, 운동요법, 식이요법 등이 이용되고 있다. 약물요법과 수술요법의 경우 부작용에 대한 위험성이 발생하는 단점이 있다[2]. 운동요법과 식이요법의 경우는 장기간에 걸쳐 지속해서 시행하여야 하며, 단기간의 효과를 보기에는 어려운 것이 현실이다[3]. 이러한 요법들은 체중 감량에 있어 장기간 지속하는 데에 어려움이 많고, 5년 안에 실패할 확률이 95% 정도로 높은 것으로 보고 되고 있어 보조적인 방안의 필요성이 대두되고 있다[4]. 보조적인 방안으로는 온열요법과 마사지 요법이 있다. 온열요법의 경우에는 비침습적 통증 관리 방법으로 주목받고 있으며, 목, 등 근막 이완 및 통증 완화에 미치는 영향에 대한 연구가 있다[5]. 마사지는 혈액의 림프의 순환을 증가시키고, 통증을 감소시키며[6], 근육을 자극하여 혈액순환을 원활하게 하고 기의 순환을 좋게 하여 운동과 마찬가지로 지방 연소가

활발해져 체중감소에 효과가 있어 수동적인 운동으로써의 역할을 가지고 있다고 보고 되고 있다[7].

따라서 본 연구에서는 비만으로 인한 통증과 체지방 증가를 관리하기 위하여 온열과 기계적 마사지를 조합한 수직 상승압을 제공하는 지압 침대를 이용하여 사용자의 통증 척도 및 체성분 변화를 확인하고, 어떠한 영향을 미치는지에 대해 임상적인 실험을 진행하였다.

## II. 연구 방법

본 연구에서는 비만으로 인한 통증과 체지방 증가를 관리하기 위하여 온열과 기계적 마사지가 조합된 수직 상승압을 제공하는 지압 침대(3H-920-K)를 이용하여 사용자의 통증 척도 및 체성분의 변화를 확인하기 위해 임상적인 실험을 진행하였다.



그림 1. 임상 실험 모습  
Fig. 1. Clinical Trial

임상 실험은 온열과 기계적 마사지가 통증 척도 및 체성분의 변화에 어떠한 영향을 미치는지 보기 위하여 온열요법과 기계적 마사지가 조합된 침대를 사용하기 전과 후로 나누어 통증 및 체성분에 대한 변화를 살펴보았다.

실험장비는 그림 2와 같이 온열과 기계적 마사지가 조합된 수직 상승압을 제공하는 침대(3H-920-K)를 사용하였고, 통증 척도 측정을 위해서 통증 변화 조사(VAS)를 이용하였다. 그리고 체성분의 변화를 확인하기 위해 InBody社의 InBody270을 사용하였다.



그림 2. 실험에 사용된 침대(3H-920-K)  
Fig. 2. the bed used in the experiment (3H-920-K)

온열과 기계적 마사지가 조합된 침대를 사용하였을 때 실험 변화 차이를 보기 위해서 통증 척도는 80명을 대상으로, 체성분 측정은 46명을 대상으로 침대를 사용 전과 사용 후로 나누어 측정하였다.

통계적 유의미한 결과를 확인하기 위하여 통계 분석 프로그램 KESS(Korean Educational Statistical Software)를 사용하였다.

### III. 실험 결과

본 연구에서는 온열과 기계적 마사지가 조합된 수직 상승압을 제공하는 침대를 사용하여 사용자의 통증 척도 및 체성분을 측정한 결과 그림 3의 통증 척도 그래프를 보면 측정 전 평균 2.9점에서 측정 후 평균 2.2점으로 0.7점의 차이가 보였고 체성분 그래프인 그림 4를 보면 체지방량의 측정 전 평균 21.163kg, 측정 후 평균 20.735kg으로 0.428kg의 차이, BMI는 측정 전 평균 24.150, 측정 후 평균 24.209로 0.059의 차이, 체지방률은 측정 전 평균 30.239%, 측정 후 평균 29.520%로 0.719%의 차이가 났다.

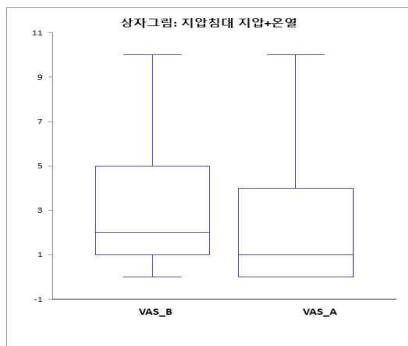


그림 3. 통증 척도(VAS) 측정 결과  
Fig. 3. Pain scale(VAS) measurement results

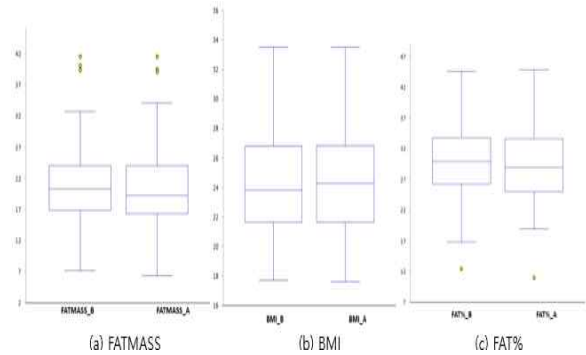


그림 4. 인체 전기 저항 측정 결과 (a) FATMASS (b) BMI (c)FAT%  
Fig. 4. Human Electrical Resistance Measurement Results (a) FATMASS (b) BMI (c)FAT%

이러한 결과로 보았을 때 온열요법과 기계적 마사지를 사용하였을 때 통증 척도, 체지방량, 체지방률에 대해서는 유의하게 감소하는 것으로 나타났지만, BMI에 대해서는 유의하게 감소하지 않는 것으로 나타났다.

### IV. 결 론

본 연구에서는 온열요법과 기계적 마사지가 조합된 지압 침대 사용 시 통증 척도, 체지방량, 체지방률이 유의하게 감소함을 확인하였다. 이는 온열과 기계적 마사지가 체지방량과 체지방률을 감소에 도움을 줌으로써 비만 예방과 관리, 다이어트에 도움이 되고, 통증 척도의 감소를 통해 스포츠 재활에 도움이 될 수 있음을 시사한다. 다만 침대 사용 시간이나 강도에 따라 통증의 감소와 체지방 감소 등이 실제 어느 정도 이루어지는가에 대한 정량적 지표를 도출하려면 추가적인 연구가 요구된다.

### 참 고 문 헌

- [1] S. H. Shin, J. I. Kang, & H. Y. Kim. (2020). "Mediating Effects of Perceived Stress in the Influence of Obesity Awareness on Health Behavior : A Comparative Study between Korean and Chinese Adults." *Journal of the Korea Contents Association*, 20(4), 441-453.
- [2] K. R. Kim & H. S. Bang. (2011). The effect of

- abdominal massage on the body composition, cardiorespiratory functions, isokinetic articular muscle functions and serum lipids in obese women. *Kor J Physical Education*, 50(5), 419-432
- [3] YONG-TAEK RHIM (2020). Effect of Mechanical Massage on Body Composition, Waist Circumference and Abdominal Fat Area in Middle-aged Obese Women. *Journal of Convergence Information Technology*, 10(3), 200-207.
- [4] Y. T. RHIM, (2020). Effect of Mechanical Massage on Body Composition, Waist Circumference and Abdominal Fat Area in Middle-aged Obese Women. , 10(3), 200-207.
- [5] J. Y. LEE, (2016). Effects of Gold Tool Heat Therapy on Myofascial Release and Pain Relief in the Neck and Back (Master's thesis, South Korea)
- [6] D. H. KIM (2013). Effects of Regional Massage on Respiratory and Circulatory Functions, Obesity Factors, and Immune Function in Middle-Aged Women (Doctoral dissertation, South Korea).
- [7] C. J. Lee, D. O. Lee, & S. Y. Park (2004). Effects of 10 Weeks Sports Massage on Blood Sugar, Blood Lipids and Body Composition in Obese Women. *Journal of Korean Association of Physical Education and Sport for Girls and Women*, 18(4), 13-25.

# 고령 우울증과 아밀로이드 베타 침착이 인지 저하에 미치는 영향에서 해마 위축의 매개 역할

석지우, 김재욱

## The Mediating Role of Hippocampal Atrophy in the Impact of Geriatric Depression and Amyloid Beta Deposition on Cognitive Decline

Jiwoo Suk\* and Jaek U. Kim\*\*

본 연구는 알츠하이머병(AD)의 병리적 진행 과정에서 나타나는 노인 우울증, 아밀로이드 베타(A $\beta$ ) 축적, 해마 위축, 인지 저하 간의 상호작용을 보다 명확히 규명하기 위해 수행되었다[1,2]. AD는 노인 인구에서 높은 발병률을 보이며, 퇴행성 뇌질환의 주요 원인으로 작용하지만, 이러한 병리적 요소들이 발현되는 순서와 상호작용에 대한 이해는 제한적이다. 이에 본 연구는 60세 이상의 노인 집단에서, 특히 경도 인지 장애(MCI)를 가진 집단을 포함하여 A $\beta$  축적과 해마 부피 감소가 인지 저하에 미치는 영향을 분석하였다. 연구 대상은 450명의 고령자로 구성되었으며, A $\beta$  축적 증가와 해마 위축, 노인 우울증이 전반적인 인지 기능 저하에 유의미한 영향을 미치는 것을 확인하였다. 특히, A $\beta$  축적과 노인 우울증이 해마 부피를 매개로 하여 인지 저하에 직접적 및 간접적으로 기여하는 경로가 나타났으나, A $\beta$  축적과 현재 노인 우울증 간에는 직접적인 연관성이 없어, 이 두 요인이 독립적으로 해마 위축과 궁극적인 인지 저하에 기여하는 경로를 형성하고 있음을 시사한다. 이러한 결과는 AD 병리적 진행 과정에서 심리적 요인과 생물학적 요인이 어떻게 상호작용하는지를 이해하는 데 중요한 단서를 제공하며, 심리적 요인을 고려한 통합적 치료 접근이 AD 위험이 높은 고령층의 인지 기능 보호에 필요함을 제안한다.

### Key words

Geriatric depression, A $\beta$  deposition, hippocampal atrophy, general cognition, serial mediation, mild cognitive impairment

### 참고문헌

- [1] J. E. Baker, Y. Y. Lim, R. H. Pietrzak, J. Hassenstab, P. J. Snyder, C. L. Masters, et al., "Cognitive impairment and decline in cognitively normal older adults with high amyloid- $\beta$ : a meta-analysis," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 6, pp. 108-121, 2017.
- [2] J. Li, X. Li, F. Chen, W. Li, J. Chen, and B. Zhang, "Studying the Alzheimer's disease continuum using EEG and fMRI in single-modality and multi-modality settings," *Reviews in the Neurosciences*, vol. 2024, no. 0, 2024.

\* 한국한의학 연구원, 디지털임상 연구부, suk6124@kiom.re.kr,

\*\* 한국한의학 연구원, 디지털임상 연구부, jaeukkim@kiom.re.kr

※ 본 연구는 한국정부가 지원하는 한국한의학연구원의 연구비(KSN1824130, KSN2312022)와 과학기술정보통신부가 지원하는 한국뇌연구원의 KBRI 기초연구사업(24-BR-03-05)을 통해 수행되었음.



# 첨단 제조 공정에서 이기종 로봇간 통신 성능 최적화를 위한 동기화 오차 추정 방법

정용안\*, 변상봉, 한동철, 조수현, 이성훈\*\*

## Synchronization Method for Optimization of Communication Performance Between Heterogeneous Robots in Advanced Manufacturing Process

Yong-An Jung\*, Sang-Bong Byun, Dong-Cheul Han, Soo-Hyun Cho, and Sung-Hun Lee\*\*

스마트 공장에는 다양한 로봇들과 첨단 설비들이 존재하며, 이들의 통합 제어를 위한 대표적인 통신 시스템으로 5G 이동통신이 있다. 스마트 공장 내 최적의 네트워크를 제공하기 위해서는 저지연, 초고속, 초연결의 서비스가 제공되어야 하며, 이를 위해 3GPP 5G NR에서는 OFDM 파형을 적용하였다. OFDM 기술은 멀티패스 페이딩 채널에 강인하며, 초고속의 데이터 전송을 가능하게 한다. 그러나, OFDM 기반의 무선통신 시스템은 시간 및 주파수 동기화 오차에 매우 민감하다. 이러한 동기화 오차는 송신기와 수신기 사이의 오실레이터 성능 차이에 의해 발생할 수 있으며, 도플러 효과에 의해 발생할 수도 있다. 동기화 오차의 존재는 직교성을 상실시켜 inter carrier interference (ICI)가 발생하기 때문에, 수신기에서 정확하게 신호를 복조할 수 없게 된다. 특히 V2X 장치가 네트워크 커버리지 영역 밖에서 서로 직접 통신해야 할 때 더욱 성능은 악화될 수 있다. 본 논문에서는 이러한 문제를 해결하기 위해 정확하게 동기화 오차를 추정할 수 있는 방법을 제안하였다. 제안된 기법은 5G NR Sidelink 시스템을 고려하였으며, synchronization signal block (SSB)의 특성을 이용하여 상관 연산을 수행함으로써 정확하게 동기화 오차를 추정할 수 있다.

### Key words

Synchronization, 5G NR, D2D Communication, Advanced Manufacturing Process,

### 참고문헌

- [1] Lu, Q., Jung, H., and Kim, K, "Optimization-based approach for resilient connected and autonomous intersection crossing traffic control under V2X communication," IEEE Trans. Intell. Veh. 2022, 7, 354-367.

---

\*구미전자정보기술원, yajung@geri.re.kr

\*\*구미전자정보기술원, leesh@geri.re.kr

※ This work was supported by the Korea Planning & Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korean government (MOTIE) (RS-2024-00445778, Development of AI-based high reliability robot-equipment and risk management integrated operation technology for digital transformation of missile manufacturing process).

# 정형외과 수술용 포지셔닝 로봇에서 인간-로봇 상호작용을 위한 제어기 개발

이현우\*, 방현희\*, 이현주\*\*, 정상현\*\*\*, 박철우\*

## Development of a controller for human-robot interaction in an orthopedic surgical positioning robot

Hyun-Woo Lee\*, Hyun-Hee Bang\*, Hyun-Joo Lee\*, Sanghyun Joung\*\*, and Chul-Woo Park\*

### 요 약

본 논문에서는 정형외과 수술에서 환자 체위를 조정할 수 있는 다자유도 포지셔닝 로봇의 제어기를 제안한다. 제안하는 제어기는 인간-로봇 상호작용을 위한 멀티모달 센서 신호 처리부, 연산부 및 구동기 출력부로 구성된다. 정형외과 수술용 포지셔닝 로봇 제어기는 의료진의 의도를 인식하여 최적의 힘으로 환부를 이동 및 고정할 수 있어 수술의 안전성과 효율성을 높일 수 있을 것으로 기대된다.

### Abstract

This paper proposes a controller for a multi-degree-of-freedom positioning robot that can adjust the patient's position in orthopedic surgery. The proposed controller consists of a multi-modal sensor signal processing unit, a computation unit, and an actuator output unit for human-robot interaction. The positioning robot controller for orthopedic surgery is expected to increase the safety and efficiency of surgery by recognizing the intention of the medical staff and moving and fixing the affected area with optimal force.

### Key words

orthopedic surgery, patient positioning robot, multi-modal sensor, human-robot interaction

## 1. 서 론

정형외과 수술은 국내 상위 10개 대학병원에서 전체 수술 건수의 평균 19.4%를 차지하고 있으며, 골다공증과 골관절염 같은 뼈 질환의 유병률 증가

와 고령 인구의 증가로 인해 지속적으로 증가하고 있다[1,2]. 이러한 수술에는 집도의 외에 수술 보조 전문의가 참여하며, 이들은 환부의 견인과 환자의 자세 조정 등 중요한 역할을 수행한다. 특히, 골절 수술에서는 전문의가 부러진 다리뼈를 맞추기 위해

\* 경북대학교 융합의료기기로봇연구소, hwlee.mdrip@gmail.com, hyuntomb@gmail.com, pcw.mdrip@gmail.com

\*\* 경북대학교 의과대학 정형외과학교실, lidmania@daum.net

\*\*\* 주식회사 에어스, shjoung@airsurgical.net

※ 본 논문은 2024년도 과학기술정보통신부의 재원으로 연구개발특구진흥재단의 지원을 받아 수행된 연구임 (2024-DG-RD-0012)

약 300N의 힘으로 견인해야 하며, 무릎 인공관절 수술에서는 적절한 자세로 환자의 다리 위치를 조정하고 고정해야 한다. 이러한 과정에서 수술 보조 전문의의 체력적 부담이 크고, 수술 시간과 환자의 회복에 영향을 미칠 수 있는 점이 문제로 지적되고 있다[3].

따라서 정형외과 수술 절차에 맞춰 환자 체위 조정이 가능한 다자유도 포지셔닝 로봇을 개발하고 있고, 본 논문에서는 환자의 신체 상태와 의료진의 수술 중 제어 의도를 인식하여 제어 명령을 생성하는 인간-로봇 상호작용용 제어기를 제안한다. 제안하는 제어기를 적용한 포지셔닝 로봇 시스템은 정형외과 수술에서 환부의 위치 이동과 고정을 최적의 힘으로 수행함으로써, 의료진이 더욱 중요한 시술에 집중할 수 있게 되며, 수술의 안전성 및 효율성을 높이는 데 중요한 역할을 할 것으로 전망된다.

## II. 정형외과 수술용 포지셔닝 로봇 제어기

정형외과 수술용 포지셔닝 로봇의 제어기는 그림 1과 같이 인간-로봇 상호작용을 위한 멀티모달 센서 신호 처리부, 연산부 및 구동기 출력부로 구성하였다.

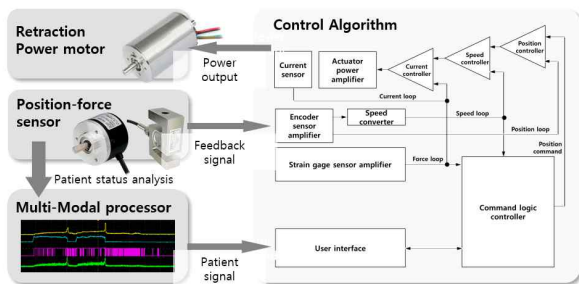


그림 1. 정형외과 수술용 포지셔닝 로봇 제어기 구성도  
Fig. 1. Block diagram of positioning robot controller for orthopedic surgery

센서 신호 처리부는 다양한 물리적 환경 변화를 감지하여 의료진의 제어 의도를 인식하고 데이터화할 수 있다. 의료진이 로봇을 잡고 움직이는 직접 교시 명령을 인식하고 사전 설정된 목표값을 실행할 수 있는 목표 지향형 인식 인터페이스를 구현하여 의료진의 의도를 더욱 명확하게 감지할 수 있도록 하였다.

또한, 다양한 센서 입력을 통해 제어 의도의 선별적 인식 및 다중 입력 처리 기술을 개발하여 로봇이 복잡한 상황에서도 적절하게 반응할 수 있다. 로봇이 수행해야 할 태스크에 따라 순차적 처리 절차와 즉각적 감응형 처리 상황을 구분하여 반응하는 제어 인터페이스를 설계하였으며, 이를 통해 로봇은 다양한 상황에 적절하게 대처할 수 있다.

그림 2에는 포지셔닝 로봇 제어기를 나타내었다. DSP 기반으로 제작된 제어기는 액추에이터의 동작 상태, 위치, 토크 센서를 측정할 수 있으며, 고속 연산과 빠른 액추에이터 제어 응답이 가능한 신호 처리 장치를 구현하였다. 의료진과 로봇이 공동으로 작업하는 공간에서 충돌과 같은 간섭 상태를 인식하고 즉각적으로 세이프티 프로파일을 수행할 수 있는 액추에이터 제어 기능과 환자의 움직임에 따른 위치변화와 환부에 가해지는 힘의 변화를 인식하여 상호작용할 수 있는 알고리즘이 구현되었다. 또한 센서 입력 상태 및 로봇 제어 출력 명령을 감시하기 위한 디버깅 시스템을 구축하여, 로봇의 동작을 실시간으로 확인하고 조정할 수 있는 기능을 제공함으로써 안전하고 효과적인 인간-로봇 상호작용을 지원할 수 있도록 하였다.



그림 2. 정형외과 수술용 포지셔닝 로봇 시스템 제어기  
Fig. 2. Controller of the positioning robot system for orthopedic surgery

## III. 결 론

제안하는 정형외과 수술용 포지셔닝 로봇 제어기는 환자의 상태에 최적의 힘으로 수술 시 자세를 유지하여 집도의를 보조하고, 과도한 힘이 환부에 가해지는 것을 원천적으로 차단하여 합병증 및 의

료사고를 방지할 수 있다. 이 시스템을 정형외과 수술에 적용할 경우 수술 보조 전문의의 역할을 로봇이 수행할 수 있어 의료 인력을 효율적으로 운영할 수 있으며, 정형외과 수술의 전반적인 품질 향상에 기여할 것으로 기대된다.

### 참 고 문 헌

- [1] H. J. Kang and Y. S. Ko, "Increasing incidence of osteoporotic fractures in Korea: a trend analysis from 2008 to 2018," *Osteoporosis International*, 33(4), pp. 1007-1014, 2022.
- [2] S. J. Park and K. J. Kim, "A Review of Total Knee Arthroplasty in the Elderly Population: Current Trends and Future Directions," *Korean Journal of Orthopedic Surgery*, 25(5), pp. 285-292, 2020.
- [3] C. W. Park, H. W. Lee, J. H. Yu, Y. K. Park, I. H. Park, & S. H. Joung, "Development of a surgical instrument insertion guidance system for fracture reduction surgery," *Journal of the Institute of Electronics and Information Engineers*, 55(6), pp. 99-106, 2018.

# 초해상화를 통한 산업 환경 내 작은 객체 탐지 성능 향상

이정민\*, 김정호\*\*

## Enhanced Object Detection for Small Items in Industrial Environments

Jeongmin Lee\*, and Jungho Kim\*\*

### 요약

산업 현장에서의 자동화 및 안전성 강화를 위해 작은 객체 인식(Small Object Detection, SOD)은 필수적인 과제이다. 그러나 작은 객체들은 해상도와 특징 정보가 제한적이기 때문에 기존 객체 탐지 모델에서 성능이 저하되는 한계를 가지고 있다. 이에 따라 본 연구는 YOLOv8과 초해상화(Super-Resolution, SR) 기법을 결합하여 산업 환경 내 작은 객체 탐지 성능을 강화하는 방법을 제안한다. YOLOv8의 첨단 백본 네트워크와 다단계 피쳐 융합 구조를 활용하여 작은 객체 탐지 성능을 극대화하며, SR 기법을 통해 고해상도 특징을 추가하여 객체 인식의 정확도를 높인다. 실험 결과, 본 연구에서 제안하는 접근법은 다양한 크기와 복잡한 배경을 포함한 산업 환경에서 작은 객체의 인식 성능을 크게 향상시키는 것으로 확인되었다. 이러한 연구는 산업 현장에서의 실시간 객체 인식과 안전성 강화에 기여할 수 있을 것으로 기대된다.

### Abstract

Small Object Detection (SOD) is essential for enhancing automation and safety in industrial environments. However, due to limited resolution and feature information, existing object detection models often underperform when detecting small objects. This study proposes a method that combines YOLOv8 and Super-Resolution (SR) techniques to enhance the performance of small object detection in industrial settings. By utilizing YOLOv8's advanced backbone network and multi-scale feature fusion structure, this approach maximizes the detection capability for small objects. Additionally, SR techniques are employed to enhance high-resolution features, improving detection accuracy. Experimental results confirm that this approach significantly improves detection performance for small objects in industrial environments with various object sizes and complex backgrounds. This research is expected to contribute to real-time object detection and safety enhancement in industrial applications.

### Key words

computer vision, small object detection, super-resolution, YOLO, industrial environments

---

\*한국전자기술연구원, jeanmichel@keti.re.kr \*\*한국전자기술연구원, jhkim77@keti.re.kr (교신저자)

※ This research was financially supported by the Ministry of Trade, Industry and Energy(MOTIE) and Korea Institute for Advancement of Technology(KIAT) through the International Cooperative R&D program. (No. P0019804, Digital twin based intelligent unmanned facility inspection solutions)

## I. 서 론

작은 객체 인식(Small Object Detection, SOD)은 산업 현장에서 자동화와 안전성을 보장하기 위해 필수적인 과제이다. 특히, 기계 부품이나 안전 표지판과 같은 작은 객체들이 중요한 정보를 포함하고 있어 이를 정확하게 탐지하는 것이 매우 중요하다. 그러나 작은 객체는 해상도가 낮고 특징 정보가 제한적이기 때문에 기존 객체 탐지 모델에서 성능이 저하되는 문제가 있다. 이러한 문제를 해결하기 위해 기존 연구들은 다중 스케일 학습, 지식 증류, 고해상도 특징 맵 사용 등을 통한 탐지 성능 개선을 시도하였다. 예를 들어, Dynamic Coarse-To-Fine Learning은 고해상도 맵에서 작은 객체를 효율적으로 탐지하기 위해 단계별 쿼리 방식을 적용하며 [1], ScaleKD는 객체 크기에 맞춘 스케일 의존 지식 증류를 통해 작은 객체 탐지 성능을 강화한다 [2]. Temporal Context는 시간적 정보의 활용을 통해 작은 객체 인식 성능을 향상시키는 방법을 제공한다 [3]. 그러나 이러한 기법들은 연산 비용이 높거나 성능 개선이 불균형적이라는 한계를 가지고 있어, 실시간 감시가 필요한 산업 현장에서의 적용이 어려운 문제가 있다.

YOLOv8 모델은 첨단 백본 네트워크와 다단계 피쳐 융합 구조를 활용하여 작은 객체 탐지 성능을 크게 향상시키며 [4], 다양한 크기의 특징 맵을 효과적으로 통합해 작은 객체 탐지 성능을 극대화한다. 또한 UAV와 같은 실시간 감시 상황에서 제한된 자원으로도 높은 정확도를 유지할 수 있으며, PIoU와 같은 정교한 손실 함수를 통해 정확도와 학습 속도를 향상시킨다 [5][6]. 하지만 YOLOv8에도 한계가 있다. 복잡한 산업 환경에서 작은 객체가 밀집된 장면에서는 탐지 정확도가 저하될 수 있으며, 계산 자원 소모와 추론 속도 면에서 최적화가 필요하다. 또한 다양한 조명 조건과 배경 복잡성이 높은 산업 장면에서의 일반화 성능이 부족할 수 있다.

본 연구에서는 이러한 한계를 극복하고자, YOLOv8과 초해상화(SR) 기법을 결합하여 산업 환경에서 작은 객체 탐지 성능을 더욱 강화하고자 한다. 이를 통해 고해상도 특징을 강화하여 연산 비용

문제를 완화하고, 높은 탐지 정확도를 달성하는 것을 목표로 한다.

## II. 본 론

### 2. 모델 선정 및 구성

본 연구에서는 산업 환경에서 작은 객체 탐지 성능을 최적화하기 위해 최신 객체 인식 모델인 YOLOv8 [7]을 사용하였다. 다단계 특징 융합과 경량화된 백본 네트워크를 통해 높은 정확도와 빠른 추론 속도를 제공하는 장점이 있다. 이러한 특성 덕분에 YOLOv8은 작은 객체 탐지에서 높은 정확도와 실시간 처리가 요구되는 산업 환경에 적합하다.

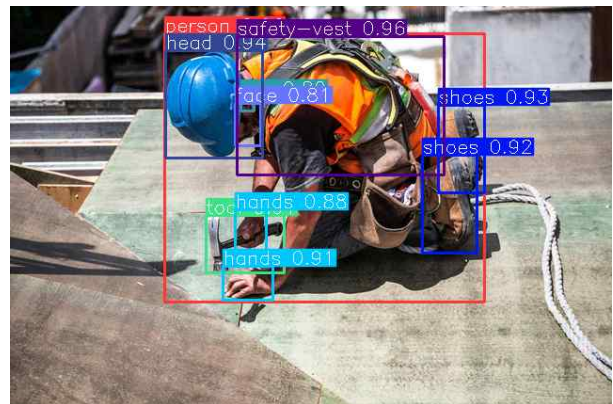


그림 1. Yolo-v8을 통한 객체 인식 예시

Fig. 1. Example of object detection with Yolo-v8

### 2.1. 데이터 셋

본 연구는 산업 환경에서의 작은 객체 인식 성능을 평가하기 위해 SH17 데이터셋 [8]을 사용하였다. SH17 데이터셋은 다양한 크기의 객체와 복잡한 배경이 포함된 실제 산업 환경을 반영하여 설계된 데이터셋으로, 특히 작은 객체 탐지 성능을 평가하는데 적합하다. 레이블이 포함된 8,099개의 이미지로 구성되어 있으며, 75,994개의 오브젝트 인스턴스 및 17가지 종류의 PPE 품목(안경, 의료용 얼굴 마스크, 얼굴 보호대, 귀마개, 장갑, 신발, 안전 조끼, 헬멧, 의료용 보호복, 안전복) 및 신체 부위(사람, 머리, 얼굴, 귀, 손, 발)를 포함한다.

SH17 데이터셋의 이미지는 전처리 과정에서 가

장 큰 변의 길이가 640이 되도록 리사이징하여 입력 데이터로 사용되었다. 또한, 초해상화 기법을 통해 한 변의 길이가 640인 이미지를 1280으로 업샘플링하여, 작은 객체의 세부 표현을 더 명확히 하고 탐지 성능을 향상시키고자 하였다. SH17 데이터셋을 통해 초해상화 기법과 YOLOv8, SAHI 기법이 결합된 모델의 성능을 평가하였으며, 이를 통해 각 기법이 실제 산업 현장에서 작은 객체 탐지 성능에 미치는 영향을 분석하였다.

## 2.2. 초해상화 모델 비교

객체 인식 성능을 강화하기 위해, 본 연구에서는 초해상화(Super-Resolution, SR) 기법을 도입하여 작은 객체의 해상도를 향상시키는 방안을 모색하였다. 초해상화 모델로는 SRGAN [9]과 Real-ESRGAN [10]을 선정하여 비교하였다. SRGAN은 적대적 생성 네트워크(GAN) 구조를 이용하여 고해상도 이미지를 생성하며, 세부 디테일이 강화된 이미지를 제공하는 장점이 있다. 이 모델은 특히 작은 객체의 특징을 더 세밀하게 표현하는 데 유리하다. Real-ESRGAN은 SRGAN을 개선하여 보다 현실적인 이미지 생성을 목표로 하며, 다양한 노이즈와 왜곡이 많은 환경에서도 안정적인 성능을 발휘하도록 설계되었다. 이는 실제 산업 현장의 복잡한 배경 속에서도 고해상도 이미지를 유지하는 데 효과적이다. 두 모델의 초해상화 결과를 그림 2에 제시하였다.

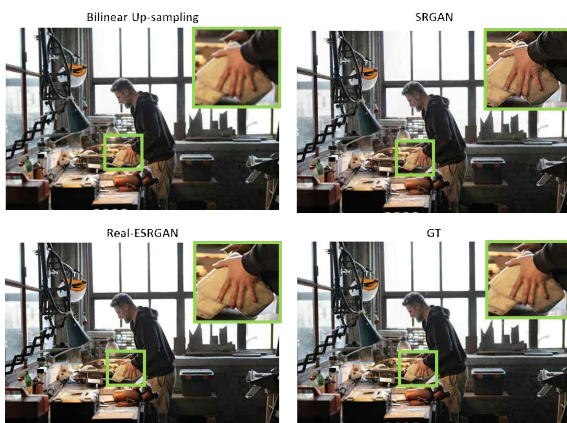


그림 2. SRGAN과 Real-ESRGAN을 이용한 초해상화 결과 비교

Fig. 2. Comparison of Super-Resolution Results Using SRGAN and Real-ESRGAN

SRGAN과 Real-ESRGAN을 통해 생성된 이미지를 비교한 결과, Real-ESRGAN은 노이즈가 많은 환경에서도 더욱 안정적인 결과를 보여주었다.

## 2.3. Slicing Aided Hyper Inference(SAHI)를 활용한 성능 평가

본 연구는 초해상화 모델을 적용한 후 YOLOv8 모델에 Slicing Aided Hyper Inference(SAHI) [11]기법을 적용하여 성능을 평가하였다. SAHI는 슬라이딩 윈도우 방식으로 분석하여 작은 객체 탐지 성능을 정밀하게 평가하는 방법이다. 이를 통해 초해상화 후 초해상화된 이미지를 4개의 타일로 나눈 뒤 타일 단위로 분할된 이미지에서 YOLOv8의 객체 인식 성능을 평가하였다.

그림 3는 초해상화 처리와 SAHI 기법을 활용한 객체 인식 결과를 시각적으로 나타낸다. 초해상화를 거친 이미지는 SAHI를 통해 보다 작은 객체를 세밀하게 인식할 수 있게 하였으며, 특히 귀와 같은 작은 크기의 객체 탐지 성능이 향상됨을 확인할 수 있다.

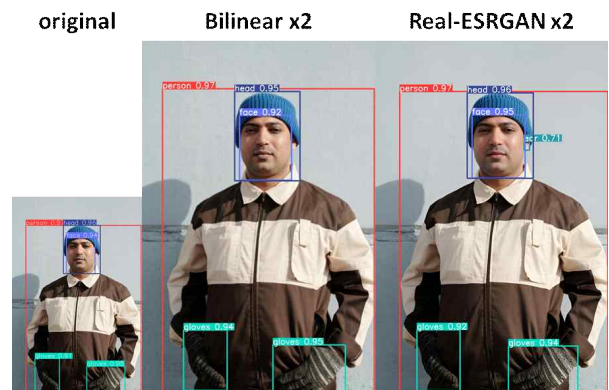


그림 3. SAHI 기법을 적용한 초해상화 기반 객체 인식 성능 비교

Fig. 3. Comparison of Object Detection Performance with Super-Resolution Using SAHI

또한, 표 1은 SAHI 기법을 활용한 YOLOv8 모델의 성능 평가 결과로, YOLOv8의 소형 모델(yolo-8-n)과 대형 모델(yolo-8-l)의 성능을 mAP@IoU=50 기준으로 비교한 것이다. 특히, 작은 크기의 객체인 ‘귀’와 ‘귀마개’에 대한 성능과 전체

클래스에 대한 성능을 함께 기록하였다. 초해상화와 SAHI를 적용한 결과, 두 모델 모두 작은 객체 탐지 성능이 개선되었으며, 특히 SAHI 기법은 타일 단위로 이미지를 세분화하여 작은 객체 탐지를 정밀하게 수행할 수 있도록 돕는 것으로 나타났다.

이러한 결과를 통해 SAHI 기법과 초해상화가 결합된 제안하는 접근법이 작은 객체 탐지 성능을 효과적으로 높이며, 산업 환경에서의 작은 객체 인식 과제에 실질적인 기여를 할 수 있음을 알 수 있다.

표 1. YOLOv8 소형 모델(yolo-8-n)과 대형 모델(yolo-8-l)의 mAP@IoU=50 성능 비교  
Table 1. Performance Comparison of YOLOv8 Small Model (yolo-8-n) and Large Model (yolo-8-l) on mAP@IoU=50

Object	Model	저해상도	Bilinear (x2)	Real-ESR GAN <sup>[10]</sup> (x2)
귀	Yolo8-n	60.93	72.09	74.59
	Yolo8-l	75.51	81.42	81.79
귀마개	Yolo8-n	41.29	44.75	44.93
	Yolo8-l	75.81	78.85	80.08
전체 클래스	Yolo8-n	56.68	59.37	59.75
	Yolo8-l	76.87	76.65	77.6

### III. 결 론

본 연구는 산업 환경에서의 작은 객체 탐지 성능을 강화하기 위해 YOLOv8 모델과 초해상화 기법을 결합한 접근법을 제안하였다. 실험 결과, 해상도와 특징 정보가 제한적인 작은 객체를 다양한 크기와 복잡한 배경 속에서도 높은 정확도로 인식할 수 있음을 확인하였다. 특히 첨단 백본 네트워크와 다단계 피쳐 융합 구조를 통한 YOLOv8의 성능을 SR 기법이 보강하여, 기존 객체 탐지 모델이 직면했던 성능 저하 문제를 효과적으로 해결하였다.

본 연구의 접근법은 산업 현장에서 실시간으로 작은 객체를 안정적으로 인식할 수 있도록 지원하여 자동화와 안전성을 높이는 데 기여할 수 있을 것으로 기대된다.

### 참 고 문 헌

[1] Lin, C., et al. "Dynamic Coarse-To-Fine Learning for Oriented Tiny Object Detection." ICCV, 2023.

[2] Huang, X., et al. "ScaleKD: Distilling Scale-Aware Knowledge in Small Object Detector." CVPR, 2023.

[3] Tang, F., et al. "Exploiting Temporal Context for Tiny Object Detection." ICCV, 2023.

[4] Jocher, G., et al. "YOLOv8: Real-Time Flying Object Detection with YOLOv8." arXiv preprint arXiv:2305.09972, 2023.

[5] Hong, H., et al. "UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios." Sensors, 2023.

[6] Khalili, B., et al. "SOD-YOLOv8: Enhancing YOLOv8 for Small Object Detection in Traffic Scenes." arXiv preprint, 2024.

[7] Jocher, G., et al. "YOLOv8: Real-Time Flying Object Detection with YOLOv8." arXiv preprint arXiv:2305.09972, 2023.

[8] Ahmad, H.M., & Rahimi, A. "SH17: A Dataset for Human Safety and Personal Protective Equipment Detection in Manufacturing Industry." arXiv preprint arXiv:2407.04590, 2024.

[9] Ledig, C., et al. "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network.

[10] Wang, X., et al. "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data." ICCV, 2021.

[11] Akyon, "Slicing Aided Hyper Inference and Fine-tuning for Small Object Detection" 2022 IEEE International Conference on Image Processing (ICIP), 2022.



# 도메인 차이 완화를 위한 자유형 스케치의 정형 표현 변환

이정민\*, 김보은\*, 김정호\*\*

## Bridging the Gap: Transforming Free-form Sketches into Formal Representations

Jeongmin Lee\*, Boeun Kim\*, and Jungho Kim\*\*

### 요약

최근 딥러닝 기술의 발전으로 사람의 스케치를 활용한 인식 및 응용 기술이 크게 발전하였다. 그러나 자유로운 형태의 스케치는 대개 합성 스케치 데이터로 학습된 모델에서 성능 저하를 유발하며, 이는 합성 스케치와 사람 손그림 간의 도메인 차이에서 기인한다. 이를 해결하기 위해, 본 연구에서는 쉽게 획득할 수 있는 합성 스케치를 활용한 보완 네트워크를 제안한다. 이 네트워크는 일부를 임의로 제거한 합성 스케치를 복원하도록 학습하여, 자유로운 스케치를 정형화된 표현으로 변환해 도메인 차이를 줄인다. 실험적으로 스케치-3D 모델 변환 작업에 적용한 결과, 기존 접근법 대비 성능이 향상됨을 확인하였다. 본 연구는 자유로운 스케치의 활용 가능성을 3D 모델링 등 고급 디자인 프로세스에서 크게 향상시킬 수 있음을 시사한다.

### Abstract

With recent advancements in deep learning, technologies leveraging human-drawn sketches for recognition and applications have significantly progressed. However, free-hand sketches often lead to performance degradation in models trained on synthetic sketch data due to the domain gap between synthetic sketches and hand-drawn ones. To address this, we propose a sketch refinement network that leverages easily obtainable synthetic sketches. This network is trained to restore partially removed synthetic sketches, allowing it to transform free-hand sketches into more structured representations and reduce the domain gap. Experimental results in a sketch-to-3D model conversion task demonstrate that our approach improves performance compared to existing methods, highlighting the potential for enhanced use of free-hand sketches in advanced design processes such as 3D modeling.

### Key words

Computer Vision, Generative Modeling, Domain Adaptation, 3D Reconstruction, Sketch-to-3D Model Generation

### 1. 서론

스케치는 직관적이며 창의적인 시각적 표현 도구

로, 디자인 및 산업 분야에서 중요한 역할을 한다. 최근 딥러닝 기술의 발전에 따라, 스케치를 활용한 인식 및 변환 응용 기술들이 CAD 설계, 가상 현실,

\*한국전자기술연구원, jeanmichel@keti.re.kr, kbe36@keti.re.kr \*\*한국전자기술연구원, jhkim77@keti.re.kr (교신저자)

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. RS-2024-00222106, 실공간 반영 지능형 콘텐츠 창작 및 실사물 필터링 기술)

메타버스와 같은 다양한 영역에서 널리 사용되고 있다. 특히, 스케치-기반 모델 생성 작업은 디자인 효율성을 크게 높일 수 있지만, 자유로운 손그림(비정형 스케치)은 대부분 정형화된 합성 스케치 데이터를 기반으로 학습된 모델들에서 성능 저하를 유발할 수 있다.

기존 연구에 따르면, 자유로운 형태의 스케치를 다룰 때 딥러닝 기반의 스케치 인식 및 변환 모델이 성능 저하를 겪는 문제가 보고되고 있다. 예를 들어, [1]은 자유롭게 그린 스케치를 3D 모델로 변환하는 과정에서 스케치의 불규칙한 선과 형태로 인해 모델이 정보 손실을 겪고, 최종 모델의 구조적 정확도가 저하됨을 확인하였다. 마찬가지로, [2]는 자유로운 스케치의 변형된 형태가 인식 및 변환 과정에서 불균형한 특성을 나타내며, 다양한 도메인에서 훈련된 모델의 일반화 성능이 감소하는 한계를 보였다.

또한, [3]은 비정형 스케치를 이용한 모델 생성에서 고유의 비대칭성을 인식하는 데 한계를 보였으며, 이는 인식 성능 저하의 주요 원인으로 작용하였다. [4]는 스케치 기반 3D 모델 생성에서 다양한 손그림 패턴이 모델 성능에 미치는 영향을 분석하며, 정형화되지 않은 자유로운 표현이 모델의 학습 가능성을 제한한다는 문제를 제기하였다. 이러한 연구들은 기존의 합성 스케치 기반 학습 모델이 비정형 스케치의 불규칙성을 충분히 반영하지 못함으로써 성능 저하가 발생하는 것을 보여준다.

본 연구는 이러한 문제를 해결하기 위해 정형화된 합성 스케치를 활용한 스케치 보완 네트워크를 제안한다. 본 네트워크는 합성 스케치의 일부를 임의로 제거하고 이를 복원하는 과정을 학습하여, 자유로운 스케치를 정형화된 표현으로 변환함으로써 도메인 차이를 완화하는 방법을 제공한다. 이를 통해 비정형 스케치의 인식 및 변환 성능을 강화하고, 특히 스케치-3D 모델 변환 작업에서의 활용 가능성을 높이려 한다.

## II. 문제 정의 및 제안 방법

스케치를 통해 3D 모델을 생성하는 연구인 sketch-to-3D model (sketch2model)[5] 은 임의의 뷰에

서 생성된 스케치 데이터를 입력으로 받아 상응하는 3D 모델을 생성하는 작업을 수행한다. 기존 연구는 3D 모델의 특정 뷰에서 이미지 에지(edge)를 스케치로 가정하여 학습을 진행했으나, 자유로운 형태로 그려진 free-hand sketch가 입력되었을 때 성능이 크게 저하됨이 관찰되었다. 이는 사람이 그린 스케치의 불규칙성과 도메인 간 차이가 주요 원인으로 작용한다.

이를 해결하기 위해, 본 연구에서는 free-hand sketch의 특성을 모방하고 복원하는 네트워크를 제안한다. 제안된 네트워크는 U-Net 구조를 기반으로 하며 [6], synthetic sketch를 입력받아 일정 비율로 랜덤하게 일부를 제거한 후 이를 복원하는 과정을 학습한다. 구체적으로, 네트워크는 synthetic sketch  $x$ 에서 일부가 제거된 변형 스케치  $x_1$ 을 생성하고, 이를 복원하는 네트워크  $f$ 를 구성해 원본 스케치와 유사한 형태로 복원하는 과정을 학습한다. 이때 네트워크의 가중치를  $\theta$ 로 나타내고, 네트워크 출력  $f(x_1; \theta)$ 과 원본 스케치  $x$  사이의 L1 손실을 최소화하여 학습을 진행하였다. 손실 함수는 다음과 같이 정의된다:

$$L = \|f(x_1; \theta) - x\|_1 \quad (1)$$

해당 손실 함수를 통해 네트워크는 주어진 스케치에서 제거된 부분을 정확히 복원하도록 학습된다. 학습이 완료된 네트워크는 자유로운 스케치가 입력되었을 때에도 이를 정형화된 표현으로 변환하여, 기존 모델이 처리할 수 있는 형태로 전환할 수 있게 한다.

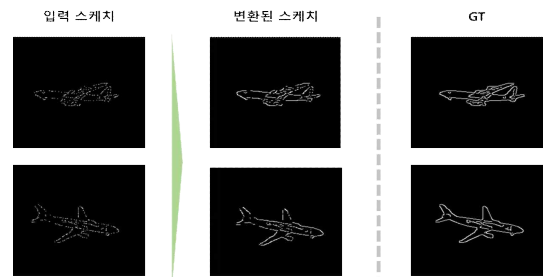


그림 1. 입력 스케치를 임의의 비정형 스케치로 변환한 예시

Fig. 1. Example of Input Sketch Transformed into a Random Free-hand Sketch

### III. 실험

성을 검증하였다.

표 1. 제안된 방법을 활용한 Sketch-to-3D 모델 성능 향상 실험 결과

Table 1. Experimental Results of Sketch-to-3D Model Performance Enhancement Using Proposed Method

객체 종류	Sketch2Model	Sketch2Model + Ours
비행기	0.4795	0.4935
벤치	0.3578	0.367
캐비닛	0.5477	0.5659
자동차	0.6498	0.6552
의자	0.3518	0.3916
디스플레이	0.4353	0.4799
라이플	0.5104	0.5321
소파	0.5287	0.5349
테이블	0.3614	0.3614
전화기	0.5512	0.5635
평균	0.4773	0.4945

본 연구에서는 ShapeNet-sketch dataset을 사용하여 학습 및 평가를 수행하였다. ShapeNet-sketch dataset은 ShapeNet [7]에서 공개된 3D 모델 데이터셋으로, 각각의 3D 모델을 ShapeNet-sketch dataset을 기반으로 다양한 스케치 데이터를 구축하였다. 학습 데이터는 ShapeNet에서 제공하는 3D 모델의 에지를 스케치로 변환한 ShapeNet-synthetic 데이터셋을 사용하였으며, 각 클래스별로 synthetic sketch를 네트워크에 학습 데이터로 활용하였다.

성능 평가는 Voxel Intersection over Union (voxel-IoU)를 통해 이루어졌다. voxel-IoU는 생성된 3D 모델과 실제 3D 모델 간의 유사성을 측정하는 지표로, 각 voxel 격자에서 예측된 3D 모델과 실제 3D 구조 간의 겹치는 비율을 계산한다. 실험 결과, 제안한 네트워크를 sketch-to-3D model 작업의 전처리 단계로 적용한 경우 기존 방법 대비 voxel-IoU 성능이 유의미하게 향상됨을 확인하였다. 제안된 방법은 free-hand sketch를 정형화된 synthetic sketch와 유사한 형태로 복원하여 모델의 학습 데이터와 유사한 입력을 제공함으로써, 3D 모델 생성의 정확도를 개선할 수 있었다. 실험 결과는 표 1에 제시되어 있으며, 이를 통해 본 연구에서 제안한 방법의 유효

### IV. 결 론

본 연구에서는 비정형 스케치가 딥러닝 기반 스케치 인식 및 응용 모델에서 성능 저하를 초래하는 문제를 해결하기 위해, 정형화된 합성 스케치를 활용한 스케치 보완 네트워크를 제안하였다. 제안된 방법은 합성 스케치 데이터의 일부를 임의로 제거하고 이를 복원하는 학습 과정을 통해, 자유로운 손그림 스케치를 정형화된 표현으로 변환하여 도메인 차이를 줄이는 데 기여한다. 이를 통해 자유로운 스케치의 불규칙성과 형태적 다양성으로 인해 발생하는 성능 저하를 완화할 수 있음을 확인하였다.

실험 결과, 제안된 네트워크는 스케치-3D 모델 변환 작업에서 기존 접근법에 비해 뛰어난 성능을 보였으며, 이는 비정형 스케치를 활용한 다양한 응용 작업의 신뢰성과 활용 가능성을 크게 향상시킬 수 있음을 시사한다. 향후 연구에서는 다양한 비정형 스케치 패턴과 도메인을 추가적으로 고려하여, 더욱 정교하고 확장 가능한 스케치 인식 및 변환 모델을 개발하는 방향으로 발전할 것이다. 본 연구의 성과는 디자인 및 산업 분야에서 자유로운 스케치를 활용한 고급 응용의 가능성을 한층 더 높일 수 있는 중요한 기초 자료가 될 것으로 기대된다.

### 참 고 문 헌

- [1] Zhou, X., et al. "Enhancing Free-Hand Sketch Recognition for 3D Model Generation through Structured Data Constraints." CVPR, 2022.
- [2] Kim, S., et al. "Domain Variability in Sketch-to-Model Conversions: Challenges with Free-Form Inputs." ECCV, 2022.
- [3] Li, Y., et al. "Understanding Asymmetric Line Patterns in Free-Hand Sketches for Model Translation." ICCV, 2023.
- [4] Zhang, H., et al. "Evaluating Pattern Complexity in Sketch-Based Model Generation: Limits of Deep Learning Models." CVPR, 2023.

- [5] Wang, X., et al. "Sketch2Model: Bridging the Gap Between Free-Hand Sketches and 3D Model Generation for Virtual Environments." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022..
- [6] Wang, X., et al. "Sketch2Model: Bridging the Gap Between Free-Hand Sketches and 3D Model Generation for Virtual Environments." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022..
- [7] Chang, A. X., et al. "ShapeNet: An Information-Rich 3D Model Repository." arXiv preprint arXiv:1512.03012 (2015).

# 스마트공장에 적용 가능한 단말간 직접통신 시스템의 성능향상을 위한 주파수 동기화 기법

정용안\*, 변상봉, 한동철, 조수현, 이성훈\*\*

## Frequency Synchronization Method for Performance Enhancement of D2D Communication in the Smart Factory

Yong-An Jung\*, Sang-Bong Byun, Dong-Cheul Han, Soo-Hyun Cho, and Sung-Hun Lee\*\*

최근 들어, 스마트 공장 구축을 위한 다양한 통신 시스템들이 제안되었으며, 대표적인 시스템으로 IEEE 802.11ax 기반 Wi-Fi, 3GPP 4G LTE, 5G NR이 있다. 3GPP 5G NR 표준화 단체에서는 private 5G 네트워크 구조를 제시하였으며, 국내에서는 4.72~4.82GHz 주파수 대역을 사용하는 이음 5G 네트워크 기술이 상용화되었다. 이러한 이음 5G 네트워크에서는 OFDM 변조 기법을 채택하여 초고속의 데이터 통신 서비스를 제공하고 있다. AMR과 같은 이동로봇 간 직접통신은 송신기와 수신기가 이음 이동성을 가지기 때문에 동기화 오차에 매우 민감하다. 특히 주파수 오차가 발생할 경우, 부반송파 간 직교성이 파괴되며, ICI가 발생하여 성능 저하가 발생한다. 본 논문에서는 이러한 문제를 해결하기 위해 송·수신기의 이음 이동성에 강인한 주파수 오차 추정 기법이 제안되었다. 제안된 기법은 3GPP 5G NR sidelink 시스템을 고려하였으며, PSSS 및 SSSS 시퀀스의 특징을 이용한 2단계 상관 기법을 적용하였다. 제안된 기법과 기존의 기법 간 성능 비교를 위해 컴퓨터 모의실험이 사용되었으며, 제안된 기법이 기존의 기법보다 좋은 mean square error (MSE) 성능을 갖는다는 것을 확인하였다.

### Key words

Smart Factory, 5G NR, OFDM, Synchronization

### 참고문헌

- [1] Y. H. You and Y. A. Jung, "complexity Efficient Sidelink synchronization Signal Detection Scheme for Cellular Vehicle-to-Everything communication Systems," MDPI mathematics, vol. 11, no. 18, pp. 1-15, Sep., 2023.
- [2] You, Y.H.; Park, J.H.; Ahn, I.Y. "Complexity effective sequential detection of secondary synchronization signal for 5G new radio communication systems," IEEE Syst. J. 2021, 15, 3382-3390.

---

\*구미전자정보기술원, yajung@geri.re.kr

\*\*구미전자정보기술원, leesh@geri.re.kr

※ This work was supported by the Korea Planning & Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korean government (MOTIE) (RS-2024-00445778, Development of AI-based high reliability robot-equipment and risk management integrated operation technology for digital transformation of missile manufacturing process).

# 반도체 센서 데이터 전처리의 중요성과 분석 정확도에 미치는 영향

윤여은\*, 심별희\*, 윤동식\*, 강정훈\*\*

## Importance of Preprocessing in Semiconductor Sensor Data and Its Impact on Analysis Accuracy

Yeoeun Yoon\*, and Byeolhee Sim\*\*, Dongsik Yoon\*\*, and Jeonghoon Kang\*\*

### 요약

본 논문은 반도체 제조 공정에서 센서 데이터 전처리의 중요성을 강조하고, 전처리 과정이 데이터 분석의 신뢰성과 정확도에 미치는 영향을 고찰한다. 웨이퍼 세척, 식각, 산화, 도핑, 박막 증착 등 주요 전처리 공정과 그 기법을 설명하고, 각 공정에서 발생하는 기술적 도전 과제를 다룬다. 또한, 딥러닝 모델을 활용하여 전처리된 데이터의 분석 정확도를 향상시킬 수 있는 가능성에 대해 논의한다. 이를 통해, 전처리의 필요성과 미래 연구의 방향성을 제시한다.

### Abstract

This paper emphasizes the importance of data preprocessing in semiconductor manufacturing processes and explores its impact on the reliability and accuracy of data analysis. Key preprocessing techniques such as wafer cleaning, etching, oxidation, doping, and thin-film deposition are detailed, highlighting their significance in maintaining high data quality. Challenges associated with each stage are discussed, and the role of deep learning models in enhancing the analysis of preprocessed data is examined. This study underscores the necessity of preprocessing for minimizing defect rates and optimizing yield in semiconductor manufacturing, providing insights for future research directions.

### Key words

Semiconductor manufacturing, Data preprocessing, Wafer Cleaning, Etching, Oxidation, Doping, Thin Film Deposition, Sensor data, Normalization

---

\*한국전자기술연구원, yeoeun278@keti.re.kr, nmjjbb@keti.re.kr, dsyoon@keti.re.kr

\*\*한국전자기술연구원, budge@keti.re.kr(교신저자)

※ "이 논문은 2024년도 산업통상자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임(20025255, 반도체 장비 지능화를 위한 공정 데이터 플랫폼 기술 개발) This paper was supported by Korea Planning & Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korea government(MOTIE)(20025255, Industry data analytics software Platform development for semiconductor equipment and process)"

## I. 서 론

반도체는 현대 전자기기에서 필수적인 요소로 자리 잡았으며, 반도체 제조 공정 중 전처리 과정은 소자의 성능과 생산성을 높이는 데 중요한 역할을 한다. 웨이퍼 표면의 불순물을 제거하고 후속 공정의 품질을 보장하기 위해 전처리 과정이 필요하다. 본 연구는 반도체 전처리 과정에서의 센서 데이터의 중요성과 이를 통해 얻는 분석의 신뢰성을 강조하며, 딥러닝 기반의 분석 모델을 활용한 전처리 데이터의 활용 가능성을 논의한다.

## II. 본 론

### II-1. 이론적 배경

반도체 전처리 과정은 웨이퍼 표면의 청결 상태를 유지하고, 물리적 및 전기적 특성을 조정함으로써 고품질 소자 제작을 위한 기초를 제공한다. 주요 전처리 공정에는 웨이퍼 세척, 식각, 산화, 도핑, 박막 증착이 포함되며, 각 공정은 다양한 기술적 도전 과제를 포함하고 있다.

딥러닝은 이러한 전처리 과정을 지원하고 데이터 분석의 정확성을 높이기 위해 활용될 수 있다. 전처리된 센서 데이터를 기반으로 딥러닝 모델은 이상 탐지, 결함 예측, 공정 최적화 등의 목적으로 사용되며, 데이터의 패턴을 효과적으로 학습하여 높은 분석 정확도를 제공할 수 있다.

### II-2. 연구 방법 및 주요 전처리 기법

반도체 전처리 과정에서 사용하는 주요 기술들은 반도체 소자의 고도화를 위해 정밀하고 균일하게 제어되어야 한다. 이를 적절히 처리하지 않으면 분석의 신뢰성과 정확도가 떨어진다. 각 전처리 단계에서의 구체적인 전처리 방법은 다음과 같다.

- 웨이퍼 세척: 웨이퍼 표면의 불순물 제거를 위해 센서 데이터에서 불필요한 신호나 노이즈를 제거하는 것이 중요하다. 이를 위해 노이즈 필터링(예: 이동 평균 필터, 저역 필터)을 사용하여 데이터를 정제한다. 결측치가 발생할 경우, 선형 보간법이

나 평균 대체법을 적용하여 데이터의 연속성을 유지한다.

- 식각: 식각 공정의 데이터는 주기적인 노이즈와 순간적인 이상치가 포함될 수 있다. 이 단계에서는 주파수 분석을 통해 주기적인 패턴을 분석하고, 파형 변환(예: 푸리에 변환)을 사용하여 노이즈를 식별 및 제거한다. 이상치는 이상치 탐지 알고리즘(예: IQR, Z-스코어 기준)을 통해 제거하거나 수정한다.

- 산화: 산화 공정에서는 데이터의 정규화가 중요하다. 산화막의 두께나 전기적 특성 데이터를 다룰 때 Min-Max 정규화 또는 Z-스코어 정규화를 적용하여 데이터를 일관된 스케일로 변환해 분석의 정확성을 높인다. 정규화는 다양한 센서에서 수집된 데이터 간 비교 가능성을 높인다.

- 도핑: 이온 주입 과정에서는 데이터의 변동성이 클 수 있으므로, 고급 스무딩 기법(예: 가우시안 스무딩)을 적용해 변동성을 줄인다. 주입된 불순물 농도와 관련된 데이터는 분할 및 조합 기법으로 처리하여 지역별 특성을 반영한다. 결측치 처리는 복잡한 패턴일 경우 KNN 보간을 통해 보정한다.

- 박막 증착: 증착 공정의 데이터는 균일성 및 박막의 결함 여부에 따라 중요하다. 피크 검출 알고리즘을 통해 박막 두께의 불균일성을 탐지하고, 이상치 수정기법으로 결함 데이터를 제거한다. 데이터의 상관성을 높이기 위해 주성분 분석(PCA)을 사용해 주요 특성을 추출하고 데이터의 차원을 축소하여 분석 효율성을 높인다.

이와 같은 전처리 기법들은 각 공정의 특성과 요구사항에 맞춰 적용된다. 이를 통해 반도체 제조 과정에서 데이터 분석의 신뢰성을 높이고, 공정의 안정성을 유지하며 불량률을 줄일 수 있다.

### II-3. 딥러닝의 적용과 분석 향상

딥러닝 기술은 전처리된 반도체 센서 데이터를 분석하는 데 중요한 도구로 사용된다. 전처리된 데이터는 이상치 제거 및 결측치 보정 등을 통해 개선된 품질을 제공하며, 이는 딥러닝 모델의 학습 데이터로 최적화된다. 특히, 합성곱 신경망(CNN)과

순환 신경망(RNN) 등은 복잡한 시간적·공간적 패턴을 탐지하고 분석하는 데 효과적이다. 이러한 모델을 통해 공정의 실시간 모니터링 및 결함 예측의 정확도가 높아질 수 있다.

딥러닝 기반의 분석은 전처리된 데이터의 패턴을 학습함으로써 예측 능력을 향상시키고, 공정 개선을 위한 효율적인 방안을 제공할 수 있다. 이를 통해 반도체 제조 공정의 생산성과 신뢰성을 높이는 데 기여할 수 있다.

### III. 결 론

반도체 제조 공정에서의 데이터 전처리는 고성능 소자 제작과 생산성 향상을 위한 필수적인 과정이다. 전처리 과정의 중요성과 각 공정의 기술적 도전 과제를 논의하였으며, 이를 통해 딥러닝 모델의 분석 향상 가능성을 함께 제시하였다. 본 논문은 이러한 전처리 및 딥러닝 기술의 발전이 반도체 제조의 핵심 요소임을 강조하며, 차세대 기술 발전의 기초 자료가 될 것으로 기대된다.



# ISO/IEC 25059 기반 AI 융합 시스템 품질 평가 방법 연구

박지훈\*, 윤성복\*, 서경덕\*

## A Study on the Quality Evaluation Method for AI-Converged Systems Based on ISO/IEC 25059

Jihoon Park\*, Seongbok Yoon\*, and Gyeongdeok Seo\*

### 요 약

AI 기술이 다양한 분야에 융합됨에 따라, AI 융합 시스템의 품질 관리가 중요한 과제로 부상하고 있다. 국제표준인 ISO/IEC 25059는 AI 시스템을 위한 품질 모델을 제시하지만, 품질모델을 기반으로 AI 융합 시스템의 품질을 평가하기 위한 구체적인 시험평가 방법에 대한 표준은 아직 없는 것이 현실이다. 본 논문에서는 ISO/IEC 25059 기반으로 AI 융합 시스템의 품질을 시험 평가할 수 있도록 AI 품질 부특성별로 시험 및 평가 방법에 대한 기준을 제시하고자 한다.

### Abstract

As AI technology becomes integrated across various fields, quality management of AI-integrated systems has emerged as a critical issue. The international standard ISO/IEC 25059 provides a quality model for AI systems; however, a standardized testing and evaluation methodology for assessing the quality of AI-integrated systems based on this model has yet to be developed. This paper aims to establish criteria for testing and evaluation methods for AI quality sub-characteristics to facilitate quality assessment of AI-integrated systems based on ISO/IEC 25059.

### Key words

artificial intelligence, quality characteristics, performance measure

## 1. 서 론

AI 기술의 발전은 다양한 산업 및 응용 분야에서 AI 융합 시스템의 개발과 적용을 급격히 확산시키고 있다. 이러한 시스템의 품질 관리는 필수적이며, 이를 위해 체계적인 품질 평가 모델과 방법론이 요구된다. ISO/IEC 25059[1]는 AI 시스템 품질 평가를 위한

국제적 기준을 제공하지만, 실제로 이를 적용하기 위한 구체적인 평가 방법론은 여전히 부족한 상황이다.

본 논문에서는 ISO/IEC 25059에서 제시한 AI 융합 시스템의 특성을 반영한 품질 모델을 기반으로 시험 평가할 수 있도록 기준을 제안한다. AI 품질의 주요 부특성별로 품질 평가를 수행할 수 있는 가이드라인을 제시한다.

\* 한국정보통신기술협회 책임연구원, 책임연구원, 선임연구원, ih91082@tta.or.kr, voonsb@tta.or.kr, sss9305@tta.or.kr

기능적합성	성능효율성	호환성	사용성	신뢰성	보안성	유지보수성	이식성
기능완전성	시간반응성	공존성	적절 인식성	성숙성	기밀성	모듈성	적응성
기능정확성	자원효율성	상호운용성	학습성	가용성	무결성	재사용성	설치성
기능적절성	용량성		운영성	결함허용성	부인방지성	분석성	대체성
기능적응성			사용자 오류 방지성	복구성	책임성	변경성	
			사용자인터페이스 심미성	강건성	인증성	시험성	
			접근성		개입가능성		
			사용자 제어 가능성				
			투명성				

수정됨

추가됨

그림 1. ISO/IEC 25059에서 제시하는 품질 모델 중 ISO/IEC 25023 대비 변경된 부분  
 Fig. 1. Changes in the quality model presented in ISO/IEC 25059 compared to ISO/IEC 25023

## II. ISO/IEC 25059 기반 품질 모델 평가 방법

ISO/IEC 25059에서 제시한 소프트웨어 품질 모델은 ISO/IEC 25023[2]과 비교 시에 부특성 중 기능 정확성에 관한 내용이 수정되었고 기능 적응성, 사용자 제어 가능성, 투명성, 강건성, 개입 가능성이 추가되었다.

기능 정확성은 기존 소프트웨어 품질 모델에서는 시스템이 요구된 기능을 정확하게 수행하는 정도를 나타낸다. 기존 SW에서는 정해진 입력에 대해 정확한 출력을 생성하는지를 기준으로 평가했다면 AI 시스템의 기능 정확성은 단순한 기능 수행 여부를 넘어, 모델의 일반화 능력과 예측의 신뢰성, 사용자 요구 충족 여부를 평가해야 한다. 이를 위한 대표적인 측정 지표[3]로는 정확도, 정밀도, 재현율, F1 Score, IoU, Confusion Matrix 등이 있다.

기능 적응성은 시스템이 변화하는 환경과 요구에 맞춰 적절히 기능을 조정하고 발전할 수 있는 정도이다. AI 시스템은 데이터를 기반으로 학습하며, 새로운 정보가 지속적으로 추가되는 특성을 가진다. 예를 들어, ChatGPT와 같은 모델은 사용자와의 상호 작용을 통해 실시간으로 학습하고 이를 바탕으로 더욱 정확하고 유용한 응답을 제공할 수 있다. 따라서 새로운 데이터 세트에 대해서 데이터 변화에 적응하기 위해 계속해서 학습하는 기능을 제공

하고 학습을 통해 정확한 결과를 제공하는지 확인해야 한다.

사용자 제어 가능성은 제품의 기능에 사용자나 다른 외부 에이전트가 적절한 때에 개입할 수 있는 정도이다. 이를 위해 사용자가 AI 시스템을 제어할 수 있는 인터페이스가 제공되어야 한다[4]. 예를 들어, 사용자가 시스템의 동작을 즉시 중단하거나 변경할 수 있는 버튼이나 명령어를 제공해야 한다. 또한, AI 시스템의 내부 동작을 세부적으로 조정할 수 있는 인터페이스도 제공되어야 한다. 예를 들어, AI 모델의 학습률, 정규화 파라미터, 의사결정나무 깊이 등의 AI 모델의 매개변수를 수정할 수 있도록 제공해야 한다.

투명성은 AI 시스템에 관한 적절한 정보가 이해관계자들에게 전달되는 정도이다. 사용자에게 AI 시스템의 작동 방식, 편향성 및 오류를 파악할 수 있게 하여 의미 있는 설명을 제공해야 한다. 이를 위해 AI 모델의 설명 가능성, 추적 가능성, 데이터 사용 투명성 등을 문서로 만들어 투명하게 공개해야 한다. 이 문서 작성 시에는 AI 생명주기와 관련된 이해관계자들을 고려하여 각자 필요한 정보를 포함하여야 한다[5]. 투명성을 보장하기 위한 평가 항목은 TTA에서 발간한 ‘2024 신뢰할 수 있는 인공지능 개발 안내서’를 통해 확인할 수 있다.

강건성은 AI 시스템이 외부의 간섭이나 극한적인

운영 환경 등에서도 사용자가 의도한 수준의 성능 및 기능을 유지할 수 있는 정도이다. 이러한 강건성을 확보하기 위해서는 데이터에 노이즈를 추가하여 의도적으로 잘못된 결과를 유도하는 상황에서도 사용자의 기대 결과와 실제 결과 간의 일치를 확인해야 한다. 예를 들어, 의도적으로 손상된 ‘STOP’ 표지판을 자율주행 차량이 ‘GO’ 표지판으로 오인하게 되면, 이는 적대적 공격에 대응하지 못하여 차량이 충돌하거나 예기치 않은 행동을 할 수 있어 심각한 사고를 초래할 수 있다. 따라서 이러한 상황을 대비하여 AI 시스템의 성능을 유지하도록 해야 한다.

개입 가능성은 AI 시스템이 위험한 상태에서 안전한 상태로 전환할 수 있도록 상태 관찰과 개입을 가능하게 하는 정도이다. 이는 AI 시스템이 자동으로 수행하는 작업이 항상 예측 가능하거나 완벽하지 않기 때문에 매우 중요하다. 특히, 자율 주행과 같은 안전에 직결된 시스템에서 오작동이 발생하거나, 외부 환경적 요인으로 AI 시스템을 통한 제어가 불가능하다고 판단될 때, 즉각적인 사용자의 개입이 이루어지지 않으면 큰 재산적 손실이나 인명 피해로 이어질 수 있다. 이를 위해 사용자가 AI 시스템 상태를 명확하고 직관적으로 파악할 수 있는지, AI 시스템의 상태 정보 및 오작동 등에 의한 경고 알림 등이 적시에 적절히 제공되는지 평가해야 한다. 또한, 사용자가 직접 개입하여 AI 시스템을 안전한 상태로 되돌리는 수 있는지도 평가해야 한다.

### III. 결 론

본 논문에서는 ISO/IEC 25059에서 제시한 부특성 중 AI 시스템에 특화된 6가지를 기반으로 AI 융합 시스템의 품질을 평가하기 위한 방법론을 제시하였다. 본문에서 다룬 AI에 특화된 6가지의 부특성 외에 나머지 부특성(기능완전성, 시간반응성, 무결성 등)들도 AI 시스템에 맞도록 개선된 평가 방법을 연구할 예정이다.

### 참 고 문 헌

- [1] ISO/IEC, "ISO/IEC 25059:2023 - Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality model for AI-based systems," pp. 1-45, March 2023.
- [2] ISO/IEC, "ISO/IEC 25023:2016 - Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of system and software product quality," pp. 1-35, October 2016.
- [3] ISO/IEC 4213, "Information technology — AI system reliability assessment — Guidelines", pp. 1-46, March 2023
- [4] ISO/IEC, "ISO/IEC 8200:2020 - Information technology — Interoperability of digital evidence systems," pp. 1-50, July 2020.
- [5] Ministry of Science and ICT, TTA, "Guidebook for Development of Trustworthy AI 2024 - General", pp. 75-76, February 2024.

# AI-시뮬레이션 워크플로우 솔루션 개발

허태상\*, 김남규\*\*, 황규현\*\*\*, 서동우\*\*\*\*

## Development of an AI-Simulation Workflow Solution

Taasang Huh\*, Namgyu Kim\*\*, Gyuhyun Hwang\*\*\*, and Dongwoo Seo\*\*\*\*

디지털 트윈 기술은 실제 세계의 물리적 요소를 컴퓨터 기반의 가상 환경에서 실시간 모니터링하고 분석함으로써, 산업 및 공공분야의 운영 효율성을 높이고, 빠른 예측과 문제 대응을 가능하게 한다[1]. 그러나 대부분의 AI와 시뮬레이션 기술은 독립적으로 활용되고 있어, 두 기술 간의 연계가 다소 어려운 상황이다. 이런 문제를 해결하고자 본 연구는 AI와 시뮬레이션을 원활하게 연계 및 활용할 수 있는 워크플로우 솔루션을 설계했다. 본 솔루션의 특징은 AI 모델과 물리 시뮬레이션 모델을 하이브리드 형태로 워크플로우에 적용함으로써, 시뮬레이션 모델을 통해 물리적 성능을 예측하고, AI 모델로 학습하여 보다 정교한 분석을 할 수 있다[2]. 사용자들은 솔루션을 통해 최소한의 코딩과 시각화 도구를 활용하여 데이터를 관리 및 모니터링하고, AI-시뮬레이션 모델을 적용한 워크플로우를 설계할 수 있다. AI 프레임워크인 PyTorch, YOLO, scikit-learn 등의 오픈소스를 활용 하여 개발하여, AI 모델과 시뮬레이션 모델을 하나의 워크플로우 인터페이스에서 이루어지도록 하였으며, 솔루션의 확장성과 유연성을 높이기 위해 Kubernetes와 Docker를 활용하였다[3]. 이를 통해 사용자들은 실시간 데이터 분석과 신속한 모델 성능 평가를 할 수 있으며, 워크플로우에서 시각적으로 결과를 피드백을 받을 수 있다. 또한, 제안한 솔루션은 사용자들이 쉽게 AI와 시뮬레이션 기술을 통합하고, 디지털 트윈 기술의 실질적인 적용을 촉진하는 데 기여할 것으로 기대한다.

### Key words

Digital Twin, AI, Simulation, workflow solution

### 참고문헌

- [1] V. Barone, C. Puzzarini, and G. Mancini, "Integration of theory, simulation, artificial intelligence and virtual reality: a four-pillar approach for reconciling accuracy and interpretability in computational spectroscopy", *Physical Chemistry Chemical Physics*, vol. 23, no. 32, pp. 17079-17096, 2021
- [2] MakinaRocks, The first step towards an AI-powered industry, available at: <https://www.makinarocks.ai>. 2024
- [3] J. Lee, S. Park, G. Nam, J. Jang and S. Lee, "Pig Image Learning for Improving Weight Measurement Accuracy", *Journal of The Korea Society of Computer and Information*, vol. 29, no. 7, pp. 33-40, 2024

---

\* 한국과학기술정보연구원, tshuh@kisti.re.kr,

\*\* 한국과학기술정보연구원, ssgyu@kisti.re.kr,

\*\*\* 한국과학기술정보연구원, ghhwang@kisti.re.kr,

\*\*\*\* 한국과학기술정보연구원, seodongwoo@kisti.re.kr

※ 본 연구는 2024년 한국과학기술정보연구원(Grant No. K24L2M1C5)의 지원을 받아 수행되었음.

# VPP 운영 플랫폼 내 출력제어 우선순위 관리에 관한 연구

최민지\*, 김동우\*\*, 고상원\*\*\*

## A Study on the Priority Management of Curtailment in VPP Operating Platform

Minji Choi\*, Dongwoo Kim\*\* and Sangwon Koh\*\*\*

### 요약

본 연구는 가상발전소(VPP) 운영 플랫폼에서 효율적인 출력제어를 시행하기 위한 우선순위 선정 방안을 제시한다. 다양한 분산에너지자원(DER)을 통합 관리하는 VPP 환경에서, 시스템 안정성과 경제성을 동시에 확보할 수 있는 최적의 출력제어 전략을 우선순위 관리방안을 통해 확보할 수 있었다. 신재생에너지 발전원의 특성을 고려한 동적 우선순위 할당 방식이 VPP 운영 효율을 향상시킬 수 있을 것이며, 나아가 데이터 기반의 정밀한 전력 운영을 통해 비상 상황에서 민첩하게 대응하여 미래 에너지 생태계를 안정화할 수 있을 것이다.

### Abstract

This study proposes a priority selection plan for efficient curtailment on a Virtual Power Plant(VPP) operating platform. It will be an optimal reduction strategy to ensure system stability and economic efficiency simultaneously in a VPP environment where various distributed energy resources(DERs) are integrated. A dynamic priority allocation method that considers the characteristics of renewable energy generation sources will improve the VPP operation efficiency, and furthermore, through precise data-based power operation, it will be possible to stabilize the future energy ecosystem by responding agilely in emergency situations.

### Key words

Virtual Power Plant, Curtailment, PV, Priority, DQN

## 1. 서론

### 1.1 VPP 운영 플랫폼 배경

신재생에너지 보급 확대에 따라 분산형 전원이 증가하였으며 이에 따른 전력수요 변동을 대비하여 전력 계통 안정성을 확보하기 위한 노력이 요구되고 있다. 가상발전소(Virtual Power Plant)는 다양한

분산에너지자원(DER)을 통합하여 ICT 기술을 이용해 하나의 발전소처럼 운영하는 시스템[1]으로, 에너지 수요와 공급을 최적화하여 전력망의 효율성과 신뢰성을 높이는 것을 목적으로 한다.

### 1.2 출력제어 필요성

태양광과 풍력과 같은 재생에너지원의 변동성을

\*한전KDN, mj\_choi20@kdn.com, \*\*한전KDN, kdw.stone.78@kdn.com, \*\*\*한전KDN, swkoh\_1995@kdn.com

관리하기 위해 출력제어 관리가 중요하게 여겨지고 있다. 낮 시간대 태양광 발전량이 급증하면서 발생하는 덩크브 현상은 전력 수요와 공급의 불균형을 나타내는 재생에너지 확대에 의해 발생하는 대표적인 사례이며, 이러한 경우에도 출력제어를 통해 과잉 생산된 전력을 조절할 수 있다.

## II. 본 론

### 2.1 시스템 개요

본 논문은 신재생에너지원의 제어 용이성을 고려하여 우선순위 선정 방법을 구현하고, 형평성을 위해 제어 시행 횟수에 따른 전략을 세워 안정적 출력제어 체계를 확보한 VPP 운영 플랫폼을 개발하고자 한다. 계통안정화를 위해 필요한 출력제어 지시는, 전력 시장과 전력 계통 운영을 전담하는 KPX의 지령을 기준으로 한다.

### 2.2 시스템 구성도

VPP 운영 플랫폼 내 출력제어 시스템 동작에 대한 구성은 그림 1과 같다. 출력제어 지령을 전송하는 신재생자료취득장치, 제어명령 및 발전단지 데이터를 수집하고 활용하는 VPP 운영 플랫폼, 제어 요청 데이터 전달 후 현장에서 제어신호를 주는 현장 EMS로 구성되어 있다.

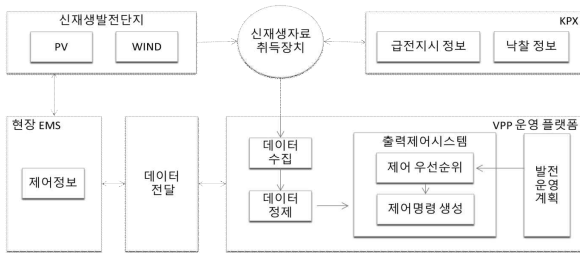


그림 1. VPP 내 출력제어 기능구성도  
Fig. 1. Functional Diagram of Curtailment Control in VPP

### 2.3 제어 시스템 설계

출력제어 시스템 내 KPX 제어 지령값에 따라 우선순위 발전기별 제어요청 데이터를 생성하기 위한 DB테이블 구조는 그림2와 같다. 우선순위가 높은 발전기부터 적용되는 제어 시나리오에 따라 제어 요청 명령을 생성하기 위한 데이터 구조와 관계를

나타낸다. 시나리오 종류[1]는 MPPT, Freq.Control, Volt.Control 등이 있지만 본 논문에서는 KPX 지령치에 응동하여 운전하는 시나리오를 활용하여 출력 제어 명령 요청문을 생성하기 위해 발전기별 우선순위 선정에 대하여 다룬다. 시스템을 VPP 운영 플랫폼 내에서 실제로 동작시키기 위해서는 발전량 경제성(SMP/REC)을 예측하는 DB테이블 구축 및 연계가 필요하다.

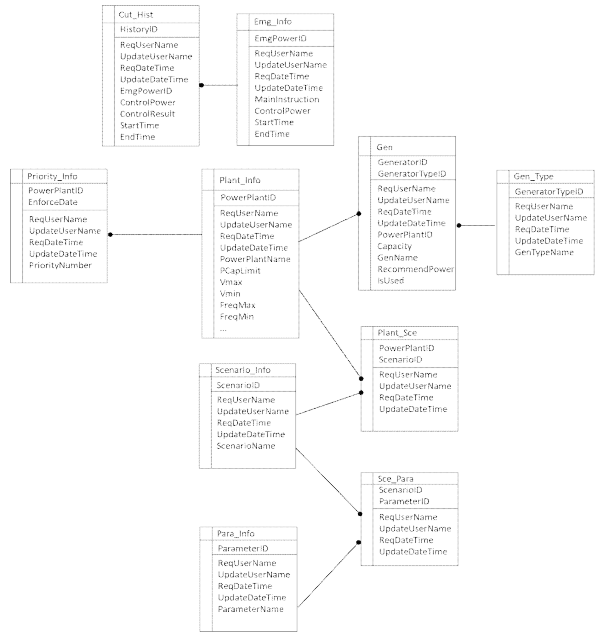


그림 2. 출력제어시스템 ERD 설계  
Fig. 2. Design of Curtailment Control System ERD

### 2.4 제어 우선순위 생성 방법

#### 2.4.1 DQN 정의

DQN(Deep Q Network)을 사용하여 최적 제어 행동 학습 및 제어 우선순위 설정에 대한 구성요소와 동작원리를 아래와 같이 구성하였다.

상태	발전량, 전력수요, REC가격, SMP가격, 운영비용, ESS 총방전량, 날씨, 발전기 수명, 고장여부
행동	출력 증가, 출력 유지, 출력 감소
보상	출력제어허용량 달성, 경제적 효율성 확보

그림 3. 우선순위 생성을 위한 DQN 정의  
Fig. 3. Definition of DQN for generating priority

기존의 Q Learning은 모든 상태와 행동에 대한 큐 함수 데이터를 저장하여 학습을 수행하고 행동을 결정한다. Deep Q Network는 CNN을 이용해 상

태를 입력받아 각 행동의 Q 함수를 근사해주기 때문에 많은 상태와 행동이 존재하는 출력제어 환경에서 학습이 가능하다. 강화학습[2]을 위한 파라미터는 그림 3과 같이 정의한다.

#### 2.4.2 Q-value 출력

하나의 통합신경망인 Q-네트워크를 설정하여 각 발전기 및 행동에 대한 Q-value를 출력한다.

##### (1) 입력층

상태벡터  $S = [\text{발전량}_i, \text{전력수요}_i, \text{운영비용}_i, \text{ESS충방전량}_i, \text{날씨}_i, \text{발전기수명}_i, \text{고장여부}_i, \text{REC}, \text{SMP}]$

행동벡터  $A = [\text{발전기}_{i,j}]$  ( $i = \text{발전기ID}, j = \text{행동ID}$ )

위 식과 같이 입력층을 설정한다.

##### (2) 은닉층

$$R_i = \alpha \text{출력목표}_i + \beta(\text{REC}_i + \text{SMP}_i + \text{운영비용}_i) + \gamma \text{ESS충방전량}_i - \delta(\text{발전기수명}_i + \text{고장여부}_i)$$

( $R = \text{보상}, \alpha, \beta, \gamma, \delta = \text{하이퍼파라미터}$ )

발전기마다 발전가능한 출력목표가 설정된 기준 테이블 DB를 확보한다. 출력목표와 현재 출력량을 비교하여 출력제어가능량을 산출하고 그에 따른 경제적 효율성이 타당한지 확인한다. 발전기의 수명이 길거나 고장난 발전기는 패널티를 적용한다.

##### (3) 출력층

모든 가능한 행동 조합에 대한 Q-value를 출력한다.

$[Q_{i,j}]$  ( $i = \text{발전기ID}, j = \text{행동ID}$ )

#### 2.4.3 Q-value 업데이트

상태-행동-보상-다음상태 튜플을 저장하고 무작위 샘플링하여 학습한다. 처음의 Q-value는 중요하지 않으며, 반복을 통해 점점 목표값에 근사해진다.

$$Q(S, A) \leftarrow Q(S, A) + \alpha \times [R + \beta \max_{A'} Q(S, A') - Q(S, A)]$$

( $\alpha = \text{학습률}, \beta = \text{할인율}, \max_{A'} Q(S, A') = \text{다음 상태에서 가능한 행동의 최대 Q값}$ )

\* DQN에서는 MSE(평균 제곱 오차)를 손실 함수로 사용한다.

$$\text{MSE} = (\text{목표값} - \text{예측값})^2 = (R + \beta \max_{A'} Q(S, A', \theta^-) - Q(S, A, \theta))^2 \quad (\theta = \text{매개변수})$$

손실 함수를 최소화하도록 가중치를 업데이트하면서  $\alpha$ (학습률)를 변경하여 Q-value를 예측한다. 할인율은 0과 1사이의 값으로, 시간이 많이 지날수록 보상값을 줄이기 위해 사용한다.

#### 2.4.4 우선순위 출력

Q-value 벡터가 추출되면, Q-value가 높을수록 발전기별 우선순위를 높게 설정한다. 그림4로 강화학습 기반으로 제어 우선순위를 선정하는 과정을 도식화하였다.

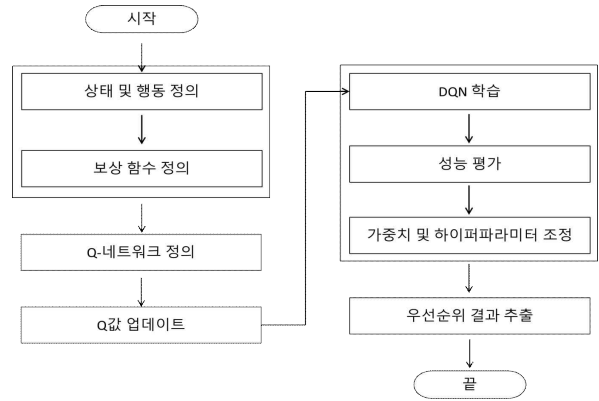


그림 4. 제어 우선순위 선정방법  
Fig. 4. Control priority order chart

### III. 결 론

변화하는 날씨와 전력 시장에 능동적으로 대응할 수 있는 출력제어 시스템에서 우선순위 설정 방법을 제안하였다. 전력시장에서 수익성이 높은 자원의 출력량을 조절함으로써 경제적 효율성을 증대시킬 수 있었다. 이러한 출력제어 기술을 연구함으로써, 가상발전소에 포함된 발전사업자에게 신뢰성과 안정성을 보장하고, 앞으로 더 많은 분산전원을 효과적으로 통제하는 미래 에너지 산업의 중심이 될 VPP 운영 플랫폼의 확산에 기여할 것으로 기대된다.

### 참 고 문 헌

[1] Park Sung Wan, Kim Dong Woo, Kim June Ho and Kim Su Hyun, "VPP Platform Development for DER(Distributed Energy Resource) Power Control", 대한전기학회 정보 및 제어 논문집, pp. 228-229, October 2023.

[2] Kim Joong-heon, Korea University, "심층 강화 학습 기술 동향", 딥러닝 학습 및 응용 기술, p31-32, April 2022.

# 반도체 제조 공정간 데이터 수집 시스템 구현

전보광\*, 윤여은\*, 윤동식\*, 강정훈\*\*

## Implementing a semiconductor manufacturing interprocess data collection system

Bokwang Jeon\*, Yeoeun Yoon\*, Dongsik Yoon\*, and Jeonghoon Kang\*\*

### 요 약

본 논문에서는 반도체 제조 공정에서 발생하는 데이터를 수집하고, 안전하게 저장 및 전송할 수 있는 엣지 컴퓨팅 기반의 데이터 수집 시스템을 설계하였다. 이 시스템은 우분투 기반의 GPU 장착 PC를 데이터 수집 장치로 활용하고, 도커 기반의 MinIO 오브젝트 스토리지를 통해 수집된 데이터를 관리하며, ZeroTier를 사용하여 보안이 강화된 가상 네트워크를 구축하여 데이터를 전송한다. 이를 통해 보안 구역에서도 데이터 전송이 가능하며, 분석과 활용을 위한 안전하고 효율적인 데이터 관리 환경을 제공한다.

### Abstract

In this paper, we designed an edge computing-based data collection system that can collect, securely store, and transmit data generated in the semiconductor manufacturing process. The system utilizes Ubuntu-based GPU-equipped PCs as data collection devices, manages the collected data through Docker-based MinIO object storage, and transmits the data by building a security-enhanced virtual network using ZeroTier. This enables data transfer even in secure areas and provides a secure and efficient data management environment for analysis and utilization.

### Key words

Semiconductor Manufacturing Process, Data Collection, MinIO Object Storage, Secure Network

## 1. 서 론

반도체 제조 공정에서 데이터의 수집과 분석은 공정 내 발생하는 이상을 신속하게 감지하고 공정

효율을 높이는 데 중요한 역할을 한다. 장비 상태의 실시간 모니터링, 품질 평가, 잠재적 문제 예측을 위해서는 다양한 데이터가 필요하며, 이를 효과적으로 활용하려면 데이터의 적시 수집과 안전한 보관,

\*한국전자기술연구원, azxcv24@keti.re.kr, yeoeun278@keti.re.kr, dsyoon@keti.re.kr

\*\*한국전자기술연구원, budge@keti.re.kr(교신저자)

※ "이 논문은 2024년도 산업통상자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임(20025255, 반도체 장비 지능화를 위한 공정 데이터 플랫폼 기술 개발) This paper was supported by Korea Planning & Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korea government(MOTIE)(20025255, Industry data analytics software Platform development for semiconductor equipment and process)"



관리, 분석이 필수적이다.

기존 데이터 수집 시스템은 반도체 공정의 외부 네트워크와 원활히 연결되지 않거나 포트포워딩과 같은 보안 제약이 있어 실시간 데이터 수집 및 전송에 어려움이 발생한다. 따라서 반도체 공정 데이터의 효율적인 수집과 전송을 위해 별도의 데이터 수집 장비가 필요하다.

본 논문에서는 반도체 제조 공정에서 발생하는 데이터를 실시간으로 수집하고 안전하게 저장하며, 보안성을 유지하면서 데이터를 전송할 수 있는 데이터 수집 시스템을 설계하고자 한다. 특히 데이터의 일관성을 보장하기 위해 MinIO를 활용한 오브젝트 스토리지와 보안 네트워크 환경 구축을 위해 ZeroTier를 결합하여 안전하고 효율적인 데이터 전송 경로를 제공한다. 이를 통해 공정 효율성을 향상시키고, 수집된 데이터를 바탕으로 분석 및 최적화할 수 있는 환경을 구축하는 데 기여하고자 한다.

## II. 데이터 수집을 환경 구현

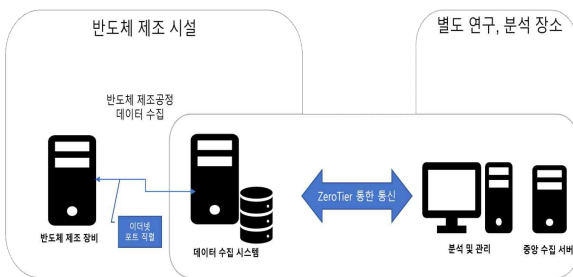


그림 1. 시스템 통신 요식도

본 논문에서는 반도체 제조 공정에서 발생하는 데이터를 수집하고, 안전하게 저장 및 전송할 수 있는 데이터 수집 시스템을 설계하였다. 제안된 시스템은 우분투 기반의 GPU가 장착된 PC를 데이터 수집 장치로 사용하여, 데이터 분석 및 학습에 필요한 높은 연산 능력을 제공한다. 데이터 수집 장치는 반도체 제조 장비의 기존 모니터링 장비로부터 데이터를 이더넷 직렬연결을 통해 가져와 수집하며, 수집된 데이터는 시스템 내 로컬 MinIO에 저장되어 관리한다.

본 시스템에서는 데이터의 효율적인 저장과 관리를 위해 MinIO를 오브젝트 스토리지로 활용하며, 데이터 수집 장치에서 발생하는 데이터를 안정적으로 처리하기 위해 MinIO를 도커 컨테이너로 배포하였다. MinIO는 Amazon S3와 호환되는 API를 제공하여 수집 시스템과 중앙 서버 간 데이터 전송을 간소화하고, 다양한 응용 프로그램과의 호환성을 보장한다. 오브젝트 스토리지 방식으로 데이터를 저장함으로써 데이터의 효율적 저장과 일관된 조회 및 접근이 가능하며, 시스템의 확장성과 관리 편의성도 높아진다.

반도체 제조 공정 장비가 위치한 구역은 주로 보안 구역으로 설정되어 외부 네트워크와의 연결이 제한적이다. 이를 해결하기 위해 ZeroTier를 활용하여 보안성이 강화된 가상 네트워크 환경을 구축에 활용하였다. ZeroTier는 가상 사설 네트워크를 제공하여 포트 포워딩 없이 안전한 데이터 통신을 지원한다. ZeroTier 서버는 비용 부담이 없는 오픈소스 버전인 ZeroTierOne의 Self-hosting 방식으로 구현하였으며, 이 버전의 경우 관리용 웹 UI가 제공되지 않는 제한 사항이 있다. 이러한 한계를 보완하기 위해 zero-ui 프로젝트를 활용하여 웹 관리 페이지를 구축하였고, 이를 통해 수집 시스템이 가상 네트워크에 안전하게 접속할 수 있도록 접근 권한을 부여하고 네트워크 상태를 모니터링할 수 있도록 하였다. 또한, 데이터 수집 장비에 ZeroTier 클라이언트를 설치하여 가상 네트워크에 연결하고, 웹 UI에서 접근 권한을 관리함으로써 데이터 전송의 보안성을 더욱 높였다.

시스템 구현 과정에서는 파이썬 기반 코드를 활용하여 기존 모니터링 장비에서 수집된 데이터를 데이터 수집 시스템으로 전송하였으며, 도커로 구동 중인 로컬 MinIO 서버에 업로드가 진행된다. 또한, 별도의 다른 오브젝트 스토리지와 버킷 동기화를 통해 수집된 데이터의 관리와 접근성을 더욱 용이하게 하여 전체 데이터 관리 과정의 효율성을 향상시켰다.

### III. 결 론

본 논문에서 설계한 시스템은 반도체 제조 공정에서 발생하는 데이터를 실시간으로 안전하게 수집, 저장 및 전송할 수 있는 효율적인 솔루션을 제공한다. GPU 장착 우분투 기반의 데이터 수집 장비와 도커 기반 MinIO를 활용하여 데이터 수집 및 관리를 용이하게 하고, ZeroTier를 사용해 보안 네트워크를 구축하여 보안 구역 내 원활한 접근과 통신을 지원함으로써 데이터 분석 및 활용의 편의성을 높였다.

제안된 시스템은 반도체 제조 공정에서 실시간 데이터 수집과 보안성을 동시에 강화할 수 있는 효과적인 솔루션으로, 다양한 환경 및 격오지 데이터 수집에도 도움을 줄 수 있을 것으로 기대된다.

# GPS 기반 차량 데이터와 LLM을 활용한 데이터 품질평가 자동화 연구

김태경\*, 권산\*, 최민호\*\*, 신청철\*\*\*, 최호식\*\*\*\*

## Automated Data Quality Assessment Using GPS-Based Vehicle Data and Large Language Models

Taekyoung Kim\*, Gwon San\*, Minho Choi\*\*, Sungchul Shin\*\*\*, and Hosik Choi\*\*\*\*

본 연구는 대형 언어 모델(LLM)을 활용하여 차량의 위치 좌표와 속도 데이터를 기반으로 복잡한 업무규칙을 자동으로 평가하는 시스템을 제안한다. GPS 데이터는 물류, 교통 관리, 도시 계획 등 다양한 산업에서 중요한 인사이트를 제공하지만, 높은 품질의 데이터를 유지하기 위해서는 철저한 규칙 기반 품질 검증이 필수적이다. GPS 수집 장치를 통해 확보한 차량 데이터의 품질을 평가하기 위해, 본 연구에서는 GPS 데이터의 완전성, 일관성, 정확성 등을 검토하는 LLM 기반 품질평가 모델을 설계하였다[1][2]. 또한 데이터의 결측치와 이상치를 탐지하고 업무규칙에 따라 수집된 차량 위치 데이터의 품질을 평가할 수 있도록 전처리와 미세 조정 과정을 포함하였다[3]. 다양한 조건에서의 실험 결과를 통해 해당 시스템의 유효성을 검증하였다[2]. 제안된 시스템은 교통 모니터링과 물류 최적화 등 다양한 분야에 활용될 수 있으며, 실시간 데이터 스트리밍 환경에서 품질 보증 프로세스를 자동화함으로써 데이터 처리의 효율성을 크게 향상시킬 것으로 기대된다[1][2]. 이를 통해 교통 흐름 분석, 안전성 향상, 운행 경로 최적화와 같은 다양한 응용 분야에 기여할 수 있다.

### Key words

data quality assessment, large language model, vehicle tracking

### 참고문헌

- [1] Lu, Ruci-Shan, Ching-Chang Lin, and Hsiu-Yuan Tsao. "Empowering Large Language Models to Leverage Domain-Specific Knowledge in E-Learning." *Applied Sciences* 14.12 (2024): 5264.
- [2] Jeong, Cheonsu. "Fine-tuning and utilization methods of domain-specific llms." *arXiv preprint arXiv:2401.02981* (2024).
- [3] Wang, Zige, et al. "Data management for large language models: A survey." *arXiv preprint arXiv:2312.01700* (2023).

---

\* 서울시립대학교 도시빅데이터융합학과 석사과정, tkdl0324@uos.ac.kr, jmt30269@uos.ac.kr

\*\* (주)크립토탭 데이터사업개발실 연구원, minho8173@cryptolab.co.kr

\*\*\* (주)크립토탭 데이터사업개발실 이사, scshin73@cryptolab.co.kr

\*\*\*\* 서울시립대학교 도시빅데이터융합학과 교수, choi.hosik@uos.ac.kr(교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NO.2022M3J6A1084845)

# 차량 탑승 상황에서의 스트레스 상태 분류 시스템 개발

김동연\*, 변상봉\*, 이성훈\*, 한동철\*, 조수현\*

## Development of a Stress State Classification System for Vehicle Occupancy

Dong-Eon Kim\*, Sang-Bong Byun\*, Sung hun Lee\*, Dong Cheul Han\* and Soo-Hyun Cho\*

### 요 약

본 논문은 차량 탑승 상황에서의 스트레스를 모니터링하기 위하여 데이터 전처리 작업 및 AI 모델 개발을 통한 시스템을 구축하는 방법을 제안한다. 감정 기반의 분류 모델들과 다르게, 스트레스를 분류하기 위하여 러셀 모델을 응용한 분류 방법을 제안한다. 또한, 개별 연구마다 지니는 자체 데이터-셋의 한계를 극복하기 위하여 DEAP 공개 데이터-셋을 활용하여 스트레스 상태를 분류하는 모델을 제안한다. 딥러닝의 학습을 위하여 데이터의 취득과 전처리 과정을 포함한 상태 분류 시스템을 제안한다. 상태 분류에 사용된 모델은 해당 도메인에서 가장 많이 활용되거나 SOTA (State-of-the-Art) 모델에 해당되는 딥러닝 모델들을 선택한다. 해당 모델들 간의 성능 비교를 통하여, 해당 도메인에서의 최적화된 모델을 판별한다.

### Abstract

This paper proposes a system for monitoring stress during vehicle occupancy by building a system through data preprocessing and AI model development. Unlike emotion-based classification models, it introduces a classification method for stress based on the Russell model. Furthermore, to overcome the limitations of individual datasets used in separate studies, this paper suggests utilizing the publicly available DEAP dataset to develop a model that classifies stress states. For deep learning training, it proposes a state classification system that includes data acquisition and preprocessing processes. The models used for state classification are selected from the most commonly used or state-of-the-art (SOTA) deep learning models in the field. By comparing the performance of these models, the study identifies the optimized model for this domain.

### Key words

Stress monitoring, DEAP dataset, EEG, Russell's model, Tception model

## 1. 서 론

국토교통부에서는 2025년도에 완전자율주행(FSD, 레벨4) 버스와 셔틀을 서비스하고, 2027년도에 FSD

\* 구미전자정보기술원 ICT디바이스연구센터, email : dekim@geri.re.kr(교신저자), sbbyun@geri.re.kr, leesh@geri.re.kr, cataegu07@geri.re.kr, shcho@geri.re.kr

※ 이 연구는 2024년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임('20017474')

상용화를 목표로 모빌리티 혁신 로드맵을 2022년 9월 19일에 발표하였다. 이러한 완전 자율주행이 대중화(자율주행 신차 보급률이 50% 이상)될 경우에 교통사고 사망률이 1,000명 이하로 감소하여 2021년도 2,916명과 대비해 65.7% 낮아질 것으로 예상된다.

이러한 목표에도 불구하고 FSD 상용화를 위해서는 규제 조건 및 기술적인 성숙도 문제가 남아있으며, 자율주행 차량 탑승자의 불안감을 해소하기 위한 개인적인 주행 스타일의 선호도 및 경험을 반영하는 시스템 개발이 요구된다.

## II. DEAP data-set과 Russell의 circumplex 모델

본 논문에서는 개별 연구 집단에서 자체 데이터셋을 통한 연구의 한계를 보완하기 위하여, IEEE Transactions on effective computing 논문지에 발표되어 4143회 인용되어 해당 분야에서 공개 데이터셋으로 널리 사용중인 DEAP 데이터셋을 활용한다 [1][2].

기존에 Emotion의 상태 분류에 비하여 Stress 상태 분류의 연구가 적은 이유는 EEG 기반의 공개 데이터셋들이 Emotion 기반의 데이터를 제공하기 때문이다[3][4].

이러한 점을 고려하여 Emotion 기반의 데이터셋을 Stress 상태 분류를 위한 데이터로 활용하기 위해 James Russell이 발표한 원형 모델을 응용한다 [5]. Russell은 통계 기법을 사용하여 긍정적인 상관관계를 기반으로 감정 평가를 그룹화하여, 감정을 Valence-Arousal 공간에 매핑하는 방법을 제안하였다. 이러한 방법을 응용하여 IEEE Transactions on Affective computing 논문지에서 감정의 원형 모델에서 스트레스 상태의 매핑에 대한 개념을 응용하여 그림 1의 왼쪽과 같은 분류 방법이 제시되었다. 즉, Valence와 Arousal의 두 축을 기반으로 2사분면의 일부분을 스트레스로 정의하여, 스트레스와 비 스트레스의 2가지로 분류한 방법이 제시되었다[6].

본문에서는 위에서 제시된 스트레스 분류 방법을 확장하여, stress를 고강도, 저강도, 없음의 3단계로 구분하는 방법을 그림 1의 우측과 같이 제안한다.

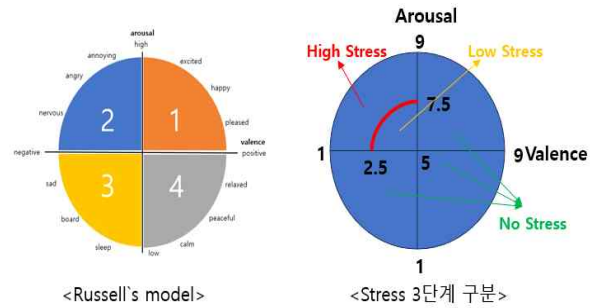


그림 1. 러셀의 원형 모델에서 착안한 Stress 3단계 구분  
Fig. 1. Three-level stress classification inspired by Russell's circumplex model

즉, 해당 모델에서 Arousal이 7.5 이상, Valence가 2.5 이하인 경우를 High stress 상태, Arousal의 지표가 5인 이상인 경우와 Valence의 지표가 5 이하인 경우를 세분화하여 Low stress 상태, 그 이외의 사분면을 No stress 상태로 분류한다.

## III. 딥러닝 모델 간 성능 비교

EEG 딥러닝 모델의 학습을 위하여, 데이터의 전처리 과정을 수행한다. 근육의 움직임과 drift 구성 요소 그리고 power-line 등으로부터 유발될 수 있는 잡음 등을 필터링하는 기술을 그림 2와 같이 설계한다.

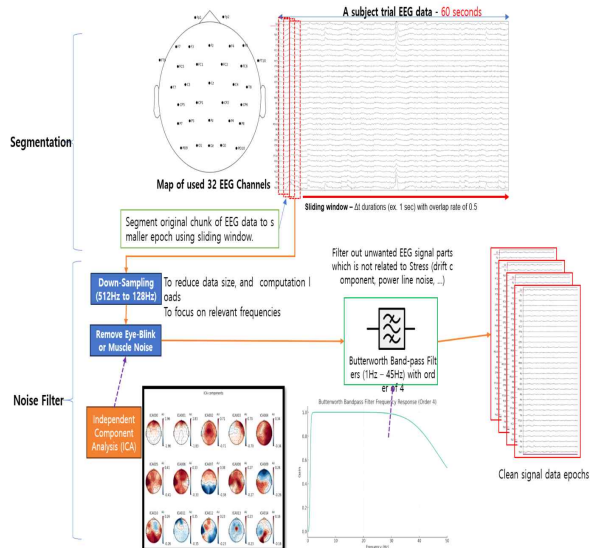


그림 2. EEG 학습을 위한 전처리 과정 설계  
Fig. 2. Design of preprocessing steps for EEG learning

이러한 전처리 과정을 거친 데이터는 그림 3와 같이 EEG 학습을 위한 딥러닝 모델에 입력된다. 본문에서는 SOTA 모델 및 해당 분야에서 널리 사용되는 모델들인 EEGNet, EEGNeX, TSception들을 구현하고, DEAP 데이터-셋의 적용을 통한 성능을 비교한다[7][8][9]. 이후, 학습을 위하여 K-fold split 방법을 이용하여 3단계의 스트레스 분류를 위해 모델이 훈련되고 평가된다.

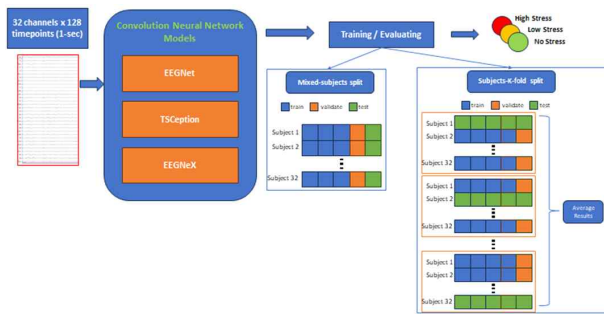


그림 3. 스트레스 단계 모니터링을 위한 AI 모델 학습  
Fig. 3. AI model learning for stress level monitoring

#### IV. 결 론

3장에서 소개된 딥러닝 모델들이 DEAP 데이터-셋에 적용된 성능은 아래 그림 4, 표 1과 같다.

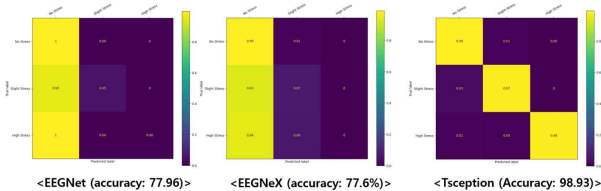


그림 4. 대표적인 딥러닝 모델들 간의 성능 비교  
Fig. 4. Performance comparison among representative deep learning models

표 1. 대표 딥러닝 모델들 간의 accuracy 성능 비교  
Table 1. Accuracy performance comparison between major deep learning models

모델	EEGNet	EEGNeX	TSception
Accuracy	77.96	77.6	98.93

본문에서는 차량 탑승 상황에서의 스트레스 모니터링 감지를 위한, 스트레스 상태의 3가지 분류 방법을 제안하고, 이를 통한 스트레스 상태 분류 모델

을 개발했다. 자체적으로 구축된 데이터-셋의 한계를 보완하기 위한 DEAP 데이터-셋을 활용하고, 대표적인 딥러닝 모델들에 적용하여 스트레스 상태의 분류 성능을 비교한다. 해당 도메인에서는 TSception의 성능이 98.93%로 가장 우수하다. 이러한 딥러닝 모델의 적용으로 탑승자가 편안함을 느끼는 주행 스타일을 파악할 수 있다. 향후 이러한 기술의 적용으로, 탑승자의 선호도에 따른 주행 파라미터 설정을 통해 자율주행 모드에서의 불안함 또는 불편함을 개선 시킬 수 있다.

#### 참 고 문 헌

- [1] Ding, Yi, et al. "Tsception: Capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition." IEEE Transactions on Affective Computing (2022).
- [2] Du, Xiaobing, et al. "An efficient LSTM network for emotion recognition from multichannel EEG signals." IEEE Transactions on Affective Computing 13.3 (2020): 1528-1540.
- [3] Gao, Zhongke, et al. "A channel-fused dense convolutional network for EEG-based emotion recognition." IEEE Transactions on Cognitive and Developmental Systems 13.4 (2020): 945-954.
- [4] Xu, Tao, et al. "E-key: An EEG-based biometric authentication and driving fatigue detection system." IEEE Transactions on Affective Computing (2021).
- [5] Fan, Chaojie, et al. "Detection of train driver fatigue and distraction based on forehead EEG: a time-series ensemble learning method." IEEE transactions on intelligent transportation systems 23.8 (2021): 13559-13569.
- [6] Song, Tengfei, et al. "Graph-embedded convolutional neural network for image-based EEG emotion recognition." IEEE Transactions on Emerging Topics in Computing 10.3 (2021): 1399-1413.

- [6] Song, Tengfei, et al. "Graph-embedded convolutional neural network for image-based EEG emotion recognition." *IEEE Transactions on Emerging Topics in Computing* 10.3 (2021): 1399-1413.
- [7] Ding, Yi, et al. "Tsception: Capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition." *IEEE Transactions on Affective Computing* (2022).
- [8] Chen, Xia, et al. "Toward reliable signals decoding for electroencephalogram: A benchmark study to EEGNeX." *Biomedical Signal Processing and Control* 87 (2024): 105475.
- [9] Lawhern, Vernon J., et al. "EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces." *Journal of neural engineering* 15.5 (2018): 056013. 2261 citations.

# 개인 가스 데이터 보호를 위한 Gasistor가 통합된 의료 정보 암호화 시스템

권익근\*, 김태기\*, 장유성\*, 김희동\*†

## Medical information encryption system with Gasistor integrated to protect personal gas data

Ikgeun Kwon\*, Yuseong Jang\*, Taegi Kim\*, and Hee-Dong Kim\*†

In this study, we proposed a gasistor (gas sensor + memristor) system that can encrypt the obtained data and manage it safely, and successfully demonstrated optimal gas sensing and encryption methods. First, in the case of proposed gasistor, we observed stable and liner response characteristics in the range of 10 to 50 ppm NO as well as very fast response of 12  $\mu$ s in pulse mode. Then, we analyzed the encryption mechanisms to secure the gasistor response and found that using a block cipher was superior to using a stream cipher. Especially, when analyzing the eight different block mechanisms with different algorithms and key lengths, we found that the AES-192 mechanism had the most optimized performance (encryption strength: 2128 attempts, en/decryption time: 572,150/540,650  $\mu$ s). Consequently, these results indicates that the proposed gasistor encryption system has an excellent sensing ability as NO gas sensor while an outstanding data security using the AES-192 block algorithm.

### Key words

gasistor application, encryption system, No gas sensor, fast recovery

---

\* Department of Semiconductor Systems Engineering and Convergence Engineering for Intelligent Drone, Institute of Semiconductor and System IC, Sejong University, kig0800@sju.ac.kr, nr7782942@sju.ac.kr, ktg2356@naver.com, khd0708@sejong.ac.kr(†)

※ This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00419201).



# 최신 LLM 성능 평가를 위한 벤치마크 동향 조사

김만조\*, 박시형\*, 전석훈\*\*

## A Survey of Benchmark Trends for Evaluating Latest LLMs

Manjo Kim\*, Sihyeong Park\*, and Seokhun Jeon\*\*

대형 언어 모델(Large Language Models, LLMs)의 빠른 발전은 인공지능 분야에 상당한 변화를 불러왔으며, 자연어를 이해하고 생성하는 능력이 크게 향상되었다. 이러한 LLM의 성능 평가는 수학 능력을 평가하는 MATH, 언어 이해, 추론, 문제 해결 등 일반적인 지능 수준을 평가하는 AGIEval 등 다양한 벤치마크를 사용한다[1]. 하지만 LLM이 빠르게 발전하면서 2022년 1월 이후 BoolQ, HellaSwag, MMLU를 비롯한 많은 벤치마크의 평가 지표가 포화함에 따라 LLM의 성능 차이를 구별하는 변별력이 줄어들고 있다[2]. 이를 극복하기 위한 다양한 벤치마크 개선 연구가 진행되고 있는데, 그중 주목할 만한 것은 BigBench와 MMLU-Pro이다. BigBench[3]는 기존의 단순한 언어 이해나 생성 능력뿐만 아니라, 창의적 사고, 상식적 추론, 윤리적 딜레마 해결 등과 같은 복잡한 문제에 대해 모델의 성능을 평가함으로써 LLM의 전반적인 능력을 더욱 포괄적으로 평가하였다. MMLU-Pro[4]는 기존의 MMLU에 존재하는 오류와 정확도 저하 문제를 보완하고 전문 지식수준을 더욱 확장한 최신 벤치마크이며, 선택지를 확장하여 평가 지표를 더 명확하게 하였다. 본 논문에서는 Microsoft, MistralAI, Meta 사에서 개발한 많이 사용되는 공개 LLM의 논문 및 자료를 바탕으로 기존 벤치마크의 최근 활용 동향을 조사하였다. 조사 결과, 가장 많이 사용된 벤치마크는 MMLU로 약 70% 이상의 연구에서 사용되었다. MMLU는 고등학교 수준부터 전문 자격 시험까지 다양한 교육 자료에서 추출되어 폭넓은 분야에서 모델의 전반적인 지식 수준을 평가할 수 있다. 또한, 사지선다형 선택지를 통해 명확한 평가가 가능하다는 것이 장점이다. 반면 TriviaQA는 약 20% 정도의 사용률을 보였다. 웹 크롤링을 통해 자동 수집된 데이터를 기반으로 구성되어 잡음과 오류, 모호성, 시대에 뒤떨어진 정보가 포함되어 있다. 또한, 개방형 답변방식을 채택하여 평가 기준이 불명확해 LLM의 정확한 성능 평가가 어려운 것이 특징이다. 본 논문에서는 최신 LLM 성능 평가를 위한 벤치마크 동향 조사를 Microsoft, MistralAI, Meta 사의 논문 및 공개자료를 바탕으로 진행하였다. 최신 모델들을 더욱 정확하게 평가하기 위해서는 폭넓은 분야 및 지식수준이 필요하며, 평가 지표가 명확해야 함을 확인하였다. 이러한 벤치마크 특성에 더해, 좀 더 변별력 있는 평가를 위해 창의적 사고, 상식적 추론, 윤리적 딜레마 해결 등과 같은 복잡한 문제가 포함되어야 하며, 성능이 더욱 명확하게 구분될 수 있는 지표의 개발이 필요하다.

### 참고문헌

- [1] CHANG, Yupeng, et al. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 2024, 15.3: 1-45.
- [2] LIANG, Percy, et al. Holistic evaluation of language models. *arXiv preprint arXiv:2211.09110*, 2022.
- [3] SRIVASTAVA, Aarohi, et al. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. *arXiv preprint arXiv:2206.04615*, 2022.
- [4] WANG, Yubo, et al. Mmlu-pro: A more robust and challenging multi-task language understanding benchmark. *arXiv preprint arXiv:2406.01574*, 2024.

---

Korea Electronics Technology Institute, Email(\* 9204kmj@keti.re.kr, \* sihyeong@keti.re.kr, \*\* seokhun.jeon@keti.re.kr)  
※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임  
(RS-2024-00402898, 시뮬레이션 기반 고속/고정확도 데이터센터 워크로드/시스템 분석 플랫폼 개발)

# 스마트 공장 내 초음파 기반 통신 시스템의 성능 향상을 위한 프리앰블 심볼 설계

정용안\*, 변상봉, 한동철, 조수현, 이성훈\*\*

## Preamble Symbol Design for Performance Enhancement of Ultrasound Based Communication Systems in Smart Factories

Yong-An Jung\*, Sang-Bong Byun, Dong-Cheul Han, Soo-Hyun Cho, and Sung-Hun Lee\*\*

최근 들어 초음파 통신을 이용한 데이터 송수신 기술이 제안되고 있다. 초음파 대역은 인간이 감지할 수 없는 영역인 20KHz ~ 수백MHz 대역을 의미하는데, 일반적인 산업용 초음파의 주파수 대역은 통상적으로 20KHz ~ 40KHz 대역에 해당한다. 초음파 통신은 주로 저전력, 저속 통신이 필요할 때 사용되며, 비접촉 방식이나 물속 데이터 전송, IoT 기기 등에서 활용될 수 있다. 이러한 초음파 통신에서는 다중 경로 페이딩과 주파수 선택적 페이딩을 효율적으로 처리하고, 특히 초음파의 대역폭이 상대적으로 좁을 수 있기 때문에 OFDM을 통해 다중 채널을 구성하여 효율적인 데이터 전송을 가능하게 한다. 하지만 OFDM 기반 초음파 통신 시스템은 동기화 오차에 매우 민감하며, 동기화 오차의 존재는 시스템의 심각한 성능 저하를 유발시킨다. 따라서, 본 논문에서는 초음파 기반 통신 시스템의 성능 향상을 위해 정확하게 동기화 오차를 추정할 수 있는 프리앰블 구조를 제안한다. 제안된 기법은 송신기와 수신기 사이의 동기화를 위해 서로 약속된 길이가 256인 파일럿 부반송파를 가지는 프리앰블 심볼을 제안한다. 초음파 통신 시스템의 수신기에서 제안된 프리앰블 심볼을 이용하여 동기화 오차를 추정할 경우, 보다 정확하게 시간 및 주파수 오차를 추정할 수 있다는 장점이 있으며, 컴퓨터 모의 실험을 통해 MSE 성능을 확인하였다. 시뮬레이션 결과로부터 제안하는 프리앰블 구조가 실제 초음파 통신 시스템에 적용될 수 있음을 확인하였다.

### Key words

UltraSound Communication, OFDM, Smart Factory, Synchronization

### 참고문헌

- [1] T. Bos, M. Verhelst, and W. Dehaene, "A Flexible End-to-End Dual ASIC Transceiver for OFDM Ultrasound In-Body Communication," 2022 IEEE BioCAS., Oct., 2022.

\*구미전자정보기술원, yajung@geri.re.kr

\*\*구미전자정보기술원, leesh@geri.re.kr

※ This work was supported by the Korea Planning & Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korean government (MOTIE) (RS-2024-00445778, Development of AI-based high reliability robot-equipment and risk management integrated operation technology for digital transformation of missile manufacturing process).

# 전기차 실제 주행 데이터를 활용한 확장 칼만 필터 기반(EKF)기반 SoC 추정에 관한 연구

심별희\*, 전보광\*\*, 윤동식\*\*\*, 강정훈\*\*\*\*

## A Study on SoC Estimation Based on an Extended Kalman Filter(EKF) Using Electric Vehicle Driving Data

Byeolhee Sim\*, Bokwang Jeon\*\*, Dongsik Yoon\*\*\*, Jeonghoon Kang\*\*\*\*

### 요 약

본 연구에서는 전기차 배터리 팩의 충전량을 예측하기 위하여 확장 칼만 필터(EKF) 기반의 SoC(State of Charge) 추정 모델을 제안한다. 전류, 전압, 온도, 충전 여부 등의 다양한 실제 주행 데이터를 통합하여 EKF 알고리즘을 통해 전기차 배터리 팩의 충전상태를 추정하며, 그 결과에 대하여 연구하였다.

### Abstract

This paper proposes an Extended Kalman Filter (EKF)-based State of Charge (SoC) estimation model to predict the charge level of an electric vehicle battery pack. By integrating various real driving data, such as current, voltage, temperature, and charge status, the EKF algorithm is used to estimate the state of charge of the battery pack, and the results are analyzed.

### Key words

big data, electric vehicle driving data, Extended Kalman Filter (EKF), EV battery

## 1. 서 론

국가 기간산업이자 종합산업으로 한국 경제를 견인해 온 자동차 산업은 최근 큰 기술적 변화를 거치

고 있다. 내연기관차에서 전기차 및 수소차 등 친환경차로의 기술적 이행으로 인해 그에 필요한 수많은 친환경차와 관련된 산업과 기술이 발전함에 따라 친환경차 시장이 지속적으로 성장하고 있다.[1]

\* 한국전자기술연구원 지능융합SW연구센터 연구원, nrjjbb@keti.re.kr

\*\* 한국전자기술연구원 지능융합SW연구센터 연구원, azxcv24@keti.re.kr

\*\*\* 한국전자기술연구원 지능융합SW연구센터 선임연구원, dsyoon@keti.re.kr

\*\*\*\* 한국전자기술연구원 지능융합SW연구센터 센터장, budge@keti.re.kr (교신저자)

※ 이 논문은 2024년 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0025765, 배터리 활용성 증대를 위한 BaaS(Battery as a Service) 실증 기반 구축)

그 중 전기차의 보급이 크게 확대되어 다양해진 전기차의 종류에 따라 배터리의 SoC는 주요 성능에 직접적인 영향을 미치는 중요한 지표로 이를 효과적으로 분석하기 위하여 전압, 전류 등의 간접적인 신호를 기반으로 분석이 필요하다.

이를 위해 본 연구에서는 수집된 실제 전기차 주행 데이터를 기반으로 배터리의 내부 저항 및 RC 회로 모델을 기반으로 비선형성을 가지는 출력특성을 단계마다 선형화시켜 상태를 평가하는 확장 칼만 필터(EKF) 알고리즘을 통해 SoC를 추정하는 방법에 대하여 연구하였다.[2]

## II. 본 론

확장 칼만 필터(EKF)는 비선형 시스템의 상태를 추정하는 데 사용되는 알고리즘으로 배터리의 SoC와 커패시터 전압을 추정하였다. EKF 알고리즘은 크게 ‘예측’과 ‘갱신’으로 구성된다. 예측 단계에서는 현재 상태와 전류를 통해 다음 시점의 SoC와 전압을 예측하고, 갱신 단계에서는 측정된 전압과 예측된 전압차를 기반으로 측정 잡음과 과정 잡음을 보정한다. 이를 기반으로 확장칼만필터(EKF)를 통해 SoC를 예측하였다.

EKF 모델의 초기 파라미터는 다음과 같이 설정하였다.

표 1. 배터리 모델 파라미터 [2]

파라미터	설정 값
R0	0.0020
R1	0.008
C1	3.435e+08

배터리 팩의 온도 및 충전 상태는 내부 저항에 큰 영향을 미치기 때문에 본 연구에서는 배터리 팩의 평균 온도에 따라 저항값을 동적으로 조정하였다. 또한, 충전 여부에 따라 전류의 방향을 조정하여 충전 중과 충전 중이지 않을 때에 따른 SoC 변화를 예측하고 실제 주행데이터와 비교하였다. 실주행 데이터기에 모델의 초기 SoC와 전압 값을 측정 값의 첫 번째 라인으로 설정하였다. 실험 결과, 제안된 EKF 모델은 실제 SoC와 전압에 비례하여

예측된 SoC 간의 오차를 최소화하면서 높은 정확도의 추정 결과를 보여주었다.

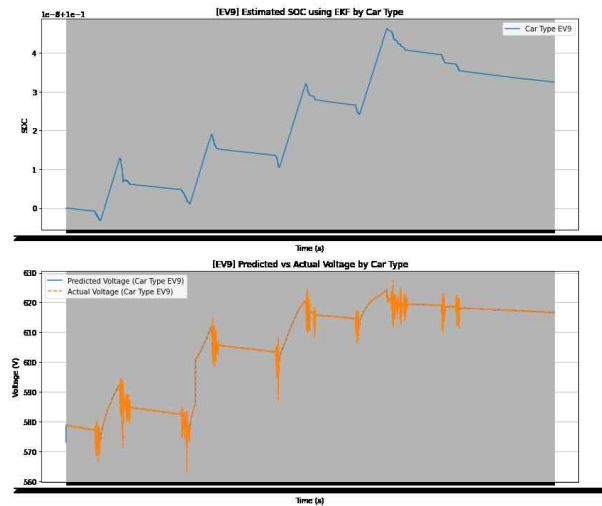


그림 1. EV9 차량의 예측된 SoC(상)와 실제 주행 데이터(하) 비교

## III. 결 론

본 연구에서는 전기차 배터리의 SoC를 예측하기 위해 확장 칼만 필터(EKF) 모델을 활용하였다. 전기차 실주행 데이터의 비선형 특성과 온도, 충전 상태에 따른 변화를 반영하여 SoC를 추정할 수 있었다. 열화된 배터리 데이터의 부족으로 많은 측면에 대해 깊은 분석과 검증이 부족하지만 전압, 충전여부, 전류값과 더불어 속도, 거리, 환경, 배터리 노후화 등 주요 데이터를 반영하여 배터리의 수명을 추정하는데에 주요 요인 파악 혹은 전기차 배터리 관리 시스템(BMS) 등 고도화하는 방향으로 지속적으로 연구한다면 전 주기적 배터리 관리 시스템 및 비즈니스 모델에 활용 되어질 것으로 기대된다.

## 참 고 문 헌

- [1] 김영룡, 김국동.(2019).전기차 및 수소차 생산을 위한 거래 협력 네트워크의 공간적 변화.경기연구원 기본연구,(0),1-70.
- [2] 유경상, 김병기, 김대진, 장문석, 고희상, 김호찬 (2017) "확장칼만필터를 이용한 전기자동차용 배터리 SOC 추정." 한국산학기술학회논문지, v.18 NO.10(2017), pp.15-23

# 전기차 충전 인프라를 활용한 전력 수요 반응에 관한 연구

전지예\*, 이성욱\*\*

## Application of Electricity DR Using Electric Vehicle Charging Infrastructure

Jeon Ji Ye\*, Lee Sung Uk\*\*\*

### 요 약

전기차 충전 인프라는 현재 전력 수요에 미미한 영향을 끼치고 있지만, 해당 인프라의 시장 규모는 점차 확대되고 있다. 시장 규모의 확대에 따라 전력 수요영향을 최소화하고 전력 계통의 안정화를 위한 방안이 필요하다. 이 논문에서는 전력 계통 안정화를 위한 전기차 충전 인프라를 활용하여 전력 수요반응 시장 적용 방안을 제안한다.

### Abstract

The electric vehicle charging infrastructure currently has a slight impact on power demand, but the market size of the infrastructure is gradually expanding. As the market size expands, measures are needed to minimize the impact of power demand and stabilize the power system. This paper proposes a plan to apply the power demand response market by utilizing the electric vehicle charging infrastructure for stabilizing the power system.

### Key words

EV, EV Charger, OCPP, Grid

## 1. 서 론

2023년 기준 국내 전기차 보급현황은 약 45만대, 전기차 충전기는 약 30만기이며, 환경부에 따르면 2030년 국내 보급 목표는 전기차 420만대, 전기차 충전기 123만기를 예상한다. [1] 전기차의 양적 성장에 따라 전기차 충전 시장 규모가 점차 확대되고 있으며, 다양한 전기차 충전 인프라는 충전기 제조사에 대한 의존성 개선과 호환성 증대를 위한 표준 프로토콜을 도입했으며, 전기차 충전 인프라는 일관

된 표준 프로토콜 기반으로 변화되고 있다. 전기차 충전 인프라는 일관된 표준 프로토콜 기반으로 활발한 고도화가 진행 중이다. 현재까지 전력시장에서 충전 인프라가 미치는 영향은 미미하지만, 앞으로 인프라의 양적 성장에 따라 전력 수요의 증가할 것으로 예상된다. 이에 따라 본 논문에서는 표준 프로토콜을 활용한 충전 인프라를 활용하여 전력 수요반응 시장 적용하여 전력 계통 안정화에 기여할 수 있는 방안을 제안한다.

\*한전KDN, jiy\_e\_0524@kdn.coml, \*\*한전KDN, lee-sung-uk05@kdn.com

## II. 본론

이 논문에서 다루는 표준 프로토콜 기반으로 충전 인프라는 충전 서비스를 제공하는 전기차 충전기, 충전 관리용 서버, 충전 인프라 관리용 운영시스템, 충전 서비스를 이용하는 사용자 APP/WEB 시스템으로 정의할 수 있다.

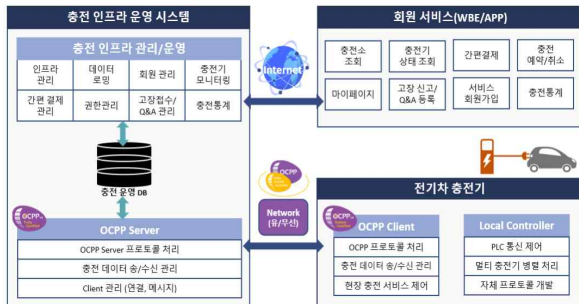


그림 1. 시스템 구성도

Fig. 1. System Configuration

또한, 본 논문에서 다루는 표준 프로토콜은 OCPP(Open Charge Point Protocol)이며, OCA(Open Charge Alliance)에서 제정한 충전스테이션의 운영, 유지관리를 목적으로 개발한 개방형 충전 프로토콜이다. 충전 관리용 서버와 전기차 충전기 사이에 프로토콜 기능 블록 항목은 다음과 같다.

표 1. 시스템 주요 기능 블록

Table 1. System Key Functional Blocks

메시지명	설 명
Authorization	충전 사용자 인증 처리
Transaction	충전 시작/종료 처리
RemoteControl	충전기 원격 제어
Availablilty	충전기 연결/상태 확인
Reservation	충전 예약 처리
SmartCharging	스마트 충전 제어
FirmwareManagement	충전 펌웨어 업데이트 처리
ISO 15118 CertificateManagement	ISO 15118 인증 처리

전기차 충전 같은 경우 설치 장소(상업/주거시설)별, 충전 용량별, 충전 시간별로 확연한 이용 패턴을 확인할 수 있다. 사용자들의 이용 패턴과 위 표에서 정의한 SmartCharging 기능 블록을 활용하여,

전력수요반응 시장에 접근할 수 있다. 전력수요반응 (Demand Response, DR) 제도는 전력수급의 안정화를 위해 동·하계 예비율 확보 등 단기 수급 대응 및 발전소 건설 불확실성을 해소하기 위한 대안이다. [3] 전기차 충전의 이용패턴을 보자면 급속 충전은 낮 시간대(09시~18시)에 휴게소와 같은 이동거점 지역에서 대부분 발생하고, 완속 충전은 저녁 시간대(23시~09시)에 아파트와 같은 생활거점 지역에서 발생하고 있다.

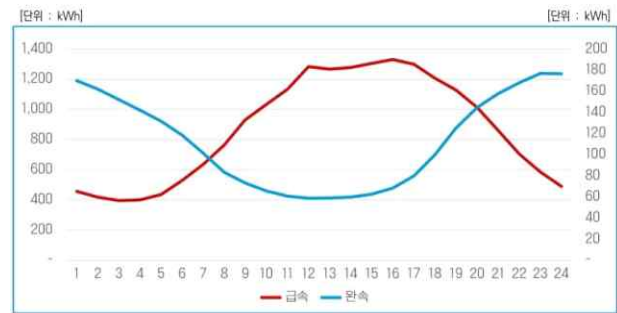


그림 2. 전기차 충전 이용 패턴 [2]

Fig. 2. Electric Vehicle Charging Usage Pattern

이때, SmartCharging 기능 블록에서 정의하는 Charging Profile 관련 메시지들을 사용하여 충전소의 전력 한도 및 전력 공급을 제어하고, 실시간으로 전력을 관리할 수 있다. Charging Profile에서 시간대, 충전 스케줄, 출력전력, 전류 등을 설정할 수 있으며, 충전 사업자는 실시간으로 전력수급의 상황을 확인하여 충전기에 Charging Profile 변경과 같은 충전 제어 명령을 내리거나, 설정된 스케줄에 따라 충전기가 Charging Profile을 변경하여 운영할 수 있다.

따라서, 재생에너지의 발전량이 많아 전력 공급량 보다 전력 수요량이 많을 경우 전기차 충전기 자체적으로 Charging Profile을 변경하여 평소보다 저렴한 가격에 전기차 충전 서비스를 제공하여, 충전 서비스 활용 고객과 DR시장 참여 사업자들은 경제성을 확보할 수 있다.

그림 3과 같이, ① 전력거래소에서 DR발령을 내리고 ② 수요관리사업자로부터 DR 정보를 수용하여 ③ DR 중개 플랫폼에서 전기차 충전 사업자로 해당 정보를 전달 받게 된다. ④ 마지막으로 전기차 충전 사업자는 해당 충전기 제어를 통해 DR 자원으로써 충전기를 활용할 수 있게 된다.



그림 3. 전기차 충전 인프라-DR 시장 Flow  
Fig. 3. EV Charging Infrastructure-DR Market Flow

### III. 결 론

본 논문을 통해 전기차 충전 인프라를 활용한 전력 수요 반응에 대한 전력계통 안정화 방안을 제시했다. 해당 기술을 적용한 전기차 충전 인프라의 보급을 통해 탄소중립 및 에너지 전환에 앞장 설 수 있으며, 전력 계통 안정화에 기여할 수 있다.

### 참 고 문 헌

- [1] Energy Planning Office Demand Forecast Team, "Concept and Status of Demand Response System", KPX., pp. 1-2, August 2023.
- [2] Energy Planning Office Demand Forecast Team, "Concept and Status of Demand Response System", KPX., pp. 36, August 2023.
- [3] Growth Technology Office Future Strategy Team, "Concept and Status of Demand Response System", KPX., pp. 1-2, August 2011.

# 5G 시험망 테스트베드 이용기업 비용절감에 관한 연구

추 인 오\*

## A study on cost reduction for companies using 5G test network testbeds

In-Oh Chu\*

### 요 약

5G 시험망 테스트베드는 글로벌 주요 이동통신사 동일 수준의 5G 시험망 테스트베드 구축하여, 5G 융합제품과 이동통신 네트워크 간의 연동 시험 환경을 제공하여 중소·중견 기업의 제품 경쟁력 강화 및 국내외 시장 진출 지원하고 있으며, 이를 이용한 5G 융합제품 개발기업들이 해외를 가지 않고 국내에서 테스트를 진행함으로써 발생하는 비용절감에 대한 조사를 실시하여 결과를 도출하였다.

이 논문에서는 5G 시험망 테스트베드를 이용함으로써 개발기업들의 비용절감을 통해 제품 신뢰성 확보, 개발기간 단축, 경쟁력 강화 등을 효과를 제시하고 있다.

### Abstract

By constructing a 5G test network test bed at the same level as that of major global mobile carriers, we are providing an interconnected test environment between 5G convergence products and mobile communication networks to help small and medium-sized companies strengthen their product competitiveness and advance into domestic and overseas markets. , companies developing 5G convergence products using this conducted a survey on the cost savings that would occur by conducting tests domestically rather than going overseas, and the results were derived.

In this paper, the use of a 5G test network test bed suggests the effects of securing product reliability, shortening the development period, and strengthening competitiveness through cost reduction for development companies.

### Key words

5G, 5G test-bed , Inter-Operability Test, 5G convergence service

## 1. 서 론

4차 산업혁명은 인공지능, 빅데이터 등 디지털 기술로 촉발되는 초연결 기반의 지능화 혁명으로 산업·소득분배·문화 등 사회 전반이 변화시키고

있으며, 디지털 전환(Digital Transformaion)은 IoT, 클라우드, 인공지능, 빅데이터 등 ICT 플랫폼으로 구축·활용하여 기존 전통적인 운영 방식과 서비스 등을 혁신을 통한 생산성과 효율성 제고가 가능하며, 5G는 대량의 데이터를 효과적으로

---

\* 구미전자정보기술원, iochu@geri.re.kr



수집-축적-분석-활용 하기 위한 기반 기술이다. IoT, Big Data, AI 등 신기술과 융합을 통해 4차 산업혁명과 기존 산업의 디지털 전환 활성화 가능하며, 우리나라는 2019년 세계 최초 5G 상용화에 돌입함에 따라, 5G 융합제품을 상용화하기 위한 국가 차원의 시험·검증 체계 마련 시급 하였으며, 국내 중소·중견기업은 5G 융합제품 시험·검증을 위한 인프라 이용에 어려움이 존재하여 5G 상용화 시점에 맞춰 5G 융합 디바이스 및 서비스 관련 중소기업 개발제품에 대한 국내외 시장진출을 위해 주요 이동통신사업자 상용망 수준의 5G 시험망 테스트베드를 구축하였다. 중소기업이 개별적으로 시험망 장비를 구축하기 어려운 현실 속에서 기업이 자유롭게 이용할 수 있는 개방형 테스트 환경을 지원하고 이를 이용한 기업들의 제품 신뢰성 확보, 제품경쟁력 강화를 위한 비용절감액 산출을 제시하고자 한다.

## II. 5G 시험망 테스트베드 비용절감액 도출

구미전자정보기술원은 5G 시험망 테스트베드를 통한 디바이스와 네트워크 시스템과의 망연동 시험을 지원하기 위해 5G NSA[비단독모드], SA[단독모드] 인프라를 구축, 운영, 보유하고 있으며, 최근 활성화 되고 있는 이음5G[특화망] 시험도 지원 가능하며, 이를 이용한 기업들의 제품 경쟁력 강화를 위한 비용절감액을 도출하였다.

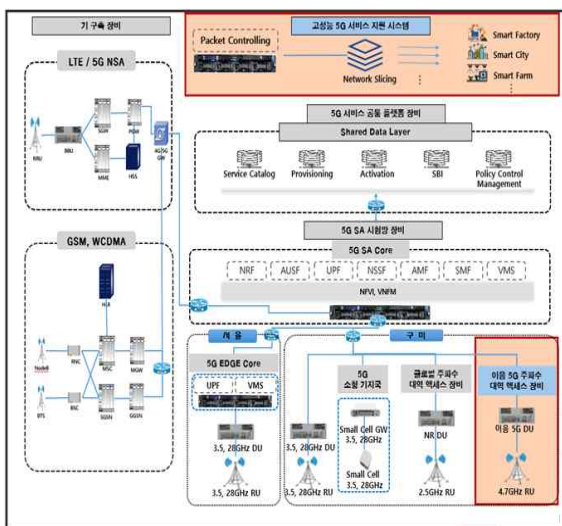


그림 1. 5G 시험망 테스트베드 구성도

Fig. 1. 5G test network testbed configuration diagram

5G 시험망 테스트베드는 중소기업의 상용화를 위한 신뢰성 확보를 위해 필요하며, 이동통신사는 통신 망 품질저하를 우려로 개발단계의 통신제품의 상용망 접속을 제한하고 있는 실정이며, 5G 서비스 상용화를 준비 중인 기업들은 신뢰성 확보를 위해 제품의 문제점 또는 사용자 체감 불만 등의 요소를 사전에 발견하고 이를 보완이 필요하므로, 개방형 테스트베드를 이용하여 다양한 망연동 시험이 가능하다.

또한 국내기업이 해외시장 진출하기 위해서는 해외 통신환경에서 제품의 성능검증으로 경쟁력 입증 필요하며, 해외 국가별 주파수 환경과 통신정책에 따라 서비스 환경이 다를 수 있어, 국내 상용화에 성공한 제품이더라도 해외 통신망 환경에서 시험은 필수여서 테스트베드를 활용하여 지원이 가능하다.



그림 2. 5G 시험망 테스트베드 시험실

Fig. 2. 5G test network test bed test room

5G 시험망 테스트베드의 비용절감을 도출하기 위해 2023년 1월~2023년 12월까지 총 107개 기업, 2682건의 시험 비용을 조사 분석하여 도출하였다.

비용절감분석을 위해 해외 대비 국내 테스트베드를 이용할 경우의 비용절감액은 해외비용(항공료, 출장인원, 체류기간, 체재비 등)에서 국내비용(체재비, 출장인원 등)의 비용 차액을 계산하여 평균을 산출하여 이용건수를 곱하여 도출하였다.

표 1. 비용절감액 산출식

Table 1. Cost savings calculation formula

(단위 : 원)

구분	비용절감액 산출식
비용절감액	(해외비용(A)-국내비용(B)) *이용 건수
해외비용(A)	비용①(항공료*출장인원/체류기간)+ 비용②(체재비(1일)*출장인원)
국내비용(B)	비용②(체재비(1일)*출장인원)

표 2. 시험 1건당 평균 비용절감액 분석 결과

Table 2. Analysis results of average cost savings per test

구분	해외	국내	비용절감
항공료의 평균	약 2,957,000원 (천원절삭)	-	약 2,957,000원
출장인원 평균	약 3.4명	약 2.6명	약 0.8명
해외 체류기간 평균	약 6일(소수점 아래 반올림)	-	약 6일
체재비(1일) 평균	약 290,000원 (소수점 아래 반올림)	약 100,000원 (소수점 아래 반올림)	약 190,000원
비용①의 평균	1,833,000원 (천원절삭)	-	1,833,000원
비용②의 평균	1,028,000원 (천원절삭)	274,000원 (천원절삭)	754,000원
비용①+②	2,861,000원 (천원절삭)	274,000원 (천원절삭)	2,587,000원

\*비용① : (항공료\*출장인원/체류기간)  
 \*\*비용② : (체재비(1일)\*출장인원)

시험 1건당 평균비용이 해외가 2,861,000원, 국내가 274,000원으로 시험 1건당 비용 총 2,587,000원이 절감되는 것으로 분석되었으며, 해외시험에 필요한 출장인원은 3.4명, 출장에 필요한 평균인원을 기준으로 1인당 1일 체재비는 290,000원의 비용이 소요되며, 국내시험에 필요한 출장인원은 2.6명으로, 출장에 필요한 평균인력을 기준으로 1인당 1일에 체재비는 100,000원의 비용이 소요될 것으로 산출되었다.

3GPP 국제표준 규격을 기반으로 제품의 디바이스의 특성에 맞게 제품에 대한 5G 성능검증항목을 기반으로 5G 성능분석 및 트래픽감시와 모니터링을 실시하여, 시험검증항목에 대한 신뢰도를 확보하였다.

2023년 12개월간 월평균 이용건수는 25건, 최대 이용 건수는 253건으로 분석되었으며, 107개사 2,682건의 이용건수를 기반으로 약 69억원의 비용절감액이 산출되었으며, 기업의 평균 비용절감효과는 90.4%로 나타났다.

표 3. 전체 이용 건수의 비용절감액 산출  
 Table 3. Calculate cost savings for the total number of uses

구분	결과값
전국 기업의 비용절감액(해외비용 - 국내비용) 평균	2,587,508원
기간 내 전체 이용 건수	2,682건
비용절감액(비용절감액 평균 * 이용 건수)	6,939,695,879원

본 논문에서는 2023년 5G 시험망 테스트베드를 이용한 기업들의 해외대비 국내 비용의 차액을 통한 비용절감액을 산출하였으며, 연간 약 69억원의 기업지원 효과가 있었으며, 기업 당(107개 기업) 평균 약 64백만원 경제적인 비용절감효과가 나타났음을 알 수 있었다.

또한 비용절감액에 생산유발계수를 곱하여 생산유발효과를 산출할수 있었으며, 이는 69억원의 비용절감액이 다른 상품이나 서비스의 최종수요가 되었을 경우 발생하는 효과이며, 산업연구원 통계서비스에서 제공하는 생산유발계수(2019년) 1.51을 적용하면 약 104억원의 생산유발효과를 나타내며, 이는 비용절감액을 다른 소비재나 서비스에 투자했을 경우 경제 전체적으로 약 104억원이 간접적으로 절감되었다고 해석이 가능하다.[1]

### III. 결 론

본 논문에서는 5G 시험망 테스트베드를 이용함으로써 국내 중소기업들의 비용절감을 통해 제품 신뢰성 확보, 개발기간 단축, 경쟁력 강화 등 5G 융합제품 개발기업들의 글로벌 경쟁력 강화에 많은 도움이 되며 시장에 선도적으로 진출 할 수 있을 것으로 판단된다.

### 참 고 문 헌

[1] 산업연구원, 주요산업동향지표, p201, June 2023.

# GPT 모델을 활용한 Dose-Response 생성 모델 연구

고재현\*, 조유진\*, 정지용\*\*, 정용민\*\*, 최진희\*\*, 최호식\*\*\*

## Dose-response generation model study using GPT model

JaeHyun Go\*, YuJin Jo\*, JiYong Jeong\*\*, YongMin Jeong\*\*, JinHee Choi\*\*, and HoSik Choi\*\*\*

산업발달 및 화학물질 수요 증가에 따라 다양한 신물질 개발이 이루어지고 있으며, 신물질을 구성하는 화학 물질들의 안전성 문제에 대한 우려가 증가하고 있다. 최근 연구들은 ToxCast 데이터베이스가 화학물질의 기전 기반 독성 평가와 화학물질의 우선순위 설정에 기여하고 있다. 그러나 화학물질에 대한 용량-반응 곡선 (Dose-Response Curve, DR-curve) 데이터가 여전히 부족하며, 이는 위해성 평가와 규제 결정에 한계로 작용하고 있다. 본 연구에서는 GPT 모델을 미세 조정(fine-tuning)하여 7,000개 이상의 화합물에 대해 활성 여부를 분류하고, 활성 물질에 대한 최대 활성의 50%에 도달하는 농도를 나타내는 Active Concentration 50(AC50) 값을 예측하는 새로운 접근 방식을 제안한다. 이 연구는 Tox21 프로그램의 TOX21\_DT40\_100 Assay를 중심으로 수행되며, Ku70 및 Rad54 유전자 결핍이 있는 닭 립프구 세포에서 ATP 양을 측정하여 세포 생존율을 평가한다. 이를 통해 화학물질이 DNA 손상 및 세포 주기에 미치는 영향을 파악할 수 있다. 또한 화합물의 SMILES 형식과 SlogP, SMR, LabuteASA, TPSA 등 주요 분자 설명자 데이터를 사용하여 모델을 훈련하였으며, 모델은 화합물이 활성 (Active), 비활성(Inactive)인지 분류하고 활성 물질에 대해서는 AC50 값을 예측한다. 미세 조정된 GPT 모델은 활성/비활성 분류 작업과 AC50 예측 작업 모두에서 높은 정확도를 달성했다. 이 연구는 AI 기술을 활용하여 화학물질의 독성을 예측하고 스크리닝하는 효율적인 도구를 제시하며, 화학물질 노출이 생물학적 시스템에 미치는 영향을 예측할 수 있을 것으로 기대한다.

### Key words

Large Language models, GPT, Toxcast, Dose-Response, Active Concentration 50

### 참고문헌

- [1] Chen, X., Roberts, R., Tong, W., & Liu, Z. (2022). Tox-GAN: an artificial intelligence approach alternative to animal studies—a case study with toxicogenomics. *Toxicological Sciences*, 186(2), 242-259.
- [2] Dix, D. J., Houck, K. A., Martin, M. T., Richard, A. M., Setzer, R. W., & Kavlock, R. J. (2007). The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicological sciences*, 95(1), 5-12.
- [3] Fujii, Y., Narita, T., Tice, R. R., Takeda, S., & Yamada, R. (2015). Isotonic Regression Based-Method in Quantitative High-Throughput Screenings for Genotoxicity. *Dose-Response*, 13(1).

---

\* 서울시립대학교 도시빅데이터융합학과, ko9805@uos.ac.kr

\*\* 서울시립대학교 환경공학부 jinhchoi@uos.ac.kr

\*\*\* 서울시립대학교 도시빅데이터융합학과 교수, choi.hosik@uos.ac.kr(교신저자)

※ 이 논문은 환경부 생활화학제품 안전관리 개발사업(R&D)의 지원을 받아 진행된 논문입니다 (RS-2023-00215309).

# 회전 기계 베어링의 진동 신호를 활용한 조건 분류 및 이상 탐지

송우진\*, 최승환\*, 이수웅\*

## Condition Classification and Anomaly Detection Using Vibration Signals of Rotating Machine Bearings

Ujin Song\*, Seung-Hwan Choi and Suwoong Lee\*

본 논문은 회전 기계 베어링의 진동 데이터의 정상데이터를 학습하고 운영 조건을 분류하여 이상 상태를 탐지하는 이상탐지 방법을 제안한다. 회전 기계의 고장 원인 중 큰 비율을 차지하는 베어링 결함을 조기에 감지하는 것은 산업 현장에서의 신뢰성과 유지보수 효율을 높이는 중요한 과제이다. 제안된 이상탐지 방법은 FEMTO-TI 베어링 진동 데이터를 활용한다[1]. 부하와 속도에 따른 데이터의 각 정상 데이터를 추출하여 1D CNN으로 학습한다. 새로운 데이터를 입력하였을 때 해당 데이터에 맞는 조건을 찾아 Autoencoder를 통해 이상 상태를 탐지한다[2,3]. 조건 분류에 사용되는 데이터는 FFT, Power Spectrum 등 주파수 기반의 진동 데이터의 특징을 추출하고 이상탐지에 사용되는 데이터는 RMS, Peak to Peak 등 시계열 기반으로 추출된다. 이를 기반으로 효율적인 데이터 분류와 이상 탐지를 수행한다. 본 시스템은 조건별 정상 데이터를 학습함으로써, 미지의 데이터라도 운영 조건을 분석하여 더 효율적으로 이상 탐지를 수행할 수 있도록 한다.

### Key words

Classification, Anomaly detection, Convolution neural network, Autoencoder

### 참고문헌

- [1] Nectoux, P. et al., "PRONOSTIA: An Experimental Platform for Bearings Accelerated Life Test," IEEE PHM 2012 Conference, Denver, CO, 2012.
- [2] Ige, A. O., & Sibiyi, M. "State-of-the-Art in 1D Convolutional Neural Networks: A Survey." IEEE Access, vol. 12, pp. 144082-144103, July 2024. DOI: 10.1109/ACCESS.2024.3433513.
- [3] Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X., "Autoencoder-Based Condition Monitoring and Anomaly Detection Method for Rotating Machines," IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 3, pp. 774-783, 2017.

---

\* 한국생산기술연구원, oojin1@kitech.re.kr, csw1496@kitech.re.kr, lee@kitech.re.kr(교신저자)

※ 본 논문은 한국생산기술연구원 기관주요사업 "제품 제조현장 작업파트너 로봇 기술 개발(5/6) (kitech EO-24-0007)"의 지원으로 수행한 연구입니다. 또한, 이 연구는 2024년도 산업통상자원부 및 한국산업기술평가관리원(KEIT) 연구비 지원에 의한 연구입니다('RS-2024-00422721').

# DIFAR 소노부이와 해상초계기 간의 효율적 통신을 위한 인공지능 기반 음향신호 압축 모델 개발

박연진\*, 박광현\*, 배근영\*, 홍정표\*\*

## Development of an AI-Based Acoustic Signal Compression Model for Efficient Communication between DIFAR Sonobuoys and Maritime Patrol Aircraft

Yeon-Jin Park\*, Kwang-Hyun Park\*, Geun-Young Bae\*, and Jung-Pyo Hong\*\*

본 연구는 Directional Frequency Analysis and Recording (DIFAR) 소노부이와 해상초계기 간의 통신 효율성을 증진하기 위한 인공지능 기반 음향신호 압축 기법 개발에 관한 것이다. DIFAR 소노부이는 3채널로 수집된 수중 음향 신호를 수상 부표를 통해 무선통신으로 해상초계기에 전송하며, 해상초계기에서 신호를 분석한다. 제한된 무선통신 대역폭으로 인해 Frequency Division Multiplexing (FDM) 방식을 활용해 3채널 신호를 1채널로 변조 전송하지만, 주파수 대역에서 쉽게 식별되어 전송 효율성과 보안성에 한계가 있다. 이를 극복하기 위해 오토인코더 구조를 적용하여 DIFAR 소노부이의 무선통신 효율을 증가시키는 연구가 진행되었다[1]. 오토인코더는 인코더, 잠재 벡터, 디코더로 구성되며, 인코더를 통해 입력 신호를 저차원 잠재 벡터로 압축하고 디코더를 통해 입력 신호와 유사한 형태로 복원한다. 또한, 압축된 잠재 벡터는 동일 디코더가 아닐 시 신호 복원이 불가능하다. 따라서 소노부이에 인코더를, 해상초계기에 디코더를 탑재하여 잠재 벡터를 전송하면 신호 전송 효율성과 보안성이 향상된다. 본 연구에서는 기존 DIFAR 소노부이의 오토인코더 방식[1]을 개선하고자 Residual Vector Quantization (RVQ)과 Compensation module (CM)[2] 기법을 적용한 새로운 모델을 개발하였다. RVQ 기법은 잠재 벡터를 양자화해 데이터 압축률과 전송 효율을 높이지만, 양자화 오류로 복원 신호 품질이 저하될 수 있다. 이를 보완하기 위해 CM 기법은 두 개의 얇은 합성곱신경망 층을 활용해 양자화 오류를 보완하고 복원 신호의 유사도를 증가시킨다. 이러한 개선을 객관적으로 평가하기 위해 Continuous wave (CW)와 Linear frequency modulation (LFM) 신호의 복원 실험 결과, 복원 유사도를 나타내는 스펙트럼 Mean Squared Error (MSE)가 기존 방법[1] 대비 CW, LFM이 각각 약 60.65%, 55.25% 감소함을 확인하였다.

### Key words

Acoustic signal compression, Autoencoder, DIFAR Sonobuoy, Residual vector quantization

### 참고문헌

- [1] J. Park, J. Seok, J. Hong, "Autoencoder-Based Signal Modulation and Demodulation Methods for Sonobuoy Signal Transmission and Reception", *Sensor* 22, no. 17, 6510, August 2022.
- [2] L. Xu, J. Wang, J. Zhang and X. Xie, "LightCodec: A High Fidelity Neural Audio Codec with Low Computation Complexity," *ICASSP IEEE 2024*, pp. 586-590, April 2024

\* 국립장원대학교 대학원 정보통신공학과 석사과정, 20247090@gs.cwnu.ac.kr

\*\* 국립장원대학교 대학원 정보통신공학과 부교수, hansin@changwon.ac.kr, 교신저자

※ 이 논문은 2024학년도 국립장원대학교 중견교원 연구 지원 사업 연구비에 의하여 연구되었음.

# 딥러닝을 이용한 발매 초기 음원 순위 예측

김민제\*, 최상민\*\*

## Prediction of Initial Music Rankings Using Deep Learning

Minje Kim\* and Sang-Min Choi\*\*

### 요 약

음반산업의 꾸준한 성장과 더불어 음악 스트리밍 시장도 함께 커져가고 있다. 이에 따라 수많은 음원이 출시되고, 출시를 위해서는 많은 경제적 비용이 발생할 수 있다. 만약 음원이 발매하기 전, 음원의 흥행 여부를 예측할 수 있다면 음원 발매 전략에 도움이 될 수 있을 것이다. 그러므로 본 연구는 발매 전 얻을 수 있는 가수명, 소속사명을 포함한 총 5개의 변수를 통해 발매 초기의 순위를 예측하는 실험을 진행하였다. 해당 변수들은 Bidirectional Encoder Representations from Transformer(BERT) tokenizer를 통해 토큰화를 진행하고 고차원 벡터로 임베딩한 뒤, Multi Layered Perceptron(MLP), Convolutional Neural Network(CNN), 어텐션 기반 트랜스포머 인코더를 적용하여 예측 결과를 도출하였다. 실험을 통해 어텐션 기반의 트랜스포머 인코더 모델이 가장 좋은 성능을 보여줌을 확인할 수 있고, 또한 음원 초기 순위의 범위를 다양하게 실험하였을 때는 5주 간의 순위를 예측했을 때의 성능이 가장 좋음을 알 수 있다.

### Abstract

With the continuous growth of the music industry, the music streaming market has expanded as well. As a result, a vast number of songs are released, and significant economic costs can be incurred in the process. If the success of a song could be predicted before its release, it would aid in formulating release strategies. Therefore, this study conducts an experiment to predict the initial ranking of a song using five variables, including the name of artist and agency name, which can be obtained prior to release. These variables are tokenized using the Bidirectional Encoder Representations from Transformers (BERT) tokenizer and embedded into high-dimensional vectors. Then, the results are derived by applying a Multi-Layered Perceptron (MLP), a Convolutional Neural Network (CNN), and an attention-based Transformer encoder. Through the experiment, it was confirmed that the attention-based Transformer encoder model exhibited the best performance. Furthermore, when experimenting with varying prediction timeframes of initial music ranking, we find that the model achieved the highest performance when predicting rankings over a five-week period.

### Key words

time series prediction, deep learning, music analysis

---

경상국립대학교 컴퓨터공학과, \*Email(knj5091q@gmail.com), \*\*교신저자 Email(jerassi@gnu.ac.kr)

※ 본 논문은 교육부와 한국연구재단의 재원으로 지원 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0), 정부(과학기술정보통신부) 재원의 한국연구재단 지원(RS-2022-00165785)과 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다(2021RIS-003).

## I. 서론

국제음반산업협회(The International Federation of the Phonographic Industry, IFPI)에 따르면 2023년 음반산업의 규모는 286억 달러에 달한다. 이는 전년도 대비 10%의 성장한 수치이며, 2014년 이후부터 꾸준히 성장을 보여주고 있다. 또한 스트리밍 시장의 경우에도 꾸준히 성장하여 2023년에는 전년도 대비 10.2%가량 증가한 193억 달러 규모를 가지고 있다[1]. 스트리밍 시장의 성장으로 수많은 음원들이 발매되고 있으며, 이러한 음원들의 발매는 흥행을 위해 마케팅 등에 많은 경제적인 비용을 수반한다. 음원을 발매하기 전, 수요 예측을 통해서 성과를 예측할 수 있다면 음원 발매 전략에 고려될 수 있는 요소로 사용할 수 있다.

음원의 흥행을 예측하는 연구로 이도연 외[2]는 가수 영향력, 제작사 역량 등 17개의 독립변인을 통해 음원의 흥행성적을 예측하는 형태의 딥러닝 구조를 제안하였다. 또한 김양석[3]은 흥행요인을 정량화할 수 있도록 다양한 요소를 분석하고 검증하여 흥행요인으로 정의하였다. 위 연구들을 토대로 본 연구에서는 선별한 가수명, 곡명, 앨범명, 배급사명, 소속사명, 총 5개의 독립변수 데이터를 학습하여 음원 발매 후 주간별 순위를 예측하는 형태의 딥러닝 구조를 제안한다. 적합한 모델을 찾기 위해 다양한 딥러닝 기법을 적용하여 성능을 비교 실험하였다.

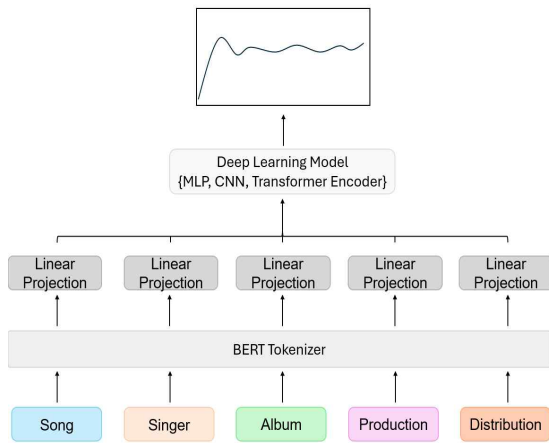


그림 1. 딥러닝 모델 구조  
Fig. 1. Deep Learning Model Structure

## II. 예측 방법론

본 연구에서 제안한 딥러닝 구조는 그림 1과 같다. 먼저, Bert Tokenizer[4]를 통해서 딥러닝 모델이 학습할 수 있도록 가수명, 곡명, 앨범명, 배급사명, 소속사명을 토큰화한 뒤에 선형 투영을 통해 잠재 공간으로 임베딩하고, 임베딩된 잠재공간을 MLP[5], CNN[6], 어텐션[7] 기반의 모델로 학습하여 초기 음원의 순위 시계열을 예측한다.

MLP의 경우, 임베딩된 5개의 독립변수를 concatenate 하고 이를 128차원, 32차원인 두 개의 은닉층을 거쳐 최종적으로 N개의 초기 음원 순위 시계열을 도출하게 된다. 어텐션 기반의 모델은 트랜스포머 인코더만을 사용하였고, 모델 차원은 64차원, 인코더는 2개를 사용하였다. 그리고 CNN은 임베딩된 값을 2개의 CNN 1D의 입력으로 사용하여 연산한 뒤, 3개의 Linear 층을 사용하여 예측값을 도출하였다.

표 1. 데이터 기초통계량

Table 1. Statistics of Dataset

구분	Contents
데이터 수집 기간	2009. 12. 27~ 2024. 9. 15
데이터 개수	11,611곡
순위 데이터 길이 평균	6.22
순위 데이터 길이 표준편차	12.42
순위 데이터 길이 최소값	1
순위 데이터 길이 중간값	2
순위 데이터 길이 최대값	340

## III. 실험

### 3.1 데이터셋

본 연구에서는 써클차트(Circle Chart)의 2010년 1주 차부터 2024년 38주 차까지의 주별 데이터 106,800개를 이용하였고, 실험을 위해 각 음원마다 차트인한 시점부터 처음으로 차트아웃한 시점까지의 순위 데이터를 추출하였다. 전처리한 순위 데이터 길이의 기초통계량은 표 1과 같다. 순위 데이터의 경우, 실험을 위해 sliding window 형태로 전처리하고, 이때 window의 크기는 표 1을 참고하여 1부터 5까지 변경하며 sliding window를 적용하였다.

### 3.2. 구현

pytorch를 이용하여 딥러닝 모델을 구현하였으며, 모델 간 성능 비교를 위해서 Mean Squared Error(MSE)를 손실함수로 사용하였다. MSE는 순위 데이터  $X = x_1, x_2, \dots, x_N$ 에서 실제 데이터의 값  $x_i$ 와 예측한 데이터의 값  $\hat{x}_i$  간 차이의 제곱을 평균한 것이다. 학습 데이터와 검증 데이터, 테스트 데이터의 비율은 6:2:2로 하였고, 최적화 알고리즘은 Adam을 사용하였다. 그리고 learning rate는 0.001 이고 모든 실험은 하나의 NVIDIA Quadro RTX 5000 16GB GPU 하에서 5번 수행하고 평균을 내어 결과를 도출하였다.

표 2. 실험 결과  
Table 2. Experiment Result

Models	N	MSE	MAE
MLP [5]	1	1012.9 ± 0.0	27.4 ± 0
	2	950.4 ± 0.0	25.5 ± 0
	3	878.7 ± 0.0	23.6 ± 0
	4	883.7 ± 0.0	22.8 ± 0
	5	1019.3 ± 0.0	24.3 ± 0
CNN [6]	1	999.7 ± 0.0	26.9 ± 0.0
	2	917.0 ± 0.0	25.2 ± 0.0
	3	799.1 ± 0.0	23 ± 0.0
	4	782.3 ± 0.0	22.4 ± 0.0
	5	832.9 ± 0.0	22.6 ± 0.0
Transformer Encoder [7]	1	<b>923.8 ± 1.4</b>	<b>26.4 ± 0.0</b>
	2	<b>837.1 ± 3.6</b>	<b>24.7 ± 0.0</b>
	3	<b>723.9 ± 1.5</b>	<b>22.4 ± 0.0</b>
	4	<b>660.5 ± 3.7</b>	<b>21.1 ± 0.1</b>
	5	<b>669.2 ± 1.5</b>	<b>21.0 ± 0.0</b>

### 3.3. 실험 결과 및 분석

실험결과는 MSE와 Mean Absolute Error(MAE)를 사용하여 성능을 측정하였다. 표 2의 실험 결과를 보면, 먼저 트랜스포머 인코더 모델이 가장 좋은 성능을 보여주었음을 확인할 수 있다. 또한 다양한 window의 크기로 실험하였을 때는, 대체적으로 window 크기가 클수록 좋은 성능을 보임을 확인할 수 있는데, 이는 모델이 음원의 정보를 통해 특정 순위를 예측하는 것보다 발매 초기의 순위의 흐름을 예측하는 것이 더 좋은 성능을 보이는 것으로

해석된다. 하지만 적은 독립 변수의 수로 인해 정확한 예측은 하지 못하는 것으로 판단된다.

## IV. 결론

본 연구에서는 음원의 흥행 여부를 판단하기 위해 발매 전에 알 수 있는 정보들로 발매 초기 시계열을 딥러닝을 통해 예측하는 연구를 수행하였다. 실험 결과로 트랜스포머 인코더를 사용했을 때와 발매 초기 5주 간의 순위를 예측했을 때 가장 좋은 성능을 보여주었다. 비교 실험과 별개로, 해당 실험 결과는 실제로 사용할 수 있을 정도의 성능을 끌어내지는 못하였지만, 향후 딥러닝을 이용한 음원 분석 연구를 통해 성공적인 음원 발매를 위한 전략을 수립하는 데 있어서 보조적인 지표로 활용될 수 있을 것을 기대한다.

## 참고 문헌

- [1] IFPI. (2024) "Global Music Report 2024", <https://globalmusicreport.ifpi.org/>
- [2] 이도연, 장병희 (2020) 딥러닝을 이용한 음악 흥행 예측모델 개발 연구. 한국콘텐츠학회 논문지, 20(8), 10-18.
- [3] 김양석. "음악콘텐츠의 흥행요인이 K-Pop의 흥행성과에 미치는 영향." 국내박사학위논문 호서대학교 벤처대학원, 2016. 충청남도
- [4] Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
- [5] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65(6), 386-408. <https://doi.org/10.1037/h0042519>
- [6] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
- [7] Vaswani, A. (2017). Attention is all you need. Advances in Neural Information Processing Systems.



# 얼굴 파싱을 위한 로짓 분포 변환 기반 지식 증류

배지호\*, 이수원\*\*

## Logit Distribution Transformation Based Knowledge Distillation For Face Parsing

Jiho Bae\* and Suwon Lee\*\*

### 요약

얼굴 조작과 같은 응용 분야에서 얼굴 파싱(face parsing)은 핵심적인 역할을 수행하며, 에지(edge) 디바이스나 실시간 추론을 위해 모델의 경량화가 절실하다. 그러나 기존 지식 증류(knowledge distillation) 기법이 얼굴 파싱에 직접 적용된 사례는 드물었다. 본 연구에서는 얼굴 파싱 태스크에 지식 증류를 적용하여 경량화된 모델의 성능을 향상시키고자 한다. 한편, 기존 로짓(logit) 기반의 지식 증류에서는 학생 모델과 교사 모델이 동일한 온도(temperature)를 사용하였다. 하지만 최근 연구에서는 공유된 온도가 학습에 문제가 될 수 있음을 지적하였다. 이를 해결하기 위한 방법으로 로짓 표준화(standardization)가 제안되었으나, 본 연구에서 얼굴 파싱에 로짓 표준화를 적용하는 과정에서 여러 한계가 발견되었다. 이러한 한계를 극복하기 위해, 우리는 새로운 로짓 분포 변환 기법을 제안한다. 제안된 방법은 CelebA-Mask-HQ 데이터셋을 활용한 실험을 통해 지식 증류의 성능을 유의미하게 향상시켰음을 보여준다.

### Abstract

Face parsing plays a pivotal role in applications such as facial manipulation, where model compression is critical for deployment on edge devices or real-time inference. However, the application of knowledge distillation (KD) techniques to face parsing tasks has been relatively rare. This study aims to enhance the performance of compressed models by applying knowledge distillation to face parsing. Traditional logit-based knowledge distillation methods typically use the same temperature for both the student and teacher models. Recent research, however, has pointed out that using a shared temperature can cause issues during training. While logit standardization has been proposed as a solution, several limitations were encountered when applying it to face parsing in this study. To address these challenges, we propose a novel logit distribution transformation method. Experimental results using the CelebA-Mask-HQ dataset demonstrate that the proposed method significantly improves the performance of knowledge distillation for face parsing.

### Key words

face parsing, knowledge distillation, logit standardization

---

\* 경상국립대학교 컴퓨터공학과, dream\_cacao\_jh@naver.com

\*\* 경상국립대학교 컴퓨터공학과, leesuwon@gnu.ac.kr (교신저자)

## 1. 서론

얼굴 파싱(face parsing)은 얼굴 이미지를 분석하여 눈, 코, 입, 귀, 헤어라인 등 얼굴의 세부 부위를 픽셀 단위로 분할하는 기술이다. 이는 얼굴의 각 부위를 정확하게 인식하고 구분하는 작업으로, 얼굴 조작, 증강 현실, 보안 시스템, 의료 영상 분석 등 다양한 응용 분야에서 필수적인 역할을 한다. 특히, 얼굴 조작이나 변형과 같은 분야에서는 얼굴의 각 부분을 정밀하게 분할해야만 자연스럽게 일관된 결과를 얻을 수 있기 때문에 얼굴 파싱의 중요성은 더욱 부각된다.

최근 몇 년간, 딥러닝 기반의 얼굴 파싱 모델이 도입되면서 그 성능은 괄목할 만한 발전을 이루었다. 그러나 고성능 모델은 대개 대규모의 연산 자원을 필요로 하며, 이는 에지 디바이스나 실시간 응용 환경에서는 여전히 중요한 문제로 남아 있다. 이와 같은 환경에서는 모델의 경량화가 필수적이며, 실시간 추론 성능을 유지하면서도 높은 정확도를 제공하는 모델이 요구된다.

경량화된 모델의 성능을 보완하기 위한 방법으로 지식 증류(knowledge distillation)는 효과적인 대안으로 주목받고 있다 [1]. 지식 증류는 고성능의 교사 모델(teacher model)로부터 학생 모델(student model)에 지식을 전달함으로써, 경량화된 학생 모델의 성능을 극대화하는 기법이다. 그러나 기존의 지식 증류 기법이 얼굴 파싱 테스트에 직접 적용된 사례는 매우 드물다.

로짓(logit)이란 분류 문제에서 소프트맥스 이전의 모델 최종 아웃풋을 말한다. 최근 연구에서 기존의 로짓 기반 지식 증류에서 교사와 학생 모델이 동일한 온도(temperature)를 사용하는 방식은 학습 과정에서 문제가 될 수 있음이 밝혀졌다[2]. 이로 인해 최근 연구에서는 로짓 표준화(standardization) [2]와 같은 기법이 제안되었지만, 우리의 실험 결과 얼굴 파싱에 이를 적용하는 과정에서 여러 한계가 발견되었다. 이러한 한계를 극복하기 위해 본 연구에서는 로짓 분포 변환(logit distribution transformation) 기반의 새로운 지식 증류기법을 제안한다. 제안된 방법은 교사 모델과 학생 모델 간의 적절한  $\tau$ 를 조정하면서 픽셀 간 분포 또한 잘 전파할 수 있다.

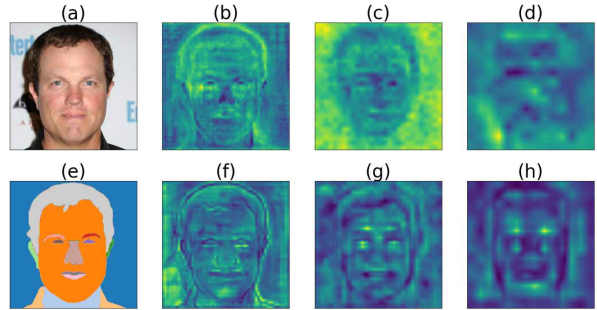


그림 1. 여러 모델에서 픽셀별 로짓의 평균과 표준편차; (a) 입력 이미지, (b) Resnet101 + FCN 모델 로짓 평균, (c) Mobilenetv3 + Deeplabv3 모델 로짓 평균, (d) Resnet18 + FCN 모델 로짓 평균, (e) 출력 레이블, (f) Resnet101 + FCN 모델 로짓 표준편차, (g) Mobilenetv4 + Deeplabv3 모델 로짓 표준편차, (h) Resnet18 + FCN 모델 로짓 표준편차

Fig. 1. Mean and standard deviation of per-pixel logit across multiple models

CelebA-Mask-HQ 데이터셋을 활용한 실험을 통해, 우리는 제안된 기법이 기존 방법에 비해 경량화된 모델의 정확도를 유의미하게 향상 시켰음을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 로짓 기반의 지식증류에 대하여 설명하고, 기존 방법인 로짓 표준화를 적용하고 분석한다. 3장에서는 제안하는 로짓 분포 변환을 설명한다. 4장에서는 기존 학습방법과 지식 증류 기반 학습방법, 제안된 학습방법을 비교하며, 5장에서 결론 맺는다.

## II. 로짓 기반 지식 증류

### 2.1 로짓 기반 지식 증류 방법

지식 증류(knowledge distillation)는 힌튼(Hinton)[1] 등이 제안한 방법으로, 큰 모델(교사 모델, teacher model)의 지식을 작은 모델(학생 모델, student model)에 전달하여 경량화된 모델의 성능을 향상시키는 기법이다. 힌튼의 지식 증류 방법에서는 교사 모델의 로짓 값을 학생 모델이 학습하게 하는데, 이때 온도값을 활용하여 로짓을 부드럽게(smooth) 만들어 준다. 소프트맥스 함수는 온도 값이 클수록 예측 확률 분포를 평탄하게 만들어 다양한 클래스 간의 미세한 차이를 더 잘 반영할 수 있게 한다. 이는 학생 모델이 단순히 정답만 학습하는 것이 아니라, 교사 모델이 예측한 다양한 클래스 간의 확률 분포를 이

해하고 이를 통해 더욱 일반화된 성능을 학습할 수 있게 돕는다. 식 (1), (2)는 지식증류 과정의 일부이다.  $z_i^t$ 와  $z_i^s$ 는 각각 교사 모델과 학생 모델의 로짓으로 소프트맥스 함수와 온도를 통해 부드럽게 만들어준다. 이후 그 결과인  $q_i, p_i$ 를 Kullback - Leibler divergence loss를 통해 학생 모델이 교사 모델과 유사한 분포를 가지도록 학습시킨다.

$$q_i = \frac{\exp(z_i^t/\tau)}{\sum_i \exp(z_i^t/\tau)} \quad (1)$$

$$p_i = \frac{\exp(z_i^s/\tau)}{\sum_i \exp(z_i^s/\tau)} \quad (2)$$

## 2.2 공유 온도 문제와 로짓 표준화

최근 연구에서는 학생 모델과 교사 모델이 동일한 온도를 사용하는 것을 문제로 제기하였다[2]. 모델의 크기와 용량 차이를 고려하면 학생모델이 교사 모델만큼 넓은 범위를 생성하지 못할 수 있다 [3]. 또한 공유 온도는 학생 모델의 표준편차를 제한한다.

이를 보완하기 위해서 [2]에서는 로짓 표준화를 사용하여 학생 모델과 교사 모델에 표준편차에 따른  $\tau$ 를 설정하여 지식 증류를 원활하게 한다.

## 2.3 얼굴 파싱에서의 로짓 표준화

$$\bar{X} = \frac{1}{K} \sum_{k=1}^K X^{(k)} \quad (3)$$

$$\sigma(X) = \sqrt{\frac{1}{K} \sum_{k=1}^K (X^{(k)} - \bar{X})^2} \quad (4)$$

$$Z(X;\tau) = (X - \bar{X})/\sigma(X)/\tau \quad (5)$$

학생 모델과 교사 모델의 공유  $\tau$ 로 인한 문제를 방지하고자 로짓 표준화가 제안되었다[2]. 식 (3), (4) 는 각각 평균과 표준편차를 의미하고, 식 (5)는 표준화 과정을 보여준다. 식  $X$ 는 모델의 최종 아웃풋인 로짓을 의미한다.  $\tau$ 는 온도를 나타낸다.

이러한 표준화 과정을 학생 모델과 교사 모델의

로짓에 적용한 후, 소프트맥스 함수에 넣어줘 지식 증류를 진행한다. 이러한 로짓 표준화를 얼굴 파싱에서 적용해 보았다.

표 1. 기존 방법의 기존 증류 성능 (F1-score)  
Table 1. Performance of the existing methods for knowledge distillation (F1-score)

미적용	vanilla-KD	KD-STD
83.97	85.37	83.51

얼굴 파싱 테스트에서는 모델의 출력이 세그먼트된 결과가 나오기 때문에 모델의 최종 아웃풋은 픽셀별 로짓이 된다. 모델의 픽셀별 로짓의 평균과 표준편차를 구하고 이를 통해서 로짓 표준화를 진행하였다. Resnet101+FCN을 교사 모델로, MobileNetV4+DeepLabV3을 학생모델로 하여 기본 학습 방법, 기본적인 지식 증류, 로짓 표준화를 위한 지식 증류 방법 3가지 케이스에 대해서 학습하였다. 표 1을 보면 지식 증류를 적용하지 않는 방법 보다 지식 증류를 실행한 vanilla-KD 방법이 더 성능이 높은 것을 알 수 있다. 반면, 로짓 표준화를 적용한 방법은 오히려 성능이 감소하였다

로짓 표준화의 성능 감소 원인을 분석하기 위해 여러 모델에서 픽셀별 로짓의 평균과 표준편차를 시각화하였다. 그림 1의 (b), (f)와 같이 Resnet101+FCN과 같이 성능이 좋은 모델은 로짓의 픽셀 별 평균과 표준편차에서 출력 레이블(label)과 비슷한 분포를 뚜렷하게 볼 수 있다. 반면, 성능이 낮은 모델인 Resnet18+FCN모델은 흐릿한 분포를 보여준다. 이를 통해서 교사 모델의 픽셀 간 로짓 분포가 중요한 정보를 가지고 있다는 것을 알 수 있다. 그러나 교사 모델과 학생 모델의 로짓에 표준화를 진행하면 유의미한 정보인 교사 모델의 픽셀 간 로짓 분포 차이가 학생 모델에 전달되지 않는다.

## III. 제안 방법

그림 2를 보면 전통적인 지식 증류 방법, 로짓 표준화 지식 증류 방법, 우리가 제안한 로짓 분포 변환 지식 증류 방법을 비교한 그림을 볼 수 있다.

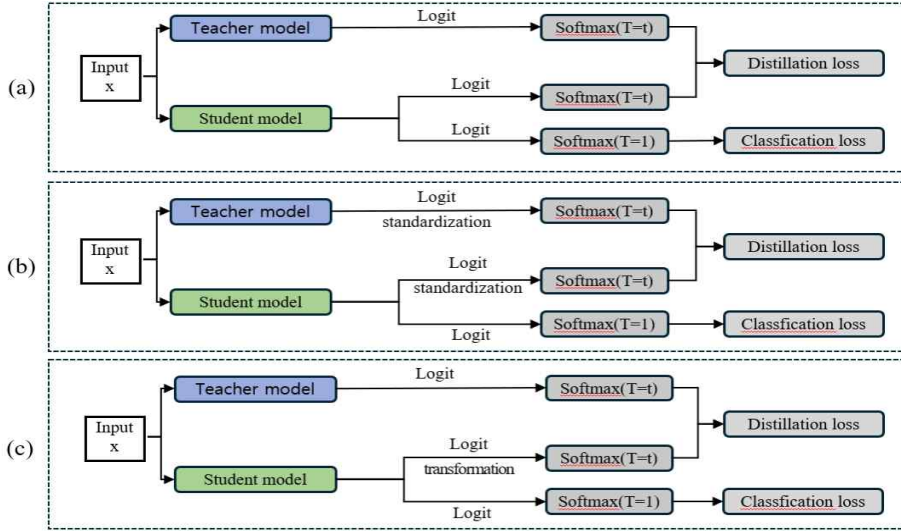


그림 2. 로짓 기반 지식 증류 방법 비교; (a) 초기 제안된 기본적인 지식 증류 방법, (b) 로짓 표준화를 적용한 지식 증류 방법, (c) 제한한 로짓 분포 변환 방법

Fig. 2. Comparing logit-based knowledge distillation methods

그림 2(a)는 기본적인 지식 증류 방법이고, 그림 2 (b)는 2장에서 설명한 로짓 표준화를 적용한 지식 증류 방법이다. 로짓 표준화 방식은 학생 모델과 교사 모델 모두에 표준화를 적용하는데, 이러한 방식은 모든 픽셀별 로짓을 평균이 0이고 표준편차가 1인 분포로 만들기 때문에 교사 모델의 픽셀 간 분포를 학생 모델에 전달하지 못한다.

$$T(X_s, \tau) = ((X_s - \bar{X}_s) / \sigma(X_s) \times \sigma(X_t) + \bar{X}_t) / \tau \quad (6)$$

우리는 이러한 점을 보완해 픽셀 간 로짓 분포를 학생 모델에 잘 전달할 수 있도록 교사 모델은 그대로 두고, 학생 모델만 로짓 분포 변환을 하여 적절한 온도 값을 찾는다.  $X_s$ ,  $X_t$  는 각각 학생 모델과 교사모델의 로짓 이다. 식 (6)과 같이  $X_s$ 을 표준화 한 뒤  $X_t$ 의 분포로 변환해준다. 이러한 과정을 통해서 [2]에서 제안한 학생 모델과 교사 모델 사이에 존재하는 적절한 온도를 비율을 유지하면서 교사모델의 픽셀 간 분포 정보를 손실 없이 전달할 수 있다.

#### IV. 실험

##### 4.1 데이터셋 및 평가지표

얼굴 파싱 문제에서 지식 증류의 효과를 검증하기 위해 CelebAMask-HQ 데이터셋 [4] 을 사용하였다. 본 데이터셋은 학습, 검증, 테스트로 분류되며, 각각 24,183개, 2,993개, 2,824개로 구성되어 있다. 하나의 데이터는 다양한 성별, 인종의 얼굴 데이터로 구성되어 있으며 얼굴 데이터에서 부분별로 19개의 클래스로 나눈 마스크가 레이블 역할을 한다. 평가지표는 Accuracy와 F1-score를 사용하였다.

##### 4.2 구현 상세

Resnet101+FCN모델을 교사 모델로 하고 Mobile-Net-v3+Deeplabv3, Resnet34+FCN, Resnet18+FCN 세 개의 모델을 학생 모델로 선정하여 학습하였다. 모든 모델에 대해 배치 사이즈는 5, 학습률은 0.01, 옵티마이저(optimizer)는 SGD를 사용하여 100 에폭 학습하였다.

##### 4.3 실험결과

표 2와 표 3은 지식 증류를 적용하지 않은 방법, 전통적인 방식의 지식 증류(vanilla-KD), 로짓 표준화를 통한 지식 증류(KD-STD)와 우리가 제안한 방법을 비교하였다.

표 2. 지식 증류 성능 비교 (F1-score)

Table 2. Knowledge distillation performance comparison (F1-score)

학생 모델	지식 증류 방법			
	미적용	vanilla-KD	KD-STD	제안방법
Resnet101+FCN	87.1	-	-	-
Deeplabv3_mobilenet	83.97	<b>85.37</b>	83.51	<b>85.37</b>
Resnet34+FCN	73.78	74.56	66.94	<b>75.42</b>
Resnet18+FCN	73.41	74.51	67.3	<b>75.01</b>

표 3. 지식 증류 성능 비교 (Accuracy)

Table 3. Knowledge distillation performance comparison (Accuracy)

학생 모델	지식 증류 방법			
	미적용	vanilla-KD	KD-STD	제안방법
Resnet101+FCN	95.04	-	-	-
Deeplabv3_mobilenet	94.34	94.72	94.32	<b>94.77</b>
Resnet34+FCN	91.89	92.72	91.48	<b>92.76</b>
Resnet18+FCN	91.56	92.53	91.15	<b>92.56</b>

우선 지식 증류를 사용한 방법(vanilla-KD)이 모든 모델과 평가 지표에서 지식 증류를 적용하지 않은 방법 보다 높은 성능을 보여주었다. 이것은 로짓 기반의 지식증류 방법의 효과를 보여준다. 로짓 표준화를 적용한 KD-STD 방식은 모든 지식 증류 모델보다 성능이 낮고, 심지어 지식 증류를 하지 않는 모델보다도 낮은 성능을 보여주었다. 이것은 2.3장에서 언급한 교사 모델의 픽셀 간 정보를 학생 모델이 정상적으로 전달받지 못해서 생긴 문제로 보인다. 우리가 제안한 방법은 vanilla-KD의 단점인 공유 온도와 KD-STD 단점인 교사 모델의 픽셀 간 정보 전달 손실 문제를 보완하여, 대부분의 모델과 평가지표에서 vanilla-KD보다 뛰어난 성능을 보여준다.

## V. 결론

본 논문에서 얼굴 파싱 분야에 지식 증류를 적용

하였다. 전통적인 지식 증류 방식의 단점인 공유 온도를 보완하기 위해, 제안된 로짓 표준화 방식을 얼굴 파싱에 적용하고, 성능 저하 원인을 분석했으며 이를 보완한 로짓 분포 변환 방법을 제안하였다. CelebAMask-HQ 데이터셋을 통한 실험을 통해 제안된 방법의 성능을 입증하였다. 향후 다양한 로짓 기반의 지식 증류 방법에 제안한 방법을 적용한다면 제안된 방법의 추가적인 적용 가능성을 입증할 수 있을 것이다.

## 참고 문헌

- [1] Hinton, Geoffrey. "Distilling the Knowledge in a Neural Network." arXiv preprint arXiv:1503.02531 (2015).
- [2] Sun, Shangquan, et al. "Logit standardization in knowledge distillation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.
- [3] Mirzadeh, Seyed Iman, et al. "Improved knowledge distillation via teacher assistant." Proceedings of the AAAI conference on artificial intelligence. Vol. 34. No. 04. 2020.
- [4] Lee, Cheng-Han, et al. "Maskgan: Towards diverse and interactive facial image manipulation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.

# 피처 중요도 기반 점진적 피처 필터링 방법의 실험적 평가

하재경\*, 오세종\*\*

## Experimental evaluation of gradual feature filtering method based on feature importance

Jaekyoung Ha\* and Sejong Oh\*\*

### 요 약

Permutation Importance는 머신러닝 모델에서 피처의 중요도를 평가하는 주요 척도이다. 본 연구에서는 Permutation Importance에 따라 중요도가 낮은 피처를 일괄 삭제하는 Permutation Filtering(PF)과 이를 개선한 점진적 삭제 방법인 Gradual Permutation Filtering(GPF)의 성능을 비교 분석하였다. 총 5개의 데이터셋을 대상으로 피처 제거 비율을 10%, 20%로 하여 필터링을 수행 한 후 모델의 예측 정확도를 비교하였다. 실험 결과 10개의 케이스중 6개에서 GPF가 더 높은 성능을 보였다. 이는 일괄 삭제 방법에 비해 점진적 삭제 방법이 더 나은 방법임을 입증한다.

### Abstract

Permutation Importance is a key measure to evaluate the importance of features in machine learning models. In this study, we compared and analyzed the performance of Permutation Filtering (PF), which batch deletes low-importance features according to permutation importance, and Gradual Permutation Filtering (GPF), an improved progressive deletion method. After filtering a total of 5 datasets with feature removal rates of 10% and 20%, the prediction accuracy of the models was compared. As a result of the experiment, GPF showed higher performance in 6 out of 10 cases. This proves that the gradual deletion method is a better method than the batch deletion method.

### Key words

feature selection, permutaion importance, gradual permutation importance, high dimensional dataset

## 1. 서 론

머신러닝 모델에서 피처의 선택은 모델 성능과 해석 가능성에 큰 영향을 미친다. 적절한 피처 선택은 모델의 성능을 향상시키고, 계산 비용을 줄이며,

차원의 저주로 인한 문제점을 해결할 수 있다. 하지만 어떤 피처가 중요한지를 평가하고 선택하는 과정은 쉽지 않다. 이를 위해 여러 가지 변수 중요도 평가 기법이 제안되었다. 그중 Permutation Importance는 간결하고, 널리 사용되는 방법이다.

\* 단국대학교 인공지능융합학부 인공지능융합학과, hajaekyung99@dankook.ac.kr,

\*\* 단국대학교 sw융합대학 소프트웨어학과, sejongoh@dankook.ac.kr (교신저자)

※ 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2023-00222191, 분산된 데이터에 대한 논리적 데이터 통합과 복합분석을 지원하는 데이터 패브릭 기술개발)

Permutation Importance는 학습된 모델에서 각 피처의 값을 임의로 섞은 후, 모델의 예측 성능 변화를 측정하는 것으로 해당 피처의 중요도를 평가한다. 이 계산된 중요도를 기준으로 임계값을 통해 중요도가 낮은 피처를 한 번에 제거하는 방식 (Permutation Filtering, PF)으로 피처 선택을 수행한다. 하지만 Permutation Filtering 방식은 실제보다 중요도가 과소평가 되거나 과대평가 될 수 있다.

이러한 문제를 해결하기 위해 Shin [1]은 박사학위 논문을 통해 점진적 제거 방식 (Gradual Permutation Filtering, GPF) 방법을 제안하였다. GPF는 중요도가 낮은 피처들을 한 번에 제거하는 대신, Permutation Importance 측정과 피처 제거를 반복하며 점진적으로 중요도가 낮은 피처를 제거하는 방식이다. 그러나 상기 논문에는 제안된 방법이 얼마나 효과가 있는지에 대한 검증 결과는 제시되지 않았다.

본 연구에서는 GPF의 효과성을 검증하기 위해 5개의 서로 다른 데이터셋을 대상으로 PF와 GPF의 성능을 비교 분석하였다.

## II. 기본 개념

### 2.1 Permutation Filtering

PF는 각 피처의 값을 임의로 섞어 모델의 예측 성능이 얼마나 떨어지는지를 평가하는 것으로 피처의 중요도를 측정하는 방법이다. 이 방법은 1회의 모델 학습으로 중요도를 측정할 수 있으며, 간단하게 피처가 모델에 기여하는 정도를 파악할 수 있다. 하지만 값을 임의로 섞는 방식으로 인해 정확한 평가를 위해 충분한 반복이 필요하다. PF는 계산된 Permutation Importance를 기준으로 임계값 아래의 피처를 한 번에 일괄 제거하는 방식이다.

### 2.2 Gradual Permutation Filtering

GPF는 PF의 단 한 번의 순위 결정으로 피처를 제거하는 방식을 개선한 방법이다. Permutation Importance를 이용하여 순위를 결정하고, 중요도가 낮은 변수를 제거하고, 다시 모델을 학습시켜

Permutation Importance를 계산하는 것을 반복하며 점진적으로 불필요한 피처들을 제거해 나가는 방법이다. 이 방법은 모델을 다시 학습시켜야 하며, Permutation Importance를 반복해서 계산하는 만큼 PF보다 상대적으로 긴 수행 시간이 필요하나, 점진적으로 불필요한 피처를 제거하는 것으로 더 정밀한 피처 선택이 가능하다.

Shin [1]이 제안한 방법은 임계 횟수에 대한 threshold를 증가시키는 방향으로 점진적으로 제거하는 방식을 제안했으나 본 논문에서는 임계치가 아닌 고정적인 비율에 따라 제거할 피처 수를 고정해 두고 반복마다 점진적으로 제거가 이루어지도록 일부 수정했다.

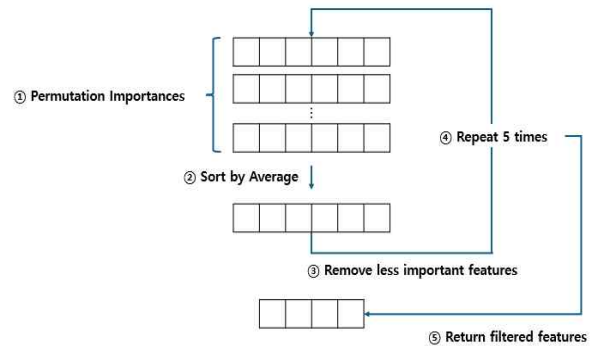


그림 1. GPFi 실행 절차

Fig. 1. Diagram of Gradual Permutation Filtering

## III. 실험 절차

본 연구에서는 5개의 서로 다른 데이터셋을 사용하여 PF와 GPF 방법의 성능을 비교하였다. 데이터셋은 피처 선택이 중요한 고차원 데이터셋을 위주로 선정하였다. 모델은 XGBoostClassifier를 사용하여 학습을 진행하였으며, Permutation은 각 피처 당 10번 반복하여 평균 중요도를 도출하였다. 모델의 성능 평가는 AUC를 통해 평가를 수행하였다. AUC는 클래스 불균형의 영향을 적게 받는 척도이다. 또한 피처 제거율에 따른 성능 변화를 관측하기 위해 피처 제거율이 10%일 때와 20%일 때로 나누어 두 방법의 성능을 비교하였다. 또한 차이율을 통해 두 방법론이 제거한 피처가 얼마나 다른가 비교하였다. 차이율은 식(1)을 이용해 산출한다.

표 1. 실험 데이터 세트

Table 1. Experimental data sets

데이터셋	Instances	Features
alon	62	2,000
gravier	168	2,905
Hill_valley	1,212	100
madelon	2,600	500
qsar_androgen_receptor	1,687	1,024

$$1 - \frac{\text{동일하게 제거한 피처수}}{\text{제거한 전체 피처수}} \quad (1)$$

#### IV. 실험 결과

총 5개의 고차원 데이터를 대상으로 피처 제거율을 10%와 20%로 나누어 실험을 수행한 결과 총 10 번 중 6번의 실험에서 GPF가 더 좋은 성능을 보였다. 하지만 20%를 제거했을 때 총 5개의 데이터셋 중 4개의 데이터셋에서 PF보다 높은 성능을 보이는 것을 확인할 수 있었다. 이는 피처 제거율이 너무 낮은 경우는 GPF의 효과가 미미하며, 적정 제거율인 경우 GPF의 효과가 증대됨을 보여준다. GPF는 반복적으로 Permutation Importance를 계산하며 정밀하게 피처를 제거하는 특성으로 인해 다수의 피처를 제거할 때 PF보다 더 나은 성능을 보이는 것으로 판단된다. 또한 차이율이 가장 낮은 모습을 보이는 상황이 madelon 데이터셋에서 피처 제거율이 20%일 때인데 이때도 0.2673으로 총 100개의 피처를 제거할 때 약 27개의 피처가 다르게 선택되었음을 의미한다. 이는 피처가 제거될 때마다 피처 간 상관관계로 인해 각 피처의 중요도가 변동되었고, GPF는 이를 고려하여 피처를 제거했음을 알 수 있다.

#### V. 한계 및 향후 연구

본 연구는 제한된 수의 데이터셋을 대상으로 실험을 수행하였으며, 다양한 도메인에서의 성능 차이를 확인하지 못했다는 한계가 있다. 또한 GPF 방식의 계산 비용이 기존 PF 방식에 비해 상대적으로 높다는 점도 고려해야 한다. 향후 연구에서는 GPF의 효율을 극대화 하는 피처 제거율에 대한 연구와

다양한 도메인에서 GPF의 적용 가능성을 검토할 예정이다.

표 2. 실험 결과

Table 2. Experiment result

데이터셋	피처 제거율	10%		20%	
		AUC	차이율	AUC	차이율
alon	PF	0.5902	0.3532	0.6069	0.2544
	GPF	<b>0.6236</b>		<b>0.6486</b>	
gravier	PF	<b>0.511</b>	0.5069	0.4587	0.5439
	GPF	0.5044		<b>0.4936</b>	
Hill_valley	PF	<b>0.5609</b>	0.5454	0.5594	0.6667
	GPF	0.5		<b>0.5623</b>	
madelon	PF	<b>0.8065</b>	0.4554	<b>0.7983</b>	0.2673
	GPF	0.8008		0.7959	
qsar	PF	0.71	0.6601	0.7198	0.3689
	GPF	<b>0.7129</b>		<b>0.7218</b>	

#### VI. 결론

본 연구에서는 PF와 GPF의 성능을 비교하고, GPF가 PF보다 더 나은 성능을 보인다는 점을 실험적으로 입증하였다. 이는 향후 변수 중요도 평가 연구에 중요한 시사점을 제공한다.

#### 참고 문헌

- [1] Shin, "An effective heuristic feature selection method for high-dimensional low-sample size datasets", doctoral thesis, Dankook University, 2024.
- [2] Li, Jundong, et al. "Feature selection: A data perspective." ACM computing surveys (CSUR) 50.6 (2017): 1-45.
- [3] Miao, Jianyu, and Lingfeng Niu. "A survey on feature selection." Procedia computer science 91 (2016): 919-926.
- [4] Chandrashekar, Girish, and Ferat Sahin. "A survey on feature selection methods." Computers & electrical engineering 40.1 (2014): 16-28.
- [5] Venkatesh, B., and J. Anuradha. "A review of feature selection and its methods." Cybernetics and information technologies 19.1 (2019): 3-26.



# Capturing Temporal Dependencies for Gait Classification using LSTM-Informer and Window Sliding Technique

Okpala Chibuike\*, Pyunghwa Park\*, Wonhyong Lee\*, and Daseong Han\*

LSTM-인포머와 윈도우 슬라이딩에 기반한 보행 분류를 위한 시간 종속성 캡처 기법

옥팔라 치부이케\*, 박평화\*, 이원형\*, 한다성\*

## Abstract

We propose a hybrid LSTM-Informer architecture for classifying gait patterns. The LSTM layers capture short-term dependencies effectively, while the Informer encoder, with its ProbSparse self-attention mechanism, effectively models long-range dependencies. This combination allows the model to handle complex sequential data, ensuring accurate temporal patterns recognition in gait sequences. The proposed model achieves an accuracy of 99.36%, across six gait classes, including one normal and five pathological type. This result demonstrates the effectiveness of combining LSTM and Informer architecture in handling time-series data.

## 요약

본 논문은 걸음걸이 패턴을 분류하기 위한 하이브리드 LSTM-인포머 아키텍처를 제안한다. LSTM 계층은 단기적 종속성을 잘 포착하는 반면, Informer 인코더는 ProbSparse self-attention 메커니즘을 통해 장기적 종속성을 효과적으로 인식한다. 이 둘을 조합하는 하이브리드 모델은 보행 시퀀스에서 시간적 패턴을 정확하게 인식하여 복잡한 시계열 데이터에 대한 처리를 가능하게 한다. 제안된 모델은 정상 보행 유형 1개와 병적 보행 유형 5개를 포함한 6개의 보행 유형에 대해 99.36%의 정확도를 달성한다. 이 결과는 시계열 데이터를 처리할 때 LSTM과 Informer 아키텍처를 결합하는 것이 효과적임을 보여준다.

## Key words

Informer; LSTM; Gait Pattern.

## 1. Introduction

Gait analysis is commonly used in clinical and rehabilitation settings to measure the movement

patterns that make up walking and to provide interpretations that assist in the rehabilitation of patients with various mobility impairments due to disease, aging and injuries [1].

---

\* Dept. of Human Ecology & Technology, Handong Global University, {myokpala, 22473003}@handong.ac.kr, {whlee, dshan}@handong.edu

※ This work was supported by the National Research Foundation(NRF), Korea, under project BK21 FOUR (No.5199990314060).

Jun et al. [2] proposed a GRU classifier for pathological gait classification. They considered various joint groups to identify significant and irrelevant joints for improving classification accuracy. Their study showed that an LSTM model achieved an accuracy of 87.25%, whereas the proposed GRU model reached an improved accuracy of 93.7% when only leg joints were considered. Zhou et al. [4] proposed an Informer Transformer model for Long Sequence Time-series Forecasting (LSTF), to address several key challenges associated with LSTF. Traditional methods, including LSTM and vanilla Transformer architectures, often struggle with scaling when it comes to long sequence inputs due to issues like quadratic time complexity and high memory usage. The Informer aims to overcome these limitations through three primary contributions: ProbSparse self-attention, self-attention distilling, and a generative-style decoder. The accurate classification of gait patterns is crucial for diagnosing, monitoring and rehabilitating various conditions. Traditional methods often fall short in capturing both long and short temporal dependencies present in gait data. To address these challenges, we proposed a hybrid LSTM-Informer architecture.

## II. Methodology

This methodology involves several key stages: data preprocessing, data representation, data augmentation, model training, and evaluation.

### 2.1 Dataset description and Preprocessing

The dataset used in this research is the Pathological Gait Datasets [2]. The skeletal gait was collected using the Kinect V2 sensors and the Microsoft SDK, which provides 3D xyz coordinates for 25 human joints. The dataset comprises the simulated pathological gaits (including normal, analgic, stiff-legged, lurching, steppage, and Trendelenburg) of 10

participants. Each participant simulated six different gait types, resulting in 120 instances per gait type. In total, the dataset contains: 10 participants, 6 gait types (including normal and 5 pathological gaits), 120 instances per gait type. The 3D coordinate data of 25 joints from a total of six Kinect V2 devices were compared and combined into a single dataset using linear interpolation which can provide higher quality training opportunities for subsequent machine learning tasks.

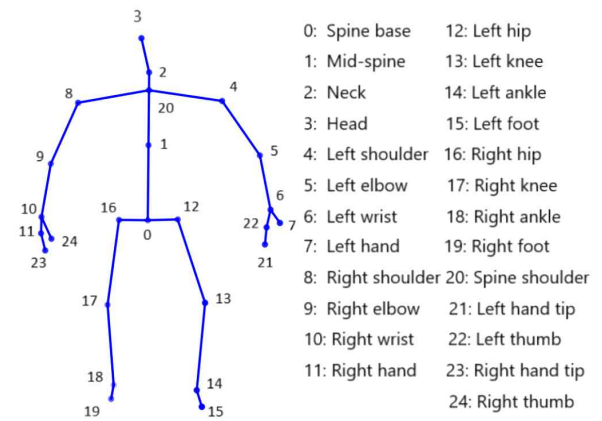


Figure 1: Skeletal structure and joint names

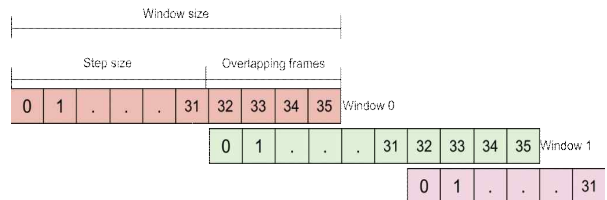


Figure 2: Window size and step size

To simulate continuous movement, a window sliding technique is employed. This technique segments continuous time-series data into overlapping windows of data points, with each window treated as a separate input to the model. We use a window size of 36 frames, which represents one gait cycle duration in the dataset, and a step size of 32, creating an overlap of 4 frames between successive windows (Figure 2). After creating the windows, each window is normalized to center the data around a reference joint, removing global positional variation and focusing on the relative movements of the joints.

## 2.2 Model Architecture

In this study, we propose an LSTM-Informer hybrid architecture presented in Figure 3 to specifically handle the complexities of sequential gait data by integrating the strengths of Long Short-Term Memory (LSTM) networks [3] and the Informer Transformer model [4]. The proposed model aims to efficiently capture both short-term and long-term dependencies, providing a comprehensive approach for analyzing and classifying gait patterns. The proposed architecture starts with a sequence of LSTM layers consisting of 512 units, which are well-known for their ability to process and retain temporal information in sequences. Following the LSTM layers, the output is fed into an embedding layer. This layer projects the sequence data into a lower-dimensional space, with an embedding dimension set to 256. To maintain the temporal context of the input sequence, a positional encoding layer is applied. The positional encoding provides the model with information about the order and relative positions of each time step in the sequence. The core component of the architecture is the Informer encoder, which comprises multiple Informer encoder blocks designed to model long-range dependencies within the sequence efficiently. The architecture includes six Informer encoder blocks, which stack multiple layers of self-attention and feed-forward networks to progressively refine and extract essential features from the input sequence. Each encoder block incorporates a ProbSparse self-attention mechanism, which optimizes the traditional self-attention mechanism by sparsely selecting attention weights. In addition to the self-attention mechanism, each Informer encoder block includes a feed-forward network composed of two dense layers. The first layer in the network has 512 units and uses a ReLU activation function. To stabilize the training process and improve generalization, each encoder block incorporates layer normalization and drop out layers. A distillation

operation is applied between the encoder blocks to reduce the sequence length. The distillation layer down-samples the sequence by a factor of 2, effectively shortening the sequence length while retaining the most relevant features. After processing through the Informer encoder, a global average pooling layer is applied. The pooled output is then passed through a dense layer with 128 units and a ReLU activation function. This dense layer acts as an intermediate processing step, refining the extracted features before they are fed into the final classification layer.

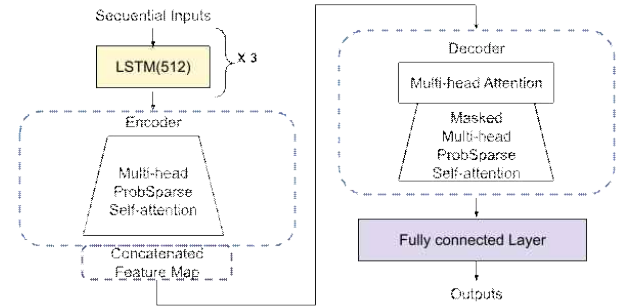


Figure 3: Proposed LSTM-Informer hybrid model architecture

## III. Results

Table 1: Comparison of proposed model variants on the classification of gait

Model	Accuracy(%)
LSTM-Informer	99.360
GRU-Informer	99.162
Informer	98.324

The results presented in Table 1 show a comparative analysis of different model variants based on accuracy. The proposed LSTM-Informer model outperforms the other variants, achieving the highest accuracy of 99.360%. In comparison, the GRU-Informer model also performs well, but slightly under performs relative to the LSTM-Informer variant, with an accuracy of 99.162%. Lastly, the stand-alone Informer model achieves an accuracy of 98.324%. The results highlight the advantage of incorporating

recurrent layers like LSTM and GRU for sequence modeling and also emphasize the effectiveness of hybrid models in improving performance over the base Informer architecture.

Table 2: Grouping of joints

Joint	Accuracy	Loss
All_joints	99.360	0.0258
All_joint_w/o Group 1	98.57	0.0491
All_joint_w/o Group 2	98.97	0.0372
All_joint_w/o Group 3	99.26	0.0425
All_joint_w/o Group 4	31.69	1.6582
All_joint_w/o Group 5	98.57	0.0593
All_joint_w/o Group 6	94.28	0.1627

The results presented in Table 2 demonstrate the impact of various joint groups on the accuracy of gait pattern classification. The proposed model, when all joints are used, achieves the highest accuracy of 99.360%, establishing an optimal performance benchmark. When Group 1 (head and neck joints) is removed, the model still performs relatively well, showing an accuracy of 98.57%. Similarly, removing Groups 2 (shoulder, elbow, and wrist) and 3 (hands, thumbs, and hand tip joints) has minimal impact, with the model maintaining high accuracies of 98.97% and 99.26%, respectively. However, the removal of Group 4 (spine-shoulder, mid-spine, and spine-base joints) leads to a substantial performance drop, with accuracy plummeting to 31.69%. On the other hand, removing Group 5 (hip and knee joints) and Group 6 (ankle and foot joints) results in moderate declines in performance, particularly for Group 6, where the accuracy drops to 94.28%. These results suggest that while spine-related joints are the most crucial, joints related to the lower body joints also play a significant role in accurate gait pattern classification.

Table 3: Comparison with state-of-the-art models

Authors	Model	Accuracy (%)
Jun et al. [2]	GRU	93.67
Jun et al. [5]	CNN-RNN encoder hybrid model	97.60
Jun et al. [6]	ST-GCN	99.03
our model	LSTM-Informer	99.36

Table 3 presents a comparison of our model's performance against state-of-the-art models in terms of accuracy. Jun et al. (2020) reported an accuracy of 93.67% with a GRU model, which was subsequently improved in 2021 with a CNN-RNN encoder hybrid model achieving 97.60%. In 2023, Jun et al. revisited the GRU model and achieved a slightly higher accuracy of 95.83%, alongside testing an LSTM model with an accuracy of 95.20%, and an ST-GCN model with 99.03%. Our model, LSTM-Informer architecture, significantly outperforms these previous models with an accuracy of 99.36%.

#### IV. Conclusion

We classify one normal gait class and five abnormal gait classes (antalgic, lurching, steppage, stiff-legged, and Trendelenburg) using an LSTM-Informer architecture. We employ a window sliding technique to represent the data, preserving its sequential and temporal nature, with each window representing one gait cycle. We achieved an accuracy of 99.36%. Despite achieving remarkable results, our work has certain limitations. We did not consider the measure of gait severity, which is crucial for monitoring and rehabilitation settings for injured or diseased individuals.

#### Reference

- [1] Baker, R. (2006). Gait analysis methods in rehabilitation. *Journal of Neuroengineering and Rehabilitation*,3(1). <https://doi.org/10.1186/1743-0003-3-4>
- [2] Jun, K., Lee, Y., Lee, S., Lee, D., & Kim, M. S. (2020). Pathological Gait Classification Using Kinect v2 and Gated Recurrent Neural Networks. *IEEE Access*, 8,139881-139891. <https://doi.org/10.1109/access.2020.3013029>
- [3] Hochreiter, S., & Schmidhuber, J. (1997). Long

- Short-Term memory. *Neural Computation*, 9(8),1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [4] Zhou, H., Zhang,S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2020). Informer:Beyond efficient transformer for long sequence Time-Series forecasting. arXiv(Cornell University). <https://doi.org/10.48550/arxiv.2012.07436>
- [5] Jun, K., Lee, S.,Lee, D., & Kim, M. S. (2021). Deep Learning-Based Multimodal Abnormal GAITclassification using a 3D skeleton and plantar foot pressure. *IEEE Access*, 9,161576-161589. <https://doi.org/10.1109/access.2021.3131613>
- [6] Jun, K., Lee, K.,Lee, S., Lee, H., & Kim, M. S. (2023). Hybrid Deep Neural network frameworkcombining skeleton and GAIT features for pathological GAIT recognition.*Bioengineering*, 10(10), 1133. <https://doi.org/10.3390/bioengineering10101133>

# RhyGEN: 다양한 박자를 고려한 CNN-LSTM 기반 리듬게임 패턴 악보 제작 모델

하현후, 윤상연, 이진희, 부석준\*

## RhyGEN: A CNN-LSTM-based Rhythm Game Chart Generation Model Considering Diverse Tempo

Hyun-Hoo Ha, Sang-Yeon Yoon, Jin-Hee Lee, and Seok-Jun Buu\*

### 요 약

리듬게임의 패턴 제작은 음원의 다양한 음악적 요소를 고려해야 하며 음원의 난해한 박자와 리듬에 대응되는 패턴을 제작하기 위하여 상당한 시간이 요구된다. 이를 해결하기 위해 본 논문에서는 컨볼루션-장단기 메모리 (CNN-LSTM) 기반 딥러닝 모델을 통해 BMS(Be-Music Script) 포맷에 대응하는 패턴 악보를 자동 제작하는 방법을 제안한다. 제안하는 방법은 마디 단위로 샘플링된 스펙트로그램 데이터셋을 기반으로 학습하므로, 특수 박자 및 박자의 변화에서 타 모델 대비 유연한 대응이 가능하다. 이는 제2 연구를 통하여 본 방법의 우수성을 실험적으로 입증하였으며 구조 유사도 지수 비교 실험에서 타 모델 대비 0.2%p의 유의미한 성능개선을 달성하였다. RhyGEN의 시연 영상은 <https://aisalab.gnu.ac.kr/ai-applications/rhygen>에서 확인 가능하다.

### Abstract

Creating rhythm game patterns requires considerable time, as it involves accounting for various musical elements and developing patterns that match complex beats and rhythms. This paper proposes a CNN-LSTM-based deep learning model for automatic pattern generation in the BMS (Be-Music Script) format. Trained on measure-level spectrogram data, the model adapts flexibly to unique time signatures and tempo changes, outperforming other models in versatility. Experimental validation through ablation studies shows a 0.2 percentage-point improvement in Structural Similarity Index (SSIM) compared to similar models. A demonstration video of RhyGEN is available at <https://aisalab.gnu.ac.kr/ai-applications/rhygen>.

### Key words

long short-term memory, convolution network, artificial neural network, rhythm game, be-music script

### 1. 서 론

리듬게임의 패턴은 음원을 기반으로 음정, 박자,

리듬감 등 다양한 부분이 고려되어 제작되며 구성된 리듬게임의 패턴 악보를 차트(chart)라 한다. 리듬게임 음원은 특수박자, 변박자, 엇박 등이 포함된

경상국립대학교 컴퓨터공학과, wldyd423@gmail.com, dmsqkf@gnu.ac.kr, gsm04548@gnu.ac.kr, sj.buu@gnu.ac.kr\*

※ 감사의글: 이 논문은 항공 우주 협동가 조합의 지원을 받아 작성되었음.

경우가 많아, 차트 제작의 난이도를 높이며 많은 시간적 자원을 필요로 한다. 이러한 문제를 해결하기 위하여 본 연구에서는 음원 입력 시, 리듬게임 스크립트 BMS(Be-Music Script) 기반 리듬게임 차트를 자동으로 제작하는 딥러닝 모델 RhyGEN을 제안한다. RhyGEN은 기존 노래들과 차트를 마디 단위로 샘플링 후 STFT(Short Time Fourier Transform)으로 전처리하며, 컨볼루션-장단기 메모리 신경망(CNN-LSTM) 모델 구조로 학습된다. 학습된 모델은 마디 단위로 구성된 차트를 출력하며, 출력 시간은 한 곡당 1분 이내로 기존 인간 대비 압도적으로 제작 속도를 개선한다. 또한 마디 단위 패턴 데이터를 학습하므로 특수 박자 및 변박자 등 타 모델에서 다루기 어려운 리듬 요소들에 대하여 유연하게 대응한다. 실험 결과, RhyGEN은 구조 유사도 지수, 코사인 유사도, MSE 등에서 기존 모델 대비 유의미한 성능 향상을 보였다.

## II. 관련 연구

딥러닝 모델을 응용하여 리듬 게임에 접목한 연구들을 표 1과 같이 정리하였다. DDC(Dance Dance Convolution) 모델은 컨볼루션 신경망(CNN)과 장단기 메모리 신경망(LSTM)을 결합하여 음악 구조와 일치하는 DDR(Dance Dance Revolution) 게임의 차트(악보)를 제작한다[1]. BMS 기반 리듬게임 차트를 제작하기 위하여 키 사운드 추정방식 개선하는 클러스터링 접근방식을 사용하며 SMO (Sequential Minimal Optimization) 학습 알고리즘을 사용한다[2]. GenerationMania는 음악의 각 오디오 샘플을 플레이어가 직접 연주해야 할 소리와 자동으로 재생될 배경음을 분류 후 LSTM, CNN 구조를 통하여 BMS를 기반으로 차트를 제작한다[3]. GenéLive! 모델은 리듬 액션 게임 “Love Live!”의 차트를 자동 제작하기 위해 고안된 딥러닝 모델로, CNN과 Bi-LSTM 구조를 토대로 멜 스펙트로그램 (mel-spectrogram) 및 비트 정보를 입력으로 받는다[4].

RhyGEN은 마디 기반의 데이터를 전처리하여 CNN-LSTM 구조로 학습하므로 다른 모델에서 고려하지 않은 특수 박자, 변박자에서 대응이 가능하다.

표 1. 리듬게임 악보 제작 관련 연구 조사  
Table 1. Research on Rhythm Game Chart Generation

Game	Signal representation	Modeling method	Dataset	Performance		Year
				F1	Acc.	
DDR	Raw waveform	LSTM, CNN	Fraxtil, ITG	0.69	-	2017[1]
Beatmania	Log Energy, F0, HNR ...	SMO	BMS	-	77.57	2019[2]
Beatmania	STFT	LSTM, CNN	BMS	0.78	-	2021[3]
Love Live!	STFT + Mel spectrogram	Bi-LSTM, CNN	In game Data	0.80	-	2023[4]

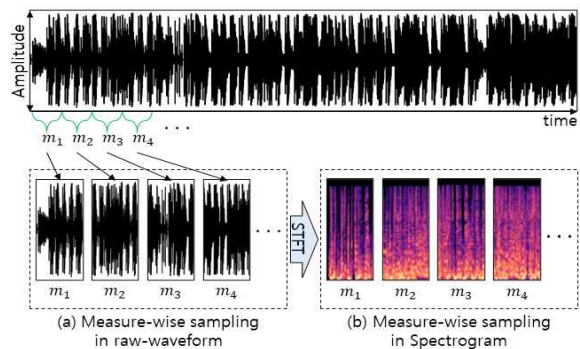


그림 1. 마디 단위 스펙트로그램의 전처리 시퀀스  
Fig. 1. Preprocessing of Measure-level Spectrograms

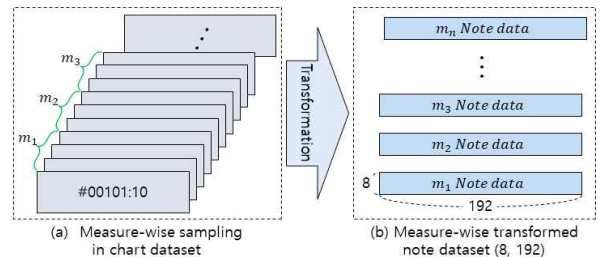


그림 2. 마디 단위 패턴정보의 전처리 시퀀스  
Fig. 2. Preprocessing of Measure-level Pattern Information

### 3.1. 데이터 전처리 및 학습, 타겟 데이터 설정

제안하는 방법에서는 BMS 데이터를 헤더 필드와 데이터 필드로 구분하여 전처리를 진행하였다. 헤더 필드에는 곡의 정보, BPM, 차트에 필요한 이미지 및 사운드 파일이 문자열로 매핑되어 있으며, 각 필드는 곡의 메타데이터를 포함한다. 데이터 필드는 콜론(:)을 기준으로 두 개의 구간으로 나뉘어 있다. 첫 번째 구간을 ‘헤더 구간’이라 칭하며, 이 구간은 오브젝트가 배치될 위치 정보를 담고 있다.

그림 1과 같이 음원 파일은 BPM을 기준으로 마디별로 분할한 후 STFT를 거쳐 스펙트로그램으로 변환 및 저장하여 학습 데이터셋 X를 구성하였다.

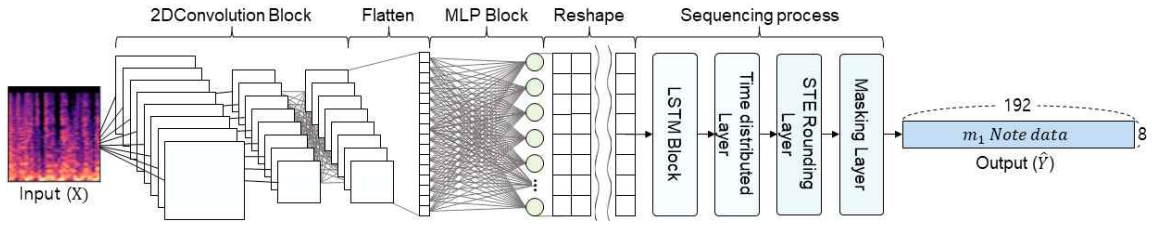


그림 3. CNN-LSTM 기반 리듬게임 차트 제작 모델 구조

Fig. 3. CNN-LSTM-based Rhythm Game Chart Generation Model Architecture

그림 2과 같이 마디의 기준은 다양한 박자와 리듬을 고려하기 위하여 최대 박자를 분할할 수 있는 길이인 192를 차원의 길이로 설정하였으며, 데이터 구간의 빈 공간은 시퀀스의 끝을 나타내는 -1로 패딩하였다. 그리고 가변길이에 해당하는 라인 데이터를 모아 하나의 마디의 해당하는 노트 정보데이터로 구성하여 이를 타겟 데이터 Y로 설정하였다.

제안하는 데이터셋 구성 방법은 특수 박자, 변하는 박자에 대응이 가능해지므로 일반적인 데이터셋 구성 방법 대비 다양한 장르, 스타일의 음원에 대처할 수 있다.

### 3.2. CNN-LSTM 기반 차트 제작 신경망

제안하는 RhyGEN 모델은 그림 3과 같이 마디 단위로 샘플링된 스펙트로그램 데이터를 입력으로 받아 컨볼루션-장단기 메모리 신경망(CNN-LSTM) 구조를 통해 2차원 차트 데이터를 생성하는 방식으로 설계되었다. 먼저 컨볼루션 신경망은 음악의 주파수 특징을 추출하는 단계로, 입력 데이터의 패턴과 리듬에 대한 저차원 표현을 효과적으로 학습한다. 이 단계는 지역적 특성을 잘 포착하여 다양한 리듬 패턴을 정확히 분류하고 파악할 수 있도록 한다. 이후 추출된 특징 정보는 장단기 메모리 신경망 레이어로 전달된다. 이는 시간적 순서를 처리함에 따라 곡의 리듬 변화와 특수 박자 등의 연속적 특징을 파악하여, 보다 자연스러운 차트 생성을 가능하게 한다. 이후 STE Rounding Layer를 통해 노트 값에 해당되는 정수값으로 변형 후 마스킹을 통하여 불필요한 값들을 -1로 처리하여 출력 형태 (8, 192)로 변환한다.

모델 최적화에는 평균 제곱 오차(MSE)를 손실 함수로 사용하였으며, MSE 손실 함수는 다음 수식 1과 같이 정의된다.

$$L(x, y) = \frac{1}{N} \sum_{i=1}^N (y - f(x_i; \theta))^2 \quad (1)$$

## III. 실험 및 결론

### 4.1. 데이터셋 수집 및 실험 방법

대형 BMS 창작 대회인 BOF(Booth of Fighters)의 음원을 모은 Primrose - Korea BMS Starter Pack을 전처리하여 이를 데이터셋으로 구성하였다. 실험에는 BMS 포맷의 데이터 길이에 따른 오차를 감안하여 차트 데이터를 게임 내에서 이미지화한 차트를 비교 평가하였다.

### 4.2. 다른 딥러닝 기반 방법 대비 성능 개선

표 2에서 기존의 딥러닝 기반 리듬게임 차트 제작 모델과 본 논문에서 제안한 방법을 구조 유사도 지수(SSIM), 코사인 유사도, MSE, L2-정규화를 통한 성능평가를 진행하였다. 엄밀한 실용성 평가를 위하여 비교 단위를 음원 단위로 설정하며, 2D 행렬 상에서 성능 측정을 진행하였다. 제안하는 방법은 비교 모델 DDC 모델 대비 SSIM, MSE, L2-정규화에서 최고 성능을 달성하였다.

### 4.3. CNN+LSTM 구조의 제거 연구 기반 성능 검증

표 3에서 CNN+LSTM 모델 구조의 제거 실험을 통하여 본 모델의 구조가 합당한지 검증하였다. 제안하는 모델은 SSIM, 코사인 유사도, MSE, L2-노름(L2-Norm) 등의 지표에서 타 모델 대비 최고 성능을 달성하였다. 이는 CNN+LSTM 모델 구조가 제안하는 리듬게임 차트 생성 방식에 적합하며, 마디 단위 데이터 처리에 있어 가장 효율적인 구조임을 실험적으로 시사한다.



표 2 RhyGen, DDC의 곡 별 패턴 제작 성능평가

Table 2. Performance Evaluation of Pattern Generation by Song for RhyGen and DDC

	RhyGen(ours)	DDC
SSIM	<b>0.94</b>	0.91
Cosine Similarity	0.91	<b>0.99</b>
MSE	<b>0.01</b>	0.03
L2-Norm	<b><math>4.83 \times 10^4</math></b>	<b><math>9.62 \times 10^4</math></b>

표 3 RhyGEN 모델 제거 실험

Table 3. Ablation Study of the RhyGEN

	SSIM	Cosine Similarity	MSE	L2-Norm
LSTM	0.9301	0.9300	0.0083	$4.40 \times 10^4$
Bi-LSTM	0.4779	0.3573	0.0731	$1.36 \times 10^5$
GRU	0.4800	0.3589	0.0730	$1.36 \times 10^5$
CNN+GRU	0.9610	0.9645	0.0044	$3.18 \times 10^5$
Ours				
CNN+LSTM	<b>0.9689</b>	<b>0.9699</b>	<b>0.0037</b>	<b><math>2.84 \times 10^4</math></b>

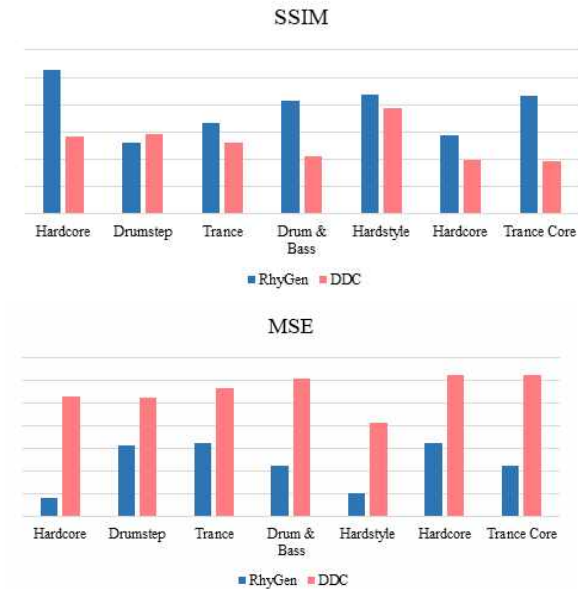


그림 4. 리듬게임 장르 별 SSIM, MSE 실험

Fig. 4. SSIM and MSE Experiments by Rhythm Game Genre

#### 4.4. 음악 장르 별 모델 성능 비교 평가

그림 4에서 음악 장르 별 모델 성능 비교 평가를 진행하였다. 실험 결과, 제안하는 RhyGEN 모델은

모든 장르에서 DDC 모델 대비 일관되게 높은 SSIM 값과 낮은 MSE 값을 기록하였다. 이는 RhyGEN 모델이 리듬게임의 다양한 장르에 대해 원본 패턴과 유사한 품질의 차트를 생성함을 의미한다.

#### 4.5. 결론 및 향후 연구

본 논문에서는 다양한 박자에 대응할 수 있는 CNN-LSTM 기반의 모델 RhyGEN을 제안한다. BMS 포맷 상에서 다양한 박자와 리듬을 고려하여 데이터셋을 구성하여 CNN-LSTM 모델을 학습시켰다. 제안하는 방법은 타 모델과의 비교, 제거 연구, 장르별 비교 실험에서 최고성능을 달성하였으며 마디로 구성된 데이터셋에 제안하는 구조의 타당성에 대하여 입증하였다. 이는 차트 제작에 시간적 자원 절감을 유도할 수 있으며, 복잡한 패턴을 가지는 리듬 게임의 박자, 리듬을 추론할 수 있다.

다만 곡의 분위기에 적절한 노트 배치는 비교적 인간에 비해 만족도가 떨어지므로 차후 연구에서는 모델 구조 개선 및 음원의 마디의 스펙트로그램을 추가 입력으로 개선할 것을 기대한다.

#### 참 고 문 헌

- [1] Donahue, C., Z. C. Lipton, and J. McAuley, "Dance dance convolution," *Int. Conf. on Machine Learning*, PMLR, pp. 1039-1048, 2017.
- [2] Fukunaga, D., K. Ochi, and Y. Obuchi, "Training Data Clustering for Key-Sound Estimation in Rhythm Action Games," *Nicograph Int.*, pp. 70-73, 2019.
- [3] Lin, Z., K. Xiao, and M. Riedl, "Generationmania: Learning to semantically choreograph," *Proc. AAAI Conf. on Artificial Intelligence and Interactive Digital Entertainment*, vol. 15, no. 1, pp. 52-58, 2019.
- [4] Takada, A., D. Yamazaki, Y. Yoshida, N. Ganbat, T. Shimotomai, N. Hamada, and D. Sakurai, "Genélive! generating rhythm actions in love live!," *Proc. AAAI Conf. on Artificial Intelligence*, vol. 37, no. 4, pp. 5266-5275, 2023

# 연속 상태 공간에서 강화 학습 정책의 군집 기반 정형 검증

권령구\*, 권기현\*, 이창율\*\*

## Clustering-Based Formal Verification of Reinforcement Learning Policy in Continuous State Space

Ryeonggu Kwon\*, Gihwon Kwon\*, and Changyul Lee\*\*

### 요약

본 연구는 연속 상태 공간에서 강화 학습 정책의 안전성을 검증하기 위한 군집화 기반 정형 검증 방법을 제안한다. KMeans++ 알고리즘을 통해 상태 공간을 군집화하고, 각 군집 간 전이 확률을 계산하여 이산 시간 마코프 체인 모델을 생성하였다. 생성된 모델은 확률 모델 체킹 도구인 PRISM을 사용해 정책의 안전 속성을 평가하며, 군집 내 전이 확률의 분포는 Kolmogorov-Smirnov 검정을 통해 통계적으로 검증하여 정확성과 효율성을 확인하였다. 실험 결과, 제안된 기법이 정책 검증의 정확성을 유지하면서도 상태 공간을 효과적으로 축소 하여서, 강화 학습의 안전성 검증에 기여하고 있다.

### Abstract

This study proposes a clustering-based formal verification method to assess the safety of reinforcement learning policies in continuous state spaces. The state space is clustered using the KMeans++ algorithm, and the transition probabilities between clusters are calculated to generate a discrete-time Markov chain model. The generated model evaluates the safety properties of the policy using PRISM, a probabilistic model-checking tool, while the distribution of transition probabilities within clusters is statistically validated for accuracy and efficiency through the Kolmogorov-Smirnov test. Experimental results demonstrate that the proposed method effectively reduces the state space while maintaining verification accuracy, contributing to the advancement of safety verification in reinforcement learning.

### Key words

formal verification, safety verification, clustering method, reinforcement learning

## 1. 서론

강화 학습(Reinforcement Learning, RL)은 자율 주행, 로봇 공학, 의료 진단 등 다양한 분야에서 최적

\* 경기대학교 컴퓨터과학과, rkkwon@kyonggi.ac.kr(교신저자), khkwon@kyonggi.ac.kr,

\*\* (주)에스피아이디, yulee@espid.com

※ 본 연구는 (주)에스피아이디의 지원을 받아 수행된 과제번호 2024-0316, "LLM 기반 컨설팅 지식 관리 및 의사결정 지원 시스템 구축"의 결과입니다.

의 행동을 학습하여 수행하는 데 활용되고 있으며, 최근 들어 그 중요성이 더욱 부각되고 있다 [1][2]. 이러한 기술은 환경과 상호작용하며 최적의 의사 결정을 내리는 능력을 통해 복잡한 문제를 해결할 수 있는 잠재력을 가지고 있다. 하지만 RL 시스템이 높은 보상을 얻는다는 것이 항상 안전한 행동을 보장하지는 않으며, 특히 연속적인 상태 공간을 가진 환경에서 정책의 안전성을 유지하는 것은 여전히 큰 과제로 남아 있다 [3][4]. 예를 들어 자율 주행 환경에서는 학습된 정책이 높은 보상에 도달할 수 있지만, 예상치 못한 도로 상황이나 급격한 장애물의 출현 등으로 인해 충돌, 경계 이탈과 같은 위험 상황이 발생할 수 있다 [5]. 이러한 이유로 강화 학습에서 정책의 성능뿐만 아니라 안전성을 보장하기 위한 체계적이고 정밀한 검증이 필수적이다 [6][7].

안전성을 확보하기 위한 기존 연구들은 주로 모델 예측 제어(Model Predictive Control, MPC)와 RL을 결합하여 정책의 안전성을 강화하려는 방식 [8], 특정 상태별 안전성을 개별적으로 평가하여 위험을 예측하는 방식 [9] 등을 제안하였다. 하지만 연속 상태 공간을 모두 검토하는 것은 실질적으로 계산 비용이 매우 크기 때문에, 모든 상태를 일일이 평가하는 것은 한계가 있다 [10]. 이와 같은 문제를 해결하기 위해 최근 연구들은 상태 공간을 압축하는 방법으로 군집화와 상태 축소 기법을 적용하여 중요한 상태만을 검토하는 방법을 시도하고 있으나, 검증의 정확성과 효율성을 동시에 확보하는 데에는 여전히 도전 과제가 존재한다. 따라서, 연속 상태 공간을 효과적으로 축소하면서도 정책의 안전성을 정밀하게 검증할 수 있는 새로운 접근법이 요구된다.

본 연구에서는 이러한 문제를 해결하기 위해 연속 상태 공간에서 강화 학습 정책의 안전성을 검증할 수 있는 군집 기반 정형 검증 기법을 제안한다. 본 논문에서 제안하는 방법은 KMeans++ 알고리즘을 통해 상태 공간을 군집화하여 주요 상태들 간의 전이 확률을 계산하고, 이를 기반으로 이산 시간 마코프 체인(Discrete-Time Markov Chain, DTMC) 모델을 구축하여 PRISM 툴을 통해 정책의 안전 속성을

정량적으로 평가한다. 이를 통해 연속 상태 공간을 효과적으로 축소하면서도 중요한 상태 전이 정보를 유지할 수 있다. 또한, 군집화된 상태의 통계적 유효성을 검토하기 위해 Kolmogorov-Smirnov (KS) 검정을 통해 군집화의 적절성을 평가하여 상태 공간 축소가 검증 정확성에 미치는 영향을 분석하였다.

논문의 구성은 다음과 같다. 2장에서는 본 연구에서 제안하는 군집 기반 정형 검증 기법의 구체적인 방법을 설명한다. 3장에서는 LunarLander 환경에서 실험을 통해 제안 기법의 검증 성능과 안전성을 평가하고, 군집 수에 따른 검증 결과의 변화와 효율성을 분석한다. 마지막으로 4장에서는 연구의 결론과 향후 연구 방향을 제시하며, 연속 상태 공간을 다루는 강화 학습 정책 검증의 적용 가능성과 확장 가능성에 대해 논의한다.

## II. 제안 방법

본 연구는 연속 상태 공간에서 강화 학습 정책의 안전성을 효과적으로 검증하기 위해 KMeans++ 알고리즘을 활용한 군집화 기반 검증 방법을 제안한다. 연속적인 상태 공간에서 개별 상태를 모두 검토하는 것은 계산 비용이 크기 때문에, 본 연구는 상태를 유사한 군집으로 묶어 이산화된 모델을 생성함으로써 검증의 효율성을 높이고자 한다.

먼저, KMeans++ 알고리즘을 사용하여 상태 공간을 군집화하였다. KMeans++ 알고리즘은 초기 중심점을 확률적으로 선택하여 군집의 균일성과 일관성을 유지하며, 이를 통해 비효율적인 군집화를 방지한다. 군집화 결과의 품질은 각 군집 내 응집도(Within-Cluster Sum of Squares, WCSS)를 통해 평가되며, 군집 수  $k$ 가 증가할 때 WCSS의 감소 속도가 급격히 완화되는 지점을 찾는 엘보 방법을 통해 최적의 군집 수  $k^*$ 를 결정하였다. 이를 수식으로 표현하면 다음과 같다:

$$WCSS = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2$$

여기서  $C_i$ 는  $i$ -번째 군집,  $\mu_i$ 는 군집  $C_i$ 의 중심

점,  $x$ 는 해당 군집의 상태를 나타낸다. 최적 군집 수  $k^*$ 는 WCSS의 감소 곡선이 완화된 엘보 포인트에서 결정된다.

군집화된 상태 공간을 바탕으로 각 군집 간 전이 확률을 계산하여 DTMC 모델을 생성하였다. DTMC 모델은 군집 간 전이 확률 행렬  $P$ 로 정의되며, 행렬의 각 요소  $P_{ij}$ 는 군집  $C_i$ 에서  $C_j$ 로의 전이 확률을 나타낸다. 전이 확률  $P_{ij}$ 는 각 군집에서 다른 군집으로 전이되는 빈도를 기반으로 다음과 같이 계산된다:

$$P_{ij} = \frac{n(C_i, C_j)}{\sum_k n(C_i, C_k)}$$

여기서  $n(C_i, C_k)$ 는 군집  $C_i$ 에서  $C_k$ 로의 전이 횟수, 분모는 군집  $C_i$ 에서 모든 군집으로의 전이 횟수를 합산한 값이다. 이와 같은 전이 확률 행렬은 정책이 다양한 상태에서 어떻게 전이되는지를 모델링하여, 군집 간 전이 경로의 신뢰성을 평가하는 데 중요한 역할을 한다.

생성된 DTMC 모델을 활용하여 PRISM 툴을 통해 정책의 안전 속성을 검증하였다. PRISM은 특정 상태에서 실패 상태(예: 충돌, 경계 초과)로 전이될 확률을 계산할 수 있으며, 이를 통해 정책의 위험도를 정량적으로 평가할 수 있다. 본 연구에서는 확률적 계산 트리 논리(PCTL)를 적용하여 정책이 실패 상태에 진입할 확률을 분석하였다. 예를 들어, 정책이 실패 상태에 도달할 확률을 다음과 같이 PCTL 형식으로 나타낼 수 있다:

$$P = ? [F \text{“fail”}]$$

여기서  $P$ 는 실패 상태로 전이될 확률을 의미한다. 마지막으로, 군집 내의 충돌 확률 분포가 정규성을 띠는지 검토하여 군집화의 통계적 적절성을 평가하였다. Kolmogorov-Smirnov 검정을 통해 각 군집 내 분포의 정규성을 확인함으로써, 군집화가 실제 상태 공간의 동적 특성을 충실히 반영하고 있음을 검증하였다. 정규성 검증 결과는 제안 기법이 신

뢰할 수 있는 검증 결과를 제공하는 데 중요한 근거이다.

### III. 실험 및 결과

본 연구는 LunarLander 환경에서 제안된 군집화 기반 검증 기법의 성능과 안전성을 평가하였다. LunarLander는 연속적인 상태 공간에서 정책이 다양한 환경 조건에서 안전하게 착륙할 수 있는지를 평가하기에 적합한 강화 학습 과제이다.

연구에서는 KMeans++ 알고리즘을 사용해 상태 공간을 군집화하고 DTMC 모델을 구축하였다. 최적 군집 수는 115개로 설정되었으며, 이는 상태 전이 패턴을 유지하면서도 검증의 계산 복잡성을 효율적으로 줄일 수 있는 값이다. 이후, 최적 군집 수의 1/4, 1/2, 2배, 4배 등의 다양한 군집 크기를 표 1부터 표 3과 같이 설정(지면상 나머지 생략)했고 상태 공간 축소가 검증 정확성에 미치는 영향을 분석하였다.

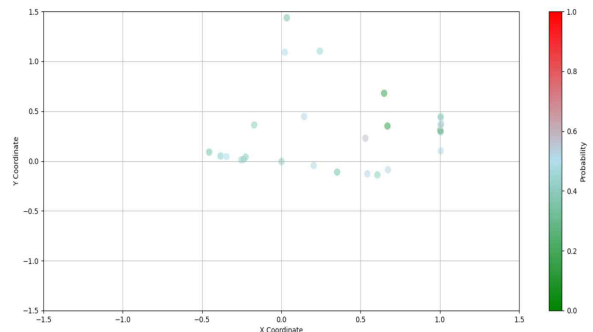


그림 1. 1/4배 군집 설정에서 착륙 중 상태에서 충돌 상태로 전이될 확률

Fig. 1. Landing to crashed probability for the 1/4x cluster configuration

표 1. 1/4배 클러스터링 결과

Table 1. 1/4x Clustering Results

라벨	상태 수	클러스터 수	축소 비율(%)
착륙 중	251,566	11	0.0044
착륙 성공	191	3	1.5707
충돌	418	6	1.4354
경계 이탈	374	7	1.8717
총합	252,549	27	0.0107

그림 1은 1/4배 군집 설정에서 착륙에서 충돌 상태로 전이 확률을 보여주며, 군집 수가 너무 적을 경우 중요한 전이 정보를 상실하여 충돌 확률이 증가하는 양상을 나타낸다.

군집 수가 최적 수준 또는 2배로 증가할 때 정책의 안전성은 개선되었다. 그림 2에서 최적 군집 수(115개) 설정에서의 충돌 상태 전이 확률이 안정적으로 낮게 유지되는 것을 볼 수 있으며, 이는 정책 검증의 효율성과 안전성을 높이는 데 효과적임을 확인시켜준다. 그러나 4배 군집 수로 증가한 경우 그림 3을 보면 추가적인 안전성 향상은 미미하고 계산 복잡도가 증가하는 경향을 보인다. 이로써 적정 군집 수의 설정이 정책 검증 성능에 중요하다는 점이 다시 강조된다.

표 2. 최적 클러스터링 결과

Table 1. Optimal Clustering Results

라벨	상태 수	클러스터 수	축소 비율(%)
착륙 중	251,566	45	0.0179
착륙 성공	191	15	7.8534
충돌	418	25	5.9809
경계 이탈	374	30	9.0214
총합	252,549	115	0.0455

표 3. 4배 클러스터링 결과

Table 1. 4x Clustering Results

라벨	상태 수	클러스터 수	축소 비율(%)
착륙 중	251,566	180	0.0716
착륙 성공	191	60	31.4136
충돌	418	100	23.9234
경계 이탈	374	120	32.0856
총합	252,549	460	0.1821

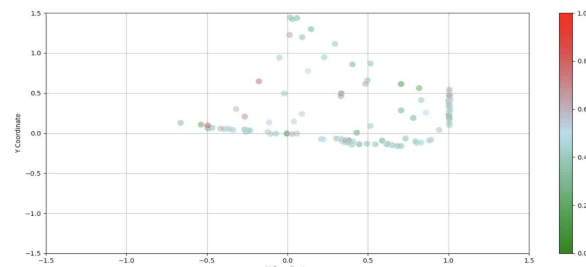


그림 2. 최적 군집 설정에서 착륙 중 상태에서 충돌 상태로 전이될 확률

Fig. 2. Landing to crashed probability for the optimal cluster configuration

정규성 검사는 군집화된 상태 분포가 실제 상태 공간의 특성을 적절히 반영하는지 확인하기 위한 중요한 절차로, 모델의 예측 정확도와 검증 신뢰성에 영향을 미친다. 만약 군집화가 잘 수행되어 군집 내 데이터가 정규성을 따른다면, 각 군집이 상태 공간을 대표하는 모델로서 역할할 수 있어 정책 검증 결과의 신뢰성이 높아진다. 본 연구에서는 군집화 과정의 적절성을 평가하기 위해 Kolmogorov-Smirnov 검정을 수행하였다. 최적 군집 설정에서는 p값이 0.3238로, 분포가 정규성을 유지하여 안정적인 결과를 나타냈다. 이는 군집이 연속 상태 공간의 특성을 잘 보존하고 있음을 시사한다. 2배 군집 설정의 경우, p값이 0.2248로 약간의 복잡성이 추가되며 상태 공간의 정규성에서 벗어나는 경향을 보였으나, 여전히 정규성을 유지할 수 있는 범위였다. 반면, 4배 군집 설정에서는 p값이 0.8321로 가장 높은 정규성을 보이며, 이는 분할과 추상화 간의 균형을 잘 유지하는 이상적인 설정임을 나타낸다.

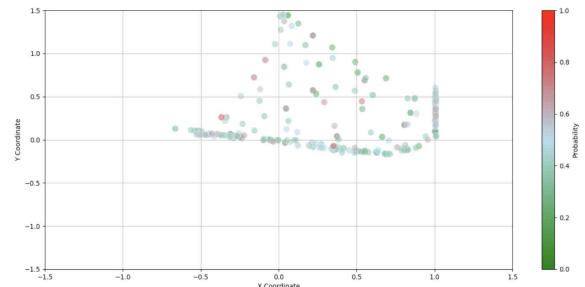


그림 3. 4배 군집 설정에서 착륙 중 상태에서 충돌 상태로 전이될 확률

Fig. 3. Landing to crashed probability for the 4x cluster configuration

그러나 군집 수가 8배 이상으로 증가할 경우, 과도한 세분화로 인해 오히려 정규성이 떨어졌다. 8배 군집 설정에서 p값이 0.0287로 감소하며 비정규성을 나타냈고, 16배 설정에서는 p값이 7.54e-06으로 크게 하락하며 통계적으로 유의미한 비정규성을 보였다. 이러한 결과는 지나친 군집화가 오히려 과적합을 유발하고, 충돌 확률 분포의 왜곡을 초래할 수 있음을 시사한다. 즉, 과도한 군집화는 상태 공간의 불필요한 복잡성을 증가시키며 검증의 안정성을 저해할 수 있다.

결론적으로, 최적 군집 수(115개)와 4배 군집 설

정에서 가장 안정적이고 효율적인 상태 분포가 유지되었다. 최적 군집 수 설정을 통해 LunarLander 환경에서 정책 검증의 효율성과 안전성을 균형 있게 달성할 수 있으며, 연속 상태 공간을 다루는 다양한 강화 학습 문제에 적용할 수 있는 실질적인 방법임을 실험을 통해 확인하였다.

#### IV. 결론

본 연구는 연속 상태 공간에서 강화 학습 정책의 안전성을 효율적으로 검증하기 위해, 상태 공간을 효과적으로 추상화하고 축소하는 군집화 기반 정책 검증 기법을 제안하고 이를 LunarLander 환경에서 평가하였다. KMeans++ 알고리즘을 통해 복잡한 연속 상태 공간을 유사한 상태끼리 군집화하여 이산화된 모델을 생성함으로써, 계산 복잡도를 줄이면서도 주요 전이 정보를 보존하는 방안을 제시하였다. 이 과정에서 상태 공간의 복잡한 구조를 단순화하여 정책의 안전성을 정량적으로 평가하는 데 유의미한 기여를 하였다. 실험 결과, 최적 군집 수(115개) 설정 시 정책 검증의 정확성과 효율성 간의 균형이 가장 잘 유지되었으며, 군집화 과정을 통해 연속 상태 공간을 효과적으로 추상화한 덕분에 계산 자원 소모가 줄어드는 동시에 안정적인 검증 결과를 얻을 수 있었다. 특히 Kolmogorov-Smirnov 검정을 통해 군집화된 상태 분포의 정규성을 검토한 결과, 최적 군집 수 설정과 4배 설정에서 각 군집이 상태 공간의 특성을 적절히 반영함을 확인하였다. 반면 과도한 세분화는 오히려 과적합을 초래하여 충돌 확률 분포의 왜곡과 검증의 불안정성을 유발할 수 있음을 시사하였다.

결론적으로, 본 연구는 복잡한 상태 공간을 효과적으로 추상화하고 축소함으로써 강화 학습 정책의 안전성을 효율적으로 검증할 수 있는 가능성을 입증하였다. 이 방법은 자율 주행 및 로봇 제어와 같은 다양한 응용 분야에서 강화 학습의 안전성을 보장하는 데 기여할 수 있으며, 향후 더 복잡한 상태 공간에서도 성능을 유지할 수 있는 확장 연구가 요구된다. 또한, 본 연구의 접근법은 실시간 제어 시스템에서도 적용 가능성이 높아, 안전성 검증이 필

수적인 산업 분야에도 실질적인 기여를 할 수 있을 것으로 기대된다.

#### 참 고 문 헌

- [1] A. E. Sallab, M. Abdou, E. Perot, and S. Yogamani, "Deep reinforcement learning framework for autonomous driving," arXiv preprint, arXiv:1704.02532, 2017.
- [2] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909-4926, 2021.
- [3] M. Zanon and S. Gros, "Safe reinforcement learning using robust MPC," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3638-3652, 2020.
- [4] W. Zhao, T. He, R. Chen, T. Wei, and C. Liu, "State-wise safe reinforcement learning: A survey," arXiv preprint, arXiv:2302.03122, 2023.
- [5] K. P. Wabersich and M. N. Zeilinger, "A predictive safety filter for learning-based control of constrained nonlinear dynamical systems," *Automatica*, vol. 129, p. 109597, 2021.
- [6] H. Chen and C. Liu, "Safe and sample-efficient reinforcement learning for clustered dynamic environments," *IEEE Control Systems Letters*, vol. 6, pp. 1928-1933, 2021.
- [7] T. P. Gros, J. Groß, D. Höller, J. Hoffmann, and V. Wolf, "DSMC evaluation stages: Fostering robust and safe behavior in deep reinforcement learning," *ACM Transactions on Modeling and Computer Simulation*, vol. 33, no. 4, pp. 1-28, 2023.
- [8] R. A. N. Starre, M. Loog, E. Congeduti, and F. A. Oliehoek, "An analysis of abstracted model-based reinforcement learning," arXiv preprint, arXiv:2208.14407, 2022.

- [9] R. Wang, Y. Zhang, H. Sun, and J. Liu, "Dependable reinforcement learning via timed differential dynamic logic," in IEEE Symposium on Computers and Communications (ISCC), 2021, pp. 1-6.
- [10] L. Lehnert, M. J. Frank, and M. L. Littman, "Reward-predictive clustering," arXiv preprint, arXiv:2211.03281, 2022.

# 심층 오디오 분류기 성능 개선을 위한 확산-모델 기반 데이터 증강

이현주\*, 부석준\*

## Diffusion-model-based Data Augmentation for Enhanced Deep Acoustic Classifier

Hyeon-Ju Lee\* and Seok-Jun Buu\*

### 요약

오디오 분류는 환경 소음 모니터링, 음성 인식, 감정 분석 등 여러 분야에서 중요한 역할을 한다. 하지만, 오디오 데이터는 환경적 요인, 발음의 미세한 차이, 다양한 소스에서 발생하는 소음 등으로 인해 복잡성과 다양성을 가지며, 제한된 데이터로 학습된 모델은 성능이 저하될 수 있다. 본 연구는 이러한 문제를 해결하기 위해 Denoising Diffusion Probabilistic Models (DDPM)을 활용한 생성 모델 기반 데이터 증강 기법을 제안한다. DDPM은 기존 기법에 비해 고품질의 다양한 오디오 데이터를 생성하여 모델의 일반화 능력을 크게 향상시키는 데 효과적임을 입증하였다. 실험 결과, DDPM 기반 증강은 기존의 GAN 및 VAE 기법을 능가하는 성능을 보였으며, 이는 DDPM이 오디오 데이터의 복잡한 특성을 효과적으로 보존하고 재현할 수 있음을 시사한다.

### Abstract

Audio classification plays a crucial role in various fields such as environmental noise monitoring, speech recognition, and emotion analysis. However, audio data is inherently complex and diverse due to environmental factors, subtle differences in pronunciation, and noise from various sources, which can lead to decreased performance in models trained on limited data. To address this issue, this study proposes a generative model-based data augmentation method using Denoising Diffusion Probabilistic Models (DDPM). DDPM has been demonstrated to be effective in significantly improving model generalization by generating high-quality, diverse audio data compared to existing methods. Experimental results show that DDPM-based augmentation outperforms traditional GAN and VAE methods, suggesting that DDPM can effectively preserve and replicate the complex characteristics of audio data.

### Key words

Data Augmentation, Audio classification, Denoising diffusion probabilistic model, Convolution Neural Network

---

\* 경상국립대학교 컴퓨터과학과 (Dept. of Computer Science and Engineering, Gyeongsang National University)  
{ouo\_ofo, sj.buu}@gnu.ac.kr



## I. 서론

오디오 분류는 환경 소음 모니터링, 음성 인식, 감정 분석 등 다양한 응용 분야에서 핵심적인 역할을 담당한다. 그러나 오디오 데이터는 그 특유의 복잡성과 다양성으로 인해, 모델이 제한된 데이터에 의존해 훈련될 경우 일반화 능력과 성능이 크게 저하될 수 있다. 특히, 오디오 데이터는 환경적 요인, 발음의 미세한 차이, 그리고 다양한 소스에서 발생하는 소음 등 다양한 변수에 영향을 받는다. 이러한 데이터 특성은 모델이 광범위한 오디오 특성을 학습하는 데 있어 중요한 도전 과제를 제시한다.

이러한 배경에서, 데이터의 복잡성과 다양성을 충분히 반영할 수 있는 효과적인 데이터 증강 기법이 필요하다. Denoising Diffusion Probabilistic Models (DDPM)은 이러한 복잡성과 다양성을 효과적으로 반영하여 오디오 데이터를 생성할 수 있는 잠재력을 지니고 있다. DDPM은 고품질의 다양한 오디오 데이터를 생성함으로써, 모델이 보다 폭넓은 오디오 특성을 학습하고 이를 통해 일반화 능력을 향상시킬 수 있다.

제안하는 방법은 기존 방법 대비 약 3.17%의 정확도 향상을 보임으로 보다 높은 성능의 오디오 분류 모델을 개발할 수 있는 가능성을 입증하였다.

## II. 관련 연구

최근 오디오 데이터 증강 기법에 대한 연구는 다

양한 응용 분야에서 활발히 진행되고 있다. 2020년에는 신호 변환 기법을 활용한 데이터 증강이 이루어져, 환경 소음 분류의 성능 향상에 기여했다[1]. 같은 해, GAN과 VAE와 같은 생성 모델을 사용한 화자 인식 데이터 증강 연구가 수행되었다[2]. 2023년에는 기계 소음 분류를 위한 신호 변환 기법이 [3], 음성 감정 인식을 위한 신호 변환 기법과 WaveGAN을 결합한 하이브리드 데이터 증강 접근법이 도입되었다[4]. 이러한 흐름은 2024년까지 이어져, 유사한 텍스트-오디오 쌍을 검색하는 Retrieval-augmented 접근법을 통해 드문 소리 생성을 위한 데이터 증강이 시도되었다 [5]. 이러한 연구들은 오디오 데이터 증강의 중요성과 그 응용 가능성을 넓히며, 다양한 오디오 분석 작업에서 모델의 성능을 향상시키고 있다.

## III. 방법

### 3.1 DDPM 기반 오디오 데이터 증강을 통한 오디오 분류

본 연구에서는 Denoising Diffusion Probabilistic Models (DDPM)을 기반으로 한 오디오 데이터 증강 기법을 통해 분류 성능을 향상시키는 방법을 제안한다. 그림 1과 같이 제안하는 방법은 크게 데이터 전처리, DDPM 훈련, 데이터 증강, 분류기 학습의 네 단계로 구성된다.

표 1. 오디오 분석을 위한 데이터 증강 연구

Index	Data	Task	Augmentation Method	Desc.
2020 [1]	ESC-50, ESC-10, Urban sound	분류	Signal Transformation Techniques	환경 소음 분류를 위한 신호 변환 기반 데이터 증강
2020 [2]	Switchboard, NIST SRE	인식	GAN, VAE	화자 인식을 위한 생성 모델을 기반 데이터 증강
2023 [3]	Squeak & Rattle dataset	분류	Signal Transformation Techniques	기계 소음 분류를 위한 신호 변환 기반 데이터 증강
2023 [4]	EmoDB, ERC	인식	Signal Transformation Techniques, WaveGAN	음성 감정 인식을 위한 하이브리드 데이터 증강
2024 [5]	AudioCaps	생성	Retrieval-augmented approach	드문 소리 생성을 위한 유사한 텍스트-오디오 쌍 검색을 통한 데이터 증강

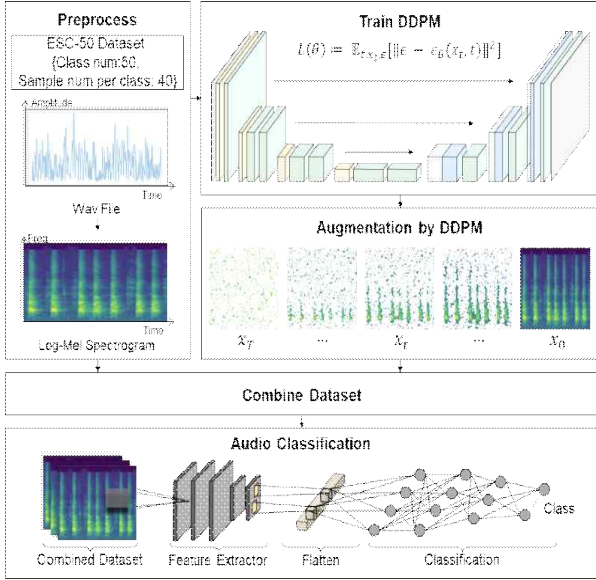


그림 1. DDPM 기반 데이터 증강을 통한 오디오 분류

우선 오디오 데이터를 Log-Mel 스펙트로그램으로 변환하여 오디오 신호의 중요한 주파수 정보를 추출한다. 이 변환된 데이터는 DDPM과 분류기의 입력으로 사용된다. 이후 DDPM은 노이즈가 추가된 데이터에서 원래의 신호를 복원하는 과정을 다양한 노이즈 수준에서 학습한다. 이렇게 훈련된 DDPM을 통해 생성된 데이터는 확률적이고 점진적인 노이즈 제거 과정을 통해 생성되기 때문에 다양하고 복잡한 오디오 데이터의 특성을 효과적으로 반영할 수 있다. DDPM의 이러한 특성으로 분류 모델이 보다 넓은 범위의 오디오 특성을 학습할 수 있도록 도와준다. 이후 생성된 증강 데이터를 기존의 오디오 데이터와 결합하여 분류기를 학습시킨다. 증강된 데이터는 모델의 일반화 성능을 향상시키는 역할을 하며, 다양한 오디오 분류 작업에서 더 높은 정확도를 달성하는 데 기여한다.

### 3.2 오디오 데이터 증강을 위한 DDPM의 학습과 생성

그림 2와 같이, 원본 데이터  $x_0$ 은 Log-Mel 스펙트로그램으로 변환되며, 시간 단계  $t$ 에 따라 점진적으로 노이즈가 추가된  $x_t$ 로 변환된다. 이 과정을 Forward Process라 부르며 수식 (1)을 통해 계산된다.

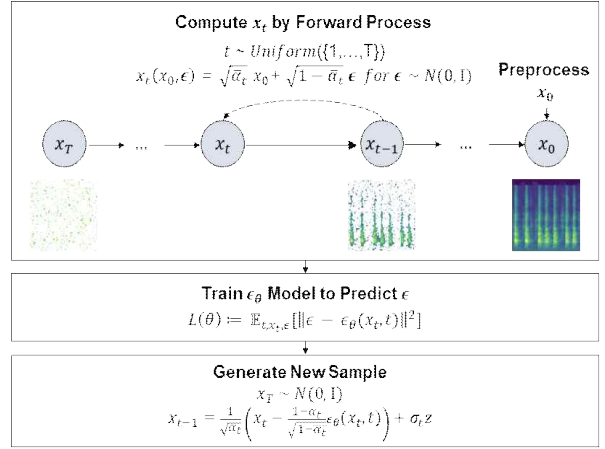


그림 2. DDPM을 통한 데이터 증강 과정

$$t \sim \text{Uniform}(\{1, \dots, T\}) \quad (1)$$

$$x_t(x_0, \epsilon) = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon \text{ for } \epsilon \sim \mathcal{N}(0, I)$$

Forward Process로 계산된  $x_t$ 는 디러닝 모델  $\epsilon_\theta$ 의 입력으로 사용되며, 실제 노이즈와 예측한 노이즈의 차이인 수식 (2)를 최소화하는 과정을 통해  $\epsilon_\theta$  모델이  $x_t$ 에서  $x_{t-1}$ 로 복원하기 위한 노이즈  $\epsilon$ 을 예측하도록 학습하게 된다.

$$\text{Loss}(\theta) := E_{t, x_t, \epsilon} [||\epsilon - \epsilon_\theta(x_t, t)||^2] \quad (2)$$

이렇게 학습한 모델  $\epsilon_\theta$ 을 통해 Forward Process의 역방향 과정인 Reverse Process를 거치면 새로운 오디오 데이터를 얻을 수 있다. 먼저, 완전히 무작위한 노이즈 벡터를 가우시안 분포  $\mathcal{N}(0, I)$ 에서 샘플링하여 Forward Process의 마지막 단계에 대응되는  $x_T$ 를 설정한다. 이후 DDPM의 학습 과정에서 얻은 모델  $\epsilon_\theta$ 을 사용하여,  $t = T$ 에서 시작하여  $t = 0$ 이 될 때까지 수식 (3)을 반복하여 계산한다.

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( x_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \epsilon_\theta(x_t, t) \right) + \sigma_t z \quad (3)$$

이 과정을 반복하여  $t = 0$ 에 도달하면, 최종적으로 노이즈가 제거된 상태의  $x_0$ 을 얻게 된다. 이  $x_0$ 이 생성된 오디오 샘플의 스펙트로그램이며, 이 생성된 샘플은 학습된 데이터 분포를 반영하여, 기존의 데이터와 유사하면서도 새로운 특징을 지닌 오

디오 데이터로 사용할 수 있다. 이러한 이유로, DDPM을 통한 샘플 생성 기반 데이터 증강은 오디오 분류 모델의 성능을 향상시키는데 기여할 수 있다.

따라서, DDPM을 통한 오디오 데이터 생성 과정은 고차원적인 데이터 분포를 효과적으로 모델링하며, 딥러닝 모델을 통해 노이즈를 제거하는 방식으로 새로운 오디오 샘플을 생성한다. 이 방법은 기존 데이터와 유사한 특징을 가지면서도 새로운 변화를 포함한 데이터를 얻을 수 있도록 하며, 이러한 샘플들은 오디오 데이터 증강의 일환으로 활용될 수 있다.

#### IV. 실험 및 결론

##### 4.1 실험 오디오 데이터

본 연구에서는 다양한 환경 소리와 인간 음성 이외의 소리를 포함하는 오디오 데이터셋인 ESC-50를 사용하여 실험을 진행하였다. ESC-50 데이터셋은 총 50개의 클래스로 구성되어 있으며, 샘플들은 모두 5초 길이의 오디오 데이터로, 44.1kHz의 샘플링 레이트를 가지는 WAV 파일 형식으로 제공된다.

##### 4.2 DDPM 기반 데이터 증강을 통한 오디오 분류 성능

표 2에서 다양한 데이터 증강 방법을 사용하여 오디오 분류 성능을 비교한 결과를 정리하였다. 데이터 증강을 하지 않은 CNN의 경우 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1-Score 모두 생성 모델 기반 데이터 증강 방식에 비해 낮은 것을 확인할 수 있었다. 반면, 기하학적 변환을 통해 데이터 증강을 한 경우는 성능이 더욱 하락한 것을 볼 수 있으며, 이는 이미지 데이터에서와 달리, 기하학적 변환이 오디오 데이터의 복잡성과 다양성을 반영하지 못하고 부적절하게 적용되었음을 시사한다. 반면, 생성 기반 데이터 증강 방식은 CNN에 비해 모든 지표에서 높은 성능을 보였으며, 그 중 제안하는 방법인 DDPM은 정확도 0.6516, 정밀도 0.6827, 재현율 0.6516, F1-Score 0.6496로 가장 높은 분류 성능을 보였다. 이는 DDPM이 데이터

의 다양성과 복잡성을 잘 반영한 샘플을 생성하여, 모델이 더 다양한 패턴을 학습하고 일반화 능력을 향상시킨 결과로 해석할 수 있다.

표 2. 방법 별 분류 성능 비교

Metrics	Accuracy	Precision	Recall	F1-Score
Base Network				
CNN	0.6199	0.6702	0.6199	0.6091
Geometric Transformations	0.2529	0.2543	0.2529	0.2031
Generative Augmentation				
GAN	0.6344	0.6725	0.6344	0.6303
VAE	0.6402	0.6697	0.6402	0.6384
DDPM (ours)	0.6516	0.6827	0.6516	0.6496

표 3에서는 데이터 증강 샘플의 수에 따른 성능을 비교하였다. 데이터 증강 없이 원본 데이터만 학습했을 때는 가장 낮은 성능을 기록했으며, 데이터 증강 수가 증가함에 따라, 모든 성능 지표에서 지속적으로 향상되는 것을 확인할 수 있다. 이 결과는 더 많은 증강 데이터가 모델의 일반화 능력을 높여, 더 정확하고 균형 잡힌 예측을 가능하게 했음을 시사한다. 이처럼 데이터 증강은 오디오 분류 작업에서 모델의 성능을 향상시키는데 효과적인 접근임을 알 수 있다.

표 3. 확산-모델로 생성된 데이터 수에 따른 성능 비교

Metrics	Accuracy	Precision	Recall	F1-Score
0 (non-augmented)	0.6199	0.6702	0.6199	0.6091
100	0.6380	0.6550	0.6380	0.6245
200	0.6430	0.6725	0.6430	0.6325
300	0.6471	0.6874	0.6471	0.6461
400	0.6516	0.6827	0.6516	0.6496

그림 3에서는 원본 데이터와 GAN, VAE, DDPM으로 생성한 스펙트로그램을 비교하였다. GAN과 VAE로 생성된 데이터는 원본 데이터와 비교했을 때 주파수 및 시간적 패턴에서 왜곡이 발생하거나 노이즈가 포함된 반면, DDPM이 생성한 데이터는 원본 데이터의 주파수 및 시간적 패턴을 훨씬 더 잘 보존하며, 고품질의 데이터를 생성하는 것을 확인할 수 있다. 이는 DDPM이 오디오 데이터의 복잡한 특성을 더욱 정확하게 재현할 수 있음을 보여준다.

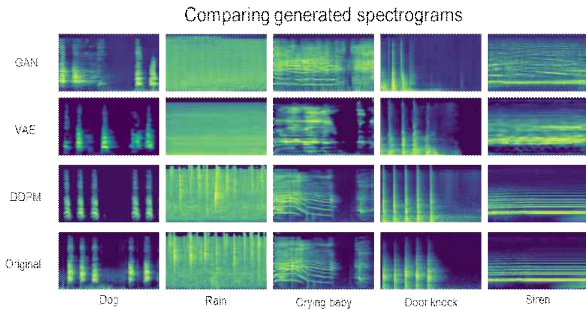


그림 3. 생성된 오디오 스펙트로그램 비교

### 4.3 결론 및 향후 연구

본 연구에서는 Denoising Diffusion Probabilistic Models (DDPM)을 기반으로 한 오디오 데이터 증강 기법을 제안하고, 이를 통해 오디오 분류 성능을 향상시키는 방법을 탐구하였다. 실험 결과, DDPM 기반 데이터 증강은 기존의 GAN과 VAE와 같은 생성 모델 기반 증강 기법에 비해 우수한 성능을 보여주었다. 이는 DDPM이 오디오 데이터의 복잡성과 다양성을 잘 반영하여 고품질의 샘플을 생성할 수 있음을 입증한다.

제안하는 방법에 사용된 DDPM은 클래스 정보를 조건으로 사용하지 못하는 한계가 있어, 각 클래스마다 별도의 생성 모델을 만들어야 했다. 이러한 접근법은 클래스 수가 많아질수록 비효율적일 수 있기 때문에 조건부 DDPM을 통한 데이터 증강 연구가 필요하다.

### 감사의 글

This work was supported by POSCO Technical Research Laboratories (2024P023).

### 참고 문헌

[1] Z.Mushtaq and S.F.Su, "Environmental sound classification using a regularized deep convolutional neural network with data augmentation," *Applied Acoustics*, vol. 167, pp.107389, 2020.

[2] S.Wang, Y.Yang, Z.Wu, Y.Qian and K.Yu, "Data augmentation using deep generative models for

embedding based speaker recognition," *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, vol.28, pp.2598-2609, 2020.

[3] A.Abeysinghe, S.Tohmuang, J.L.Davy and M.Fard, "Data augmentation on convolutional neural networks to classify mechanical noise," *Applied Acoustics*, vol. 203, pp.109209, 2023.

[4] N.T.Pham, D.N.M.Dang, N.D.Nguyen, T.T.Nguyen, H.Nguyen, B.Manavalan and S.D.Nguyen, "Hybrid data augmentation and deep attention-based dilated convolutional-recurrent neural networks for speech emotion recognition," *Expert Systems with Applications*, vol.230, pp.120608, 2023.

[5] Y.Yuan, H.Liu, X.Liu, Q.Huang, M.D.Plumbley and W.Wang, "Retrieval-augmented text-to-audio generation," *Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 581-585, 2024.

# 디지털 트윈용 범용성 보장 실시간 네트워크 통신 서버 프로그램 설계 연구

황교찬\*, 기재석\*\*, 권재민\*\*\*, 이찬영\*\*\*, 박진섭\*\*\*,

## Research on Universal Real-time Network Server Architecture for Digital Twin Systems

Hwang, Kyo-Chan\*, Ki, Jae-Sug\*\*, Kwon, Jae-Min\*\*\*, Lee, Chanyoung\*\*\*, Park, Jinseop\*\*\*

### 요 약

디지털 트윈(Digital Twin)은 물리적 시스템의 가상 모델을 실시간으로 동기화하고 모니터링하여, 다양한 산업 분야에서 시스템의 성능을 최적화하고 예측 가능한 문제를 사전에 대응할 수 있도록 하는 중요한 기술이다. 특히, 제조업, 스마트 시티, 헬스케어, 물류 등 다양한 산업에서 디지털 트윈 기술은 효율성 향상, 비용 절감, 그리고 운영 리스크 감소를 목표로 활발히 도입되고 있다. 디지털 트윈의 핵심은 실제 물리적 시스템과의 실시간 연계 및 데이터의 실시간 처리 능력에 있다. 이러한 연계성을 통해 디지털 트윈은 물리적 환경에서 발생하는 변화를 분석하여 의사결정 과정을 최적화할 수 있다. 하지만, 디지털 트윈을 위해 연결되는 장치가 다양해지면서 제공받는 데이터의 형태가 일률적이지 않다는 문제를 안게 되었다. 이로 인해 각각의 장치에 맞게 통신 시스템을 구축하게 되면서 개발 기간 및 비용의 증가로 이어지게 된다. 본 논문에서는 디지털 트윈 환경에서의 다양한 장치의 신호를 받을 수 있는 범용성이 보장된 서버 프로그램 설계를 다룬다. 이를 통해 보다 쉽게 디지털 트윈 환경을 구축할 수 있는데 기여할 것으로 기대된다.

### Abstract

Digital Twin technology is a crucial innovation that synchronizes and monitors virtual models of physical systems in real-time, enabling various industries to optimize system performance and proactively address predictable issues. Particularly in manufacturing, smart cities, healthcare, and logistics, digital twin technology is being actively implemented to improve efficiency, reduce costs, and minimize operational risks. The core of digital twin lies in its real-time connectivity with physical systems and its ability to process data in real-time. Through this connectivity, digital twins can analyze changes in the physical environment to optimize decision-making processes. However, as the devices connected for digital twins become more diverse, the problem arises that the data formats received are not uniform. This leads to increased development time and costs as communication systems must be built specifically for each device. This paper addresses the design of a universal server program that can receive signals from various devices in a digital twin environment. It is expected to contribute to easier construction of digital twin environments.

### Key words

realtime network, digital twin, flexible protocols, IoT, universal server program

---

\* (주)트라이텍 이사, eekdro@gmail.com, \*\* (주)트라이텍 대표이사, ceo@tritech.co.kr(교신저자),  
\*\*\* 핵융합에너지연구원, cylee@kfe.re.kr

※ 이 논문은 2024년도 한국핵융합에너지연구원 주요사업 “고성능 핵융합 시뮬레이션 연구”의 지원을 받아 수행된 연구임

## 1. 서 론

디지털 트윈을 시도하거나 준비하는 산업군이 점차 증가하고 있다.[1] 특히, 설비 및 공정을 실시간으로 모니터링하고 분석하고자 하는 연구는 다양하게 이루어지고 있다.[2] 이러한 모니터링을 통해 제조 프로세스를 개선하나 생산 공정에 소요되는 시간을 절감하게 된다.[3] 원자재와 같은 고위험군의 경우, 정확한 데이터를 각각의 Devices에서 수집해야만 정밀한 모니터링이나 시뮬레이션이 가능하다.[4][5] 이처럼 디지털 트윈을 위한 데이터 수집은 설비의 규모나 복잡도에 따라 적게는 한 두개에서 많게는 수 백개에 이르는 데이터를 정밀하게 수집해서 분석해야 된다. 이러한 상황에서 각기 다른 Devices에서 제공되는 데이터의 형태가 일률적이지 않아 디지털 트윈을 구축하기 위해서는 각각의 Devices에서 발생하는 데이터에 맞게 Server를 설계하고 개발해야 되는 문제를 안게 된다.

이에 본 논문에서는 다양한 Devices에서 사용 가능한 범용성이 보장된 디지털 트윈용 서버 설계를 제안함으로써 개발 시간 단축에 기여하고자 한다.

## II. 본 론

### 1. 용어

표 1. 용어 및 설명

용어	설명
장치 (Device)	물리적 하드웨어로 아날로그 또는 디지털 데이터를 외부로 송출할 수 있는 기능을 갖춘 기기
Server 시스템 (Server System)	네트워크 처리에 특화된 컴퓨터, Server 프로그램이 실행되는 장치
Server 프로그램 (Server Program)	다수의 클라이언트와 네트워크 상에 연결되어 데이터 송. 수신을 제공 할 수 있는 프로그램
클라이언트 (Client)	Server 프로그램과 연결하여 데이터를 송. 수신하는 프로그램
패킷 (Packet)	네트워크 통신 시 데이터 전송 형태

### 2. 범용성이 보장

본 논문에서 제시하는 범용성이 보장된 Server Program (이하 Server)은 여러 Device에서 발생하는 데이터를 수집하여 각각의 수신측 Client가 원하는 데이터만 선별적으로 제공하는 것을 목표로 한다.



그림 1. 디지털 트윈 데이터 흐름

그림1은 가장 일반적인 디지털 트윈의 데이터 흐름을 나타내고 있다. Sensor, Manipulator, IoT 등에서 발생하는 데이터를 Network을 통해 Server 로 전송하게 된다. Device에 따라 한 개 이상의 데이터를 동시에 발생할 수도 있다. 이렇게 Server로 전송된 데이터는 해당 데이터를 필요로 하는 Clients에게 전송하게 된다. 수신된 데이터를 기반으로 Simulation, Monitoring, Dashboard, Analysis 등 가시화 및 분석용으로 사용하게 된다. 이렇게 다양한 Devices의 데이터 전송을 수용할 수 있게 하기 위해 3가지 설계안을 제안한다.

- 다양한 Devices의 데이터를 수용할 수 있는 패킷 구조
- Server는 한 대 이상의 Devices에서 전송되는 데이터를 병렬적으로 수용 가능
- 다수의 수신측 Client는 각기 원하는 Devices의 데이터를 선별적으로 수신 가능

### 3. 유동성이 보장된 패킷 설계

다양한 Device에서 발생하는 데이터 수용이 가능한 범용적 패킷 설계를 위해서는 전체 Devices에서 발생하는 다양한 데이터를 감싸는 구조의 유동적인 패킷 설계가 필요하다.

1 요청			pkcRequest
Name	Type	Field	Detail
패킷키	ePacketKey	PacketKey	
데이터	string	Data	장비에서 얻어진 데이터를 Json 형식으로 보관한다.
2 결과			pkcResult
Name	Type	Field	Detail
성공여부	bool	success	요청한 처리 결과 성공 여부
패킷키	ePacketKey	packetKey	pkcRequest 에 넣은 PacketKey 값을 리턴한다
데이터	string	data	success == false 이 경우 pkError 을 결과 값으로 리턴된다.
3 예러			pkcError
Name	Type	Field	Detail
ErrorKey	ePacketError	ErrorKey	
ErrorDetail	string	ErrorDetail	오류 상세 상황

그림 2. 최상위 패킷 구조

그림 2는 Server로 전송되는 최상위 패킷의 구조가 기술되어 있다. Server는 pkcRequest 형태로만 정보를 수신한다. Type 중 ePacketKey는 Enum 형으로 pkcRequest를 통해 발송하는 패킷의 종류를 기록한다. ePacketKey에 해당하는 데이터는 pkcRequest.Data 안에 Json(JavaScript Object Notation) 형태로 전달 된다. Json을 사용함으로써 유동적인 형태의 데이터 구조를 갖더라도 Client에서 ePacketKey로 어떤 데이터인지 확인이 가능하다. Server가 pkcRequest 에 대한 처리를 완료했을 경우 pksResult 로 결과를 Client에게 제공하며 세부적인 내용은 pksResult.Data 에 Json 형태로 전달된다. 이를 통해 Server 프로그램에서 일관성 있는 패킷 관리가 가능하다.

10 데이터 송신 시작			pkcSendStart
Name	Type	Field	Detail
데이터 키	string	DataKey	장치마다 고유한 키를 부여
데이터 구분	int	SubKey	장치에서 발생하는 데이터 구분 키
11 데이터 송신			pkcSendProgress
Name	Type	Field	Detail
데이터 구분	int	SubKey	
누적 순서	long	SequenceNum	데이터 정렬 및 무결성 검증용
Data	string	Data	Byte[] 지원을 위해 Base64 로 인코딩 전송
12 송신 완료			pkcSendDone
Name	Type	Field	Detail
데이터 구분	int	SubKey	

그림 3. 데이터 송신 측 패킷 구조

그림 3은 Device에서 Server로 데이터를 전송하는 송신측 패킷의 구조가 기술되어 있다. 송신측의 모든 패킷은 pkcRequest의 Data에 Json 형태로 저장되어 전송된다. pkcSendStart는 전송 시작을 Server에 알리게 된다. pkcSendStart.DataKey는 Device에 따라 고유하며 Server는 DataKey를 기준으로 Device 그룹을 생성하여 관리하게 된다. 이를 통해 다양한

Device의 데이터를 유동적으로 수용할 수 있는 구조를 갖게 된다.

20 실시간 데이터 요청			pkcRevLiveStart
Name	Type	Field	Detail
데이터 키	string	DataKey	
데이터 구분	int	SubKey	
21 데이터 수신			pkcRevProgress
Name	Type	Field	Detail
데이터 구분	int	SubKey	
순서	long	SequenceNum	
Data	string	Data	
22 데이터 전송 완료			pkcRevProgressDone
Name	Type	Field	Detail
데이터 키	string	DataKey	
데이터 구분	int	SubKey	
23 데이터 수신 중지			pkcRevStop
Name	Type	Field	Detail
데이터 구분	int	SubKey	

그림 4. 수신 측 패킷 구조

그림 4는 Server로 데이터를 수신하는 패킷의 구조가 기술되어 있다. 수신측 Client에서 pkcRevLiveStart 는 Device에서 Server로 보내는 신호 수신 시작을 알리게 된다. Server는 DataKey 와 SubKey를 기준으로 데이터를 관리한다. 실제 Device에 대한 데이터는 pksRevProgress.Data에 Base64로 인코딩된 형태로 저장되어 전송된다. 이는 Text와 Binary 둘 다 전송이 가능하게 하기 위함이다. Client는 pksRevProgress.Data를 Device의 정보에 맞게 Parsing 하여 사용하게 된다.

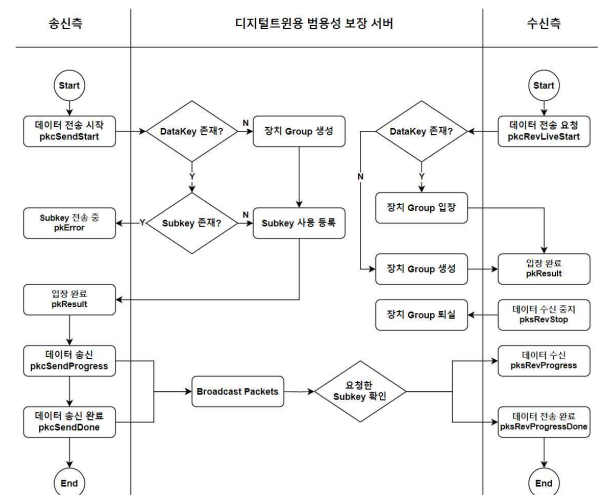


그림 5. 송.수신 패킷 흐름도

그림 5는 데이터 송.수신 패킷의 흐름을 나타내고 있다. Device Group에 접속된 Client는 송.수신

패킷 구조에 따라 데이터를 전송하게 된다. 송신측 Client에서 pksSendProgress.Data 에 Device에서 얻은 신호 또는 데이터 값을 저장해서 Server로 보내면 해당 Device Group 속에 연결되어 있는 수신측에 Client에 Subkey 일치 확인 후 pksRevProgress.Data에 데이터를 담아 전송하게 된다.

#### 4. 결과

본 논문에서는 다양한 Device의 데이터 구조를 수용하여 수신 측에 제공하고자 했다.

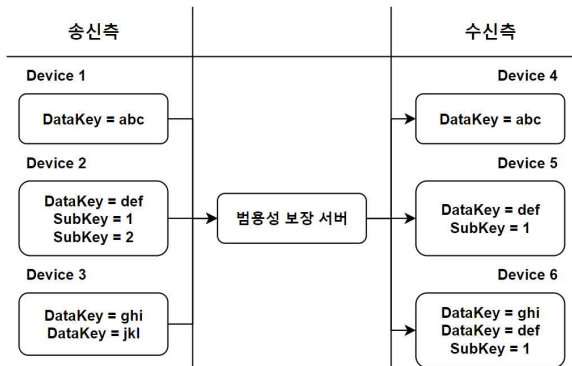


그림 6. 데이터 전달의 유연성 보장 흐름도

본문 3의 유동성이 보장된 패킷 설계를 활용하면 그림 6처럼 데이터를 주고받을 수 있다. 송신측 Device 1,2,3 에서 각각 다양한 형태의 데이터가 pksRevProgress의 Data에 넣어서 발송되면 수신측 Device 4,5,6 에서는 특정 Device에서 발생된 정보 중 Subkey 를 통해 원하는 정보만을 수신할 수 있다. 또한 수신측 Device처럼 송신 측에 발생하는 다양한 데이터를 동시에 수용하는 것도 가능하다. 이를 통해 디지털 트윈 구축 시 발생될 수 있는 다양한 Devices와의 네트워크 통신을 유연하게 구축할 수 있다.

### III. 결 론

본 연구에서는 다양한 Devices의 데이터를 수집하고 전송할 수 있는 디지털 트윈용 네트워크 통신 서버 프로그램을 설계하였다. 제안한 Server는 범용성과 유동성이 보장된 패킷 구조를 통해 다수의

Devices에서 발생하는 데이터를 실시간으로 수집하고, Clients가 필요로 하는 특정 데이터를 효율적으로 제공할 수 있도록 설계되었다. 이로써 디지털 트윈 시스템 구축 시 발생하는 데이터 형식의 불일치 문제를 해결하고, 개발 비용과 시간을 절감하는 데 기여 할 수 있을 것으로 기대된다.

본 연구의 결과는 제조업, 스마트 시티, 헬스케어 등 다양한 산업 분야에서 디지털 트윈 시스템의 효율적인 구축을 지원할 수 있는 중요한 토대가 될 것으로 보인다. 향후에는 이 시스템의 성능을 더욱 향상시키고, 다양한 실제 응용 사례에 적용하여 디지털 트윈 기술의 효과성을 입증하고자 한다.

### 참 고 문 헌

- [1] G.N. Schroeder, C. Steinmetz, R.N. Rodrigues, R.V.B. Henriques, A. Rottberg, and C.E. Pereira, "A Methodology for Digital Twin Modeling and Deployment for Industry 4.0" Proceedings of the IEEE, vol. 109, No. 4, pp. 556-567, 2020.
- [2] C.Y. Lim, C.E. Yeo, S.Y. Ahn, M.O. Lee, H.J. Sung "Design and Performance Evaluation of Digital Twin Prototype Based on Biomass Plant" The Journal of the Convergence on Culture Technology (JCCT) 9, no.5 (2023) 789-795.
- [3] S. H. Moon, "Big Data Platform Construction and Application for Smart City Development" The journal of the convergence on culture technology, vol. 6, no. 2, pp. 529-534, May 2020
- [4] K. C. Hwang, "A Study on the Reliability of Time Interval by the Application of Digital Twin Technology in Actual Equipment and Simulator Developed as Unity." The Korean Society of Disaster Information. 2020.11 (2020): 395-400.
- [5] K. C. Hwang, "A Study on the Development of High Sensitivity Collision Simulation with Digital Twin." The Korean Society of Disaster Information. vol. 16.4 (2020): 813-823.



# 화약 총을 이용한 혼합 현실 FPS 게임에서의 몰입형 감각 경험 향상

최보규\*, 김웅식\*\*, 이수원\*\*\*

## Enhancing Immersive Sensory Experiences in Mixed Reality FPS Games Using Gunpowder Guns

Bogyu Choi\*, Ungsik Kim\*\*, and Suwon Lee\*\*\*

### 요약

최근 혼합 현실(MR) 환경에서의 사용자 상호작용은 점점 더 현실감 있고 몰입감 있는 경험을 제공하기 위해 발전해 오고 있다. 하지만 전통적인 컨트롤러는 각 사물을 정확하게 모방하지 못해 현실적인 경험을 충분히 전달하지 못한다는 한계가 존재한다. 이를 개선하기 위해, 우리는 MR 환경에서 화약총을 사용하여, 보다 직관적이고 몰입감 있는 상호작용 방식을 제안한다. 화약총의 실제적인 발사 경험을 MR 내에서 구현함으로써 사용자는 실감 나는 촉각 피드백과 현실감 있는 상호작용을 경험할 수 있다. 본 연구에서는 화약총의 하드웨어와 MR 소프트웨어 간의 동기화를 위한 기술적인 구현 방안을 소개한다. 제안된 방법은 기존 컨트롤러 대비 높은 몰입감을 제공함으로써 MR 환경에서의 새로운 가능성을 제시한다.

### Abstract

In recent years, user interaction in mixed reality (MR) environments has evolved to provide increasingly realistic and immersive experiences. However, traditional controllers are limited in that they are not intuitive to use and do not fully convey the physical experience. To improve this, we propose a more intuitive and immersive way of interacting with the MR environment, using a gunshot gun. By implementing the physical experience of firing a gun in MR, users can experience realistic tactile feedback and realistic interaction. In this study, we present a technical implementation for synchronization between the hardware of the gun and the MR software. The proposed method offers new possibilities for MR environments by providing higher immersion compared to existing controllers.

### Key words

mixed reality, gunpowder gun, 3D pose estimation, firing event detection

## 1. 서론

혼합 현실(MR)은 가상 요소를 현실 세계와 융합

하여 플레이어에게 몰입적이고 상호작용적인 경험을 제공한다[1, 2]. 이러한 MR 환경에서 몰입감을 높이기 위해선 감각적 피드백이 중요한 요소 중 하

\* 경상국립대학교 컴퓨터공학과, chlqhrb8769@naver.com

\*\* 경상국립대학교 컴퓨터공학과, blpeng@gnu.ac.kr

\*\*\* 경상국립대학교 컴퓨터공학과, leesuwon@gnu.ac.kr (교신저자)

나이다[3]. 전통적인 컨트롤러는 MR 환경에서 여러 사물로의 호환성은 뛰어나지만, 각각의 사물들을 정확하게 모방하지 못한다. MR 환경에 대한 몰입감을 높이기 위해 더 현실적인 대안을 찾는 연구가 진행되고 있다[4, 5]. 본 연구에서는 슈팅 게임 환경에서 플레이어의 감각적 몰입도를 높이기 위해 컨트롤러를 대신해, 화약총을 활용한 MR 기반 1인칭 슈팅 게임을 개발하였다. 헤드 마운트 디스플레이(HMD)를 사용해 총기의 위치를 정확하게 예측하고, 발사 이벤트를 감지하여 실시간 피드백을 제공한다.

## II. 개발 과정

이 연구에서는 MR 환경에서 플레이어의 슈팅 경험을 향상하기 위해 전통적인 컨트롤러 대신 화약총을 선택하였다. HMD 카메라를 사용하여 총의 위치와 방향을 예측하고, 발사 소리의 크기를 통해 발사 이벤트를 감지한다. 개발 과정은 총기의 3D 자세 추정, 발사 이벤트 감지의 두 가지 주요 단계로 이루어진다.

### 2.1 총의 3D 포즈 예측

HMD 카메라를 사용하여 손목, 중수골, 근위 지

간, 원위 지간 등 관련 손 관절의 위치에 대응했다. 관절 지점을 선으로 연결하여 손의 골격을 만들었다. 그림 2는 총을 들고 있는 손의 관절의 상대적 위치와 방향을 분석하여 총의 6DOF 자세를 계산하고 실시간으로 업데이트하는 과정을 보여준다.

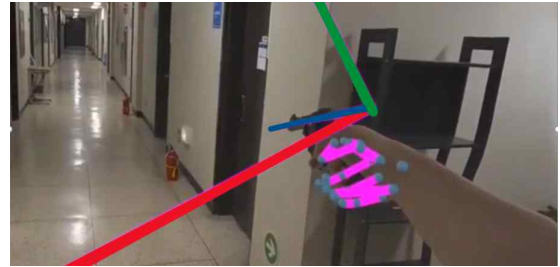


그림 2. 화약총의 6DOF 자세 추정  
Fig. 2. Estimating 6DOF pose of the gun

### 2.2 격발 이벤트 감지

총알이 발사되는 이벤트 감지 시스템을 개발했다. 총기는 일상 소음보다 훨씬 더 큰 소리를 낸다. 여러 차례 데시벨(dB)을 측정한 결과, HMD에 포착된 화약총 소리의 dB가 약 100dB 이상이라는 사실을 발견했다. 따라서 임계값을 100dB로 설정했다. 시스템은 HMD의 마이크를 사용하여 소리를 감지하고 임계값을 초과하면 총알을 발사한다.

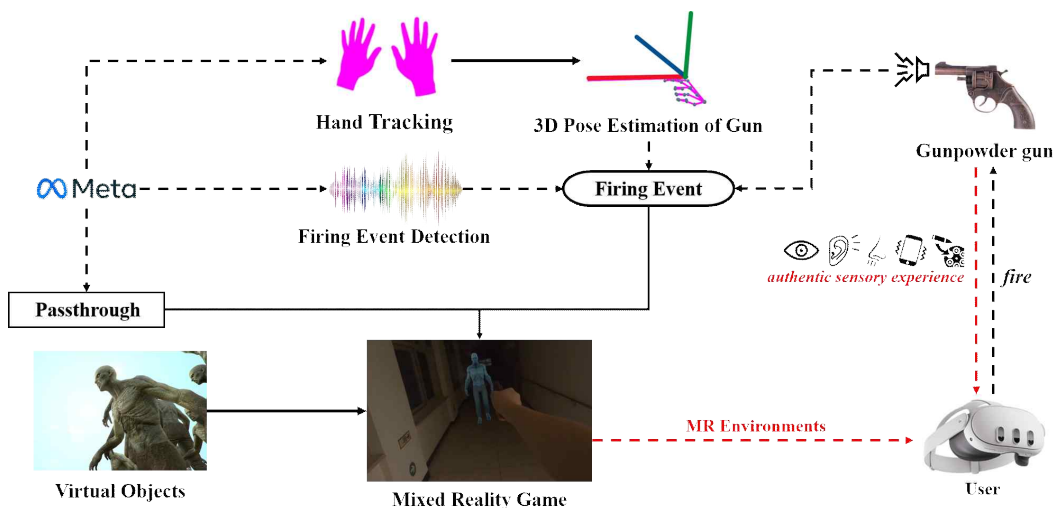


그림 1. 개발 및 시연 과정  
Fig. 1. Development and demonstration process

### III. 감각 경험 향상

이 연구에서는 기존 전통적인 컨트롤러에서 느낄 수 없는 시각 효과, 후각 효과, 청각 효과, 물리적 효과, 수동 재장전 경험 등 실제와 같은 감각적 경험을 플레이어에게 제공한다. 플레이어가 화약총을 발사하면 연기, 총성, 총기 반동들이 발생한다. 연기는 플레이어에게 시각적으로 보임과 동시에 화약 냄새를 나게 해 후각을 자극한다. 총성은 스피커로는 낼 수 없는 현실감 있는 소리를 전달하여 플레이어의 청각을 자극한다. 화약총을 격발했을 때 생기는 반동은 전통적인 컨트롤러로는 느낄 수 없는 새로운 물리적 감각을 더해 준다. 마지막으로 이 시스템의 독특한 특징은 플레이어가 직접 무기의 화약을 재장전해야 하는 수동 재장전 메커니즘이다. 이 요소는 게임 플레이에 촉각적인 경험을 도입하여 전반적인 사실감을 크게 향상한다.

### IV. 게임 방식

데모용 MR 기반 1인칭 슈팅 게임을 개발하였다. 이 게임은 메타퀘스트와 같은 MR 장비를 사용하여 플레이할 수 있다. 게임에서 플레이어는 제한된 시간 동안 건물 안에 있는 좀비를 쏘거나 피해서 생존해야 한다. 플레이어의 편의를 위해 좀비는 플레이어의 카메라 시야각에 따라 생성되므로 플레이어는 좀비가 나타나는 방향을 예측할 수 있다. 그림 3과 그림 4는 각각 게임의 1인칭 시점과 게임을 플레이하는 플레이어의 모습을 보여준다.



그림 3. 1인칭 시점의 게임플레이 장면  
Fig. 3. First-person view of the gameplay



그림 4. 게임을 플레이하는 플레이어  
Fig. 4. Player playing the game

### V. 실험

#### 5.1 실험 목적

본 실험은 기존 컨트롤러 대신 실제 화약총을 사용한 혼합 현실 게임에서 정확성을 평가하는 것을 목적으로 한다. 실제 화약총을 사용한 인터페이스가 정확도 측면에서, 컨트롤러를 대체해, 사용이 가능하다는 가설을 검증하고자 한다.

#### 5.2 실험 방법

실험은 100발의 총알을 발사하여 MR 환경 내 좀비 목표물에 대한 명중 횟수를 기록하는 방식으로 진행되었다. 각 발사는 혼합현실 환경에서 시각화된 좀비를 대상으로 이루어졌으며, 목표물은 일정한 경로를 따라 움직이도록 설정되었다. 총기의 조준 방식과 발사 방식은 모두 실제 화약총을 기준으로 했으며, 명중 횟수를 기준으로 정확성을 평가하였다.

#### 5.3 실험 결과

표 1은 화약총과 컨트롤러를 비교한 실험에서 사용자가 좀비를 명중시킨 횟수를 보여준다. 실험 결과, 실제 화약총을 사용한 MR 환경에서 사용자의 명중 정확성이 컨트롤러를 사용할 때보다 높게 나타났다.

표 1. 화약총과 컨트롤러의 명중률  
Table 1. Gunpowder gun and controller hit rate

	화약총	컨트롤러
명중 횟수	39.75	38.75

이는 MR 환경에서 화약총이 컨트롤러를 대체할 수 있는 가능성을 시사한다. 하지만 한계점으로, 실험에 참여한 사용자들 중 메타퀘스트 자체가 처음인 경우가 많았고, 컨트롤러가 실제 총기처럼 직관적이지 않았기 때문에 화약총을 사용할 때 더 높은 명중률을 보였을 가능성이 있다.

Popova. "Spatial presence and perceived reality as predictors of motion-based video game enjoyment." *Presence: Teleoperators and Virtual Environments* 20.6 (2011): 591-619.

## VI. 결 론

이 연구에서 제안한 MR 기반 1인칭 슈팅 게임은 MR 경험의 사실성과 몰입도를 향상시키는 데 있어 다감각 통합의 중요성을 강조한다. 화약총과 HMD를 활용한 게임은 시각, 후각, 청각, 물리적 피드백을 통합하여 플레이어의 몰입감을 크게 향상시켰다. 수동 장전 과정은 게임 플레이에 독특한 촉각적 차원을 추가하며, 본 시스템은 MR 기술이 게임 산업에 미치는 잠재력을 보여준다. 향후 연구에서는 실제 물체나 다른 사용자와의 상호작용이 더욱 몰입감을 높일 수 있을 것이다.

## 참 고 문 헌

- [1] Costanza, Enrico, Andreas Kunz, and Morten Fjeld. *Mixed reality: A survey*. Springer Berlin Heidelberg, 2009.
- [2] Milgram, Paul, and Fumio Kishino. "A taxonomy of mixed reality visual displays." *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994): 1321-1329.
- [3] Slater, Mel, and Sylvia Wilbur. "A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments." *Presence: Teleoperators & Virtual Environments* 6.6 (1997): 603-616.
- [4] Wei, Tzu-Yun, et al. "Elastilinks: Force feedback between vr controllers with dynamic points of application of force." *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 2020.
- [5] Shafer, Daniel M., Corey P. Carbonara, and Lucy

# 손 끼임 사고 방지를 위한 듀얼 카메라 기반 위험영역 침범 판별 시스템

왕태수\*, 장시웅\*\*

## Dual Camera-based Risk Zone Intrusion Detection Algorithm to Prevent Hand Entrapment Accidents

Tae-su Wang\* and Si-Woong Jang\*\*

### 요 약

본 연구는 산업 현장에서 빈번히 발생하는 손 끼임 사고를 방지하기 위해 듀얼 카메라를 활용한 저비용 시스템을 제안한다. 단일 카메라 시스템의 한계를 극복하고, 다각도에서 수집한 이미지를 바탕으로 손의 위치와 움직임을 3차원에서 추적하여 위험 구역과의 거리를 실시간 계산한다. 이를 통해 잠재적 사고를 예측하고 작업자에게 경고를 제공하여 안전을 강화한다. 저가형 카메라로도 실시간 모니터링이 가능하며, 고성능 카메라를 사용할 경우 응답 속도가 향상된다.

### Abstract

This study proposes a low-cost system using dual cameras to prevent hand entrapment accidents, which frequently occur in industrial settings. It overcomes the limitations of single-camera systems by using images captured from multiple angles to accurately track hand positions and movements in 3D space, calculating the distance from danger zones in real time. The system predicts potential accidents and provides warnings to ensure worker safety. It enables real-time monitoring even with low-cost cameras, with improved response times when using high-performance cameras.

### Key words

Camera, Hand Entrapment Prevention, Hand Estimation, Geometry Information

## 1. 서 론

산업 현장에서 손 끼임 사고는 작업자의 안전을 위협하는 주요 사고 유형이며, 특히 고속 기계와 중

장비를 다루는 환경에서 빈번하게 발생한다[1,2]. 기존의 단일 카메라 시스템은 시야의 제한으로 인해 예방 효과에 한계가 있었다[3,4]. 그러나 카메라 기술의 발전으로 손의 위치와 움직임을 더 정확하게

\* 동의대학교 컴퓨터공학과, tswang@office.deu.ac.kr,

\*\* 동의대학교 스마트자동차연구소 연구소장, swjang@deu.ac.kr(교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업임(IITP-2024-RS-2020-II201791, 50%). 또한 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학연협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다(50%).

추적할 수 있게 되었으며, 본 연구는 저비용 듀얼 카메라를 활용하여 실시간으로 손과 위험 구역 간의 거리를 모니터링하는 사고 예방 시스템을 제안한다.

## II. 관련 연구

산업 현장에서 손 끼임 사고를 방지하기 위해 다양한 도구와 기술이 사용된다. 주요 방법으로는 Machine Guards, Trip-nip Bars, Safety Light Curtains, Presence-Sensing Devices 등이 있다[5-8]. 단일 카메라 기반 예방 기술은 저비용 및 설치 용이성의 장점이 있지만, 깊이 정보 부족과 정확한 위치 추적의 어려움이라는 한계를 가진다[3].

## III. 시스템 설계

본 장에서는 제안한 시스템의 설계 과정을 설명한다. 두 대의 카메라를 통해 작업자의 손과 기계를 실시간 모니터링하여, 손의 움직임을 바탕으로 위험 영역 침범 여부를 판단해 손 끼임 사고를 예방한다.

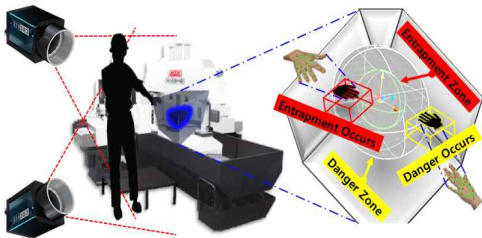


그림 1. 시스템 구성도  
Fig. 1. System Configuration Diagram

그림 1의 시스템 구성도에 따르면, 두 카메라는 서로 다른 각도로 배치되어 작업자의 손을 추적하며, 저가형 카메라를 사용하여 비용을 절감하고 사각지대를 최소화했다.

그림 2의 흐름도는 시스템의 동작 과정을 설명한다. (1) 두 카메라가 실시간 프레임 수집하고, (2) 수집된 영상에서 손의 Pose Estimation을 수행한다. (3) 손가락 관절 좌표를 추출해 3D 좌표로 변환하고, (4) 이를 2D 이미지로 투영하여 위험영역 침범

을 실시간으로 모니터링한다. 손이 위험영역에 접근하면 즉시 경고가 발생된다.

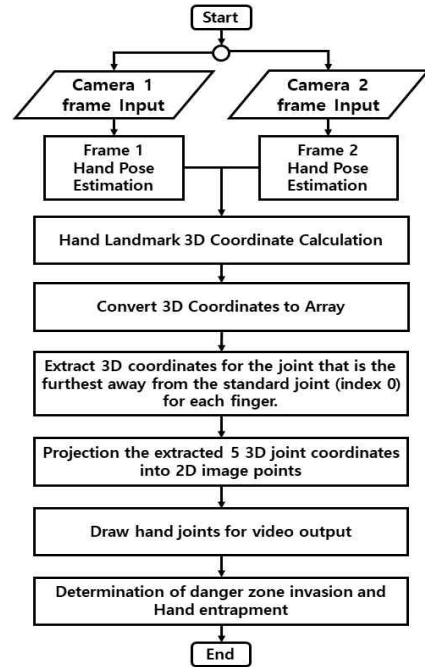


그림 2. 시스템 흐름도  
Fig. 2. System Flow Chart

위험영역과 손 끼임 영역은 xyz 축 기준으로 육면체로 정의되며, 손의 3D 좌표가 해당 영역 내에 있으면 경고가 발생된다. MediaPipe를 사용해 손의 21개 관절을 추정하며, 수식을 통해 두 카메라로부터 추출한 3D 좌표를 계산하고 이를 픽셀 좌표로 변환해 실시간 모니터링을 수행한다.

$$X_i = \frac{X_{i0} + X_{i1}}{2}, Y_i = \frac{Y_{i0} + Y_{i1}}{2}, Z_i = \frac{Z_{i0} + Z_{i1}}{2} \quad (1)$$

$$x_{px} = \min([x \times img\_width], img\_width - 1) \quad (2)$$

$$y_{px} = \min([y \times img\_height], img\_height - 1)$$

$$Inside = (|x_p| \leq x) \wedge (|y_p| \leq y) \wedge (|z_p| \leq z) \quad (3)$$

## IV. 실행 결과

그림 3은 시스템 실행 결과를 보여준다. 상단 이미지는 손 관절을 검출하고 가장 먼 관절을 표시한

것이며, 하단 왼쪽 이미지는 손이 위험영역에 침범하지 않은 상태로, 경계선이 흰색으로 표시된다. 오른쪽 이미지는 손이 위험영역에 침범한 상태로, 경계선이 파란색으로 출력된다.

실험을 통해 저가형 카메라로 평균 19.5 FPS, 고성능 카메라로는 평균 30 FPS를 기록하며, 실시간 성능을 입증했다.

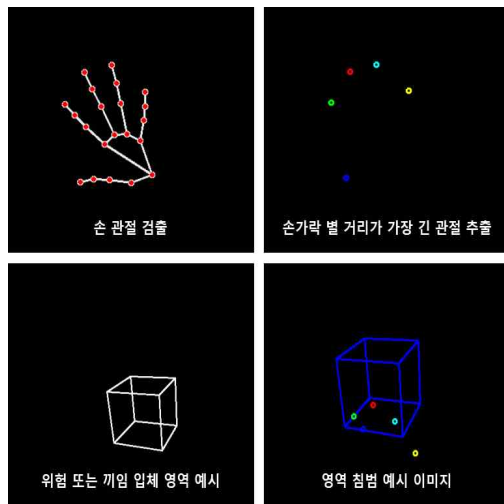


그림 3. 시스템 실행 결과  
Fig. 3. System Execution Results

## V. 결 론

본 연구에서는 듀얼 카메라를 활용한 손 끼임 사고 예방 시스템을 설계하고 구현했다. 두 대의 저가형 카메라를 통해 작업 공간을 다양한 각도에서 모니터링하여, 단일 카메라의 시야 제한과 깊이 정보 부족 문제를 개선할 수 있었다. 실험 결과, 시스템을 통해 실시간으로 위험영역에 대한 침범 여부를 판단하여 손 끼임 사고를 예방할 수 있고, 고성능 카메라 사용 시 30 FPS로 실시간 모니터링이 가능하다. 향후 연구에서는 추가 카메라를 도입해 시스템의 정확도를 높일 계획이다.

## 참 고 문 헌

[1] T. B. Hansen and O. Carstensen, "Hand injuries in agricultural accidents", *The Journal of Hand Surgery: British & European*, vol. 24, no. 2, pp.

190-192, April 1999. DOI: 10.1054/JHSB.1998.0189

[2] R. M. Reyes, J. de la Riva, A. Maldonado, and A. Woocay, "Association between human error and occupational accidents' contributing factors for hand injuries in the automotive manufacturing industry", *Procedia Manufacturing*, vol. 3, pp. 6498-6504, 2015. DOI: 10.1016/j.promfg.2015.07.936

[3] Y. H. Bae, Y. T. Lee, and H. C. Kim, "A Study on Worker Risk Reduction Methods using the Deep Learning Image Processing Technique in the Turning Process", *The Journal of the Korean Society of Manufacturing Process Engineers(KSMPE)*, vol. 20, no. 12, pp. 1-7, December 2021. DOI: 10.14775/ksmpe.2021.20.12.001

[4] F. Zhao, Z. Cao, Y. Xiao, J. Mao and J. Yuan, "Real-Time Detection of Fall From Bed Using a Single Depth Camera", *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 3, pp. 1018-1032, July 2019. DOI: 10.1109/TASE.2018.2861382.

[5] SafetySkills Homepage, "Prevent Workplace Injuries: A Guide to Machine Guarding". September 2018. URL: <https://safetyskills.com/machine-guarding>

[6] E. Milnes, "A Method for Testing Pressures on Fingers in Entrapment Haz-ards(Contemporary Ergonomics 2006)", 1st Edition Taylor & Francis, vol. 1, July 2020. DOI: 10.1201/9781003072072

[7] A. Siddharth, et al. "Active safety envelopes using light curtains with probabilistic guarantees", arXiv, July 2021. DOI: 10.48550/arXiv.2107.04000

[8] Canadian Occupational Safety Homepage, "Selecting presence-sensing devices for machine guarding". September 2017. URL: <https://www.thesafetymag.com/ca/topics/technology>

# 3D LiDAR 환경에서의 평면 탐지 필터 기반 동적 객체 분할

이민재\*, 배지호\*\*, 이수원\*\*

## Moving Object Segmentation in 3D LiDAR Environment based on Plane Detection Filter

Minjae Lee\*, Jiho Bae\*\*, and Suwon Lee\*\*

### 요 약

최근 자율주행 기술의 급속한 발전에 따라 3D LiDAR 데이터를 기반으로 한 동적 객체 분할의 중요성이 더욱 커지고 있다. 자율주행 시스템은 주변 환경을 정확히 인식하고 장애물 및 객체의 움직임을 신속히 판단하는 능력이 필수적이다. 그러나 기존 연구에서는 비동적 객체를 동적 객체로 잘못 탐지하는 문제가 있어, 동적 객체 분할의 정확성을 개선할 필요가 있다. 본 연구에서는 평면 탐지 기법을 활용하여 비동적 객체를 효과적으로 제거함으로써 3D LiDAR 데이터에서 잘못 탐지된 동적 객체 분할 결과를 개선하는 방법을 제시한다. 제안된 방법은 자율주행 시스템의 안전성을 강화하는 데 기여할 것으로 기대된다.

### Abstract

With the rapid advancement of autonomous driving technology, the importance of dynamic object segmentation based on 3D LiDAR data has grown significantly. Autonomous driving systems require accurate perception of their surroundings and quick assessment of the movement of various obstacles. However, existing studies have faced issues such as misclassifying static objects as dynamic ones, highlighting the need to improve dynamic object segmentation accuracy. In this study, we propose a method to enhance the segmentation of dynamic objects in 3D LiDAR data by effectively removing static objects using plane detection techniques. The proposed method is expected to enhance the safety of autonomous driving systems.

### Key words

moving object segmentation, plane detection, 3D LiDAR, LMNet

## 1. 서 론

자율주행 기술이 발전함에 따라, 3D LiDAR 데이터를 활용한 동적 객체 분할(Moving Object Segmentation, MOS)의 중요성은 점점 더 커지고 있다. 자율주행 시스템은 실시간으로 주변 환경을 정

확히 인식하고, 다양한 장애물과 객체의 움직임을 구분하는 능력이 필수적이다. 최근 연구에서 제안된 LiDAR MOS Network(LMNet)[1]는 동적 객체 분할에서 준수한 성능을 보인다. 그러나 LMNet은 LiDAR 데이터의 그림자 영역이나 LiDAR 센서로부터 거리가 먼 지표면과 같은 비동적 요소들이 동적

\* 경상국립대학교 컴퓨터공학과, wjdchs0129@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, dream\_cacao\_jh@naver.com

\*\*\* 경상국립대학교 컴퓨터공학과, leesuwon@gnu.ac.kr (교신저자)



객체로 오탐지되는 문제가 있다. 예를 들어, 그림자 부분이 움직이면서 그 영역이 마치 동적 객체인 것처럼 잘못 탐지되거나, 먼 거리에 있는 지표면 영역이 노이즈를 포함하여 동적 객체로 오인되는 경우가 발생한다. 이는 자율주행 시스템의 정확한 환경 인식에 부정적인 영향을 미칠 수 있으며, 안전성의 측면에서 매우 중요하다. 본 연구는 이러한 오탐지 문제를 해결하기 위해 평면 탐지 기법을 활용하여 비동적 객체를 더욱 정교하게 구분하고, 동적 객체 분할의 정확성을 개선하는 방법을 제안한다. 이를 통해 자율주행 시스템의 인지 성능을 개선하고, 신뢰할 수 있는 주행 환경을 제공하는 데 기여할 것으로 기대된다. 기존의 많은 연구들과 같이, 본 연구도 Semantic KITTI[2] 데이터셋의 벨로다인 LIDAR 데이터 시퀀스를 활용하였다.

## II. 평면 탐지

LiDAR 데이터의 특성상, 동적 객체(차량, 사람, 자전거 등)가 이동할 때 해당 객체에 생기는 그림자 또한 함께 이동하게 된다. LMNet을 통한 출력 결과는 그림 1(a)에서 확인할 수 있듯이, 실제 차량뿐만 아니라 차량에 의해 생긴 그림자까지 동적 물체로 분류된 모습을 보인다. 본 논문에서는 LiDAR 데이터에서 그림자 영역이 대부분 지표면에서 발생한다는 점에 착안하여, RANSAC 기반의 평면 탐지 기법을 도입함으로써 이러한 오분류된 결과를 정제하는 방법을 제안한다. 평면 탐지를 이용하면 불필요한 그림자 영역을 탐지할 수 있으며 그림 1(b)는 평면 탐지를 수행한 결과로, 지표면을 빨간색으로, 나머지 영역을 흰색으로 시각화한 예시를 보여준다.

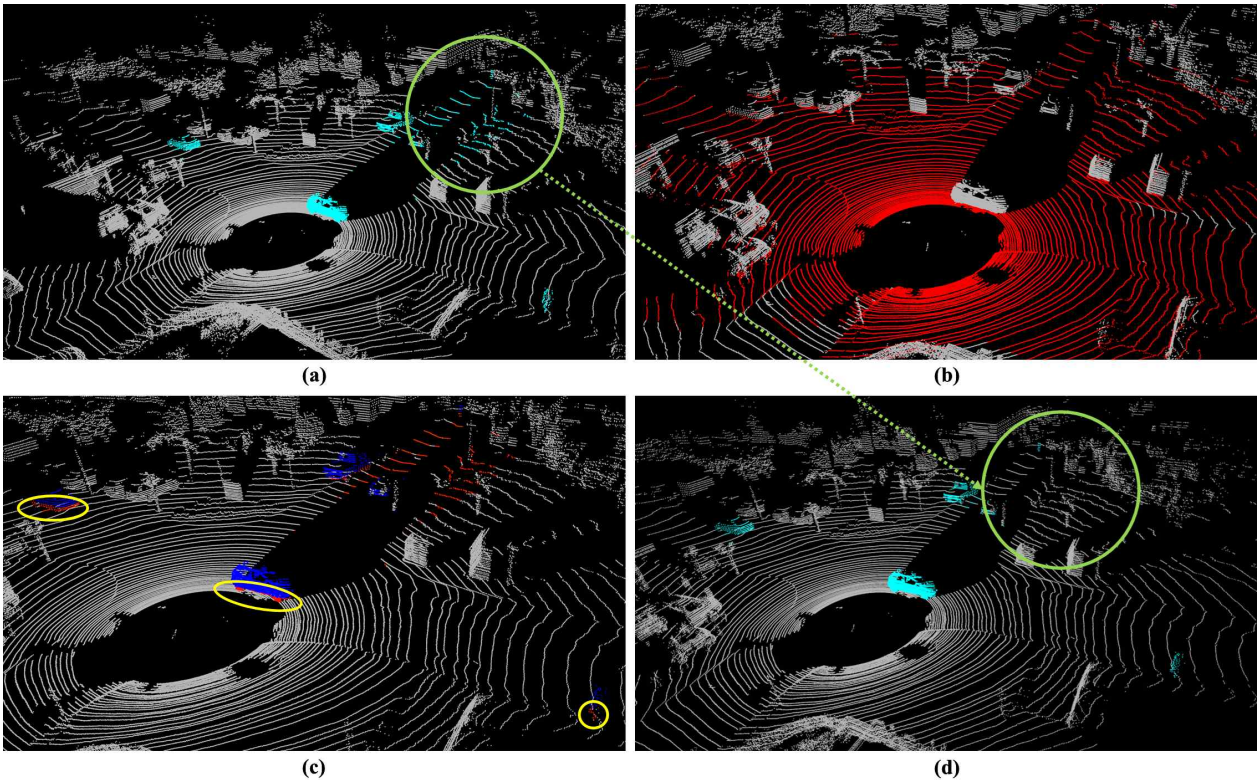


그림 1. 평면 탐지를 기반으로 한 동적 물체 분할 개선 과정. (a) LiDAR 데이터에서 그림자 영역을 동적 물체로 잘못 분류한 결과; (b) RANSAC 기반 평면 탐지 결과; (c) 동적 물체 분할 결과와 평면 탐지 결과의 단순 교차; (d) 제안된 방법을 적용하여 그림자 영역을 제거하고 바퀴 및 발 부위의 동적 특성을 유지한 결과.

Fig. 1. Enhancement process of moving object segmentation based on plane detection. (a) Result of misclassifying shadow areas in LiDAR data; (b) RANSAC-based plane detection result; (c) Simple intersection of moving object segmentation and plane detection results; (d) Result of the proposed method, removing shadow areas while preserving the dynamic characteristics of wheels and feet.

### III. 프로젝션 거리 기반 필터

II에서 제시한 평면 탐지 결과를 바탕으로 제거 대상을 설정할 수 있다. 그림 1(c)에서 파란색 영역은 유지할 부분이며, 빨간색 영역은 제거할 대상으로 설정된다. 그러나 단순히 평면 탐지 결과만을 바탕으로 오분류 영역을 제거할 경우, 그림 1(c)의 노란색 타원으로 표시된 것처럼 실제로 제거하지 말아야 할 영역(예: 자동차나 자전거의 바퀴, 사람의 발)까지 제거하는 문제가 발생할 수 있다.

이 문제를 해결하기 위한 한 가지 방법은 LMNet의 결과를 단순 이진 정보(Binary Mask)가 아닌 의미론적 정보(Semantic Mask)를 포함하도록 확장하여, 지표면 라벨을 갖는 객체들만 제거하는 것이다. 그러나 이 방법은 새로운 딥러닝 모델의 설계와 학습이 필요하므로 자원 소모가 크다. 이에 본 논문에서는 보다 직관적인 해결 방법으로, 프로젝션 거리 기반 필터를 제안한다. 그림 2는 평면 탐지를 통해 얻은 프레임별 평면 모델에 정분류된 포인트 클라우드(하늘색)와 오분류된 포인트 클라우드(빨간색)를 프로젝션한 결과로, 각각 노란색과 초록색으로 투영된 모습을 보여준다. 즉, 제거 대상인 그림자 영역(빨간색)의 경우 초록색 포인트로 투영되고, 유지 대상(하늘색)의 경우 노란색으로 투영된다.

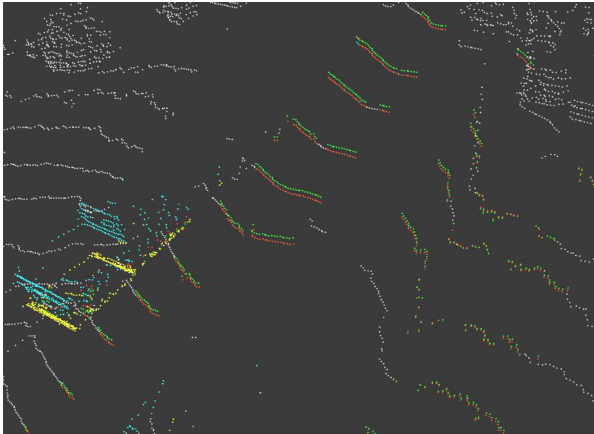


그림 2. 제거 대상(그림자 영역)과 제거 불필요 대상(동적 객체의 지표면 근처 점)에 대한 비교

Fig. 2. Comparison between removal targets (shadow areas) and non-removal targets (points near the ground of dynamic objects)

필터링의 주체가 되는 초록색 점 근처에 노란색 포인트가 존재하면 초록색 점의 근원인 빨간색 점은 유지(하늘색), 존재하지 않는다면 그 근원인 빨간색 점은 제거(흰색)하도록 한다. 색상별 (빨간색, 초록색, 하늘색, 노란색) 포인트 클라우드를 각각  $\mathbf{R} = \{r_1, \dots, r_n\}$ ,  $\mathbf{G} = \{g_1, \dots, g_m\}$ ,  $\mathbf{S} = \{s_1, \dots, s_m\}$ ,  $\mathbf{Y} = \{y_1, \dots, y_m\}$ , ( $\forall r_i, g_i, s_j, y_j \in \mathbf{R}^3$ ) 라고 하였을 때, 다음과 같은 과정을 거친다.

$$d_{ij} = \|g_i - y_j\|_2, \forall g_i \in \mathbf{G}, \forall y_j \in \mathbf{Y} \quad (1)$$

$$r_i \in \begin{cases} \mathbf{S}, & \text{if } \exists g_i \text{ s.t. } d_{ij} < \delta \\ \emptyset, & \text{otherwise} \end{cases} \quad (2)$$

$$N_k(g_i) = \{y_j \in \mathbf{Y} \mid d_{ij} \leq \delta, \forall y_j \in \text{NNs}\} \quad (3)$$

수식 (1)을 사용하여 초록색 포인트( $g_i$ )와 노란색 포인트( $y_j$ ) 사이의 거리( $d_{ij}$ )를 계산하고, 수식 (2)를 사용하여 특정 임계값( $\delta$ ) 이상인 거리에 위치한 빨간색 포인트 클라우드( $r_i$ )를 흰색( $\emptyset$ )으로 변경(제거)하며, 그 외의 경우에는 하늘색( $s$ )으로 변경(유지)하는 방식으로 오탐지 영역을 정제한다.

그러나 거리를 측정하는 과정에서 단순히 초록색 포인트( $g_i$ )와 노란색 포인트( $y_j$ ) 사이의 거리( $d_{ij}$ )를 측정할 때 모든 경우의 수에 대해 조사하게 된다면  $O(nm)$ 의 시간복잡도를 갖게 되므로  $n$ 과  $m$ 이 커진다면 매우 비효율적일 것이다. 이를 해결하기 위해 수식 (3)을 따르는 KD-Tree 기반의 근접 이웃(Nearest Neighbors)을 이용하였다.  $O(n \cdot \log m)$ 으로 KD-Tree를 구성하고  $O(m \cdot \log m)$ 으로 탐색함으로써 빠른 속도를 보장할 수 있게 하였다.

### IV. 실험 및 평가

본 논문에서는 수식 (4)에 따른 Intersection over Union (IoU) 지표를 사용하여 성능을 평가하였으며, 실험에는 Semantic KITTI의 벨로다인 시퀀스 8번 데이터를 활용하였다.

$$IoU = \frac{TP}{TP + FP + FN}, \quad (4)$$

where  $P$ : Moving,  $N$ : Static

최적의 거리 임계값  $\delta$ 를 구하기 위해  $\delta$  값을 변화시키며 IoU 변화를 분석하였다. 그림 3은  $\delta$  값에 따른 IoU 변화를 그래프로 나타낸 것이며, 최적의 IoU는  $\delta = 0.2$  또는  $\delta = 0.3$ 에서 달성되는 것을 확인하였다. 이후 실험에서는  $\delta = 0.2$ 를 기준으로 다양한 방법에 필터를 적용하여 성능을 비교 분석하였다.

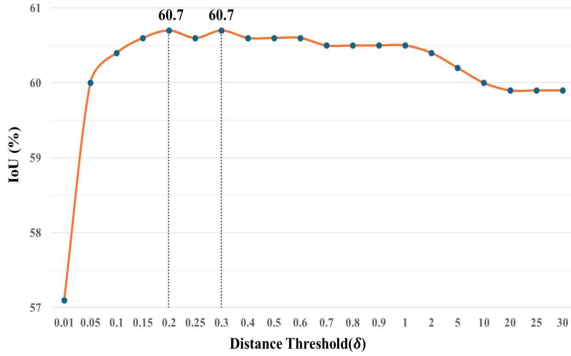


그림 3 임계값  $\delta$ 에 따른 IoU 그래프  
Fig. 3. IoU graph according to threshold  $\delta$

표 1은 다양한 MOS 방법론에 따른 IoU 성능을 비교한 결과를 보여준다. 여기에는 (i) LMNet 방법별 성능, (ii) LMNet에 제안한 필터를 적용한 성능, (iii) 4DMOS의 성능, (iv) 4DMOS에 필터를 적용한 성능이 포함된다. LMNet의 저자가 제안한 방법(N=1, N=8, N=8+Semantics)과 4DMOS에 모두 필터를 적용한 결과, 유의미한 성능 개선이 나타났다.

표 1. 동적 객체 분할의 방법론 별 IoU 개선  
Table 1. IoU enhancement for different MOS methods

방법	IoU(%)
LMNet(N=1)[1]	52.0
LMNet(N=1) + 제안 필터	<b>60.7</b>
LMNet(N=8) + 제안 필터	<b>62.6</b>
LMNet(N=8+Semantics)[1]	62.5
LMNet(N=8+Semantics) + 제안 필터	<b>65.1</b>
4DMOS[3]	65.2
4DMOS + 제안 필터	<b>71.5</b>

특히, LMNet의 N=8 모델의 경우 기존에 정보가 부족해 직접 학습한 후 필터를 적용한 결과 IoU가 62.6%를 기록하였다. 이는 필터를 적용하지 않은 N=8+Semantics 모델(62.5%)보다 높은 수치이다. 또한, N=8+Semantics 모델에 필터를 적용했을 때는 2.6%p 증가하였고, N=1 모델에서는 8.7%p로 가장 큰 상승폭을 보였다. 4DMOS 모델에도 필터를 적용한 결과 6.3% 포인트의 성능 향상이 확인되었다.

## V. 논의 및 향후 연구

본 논문에서 제시한 ‘평면 탐지 후 프로젝션 거리 기반 필터 기법’은 대부분의 그림자 영역이 지표면에 발생한다는 점에서 착안하여 실험되었고 개선된 결과를 확인할 수 있었다. 하지만, 실제로는 건물의 벽, 나무 등과 같이 지표면이 아닌 물체에도 일부 그림자가 발생할 수 있다. 그림 4는 (i) 지표면이 아닌 영역에 생긴 그림자와 (ii) 평면 탐지로 필터링되지 않은 오분류 포인트 클라우드를 보여준다. (i)의 경우에는 기존 방법론으로 개선이 어렵기 때문에 새로운 접근 방식이 필요할 것이다. 한편, (ii)의 경우는 평면 탐지 과정에서 평면으로 간주되지 않아 발생하는 문제로, 이는 그림 5에서 더 명확히 확인할 수 있다. 그림 5는 그림 4의 자전거 운전자 2명을 측면에서 바라본 모습으로, RANSAC 기반으로 추론된 평면 모델이 실제 지표면보다 기울어지게 추론된 경우를 나타낸다. RANSAC 알고리즘의 무작위성으로 인해 평면이 잘못 추정되면, 그림 5의 노란색 평면처럼 지표면과 다르게 모델링되며, 그 결과 필터링이 되어야 할 하늘색 점들이 그대로 남아 동적 객체로 분류된다. 이는 개선될 여지가 있으며 추후 MOS 분야에 더 기여할 것으로 생각된다.

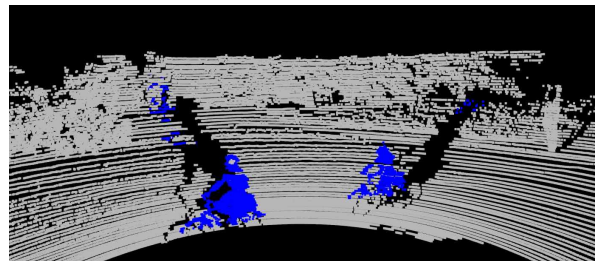


그림 4 필터링 되지 않은 그림자 영역 포인트 클라우드  
Fig. 4. Unfiltered shadow areas point cloud

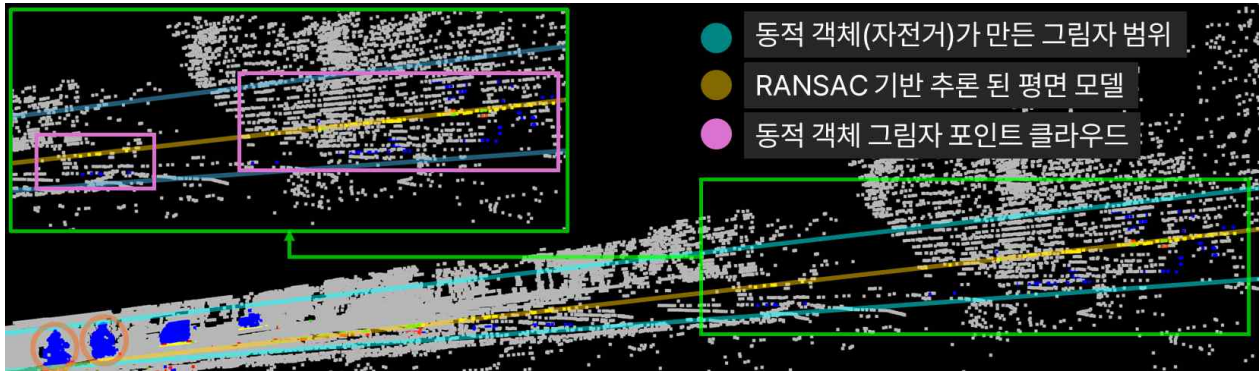


그림 5 RANSAC 오류로 인한 잘못 추론된 평면(노란색)과 필터링 되지 않은 포인트 클라우드(파란색)  
Fig. 5. The incorrectly inferred plane due to RANSAC errors (yellow) and the unfiltered point cloud (blue)

## VI. 결 론

본 논문에서는 3D LiDAR 데이터를 이용한 동적 객체 분할에서 발생하는 오탐지 문제를 해결하기 위해 평면 탐지 기법을 제안하였다. 기존의 MOS 모델인 LMNet은 그림자 영역이나 지표면과 같은 비동적 요소를 동적 객체로 오인하는 문제가 있었으며, 이는 자율주행 시스템의 인지 성능에 큰 영향을 미쳤다. 이를 개선하기 위해 RANSAC 기반의 평면 탐지 기법을 적용하여 이러한 오탐지 영역을 효과적으로 제거하였다. 또한, 평면 탐지와 수식 기반의 필터링 기법을 결합하여 동적 객체와 비동적 객체를 더욱 정교하게 분리함으로써 MOS의 성능을 향상시켰다. 실험 결과, 본 논문에서 제안한 방법은 기존 LMNet 대비 IoU 성능에서 큰 향상을 보였으며, 임계값( $\delta$ )을 조정함으로써 개선된 성능을 얻을 수 있음을 확인하였다. 그러나 평면 탐지 기법은 건물의 벽이나 나무와 같은 비지표면 영역의 그림자까지 완벽하게 해결하지 못하는 한계가 있었다. 또한, RANSAC 기반의 평면 추론 과정에서 발생하는 무작위성으로 인해 오탐지가 발생할 수 있음을 확인하였으며, 이를 개선하기 위한 추가적인 연구가 필요하다.

## 참 고 문 헌

[1] Chen, X., Li, S., Mersch, B., Wiesmann, L., Gall, J., Behley, J., & Stachniss, C. (2021). Moving object segmentation in 3D LiDAR data: A

learning-based approach exploiting sequential data. *IEEE Robotics and Automation Letters*, 6(4), 6529-6536.

- [2] Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., & Gall, J. (2019). Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 9297-9307).
- [3] Mersch, B., Chen, X., Vizzo, I., Nunes, L., Behley, J., & Stachniss, C. (2022). Receding moving object segmentation in 3d lidar data using sparse 4d convolutions. *IEEE Robotics and Automation Letters*, 7(3), 7503-7510.

# GPT-4o 프롬프트 기반 파노라마 생성을 통한 공간 디자인 및 활용

시종욱\*, 조재혁\*\*, 김성영\*\*

## Space Design and Utilization through GPT-4o Prompt-Based Panoramic Generation

Jongwook Si\*, Jaehyuk Jo\*\*, and Sungyoung Kim\*\*

메타버스 기술의 발전으로 가상 현실과 증강 현실이 점차 중요해지면서, 이러한 가상 공간을 더욱 현실감 있게 구성하기 위한 생성형 인공지능의 역할이 크게 부각되고 있다[1]. 그러나, 이러한 생성형 인공지능의 입력으로 하는 프롬프트에 따라 결과의 품질이 많이 달라지며, 수동으로 프롬프트를 작성하기에는 많은 시간과 노력이 소요된다. 따라서, 본 논문에서는 GPT-4o[2]를 활용하여 사용자 요청에 맞춘 프롬프트를 생성하고, 이를 Stable Diffusion[3] 모델에 적용하여 카페나 쇼핑몰 등의 3D 가상 공간을 구성하는 방법을 제시한다. Stable Diffusion은 다양한 체크포인트(ckpt)와 LoRA를 활용하여 유연하게 파노라마 영상을 생성할 수 있으며, 실험 결과 여러 체크포인트를 적용한 경우 현실감 있는 영상을 성공적으로 생성할 수 있었다. 특히, 파노라마 생성에 특화된 LoRA를 추가로 적용했을 때 더욱 향상된 결과를 얻을 수 있음을 확인하였다. 생성된 파노라마는 초고해상도 기술을 통해 8K 해상도로 확장되었으며, 이를 Unity와 Unreal Engine에 적용하여 실제와 유사한 3D 가상 공간을 구축할 수 있음을 실험을 통해 입증하였다. 두 엔진의 실험 결과, 그래픽 엔진에 따라 선명도와 현실감에서 차이는 있었지만, 모두 실제와 매우 유사한 공간을 효과적으로 재현할 수 있었다. 이러한 접근법은 수동적인 가상 공간 구축 과정에서 발생하는 시간과 노력을 크게 절감하면서도, 더 정교하고 사실적인 3D 가상 공간을 생성할 수 있음을 증명한다.

### Key words

Space Design, GPT, Prompt, Panoramic Generation, O2O Stores

### 참고문헌

- [1] D. Kim, J. Si, and S. Kim, "Generation of High-Resolution Virtual Interview Rooms Using LoRA-applied Stable Diffusion and Real-ESRGAN", Proceedings of KIIT Conference, May 2024.
- [2] gpt-4o, <https://openai.com/index/hello-gpt-4o/>
- [3] R. Rombach, B. Ommer, *et al.*, "High-Resolution Image Synthesis with Latent Diffusion Models", In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684-10685, Jun. 2022.

---

\* 국립금오공과대학교 컴퓨터·AI융합공학과 박사과정

\*\* 국립금오공과대학교 컴퓨터공학과 학부생

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

※ 본 연구는 2024년도 과학기술정보통신부와 한국전파진흥협회의 메타버스랩지원사업에 의한 연구임

# 해면골 두께가 저강도 경두개 집속 초음파에 미치는 영향에 대한 시뮬레이션 연구

유지훈\*, 임철기\*\*, 서현\*

## A Simulation Study on Effect of Cancellous bone thickness on Low-Intensity transcranial focused ultrasound

Jihun Yu\*, Cheolki Im\*\* and Hyeon Seo\*

### 요 약

저강도 집속 초음파 기술(tFUS : transcranial Focused UltraSound)은 뇌의 특정 영역에 에너지를 집중시켜 신경을 조절하는 기술이다. tFUS 사용 시에는 두개골의 불균일성으로 인해 에너지 전달 예측이 어려워 시뮬레이션이 필수적이며 CT 또는 MRI를 활용한다. 이때 두개골 중간에 위치한 해면골의 존재가 중요한 요인으로 보고되었다. 본 연구팀은 해면골이 초음파 에너지 전달에 미치는 영향을 분석하기 위해, 두개골의 단순 모형인 구를 사용하여 다른 요인의 영향을 배제하고 시뮬레이션을 통해 해면골의 영향을 조사했다. 두개골의 두께는 5mm와 7mm로 설정했으며 해면골의 비율을 0%에서 80%까지 10% 간격으로 증가시켰다. 두개골 두께가 7mm일 때는 5mm에 비해 전반적으로 낮은 최대 음압이 관측되었다. 해면골 비율이 증가함에 따라 최대 음압은 잠시 감소한 후 증가하는 경향을 보였고, 집속 부피는 해면골 비율이 낮을 때 잠시 증가했다가 다시 감소하는 경향을 나타냈다. 이는 해면골 비율이 낮을 때 두개골의 불균일성이 에너지 전달을 방해하지만, 일정 비율을 초과하면 에너지 전달이 효과적임을 시사한다.

### Abstract

Low-intensity focused ultrasound technology (tFUS: transcranial Focused UltraSound) is a technique that focuses energy on specific brain regions to modulate neural activity. When using tFUS, predicting energy delivery is challenging due to skull heterogeneous, making simulation essential, utilizing CT or MRI. The presence of cancellous bone located in the middle of the skull has been identified as a critical factor. Our research team investigated the impact of cancellous bone on ultrasound energy transmission by using a simplified spherical model of the skull to neglect other factors and conducted simulations. The skull thickness was set as 5mm and 7mm, with the cancellous bone ratio increasing from 0% to 80% in 10% increments. A generally lower peak acoustic pressure was observed when the skull thickness was 7mm compared to 5mm. As the cancellous bone ratio increased, the peak acoustic pressure showed a tendency to decrease momentarily before increasing, while the focal volume increased slightly at low cancellous bone ratios and then decreased. This suggests that at low cancellous bone ratios, skull heterogeneous hinders energy transmission, but beyond a certain ratio, energy delivery becomes more effective.

### Key words

transcranial Focused UltraSound, cancellous bone, three layer model, acoustic simulation

\*경상국립대학교 IT공과대학 컴퓨터공학과 dbwlgns0923@naver.com, hseo0612@gnu.ac.kr (교신저자)

\*\*경상국립대학교 기초과학연구소 chelgj47@gmail.com

※ 본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (RS-2023-00280241)

## I. 서 론

뇌 신경조절술(Neuromodulation)은 뇌 특정 부위에 외부 자극을 가함으로써 신경 세포의 활성화 정도를 조절하는 기술로 주로 전기나 자기를 통해 자극을 전달한다[1].

이러한 방법은 낮은 공간 해상도를 가지기 때문에 국소 부위를 자극하기 어렵지만 tFUS 기술은 전기나 자기 대비 높은 공간 해상도를 가진다는 장점이 있다[1]. 고강도 집속 초음파의 경우 초음파가 집속된 영역에서 발생한 열을 이용한 조직 절제술, 저강도 집속 초음파의 경우 신경 조절에 사용되며 안전성 또한 입증됐다.

tFUS 사용 시 목표 영역에 원하는 양의 에너지를 집속 전달하는 것이 중요하지만 두개골로 인해서 이를 정확하게 예측하기 어렵다. 이는 두개골 내의 불균일성에 의해 두개골을 통과하는 초음파의 전파 과정에서 반사, 굴절 및 감쇠가 일어나기 때문이다. 따라서 환자에게 적용하기 전 시뮬레이션 과정이 필수적이며, 두개골 구조를 상세하게 보여주는 CT 영상을 주로 사용한다[1-3].

그러나 CT 영상은 방사선 노출 등의 이유로 획득하기 어렵고 MRI가 CT 사진보다 더 상세하게 뇌 조직을 분별할 수 있어 최근 MRI를 이용한 tFUS 시뮬레이션 연구들이 발표되었다. 예를 들면, Y. Huang은 MRI로 상세한 뇌 구조를 모델링하고 뇌 조직별로 tFUS 시뮬레이션 결과에 미치는 영향을 조사했다[4]. 이때 두개골은 해면골(cancellous bone)을 하나의 층으로 그 외 치밀뼈(compact bone)를 다른 층으로 설정한 후 균일한 물성으로 부여해 시뮬레이션을 진행한다. 최근 본 연구팀은 MRI를 사용한 tFUS 시뮬레이션에서 두개골에 균일한 물성을 주는 방식과 치밀뼈와 그 사이의 해면골을 고려해 3개의 레이어로 두개골 모델을 설정한 방식 간의

표 1. 물과 두개골의 음향 특성

Table 1. acoustic properties of water and skull

	음속[ms <sup>-1</sup> ]	밀도[kgm <sup>-3</sup> ]	감쇠계수[dBMHz <sup>-1</sup> cm <sup>-1</sup> ]
물	1482	1000	4.2050×10 <sup>-5</sup>
두개골(inner, outer)	2813.692	1908	27.2765
두개골(cancellous)	2117.553	1178.33	20.4579

CT 대비 발생하는 오차를 보고했다[5]. 그 결과 해면골의 물성값에 따라 결과가 다양하게 변하는 현상을 관찰했고 tFUS 시뮬레이션에서 해면골의 중요성을 시사했다.

해면골의 존재 여부가 tFUS 시뮬레이션에 있어 중요한 요인지만, 그 영향은 아직 체계적으로 조사되지 않았다. 이에 본 연구에서는 두개골의 단순 모형인 구를 사용해서 다른 요인을 배제하면서 해면골의 두께가 tFUS 시뮬레이션에 미치는 영향을 분석한다.

## II. 연구 방법

tFUS 유도 음압 계산을 위해 MATLAB의 오픈 소스인 k-Wave MATLAB toolbox[6]를 이용했다. 최대한 사람의 머리와 비슷한 곡률의 단순 구형 모델을 구현하고[7] 0.42mm<sup>3</sup> 크기의 등방성 복셀로 500 × 500 × 500 그리드 공간을 구성했다. 0.5MHz의 단일 요소 집속 트랜스듀서(Blatek, Inc., State College, PA)의 규격에 따라 모델링했으며, 진단과의 영향을 무시하기 위해 두개골과 수직으로 10mm 떨어진 지점에 배치했다[8].

해면골은 SimNIBS-4.1의 charm[9]을 통해서 식별되고 이후 IT'IS의 데이터 셋을 참고해서 두개골과 물의 물성(음속, 밀도, 감쇠계수)을 표1과 같이 설정했다[10]. 마지막으로 두개골 내의 뇌 조직은 tFUS에 영향은 거의 주지 않기 때문에 물로 대체했다[7]. 전체 두개골 두께(outer + cancellous + inner)를 5mm와 7mm로 설정하고 두개골 중심에서부터 해면골의 두께를 전체 두개골 두께의 0%부터 80%까지 10% 간격으로 증가시키며 3D 시뮬레이션을 진행했다 (그림 1).

## III. 결 론

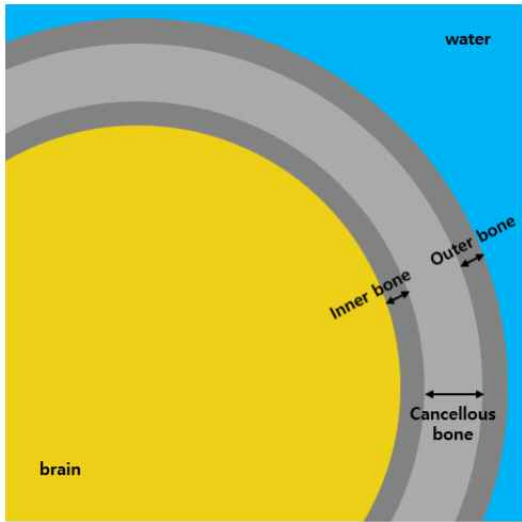


그림 1. 시뮬레이션 세팅  
Fig. 1. simulation setting

본 논문에서는 두개골 내에 존재하는 해면골이 tFUS 시뮬레이션 결과에 미치는 영향을 최대 음압, FWHM(Full Width at Half Maximum)의 장축, 단축, 부피 그리고 초점거리 측면에서 분석했다(그림 2). FWHM이란 어떤 함수에서 최댓값의 절반이 되는 두 지점의 거리 차이로 본 논문에서 이를 활용하여 집속 영역의 부피를 정량화했다.

비교를 위해 두개골이 없는 순수 물에서 집속되는 음압 분포를 계산하였고, 최대 음압은 8.1899MPa, FWHM의 장축, 단축, 부피는 각각 55.43mm, 9.06mm, 211.89mm<sup>3</sup> 그리고 초점 거리는 36.62mm였다. 이를 바탕으로 두개골 모델에서의 결과를 비교할 때 최대 음압은 물 대비 10% ~ 33% 정도에 해당하는 수치로 두개골에 의해 감소했으며 초점 거리도 순수 물에서 결과 대비 감소된 수치를 보여준다. 이와 반대로 FWHM의 장축, 단축, 부피는 순수 물에서 결과 대비 모두 증가했다.

최대 음압의 경우 두개골 두께가 7mm일 때 5mm인 경우에 비해 전반적으로 더 낮은 음압이 관찰되었다. 동일한 두개골 두께에서 해면골의 비율이 커질수록 최대 음압은 잠시 감소한 후 다시 증가하는 경향을 보였다.

FWHM 부피는 해면골의 비율이 커질수록 일반적으로 감소하지만 5mm의 경우 해면골의 비율이 20%, 80% 그리고 7mm의 경우 10%, 50%일 때 이러한 경향을 잠시 벗어난다. FWHM의 장축, 단축에

서도 비슷한 경향이 나타난다.

초점 거리는 5mm인 경우 0%일 때 26.54mm에서 20%~30%일 때 28.22mm까지 증가하며 이후 서서히 감소하다가 60%가 넘어가면 25.7mm에 수렴하는 현상을 보이고 7mm인 경우 0%일 때 27.8mm부터 증가와 감소를 반복하다 80%일 때 24.86mm로 전반적으로 감소하는 경향을 보인다.

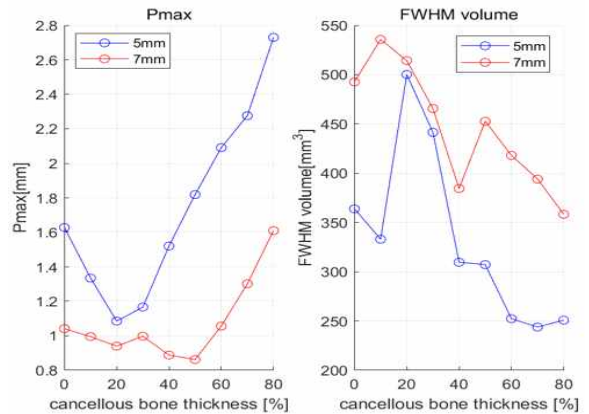


그림 2. 두개골 두께와 cancellous bone의 비율에 따른 최대 음압과 FWHM 부피

Fig. 2. P-max and FWHM volume according to skull thickness and cancellous bone ratio

두개골 두께와 해면골의 비율에 따른 빔 패턴을 가시화했다 (그림 3). 두개골 두께가 5mm인 경우 최대 음압이 감소하고 이후 증가하는 구간인 10%, 20%의 해면골 비율과 70%를 확인했다. 두개골 두께가 7mm인 경우 해면골 부피가 증가하지만 집속 영역의 부피가 감소하고 크게 증가하는 구간인 10%, 40%의 해면골 비율과 80% 해면골 비율의 결과를 살펴보았다. 일반적으로 타원체 모양의 집속 영역을 확인할 수 있었으며 해면골 비율이 높은 경우 최대 음압이 크기 때문에 상대적으로 작은 집속 부피를 확인할 수 있다.

해면골 비율에 따른 최대 음압, FWHM의 장축, 단축, 부피 그리고 초점 거리에 대한 시뮬레이션을 분석한 결과 비선형적이긴 하나 결국 일관적으로 증가하거나 감소하는 경향을 보이는 것을 관찰했다. 따라서 추후 다양한 두개골 두께와 해면골 비율을 고려해서 이러한 현상을 해석하고 실제 두개골 모델에 적용해 일반화할 계획이다.



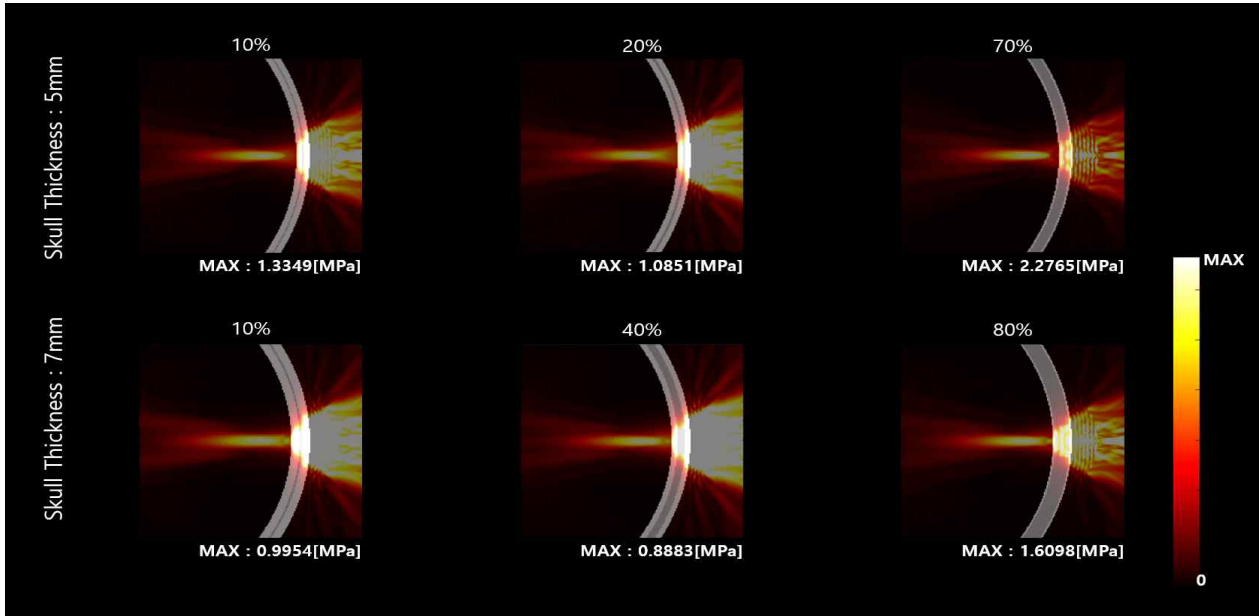


그림 3. cancellous bone과 두개골 두께에 따른 빔패턴  
Fig. 3. Beam patterns according to cancellous bone and skull thickness

### 참 고 문 헌

- [1] Hyeon. S, E. H. Lee, "Comparison of Analysis Results According to Heterogeneous or Homogeneous Model for CT-based Focused Ultrasound Simulation", Journal of Biomedical Engineering Research, vol. 43, pp. 369-374 11 2022.
- [2] J. H. Yu et al., "A simulation study of low-intensity focused ultrasound simulation according to diameter of transducer", Journal of KIIT, Jul 2024.
- [3] M. Chen et al., "Numerical and experimental evaluation of low-intensity transcranial focused ultrasound wave propagation using human skulls for brain neuromodulation", Medical Physics, Vol. 50, No. 1, pp. 38-49, Jan., 2023.
- [4] Y. Huang et al., "Numerical investigation of the energy distribution of low-intensity transcranial focused ultrasound neuromodulation for hippocampus", Ultrasonics, Vol. 124, Aug., 2022.
- [5] Hyeon. S et al., "Numerical Investigation of Layered Homogeneous Skull Model for Simulations of Transcranial Focused Ultrasound", Neuromodulation: Technology at the Neural Interface, April 2024.
- [6] B. T. Cox et al., "K-space propagation models for acoustically heterogeneous media: application to biomedical photoacoustics", The Journal of the Acoustical Society of America, Vol. 121, No. 6, pp. 3453-3464, Jun., 2007.
- [7] P. C. Miranda, M. Lovarev and M. Hallett, "Modeling the current distribution during transcranial direct current stimulation". Clinical Neurophysiology, vol. 117, pp. 1623-1629 , July 2006.
- [8] J. K. Mueller et al., "Numerical evaluation of the skull for human neuromodulation with transcranial focused ultrasound", Journal of neural engineering, Vol. 14, No. 6, Nov., 2017.
- [9] O. Puonti et al., "Accurate and robust whole-head segmentation from magnetic resonance images for individualized head modeling", NeuroImage, Vol. 219, Oct., 2020
- [10] <https://itis.swiss/who-we-are/> [Accessed: Oct. 14, 2024]

# 1인칭 혼합 현실과 실시간 전략이 결합된 멀티플레이 시스템

김웅식\*, 이민재\*\*, 이수원\*\*\*

## Integrating First-person Mixed Reality and Real-time Strategy in a Novel Multiplayer Gaming System

Ungsik Kim\*, Minjae Lee\*\*, and Suwon Lee\*\*\*

### 요 약

몰입형 게임 기반 학습은 학습자에게 현실과 유사한 환경을 제공하고, 몰입형 장비를 착용한 사용자간의 멀티플레이는 많이 출시된 상황이다. 본 논문은 비 몰입형 사용자 또한 함께 할 수 있도록 물리적 공간을 탐색하는 1인칭 혼합현실(MR) 사용자와 PC에서 게임 요소를 조작하는 실시간 전략 사용자를 통합하는 멀티플레이 시스템을 소개한다. 또한, MR 사용자의 물리적 움직임과 위치가 실시간 전략 사용자의 게임 환경에 동적으로 통합되는 프로토타입을 개발하였다. 해당 프로토타입은 2인용 시뮬레이션 연습과 같은 미션 기반 게임에 적합하다. 게임플레이 경험이 기술과 몰입감을 향상시키는 것에 반하여 해당 프로토타입 개발에 있어 한계점을 밝히고 향후 과제로 남기고자 한다.

### Abstract

Immersive game-based learning provides learners with life-like experiences, and multiplayer between users wearing immersive equipment is widely available. In this paper, we present a multiplayer system that integrates a first-person mixed reality (MR) user exploring a physical space with a real-time strategy (RTS) user manipulating game elements on a remote PC to allow non-immersive users to join in. We developed a prototype in which the MR user's physical movements and position are dynamically integrated into the RTS user's game environment. The prototype is suitable for mission-based games such as two-player simulation exercises. While the gameplay experience enhances tactics and immersion, we also discuss the limitations of the prototype development and future work.

### Key words

mixed reality, real-time strategy, multi-player, gaming system

## 1. 서 론

최근 몰입형 게임 기반 학습에서 가상현실(VR)과 증강현실(AR)의 활용이 교육 분야에서 혁신적인 접근 방식으로 부상하고 있다. 연구자들은 VR/AR 기반 학습의 잠재력을 탐구하고 이러한 기술이 학습자의 동기 부여와 성취도를 어떻게 향상시킬 수 있

는지에 대해 다양한 연구를 진행하고 있다. [1]

기존 연구는 VR 사용자가 몰입형 헤드 마운트 디스플레이(HMD)를 사용하는 것에 초점을 맞췄다. 이러한 접근은 PC 사용자가 VR/AR 게임 기반 학습에 참여할 수 있는 기회를 제한하고 전반적인 학습 경험의 다양성과 접근성을 떨어뜨린다. 본 논문은 이러한 문제를 인식하고 PC 사용자와 메타 퀘스

\* 경상국립대학교 컴퓨터공학과, blpeng@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, wjdchs0129@gnu.ac.kr

\*\*\* 경상국립대학교 컴퓨터공학과, leesuwon@gnu.ac.kr (교신저자)

트 사용자가 네트워크 통신을 통해 상호작용할 수 있는 교육 개념을 제안한다.

이러한 교육 개념은 전략가와 활동가 간의 협력 학습에 효과적이다. 다양한 플랫폼의 사용자가 협력하여 문제를 해결하고, 서로 다른 관점과 전략을 공유함으로써 학습 효과를 극대화할 수 있다. 전략 사용자는 상황을 관찰하고 1인칭 MR 사용자에게 전략 지시를 내리거나 유용한 게임 요소를 배치할 수 있다. 1인칭 MR 사용자는 전략 사용자의 지시와 지원을 받아 실제 환경에서 게임을 플레이한다.

## II. 관련 연구

메타 퀘스트 2는 다른 상용 HMD 장비에 비해 트래킹 정확도가 높다[2]. 메타 퀘스트 3 역시 마찬가지로 정확한 트래킹을 기대할 수 있다. 그러나 메타 퀘스트는 햇빛에 노출되면 영구적으로 센서와 디스플레이가 손상될 위험이 있어 야외에서 MR을 즐기기에는 무리가 있다.[3].

[4]는 VR과 MR 사용자 간의 상호작용을 중심으로 몰입형 콘텐츠 제작 공정을 정리하였다. [5]는 메타 퀘스트와 같은 상용 HMD를 야외에서 사용할 수 있는 GPS 기반 접근방식을 제안하였다. 앞선 연구들은 대부분 HMD 사용자만을 대상으로 했다. [6, 7]은 HMD 사용자와 비 HMD 사용자 간의 상호작용을 포함하는 게임 콘셉트를 제안하였다. 하지만 이러한 연구는 비 HMD 사용자도 신체적 움직임에 참여해야 한다고 가정한다. 이는 PC 사용자가 신체적 움직임을 수행하지 않아도 되는 PC와 1인칭 MR 사용자 간의 협업 시나리오와 대조적이다. 이러한 차이는 게임 환경의 다양한 사용자 유형에서 인터랙티브 경험을 향상하기 위한 다양한 접근 방식을 강조한다.

## III. 게임 컨셉

### 3.1 1인칭 시점 MR 사용자

MR 구현에서 1인칭 시점 MR 사용자는 컨트롤러 또는 손 트래킹을 사용하여 플레이할 수 있으며, 게임 플레이 중에 두 가지 모드를 전환할 수 있다.

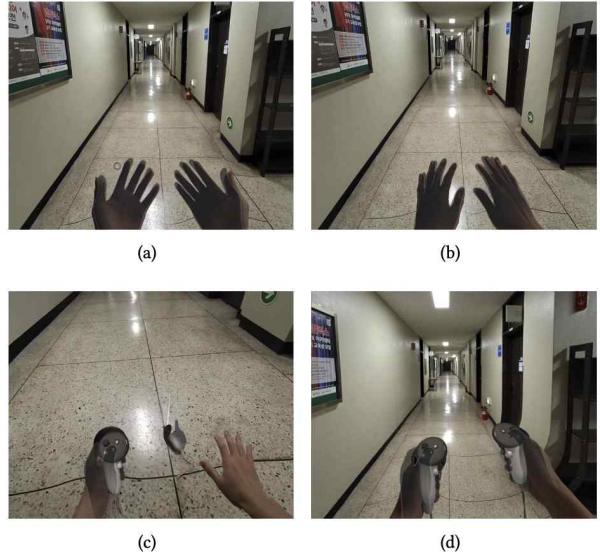


그림 1. MR 플레이어가 경험하는 네 가지 시나리오  
Fig. 1. Four scenarios experienced by the MR player during gameplay

그림 1에서 컨트롤러 방식은 버튼 클릭 이벤트를 통해 기능을 개발하는 데 용이하지만, 현실 세계에서 움직이는 사용자의 MR 경험이 저하될 수 있다. 그림 1의 (a) 와 (b)는 사용자가 손을 자유롭게 움직일 수 있어 잡거나 던지는 등 다양한 동작을 구현할 수 있다. 하지만 이를 위해서는 기존에 버튼 클릭으로 실행되던 기능들을 사용하기 위해 손동작을 익혀야 하며, 수많은 손동작을 요구하면 사용자의 MR 경험이 저하될 수 있다. 사용자가 한 손에 컨트롤러를 들고 있으면 그림 1의 (c)처럼 양손에 컨트롤러를 들고 있는 것으로 인식된다. 따라서 한 손으로 제어할 수 없다.

### 3.2 실시간 전략 사용자

실시간 전략 사용자는 2.5D 시점 또는 항공 시점에서 맵을 관찰한다. 이러한 구성을 통해 실시간 전략 사용자는 전체 맵과 1인칭 MR 사용자의 위치를 감독하여 전략적인 명령을 내릴 수 있다. 실시간 전략 사용자는 PC의 주요 장비인 모니터와 마우스 및 키보드를 사용하여 조작한다.

## IV. 개발

1인칭 MR 경험을 제공하기 위해 메타에서 제작한 MR 헤드장비인 메타 퀘스트 3를 사용하였다.

유니티 2022.3.16.fl, Meta XR All-in-One SDK, 포톤 퓨전 2.0을 구현하는 데 사용하였다. 실시간 전략 사용자가 사용하는 3D 지도 모델링에는 블렌더 3.6을 활용하였다.

#### 4.1 맵 렌더링

그림 2는 LiDAR 스캐닝 앱을 사용하여 지도를 렌더링한 것과 블렌더를 사용하여 건축 청사진을 참조하여 실제 공간을 복제한 맵이다. LiDAR 스캐닝 방법은 지도의 길이가 길기 때문에 스캔하는 동안 카메라가 트래킹을 잃고 지도가 부적절하게 렌더링 되었다.

블렌더 맵을 유니티 에디터에 업로드할 때 유니티의 측정 단위가 현실 세계의 미터에 대응하도록 설계하였다. 1인치 MR 사용자가 현실 세계에서 1m를 이동하면 실시간 전략 사용자가 관찰한 대로 유니티 좌표에서도 1m씩 이동한다.

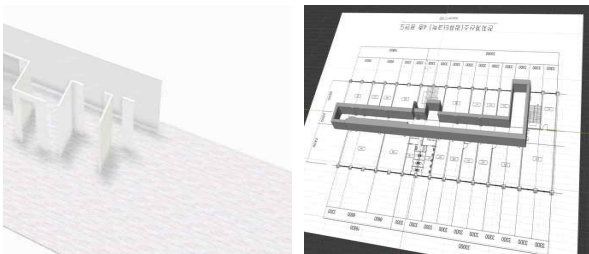


그림 2. LiDAR 스캐닝 앱 결과와 Blender를 통한 맵핑  
Fig. 2. LiDAR scanning app results and mapping with Blender

#### 4.2 패스스루

패스스루는 메타 퀘스트 사용자가 실제 환경을 화면에서 볼 수 있도록 하는 기능이다. 하지만 개발자는 사용자의 실제 환경의 이미지나 동영상에 액세스할 수 없기 때문에 유니티 에디터에서 실행할 때 패스스루 레이어가 정상 작동하지 않는다.[8] 애플리케이션을 빌드하고 실행하면 올바르게 작동한다. 유니티에서 제공하는 기본 혼합현실 템플릿에서도 실행 중에 해당 문제를 나타내는 메시지가 표시된다.

#### 4.3 좌표계 세팅

메타 퀘스트 앱을 실행할 때, 사용자의 시작 지점이 유니티 공간의 원점으로 설정되므로 앱이 실행

될 때마다 원점의 좌표가 동적으로 변경된다. 이러한 가변성은 MR 오브젝트를 여러 위치에 정확하고 일관되게 다양한 위치에 배치하는 것이 어려워진다. 이러한 좌표계 문제에는 두 가지 해결책이 있다. 첫 번째로 공간 앵커를 사용하여 이 문제를 해결할 수 있다. 이 기술은 포인트 클라우드들의 특징을 기반으로 앵커의 위치를 기억한다. 따라서 유니티에서 공간 앵커의 위치가 변경되더라도 현실 세계에서의 좌표는 고정된 상태로 유지된다. 두 번째는 메타 퀘스트 애플리케이션을 실행할 때 시작 위치와 방향을 통제하는 것이다. 애플리케이션을 실행할 때 초기 위치와 방향을 관리하면 비교적 동기화된 좌표계를 얻을 수 있다.

공간 앵커를 사용하여 시작 지점을 제어하는 것이 항상 유리하다고 생각할 수 있지만, 이 경우 1인치 MR 사용자가 앱 실행 후 공간 앵커의 시작 지점을 수동으로 설정하는 과정에서 인적 오류가 발생할 수 있다는 점이 시작 위치를 통제하는 것과 동일하다. 마커 인식을 사용하여 이 작업을 메타 퀘스트에 위임하면 인적 오류를 해소할 수 있지만, 카메라를 통해 이미지 스트림을 수신할 수 없어 마커 인식이 불가능하기 때문에 이 접근 방식은 적합하지 않다.

#### 4.4 네트워킹

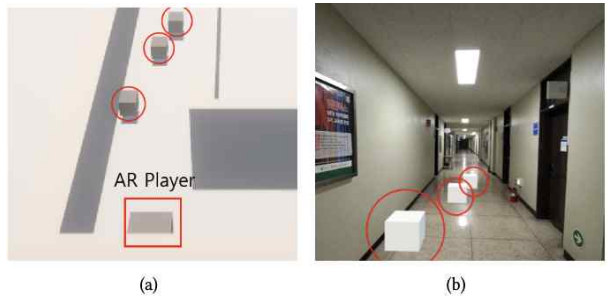


그림 3. 서로 공유되는 게임 오브젝트  
Fig. 3. Shared game objects

그림 3은 포톤 퓨전을 사용하여 실시간 전략 사용자와 1인치 MR 사용자 간의 통신을 개발하였다. 1인치 MR 사용자가 퓨전 서버에 성공적으로 연결되면 사용자의 머리에 큐브가 생성된다. 이 큐브는 OVRCameraRig의 CenterEyeAnchor의 위치와 회전을 추적하기 위해 추가하였다. 이 큐브는 실시간 전략 사용자의 모니터에서 포톤 네트워크 통신을 통해

볼 수 있다. 그런 다음 실시간 전략 사용자가 맵의 아무 지점이나 클릭하여 큐브를 인스턴스화 하도록 구현하였다. 인스턴스화된 큐브는 포톤 네트워크 통신을 통해 1인칭 MR 사용자의 화면에 표시된다.

## V. 실험

그림 4는 실시간 전략 사용자와 1인칭 MR 사용자의 좌표계가 얼마나 차이가 나는지 실험하였다. 실험 결과, 좌표계의 오류로 인해 MR 1인칭 사용자가 벽에 붙을 시, 실시간 전략 사용자의 화면에서 벽과 충돌하는 현상이 발생하였다. 이는 광범위한 동작이 아닌 섬세한 움직임이 필요한 게임에서 향후 과제가 남아있음을 시사한다.



그림 4. 좌표계 동기화 실험  
Fig. 4. Coordinate System Synchronization Test

## VI. 결론 및 향후 과제

본 논문은 1인칭 MR 사용자와 실시간 전략 사용자를 통합하는 멀티플레이 게임 시스템을 제안하고 시스템의 프로토타입을 개발하였다. 우리가 제안한 시스템은 이전 연구에서 다루지 않은 전략가와 행동가 간의 협동을 제공할 수 있다. 이 시스템은 2인용 시뮬레이션 연습과 같은 미션 기반 게임에 적합하다. 제안된 개념의 효과를 입증하기 위해서는 다양한 평가 도구를 사용하는 것이 필수적이다. [9] 향후에는 제안한 시스템의 효과를 평가하기 위해 실시간 전략 사용자와 1인칭 MR 사용자가 공동으로 수행할 수 있는 학습 과제를 만들어 테스트하고, 제안한 시스템의 효율성을 추가로 평가할 것이다. 또한, 좌표계 통일 문제를 해결하기 위해 아두이노와 같은 소형 모듈과의 통합을 후속 연구로 수행할 계획이다.

## 참 고 문 헌

- [1] Gugenheimer, Jan, et al. "Sharevr: Enabling co-located experiences for virtual reality between hmd and non-hmd users." Proceedings of the 2017 CHI conference on human factors in computing systems. 2017.
- [2] Holzwarth, Valentin, et al. "Comparing the accuracy and precision of steamvr tracking 2.0 and oculus quest 2 in a room scale setup." Proceedings of the 2021 5th International Conference on Virtual and Augmented Reality Simulations. 2021.
- [3] Meta. "How to protect your Meta Quest headset from sun damage". Meta, <https://www.meta.com/help/quest/articles/headsets-and-accessories/product-care-and-best-practices/avoiding-sunlight-damage/>. accessed 9 October 2024.
- [4] Jongseon Kim, Sumin Kong, Moonsu Jang, & Jinmo Kim (2024). A Study on Production and Experience of Immersive Contents based on Mixed Reality and Virtual Reality using Meta Quest Pro. Journal of the Korea Computer Graphics Society, 30(3), 71-79.
- [5] Ahyeon Park, Sangjoon Bae, Mijin Kim, Woohyun Lee, & Dongho Kim (2024). Location-based Space-Time Interlinked Extended Reality using HMD. Journal of the Korea Computer Graphics Society, 30(3), 133-140.
- [6] Kamei, Ikuo, et al. "CoVR: Co-located virtual reality experience sharing for facilitating joint attention via projected view of HMD users." SIGGRAPH Asia 2020 Emerging Technologies. 2020. 1-2.
- [7] Lin, Hong, et al. "Metaverse in education: Vision, opportunities, and challenges." 2022 IEEE International Conference on Big Data (Big Data). IEEE, 2022.
- [8] Passthrough API Overview, <https://developers.meta.com/horizon/documentation/unity/unity-passthrough/>. accessed 9 October 2024.
- [9] Radianti, Jaziar, et al. "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda." Computers & education 147 (2020): 103778.

# 제주시 전력 수요 예측을 위한 RNN 계열 모델과 기상 변수 통합 분석

정현수\*, 박지수\*\*, 길준민\*\*

## Integration of Weather Variables and RNN-based Models for Electricity Demand Forecasting in Jeju-si

Hyun-su Jeong\*, Ji Su Park\*\*, and Joon-Min Gil\*

### 요약

본 연구에서는 제주도 전력 수요 예측을 위해 RNN, LSTM, BiLSTM 모델을 사용하고, 온도와 지중 온도 등 전력 수요와 높은 상관관계를 가진 주요 기상 변수를 통합하여 예측 성능을 개선하였다. 연구 결과, 기상 데이터를 포함한 모델이 기상 데이터를 포함하지 않은 모델에 비해 높은 예측 정확도를 보여 주었으며, 이를 통해 기상 조건 변화에 따른 전력 수요 변동성을 효과적으로 반영함을 확인하였다.

### Abstract

In this paper, RNN, LSTM, and BiLSTM models are employed to forecast electricity demand in Jeju City, incorporating key weather variables such as temperature and ground temperature, which have a high correlation with electricity demand. The results demonstrate that the models integrated with weather data achieve higher prediction accuracy than the ones without weather data, effectively capturing the variability in electricity demand driven by changes in weather conditions.

### Key words

Electricity demand forecasting, Pearson correlation coefficient, Weather data, RNN, LSTM, BiLSTM

### I. 서론

기후 변화에 대한 대응과 지속 가능한 환경에 대한 요구가 증가함에 따라, 전 세계적으로 신재생

에너지의 도입이 빠르게 확산되고 있다. 이에 발맞추어, 제주도는 '탄소 없는 섬 2023(Carbon Free Island 2023)[1] 프로젝트를 통해 2030년까지 전력 수요의 100%를 신재생에너지로 충당하여 탄소 배출

\* 제주대학교 컴퓨터공학과, jhs990909@stu.jejunu.ac.kr, jmgil@jejunu.ac.kr(교신저자)

\*\* 전주대학교 컴퓨터공학과, jisupark@jj.ac.kr

※ 본 결과물은 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신사업의 결과입니다(2023RIS-009). 그리고 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022R1A2C1092934).

없는 지역을 구현하는 것을 목표로 한다. 이를 위해 제주도는 총 4,085MW의 신재생에너지 설비(태양광 1,411MW, 풍력 2,345MW)를 구축할 계획이다. 또한, 제주도는 장기적으로 고전압 직류 송전(HVDC) 의존도를 줄이고, 대규모 에너지 저장 시스템(ESS)을 적극 활용하여 전력 계통의 안정성과 유연성을 높이고자 한다. 신재생에너지는 기상 조건에 따라 발전량이 크게 변동되는 특성상, 발전 설비가 분산된 제주도에서 수요와 공급을 균형 있게 맞추기 위해서는 지역별 전력 관리가 필수적이다. 특히 제주도와 서귀포시와 같은 주요 지역에서 생산되는 전력을 중앙 송전 방식 대신 각 지역에서 직접 소비하는 것이 효율적이다. 그러나 각 지역의 수요를 정확하게 예측하지 못하면 출력 제한 조치를 취해야 하며, 이는 전력의 낭비와 효율성 저하로 이어질 수 있다. 출력 제한은 수요와 공급의 균형을 유지하고 전력 계통의 안정성을 확보하기 위한 필수 조치이지만, 불필요한 전력 손실을 방지하고 전력 효율성 극대화를 위해서는 수요 예측이 중요하다.

이러한 수요 예측을 위해 본 연구에서는 제주도의 특성에 맞는 기상 변수를 선별하여 RNN 계열 모델을 학습하는 방법을 제안한다. 구체적으로, 전력 수요와 높은 상관 관계를 가지는 온도와 지중 온도를 주요 특징으로 사용하여 예측 모델의 성능을 향상시키고자 한다. 이를 통해, 지역적 특성을 반영한 기상 데이터의 활용이 전력 수요 예측에서 중요한 요인임을 입증하고자 한다. 성능 평가를 위해 RNN, LSTM, BiLSTM 모델을 사용하여 실험을 수행하였으며, 예측 성능을 MAE, MSE,  $R^2$  지표로 평가한다. 실험 결과는 지역적 기상 변수를 포함한 모델이 전력 수요 데이터만 사용한 모델에 비해 우수한 예측 정확도를 나타낸다.

## II. 관련 연구

### 2.1 기존 연구

문형빈 외(2023)의 연구에서는 머신러닝 및 DNN(Deep Neural Network)을 활용하여 한국전력거래소의 제주도 전력 거래 데이터를 기반으로 제주도의 전력 수요를 예측하였다[2]. 이 연구에서는 제주도

전체를 대상으로 하여 평일, 주말, 휴일 및 계절별로 데이터를 세분화해 학습을 수행하였으며, 기상 데이터를 전력 수요 예측의 독립 변수로 활용한다. 이를 통해 전력 수요의 주기적 변동성을 반영하였다. 반면, 본 연구는 RNN(Recurrent Neural Network) 계열 모델을 사용하고 기상 데이터를 분석 및 상관관계가 높은 기상데이터만 함께 학습에 포함시켜 제주도 전력 수요를 예측하고자 한다. RNN 계열 모델은 DNN에 비해 시계열 데이터의 순차적 패턴을 학습하는 데 유리한 장점을 가지고 있다. 전력 수요 데이터는 시간에 따른 연속적 변동성을 가지기 때문에, 이를 효과적으로 예측이 가능하며, 기상 데이터를 함께 사용하여 RNN 계열 모델이 시간에 따른 날씨의 변화와 전력 수요 간의 관계를 효과적으로 학습하도록 설계하였다.

Sujan Ghimire et al.(2023)는 전력 수요 예측에서 지역 기후 변수를 통합하는 TNET(Multi-Head Self-Attention Transformer)의 중요성을 강조하면서, 온도, 강수량, 습도 등의 기후 요소가 전력 소비 패턴에 중요한 영향을 미친다고 주장한다[3].

김동관 외(2024)의 연구에서는 3일 이동평균과 유사일 기반 전력 수요 예측 알고리즘을 통해 제주 계통의 요일 구분을 최적화하여 예측 성능을 향상시켰다[4]. 그러나 이러한 방법은 시계열 데이터 내의 장기적 패턴이나 비선형적인 관계를 반영하는데 제한적일 수 있다.

윤상철 외(2023)는 태양광 발전량을 전력 수요 예측 요소로 사용할 수 있다고 주장하며, 제주도 하계 전력 수요 예측의 정확도를 높이기 위해 태양광 발전량과 운량(구름 양)을 독립 변수로 추가하였다[5]. 그러나 이 접근법은 전력 시스템에서의 수요와 공급 관계를 오해한 결과일 수 있다. 실제로 발전량은 전력 수요에 맞추어 조정되며, 수요가 증가하면 발전량이 늘어나고, 수요가 감소하면 줄어드는 방식으로 균형을 맞춘다. 따라서, 태양광 발전량을 수요 예측의 독립 변수로 사용하는 것은 수요에 종속된 변수를 사용하는 것으로 적절하지 않다.

Amir Miraki et al.(2024)는 전력 수요 예측에서 과거 데이터의 중요성을 강조하면서, 1시간 전 데이터가 예측에 가장 중요한 요소로 작용하며, 24시간

전과 48시간 전의 데이터도 주기적인 전력 소비 패턴을 반영하여 예측 정확성을 높이는 데 유용하다고 제안한다[6]. 본 연구에서는 가정 전력 소비가 하루 단위로 유사한 패턴을 보인다는 점에서, 과거 24시간 데이터를 활용해 다음 날 24시간의 수요를 예측하는 방식을 사용한다.

## 2.2 전력 수요 데이터

본 연구에서 사용된 시간별 전력 사용량 데이터는 제주도 지역을 대상으로 하며, 원격검침 인프라(AMI)를 통해 수집된 데이터이다. 연구 데이터는 한국전력공사(KEPCO)[7]로부터 제공받았으며, 2021년 1월 1일에서 2023년 12월 31일까지의 시간당 제주도 전력 수요량 데이터이다.

## 2.3 기상 데이터

본 연구에서 사용된 시간별 기상 데이터는 기상청의 기상자료개방포털에서 제공하는 종관기상관측(ASOS)[8] 데이터를 활용하였으며, 이는 제주지방기상청에서 관측한 데이터이다. 2021년 1월 1일에서 2023년 12월 31일까지의 시간당 기온, 지중온도, 풍속, 풍향, 습도, 증기압, 이슬점온도, 현지 기압, 해면기압, 전운량, 중하층운량을 포함하고 있다.

### Ⅲ. 지역 특성 기반 전력 수요 예측 기법

#### 3.1 기상 데이터 분석

본 연구에서는 2021년부터 2022년까지의 기상 데이터와 제주시의 전력 수요 데이터 사이의 상관 관계를 분석하기 위해 피어슨 상관 계수를 활용한다. 피어슨 상관 계수는 두 변수 간의 선형 관계의 강도와 방향을 나타내는 지표로, 값의 범위는 -1에서 1까지이다. 상관 계수가 1에 가까울수록 두 변수는 강한 양의 선형관계를 가지며, -1에 가까울수록 강한 음의 선형 관계를 가진다. 0에 가까우면 값은 두변수 간에 선형적 상관 관계가 거의 없음을 의미한다. 본 연구의 분석 결과, 그림 1은 2021년 8월의 상관계수를 나타낸다. 여름철에 특히 온도와 지중 온도가 전력 수요와 높은 상관 관계를 보이는 것으로 나타났다. 이러한 결과는 여름철 기온 상승이 전력 수요 증가에 큰 영향을 미친다는 점을 나타낸

다. 반면, 기온과 지중 온도 이외의 변수들은 전력 수요와 낮은 상관계수를 보여, 전력 수요에 미치는 영향이 상대적으로 적은 것으로 분석되었다. 그리하여 본 연구에서는 온도와 지중 온도를 주요 특징으로 사용하여 전력 수요 예측 모델을 학습한다.

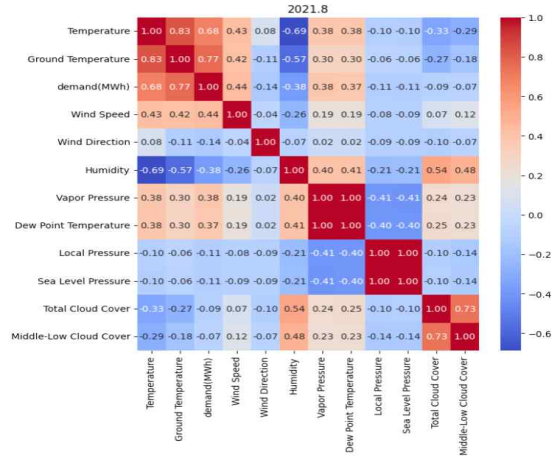


그림 1. 8월 기상 변수별 전력 수요 상관 계수 분석

#### 3.2 모델 학습 및 평가

본 연구에서는 전력 수요 예측을 위해 RNN, LSTM, BiLSTM 모델을 사용하여 시계열 데이터를 학습하였다. 학습 데이터는 2021년부터 2022년까지의 전력 수요 및 기상 데이터를 사용하였으며, 2023년 데이터를 테스트 데이터로 활용하였다. 모델의 입력 시퀀스는 24시간 단위로 설정하여, 24시간의 입력 데이터를 기반으로 다음 24시간의 전력 수요를 예측하도록 설계한다. 또한, 온도 데이터를 포함한 모델과 포함하지 않은 모델의 예측 성능을 평균 절대 오차(MAE), 평균 제곱 오차(MSE), 결정 계수( $R^2$ )를 기준으로 평가한다. 표 1은 전력 수요 예측성능을 보여주며, 온도와 지중 온도를 함께 학습에 사용한 모델이 전력 수요 데이터만 사용한 모델에 비해 우수한 성능을 보인다.

표 1. 전력 수요 예측 성능

Model	MAE	MSE	$R^2$
RNN w/o Temp.	17.079	558.946	0.894
RNN with Temp.	14.422	427.011	0.919
LSTM w/o Temp.	15.756	503.965	0.904
LSTM with Temp.	12.451	346.091	0.934
BiLSTM w/o Temp.	13.4575	394.234	0.925
BiLSTM with Temp.	11.945	323.864	0.938



그림 2는 온도 변수들을 포함한 모델과 포함하지 않은 모델의 성능 차이를 비교한 그래프이다. RNN, LSTM, BiLSTM 모델을 대상으로 MAE, MSE, R<sup>2</sup> 기준에서 성능 향상을 상대적 개선율(%)로 나타낸다. 상대적 개선률 RI는 다음과 같이 정의된다.

$$RI = \frac{\text{with Temp} - \text{w/o Temp}}{\text{w/o Temp}} \times 100 \quad (1)$$

그림 2의 결과를 살펴보면, 전반적으로 온도 변수를 포함한 모델이 포함하지 않은 모델에 비해 모든 지표에서 성능이 개선된 것을 확인할 수 있다.

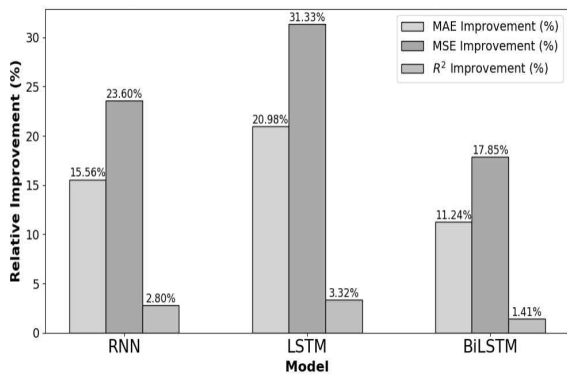


그림 2. 온도 변수 포함에 따른 성능 개선율

#### IV. 결 론

본 논문에서는 제주시 전력 수요 예측을 위해 RNN 계열 모델을 활용하였으며, 기상 데이터를 통합하여 예측 성능을 개선한다. 기존 연구와 달리 본 연구는 기상 변수 중 전력 수요와 높은 상관 관계를 가지는 온도와 지중 온도를 주요 특징으로 삼아 예측 모델을 학습한다. 연구 결과, 온도 변수를 포함한 RNN, LSTM, BiLSTM 모델이 보다 높은 예측 정확도를 보였으며, 이는 제주시 전력 수요 데이터가 온도와 지중 온도를 사용했을 때 전력 수요 변동성을 효과적으로 반영할 수 있음을 보여준다.

추후 연구에서는 제주시뿐만 아니라 서귀포시의 전력 수요 예측에도 본 연구의 예측모델을 확장적용하고자 한다. 또한, 전력 수요와 관련성이 높은 추가 변수를 분석하고 이를 모델에 적용함으로써, 예측 성능을 더욱 향상시킬 방안을 모색할 예정이

다. 이러한 확장은 제주도 전역의 전력 수요 예측의 정확도를 높이는 데 기여할 수 있으며, 신재생 에너지의 효율적인 운영을 위한 데이터 기반 의사 결정 지원에 중요한 자료로 제공할 것이다.

#### 참 고 문 헌

- [1] Carbon Free Island Jeju by 2030. Accessed: Nov. 1, 2024. [Online]. Available: <https://oecd-opsi.org/innovations/carbon-free-island-jeju-by-2030/>
- [2] Moon, HyungBin, Ahn, Jaekyun, and Lee, Chul-Yong, "Forecasting electricity demand of Jeju island using machine learning," *Innovation Studies*, vol. 18, no. 4, pp. 49-63, 2023.
- [3] S. Ghimire, T. Nguyen-Huy, M. S. AL-Musaylh, R. C. Deo, D. Casillas-Pérez, and S. Salcedo-Sanz, "Integrated Multi-Head Self-Attention Transformer model for electricity demand prediction incorporating local climate variables," *Energy and AI*, vol. 14, 2023.
- [4] DongKwan Kim, Woo-Jin Jang, Young-Min Wi, and Sung-Kwan Joo, "Categorizing Electrical Day Types in Jeju Grid," *Proceeding of the 55th KIEE Summer Conference 2024*, 2024, pp. 638-639.
- [5] Sang-Cheol Yun, Ho-Jun LEE, Byung-Ho Kim, and Hong-Rae Kim, "Deep Learning Based Load Forecasting Considering Power Generation in Jeju," *Proceeding of the 54th KIEE Summer Conference 2023*, 2023, pp. 113-114.
- [6] A. Miraki, P. Parviainen, and R. Arghandeh, "Electricity demand forecasting at distribution and household levels using explainable causal graph neural network," *Energy and AI*, vol. 16, 2024.
- [7] KEPCO. Accessed: Nov. 1, 2024. [Online]. Available: <https://home.kepco.co.kr/kepco/main.do>
- [8] KMA. Accessed: Nov. 1, 2024. [Online]. Available: <https://data.kma.go.kr/data/grnd/selectAsosRltmlist.do>



# IT계열 학부생 대상 클라우드 컴퓨팅과 MSA 기반 서비스 구축에 대한 이해를 돕기 위한 교과목 개발에 관한 제안

김민영\*, 한유정\*\*, 김남우\*\*\*, 이진형\*\*\*, 이현섭\*\*\*\*

## Proposal for the development of a course to help IT undergraduate students understand cloud computing and MSA-based service construction

Minyoung Kim\*, Yujung Han\*\*, Jinhyeong Lee\*\*\*, Namu Kim\*\*\*, and Hyoun-Sup Lee\*\*\*\*

### 요 약

본 논문은 IT학과 학부생들이 최신 클라우드 컴퓨팅 기술 중 하나인 컨테이너 플랫폼을 활용하여 마이크로 서비스 아키텍처(MSA)를 구현하는 방법을 학습할 수 있도록 설계된 전공 교과목 개발을 목적으로 한다. 본 교과목은 3학점으로 구성되며, 중간·기말고사를 제외한 13주 동안 매주 3시간씩 진행된다. 학생들은 이 교과목을 통해 컨테이너 기반의 MSA를 이해하고 이를 실제 IT서비스에 적용할 수 있는 역량을 갖추게 되며, IT서비스 분야의 개발자로 취업 경쟁력을 높이고 기술 커리어를 발전시킬 기회를 제공받을 수 있다.

### Abstract

This paper aims to develop a specialized course designed for undergraduate IT students to learn how to implement Microservice Architecture (MSA) using container platforms, a key technology in modern cloud computing. The course is structured as a 3-credit class, running for 13 weeks with 3 hours of instruction per week, excluding mid-term and final exams. Through this course, students will gain a general understanding of container-based MSA and the ability to apply it in real-world IT services. Additionally, the course offers students the opportunity to enhance their employability as developers in the IT service industry and advance their technical careers.

### Key words

Microservice Architecture, Cloud Computing, Container Platform, Web Service

## 1. 서 론

현재 많은 IT서비스는 마이크로서비스 아키텍처

(MSA)를 기반으로 구축되고 있으며, 이를 통해 사용자 요구사항에 신속하게 대응할 수 있어 사용자 만족도를 크게 향상시키고 있다. MSA가 이러한 장

---

\* 동의대학교 ICT융복합연구소[kmyco@deu.ac.kr](제1저자),

\*\* 동의대학교 ICT융복합연구소,

\*\*\* Rama(동의대 창업교육센터 2024년 창업동아리),

\*\*\*\* 동의대학교 게임공학과[lhskmi@deu.ac.kr](교신저자)

점을 제공할 수 있는 이유는 클라우드 컴퓨팅 기술 중 하나인 컨테이너 플랫폼의 발전 덕분이다. MSA에서는 각 기능을 개별 서비스로 나누어 구현하며, 이때 각 서비스를 하나의 컨테이너로 분리하여 관리할 수 있어 유연성과 확장성을 극대화할 수 있다 [1-3].

그러나 현재 많은 IT학과에서는 클라우드 컴퓨팅과 MSA를 다루는 교과목이 부족하며, 개설된 과목조차 컨테이너 생성 및 관리에 그치는 수준에 머무르고 있다. 이에 따라 학생들은 최신 기술 트렌드를 충분히 경험하지 못해 IT서비스 분야에서의 취업 경쟁력과 기술 커리어 발전에 어려움을 겪고 있다.

본 논문은 이러한 문제를 해결하고자, IT계열 학부생들이 최신 클라우드 컴퓨팅 기술 중 하나인 컨테이너 플랫폼을 활용하여 MSA를 이해하고 구현할 수 있는 능력을 배양하는 전공 교과목을 개발하기 위한 내용을 제시한다.

## II. 본 론

본 논문에서 제시하는 교과목은 3학점(주당 3시간)으로 구성되며, 총 15주 동안 진행된다. 중간 및 기말고사 기간을 제외하고 13주 동안 실제 강의가 진행되며, 학생들이 무리 없이 따라올 수 있도록 웹 개발 언어를 이해하는 학생을 대상으로 가정한다. 이 교과목은 최신 클라우드 컴퓨팅 기술인 컨테이너 플랫폼과 MSA(Microservice Architecture)를 학습하는 것을 목표로 구성되었다.

표 1은 교과목의 주요 학습 단계를 나타내며, 첫 7주 동안은 중간고사를 기준으로 클라우드 시스템과 컨테이너 플랫폼의 기본 개념을 다룬다. 이 기간에 학생들은 클라우드와 네트워크의 기초, 리눅스 운영체제, 그리고 Docker를 중심으로 한 컨테이너 플랫폼을 학습하게 된다. 이 과정에서 학생들은 Docker와 같은 컨테이너 기술을 통해 클라우드 환경에서 서비스를 어떻게 구성하고 운영하는지 실습으로 이해하게 된다.

이후 기말고사 이전 6주 동안은 MSA 기반의 마이크로서비스 설계 및 구현에 중점을 둔다. 학생들은 MSA의 이론적 배경을 학습한 후, 자신이 익숙

한 웹 개발 언어를 이용해 MSA 기반 마이크로서비스를 설계하고 이를 컨테이너 플랫폼에 적용하는 실습을 진행한다. 이 과정에서는 각 서비스의 독립성을 유지하면서도 전체 시스템의 효율성을 높이는 방법을 학습하게 된다[4].

표 1. 본 논문에서 제안하는 교과목 내용 구성

Table 1. The composition of the course content proposed in this paper

진행 순서	학습 내용	실시시간 (주)
1	클라우드 개요	1
2	TCP/IP 기반 네트워크	1
3	리눅스	2
4	컨테이너 플랫폼	3
5	MSA(Microservice Architecture) 개요	2
6	클라우드와 MSA기반 웹서비스	4

본 교과목에서 컨테이너 플랫폼으로는 주로 Docker를 권장한다. Docker는 현재 가장 많이 사용되는 컨테이너 플랫폼으로, 다른 플랫폼도 유사한 기능을 제공하지만, Docker는 학습과 적용에 있어 상대적으로 쉬운 접근성을 제공한다. Docker를 통해 컨테이너를 생성하고 관리하기 위해서는 리눅스 환경에서 Shell 명령어를 사용할 줄 알아야 하며, Docker의 인터페이스가 CUI(Character User Interface)로 구성되어 있어 이에 대한 기초적인 명령어 사용이 필요하다. 더불어, 컨테이너 내부는 Linux 기반으로 제작되어 Shell 명령어에 대한 숙련도가 높을수록 효율적인 시스템 운영이 가능하다[4].

MSA 기반 마이크로서비스 개발을 위해서는 각 웹 언어에서 필요한 WAS(Web Application Server)를 구축하는 방법을 교수자가 안내해야 한다. 예를 들어, Spring을 사용하여 마이크로서비스를 개발하는 경우, 컨테이너에는 Java JDK(Java Development Kit)와 Apache Tomcat을 설치해야 하며, 학생들이 자신이 선택한 웹 언어에 맞는 WAS 환경을 구성할 수 있도록 한다. 또한, 컨테이너의 용량을 줄여 클라우드 플랫폼에서 최적화된 방식으로 운영할 수 있도록 경량화된 설치 방식을 제시하는 것이 필수적이다[5].

### III. 결 론

본 논문에서는 IT계열 학생들이 클라우드 컴퓨팅 기술 중 컨테이너 플랫폼을 기반으로 MSA (Microservice Architecture) 서비스를 구현하는 방법을 학습할 수 있는 3학점(주당 3시간, 총 13주) 전공 교과목을 개발하는 과정을 제시했다. 이 교과목은 학생들이 컨테이너 플랫폼의 기초 개념을 학습하고, 이를 바탕으로 MSA를 설계 및 구현하는 실습을 통해 실무 역량을 강화하는 데 중점을 두고 있다. 특히 Docker를 활용한 컨테이너 환경 구성과, 웹 언어를 이용해 마이크로서비스를 구축하는 방법을 통해 수강생이 클라우드 기반 IT서비스를 실제로 운영할 수 있는 경험을 제공한다.

학생들은 이 교과목을 통해 최신 클라우드 컴퓨팅 기술과 MSA를 적용하여 IT서비스 운영에 필요한 기초 역량을 쌓을 수 있으며, 이는 향후 IT서비스 관련 기업에서 요구하는 기술을 습득함으로써 취업 경쟁력을 높이는 데 기여할 것이다.

또한, 본 교과목에서 배운 내용을 바탕으로 추후 더 심화된 실습 과정인 캡스톤 프로젝트와 같은 과목에 연계할 필요가 있다. 이는 수강생의 기술적 성장과 취업 경쟁력을 더욱 강화할 수 있는 중요한 교육적 자원이 될 것이다.

### 참 고 문 헌

- [1] A. Raj and K. S. Jasmine, "Building Microservices with Docker Compose," *The International journal of analytical and experimental modal analysis*, Vol. 8, Issue 5, pp.1215-1219, May 2021.
- [2] D. Hossain, T. Sultana, S. Akhter, N. T. Thu, L. N. T. Huynh, G. Lee and E. Huh, "The role of microservice approach in edge computing: Opportunities, challenges, and research directions," *ICT Express from KICS*, Vol. 9, Issue 6, pp.1162-1182, Dec. 2023. DOI : <https://doi.org/10.1016/j.icte.2023.06.006>
- [3] *Microservices in Practice - Key Architectural Concepts of an MSA*, <https://wso2.com/whitepapers/microservices-in-pra>

[ctice-key-architectural-concepts-of-an-msa/](https://wso2.com/whitepapers/microservices-in-practice-key-architectural-concepts-of-an-msa/) [accessed: Oct. 12, 2024]

- [4] W. Jang, "Learn Docker & Kubernetes in One Book", Hanbit Publishing, Apr. 2024.
- [5] Amazon Corretto Lightweight Java-based application containers using OpenJDK, <https://aws.amazon.com/ko/blogs/tech/amazon-corretto-base-container-diet/> [accessed: Oct. 12, 2024]

# 건설사업정보 기반의 지식서비스 도출방안 연구

김 성 진\*

## A Study on the Method of Deriving Knowledge Service based on Construction Project Information

Seong-Jin Kim\*

AI 기술이 사회 각 분야에 급속히 확산되면서, 건설분야도 AI 기술을 도입하여 디지털업무로 전환하기 위한 기반 연구가 활발히 진행되고 있다. 그동안 축적된 건설데이터를 바탕으로 AI 환경에 맞춰 발주청 및 건설현장에 도움이 되는 지식서비스 요구가 증가 추세에 있다. 건설분야는 지난 20여년동안 건설정보화를 통해 종이·수기·대면 업무방식에서 전자문서·디지털·온라인 업무방식으로 전환하는데 노력하였다. 그 결과, 발주청과 건설현장간에 온라인으로 업무처리가 가능하게 되었으나, AI 기술을 통해 축적된 건설데이터를 분석하기 위해서는 데이터 전처리 등 사전 준비작업이 필요하며, 불특정다수를 대상으로 서비스함으로써 사용자별로 필요로 하는 서비스 제공이 미흡한 실정이다. 일본(JACIC: Japan Corporation Information Center), 미국(CIS: Construction Information Service), 싱가포르(CORENET: Construction and Real Estate Network), 영국(NBS: National Building Specification), 독일(Building radar) 등 해외에서는 건설사업 및 현장정보를 기반으로 지식서비스하기 위한 플랫폼 및 서비스를 제공하고 있다. 따라서, 해외사례와 같이 건설사업정보를 기반으로 정보활용주체별 지식서비스 개발이 필요하다.

본 연구는 건설사업정보를 지식화하여 주체별·단계별로 필요로 하는 서비스 방안과 활용 시나리오를 제시한다. 주체별·단계별 서비스 방안으로, 1단계는 기존 축적된 건설데이터를 활용한 서비스 범위를 제시하고, 2단계는 외부 시스템·데이터와 연계한 서비스 범위를 제시한다. 활용 시나리오로는, 과거 설계변경자료 분석결과를 현재 설계변경 가이드라인으로 활용하고, 과거 공사현황정보를 참고하여 향후 수행할 공사규모·공종에 따른 적정업체 검색이 가능한 활용 시나리오를 제시한다. 또한, 건설중 발생하는 민원을 분류하여 유사 공종시 참고자료로 활용하거나 지역 특성에 따라 발생하는 민원정보를 사전에 파악하여 공사 수행간 공기지연을 최소화할 수 있는 시나리오를 제시한다.

본 연구 결과를 통해, 다양한 분야에서 지식서비스가 활용될 수 있다. 공사 유형에 따른 지가 상승률 및 지역 총생산 등 지역 경제 활성요인 분석정보를 제공받거나, 거주지 인근의 SOC 개발정보를 서비스함으로써 국민의 알권리를 보장받을 수 있으며, 다양한 정보를 필요로 하는 창업·스타트업에도 도움이 될 것으로 기대된다.

### Key words

건설정보, 디지털전환, 지식서비스, 활용시나리오, AI건설

### 참고문헌

[1] 국토교통부, 19 건설사업정보시스템 운영 및 기능 개선(1), 2019.

\* 한국건설기술연구원, sjkim72@kict.re.kr

※ 본 연구는 국토교통부 건설사업정보화(CALS) 사업(24 건설사업정보시스템 운영 및 기능개선)으로 수행되었음

# NTIS 홈페이지 사용자 유입 현황 분석

김은경\*, 김도균\*\*

## Analysis of User Traffic Trends on the NTIS Website

Eungyeong Kim\* and Do Gyun Kim\*\*

본 연구의 목적은 NTIS(국가과학기술정보서비스)의 웹로그를 분석하여 사용자들의 이용행태를 파악하는 것이다[1]. 이를 통해 NTIS의 사용자 유입 현황, 트래픽 소스별 특징 및 성과를 분석하고, 유입 확대와 체류 시간 증대를 위한 전략 수립의 기반을 마련하여 NTIS 서비스 개선에 기여하고자 한다. 연구는 구글 애널리틱스를 활용하여 2024년 3월 부터 8월의 로그를 분석하였다. 사용자 분석 결과, 연령대별 활성 사용자는 45~54세 비율이 유지되었으나, 25~34세 비율은 감소 추세를 보였다. 지역별로 활성 사용자는 서울 사용자가 감소하였고, 신규 사용자(가입 기간 30일 이내)의 추이는 하락세를 지속하였으나 평균 조회수와 세션 시간(사용 시간)은 증가하였다. 주요 사용국은 한국, 미국, 일본, 중국 순으로 나타났으며, 한국의 총 사용자는 25.91%, 신규 사용자는 28.29% 감소하였다. 경로, 채널, 캠페인을 통해 웹사이트 방문 유입 경로 분석 결과, Direct(직접 URL 입력, 북마크, 이메일 링크 등) 사용자가 감소하였다. 페이지 조회수, 세션 지속 시간, 이벤트 등을 통해 사용자의 상호작용을 분석한 결과 사용자 수가 크게 감소, 이벤트 수도 크게 줄어들었다. 이는 신규 사용자의 유입 감소와 기존 사용자 참여도 저하를 의미한다. 새 사용자 수는 28.0%, 재사용자 수는 27.3% 감소하여 새 사용자 수의 감소와 활성 사용자의 이탈로 인해 전체적인 데이터가 감소를 보였다. 이러한 연구 결과는 NTIS 서비스 개선을 위한 기초자료로 활용될 수 있을 것이다.

### Key words

NTIS(National Science & Technology Information Service), Weblog, Web Traffic, Website using behavior

### 참고문헌

[1] <https://www.ntis.go.kr/>

---

\* 한국과학기술정보연구원(Korea Institute of Science and Technology Information), eungyeong@kisti.re.kr,

\*\* 한국과학기술정보연구원(Korea Institute of Science and Technology Information), koras@kisti.re.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 '국가과학기술 지식정보서비스 사업' 연구과제(과제번호: N24NM014-241)의 연구비 지원에 의함

# 과학기술 관련 질의 응답 과정에서의 사용자 경험향상을 위한 ChatGPT 기반 프롬프트 엔지니어링 기법 연구

이지민\*

## A Study on ChatGPT-based Prompt Engineering Techniques to Improve User Experience in Science and Technology-Related Q&A Process

Jimin Lee\*

본 논문은 대표적인 Large Language Model (LLM)의 ChatGPT를 활용하여, 과학기술 관련 질의 응답 과정에서 AI비전문가인 일반 사용자의 경험을 향상시키기 위한 프롬프트 엔지니어링 전략을 제안하고자 한다. LLM은 방대한 과학 문헌 정보를 빠른 속도로 리뷰하여 요약 내용을 제공하는 등 이례없이 효율적인 과학 정보 접근 방식을 제공하지만 [1], LLM의 매커니즘을 알지 못하는 일반사용자의 경우, 원하는 결과를 얻기 위한 프롬프트 작성에 어려움을 겪어 일반적이거나 정확하지 않은 응답 결과를 얻는 경우가 있다[2]. 본 연구는 LLM 사용자들이 겪는 주요 이슈를 관련 연구를 통해 분석하고, 정보 품질 기준 (정확성, 정밀성, 타당성, 신뢰성)[3]을 고려하여 사용자의 과학기술 관련 ChatGPT 응답 경험을 향상할 수 있는 프롬프트 엔지니어링 기법(Role prompting, Metacognitive Prompting 등) 적용 전략을 제안한다. 이를 바탕으로, 본 연구는 GPTs를 개발하여 과학기술 정보 습득 과정에서 비전문가 사용자들의 AI 활용 경험을 개선하기 위한 가능성과 효과를 탐구하였다.

### Key words

Large Language Model (LLM), ChatGPT, prompt engineering, scientific information, user experience

### 참고문헌

- [1] Birhane, Abeba, et al. "Science in the age of large language models." *Nature Reviews Physics* 5.5 (2023): 277-280.
- [2] J. Doe, A. Brown, and M. Green, "Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail)," *IEEE Transactions on AI*, vol. 12, no. 3, pp. 33-40, 2023.
- [3] Taleb, Ikbal, et al. "Big data quality framework: a holistic approach to continuous quality management." *Journal of Big Data* 8.1 (2021): 76.

---

\*한국과학기술정보연구원(KISTI), jimmin2024.lee@kisti.re.kr

※이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 자체사업으로 수행된 연구입니다.(과제번호: J24JR042-241)



# 원전 운전경험 및 교육 분야 써제스트 AI 서비스 설계 방법론

이수일\*, 이상현, 한상재, 김남현, 유기중, 조창현

## Design Methodology of Seargest Artificial Intelligence Services in the Fields of Operation Experience and Training in NPPs

Soo-Il Lee\*, Sang Hyun Lee, SangJae Han, Nam-Hyun Kim, Ki-Jung Ryu, Chang-Hyun Cho

생성형 AI의 이슈화와 함께 콘텐츠 산업, 유통 등의 분야에서는 사용자 맞춤형 기술인 초개인화(Hyper-Personalization)에 대한 관심이 급증하고 있으며, 유튜브, 넷플릭스 등에서 이러한 기술을 활용하고 있다 [1]. 국내 원전은 운전 뿐만 아니라 설비 고장정비, 고장 원인분석, 경향분석, 교육 등의 수행을 통해 수십년 간의 운전경험을 보유하고 있으며, 원전 운영효율성, 안전성 등 향상을 위해 과거 운전경험의 디지털변환을 통한 빅데이터 구축, 써제스트(Seargest) AI 서비스 개발 등을 추진하고 있다[2].

원전 운전경험 활용도 향상을 위해서 빅데이터 구축 및 AI 서비스 개발 단계에서부터 2가지 요건을 고려하고 있다. 첫 번째는 설비 고장정비 및 운전 업무의 모든 프로세스에 대한 적용 가능성을 고려하는 것이다. 설비 고장통지 - 운영개선 - 작업오더 - 원인분석 - 운전경험 관리 - 직무교육 등 발전소 설비 고장 및 운전 전 단계에서 활용가능하도록 빅데이터를 융합한 서비스를 제공하는 것이다. 두 번째는 사용자의 개인 이력, 경험 및 선호도 등을 추적하여 정확한 정보를 빠른 시간 내에 제공하는 것이다. 단순 키워드 검색 기법이 아닌 실시간으로 사용자 맞춤형 검색, 추천 및 계획 수립 기능을 제공한다.

현재 고려하고 있는 원전 운전경험 및 교육 분야의 고수준 기술데이터는 주로 원전 운전경험 보고서, 설비 고장조치 이력, 국내외 원자력분야 기술보고서, 기술현안 지원 경험 자료, 원전 직무 교육자료 등이다. 첫 번째 요건을 위해서는 데이터 디지털변환 및 빅데이터 구축단계에서 빅데이터를 서로 연계 융합하는 것이다. 일부 데이터는 설비 고유번호를 통해서 연계하고, 설비 고유번호가 없는 경우에는 의미기반 유사도를 통해 빅데이터 내 의미있는 정보를 상호 연계한다. 두 번째 요건을 위해서는 실시간 사용자 정보와 발전소 상태정보를 활용한다. 사용자 정보는 직무, 보직, 경력, 교육이력, 담당업무, 부서명, 검색 및 열람 로그 등이며, 발전소 상태정보는 발전소 운전상태, 설비 고장상태, 발전소 시험현황 및 기상상태 등으로 구성된다. 원전 운전경험 및 교육 분야 사용자 맞춤형 AI 서비스는 ① 빅데이터 융합모델, ② 사용자 AI 모델, ③ 콘텐츠 AI 모델과 ④ 사용자와 콘텐츠 모델 간의 통합 AI 모델로 구성되도록 설계하였다. 빅데이터 융합모델은 빅데이터 상호간과 각 AI 모델(② ~ ④)과 연계된다. 통합 AI 모델은 ②, ③ AI 모델을 통합하고, 실시간으로 결과를 서비스하는 모델이다.

본 논문은 원전 운전경험 및 교육분야에서 설비 고장정비 및 운전 업무의 모든 프로세스에서 활용 가능하며 사용자 맞춤형으로 실시간 정보를 제공하는 써제스트 AI 서비스 설계 방법론에 대해 제시하였고, 현재 이러한 방법론에 따라 빅데이터 구축 및 AI 서비스를 개발 중에 있다.

### Key words

NPP Operation Experience, NPP Training, Personalized AI Services, Hyper-Personalization, Seargest AI

### 참고문헌

- [1] Paul Covington, et al., "Deep Neural Networks for Youtube Recommendations," RecSys, 2016.
- [2] 이수일, et al., "원전 빅데이터 기반 사용자 맞춤형 AI 서비스 개발 방안," 한국산업정보학회 추계학술대회, 2023

\*한국수력원자력 중앙연구원, sooill.lee@khnp.co.kr

# 미국 “연구용 원자로를 위한 보안 사례집” 분석

김태희\*, 이현주\*\*, 이수봉\*\*\*,

## An Analysis of Effective Cybersecurity Practices for U.S. Research Reactors

Taehee Kim\*, Hyunjoo Lee\*\*, and Subong Lee\*\*\*

### 요 약

미국은 법률에 따라 상업용 원자력발전소에 대한 사이버보안 규제를 적용하고 있으나, 연구용 원자로 역시 방사선적 영향을 발생시킬 수 있음에 따라 보안 사례집을 배포하고 이를 따르도록 권고하고 있다. 이 논문에서는 향후 우리나라 원자력시설의 사이버보안 규제에 미국의 보안 사례집을 참고할 수 있도록 그 구성과 성격을 분석하여 제시한다.

### Abstract

The U.S. enforces cybersecurity regulations on commercial nuclear power plants through legal frameworks, However, recognizing the potential radiological consequences of research reactors, it also distributes effective practices and recommends adherence to these practice. This paper analyzes the structure and characteristics of U.S. research reactor practices and offers foundational insights that may inform future cybersecurity regulations for nuclear facilities in R.O.K.

### Key words

Non-power Reactor, NPR, cybersecurity, cybersecurity practice

### 1. 배경: 미국에서 상업용 원자력발전소와 연구용 원자로의 규제 접근 방식 차이

미국은 9.11테러[1] 이후, 원자력발전소의 사이버 보안 체계 강화를 위하여 관련 법률[2]과 규제 가이드[3]를 제정하고, 상업용 원자력발전소로 하여금 이를 준수하도록 규제하고 있다. 그러나, 상업용 원

자력발전소 뿐만 아니라 연구용 원자로에 대해서도 보안 사례집[4]을 발간하고 이를 지키도록 권고하고 있다.

이와는 다르게 우리나라는, 물론 규제 대상 원자력시설의 방사선적 영향에 따라 다소 간의 보안조치 차등 적용을 인정하고 있으나, 기본적으로는 상업용 원자력발전소와 연구용 원자로를 모두 같은

\*한국원자력통제기술원, kimtaehee@kinac.re.kr(교신저자), \*\*한국원자력통제기술원, sn220865@kinac.re.kr, \*\*\*한국원자력통제기술원, sblee@kinac.re.kr

방사능방재법의 테두리 안에서 규제하고 있다. 이는 어느 나라의 접근 방식이 절대적으로 옳고 그르다고 판단할 문제는 아닌 것으로 판단된다. 다만 원자력시설 사이버보안 규제를 선도하고 있는 미국의 사례를 살펴보는 것은 향후 우리나라의 규제 방향을 결정하는데 있어 참고가 될 수 있을 것이다.

이에 따라 본 논문에서는 미국에서 연구용 원자로를 대상으로 발간된 보안 사례집의 구조와 각 보안 사례들을 살펴본다.

## II. 보안 사례집의 구성

보안 사례집은 Executive Summary, Overview and Purpose와 용어집 등의 부록을 제외하면, 표 1과 같은 목차로 구성되어 있다. 목차에 제시된 각 섹션들은 각기 규제 가이드 대비 간략화 된 보안 조치 카테고리로서 볼 수 있으며, 각 카테고리별 보안조치 적용 필요성, 방법 등을 사례 위주로 설명하고 있다.

예를 들어 무선 보안조치와 관련하여, 규제 가이드의 요건(B.1.17 “Wireless Access Restrictions”)은 단순히 다음과 같은 요건을 기술하고 있다.

- 안전 관련 기능(safety-related) 또는 안전에 중요한(important-to-safety) 기능을 수행하는 필수

디지털자산에 대한 무선 기술의 사용을 금지한다.

- ※ 허용되지 않는 행위에 대해 명확히 금지하는 의미의 “prohibit” 사용

반면, 보안 사례집에서는 관련 내용을 단정적이기 보다는, 다음과 같이 완곡한 표현의 사례 위주로 기술하고 있다.

- 필수적이고 핵심적이며, 원자로 안전 기능(reactor-safety functions)을 위한 무선 통신의 사용은 가급적 지양한다. 이러한 기능을 위해서는 유선 연결을 사용해야 한다.

- ※ 특정 행동을 피하거나 하지 않는 것이 바람직하다는 의미의 “avoid” 사용

위 사례에 따르면 원자로 안전 기능 자체에 대한 무선 사용만 금지될 뿐 그 외 기능을 수행하는 필수 디지털자산에 대한 무선 사용이 원칙적으로 금지된 것은 아니라는 사실을 알 수 있게 된다.

또한, 안전 기능 자체에 대해서도 “지양한다”는 완곡한 표현이 사용됨에 따라 필요에 따라서는 무선 기술을 사용할 수 있는 여지를 남겨놓았다.

표 1. 미국 연구용 원자로를 위한 보안 사례집[4]의 목차 및 주요 내용

Table 1. Contents and Key Topics of Effective Cyber security Practices[4] for U.S. Research Reactors

Section	목차	주요 내용(요약)
4	Cyber assets that require protection	보호가 필요한 필수디지털자산의 식별과 보호기준 설정
5	Physical Security Requirements and Cyber Security	물리적 보안과 사이버보안의 통합 관리, 보안 조치의 상호 보완
6	Network architectural-interconnectivity issues	네트워크 구조와 연결성 문제를 고려한 사이버 공격에 대비한 아키텍처 설계
7	System administration and support issues	시스템 관리, 보안 패치, 접근 제어를 포함한 운영 지원 요구사항
8	Personnel cyber security training issues	직원 대상 사이버 보안 훈련 및 보안 인식 강화
9	Use of portable media and portable digital devices	이동식 미디어, 장치 사용 시 보안 위협 완화를 위한 정책과 절차
10	Use of digital I&C technologies, including wireless versions	무선 및 디지털 제어 기술 사용 시 보안 요건과 적용 가이드라인
11	Upgrades, replacements, and retrofits	시스템 업그레이드와 개보수 시 보안 요건 준수 및 관리 조치

다른 예시로, 휴대용 저장매체(USB 메모리 스틱 등)와 관련하여, 규제 가이드의 요건(B.1.19 “Access Control for Portable Media and Mobile Devices”, C.3.3 “Malicious Code Protection”)에는 “휴대용 저장 매체는 하나의 보안 등급(security level)에서만 사용되어야 한다.”는 요건과 함께 “악성코드가 검색되어야 한다.”는 요건이 제시되어 있어 악성코드 검색을 위한 상세 방법을 제시하고 있지 않다.

반면, 보안 사례집에서는 보안 등급의 개념 자체가 도입되어 있지 않으며, 악성코드 검사에 대해서는 별도의 지정된 시스템(“a separate system designated and configured for that purpose”)에서 수행하도록 명확한 지침을 제시하고 있다.

### III. 결론

이와 같이 보안 사례집은 표 1의 카테고리 8개에 대해 각 81개의 사례를 제시하고 있다. 이는 단순히 요건만을 제시하고 있는 규제 가이드와는 다른 접근 방식이다.

이러한 접근 방식은 아직 연구용 원자로가 사이버보안 관련 법률[2]에 포함되지 않은 미국의 특성을 반영한 것으로 보이며, 향후 연구용 원자로가 법률에 명시적으로 포함되게 될 경우에는 기존의 규제 가이드와 같이 연구용 원자로가 도달해야 할 요건만 제시하는 방향으로 전환될 가능성이 높다.

우리나라는 앞서 “I. 배경”에서 살펴본 바와 같이 다수의 원자력시설을 규제 대상에 포함하고 있기 때문에, 규제 사례집 보다는 미국의 규제 가이드에 해당하는 심·검사 지침서만을 규제에 적용하고 있다. 그러나 각 원자력시설의 방사선적 영향에 따라 향후 미국과 같은 보안 사례집의 발간 역시 검토해볼 필요성이 있을 것으로 판단된다.

### 참 고 문 헌

- [1] S. D. Brunn, Ed., *11 September and Its Aftermath: The Geopolitics of Terror*. London: Routledge, 2004.
- [2] U.S. Nuclear Regulatory Commission, *Protection of digital computer and communication systems*

*and networks*, Code of Federal Regulations, Title 10, Part 73.54, 2009.

- [3] U.S. Nuclear Regulatory Commission, *Cyber security programs for nuclear facilities*, Regulatory Guide 5.71, Rev. 1, 2023.
- [4] U.S. Nuclear Regulatory Commission, *Effective Practices for the Establishment and Maintenance of Adequate Cyber Security at Non-Power (Research and Test) Reactor Facilities*, Jan. 2016. Available: <https://www.nrc.gov>.

# 클라우드 서비스 품질평가를 위한 ISO/IEC 25023 기반 시험방법 적용 및 개선에 관한 연구

최재은

## Improvement of Test Methods for Cloud Service Based on ISO/IEC 25023

Jaeun Choi

### 요 약

국제표준 ISO/IEC 25023 평가모델은 기능적합성, 성능효율성, 호환성, 사용성, 신뢰성, 보안성, 유지보수성, 이식성 총 8가지 품질 특성을 기반으로 소프트웨어 평가 방법을 제시한다. 그러나 클라우드 서비스는 ISO/IEC 25023에서 제시한 소프트웨어 품질 특성 외에도 확장성, 탄력성, 서비스 측정성 등 클라우드 서비스에 특화된 품질 특성이 필요하므로, 기존 소프트웨어 평가모델을 클라우드 서비스에 그대로 적용하는 데 한계가 있을 수 있다. 따라서 클라우드 서비스를 평가하기 위해서는 기존 소프트웨어 품질 특성을 클라우드 특성에 맞게 수정하여 적용된 'ISO/IEC TS 25052-1 평가모델'에 따라 소프트웨어와는 다른 클라우드 독자적인 품질평가 방법이 제공되어야 한다.

본 논문에서는 ISO/IEC 25023 평가모델과 클라우드 서비스를 평가하는 ISO/IEC TS 25052-1 평가모델을 비교 분석하여 각 평가모델에서 제공되는 대표적인 주특성 및 부특성을 살펴보고, 클라우드 서비스에 맞는 테스트 시나리오를 제시한다.

### Abstract

The international standard ISO/IEC 25023 evaluation model provides a software evaluation methodology based on eight quality characteristics: functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, and portability. However, cloud services require additional cloud-specific quality characteristics, such as scalability, elasticity, and service measurability, beyond those presented in ISO/IEC 25023 for software quality. This presents limitations in directly applying the existing software evaluation model to cloud services. Therefore, to assess cloud services, it is essential to modify traditional software quality characteristics to align with cloud-specific traits and adopt the "ISO/IEC TS 25052-1 evaluation model" that offers a distinct quality evaluation approach tailored to cloud services.

This paper compares and analyzes the ISO/IEC 25023 evaluation model with the ISO/IEC TS 25052-1 evaluation model for cloud service evaluation, examining the primary and subcharacteristics provided by each model. It also proposes test scenarios tailored to cloud services.

### Key words

ISO/IEC 25023 evaluation model, cloud-specific quality characteristics, test scenarios for cloud services

---

\* 한국정보통신기술협회, kkseru10@tta.or.kr

## 1. 서 론

현대 사회에서 소프트웨어는 필수 요소로 자리 잡고 있다. 그중 '클라우드 서비스'는 최적의 리소스를 활용해 소프트웨어를 개발할 수 있는 주요 방안으로 현재 많은 주목을 받고 있다.

멀티 클라우드 활용 현황 조사(Multicloud in the Mainstreams, 451 Research, 2023)에 따르면, 북미 지역에서 정규직 직원 1,000명 이상을 보유하거나 다른 지역에서 500명 이상의 직원을 보유한 기업 응답자 1,500명 중 약 98%가 2개 이상의 클라우드 인프라 서비스를 활용하거나 이를 계획하고 있다고 밝혔다. 또한, 전 세계적으로 클라우드 서비스 시장이 확대되면서 클라우드 사업은 아마존, 구글, MS와 같은 글로벌 빅테크 기업들의 핵심 분야로 자리 잡았다. 해당 사업의 규모는 2033년까지 약 2,974억 달러에 이를 것으로 예상되며, 이러한 흐름 속 많은 소프트웨어 기업은 클라우드 서비스의 강점을 바탕으로 소프트웨어 개발 환경을 개선하는 데 주력하고 있다.

소프트웨어의 이용률이 증가함에 따라, 소프트웨어 품질 평가 기준의 중요성도 함께 커지고 있다. 국제표준 'ISO/IEC 25023[1](시스템/소프트웨어 제품 품질)'은 기능 적합성(Functional suitability), 성능 효율성(Performance efficiency), 호환성(Compatibility), 사용성(Usability), 신뢰성(Reliability), 보안성(Security), 유지보수성(Maintainability), 이식성(Portability) 총 8가지로 구성된 주요 특성(주특성)과 하위 특성(부특성)을 소프트웨어 품질측정을 위한 평가모델로 제시하고 있다. ISO/IEC 25023 기반 평가모델은 소프트웨어 서비스 품질을 객관적이고 일관되게 평가할 수 있는 기준을 제공함으로써 소비자에게 높은 품질의 소프트웨어를 효과적으로 제공하는 데 기여한다. 그러나 클라우드 서비스의 경우, 기존 패키지형 소프트웨어와는 다른 클라우드 컴퓨팅만의 특징이 있으므로 ISO/IEC 25023 기반 평가모델을 개선할 필요성이 있다.

## II. 관련 연구

이런 필요성으로 인해 국제 표준화 기구인 ISO에서는 기존 ISO/IEC 25023 기반 평가 모델을 클라우드 서비스에 맞게 작성하여 ISO/IEC TS 25052-1[2](클라우드 서비스 품질)을 발표하였다. ISO/IEC TS 25052-1에는 그림 1과 같이 1)기존 ISO/IEC 25023 평가모델에 더해 '결합된 자원 이용률(Aggregated resource utilization)', '확장성(Scalability)', '접근성(Accessibility)', '서비스 측정성(Service Measurability)', '셀프서비스 지원성(Self-provisioning)' 등 클라우드 서비스에 특화된 품질 특성이 추가되었고, 2)기존 소프트웨어 특성을 클라우드 환경에 맞게 조정하였다. 그림 1에서 유사 항목은 노란색, 추가 항목은 분홍색, 클라우드 서비스 품질에서 삭제된 항목은 회색으로 표시하였다.



그림 1. ISO/IEC 25023과 ISO/IEC TS 25052-1 품질평가모델 비교

Fig. 1. Comparison of ISO/IEC 25023 and ISO/IEC TS 25052-1 Quality Evaluation Models

## III. 개선된 주요 부특성

### 2.1. 유사한 주요 부특성

#### 2.1.1. 자원효율성 VS 결합된 자원 이용률

ISO/IEC 25023에서 자원 효율성이란 소프트웨어 제품 기능 실행 시 제품이 사용하는 자원의 양이 요구사항을 충족하는 정도를 나타내는 부특성이다. 이와 달리 ISO/IEC TS 25052-1에서는 해당 부특성이 결합된 '결합된 자원 이용률(Aggregated resource utilization)'로 명칭이 변경되어 가상화 기술을 통해 결합된 가상 자원(예: CPU, 메모리, 디스크 등)의 활용도를 평가하는 항목으로 그 의미가 변경되었다.

표 1. 결합된 자원 이용률 테스트 시나리오(예)

Table 1. Aggregated Resource Utilization Test Scenario (Example)

목적	여러 테넌트가 동시에 클라우드 서비스에서 자원을 할당받고 사용할 때 리소스의 이용률을 측정
시나리오	<ol style="list-style-type: none"> <li>1. 여러 개의 테넌트가 클라우드 환경에서 다양한 크기의 리소스를 요청하도록 함</li> <li>2. 각 테넌트는 CPU, 메모리, 네트워크 대역폭 등의 자원을 할당받음</li> <li>3. 각 테넌트가 요청한 작업 실행 후, 자원의 이용률(CPU, 메모리, 네트워크 사용량 등)을 모니터링</li> <li>4. 전체 리소스 풀에서 사용 중인 리소스의 백분율을 계산하여 결합된 자원 이용률 측정</li> </ol>

표 2. 서비스 상호운용성 테스트 시나리오(예)

Table 2. Cloud Interoperability Test Scenario (Example)

목적	다른 클라우드 서비스 간 데이터 이동의 상호운용성을 평가
시나리오	<ol style="list-style-type: none"> <li>1. A 클라우드 서비스에 저장된 데이터를 B 클라우드 서비스로 이동</li> <li>2. 데이터 포맷과 전송 프로토콜에 차이가 있더라도 무결하게 데이터가 전송되고 활용이 가능한지 확인</li> <li>3. 이동한 데이터가 B 클라우드 서비스에서 그대로 사용될 수 있는지 확인</li> <li>4. 데이터 이동 중 발생할 수 있는 호환성 문제(인코딩, 파일 구조 등)를 테스트하고 오류 발생 여부를 기록</li> </ol>

표 3. 확장성 테스트 시나리오(예)

Table 3. Scalability Test Scenario (Example)

목적	자원이 부족한 상황에서도 서비스가 어떻게 유지되고, 사용자가 느끼는 성능 저하가 얼마나 발생하는지 평가
시나리오	<ol style="list-style-type: none"> <li>1. 특정 시점에 다수의 사용자가 동시에 접속하여 리소스 한계에 도달하도록 시스템에 의도적으로 과부하 상태를 유도(예: CPU 사용량 90% 이상, 메모리 거의 고갈 등)</li> <li>2. 자원이 부족한 상황에서 각 사용자의 가상 데스크톱 환경이 어떻게 반응하는지 확인(예: 응답 지연, 화면 끊김, 접속 실패 등)</li> <li>3. 시스템이 동적으로 자원을 할당하거나 대체할 수 있는지 확인(예: 새로운 가상머신을 추가하거나 (스케일 아웃), 현재 사용 중인 자원의 리소스를 더 확보(스케일 업) 등)</li> </ol>

### 2.1.2. 상호운용성 VS 서비스 상호운용성

ISO/IEC 25023에서 상호운용성(Interoperability)은 서로 다른 소프트웨어 시스템이나 컴포넌트가 데이터를 교환하고, 교환된 정보를 성공적으로 사용하는 것을 의미한다. 예를 들어, 두 개의 소프트웨어 시스템이 서로 다른 데이터베이스를 사용하더라도, 데이터를 성공적으로 교환하거나 연동할 수 있다면 이 시스템들은 상호운용성을 만족한다고 할 수 있다.

ISO/IEC TS 25052-1에서 서비스 상호운용성(Cloud Interoperability)이란 서로 다른 클라우드 서비스 간에 데이터가 교환되는지 확인하기 위한 부특성이다. 예를 들어, 사용자가 Google Drive에 저장된 파일을 Dropbox로 이동하거나 복사할 수 있다면 해당 서비스는 서비스 상호운용성을 만족한다고 볼

수 있다.

## 2.2. 추가된 주요 부특성

### 2.2.1. 확장성

ISO/IEC TS 25052-1에서 확장성(Scalability)은 클라우드 서비스가 물리적 및 가상 자원을 확장하거나 축소할 수 있는 능력을 측정한다. 예를 들어, 여러 사용자가 가상 데스크톱을 생성할 때 클라우드 서비스의 CPU 사용률, 메모리 사용량, 처리량(TPS) 등을 측정함으로써 가상 자원이 수평적으로 확장되는지(스케일 아웃, Scale-out) 또는 수직적으로 확장되는지(스케일 업, Scale-up)를 확인함으로써 확장성을 평가할 수 있다.

표 4. 탄력성 테스트 시나리오(예)

Table 4. Elasticity Test Scenario (Example)

목적	갑작스러운 트래픽 증가에 따른 자원의 확장 속도 및 정확도 측정
시나리오	1. 예기치 않게 사용자 수를 증가시키는 시뮬레이션 실행 후, 자원 확장 요청까지 발생하는 시간을 기록 2. 자원 확장이 완료된 시간을 기록 3. 확장된 자원이 사용자 요구를 충족하는지 확인

표 5. 고립성 테스트 시나리오(예)

Table 5. Isolation Test Scenario (Example)

목적	클라우드 환경에서 서로 다른 테넌트 간의 데이터 분리가 제대로 이루어지는지 확인
시나리오	1. 클라우드 서비스에 두 개 이상의 테넌트(예: Tenant A, Tenant B)를 설정하여 각 테넌트에 로그인할 수 있는 사용자 계정 및 독립적인 데이터베이스/애플리케이션 인스턴스 등을 준비 2. 두 테넌트에서 동시에 데이터를 입력하고 조회하는 시나리오를 실행하여 기능 동작 확인(예: 동시 요청 시, 결과값에 다른 테넌트의 데이터가 포함되지 않는지 확인 등)

스케일 아웃(Scale-out)이란 기존 서버에 추가 서버를 더해 자원을 확장하는 방식이다. 한 대의 서버로 처리할 수 없는 작업을 여러 대의 서버로 분산시켜 더 많은 사용자 요청을 처리하는 것이다. 반면, 스케일 업(Scale-up)은 하나의 서버에 더 강력한 CPU나 메모리를 추가하여 성능을 향상시키는 방식이다. 즉, 같은 서버에서 성능을 높여 더 많은 작업을 처리할 수 있게 만드는 것이다.

### 2.2.2. 탄력성

다음으로, ISO/IEC TS 25052-1에서 탄력성(Elasticity)이란 클라우드 서비스가 계약 조건에 따라 필요할 때 가상 자원을 사용자가 즉시 사용할 수 있도록 제공하는지를 측정한다. 탄력성이 제공되지 않는 클라우드 서비스의 경우, 갑작스러운 부하가 발생했을 때 필요한 가상 자원을 사용자에게 신속하게 제공하지 못기 때문에 자원이 확장되기 전까지 사용자가 서비스를 이용할 수 없는 문제가 발생할 수 있다. 해당 부특성은 사용자가 자원을 수동으로 확장하거나 서비스 내에서 자동 확장이 이루어질 때, 확장 속도 및 정확도를 측정함으로써 평가할 수 있다.

### 2.2.3. 고립성

ISO/IEC TS 25052-1에서 고립성(Isolation)이란, 클라우드 환경에서 여러 사용자나 테넌트(tenant)가 동일한 물리적 자원을 공유하더라도 각 사용자나 테

넌트의 데이터가 안전하게 분리되어 다른 사용자에게 영향을 주지 않는지를 측정하는 보안 부특성이다. 테넌트는 클라우드 컴퓨팅에서 서비스를 이용하는 개별 사용자나 조직을 의미한다. 고립성은 멀티 테넌트 환경에서 매우 중요한 요소로, 자원이 공유되는 동안 클라우드 서비스에서 연산과 데이터가 격리되고 다른 테넌트가 접근할 수 없는지를 확인하여 그 정도를 평가한다.

예를 들어 각 테넌트의 가상 머신(VM)이나 컨테이너가 같은 물리적 서버에서 실행될 때 한 테넌트가 다른 테넌트의 데이터나 자원에 접근할 수 없는지를 확인하여 고립성을 테스트할 수 있다. 또한, 데이터 암호화, 가상 네트워크 분리, 보안 로그 모니터링을 사용하여 다른 테넌트가 의도치 않게 시스템의 다른 부분에 접근하는 일이 없는지를 확인하여 평가할 수도 있다.

## III. 결론

본 논문에서는 국제표준 ISO/IEC 25023과 클라우드 서비스를 평가하는 ISO/IEC TS 25052-1을 비교 분석하였다. 또한, 기존 소프트웨어 품질 모델에서 변형된 클라우드 품질 모델의 부특성을 비교하였다. 아울러 클라우드 품질 모델만의 독자적인 테스트 시나리오를 제시하였다. 향후 클라우드 품질 평가 방안을 활용하여 다양한 클라우드 제품을 테스트 적용할 계획이다.



참 고 문 헌

- [1] ISO/IEC 25023, “Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of system and software product quality”, 2016
- [2] ISO/IEC TS 25052-1 “Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE): cloud services”, 2022

# 양자 컴퓨터 및 양자 내성 암호 NIST 표준화 동향

최승주\*, 이승민\*, 김도현\*, 장예지\*

## Trend of Quantum Computer and NIST Post-Quantum Cryptography Standardization

Seung-Joo Choi\*, Seung-Min Lee\*, Do-Hyun Kim\*, and Ye-Ji Jang\*

### 요 약

최근 양자 컴퓨터 개발이 활발하게 진행됨에 따라, 양자 컴퓨터 시대에 대비한 양자 내성 암호에 대한 연구도 활발히 이루어지고 있다. 이에 발맞추어 미국 국립표준기술연구소(NIST)는 2016년 10월 20일부터 양자 내성 암호(Post-Quantum Cryptography, PQC)의 표준화 작업을 시작했으며, 2024년 8월 13일에는 양자 암호 표준을 발표했다. 이 논문에서는 양자 컴퓨터와 양자 내성 암호 표준화 연구의 최신 동향을 다루고자 한다.

### Abstract

With the rapid development of quantum computers, research on post-quantum cryptography is also progressing actively in preparation for the quantum era. In line with this, the U.S. National Institute of Standards and Technology (NIST) initiated a standardization project for post-quantum cryptography on October 20, 2016, and officially announced the quantum-resistant cryptography standards on August 13, 2024. This paper aims to explore the latest trends in the quantum computer and standardization of post-quantum cryptography.

### Key words

quantum computing, post quantum cryptography, national institute of standards and technology

## 1. 서 론

최근 양자 컴퓨터의 활발한 개발과 더불어 양자 내성 암호에 대한 많은 연구가 이루어지고 있다. 미국 국립 표준 기술 연구소(National Institute of Standards and Technology, 이하 NIST)에서는 2016년 양자 내성 암호 표준 계획안을 발표하였고 8년간의 평가 기간을 거쳐 세계 여러 각국에서 제출한 양자 내성

암호 표준 후보를 선출해 나갔다. 이런 평가 기간 동안 표준 공모전에 제안된 기법들에 소프트웨어 최적화, 안전성 분석 등 여러 가지 관점에서 연구가 이루어졌으며, 2024년 8월 13일 드디어 양자 암호에 대한 연방 정보 처리 표준(FIPS)을 발표하였다. 본 논문에서는 발표된 미국 NIST 양자 내성 암호 표준에 대해 살펴보고자 한다.

\* 정보통신기획평가원 (Institute of Information & Communications Technology Planning & Evaluation), bs3@iitp.kr

## II. 양자 컴퓨터와 큐비트

양자 컴퓨터 개념은 1980년대 초, R. 파인만이 양자 물리학의 원리를 컴퓨터에 적용할 수 있다는 아이디어를 제안하면서 등장했다[1]. 이 아이디어는 양자 물리학의 중첩 상태를 기존 디지털 컴퓨터로 계산하는 데 한계에 봉착했기 때문이다. 비록 처음에는 이러한 제안이 큰 주목을 받지 못했지만, 1990년대에 소인수 분해를 위한 양자 알고리즘이 제안되면서 이를 실제로 구현하려는 연구가 활발해졌다.

기존의 디지털 컴퓨터는 0과 1이라는 두 가지 상태를 비트로 표현하고 처리한다. 반면에 양자 컴퓨터는 큐비트라는 단위를 사용하며, 이 큐비트는 0과 1을 동시에 가질 수 있는 중첩 상태를 특징으로 한다. 이로 인해 큐비트는 더 많은 정보를 한 번에 처리할 수 있으며 양자 컴퓨터의 큐비트는 얽힘 현상을 통해 서로 연결되어 있는데, 한 큐비트의 상태가 결정되면 얽혀 있는 다른 큐비트의 상태도 자동으로 결정된다. 이런 특성 덕분에 양자 컴퓨터는 고전적인 컴퓨터보다 훨씬 더 빠르고 효율적으로 문제를 해결할 수 있게 된다.

이러한 특성 덕분에 기존 컴퓨터로는 수백 년이 걸릴 연산도 양자 컴퓨터는 병렬 처리 방식을 통해 몇 분 안에 해결할 수 있다. 대표적인 사례로는 수학자 쇼어(Shor)가 제안한 Shor 알고리즘이 있는데, 이 알고리즘은 중첩된 상태의 숫자에 푸리에 변환을 적용해 한 번의 연산으로 소인수 분해에 필요한 주기를 계산할 수 있다. 만약 양자 컴퓨터가 널리 보급된다면, 기존의 공개키 암호화 방식은 큰 위협을 받게 될 것이다.

## III. 양자 컴퓨터 개발 현황

양자 컴퓨터를 개발하려는 세계적인 기업들의 노력은 계속 해서 진행 중이다. 마이크로소프트는 Q#이라는 양자 프로그래밍 언어를 출시했으며, Azure Quantum이라는 클라우드 플랫폼을 통해 다양한 양자 하드웨어에 접근할 수 있는 환경을 제공하고 있

다. 구글은 70큐비트를 탑재한 Sycamore 프로세서를 보유하고 있으며, IBM은 433큐비트의 Osprey 프로세서를 선보였다. 최근 우리나라에서도 8큐비트 칩을 개발하는 등 양자 컴퓨터를 개발하기 위한 노력은 활발하게 이뤄지고 있다.

다만 이러한 양자 컴퓨터는 전자를 기반으로 논리적 큐비트를 구성하므로, 외부 환경의 미세한 변화에도 매우 취약하다는 특징이 있다. 이러한 특징과 상황 때문에 양자 컴퓨터를 개발하고 있는 기업과 기관들은 큐비트의 수를 늘리는 것뿐만 아니라 양자 프로세서의 오류율을 감소시키는 데에도 힘쓰며 양자 컴퓨터 시대를 준비하고 있다.

## IV. NIST 양자내성암호 표준화 현황

2016년 초 미국 국립표준기술연구소 NIST는 양자 컴퓨터의 도래가 기존의 많은 암호 알고리즘의 안전성을 위협할 것이라는 초안을 발표하였다[2]. NIST는 양자 내성 암호 표준화 공모전을 <표1>과 같이 진행하며 세계 각국의 암호 연구 기관에 양자 내성 암호 표준 후보를 선출하였고 공모전에는 한국을 포함한 다양한 기관이 참여를 하였다.

표 1. NIST 양자 내성 암호 공모전 진행현황

Table 1. NIST Post Quantum Cryptography Submission History

일정	내용
2016.02	NIST 양자 내성 암호 공모전 발표
2016.10	양자 내성 암호 후보군 요청
2017.12	Round 1 양자 내성 암호 후보 발표
2019.01	Round 2 양자 내성 암호 후보 발표
2020.07	Round 3 양자 내성 암호 후보 발표
2022.07	양자 내성 암호 4개 후보군 발표
2023.08	양자 내성 암호 표준 초안 공청회
2024.08	양자 내성 암호 표준 발표

제출된 알고리즘의 유형으로는 부호 기반, 격자 기반 또 다변수 다항식, 해시 기반 등 다양한 유형의 알고리즘이 제출되었다. 오랜 공모 끝에 2024년 8월 공개키 암호화를 위한 1개의 알고리즘과 전자 서명을 위한 2개의 알고리즘이 표준으로 발표되었다[3].

표 2. 발표된 양자 내성 암호 표준  
Table 2. Announced Post Quantum Cryptography Standards

알고리즘	표준	용도
격자기반 키 암호화	FIPS 203 ML-KEM	키 교환
격자기반 전자서명	FIPS 204 ML-DSA	전자서명
해시기반 전자서명	FIPS 205 SLH-DSA	해시기반 전자서명

일반 암호화에 대한 기본 표준 방식으로 선출이 된 ML-KEM은 격자 기반 알고리즘인 CRYSTALS-Kyber을 기반으로 만들어졌다. ML-KEM이 최종적으로 선출이 된 주요 이유로는 비교적 작은 암호화 키 사이즈, 알고리즘 동작 속도, 그리고 암호화 시 여러 매개변수 세트(ML-KEM-512, ML-KEM-768, ML-KEM-1024)를 제공하여 사용자가 필요한 보안 수준에 맞춰 연산 수준을 유연하게 선택할 수 있게 제공한다는 점이다.

전자 서명을 위한 기본 표준 방식으로는 격자 기반 알고리즘인 CRYSTALS-Dilithium을 기반으로 ML-DSA가 선출이 되었는데 미래에 ML-DSA 방식에 취약점이 발견될 경우를 대비하여 SLH-DSA도 같이 선출 하였다. 다만 전자 서명 알고리즘에 제안에 대한 추가적인 공모는 아직 진행되고 있다[4]. 진행되고 있는 공모에서는 격자 기반 서명 방식보다 서명 길이도 짧고 처리 속도도 더 빨라 실시간 통신이나 고빈도 거래를 처리하는 시스템 등에 적용할 수 있는 방식을 계속해서 모색하고 있다.

## V. 결 론

양자 컴퓨터 시대가 다가옴에 따라 기존 시스템에 양자 내성 암호의 빠른 적용이 필요해진 상황이다. 이에 미국 국립 표준 기술 연구소에서는 양자 내성 암호 공모전을 개최하여 공개키 암호 알고리즘 1개, 전자 서명 알고리즘 2개, 총 3개의 표준을 발표 하였다. 이런 암호의 선출 기준으로는 암호의 안전성과 현실적인 암호의 실용을 고려한 최적화를 두고 있다.

양자 컴퓨터의 시대가 현실적으로 다가오고 있는 지금 앞으로도 계속해서 진행되는 양자 내성 암호 표준화 사업에 관심을 갖고 지켜봐야 할 것이다.

## 참 고 문 헌

- [1] R. Feynman, Simulating Physics with Computers, International Journal of Theoretical Physics 21, 467, 1982.
- [2] Lily Chen, Stephen Jordan, Yi-Kai Liu, Dustin Moody, Rene Peralta, Ray Perlner, Daniel Smith-Tone, Report on Post-Quantum Cryptography, NIST IR 8015 DRAFT, 2016.
- [3] NIST ‘Announcing Approval of Three Federal Information Processing Standards (FIPS) for Post-Quantum Cryptography’, available at <https://csrc.nist.gov/news/2024/postquantum-cryptography-fips-approved>
- [4] NIST ‘Request for Additional Digital Signature Schemes for the Post-Quantum Cryptography Standardization Process’, available at <https://csrc.nist.gov/News/2022/request-additional-pqc-digital-signature-schemes>

# 중국 인공지능산업정책 확산 메커니즘

이 경 우\*

## Mechanism of China Artificial Intelligence Industry Policy Diffusion

Kyungwoo Lee\*

본 연구의 목적은 미국과 더불어 세계 인공지능산업을 견인하고 있는 중국 인공지능산업의 구조를 살펴보고, 산업확산에 영향을 미치는 정책적 요인은 무엇이며, 어떠한 메커니즘을 통해 확산하는가 규명하는 것이다. 이를 위하여 중국 정부에서 2017년 제정한 통합 인공지능정책인 ‘차세대 인공지능 발전계획’을 기준으로 하여 2017년부터 2023년까지 31개 성·시를 대상으로 20종의 패널데이터를 수집하였다. 수집된 데이터는 국제적으로 활용되는 정책확산분석 모형인 Berry & Berry(1990)의 ‘Unifide Analysis Model’에 적용하여 이산시간 사건사분석법(discrete-time event history analysis: EHA)을 통해 실증 분석하였다.

본 연구에서는 먼저 문헌분석을 통하여 기초층, 기술층, 응용층에 세부 30여 종의 기술로 구성된 중국 인공지능산업사슬 구조에 최근의 기술 변화의 요인들을 반영하였다. 또한, 중국 인공지능산업의 확산에 미치는 정책 요인으로 10개를 선별하여 가설을 정하고 EHA 분석으로 검증하였으며, 분석 결과 지도자의 정치적 영향, 지리적위치, 재정자립률, 소득수준, 정보화수준, 기술수준이 주요 요인임을 찾고 각 요인 간의 상대적 크기를 비교하였다. 아울러, 이러한 확산에 작용하는 메커니즘으로 ‘약한 규범형(weak instruction)’의 수직 메커니즘이 작동함을 정책 네트워크 분석과 정책내용 분석을 통하여 규명하였다.

### Key words

artificial intelligence, China AI industry, policy diffusion, diffusion mechanism, event history analysis

### 참고문헌

- [1] Berry F. S. Berry W. D., “State Lottery Adoptions as Policy Innovations: An Event History Analysis”, *The American Political Science Review*, 84(2), pp. 395-415, 1999
- [2] Walker, J. L., “The Diffusion of Innovations among the American States”, *American Political Science Review*, 63(03), pp. 880-889, 1969
- [2] 김진용, “중국 인공지능 발전과 지방 행위자”, *국방연구* 제65권 제4호, pp. 87-117, 2022
- [2] 王浦劬, 赖先进, “中国公共政策扩散的模式与机制分析”, *北京大学学报* 第50卷 第6期, pp. 14-23, 2013
- [3] 刘红波, 林彬, “人工智能政策扩散的机制与路径研究—一个类型学的分析视角”, *中国行政管理* 2019年第4期, pp. 38-45, 2019

---

\* 연세대학교, kwlee7226@naver.com

# 국가R&D정보서비스에 대한 개인정보 영향평가 사례 연구 -NTIS 개인정보 영향평가 사례를 중심으로-

박남규\*, 박근희\*\*, 신동구\*\*\*

## A Case Study on Privacy Impact Assessment for National R&D Information Services -Focus on NTIS Privacy Impact Assessment case-

Namkyu Park\*, Keun-Hee Park\*\*, and Dong-Gu Shin\*\*\*

NTIS(National Science & Technology Information Service)는 대국민 국가R&D지식포털서비스로써 국가R&D 사업 1.7만건, 과제 112.8만건, 국가연구개발성과 751.2만건 등 국가과학기술 분야 전반에 걸친 주요 정보와 더불어 회원정보 20.8만건, 연구자정보 82.4만건, 국가연구자번호 223.5만건의 개인정보를 보관·운영해오고 있다[1]. 개인정보보호법에서는 100만명 이상의 정보주체에 관한 개인정보 파일 등을 보유한 경우에 개인정보 영향평가 대상으로 규정하고 있으며, 잠재적인 개인정보 침해 위험 요인을 미리 분석하고 개선방안을 도출해 안전한 개인정보 처리 과정 설계를 유도하고 있다[2]. 본 연구에서는 NTIS를 통하여 국가R&D 사업 및 과제를 수행하는 회원들과 연구자들의 개인정보 흐름과 유통과정에 대하여 2024년 개인정보 영향평가를 수행하여 보안 위험 요소를 분석하였다. 4개 영역, 21개 분야, 총 77개 항목에 대한 평가를 수행한 결과 이행률은 99.2%로 매우 우수한 수준으로 분석되었다. 이를 통하여 NTIS가 보관·운영하고 있는 회원정보, 연구자정보, 국가연구자번호 3개 개인정보에 대한 지속적인 개인정보 보호 활동을 수행하고 있을 뿐만 아니라, 보완이 요구되는 사항에 대한 개선 의지 또한 높은 상태임을 확인할 수 있다.

### Key words

Privacy Impact Assessment (PIA), Privacy, National R&D, Information Security

### 참고문헌

- [1] 국가과학기술 지식정보서비스(NTIS), <https://www.ntis.go.kr>
- [2] 개인정보보호위원회, “개인정보 보호법” [시행 2024. 3. 15.] [법률 제19234호, 2023. 3. 14., 일부개정]
- [3] 개인정보보호위원회&한국인터넷진흥원, “개인정보 영향평가 수행 안내서”, 2024
- [4] 정해산, “공공기관 개인정보영향평가제도의 실효성 확보방안 연구”, 석사학위 청구논문:고려대학교, 2018
- [5] 정우송, “클라우드 환경에 적합한 개인정보 영향평가 기법 C-PIA”, 석사학위 청구논문:송실대학교, 2023.6

\*한국과학기술정보연구원, namkyu.park@kisti.re.kr(제1저자), \*\*한국과학기술정보연구원, pkh7514@kisti.re.kr,

\*\*\*한국과학기술정보연구원, lovesin@kisti.re.kr(교신저자)

※ 이 논문은 2024년도 과학기술정보통신부 『국가과학기술 지식정보서비스』 사업의 지원을 받아 수행된 연구입니다.(NTIS 과제번호: 1711195990)

# 나노팩 공정·장비 데이터 공유 플랫폼의 설계

이민호\*, 신성호\*\*

## Design of a NanoFab Process and Equipment Data Sharing Platform

Minho Lee\* and Sungho Shin\*\*

### 요 약

디지털 전환의 가속화는 나노 기술과 반도체 산업에서 공정·장비 데이터의 공유 및 관리 플랫폼 필요성을 높이고 있다. 특히 스마트 나노팩에서 자동화 및 디지털화된 공정 데이터를 활용하여 연구 효율성을 극대화하고 데이터 접근성을 향상 시키는 방안이 필수적이다. 본 논문에서 제안하는 플랫폼은 데이터를 자동 수집하고 메타데이터를 통합 관리하며, 다양한 필터링과 검색 기능을 통해 연구자들이 특정 조건에 맞는 데이터를 찾을 수 있도록 한다. 또한, 연구자들은 플랫폼을 통해 데이터 다운로드 또는 분석 공간 할당과 같은 기능을 활용하여, 연구 목적에 맞는 분석을 수행할 수 있다. 본 논문은 공정 및 장비 데이터의 체계적 관리와 분석을 통해 연구 효율을 높이고, 협력 연구 기반을 마련하는 스마트 나노팩 데이터 플랫폼의 설계를 목표로 한다.

### Abstract

The acceleration of digital transformation has increased the need for platforms to share and manage process and equipment data within nanotechnology and semiconductor industries. Specifically, utilizing automated and digitized process data from smart nano fabs is essential to maximize research efficiency and enhance data accessibility. This platform automates data collection, manages metadata integratively, and offers diverse filtering and search functions, enabling researchers to locate data that meets specific conditions. Researchers can also leverage platform features such as data download and dedicated analysis spaces to perform research-specific analyses. This paper aims to design a smart nano fab data platform that facilitates systematic management and analysis of process and equipment data, ultimately enhancing research efficiency and establishing a foundation for collaborative research.

### Key words

Digital Transformation; Smart Nano Fab ;Data Sharing Platform; Research Efficiency

## 1. 서 론

산업 전반에서 디지털 전환이 가속화됨에 따라 나노 기술 분야의 데이터 생태계에 대한 수요도 지

\*한국과학기술정보연구원, 과학기술연합대학원대학교 cokeman@kisti.re.kr,

\*\*한국과학기술정보연구원, maximus74@kisti.re.kr

※ 본 연구는 '나노팩 공정·장비 데이터의 K-MDS연계활용·확산(N24NM038-24)' 과제의 지원으로 수행되었음

속적으로 증가할 것으로 예상된다. 특히 반도체 산업은 국가의 핵심 산업으로, 나노 기술 분야에서 국가 공공 인프라로 부터 생산된 공정·장비 데이터의 공유가 점차 중요해지고 있다. 플랫폼에 축적된 데이터를 통해 데이터 기반 연구 및 개발 효율성을 높일 수 있다.

스마트 나노팸은 공정 데이터를 기반으로 자동화와 디지털화를 통합한 개념으로, 생산된 데이터는 플랫폼화를 통해 접근성을 높이고, 데이터를 축적, 정제, 공유할 수 있는 시스템을 구축함으로써 중복 실험이나 인력 교체로 인한 경험 단절 등의 비효율을 해소할 수 있다[1][2]. 또한, 소재 연구 데이터 플랫폼과의 연계를 통해 연구 시너지 효과를 극대화할 수 있다[3]. 이에 따라 본 논문에서는 소재, 반도체, 장비 개발 연구자들이 손쉽게 데이터를 검색, 공유, 활용할 수 있도록 플랫폼이 갖추어야 할 기능들을 정의해 보고자 한다..

## II. 나노팸 공정·장비 데이터 플랫폼의 주요 기능

먼저 스마트 나노팸에서 생산되는 공정 및 장비 데이터의 자동 수집과 연관된 공정 및 장비 데이터를 세트로 관리할 수 있도록 메타데이터를 등록하기 위한 기능은 필수적이다. 이를 통해 데이터는 체계적으로 관리되며, 사용자들은 필요한 데이터를 검색하고 상세 정보에 접근할 수 있다.

반도체 제조 장비는 공정별로도 다양하지만, 같은 공정이라도 제조사별, 모델별로 천차만별이다. 장비에서 생산된 데이터 역시 용어, 의미, 단위 등이 다른 경우가 많기 때문에, 이러한 데이터 스키마의 표준화와 변환은 필수적이다. 플랫폼에서는 표준 스키마에 맞도록 자동 수집된 데이터를 변환하는 기능을 반드시 제공하여야 한다.

다음으로는 저장된 데이터 및 메타데이터를 검색하는 기능이다. 대부분의 반도체 연구자들은 제품 생산에서 사용된 일련의 공정 및 조건 (Recipe)에 따른 데이터를 필터링하기를 원한다. 따라서 메타데이터에는 해당 정보가 반드시 포함되어야 하며, 이

를 기반으로 다양한 패킷으로 검색할 수 있어야 한다. 또한 장비 설정 조건, 모니터링 값, 공정 후 측정값 등 데이터 내에 특정 조건별로 검색하여 동일 조건에서 어떻게 장비가 운용되고 제품의 특성이 변화되는지를 검색할 필요도 있다. 이러한 검색은 메타데이터 수준의 데이터 관리 및 검색에서는 할 수 없기 때문에, 다차원 데이터베이스와 같은 특별한 데이터 관리시스템과 데이터 모델링이 필요하다.

검색을 통하여 원하는 데이터들을 식별하였으면, 데이터 연구를 수행할 공간으로 가져와야 한다. 크게 두 가지 공간을 생각해 볼 수 있는데 하나는 사용자의 로컬 영역 (PC)으로 데이터를 다운로드하는 것이고, 다른 하나는 플랫폼 상에 개인 데이터 연구 공간을 할당하여 그 공간에 데이터를 복사하는 것이다. 이것은 플랫폼 운영 H/W 상황, 인적 역량 등과 사용자의 연구 크기나 선호도 등에 따라 선택하여야 한다.

이후는 플랫폼에서 연구 공간을 제공하는 경우에 필요한 기능들을 서술한다.

플랫폼 상에 개인 연구 공간으로 데이터를 복사한 다음에는, 데이터의 형태, 특징 등을 살펴보는 탐색적 분석 (EDA) 과정이 필요하다. 데이터 기반 연구는 목적에 따라 매우 다양한 형태로 진행될 수 있다. 하지만 본격적인 분석 전에 수행되는 탐색적 분석 과정에서 데이터의 크기, 형태, 분포 등을 살펴보는 과정은 대부분 유사하기 때문에, 주로 사용하는 탐색적 분석 기능의 구현, 가시화 도구 등을 통해 탐색 시간과 작업을 줄여줄 필요가 있다.

탐색적 분석 이후 연구는 목적에 따라 매우 다양하기 때문에 공통된 기능을 제공하기가 매우 어렵다. 하지만 많이 사용하는 분석 도구, 예측을 위한 공통 AI 모델 등은 구매 혹은 구현하여 제공하거나, 사용자들이 직접 공유할 수 있도록 마켓플레이스를 제공할 필요가 있다. 또한 많이 사용하는 분석 환경, 예를 들면 JupyterLab을 통한 python 개발환경이나 R 개발환경 등을 제공하는 것도 플랫폼 및 데이터 활용도를 높이기 위한 방안이다.

공정·장비 데이터 플랫폼의 목적은 데이터 및 도구를 공유하여 연구 혁신을 이루고자 함이다. 협력 연구도 연구 혁신을 가속화하기 위한 좋은 전략이다.



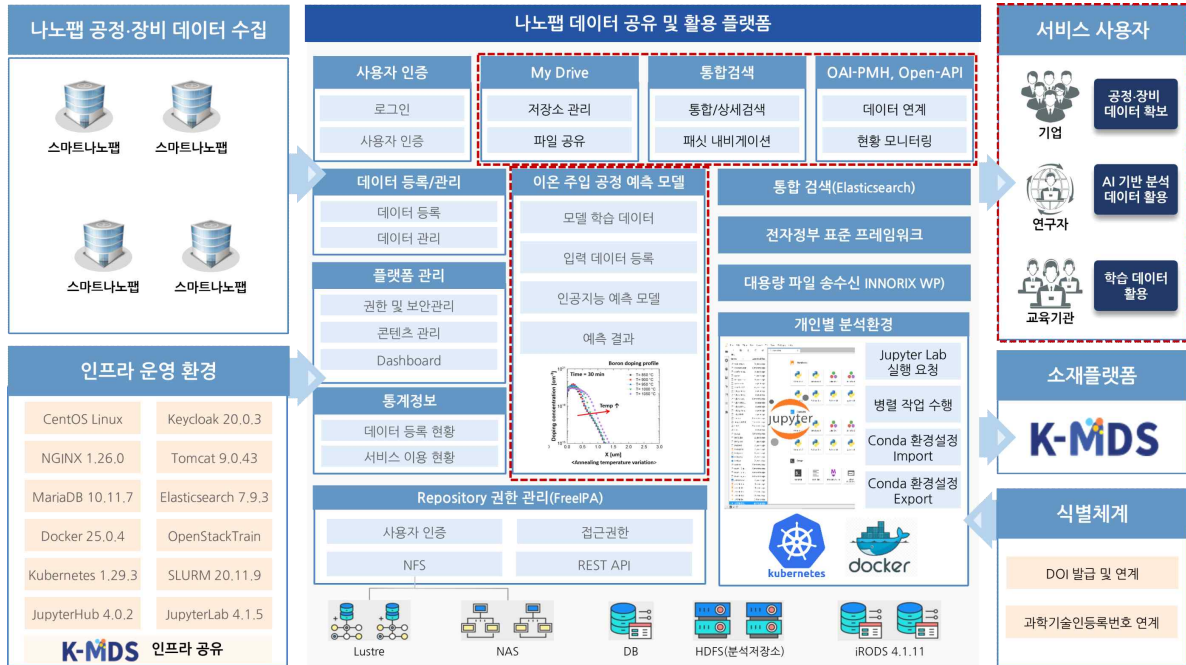


그림 1. 나노랩 공유 및 활용 플랫폼 기능 구성도

따라서 플랫폼에서는 연구 커뮤니티를 쉽게 구성하고, 개인 연구 공간을 할당하는 것처럼 연구 커뮤니티도 협력 연구를 수행할 수 있는 공간을 제공하여야 한다. 플랫폼 상의 연구 공간 제공을 통한 연구 이력이 축적되며, 자주 사용되는 데이터나 도구 등의 추천 서비스나 협력 가능한 연구자들의 매칭 서비스 등 다양한 파생 서비스를 제공할 수 있게 된다.

### III. 결 론

논문에서는 디지털 전환 가속화 속에서 나노 기술과 반도체 분야의 데이터 공유 및 관리 플랫폼의 필요성을 강조하고 필요한 기능 정의를 제안하였다. 스마트 나노랩의 공정 및 장비 데이터를 기반으로 플랫폼은 데이터 접근성을 높이고, 연구 효율성을 향상시키고, 협력 연구 기반을 마련하는 것을 목표로 한다.

### 참 고 문 헌

[1] 이정엽, "[확대경] 나노인프라의 이용자 지원 현황." 진공이야기 6(4), 32-35, 2020

[2] 황미녕, 박상배, 신영호, 신성호, "연구데이터 공유·관리를 위한 국가연구데이터플랫폼(DataON) 구축", 한국콘텐츠학회 종합학술대회 논문집, 2020년 7월.

[3] 소재연구데이터플랫폼, "<https://kmds.re.kr/>", 2024

# 이종 센서 간 데이터 퓨전을 이용한 3차원 디지털 부피 측정

이 윤 주\*

## 3D Digital Volume Measurement Using Data Fusion of Heterogeneous Sensors

Yunju Lee\*

본 논문에서는 이종 센서 간 데이터 퓨전을 이용한 3차원 디지털 부피 측정 기술을 제안한다. 물류 산업의 완전 자동화 및 스마트 물류 시스템 구현을 위해서는 측정 대상에 대한 정확한 3차원 데이터 측정이 필수적이다[1]. 정확한 치수, 부피, 모양 등의 빅데이터는 전자상거래 공급망, 3PL(3자 물류 대행) 및 창고 운영의 효율성을 극대화하는 데 중요한 역할을 한다[2]. 본 논문에서 제안하는 기술은 이종 센서 데이터를 캘리브레이션하여 융합하고, 이를 통해 대상 객체를 정밀하게 3차원으로 복원하는 것을 목표로 한다. 센서 퓨전 기술은 RGB 이미지와 Depth 이미지를 활용하여 객체의 특성을 명확히 파악하며, ToF 및 LiDAR 센서의 포인트 클라우드 정보에 칼라 텍스처를 매핑함으로써 시각적 품질을 향상시킨다[3,4]. 복원된 3D 데이터는 메쉬 작업을 거쳐 ply 파일로 저장될 수 있으며, 이를 다양한 플랫폼에서 불러들여 시각화할 수 있다. 본 연구는 이러한 기술적 발전을 통해 물류 자동화 시스템의 경쟁력을 강화하고, 소비자의 다양한 요구에 부응하는 효율적인 솔루션을 제공할 것으로 기대된다. 특히, 이 기술은 임베디드 엣지 디바이스 프로세싱 기반 3D 스캐너를 활용하여 물류 운영 비용을 효율화하고, 수요처의 물류 시스템에 적용 가능한 시스템 개발을 통해 고객 만족과 매출 증대 효과를 달성할 것으로 기대된다.

Key words

digital volume measurement, 3d, data fusion, heterogeneous sensors

### 참고문헌

- [1] M. Chen, Y. Liu, and Z. Zhao, "Deep Learning for Point Cloud Processing: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 5, pp. 2762-2780, 2022.
- [2] T. Smith and R. Johnson, "Optimizing 3D Scanning for Logistics Applications," *International Journal of Logistics Management*, vol. 33, no. 3, pp. 617-635, 2022.
- [3] L. Wang and M. Li, "Real-time 3D Object Reconstruction Using RGB-D Cameras," *IEEE Transactions on Image Processing*, vol. 29, pp. 5573-5587, 2020.
- [4] H. Zhang, "Calibration and Fusion of Sensor Data for 3D Reconstruction," *Sensors*, vol. 21, no. 3, p. 876, 2021.

---

\*한국전자기술연구원, topgun@keti.re.kr.

※ 본 논문은 2023년도 중소벤처기업부의 재원으로 중소기업기술정보진흥원의 지원을 받아 수행된 연구임 (RS-2023-00270156).

# 생체 데이터 기반 음파진동기 전용 모바일 어플리케이션 개발

이윤주\*

## Development of a Biometric Data-Based Sonic Vibration Device Dedicated Mobile Application

Yunju Lee\*

스피커를 이용해 음파를 진동으로 변환하는 음파진동 기술은 연속된 주파수에서 다양한 진동 효과를 낼 수 있기 때문에 마사지, 안마 의자 등 건강용품에서 기술 요구가 증대되고 있다[1]. 그러나 음파 진동에 대한 현재 연구는 개인 스트레스 완화[2]와 신체 특성에 따른 맞춤형 음파진동을 고려하지 않고 기계에서 생성된 음파 진동 요법의 개발에만 초점을 맞추어 음파진동을 제어하는 어플리케이션에 관한 연구가 부족하다[3]. 본 논문에서는 사용자가 건강 관리를 하면서 몰입감 높은 엔터테인먼트 경험을 즐길 수 있도록 신체의 각 부분에 개인 맞춤형 음파 진동을 생성하는 음파진동기의 전용 모바일 어플리케이션 개발을 제안한다. 제안된 음파진동기 전용 모바일 어플리케이션은 영화, 음악, 게임 등 다양한 콘텐츠에 맞춰 음파진동기의 진동 패턴과 강도를 조절하여, 사용자에게 현실감 있는 감각 자극을 제공한다. 특히 스마트 워치와의 실시간 연동을 통해 사용자 심박수를 모니터링하고, 이를 기반으로 안정 상태를 유지하는 제어 기능이 탑재되어 있다. 모바일 어플리케이션은 블루투스 기반 통신을 통해 스마트 워치와 진동기를 연결하여, 실시간 데이터를 빠르고 정확하게 전송하여 개인의 취향과 안정성을 고려한 최적의 진동 강도와 패턴을 제공할 수 있다. 생체 데이터와 스마트 기기 간의 연동을 활용하여 음파진동을 제어함으로써 건강 관리와 엔터테인먼트 기능을 동시에 제공할 수 있다.

### Key words

sonic vibration device, mobile application, biometric data, health management

### 참고문헌

- [1] J. H. Lee, J. Y. Beag, S. J. Chang, and G. M. Pil, An overview on vibration or wave therapy in Korea, The Korean Academy of Medical Gi-Gong, Vol. 20, No. 1, pp. 15-67, 2020.
- [2] H. Mori, H. Ohsawa, T. H. Tanaka, E. Taniwaki, G. Leisman, and K. Nishijo, Effect of massage on blood flow and muscle fatigue following isometric lumbar exercise, Med. Sci. Monit, Vol. 10, No. 5, pp. 173-178, 2004
- [3] D. S. Choi, Effects of whole-body vibration exercise on soleus neuro-muscular activity in healthy adults, The Korean Society of Living Environmental System (ksles), vol. 26, No. 5, pp. 735-741, 2019.

---

\*한국전자기술연구원, topgun@keti.re.kr.

※ 본 논문은 2022년도 중소벤처기업부의 재원으로 중소기업기술정보진흥원의 지원을 받아 수행된 연구임.  
(No. S3310926)



# 미사일 이미지 판정을 위한 강화 학습에 대하여

김 세 라

## On Deep Learning for Missile Image Classification

Kim, Sera

본 연구에서는 디지털 이미지에서 중요한 경계 정보를 추출하기 위해 Canny 에지 검출 기법을 적용하는 방법[1]을 재해석하고 이를 확장하고자 한다. Canny function을 이용한 경계선 검출은 영상 처리에서 널리 사용되는 방법으로, 이미지의 가장자리나 경계를 강조해 주요 구조적 특징을 추출할 수 있다는 점에서 매우 유용하다. 이 기법을 통해 원본 이미지의 세부적인 경계를 파악하고, 이후에 진행될 학습 과정에서 필요한 정보만을 효과적으로 필터링하여 후에 적용할 Convolutional Neural Network 기반 딥러닝 모델이 최적의 성능을 발휘할 수 있도록 전처리 과정을 수행한다.

이렇게 추출된 에지 이미지를 특정 데이터 셋을 사용하여 사전 학습된 Convolutional Neural Network 모델에 적용하였다. 사용된 데이터셋은 손으로 쓴 숫자 이미지가 포함된 대규모 데이터셋인 MNIST data set으로, 딥러닝과 머신러닝 초기 연구에서부터 기본적이고도 중요한 실험 자료로 활용되고 있다. 이를 Adam optimizer 라는 기울기 하강법을 개선한 최적화 알고리즘을 기반으로 학습시켰다. 이 결과 학습 속도와 수렴 속도 면에서 우수한 성능을 제공하며 가중치 조정을 통해 해당 모델의 이미지 일반화 능력을 높였다. 저자는 이전 단계에서 얻은 에지 이미지를 이 모델에 적용하여 미사일 이미지에서 얻을 수 있는 결과가 무엇인지를 설명하고자 한다. 이러한 접근 방식을 통해 일반적인 미사일 피격으로 인해 발생하는 이미지와 기타 이미지 간의 시각적 차이를 구별할 수 있는 가능성을 탐구하고자 한다. 미사일 피격 이미지와 시각적으로 유사한 요소를 포함할 수 있는 이미지에서 해당 딥러닝 모델을 활용하여 두 유형의 이미지가 갖는 미세한 특징적 차이를 학습하고 이를 통해 거의 정확한 분류를 시도한다.

본 연구의 결과는 다양한 이미지 분류 문제에 있어 중요한 기여를 할 수 있으며, 특히 비슷한 시각적 특징을 가진 이미지들 간의 차이를 인공지능이 학습을 통해 구분할 수 있는 새로운 가능성을 시사한다. 본 연구는 단순히 미사일 이미지 분류 문제에 국한되지 않고, 나아가 다양한 환경에서 이미지 분류의 정확도를 높이기 위한 전처리 방법과 학습 알고리즘의 효과적인 조합을 제안함으로써 딥러닝 모델의 활용 범위를 확장하는 데에 의미 있는 기여를 제공하고자 한다.

### Key words

Missile strike image, Canny function, Deep learning, MNIST data set, Adam optimizer, Convolutional Neural Network

### 참고문헌

[1] 김세라. 파이썬과 MOY graph를 이용한 항적 분석방법. Journal of KNST (2023) Vol. 6, No. 3 pp 235-239.

---

\*해군사관학교, srkim85@gmail.com (교신저자)

※ 이 논문은 2023년 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임(KRIT-CT-23-030)

# 실시간 드론 촬영 기반 콘크리트 균열 탐지 모델 성능 비교

김영규\*, 최대영\*\*, 김대년\*\*

## Performance Comparison of Real-Time Concrete Crack Detection Models Based on Drone Footage

Young-Kyu Kim\*, Daeyoung Choi\*\*, and Dae-Nyeon Kim\*\*\*

콘크리트 구조물의 유지 관리에서 균열 탐지는 매우 중요한 작업으로, 균열의 조기 발견과 처리를 통해 구조물의 안전성을 보장할 수 있다. 본 논문에서는 드론을 이용한 실시간 콘크리트 균열 탐지의 효율성을 높이기 위해 다양한 인공지능 기반 모델들의 성능을 비교 분석하였다. PIDNet을 기반으로 한 균열 탐지 모델과 기존의 U-Net, Attention U-Net, U-Net++, Mask R-CNN 모델을 비교하여 각 모델의 균열 탐지 성능을 평가하였다. 성능 지표로는 mIoU (Mean Intersection over Union), 정확도 (Accuracy), 정밀도 (Precision), 재현율 (Recall), 모델 크기 (Model Size), 처리 속도 (Processing Speed)를 사용하였다. 실험 결과, PIDNet은 상대적으로 가벼운 모델 구조로 인해 높은 처리 속도를 유지하면서도 U-Net, Attention U-Net, U-Net++에 비해 우수한 mIoU 및 정확도를 기록하였다. Mask R-CNN은 미세한 균열을 정확하게 감지하는 데 강점을 보이며 mIoU 및 정확도에서 가장 높은 성능을 나타냈지만, 모델 크기와 처리 속도 측면에서는 다른 모델들에 비해 불리한 결과를 보였다. PIDNet은 경량성과 실시간 처리 측면에서 우위를 점하며, 균열 탐지의 정확도와 성능 간의 균형을 잘 유지하는 모델임을 확인하였다. 본 논문의 연구는 실시간 드론 촬영을 통한 콘크리트 균열 탐지에서 각 모델의 장단점을 명확히 하고, 적합한 모델 선택을 위한 기초 데이터를 제공하는 데 그 의의가 있다. 이러한 비교 분석을 통해 PIDNet과 Mask R-CNN의 특성을 활용한 하이브리드 접근법이나 특정 환경에 최적화된 모델 설계 가능성을 제안한다.

### Key words

crack, concrete crack, crack detection, crack detection model,

### 참고문헌

- [1] S. K. Nayar and Y. Nakagawa, "Shape from focus", IEEE Trans. Pattern Anal. Machine Intell., vol. 16, pp. 824-831, August 1994.
- [1] Oliveira, H., & Correia, P. L., "Automatic Road Crack Detection and Characterization", IEEE Transactions on Intelligent Transportation Systems, Vol. 15, No. 4, pp. 155-165, October 2014
- [2] Cha, Y. J., Choi, W., & Büyüköztürk, O., "Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks", Computer-Aided Civil and Infrastructure Engineering, Vol. 32, No. 5, pp. 361-378, May 2017.
- [3] Li, S., Zhao, X., Xu, Y., & Du, S., "Image-Based Concrete Crack Detection Using Convolutional Neural Network and Exhaustive Search Method", Advances in Civil Engineering, pp. 1-12, June 2019.

\*(재)경북IT융합산업기술원 ykkim@gjtc.or.kr, \*\*(주)파미티 dychoi@farmity.co.kr, \*(재)경북IT융합산업기술원 dnkim@gjtc.or.kr

※ 본 논문은 산업통상자원부의 '지역혁신클러스터육성(R&D) 거점기관 개방형혁신' 사업의 연구비지원(P0025720)으로 수행된 연구결과입니다.

# 유·무인 복합 전투체계 구현을 위한 네트워크 발전 방향에 관한 연구

황재룡\*

## A Study on the Direction of Network Development for the Realization of Manned and Unmanned combined Combat System

Jaeryong Hwang\*

### 요 약

무인시스템과 인공지능이 결합된 유·무인 복합 전투체계는 미래 군의 전장의 패러다임을 바꿀 중요한 개념이다. 그러나, 드론, 로봇 등 무인장비와 군인이 복합되어 있는 전장에서 작전을 성공적으로 완수하기 위해서는 초연결 기반의 네트워크로 발전이 요구된다. 이를 위하여, 군은 다양한 플랫폼으로부터 생산되는 데이터를 신속하고 정확하게 유통하는 센서네트워크의 구축, 저궤도 위성을 이용한 다계층 통합 네트워크 구축, 지능형 네트워크 관리체계가 필요하다.

### Abstract

Combined manned and unmanned combat systems, which combine unmanned systems and artificial intelligence, are an important concept that will change the paradigm of the future military battlefield. However, in order to successfully complete operations on the battlefield where unmanned equipment such as drones and robots are combined with soldiers, it is necessary to develop a hyper-connected network. To achieve this, the military needs to build a sensor network that quickly and accurately distributes data from various platforms, build a multi-layer integrated network using low-orbit satellites, and develop an intelligent network management system.

### Key words

Combined manned and unmanned combat systems, hyper-connected network, intelligent network management

## 1. 서 론

21세기 들어 군사 기술의 발전과 함께 무인 시스템과 인공지능(AI)이 결합된 유무인 복합 전투체계(Unmanned and Manned Combat System)가 주요한

연구 및 개발 주제로 떠오르고 있다[1]. 유무인 복합 전투체계는 병력의 안전을 강화하고 작전 효율성을 높이는 데 중요한 역할을 한다. 이러한 시스템이 효과적으로 운영되기 위해서는 유기적인 네트워크 통신이 필수적이라고 할 수 있다. 본 논문에서는

\*해군사관학교, jhwang@navy.ac.kr

유·무인 복합 전투체계의 구현을 위한 네트워크의 발전 방향성을 살펴본다.

## II. 유·무인 복합 전투체계

유·무인 복합 전투체계는 유인 장비와 무인 장비를 통합하여 상호 보완적인 전투 임무를 수행할 수 있는 전투 시스템이다. 이 체계는 무인 장비의 고위험 지역 투입과 유인 장비의 정밀 조작을 결합해 작전 효율성을 극대화하고 병력의 안전을 확보하는데 목적이 있다.

이와 같은 목적 달성을 위하여 유·무인 복합 전투체계는 유인 전투 장비(탱크, 전투기 등), 무인 전투 장비(드론, 로봇 등), 이들을 제어하는 통제 시스템, 그리고 이를 지원하는 통신 네트워크로 구성된다. 각각의 장비는 실시간 정보를 주고받으며 협력적인 임무 수행을 위해 AI와 센서 시스템을 통합하여 작동한다.

유·무인 복합 전투체계의 네트워크는 저지연, 고신뢰성, 높은 보안성을 바탕으로 다양한 장비의 동시 연결을 지원이 요구된다. 또한, 실시간 데이터 전송을 위한 높은 대역폭과 다수 장비 간의 호환성을 제공하며, AI 기반 자율 네트워크 관리로 장애 발생 시 신속하게 대응할 수 있어야 한다.

## III. 네트워크 발전 방향

유·무인 복합 전투체계 구현하고 효과적인 작전 수행을 위해서 군은 초연결 기반의 네트워크로 발전이 필요하다.

이를 위하여, 첩재, 전투원, 드론, 로봇, 센서 등 다양한 유·무인 전투플랫폼을 위한 센서 네트워크가 만들어져야 한다. 센서 네트워크는 드론봇, M-IoT, 유·무인 전투체계 등 다양한 전투 시스템의 복합 운용을 지원하는 기반 네트워크로서, 여러 종류의 센서와 유·무인 전투체계로부터 전송되는 데이터를 통합 관리할 수 있는 능력을 갖추어야 한다. 이와 더불어 획득한 정보의 공유, 분배, 저장 및 C4I 체계와 연동하기 위한 연구개발이 필요하다.

둘째, 공중, 해상 그리고 지상에 이르기까지 전장

의 모든 영역에서 전투에 참여하는 플랫폼들이 항상 네트워크에 연결 가능하도록 다계층 통합 네트워크가 구성될 수 있어야 한다. 이를 위하여 군은 저궤도 소형 통신 위성군 운영 개념을 정립하고, 군 위성통신체계와 상용 위성의 상호 운용성을 위한 연동 기술에 대한 발전이 요구된다.

마지막으로 자율적으로 네트워크 상태를 점검하고, 문제가 발생했을 때 빠르게 대응할 수 있는 지능형 네트워크 관리체계가 필요하다. 지능형 네트워크 관리체계는 AI 기반 최적화된 전장의 정보를 유통하고 실시간성을 보장할 수 있는 네트워크 서비스 제공이 필요하다. 따라서 SDN(Software Defined Network)과 같이 네트워크 제어 영역과 데이터 전달 영역을 분리하여 네트워크를 소프트웨어 기반으로 제어할 수 있는 아키텍처를 도입하고, NFV(Network Function Virtualization)과 같이 네트워크 기능을 소프트웨어로 가상화하여 유연하게 시스템을 설치하고 업그레이드 할 수 있어야 한다. 또한, 전장에서 발생하는 적의 전파방해 및 연합·합동 무기체계 간 상호 간섭에 대해 자동으로 대응이 가능한 자동 주파수 관리체계의 구현이 요구된다.

## III. 결 론

본 논문에서는 유·무인 복합 전투체계의 구현을 위해 요구되는 네트워크 발전 방향에 대하여 소개하였다. 다양한 플랫폼으로부터 생산되는 데이터를 신속하고 정확하게 유통하는 센서네트워크의 구축, 저궤도 위성을 이용한 다계층 통합 네트워크 구축, 지능형 네트워크 관리체계가 필요하다.

## 참 고 문 헌

[1] J. -K. Choi, Y. -T. Lee, H. Park, B. Kim and B. -W. Kim, "Challenges to the Development of Manned and Unmanned Combat Systems," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 2362-2364, doi: 10.1109/ICTC55196.2022.9952483.



# 차량용 열영상 카메라 ISP 구현 기술

이선영\*\*, 유재웅\*, 조용현\*

## Automotive Thermal Imaging Camera ISP Implementation

Seonyoung Lee\*\*, Jaewoong Yoo\*, and Yonghyeon Cho\*

차량용 열영상 카메라는 야간, 악천후 상황에서 주간 카메라와 함께 자율주행시스템 및 ADAS(Advanced Driver Assistance System, 첨단운전자보조시스템)에 사용이 요구되는 핵심 센서이다. NHTSA(National Highway Traffic Safety Administration, 미국 도로교통안전국)에서는 사각지대 감지/개입 및 보행자 자동비상 제동 기능과 관련한 ADAS기술에 열영상 카메라를 사용할 수 있는 감지 센서로 제시하고 있다.[1] 또한 최근 발표된 Euro-NCAP 비전 2030에서는 2026년부터 안전 운전, 충돌 회피, 충돌 보호, 충돌 후 안전에 대한 항목을 평가에 추가한다고 발표하였고 ADAS 및 자율주행 기술을 탑재한 차량이 일상적인 주행환경 뿐 아니라 비나 눈, 안개 등의 악천후에도 이륜차를 감지하고 충돌을 방지하기 위한 기능이 제대로 작동하는지를 평가 항목에 추가할 것이라고 발표하였다. 따라서, 주요 선진국들이 레벨 2이상의 자율주행차를 위한 교통 안전 정책을 도입 및 강화함에 따라 저조도 및 다양한 악천후 환경에서도 안정적으로 운전자를 지원할 수 있도록 연구개발과 제품의 차량 탑재가 필요하다. 현재는 자율주행자동차의 레벨 2, 3이상의 기능을 구현하기 위해 일부 고급차량에 열영상 카메라가 적용되어 있으며, 이와 관련한 열영상 IP 탑재는 필수적으로 요구되고 있다.

기존의 열영상 카메라는 영상처리를 위해 범용 SoC 및 전용 FPGA 보드를 탑재하는 방식으로 구현하였으며 이로 인한 소모전력 효율이 낮아지고 처리속도, 제품크기와 가격이 높아지는 문제가 있다. 이를 극복하기 위해 한 개의 칩에서 열영상 센서를 인터페이스하여 안정적으로 열영상을 출력할 수 있는 ISP(Image Signal Processor) 탑재가 필요하다. 본 논문에서는 차량용 열영상 카메라에 사용되는 비냉각 방식의 열영상 센서에서 발생하는 고정 패턴 노이즈를 제거하기 위해 Two-Point Correction(TPC) 알고리즘과 열영상 센서의 특징 및 온도 변화에 대응하기 위해 Thermal Drift Compensation(TDC) 알고리즘을 사용하였다.[2] 그리고 이를 가속 처리하기 위해 HDL(Hardware Description Language)을 이용하여 하드웨어 회로를 구현하여 처리 속도를 높였다. 이를 통해 XGA급 열영상 센서로부터 전달되는 데이터의 불균일성을 실시간으로 처리할 수 있는 하드웨어를 구현하였다.

### Key words

thermal camera, non-uniformity correction, ADAS, autonomous driving

### 참고문헌

- [1] Federal Motor Vehicle Safety Standard (FMVSS) No. 127, Automatic Emergency Braking(AEB) Systems for Light Vehicles, NHTSA, 2024.
- [2] S. Rong, H. Zhou, Z. Wen, H. Qin, K. Qian, and K. Cheng, "An improved non-uniformity correction algorithm and its hardware implementation on FPGA," *Infrared Physics & Technology*, vol. 85, pp. 410-420, Sep. 2017.

\*한국전자기술연구원, drleesy@keti.re.kr, \*\*한국전자기술연구원, drleesy@keti.re.kr(교신저자표시)

※ 이 연구는 2024년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임 (RS-2024-00405892, XGA급 열영상 나이트 비전 카메라용 ISP(Image Signal Processor) 반도체 개발).

# 기상 데이터를 이용한 딥러닝 기반 일사량 예측 모델 연구

김민석\*, 박형욱\*, 이은지\*, 김은주\*

## Deep learning based solar irradiance prediction model using weather data

Minseok Kim\*, Hyungwook Park\*, Eunji Lee\*, and Eunju Kim\*

대표적인 친환경 에너지원인 태양광 발전은 전 세계적으로 신규 설치 용량이 꾸준히 증가하며 주목받고 있다. 태양광 발전에 대한 안정적인 전력 공급과 수익성 확보를 위해서는 태양광 발전량에 대한 예측이 필요하지만 태양광 발전은 기상 조건에 따른 출력 변동성이 커서 태양광 발전의 핵심 요소인 일사량에 대한 예측이 선행되어야 한다[1]. 따라서 본 연구에서는 기상 데이터를 이용한 딥러닝 기반의 일사량 예측 모델을 제안한다. 본 논문에서 일사량 예측을 위해 유럽중기예보센터(ECMWF)에서 제공하는 고해상도 기상 데이터인 ERA5 재분석 데이터[2]를 이용한다. 특히, 일사량과 직접적으로 연관이 있는 '대기 상단 입사 태양 복사', '총 구름량', '비습' 등의 변수들을 입력 데이터로 활용한다. 그리고 실제 운영 중인 태양광 발전 시설에서 측정·수집된 수직 일사량 및 수평 일사량을 단위 시간별로 예측한다. 일사량 예측 모델로는 LSTM(Long Short-term Memory)[3]을 이용한다. LSTM을 이용하여 시간에 따른 기상 변화에 대한 정보를 획득하고, 이를 통해 일사량을 예측하는 것이다. 제안한 LSTM 모델에 대한 성능 평가는 RMSE(Root Mean Square Error)를 통해 검증한다. 제안한 딥러닝 기반 일사량 예측 모델을 통해 정확도 높은 일사량 예측을 수행함으로써 태양광 발전량 예측에 대한 신뢰성도 함께 높일 수 있을 것으로 기대된다.

### Key words

deep learning, LSTM, photovoltaic, solar irradiance, weather data

### 참고문헌

- [1] H. J. Jeon, M. W. Choi and O. J. Lee, "Day-Ahead Hourly Solar Irradiance Forecasting Based on Multi-Attributed Spatio-Temporal Graph Convolutional Network", *Sensors*, vol. 22, No. 19, September 2022.
- [2] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, et al. "The ERA5 global reanalysis", *Quarterly Journal of the Royal Meteorological Society*, vol. 146, No. 730, pp. 1999-2049, June, 2020.
- [3] S. Hochreiter, and J. Schmidhuber, "Long short-term memory", *Neural computation*, vol. 9, No. 8, pp. 1735-1780, 1997.

---

\*한국과학기술정보연구원, kimminseok@kisti.re.kr, hyungwook@kisti.re.kr(교신저자), eunji\_lee@kisti.re.kr, ejkim@kisti.re.kr

※ 이 논문은 중소벤처기업부의 기술개발사업 지원에 의한 연구임[S3400727]

# Wafer Edge Chip 분석 및 분류 소프트웨어 개발

박송희\*, 박종익\*\*, 김영형\*\*\*

## Development of Wafer Edge Chip Analysis and Classification Software

Song-Hee Park\*, Jong-Ik Park\*\*, and Young-Hyoung Kim\*\*\*

### 요 약

본 논문에서는 반도체 제조 공정 중 웨이퍼 엣지의 불량을 정확하게 검출하고 분류하기 위한 소프트웨어 개발 방안을 제시한다. 현재 웨이퍼 엣지 검사 공정은 검사자의 육안에 의존해 수작업으로 진행되고 있으며, 이는 불량 누락 및 품질 편차 문제를 야기한다. 이를 개선하기 위해 비전 검사 장비와 딥러닝 기반 패턴 분석 기법을 적용하여 다양한 불량 유형의 판별과 데이터 기반 품질 관리를 가능하게 한다. 제안된 소프트웨어는 불량 유형별 데이터를 수집하고 분석하여 높은 예측 정확도와 공정의 신뢰성을 향상시킬 수 있다.

### Abstract

This paper presents a software development approach for accurately detecting and classifying wafer edge defects in semiconductor manufacturing. Currently, wafer edge inspection relies on manual visual checks, leading to defect omissions and quality inconsistencies due to subjective judgment. To address these limitations, this study incorporates vision inspection equipment and deep learning-based pattern analysis techniques, enabling reliable identification of various defect types and data-driven quality management. The proposed software collects and analyzes defect data by type, enhancing prediction accuracy and overall process reliability.

### Key words

wafer, wafer edge, vision system

## 1. 서 론

반도체 및 배터리 업계에서는 수율이 높아야 시장에서의 경쟁력을 확보할 수 있으므로 수율을 높이고 노력하고 있다[1]. 웨이퍼 제조 단계 중 웨

이퍼의 모서리를 깎아 둥글게 만드는 공정인 엣지 그라인딩 공정[2]에서 다양한 형태의 엣지 파손이 발생하고 있다[3]. 현재 Wafer Edge 검사 공정은 검사자의 육안에 의존하는 수작업 방식으로 수행되고 있다. 그러나 검사자의 주관적 판단에 의해 불량이

\* 국립금오공과대학교 전자공학과 psh0203@kumoh.ac.kr

\*\* (주) 세정로봇 대표, unihost@chol.com

\*\*\* 금오공과대학교 IT융합학과, kic126@kumoh.ac.kr

※ 본 연구는 2023년도 중소벤처기업부의 기술개발사업 지원에 의한 연구임 [RS-2023-00269460]

누락되는 경우가 발생하며, 개인의 숙련도에 따라 품질에 편차가 생기는 문제도 발생한다. 비전 장비가 상용화 되어 있지만 다양한 불량 유형을 판별하기는 어렵고 불량 유형별로 분류하는 것이 불가능하다. 이에 본 논문에서는 Wafer Edge 불량의 데이터 분석과 분류가 가능한 소프트웨어 개발 방법을 제시한다.

## II. 본 론

Wafer Edge의 Chip 패턴 데이터 분석을 위해 비전검사 공정과 품질, 데이터 관리 및 활용 프로세스를 설계한다. 그림 1은 패턴 데이터 분석 프로세스를 나타낸다.

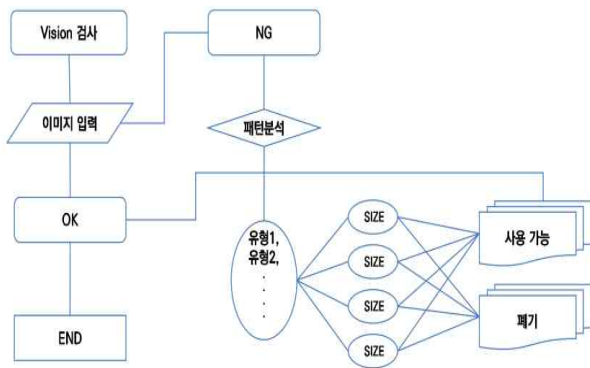


그림 1. 패턴데이터 분석 프로세스  
Fig. 1. Pattern data analysis process

비전 검사를 통해 Wafer의 이미지를 받아들여 불량량의 유무를 판단한다. 이 때 비전은 Wafer Edge의 상, 하, 측면을 다 검사 할 수 있도록 카메라 3개를 한 세트르 구성하여 설치한다. Edge Chip이 있다고 판단되면 패턴 분석 프로세스로 진행하고 없다고 판단되면 양품으로 판정하여 다음 공정으로 진행한다. 패턴 분석 프로세스에서는 Chip의 유형과 크기를 판단하여 사용가능여부를 판단하고 기준을 벗어나는 Chip이라고 판단하면 Wafer는 폐기 처리한다. 이 때 Chip의 형태와 크기, 발생 위치 등 불량 유형에 따른 각각의 데이터와 불량 이미지를 수집하고 관리 및 활용하는 프로그램을 개발 및 연계 구축하여 딥 러닝을 적용시킨다. 이를 통해 예측 정확도를 높이고 수집된 데이터를 이용한 유의미한 정보

를 도출 할 수 있다. 그림 2는 딥러닝 모델 학습과정을 나타낸다.

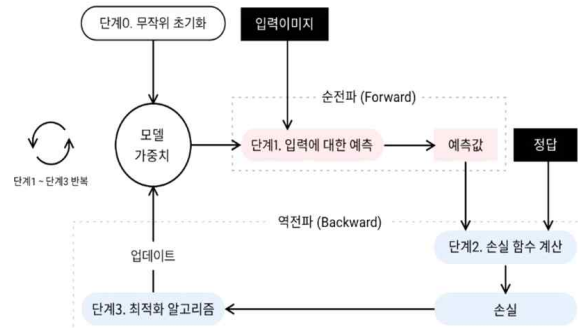


그림 2. 딥러닝 프로세스  
Fig. 2. Deep Learning Process

## III. 결 론

본 논문에서는 Wafer Edge Chip을 비전 검사를 통해 분석 및 분류하는 방법을 제시하였다. 제시한 방법을 적용하면 동으로 초점 형상 복원을 위한 인공신경망을 제안하였다. 제안된 방법은 인공 신경망의 학습 시에 깊이 맵을 FIS로 정확하게 근사시키면서 학습이 되기에 기존의 방법보다 개선된 복원 결과를 보여준다.

## 참 고 문 헌

- [1] Baek Dong-Hyeon and Nam Jeong-Gon, "Semiconductor Yield Improvement System Using Data Mining", Korean Institute of Industrial Engineers/Korean Management Science Society vol. 14, pp. 293-300, 5, 1994.
- [2] "Edge Grinding", AxusTech, 2023 인용 <https://axustech.com/processdevelopment-foundry/edge-grinding/>.
- [3] Young-Hyung Kim, Jong-Ik Park, Young-Kwang Hwang, & So-Young Kwon (2023-11-23). Vision Inspection System to Improve Wafer Yield. Proceedings of KIIT Conference, 제주

# 고품질 배경 생성을 위한 3D 퍼포먼스 기술 연구

장민혁\*, 김현주\*

## Research on 3D performance techniques for high-quality background generation

Min-Hyuk Chang\*, and Hyeon-Ju Kim\*

### 요약

초기의 3D 배경 제작 방법들은 수차례의 렌더링 과정을 거치면서 효율성이 저하되는 문제를 보였으며, 최근에는 게임엔진을 통해 실시간 렌더링 방식을 도입하였으나 학습이 쉽지 않다는 단점이 있다. 본 논문에서는 3D 퍼포먼스 기술을 연구하고, 이를 바탕으로 4가지 주요 기능을 통해 효율적이고 몰입감을 제공하는 3D 배경을 생성할 수 있는 방법을 제안한다.

### Abstract

The early methods of 3D background creation faced efficiency issues due to multiple rendering processes. Recently, real-time rendering techniques using game engines have been introduced, but they present challenges in terms of ease of learning. This paper studies 3D performance techniques and proposes a method to create efficient and immersive 3D backgrounds by utilizing four key features.

### Key words

3D Contents, Metaverse, Background generation, XR services, VR

## 1. 서론

그래픽 기술과 게임엔진의 발전으로 3D 모델링을 활용하여 웹툰, 애니메이션, 메타버스, 영화, 게임 등의 배경을 제작하는 시도들이 늘어나고 있다. 3D 배경을 사용하면 조명과 텍스처의 변화를 통해 하나의 콘텐츠로 다양한 장르의 배경을 제작할 수

있다는 장점이 있으며, 가상의 배경을 제작하여 표현의 범위를 확장할 수 있다[1]. 초기 3D 배경은 3D 소프트웨어를 활용하여 제작되었다. 원하는 결과물을 얻기 위해 수차례의 렌더링 과정이 요구되었으며, 이 과정은 많은 시간을 소모하여 효율성이 저하되었다[2]. 최근에는 효율성을 높이기 위해 언리얼 엔진(Unreal Engine)이나 유니티(Unity)와 같은

\*한국전자기술연구원, minhyuk@keti.re.kr, \*한국전자기술연구원, kemhj@keti.re.kr

※ 연구개발특구진흥재단(본 연구는 2022년도 정부(과학기술정보통신부)의 재원으로 연구개발특구진흥재단의 '지역의 미래를 여는 과학기술 프로젝트' 사업으로 수행되었습니다.(과제명: 인공지능 기반 메타버스 구현을 위한 융·복합 문화 가상 스튜디오, 과제번호: 2022-DD-UP-0312))

게임엔진이 3D 편집, 라이팅, 렌더링등에 활용되고 있으나 프로그래밍 기능을 포함하고 있어 학습에 어려움이 있다[2].

효율적인 3D 작업 방안을 모색하기 위해 가상공간에서 3D 모델링 작업을 지원하는 실감 미디어 제작 툴에 대한 연구[3]와 제작 시간을 단축하기 위해 실시간 렌더링 엔진 분석 및 가이드를 제공하는 연구[4] 등이 활발히 진행되고 있다. 아울러 3D 모델링을 다양한 업계에서 활용하기 위한 리토폴로지[5] 관련 연구 역시 지속해서 수행되고 있다.

본 연구에서는 이러한 선행 연구들과 함께 3D 배경 작업을 다양한 장르에 활용할 수 있는 고품질의 배경 생성 3D 퍼포먼스 기술을 연구하고자 한다.

## II. 3D 퍼포먼스 기술 분석

3D 퍼포먼스 기술 개발에 앞서 사용자들이 자주 사용하는 3D 배경 기술을 도출하기 위해 유니티의 에셋스토어와 언리얼 마켓플레이스에서 판매중인 Shader 태그와 카테고리를 활용 목적에 따라 분류하고 정의하였다. 에셋스토어에서는 시각효과 관련 Shader 카테고리에서 총 3,752건[6]이 판매 중이었으며 마켓플레이스에서는 Shader 키워드로 검색한 결과 총 381건[7]이 판매 중인 것으로 나타났다. 이를 인기순으로 재정렬하여 각 100건의 에셋의 주요 활용 목적에 따라 분류한 결과 전체 배경의 분위기를 변화시키는 전체 배경 효과, 배경의 일부에 특수한 효과를 적용하는 일부 배경 효과, 파티클이나 불, 번개 등과 같은 이펙트 효과, 그리고 기타로 나눌 수 있었다. 이 중 3D 배경 제작을 위한 이펙트는 배경 생성 기술에서 제외하고 기능을 도출한 결과 3D 퍼포먼스 생성 기술을 총 4가지의 기능으로 정의할 수 있었다.

또한 이전 연구에서 배경 최종 결과물을 미리 확인할 수 있는 프리뷰 기능과 스크린샷 기능의 필요성이 제기된 바 있어 이를 별도로 분석하였다. 분석 결과 제작 중인 배경의 스크린샷 및 배경 프리뷰 기술은 에셋 스토어에서 183개, 마켓플레이스에서 66개가 제공되고 있는 것으로 나타났으며, 사용자들

에게 필요한 주요 기능임을 다시 확인하였다.

표 1. 에셋스토어와 마켓플레이스의 에셋 활용 목적  
Table 1. AssetStore and Marketplace asset

분류	내용	사이트	개수
배경 최적화	아웃라인, 픽셀 등 원하는 장르를 표현하기 위한 배경 최적화	에셋 스토어	24개
		마켓플레이스	36개
날씨 시뮬레이션	비, 눈, 바람 등 날씨 현상을 구현하는 파티클 시스템 및 Shader	에셋 스토어	10개
		마켓플레이스	13개
배경 효과 적용	조명의 색상과 밝기 등을 조절하여 전체 씬 분위기를 형성	에셋 스토어	37개
		마켓플레이스	35개

표 2. 에셋스토어와 마켓플레이스의 장면 프리뷰 기능  
Table 2. AssetStore and Marketplace

분류	내용	사이트	개수
실시간 프리뷰	제작 중인 배경 스크린샷 및 배경 프리뷰 기술	에셋 스토어	183개
		마켓플레이스	66개

## III. 3D 퍼포먼스 기술 개발

본 연구에서 제안하는 3D 퍼포먼스 기술은 PC, 모바일, HMD 등 다양한 플랫폼에서 3D 모델링 에셋을 활용하여 배경을 생성하는 작업에 적용할 수 있으며, 사용자에게 보다 현실감과 몰입감을 주는 환경을 제작하기 위한 기술이다. 특히, 사용자가 배경을 커스터마이징하는 과정에서 효율성을 높여 제작 시간을 단축하고자 한다.

3D 퍼포먼스 기술은 다음의 4가지 핵심 기술로 구성되며, 이 기술들을 포함한 고품질 배경 생성 저작도구를 개발하여 지속적인 테스트 과정을 진행할 예정이다. 이를 통해, 추후 XR 스튜디오에 적용하는 것을 목표로 하고 있다.

3D 퍼포먼스 기술 : a. XR Viewer - 사용자가 실제와 같은 1:1 배율의 환경에서 생성된 배경을 탐색하고, 최종 결과물을 즉각적으로 확인할 수 있는 기술로 자유도가 높은 가상환경을 제공하며 스크린

샷 기능을 통해 결과물에 간편하게 적용, **b. XR Optimization** - 웹툰, 애니메이션, 메타버스 등 다양한 장르에 반영가능한 3D 모델링을 위한 최적화 기술로 적정 폴리곤 수 조정 및 머티리얼 최적화를 통해 성능 향상, **c. XR Weather** - 전 세계 275개국 여러 도시에서 40년 전부터 현재까지의 1시간 단위 날씨 데이터를 API로 받아와, 가상 날씨 시뮬레이션을 적용하는 기술 **d. XR Skybox** - 배경의 전체 분위기를 조성하는 요소인 태양/달, 구름, 하늘의 색조 등을 설정할 수 있는 템플릿 제공 및 세부 값을 조정할 수 있는 배경 렌더링 제어 기술

본 논문에서는 메타버스, 웹툰, 애니메이션 등 다양한 장르에 기술이 적용가능함을 확인하기 위해 배경 3D 에셋 중 건물 실외 에셋 1종, 건물 실내 에셋 1종으로 구성된 실험 데이터셋을 테스트하였다.



그림 1. 실험 데이터셋  
Fig. 1. Experimental data sets

XR Viewer는 사용자가 PC, 모바일, HMD에서 실제와 같은 1:1 배율의 환경과 자유도를 보장하는 가상환경에서 생성된 배경을 탐색하고, 작업 결과를 즉시 확인하고 조정하는 기술이다. 또한 가상 카메라를 통해 스크린샷을 저장하고 이를 즉시 적용할 수 있도록 지원하는 기능을 제공한다. PC와 HMD의 환경에서 테스트한 결과는 다음과 같다.



그림 2. XR Viewer PC 환경 및 HMD 환경 테스트  
Fig. 2. Testing XR Viewer PC environment and HMD environment

테스트 결과 PC 환경과 HMD 환경 모두에서 생성된 배경을 탐색하고 카메라 기능을 통해 최종 렌더링 결과를 즉시 확인하며 스크린샷 촬영이 가능

하였다. 특히 HMD 환경에서는 사용자가 직접 제작된 배경을 건너나 컨트롤러를 사용하여 다양한 방식으로 탐색할 수 있었으며, 배경의 활용도가 매우 높다는 점을 확인할 수 있었다.

XR Optimization는 3D 모델링에서 적정 폴리곤 수 조정 및 머티리얼 최적화를 지원하는 기술로, 3D 배경의 특정 부분을 선택하여 폴리곤 수를 조정할 수 있도록 설계되었다. 이를 통해 사용자는 배경 제작 시 각 요소의 중요도에 따라 효율적으로 작업을 수행할 수 있으며 성능과 품질의 균형을 맞출 수 있다. PC와 HMD의 환경에서 테스트한 결과는 다음과 같다.

그림 3. XR Optimization PC 환경 및 HMD 환경 테스트



Fig. 3. Testing XR Optimization PC environment and HMD environment

테스트 결과 PC 환경에서 건물 실외 에셋의 지붕 폴리곤 수가 132개에서 87개로, 건물 실내 에셋의 바닥 폴리곤 수가 251개에서 143개로 최적화되었다. 또한 HMD 환경에서 건물 실외 배경의 폴리곤 수는 307,378개에서 206,672개로, 건물 실내 배경의 폴리곤 수가 643,067개에서 365,142로 조정되었다. 폴리곤 수가 70~80% 감소하였음에도 불구하고 배경에 시각적 변화 없이 작업을 원활하게 수행할 수 있음을 확인하였다.

XR Weather는 외부 API를 통해 나라, 도시, 날짜, 시간에 날씨 데이터를 수집하여 구름 상태, 태양과 달의 위치, 다양한 날씨 효과의 변화를 시뮬레이션함으로써 사용자가 40년 전부터 현재까지 과거의 특정 시점과 위치의 날씨 상황을 재현할 수 있는 기술이다. 이 기술은 가상환경에서 과거의 기상 조

건을 구현할 수 있도록 지원한다. 건물 실내 에셋은 날씨의 영향을 받지 않으므로 테스트에서 제외하였다. PC와 HMD의 환경에서 테스트한 결과는 다음과 같다.



그림 4. XR Weather 인터페이스  
Fig. 4. XR Weather Interface



그림 5. XR Weather PC 환경 및 HMD 환경 테스트  
Fig. 5. Testing XR Weather PC environment and HMD environment

테스트 결과 PC 환경과 HMD 환경에서 나라와 도시, 날짜를 선택하여 특정 시점의 날씨를 시뮬레이션할 수 있음을 확인하였다. 날씨 변화에 따른 환경적 요소를 실시간으로 배경에 반영할 수 있었으며 개별적으로 날씨의 맑음, 눈, 비 등의 날씨 효과를 적용하여 다양한 환경 변화를 확인하였다.

XR Skybox는 태양과 달의 위치를 조정하여 그림자의 길이나 방향을 변경할 수 있으며, 하늘의 색조를 조정하여 배경의 분위기를 차갑거나 따뜻하게 변경하는 기술이다. 이를 통해 전체 배경의 조명과 분위기를 세밀하게 제어하여 다양한 시각적 효과를 제공한다. PC와 HMD의 환경에서 테스트한 결과는 다음과 같다.



그림 6. XR Skybox PC 환경 및 HMD 환경 테스트  
Fig. 6. Testing XR Skybox PC environment and HMD environment

테스트 결과 PC 환경과 HMD 환경 모두에서 배경의 태양 및 달의 위치를 조정하거나 하늘의 색조 변경, 다양한 필터를 적용하여 배경의 분위기를 다채롭게 변화시킬 수 있음을 확인하였다. 이는 배경의 특정 분위기나 시각적 연출할 수 있도록 지원한다.

#### IV. 결 론

본 연구에서는 3D 퍼포먼스 기술을 통해 다양한 플랫폼에서 효율적이고 몰입감 있는 3D 배경을 생성할 수 있는 방법을 제안하였다. 제안된 기술은 크게 XR Viewer, XR Optimization, XR Weather, XR Skybox로 구분되며, 각각의 기술이 3D 배경 제작 과정에서 중요한 역할을 수행한다.

이러한 기술들은 메타버스, 웹툰, 애니메이션 등 다양한 장르에서 배경 제작에 적용 가능하며, 본 연구에서 실험한 데이터셋을 통해 그 효과성과 실용성을 검증하였다. 향후 3D 퍼포먼스 기술은 XR 스튜디오 환경에서 더욱 효과적인 배경 제작 기술로 활용될 수 있을 것이다.

#### 참 고 문 헌

[1] 서다운, 조영대, 손우람, 장현국, 이희연, and 김형석, "가상 배경 제작을 위한 절차적 도시 생성 플러그인 UI 설계", 한국HCI학회 학술대회, pp. 1017-1020, 2024.



- [2] 권중산 and 김재용, "실시간 CG렌더링 소프트웨어를 활용한 3D 애니메이션 제작 사례 분석 연구 - 루미온을 활용한 제작 사례를 중심으로", 한국콘텐츠학회논문지, Vol. 23, No. 12, pp. 162-175, 2023.
- [3] 모달 and 김효용, "실감 미디어 제작 툴을 활용한 3D 모델링 연구", 애니메이션연구, Vol. 20, No. 1, pp. 25-43, 2024.
- [4] 김금영, "실시간 렌더링 제작 도구 비교 분석", 차세대컨버전스정보서비스기술논문지, Vol. 11, No. 5, pp. 499-509, 2022.
- [5] 안용, "리토폴로지 효율성 향상을 위한 소프트웨어의 비교분석 및 유형별 3D 모델링 사례 제작", 한국콘텐츠학회논문지, Vol. 20, No. 6, pp. 9-25, 2020.
- [6] Unity AssetStore: <https://assetstore.unity.com/vfx/shaders> [accessed: April. 09. 2024]
- [7] Unreal Engine: <https://www.unrealengine.com/marketplace/ko/assets?keywords=Shader> [accessed: April. 09. 2024]

# MFC 운영 시스템과 모빌리티 연계를 위한 API 체계 설계

김정권\*, 박준호\*\*

## Design of API system for linking MFC operating system and mobility

Kim Jeoung-Kwon\* and Park Juno\*\*

### 요 약

MFC(Micro Fulfillment Center)는 다수의 생산자 및 공급자로부터 다품종 소량의 물품을 공급받아 입고, 보관하여야 하며 접수된 출고 지시에 따라 피킹 및 패키징된 물품을 출고하는 다수의 물류 업무를 신속히 수행할 수 있어야 한다. MFC에서는 물류 공정 전반에 걸쳐 무인 모빌리티 디바이스를 활용하여 물품을 이송하는 방식을 도입하고 있다. 이러한 방식의 성공적인 구현을 위해 MFC 운영 소프트웨어는 물류 모빌리티 디바이스를 원격으로 제어할 수 있어야 한다. 본 연구에서는 MFC 운영 소프트웨어에서 물류 모빌리티 디바이스를 효과적으로 제어하기 위한 API 체계의 설계 방안을 제안한다.

### Abstract

Micro Fulfillment Center (MFC) must receive, store, and deliver a variety of small quantities of goods from multiple manufacturers and suppliers, and must be able to quickly perform a number of logistics tasks, such as picking and packing goods and shipping them out according to received shipping instructions. MFC is introducing a method of transporting goods by utilizing unmanned mobility devices throughout the logistics process. For the successful implementation of this method, MFC operating software must be able to remotely control logistics mobility devices. In this study, we propose a design method of an API system for effectively controlling logistics mobility devices in MFC operating software.

### Key words

api designg, mfc operation, mobility linkage

### 1. 서 론

MFC는 수요자 밀집 지역인 도심 인근에 소규모로 구축하여 물리적인 배송 거리를 줄여 빠른 배송

을 실현을 목표로 구축되고 있다. 작은 보관 면적에서 최대한의 물동량을 신속히 처리하기 위해서 WMS (Warehouse Management System)와 WCS (Warehouse Control System)를 활용하여 센터 내의

\* Studio 3S Korea, john@studio3s.co.kr, \*\*Studio 3S Korea, juno@studio3s.co.kr.

※ Acknowledgement(2197)

- 본 연구는 국토교통부국토교통과학기술진흥원의 지원으로 수행되었음(과제번호 RS-2023-00256323)

정보를 처리하며 AGV (Automated Guided Vehicle), AMR (Autonomous Mobile Robot) 등의 모빌리티 디바이스를 사용하여 물리적인 물류 공정을 처리할 수 있도록 자동화된 시스템을 구축한다. 이러한 요구에 따라 MFC의 핵심 운영 소프트웨어인 WCS는 API를 활용하여 모빌리티 디바이스를 제어하여야 한다. 모빌리티 디바이스 제어를 위한 API는 데이터 교환이라는 단일 요구에 집중된 일반적인 정보 시스템의 API와 다르게 모빌리티의 물리적인 동작 상황에 따라 제어의 가능 여부가 결정되어야 하는 특징이 있으며 디바이스에서 처리중인 단일 트랜잭션 소요시간이 수초에서 수분으로 상대적으로 길기 때문에 이러한 특성을 감안하여 API가 구현되어야 한다. 본 논문에서는 이를 위한 효과적이며 직관적인 API 체계의 설계 방안을 제안한다.

## II. API Server

일반적인 상황의 MFC에서는 입고, 피킹, 배송 등의 업무 특성에 적합한 다양한 종류의 모빌리티 디바이스를 도입하여 운영하게 된다. 이런 경우 도입된 디바이스의 제조사가 다른 경우가 대부분이며 디바이스별 컨트롤 방식과 API 연동 방식 및 프로토콜이 제각기 상이하서 WCS는 디바이스별 각기 다른 API를 사용하기 위하여 복수의 API 사용 모듈을 개발하여 사용하여야 한다. 이러한 상황은 시스템상 제어 체계의 복잡도를 증가시켜 개발, 테스트, 운영 등의 활동시에 리소스를 증가시키는 원인이 되기도 한다.

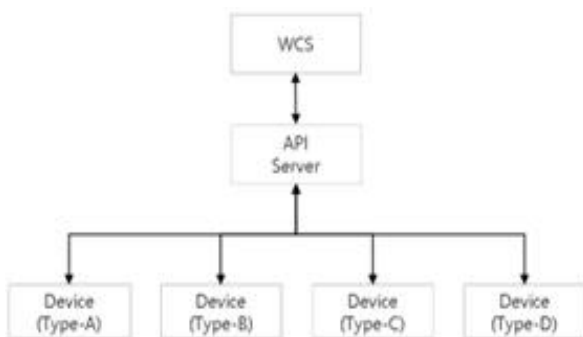


그림 1. API Server 구성  
Fig. 1. API Server Composition

이런 문제를 해결하기 위해 WCS에서 각기 다른 API를 개별적으로 연동하는 방식보다는 멀티 디바이스용 API 서버를 도입하여 WCS와 API 서버 구간의 데이터 교환 체계를 단일화하여 시스템 복잡도를 감소시키고 API 서버와 각 디바이스 구간별 데이터 교환 체계를 정의하여 다양성을 향상시키는 방식을 제안한다. 근래들어 설계 및 구축되는 대다수의 비즈니스 어플리케이션 시스템은 AWS (Amazon Web Services), MS Azure, GCP (Google Cloud Platform) 등의 Cloud Platform 기반에서 SaaS (Software as a Service), PaaS (Platform as a Service), IaaS (Infrastructure as a Service) 등의 형태로 구축되어 운영되어 있으며 WAS (Web Application Server) 상에서 구동되는 Client-Server 방식을 채택하고 있다. 또한 HTTP (Hyper Text Transfer Protocol)을 기본 프로토콜로 사용하며 외부 시스템과의 연동을 위한 API에 REST (Representational State Transfer) 아키텍처와 JSON (JavaScript Object Notation) Format을 결합한 기술을 사용하고 있다

이러한 환경에 따라 API Server는 WAS를 기반으로 구축하며 API 구현 방안으로 REST와 JSON Format을 사용하는 것을 제안한다. 산업용 모빌리티 로봇 시장에서 디바이스와 외부 시스템과 데이터 연동을 위해서 주로 사용하는 프로토콜은 전통적으로 Modbus, Profibus, CANbus, CC-Link 등을 사용하였으나 근래 들어 높은 비용 효율성과 단순한 물리적 구성과 향상된 속도 등의 이유로 Ethernet 기반의 프로토콜을 주로 사용하고 있다.

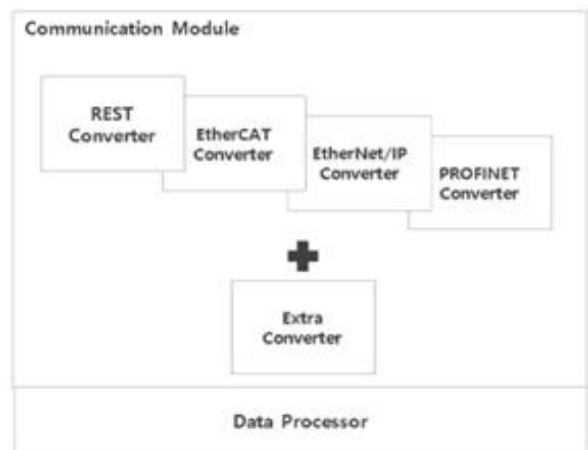


그림 2. 멀티 프로토콜 지원 체계  
Fig. 2. Multi-protocol support system

주요 Ethernet 기반의 프로토콜로 EtherCAT, EtherNet/IP, Profinet, Ethernet POWERLINK, Modbus TCP 등이 있다. API 서버는 연결하고자 하는 개별 디바이스별 통신 프로토콜을 위한 컨버터를 개발하여 이를 커뮤니케이터 모듈에 추가하는 방식으로 구축하여 호환성을 확보하고자 한다.

### III. 비동기식 데이터 교환

디바이스의 물리적인 동작을 제어하는 API 호출은 제어에 대한 결과를 호출 후 수초에서 길게는 수분 후에 알 수 있게 된다. 물리 제어 요청이 성공인지 실패인지에 대한 결과와 물리 제어의 진척 상태에 대한 주기적인 결과 보고가 이루어져야 WCS에서 해당 물리 제어의 진척 상태에 맞게 연결된 후속 제어를 준비할 수 있다.

예를 들어 운송 차량에게 물품을 목적지로 이송하라는 물리 제어를 요청한 경우 운송 차량이 목적지 인근 위치에 도달한 진척 상태를 보고하여 WCS는 지게차 또는 작업자에게 물품 하차를 위한 준비를 지시할 수 있다. 또는 해당 물리 제어를 취소하려는 경우에 물리 제어의 상태에 따라 취소 요청의 가능 여부를 판단할 수 있는데 가령 운송 차량에게 이송 제어의 취소를 요청하고 목적지를 변경하고자 한다면 목적지 도착 이전 또는 하차 도크에 접안 이전에만 이러한 취소 요청이 가능할 것이다.

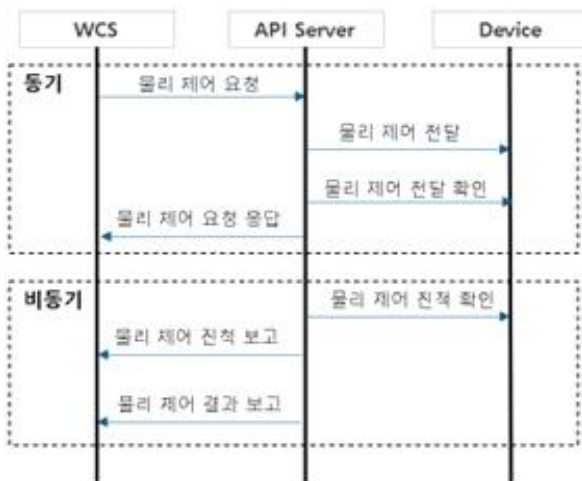


그림 3. 비동기식 데이터 교환  
Fig. 3. Asynchronous data exchange

따라서 물리 제어 요청 시 API 서버는 WCS에게 제어 요청이 정상적으로 디바이스에 전달되었는지에 대한 여부만 즉시 반환하며 요청된 제어의 진척 상태를 API 서버가 능동적으로 모니터링하여 주기적으로 WCS에게 보고하여야 한다. 또한 해당 제어 요청이 정상적으로 또는 비정상적으로 완료된 경우에 이를 WCS에게 즉시 보고하여야 한다. 위와 같은 특징과 문제점을 해결하기 위해 디바이스의 물리적 제어는 병렬로 처리될 수 있어야 하며 이를 위해서 비동기 방식으로 데이터 교환을 설계하는 것을 제안한다.

### IV. Callback Design Pattern

비동기식 데이터 교환체계를 기반으로 WCS는 모빌리티 디바이스로부터 물리적 제어의 실행 진척과 결과를 진척 상태의 변경과 성공 또는 실패로 완료 되어진 실행 결과를 수신받을 수 있다. 이러한 실행의 진척 및 결과 정보가 응답 시스템에서 요청 시스템으로 비동기로 전달되어지는 환경에서 데이터 제어 프로세스를 설계하기 위해 Callback Pattern, Promise Pattern, Async/Await Pattern 등의 설계 기법을 고려해 볼 수 있는데 본 논문에서는 기본적인 직관적인 Callback Pattern으로 소프트웨어를 설계하는 것을 제안한다.

응답 시스템으로부터 요청 시스템으로 전달되는 데이터를 Callback Data라 정의하며 이러한 데이터가 수신되는 상황을 Callback Event 발생한다고 정의한다.

Callback Event	Callback Data		Callback Event Handler
	Key	Value	
Started			OnStarted
Changed	Status	100	OnChanged100
		200	OnChanged200
Finished	Result	Success	OnSucceeded
		Failure	OnFailed

그림 4. Callback 제어 구조  
Fig. 4. Callback control structure

특정 Callback Event 발생시 이에 매핑된 Callback Event Handler가 시스템에서 자동적으로 수행되기 위해서는 일반적으로 제어 요청 시 Callback Event 별 매핑 Callback Event Handler의 식별자를 요청 매개변수에 함께 전달하도록 설계한다. 제안한 방식으로 구현된 시스템을 가정한다면 지게차가 물품을 집어 목적지로 이동하여 물품을 내려놓는 물리 제어의 실행을 요청한 경우 물품을 집어서 이동중이라는 Callback Event 에는 물품 적재를 업무를 담당하는 작업자를 목적지로 이동시키기 위한 지시를 수행하는 Callback Event Handler를 매핑하며 물품을 목적지에 내려놓은 Callback Event 발생시에는 작업자에게 지게차가 내려 놓은 물품을 적재하라는 지시를 수행하는 Callback Event Handler를 매핑할 수 있다. 제안한 방식을 활용하여 MFC 운영을 위해 연속적이며 효과적인 Control Chain을 구현할 수 있다.

## V. 가상 디바이스 모델

WCS는 주기적이며 상시적으로 모빌리티 디바이스의 현재 상태를 파악할 수 있어야 한다. 디바이스에게 물리적 제어 요청하기 이전에 해당 디바이스가 물리적 제어가 불가능한 상태임을 WCS가 인지할 수 있다면 실행 불가능한 제어를 불필요하게 요청하는 행위를 회피할 수 있으며 이를 통해 효과적인 리소스 관리를 실현할 수 있을 것이다.

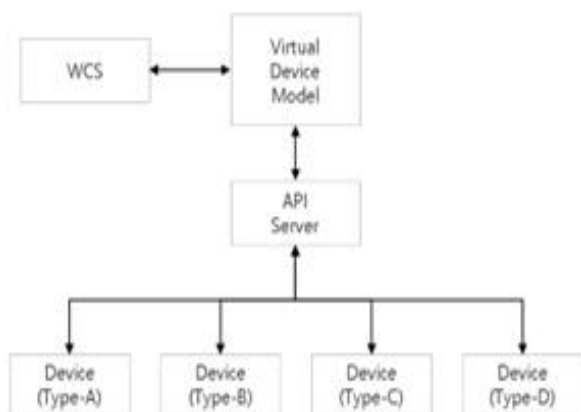


그림 5. 가상 디바이스 모델  
Fig. 5. Virtual Device Model

이를 위한 직관적인 방법으로 디바이스의 상태 파악을 위한 API를 호출하여 디바이스로부터 현재 상태를 조회하는 방식을 선택할 수 있으나 점진적으로 연계된 디바이스의 수가 증가하는 경우 상태 파악을 위한 API 호출 빈도가 동반 상승하게 되어 WCS 리소스 관리면에서 과도한 부하의 원인이 될 수 있다. 이러한 문제를 효과적으로 해결하기 위해 가상 디바이스 모델 (VDM : Virtual Device Model)을 도입하여 디바이스의 상태를 파악하는 Task를 VMD에게 위임하는 방안을 제안한다. VDM은 사전에 정의된 시간 간격에 따라 연결된 모든 실제 디바이스로부터 현재 상태를 주기적으로 모니터링하며 이를 자체 메모리에 적재한다. WCS에서 VDM에게 디바이스 상태를 요청하는 즉시 VDM은 디바이스별 상태 정보를 메모리에서 조회하여 반환하도록 설계한다. 또한 추가적인 기능으로 디바이스에서 모니터링 가능한 여러 속성의 조합을 통해 새로운 속성을 정의하여 WCS에게 제공할 수 있어야 한다. 이러한 기능을 통해 추상화(abstraction) 개념을 손쉽게 구현할 수 있게 할 수 있으며 이는 다양한 디바이스를 관리하여야 하는 WCS의 설계와 구현을 위해 OOP (Object-Oriented Programming) 방법론을 도입하여 시스템을 설계 및 구축 할 수 있도록 한다. VDM을 WCS와 API Server 사이에 위치하도록 하여 디바이스 제어를 위해 WCS에서 API를 직접 호출하지 않고 VDM에게 제어를 요청하면 VDM이 API Server에 제어 요청을 전송하도록 한다면 보다 향상된 추상화를 실현할 수 있을 것이다

## VI. 디바이스 제어 대행

특정 디바이스의 현재 상태가 물리적 제어 요청을 즉시 실행할 수 없는 상태인 경우 WCS는 디바이스의 상태가 제어가 가능한 상태로 변경되기까지 기다려야 하는 waiting 상황이 발생할 수 있다. 빈번히 발생하는 waiting 상황은 WCS의 원활한 디바이스 제어를 방해하는 요인이기 때문에 이를 회피하는 방안으로 WCS의 제어 요청을 수신받아 큐(Queue)에 적재하고 제어 대상 디바이스가 실행 가능한 상태로 변경되면 큐에 적재되어 있는 제어 요

청을 실행하는 역할을 대행하는 디바이스 제어 대행 (DCA : Device Control Agent) 역할이 필요하다. 이는 WCS를 과도한 waiting 상황으로부터 보호하는 효과적인 방안이 될 수 있다. 또한 수신 받은 제어 요청을 저장하는 큐(Queue)를 Priority Queue 방식으로 설계하여 우선순위에 따라 제어 요청의 실행 순서를 WCS에서 정의할 수 있도록 하며 예약된 제어 요청의 실행 우선순위의 변경 기능과 예약의 취소 등의 기능을 제공한다면 WCS가 효과적으로 디바이스 제어하기 위한 해결 방안이 될 것으로 판단된다.

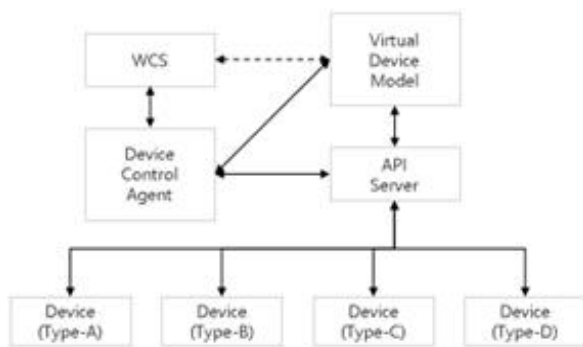


그림 6. 디바이스 제어 대행  
Fig. 6. Device Control Agent

## VII. 결 론

상시적으로 물리적 위상이 변화되는 다양한 프로토콜의 모빌리티 디바이스를 원격지 정보 시스템에서 API를 기반으로 효과적으로 제어하기 위해서 MFC 운영 시스템인 WCS와 디바이스 사이에 API Server를 도입하여 양 시스템의 데이터 교환 규칙을 단일화하고 추가적인 프로토콜 적용은 API Server에 위임시키며 시스템간 데이터 교환을 비동기식으로 통신하며 Callback Design Pattern 방식으로 프로세스를 실행하도록 하여 디바이스의 물리적 제어를 병렬 처리를 구현할 수 있는 설계안을 제안하였다. 또한 가상의 디바이스 모델을 제안하였는데 이는 MFC의 핵심 운영시스템인 WCS의 설계 및 구축에 객체 지향적인 방법론을 적용한 경우 다양한 디바이스를 추상화하여 안정적으로 제어할 수 있는 선택지를 제공한다. 아울러 본 논문에서 제안한 디바이스 제어 대행자를 기반으로 검증된 기존 운영 소프트웨어의 API 설계 방식을 물리적 위상의 변환하

는 디바이스 제어를 위한 시스템에 무리없이 적용할 수 있을 것으로 예상된다. 향후 더욱 향상된 모빌리티 디바이스 제어용 시스템 설계를 위한 심화 연구가 진행되기를 기대한다.

## 참 고 문 헌

- [1] Arnaud Lauret, "The Design of Web APIs First Edition", Manning, November 2019.
- [2] James Higginbotham, "Principles of Web API Design: Delivering Value with APIs and Microservices 1st Edition", Addison-Wesley Professional, December 2021.
- [3] Vaughn Vernon, "Implementing Domain-Driven Design 1st Edition", Addison-Wesley Professional, February 2013.
- [4] Ian Verhappen, Augusto Pereira, "Foundation Fieldbus 4th Edition", International Society of Automation, February 2012.
- [5] Peter Chipkin, "Modbus For Field Technicians", CreateSpace Independent Publishing Platform, January 2011.
- [6] Stephen R. Schach, "Object-Oriented and Classical Software Engineering 8th Edition", McGraw Hill, July 2010.
- [7] Ben Weidig, "A Functional Approach to Java: Augmenting Object-Oriented Java Code with Functional Principles 1st Edition", O'Reilly Media; June 2023

# 스마트 시스템 에어컨 청소 로봇시스템의 설계와 서비스 플랫폼 구현

이연수\*, 김능한\*\*

## Design and implementation of a service platform for a smart system air conditioner cleaning robot system

LEE YOUNSU\*, and KIM NEUNGHAN \*\*

### 요 약

본 논문은 시스템 에어컨 청소를 자동화하고 효율성을 높이기 위한 로봇 시스템 및 서비스 플랫폼을 설계하고 개발하는 것을 목적으로 한다. 연구에서는 몸체로봇 PLC 컨베이어 시스템, 고압 분사 및 공기 분사 노즐, PLC 제어 시스템, 서비스 플랫폼 등의 요소들을 통합하여 청소 효율성과 사용 편의성을 극대화하였다. 실험 결과, 각 요소가 높은 효율성과 성능을 발휘하여 청소 작업의 질을 크게 향상시켰으며, 서비스 플랫폼을 통해 사용자가 손쉽게 청소 상태를 확인하고 예약할 수 있도록 하였다. 이 연구는 향후 더 정밀하고 자동화된 청소 시스템 개발에 기여할 것으로 기대된다.

### Abstract

This paper aims to design and develop a robot system and service platform to automate and improve the efficiency of cleaning system air conditioners. The research integrates elements such as body robot PLC conveyor systems, high-pressure and air spray nozzles, PLC control systems, and service platforms to maximize cleaning efficiency and user convenience. Experimental results demonstrated that each element exhibited high efficiency and performance, significantly enhancing the quality of cleaning tasks. The service platform allowed users to easily check and schedule cleaning status. This study is expected to contribute to the development of more precise and automated cleaning systems in the future.

### Key words

Smart robot system, Robot cleaner, Refraction Robot, Air conditioner service platform, Robot cleaner service scenario

## 1. 서 론

현대 사회에서는 실내 공기의 질이 건강과 생활

의 질에 큰 영향을 미치는 것으로 알려져 있다[1]. 특히, 시스템 에어컨은 실내 공기 순환과 냉방을 담당하는 중요한 역할을 하지만, 장기간 사용 시

\*(주)효인산업, soosungtec@daum.net , \*\*\*(주)효인산업, soosungtec@daum.net (교신저자)

※이 연구는 2024년도 산업통상자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임(20019225)

내부에 먼지와 오물이 쌓여 성능 저하와 함께 건강에 해로운 영향을 미칠 수 있다. 따라서, 시스템 에어컨의 정기적인 청소와 유지보수는 필수적이다. 그러나 높은 천장에 설치된 시스템 에어컨을 청소하는 작업은 작업자에게 물리적인 부담을 주며, 효율적인 청소 방법의 필요성이 대두되고 있다.

본 연구의 목적은 시스템 에어컨 청소를 자동화하고, 효율성을 높이는 청소 로봇 시스템을 설계하고 개발하는 것이다[2]. 이를 통해 작업자의 부담을 줄이고, 청소 품질을 향상시키며, 정기적인 유지보수가 용이하게 이루어질 수 있도록 한다. 또한, 스마트 플랫폼을 통해 사용자가 청소 상태를 확인하고 예약할 수 있는 시스템을 구축하여 사용자 편의성을 극대화하고자 한다.

이 연구는 청소 로봇의 설계 및 개발, 고압 분사 및 공기 분사 노즐, PLC 제어 시스템, 서비스 플랫폼 등의 구성 요소를 통합적으로 다루며, 각각의 요소들이 조화롭게 작동할 수 있도록 설계되었다. 이를 통해 기존의 시스템 에어컨 청소 방식보다 향상된 성능과 효율성을 제공할 것으로 기대된다.

## II. 초점 형상 복원을 위한 인공지능망

본 연구에서는 몸체로봇의 클리닝 시스템을 4m 이상의 높이에서 효과적으로 제어하기 위해 PLC 기반의 제어 시스템을 설계하였다. 이를 위해 높이 센서와 연계된 피드백 루프를 구성하여 실시간으로 높이를 조절할 수 있도록 하였으며, 고정밀 모터를 사용하여 미세한 높이 변화도 정밀하게 조정할 수 있도록 하였다. 로봇의 원활한 작동을 위해 안정적인 전력공급과 전기통신이 필수적이다. 본 연구에서는 전력공급 모듈과 통신 케이블을 일체화하여 설치 및 유지보수를 용이하게 하였다. 또한, 전자기 간섭을 최소화하기 위해 차폐 케이블을 사용하였고, 케이블 연결 부분에는 빠른 연결이 가능한 커넥터를 채택하였다. 에어컨 청소 시 발생하는 오물이 외부로 튀는 것을 방지하기 위해 투명 플라스틱 커버

를 설계하였다[3]. 플라스틱 소재는 고투명도를 유지하면서도 내구성이 뛰어나 세척 시 변형이 없도록 하였다. 또한, 커버는 쉽게 탈착 가능하도록 하여 사용 후 세척이 용이하게 하였다.

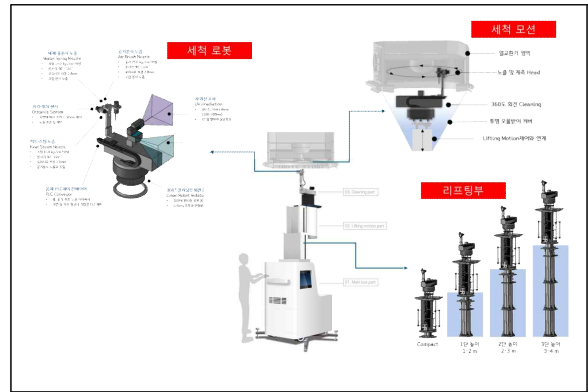


그림 1. 시스템 에어컨 로봇 청소기 구성도  
Fig. 1. System Air Conditioning Robot Cleaner Configuration Diagram

MAIN BOX 내부에는 고압펌프를 설치하여 물과 세제를 최대 4.5M 높이의 클리닝부로 이동시키도록 설계하였다. 이를 통해 고층 청소 작업에서도 안정적인 세제 공급이 가능하도록 하였다. 세척이 잘 되도록 물의 온도를 높이기 위해 MAIN BOX 내부에 2,500W 히터를 적용하였다. 이 히터는 빠르게 물을 가열하여 세척 효율을 높이며, 온도 조절 기능을 통해 다양한 세척 조건에 맞출 수 있다.

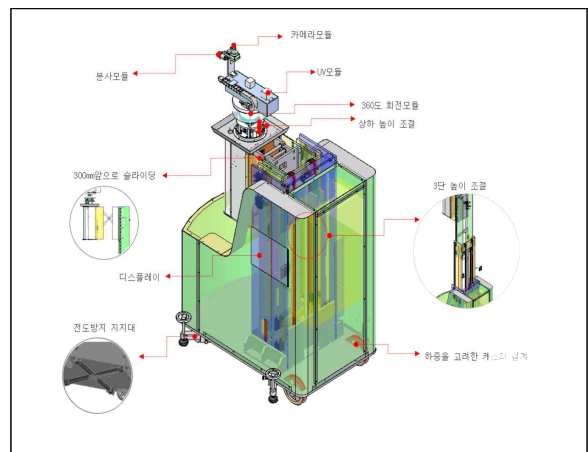


그림 2. 에어컨 로봇청소기 메인박스 설계  
Fig. 2. Air conditioner robot cleaner main box design



Touch-LCD Controller는 세척, 스팀, 행균, 건조의 청소 프로세스를 일괄적으로 제어하고 실시간으로 상태를 확인할 수 있도록 설계되었다. 사용자 인터페이스는 직관적이고 사용하기 쉽게 구성하여, 청소 작업의 효율성을 극대화하였다[4]. Controller는 청소에 필요한 압력, 물의 양, 노즐의 위치와 속도를 정밀하게 제어할 수 있는 기능을 갖추고 있다. 이를 통해 사용자는 특정 청소 조건에 맞춘 세밀한 조정을 할 수 있다.

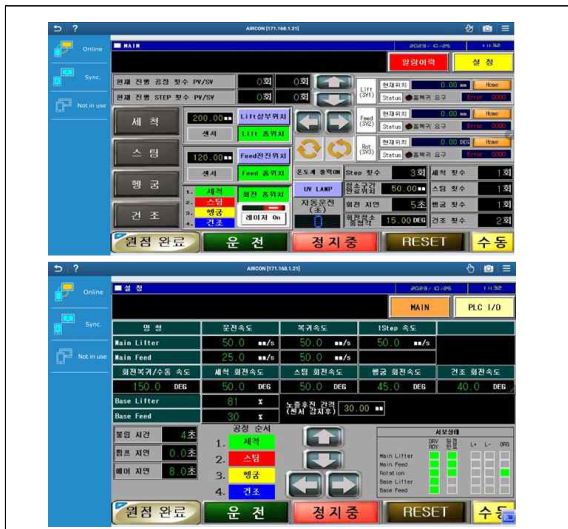


그림 2. HMI 화면  
Fig. 2. 영문제목

본 연구에서는 시스템 에어컨이 설치된 높은 천장에 닿을 수 있도록 상하 작동이 가능한 리프팅 장치를 설계하였다. 리프팅 높이는 최대 4m로 설정하여, 고층 천장에서 안정적으로 작업할 수 있도록 하였다. 리프팅 장치는 전동 모터와 정밀 기어를 사용하여 부드럽고 정확한 상하 이동을 가능하게 하였다. 앞쪽 장애물을 피해 효율적으로 청소할 수 있도록 로봇의 좌우 작동 기능을 설계하였다. 좌우 이동 범위는 30mm로 설정하였으며, 이를 통해 청소 중에 장애물을 회피하면서도 정확하게 작업을 수행할 수 있다. 좌우 작동 장치는 소형 서보모터를 사용하여 정밀한 제어를 가능하게 하였으며, 센서를 통해 장애물을 실시간으로 감지하고 회피할 수 있도록 하였다.

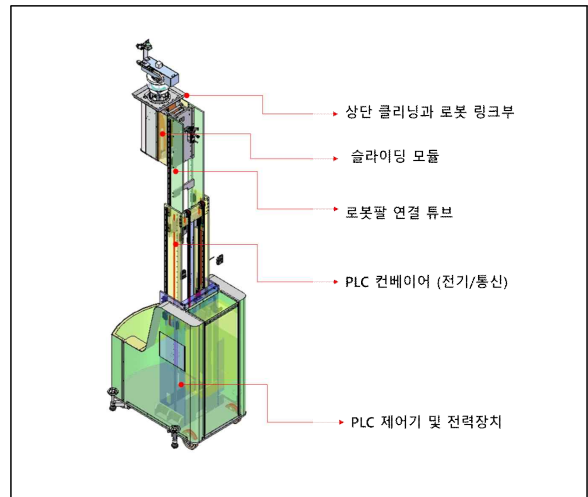


그림 3. 에어컨 로봇청소기 리프팅부 설계  
Fig. 3. Air conditioner robot cleaner lifting part design

세제물이 유량 14.0 /2 이상으로 분사되기 위해, 분사각 90°-124°를 유지하는 고압 분사 노즐을 설계하였다. 노즐의 오리피스 직경은 2.8mm로 설정되어 강력한 세제물이 넓은 범위에 고르게 분포되도록 하였다. 이 시스템은 특히 클리닝 작업 시 효율성을 극대화하기 위해 개발되었다. 공기분사(=air brush)를 위한 노즐은 유량 5.0 /2 이상으로 분사각 90°-124°를 유지하며, 오리피스 직경 또한 2.8mm로 설계되었다. 이 노즐은 주로 세밀한 청소 작업에 사용되며, 고압 공기 분사를 통해 이물질을 제거하는 데 효과적이다. Heat Stream 노즐의 경우, 곡면을 따라 거리 1-10cm를 정밀하게 제어할 수 있도록 설계되었다. 노즐은 센서와 연계되어, 실시간으로 표면의 곡률에 따라 거리와 위치를 조정할 수 있다. 이를 통해 청소 효율성을 높이고, 특정 위치를 정밀하게 조준 및 제어할 수 있다. 자외선(UV-C) lamp는 100-200nm 범위의 자외선을 3개 발생시켜 살균 효과를 구현하도록 설계되었다. 이러한 설계는 병원, 실험실 등의 환경에서 높은 살균 효과를 발휘하며, 인체에 무해한 범위에서 작동한다. 360° 클리닝부 회전 몸체는 회전에 편리한 로봇 몸체와 Lifting 로봇과 연결되어 개발되었다. 이는 다양한 위치에서 클리닝 작업을 수행할 수 있도록 하며, 높이 조절이 가능하여 작업 환경에 최적화된 청소를 가능하게 한다. 몸체의 PLC 제어 시스템은 물, 공기, 히트 노즐의 거리를 정밀하게 제어할 수 있

도록 설계되었다. 각 노즐의 작동을 실시간으로 모니터링하고 제어할 수 있으며, 이를 통해 청소 작업의 효율성과 정밀성을 높였다.

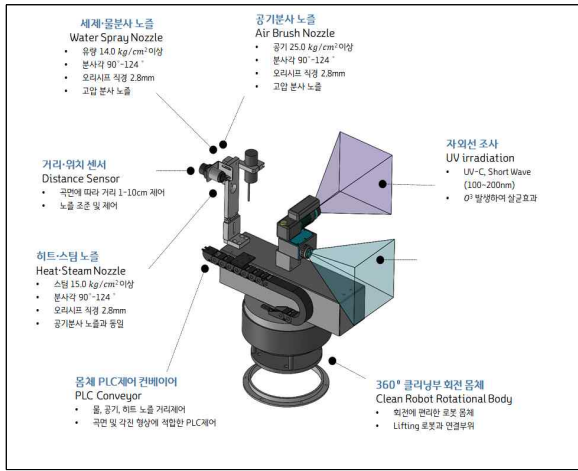


그림 4. 에어컨 로봇청소기 클리닝부 설계  
Fig. 4. Air conditioner robot cleaner cleaning part design

스마트 에어컨 청소로봇을 사업화하기 위해 플랫폼 형태의 서비스를 제공하는 구성을 설계하였다 [5]. 이 서비스 플랫폼은 청소 예약, 결과 확인, 피드백 제공 등 다양한 기능을 포함하며, 사용자가 손쉽게 접근하고 이용할 수 있도록 개발되었다. 고객은 플랫폼을 통해 청소 예약을 간편하게 할 수 있다. 사용자 인터페이스는 직관적이고, 사용자는 원하는 날짜와 시간을 선택하여 청소 예약을 완료할 수 있다. 예약 완료 후, 청소 로봇이 자동으로 스케줄에 맞춰 작동하게 된다. 고객은 청소가 완료된 후, 플랫폼을 통해 청소 결과를 영상으로 확인할 수 있다. 또한, 세척 상태를 정량적으로 확인할 수 있는 측정 결과가 제공되며, 이를 통해 청소의 효율성과 상태를 명확하게 판단할 수 있다. 소비자와 청소를 제공하는 사업자 간의 연결 기능을 제공하여, 소비자가 직접 사업자에게 피드백을 주고, 추가적인 요청을 할 수 있도록 하였다. 이를 통해 서비스 품질을 향상시키고, 사용자 만족도를 높일 수 있다.

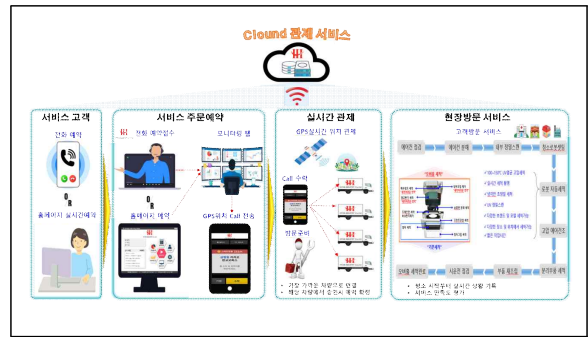


그림 5. 스마트 에어컨 로봇청소 플랫폼 구성도  
Fig. 5. Smart Air Conditioner Robot Cleaning Platform Configuration Diagram

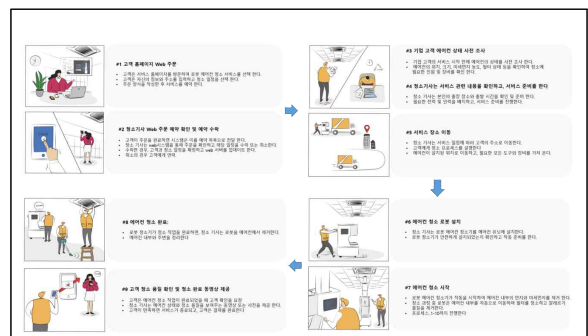


그림 6. 스마트 에어컨 로봇청소 서비스 시나리오  
Fig. 6. Smart Air Conditioner Robot Cleaning Service Scenario

### III. 결 론

본 논문에서는 시스템 에어컨 청소로봇의 설계 및 개발을 통해 다음과 같은 성과를 얻었다.

첫째, 높이 제어 시스템: 몸체로봇의 높이 제어 시스템을 4m 이상으로 구현하여 고층에서도 안정적인 청소가 가능하게 하였다.

둘째, 오물 우회관 및 투명 오물받이 플라스틱: 오물이 효율적으로 모이도록 설계된 우회관과 청소 시 외부로 오물이 튀는 것을 방지하는 플라스틱 커버를 개발하였다.

셋째, 정밀 제어기 및 차단기: HMI를 통한 정밀한 제어가 가능한 시스템을 개발하였으며, 전력 공급의 안정성을 높이는 차단기를 설계하였다.

넷째, 고압 분사 및 공기 분사 노즐: 세제와 물, 공기를 고압으로 분사할 수 있는 노즐을 개발하여

청소 효율을 극대화하였다.

다섯째, 로봇 팔 모듈 및 클리닝 모듈: 다양한 청소 작업을 수행할 수 있는 로봇 팔과 클리닝 모듈을 설계하였다.

여섯째, 서비스 플랫폼: 청소 예약 및 결과 확인이 가능한 서비스 플랫폼을 개발하여 사용자의 편의성을 높였다.

각 구성 요소의 테스트 결과, 시스템은 고층 청소 작업에서 높은 효율성과 성능을 발휘하였다. 높이 제어 시스템과 고압 노즐은 예상대로 작동하여 청소의 질을 크게 향상시켰다. 서비스 플랫폼의 경우, 사용자 인터페이스가 직관적이어서 사용자가 쉽게 접근하고 이용할 수 있었다.

본 연구에서는 시스템 에어컨 청소로봇의 설계 및 개발을 통해 고층 청소 작업의 효율성과 편의성을 향상시킬 수 있는 시스템을 제안하였다. 각 구성 요소의 개발과 통합을 통해 높은 성능을 발휘하는 청소 로봇을 구현하였으며, 서비스 플랫폼을 통해 사용자가 쉽게 이용할 수 있도록 하였다. 향후 연구를 통해 더욱 정교하고 자동화된 시스템을 개발하여, 보다 높은 수준의 청소 서비스를 제공할 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1]박영수. (2017). 에어컨 청소시스템의 발전방향. 냉동공학회논문지, 25(2), 175-190.
- [2]조수현. (2018). 자동화된 청소로봇시스템의 설계 및 개발. 로봇학회지, 10(4), 150-160.
- [3] 박정훈, 최진영. (2019). 고압노즐설계 및 응용기술. 한국기계기술학회논문집, 15(2), 120-135.
- [4]정은지, 송민준. (2023). HMI 정밀제어시스템의 구현. 한국제어학회논문지, 28(2), 50-65.
- [5] 김영호. (2022). 스마트 서비스플랫폼의 설계와 구현. 한국정보과학회논문지, 30(1), 85-100.

## 원전 현장점검 절차서 전자서식 변환방안 연구

이상현\*, 김남현, 이수일, 한상재, 유기중

### A Study on the Conversion Plan of Field Work Procedure for Mobile Solutions in Nuclear Power Plants

Sang Hyun Lee\*, Nam-Hyun Kim, Soo-Ill Lee, SangJae Han, and Ki-Jung Ryu

본 연구에서는 원전 현장점검 절차서의 디지털 전환을 통해 점검 절차의 효율성을 향상시키고자, 기존 종이 문서 형태의 점검 절차서를 전자서식으로 변환하는 방안을 모색하였다. 원전 현장점검은 안전성과 정확성을 동시에 요구하며, 다양한 점검항목을 체계적으로 관리하는 것이 필수적이다. 그러나 현재 대부분의 현장점검 절차서는 종이 문서에 의존하고 있어 실시간 데이터 입력, 정보 공유, 점검기록 관리 등에 어려움이 있다. 이에 따라 모바일 기기를 활용하여 현장에서 직접 입력이 가능한 전자서식을 개발함으로써 점검 절차를 간소화하고, 데이터의 일관성과 접근성을 높이는 방안을 제안하였다.

연구는 먼저 원전 현장점검 절차서의 구조와 점검 프로세스를 분석하고, 주요 절차서 항목들을 디지털화할 수 있는 요소들을 도출하였다. 이를 바탕으로, 현장에서 사용 가능한 모바일 전자서식의 설계방안을 제시한다. 사용자 편의성과 데이터 관리의 효율성을 고려하여 사용자 인터페이스와 사용자 경험을 최적화하고, 다양한 입력방식과 데이터 검증 기능을 포함하여 사용자의 오류를 최소화하고 데이터의 신뢰성을 확보할 방안을 검토한다. 실시간 입력이 가능한 모바일 환경의 특성을 살려, 원전에 구축된 무선 인프라 환경에서 점검 데이터를 실시간으로 수집하고 공유할 수 있는 기능도 고려가 필요하다. 현장 데이터의 실시간 전송과 데이터베이스와의 연계를 통해 효율적인 데이터 관리체계를 구축하는데 기여할 수 있다. 전자서식을 통한 현장점검 절차는 종이문서를 사용할때 발생하는 데이터 입력 및 관리의 비효율성을 개선하는 효과가 있을 것이며, 점검과정에서 발생하는 실시간으로 수집하고 분석할 수 있어, 유지보수 및 안전성 확보가 가능할 것이다. 향후 다양한 점검 절차서로 확장 적용할 수 있도록 원전 절차서 전자서식 변환모델을 개발하고자 한다.

#### Key words

digital transformation, mobile solution, field work procedure, electronic form, nuclear power plant

#### 참고문헌

- [1] J. H. Yun & Y. H. Lee, "Development of a Human Error Evaluation Method for Introducing Smart Mobile to Nuclear Power Plants", The Ergonomics Society of Korea(ESK), Academic conference paper pp.190-195, 2011.
- [2] S. J. Lee, N. K. Seong & Y. S. Jung, "The Development of Computerized Procedure System based on the Mobile Technology for the Human Error Reduction in Nuclear Power Plants", The Korean Institute of Electrical Engineers (KIEE), Academic conference paper pp.137-138, Octobers 2013.

---

\*한국수력원자력 중앙연구원, 12zero514@khnp.co.kr

# 지오펜싱 기술을 이용한 드론 시스템에 관한 연구

남윤정\*, 김희선\*\*

## Research on Drone Systems Using Geofencing Technology

Yun-Jung Nam, Heesun Kim\*\*

### 요 약

드론 활용 범위가 다양해져 비행 신청은 매년 두 배 이상으로 급증하고 있다. 일반인들은 비행 절차를 번거롭게 생각하여 별도의 승인 절차를 거치지 않아 불법 비행이 늘어나는 추세이다. 원전 인근에서 탐지된 불법 드론 중 총 533건 중에 239건은 드론 조종자의 신원을 확인하지 못한 것으로 나타났다. 조종자 파악의 어려움으로 피해 발생시 책임이 불분명하고 사고시 드론 소유자는 책임을 부담할 수 있어야 한다.

본 논문은 모바일 기기를 이용하여 개인식별코드를 생성한 후 지오펜싱(Geofencing) 구역을 설정하여 기기를 등록하였다. 드론 조종자의 위치 식별과 실시간 모니터링이 가능하도록 시스템을 제안하고자 한다.

The increasing range of drone applications has led to a surge in flight applications, doubling every year. However, many individuals find the flight application procedures cumbersome, resulting in a rise in unauthorized flights due to the omission of necessary approval steps. Among the 533 cases of unauthorized drones detected near nuclear power plants, the identities of the drone operators were unconfirmed in 239 instances. This difficulty in identifying operators leads to unclear accountability in the event of damage, and drone owners should be held responsible in case of accidents. This paper proposes a system enabling the identification and location monitoring of drone operators by generating personal identification codes through mobile devices, registering these devices, and establishing geofencing zones.

### Key words

Mobile Geofencing, Drone Geofencing, Personal Identification Code, Drone Identification Code

## 1. 서 론

드론은 조종사가 탑승하지 않고 무선으로 전파 유도되는 비행 및 컨트롤이 가능한 회전익 모양의 항공기를 총칭하며, 국내 항공법에서는 연료를 제외

한 자체 중량이 150kg 이하인 것은 무인 비행 장치로 규정하고 있다. 상용으로 판매되는 대부분의 드론은 150kg 이하의 무인비행 장치이며, 특히 무인 비행 장치 중 자체 중량이 12kg 이하이면서 엔진 배기량이 50cc는 스포츠용 무선 조정 모형항공기로

\*국립안동대학교 대학원 소프트웨어융합학과, bbomi4859@naver.com

\*\*국립안동대학교 소프트웨어융합학과, hskim@anu.ac.kr

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

간주하여 신고 없이 누구나 비행이 가능한 상황이다[1]. 항공안전법 제129조, 시행규칙 제310조 ‘조종사 준수사항’이다. 여기에 ‘휴전선, 서울 도심 상공 일부 등 국방, 보안상의 이유로 비행이 금지된 곳 비행 금지’라는 조항이 있다[2]. 사전에 비행을 신고하고, 비행 신청 과정에서 승인까지 절차적 준수를 강조하도록 항공안전법(시행 규칙)에 적시되어 있다. 드론 윈스톱 민원 포털 서비스에서 비행 승인 및 촬영 승인을 신청할 수 있으며 처리 기간이 3~7일 소요된다. 드론 활용 범위가 다양해져 비행 신청은 매년 두 배 이상으로 급증하고 있지만 일반인들은 이 절차를 번거롭게 생각하여 별도의 승인 절차를 거치지 않아 불법 비행도 늘어나는 추세이다. 최근 5년간 원전 주변에서 적발한 불법 비행 드론 건수는 총 26건이며, 그중 9건은 조종자를 발견하지 못했다[3]. 불법 비행 및 불법 촬영 시 도심에서는 사생활 침해가 될 수 있고 중요 시설 보안 위협 등의 사례 및 군사시설의 주요 핵심기지에 대한 보안에 치명적인 피해를 끼칠 수도 있다. 항공 안전법은 안전한 항행에 초점이 맞춰져 있어 불법 비행을 통한 새로운 범죄나 대응 규정은 미비한 실정이다.

본 논문에서는 모바일 기기를 이용하여 개인식별 코드 생성과 UA 지오펜싱(Geofencing) 구역을 설정하여 조종자의 위치 및 신원을 파악하며 조종자 위치를 모니터링할 수 있는 시스템을 제안하였다.

## II. 연구 동향

모바일 기기의 보편화와 공간 데이터 사용 기술의 발달로 응용분야가 늘어나고 있으며 지오펜싱 시장은 계속 성장할 것으로 예상된다.

### 2.1 지오펜싱

지오펜싱은 Geographic과 Fencing의 합성어로 특정구역에 대하여 사용자의 출입 현황을 알려주는 API(Application Programming Interface)이다. 이는 위치 추적 기술 중 하나인 위성항법장치(GPS, Global Positioning System)를 이용한 인터페이스라 할 수 있는데, 여기서 말하는 지오펜싱(Geofence)는 사용자가 지정한 특정 구역 즉, 가상의 울타리를 의미하며

설정된 GPS 울타리 안에 사용자가 드나드는 것을 감지하는 기능이 바로 지오펜싱인 것이다[4].

### 2.2 지오펜싱 활용

지오펜싱의 응용분야로는 소셜 네트워킹, 마케팅, 스마트 가정 기기, 인사관리, 텔레매틱스, 보안 등 다양한 분야에서 현재 사용이 되고 있다[5].

표 1. 지오펜싱 국내 활용 사례

Table 1. Geofencing Domestic Use Cases

지오펜싱 국내 활용 사례	
배달앱	배달의 민족 배달지역 지정에 활용 매장 점주들끼리 배달구역 겹치지 않기
전동킥보드	전동킥보드 반납구역에 활용
카드 회사	위치기반 마케팅 방법에 활용 가맹점 주변으로 쿠폰, 정보 메시지
인적 자원	회사에서 출퇴근 여부 파악에 활용 근무지 위치 확인 서비스
매장	관광객이 매장을 지날 때 홍보 메시지에 활용
공항	무료 와이파이 서비스 공항 주변 가이드 서비스 제공에 활용
군사	공항 및 군사지역 비행금지 구역에 접근하지 못하도록 제한에 활용

표 1에서와 같이 위치 정보를 담은 모바일 기기가 지정된 구역에 들어오면, 조건에 따라 텍스트 메시지, 이메일, 앱 알림 등 미리 정해진 특정 동작을 실행하게 된다. GPS는 특정 위치를 표시하기 위해 ‘점’ 혹은 위치를 이은 ‘선’으로 표현된다. 반면 지오펜싱은 위치의 범위를 표시해야 하므로 ‘면’으로 표현된다. 즉, 다양한 형태의 다각형(Polygon) 으로 표현이 가능하다. 그림 1은 GPS와 지오펜싱의 차이를 그림으로 표현한 예제이다[5].

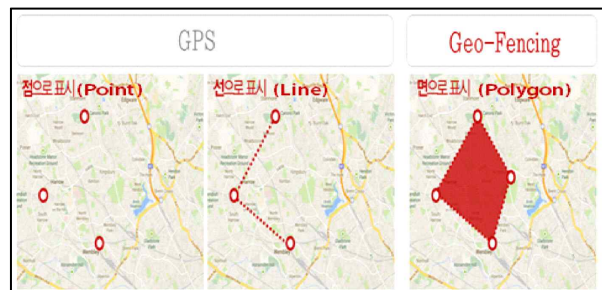


그림 1. GPS 와 Geofencing 차이점  
Fig. 1. Difference GPS and Geofencing

GPS는 특정 대상이 어느 위치에 있는지 즉, 특정 위치가 주요 목적이라 할 수 있고 지오펜싱의 경우 ‘특정 위치에 대한 출입 현황’을 위해 많이 사용된다. 특정 대상이 위치에 있는지 없는지 지오펜싱은 위치의 범위가 목적이라 할 수 있다.

### III. 비교 분석

#### 3.1 DJI GEO 활용 사례

2013년 DJI 드론 제조사가 항공기 보호와 드론 비행 시 조종자가 주변 위험 요소를 알 수 있도록 지오펜싱 시스템을 개발했다. GPS 및 위성 항법 신호를 사용해 드론이 공항, 군사지역 등 안전 및 보안 등의 이유로 비행이 민감한 지역 부근에서 자동으로 드론 사용자의 비행을 제한한다. DJI의 모든 드론은 지오펜싱이 적용된 구역 내에서 특별한 허가 없이 이착륙하는 양단의 3차원 비행 경로에 진입하는 것이 불가능하도록 세팅되어진다.

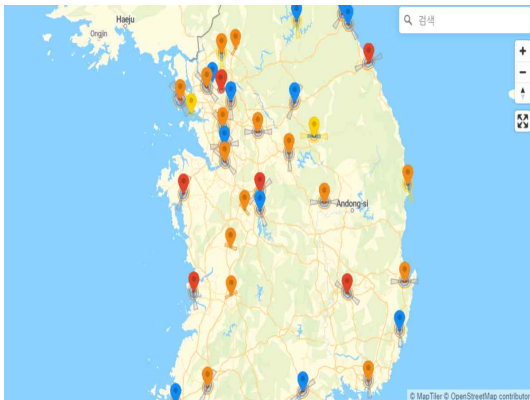


그림 2. 국내 Geo 설정구역  
Fig. 2. Domestic Geo Setting Area

그림 2는 비행 앱(DJI FLY)에서 색상으로 구분하여 지오펜싱 구역을 확인할 수 있으며, 드론의 전원이 켜진 상태일 때 이륙불가라고 알려준다. DJI가 계정을 인증 후 안전비행 메뉴(Fly-safe.dji.com) 새로운 잠금 해제 신청을 할 수 있고 비행이 필요한 경우 DJI GEO시스템에 비행 임시 허가를 받게 되면, 비행금지 지역 안에서 사용자가 DJI 계정을 사용하여 비행규제를 해제(Unlock)한 후 일시적으로 비행 허가 상태를 받을 수 있게 해준다.

표 2는 DJI 공식 사이트에 FLY SAFE GEO 구역

정보의 정책을 내용과 색깔 별로 나타내고 있다.

표 2. DJI사 GEO 구역 색상별 구분  
Table 2. DJI GEO Zone Color Classification

DJI GEO 구역	
● 제한구역	비행은 금지 사용자 알림 메시지
● 고도제한 구역	비행 고도 제한 적용 (공항 인근 회색 표시 구역)
● 허가 구역	비행제한 경고 메시지
● 경고 구역	경고 메시지가 표시 (예: E등급 공역)
● 규제제한 구역	현지 규정 및 정책에 따라 일부 특수 구역 (예: 교도소) 비행금지
● 권장구역	비행 권장하는 구역

#### 3.2 드론 존 GEO 모바일 시스템 제안

스마트 폰의 위치 데이터를 활용 한 다양한 종류의 위치기반서비스(LBS, Location Based Service)들이 개발이 되어 사용자들에 의해 활용이 되고 있다.

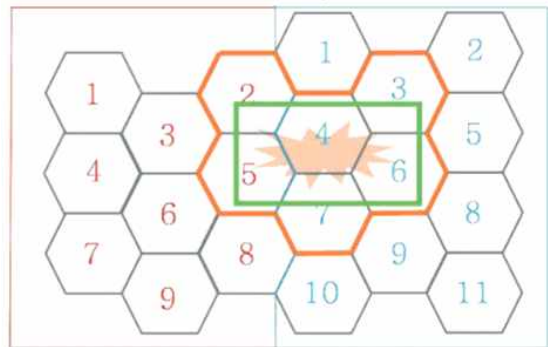


그림 3. 다각형 기반 지오펜싱 서비스  
Fig. 3. Polygon-based Geofencing Service

그림 3은 다각형 기반 지오펜싱 서비스는 위치 기반 서비스에 사용되는 기술로 가상의 경계를 설정하여 특정 지역 진입에 대한 이벤트 또는 알림을 트리거 한다. 4번 구역은 비행 금지 구역이고 2, 5, 7, 6, 3은 UA 공역(초경량 비행장치 전용공역)은 150미터 이하의 고도로 제약없이 비행이 가능한 공역이다. 이 구역을 UA 지오펜싱 구역으로 설정할 수 있다.

DJI GEO 구역의 근접거리에 현재 비행 중인 조종자의 정보가 검색이 된다. 모바일 기기에서 현재 위치 (위도, 경도, 반지름) 값을 계산해 서버에 전달하면 사용자가 지오펜스 구역에 진입할 때 스마트폰으로 전달된다. 모바일 기기를 휴대하고 있는 사용자는 날씨 정보(풍향, 강우, 자기장 지수) 등에 관한 정보를 정보들을 수신 받게 된다. UA 구역 진입 시 모바일은 알림 및 문자 등 미리 정해진 특정 동작을 실행하게 된다. 모바일 기기를 이용하여 개인 식별코드를 생성하고 사용자가 UA 지오펜싱 구역을 설정하면 조종자의 위치 및 신원을 파악할 수 있게 된다.

드론 탐지 기술 동향", 한국전자통신연구원, 2022.

- [4] Trend Insight Business. "Geo-Fencing", available: <http://trendinsight.biz/archives/21929>.
- [5] S. K. White. "What is geofencing? Putting location to work", CIO US, 2017.

#### IV. 결론 및 향후 연구

드론의 수요가 증가하면서 부작용도 함께 커지고 있다. 최근 원전 인근에서 탐지된 불법 드론 조종자의 신원을 확인하지 못한 것으로 나타났다. 드론 사고로 인적·물적 피해도 지속적으로 증가하는 추세이다. 항공 안전법은 안전한 항행에 초점이 맞춰져 있어 불법 비행을 통한 새로운 범죄나 대응 규정은 상당히 미비한 실정이다. 드론 사고는 드론 소유자 조종자에게 있으며 사고에 대한 책임은 따른다.

본 논문에서는 비사업용 드론의 기체 사고 발생 시 소유자와 조종자 파악의 어려움으로 인한 보상의 한계점을 극복하기 위한 대책 마련이 절실함을 강조하였다. 모바일기기를 이용하여 개인식별코드 생성하고 사용자가 UA 지오펜싱 구역을 설정하여 조종자의 위치 및 신원을 파악하는 방법을 모색하였으며, 추가적으로 조종자 위치를 모니터링할 수 있는 시스템을 제안하였다.

#### 참 고 문 헌

- [1] 김두환, 이윤환, "군보안상 드론위협과 대응방안", *Journal of Digital Convergence* Vol. 16. No. 10, pp. 223-233, 2018.
- [2] 엄단비, "항공안전법상 장난감 드론 규제 : 사람이 많이 모인 장소에서 비행 금지의 비현실성", *법학연구* 28(1), 2020.
- [3] 이인재 외 6명, "불법 드론 대응을 위한 저고도



# 정·주기시험 현장 점검시 QR 코드 활용방안

김남현\*, 이수일, 이상현, 한상재, 유기중

## Utilizing QR Codes for On-Site Inspection in Periodic Testing

Nam-Hyun Kim\*, Soo-Ill Lee, Sang Hyun Lee, SangJae Han, Ki-Jung Ryu

본 논문은 원자력발전소에서 QR 코드를 활용한 정·주기 시험 효율화 방안을 제시하고자 한다. 무선망이 구축된 원자력발전소에 QR 코드를 활용하면 설비 및 기기 점검의 효율성과 안전성을 향상 시킬 수 있다. 또한 현장에서 QR 코드를 스캔하여 절차서와 점검 항목을 빠르게 확인하고 점검 후, 점검결과를 공유 및 저장할 수 있다. 그리고 실시간 정보 및 피드백을 통해 설비관리자의 지속적인 개선과 예방정비 전략 수립이 가능하다.

현재 정·주기 시험시 인적오류예방, 점검결과 DB화 등을 위해 종이절차서 대신 모바일절차서 시스템을 개발 및 시험 운영 중이다. 특히 설비점검 항목이 많은 정·주기 시험은 설비점검 누락 등 인적실수를 유발할 수 있다. 그래서 모바일솔루션으로 QR 코드를 활용한 설비 점검방안을 제시한다. QR 코드를 활용한 정·주기 시험은 해당 설비의 각 위치에 QR 코드를 부착하여 정·주기 시험 수행자가 모바일 단말기 등으로 QR 코드를 스캔하면 점검지의 해당항목으로 이동하여 점검을 수행할 수 있어 신속하고 정확한 점검이 가능하다. 현재 모바일절차서 시스템을 개발하면서 QR 코드 적용은 비상사위기 및 세안시설 점검 및 유해화학물질 취급시설 점검에 시범적으로 운영 중이다. 지금은 단순히 현장에서 설비의 QR 스캔 후 점검지의 해당항목으로 이동하여 점검결과를 기록하는 단순한 기능이지만 추후 QR 코드를 활용한 설비 점검의 기대효과는 다음과 같다. QR 코드를 통해 운전원은 현장에서 필요한 절차서 및 설비 정보를 바로 확인하여 설비 점검을 신속하고 정확하게 수행할 수 있다. 또한, 설비 문제시 즉각적인 대응이 가능하고, 설비의 점검결과를 데이터화하여 설비의 상태감시 및 예방정비 계획에 도움을 줄 수 있다.

처음으로 원자력발전소에 QR 코드를 활용한 정·주기 시험 모바일절차서 시스템은 정·주기 시험의 신속성과 정확성을 동시에 향상 시킬 수 있다. 나아가 QR 코드를 통해 각 설비 및 기기에 대한 실시간 정보를 제공하여, 발전 운전원이나 작업자에게 현장에서 필요한 정보를 신속하게 확인할 수 있다. 또한, 점검결과 데이터를 DB화하여 설비를 지속적으로 감시하고 효율적인 예방정비가 가능해진다. 원자력발전소에서 QR 코드 기반 설비 점검은 원자력발전소에 효율적인 운영 환경을 조성하는 데 기여할 수 있다고 판단되며, 향후 QR 코드를 적용한 운전, 정비 프로세스 등에 적용되기를 기대한다.

### Key words

Periodic Testing, QR code, mobile solution, nuclear power plant

### 참고문헌

- [1] 김남현, et al., “원전 무선망을 활용한 발전소 현장점검 적용방안,” 한국에너지학회 춘계학술대회, 2024.

\* 한국수력원자력 중앙연구원, kimmnamhyun@khnp.co.kr

# 칼만 필터 기반의 해군 함정 전투체계 표적 추적 성능 향상 연구

한민석\*, 장원석\*\*, 최영두\*\*\*, 홍순국\*\*\*

## A Study on Enhancing Target Tracking Performance in Naval Combat Systems Using Kalman Filter

Minseok Han\*, Wonseok Jang\*\*, Youngdoo Choi\*\*\*, and Soonkook Hong\*\*\*

본 연구는 해군 함정 전투체계의 표적 추적 성능 향상을 위해 칼만 필터 기반의 제어기를 설계하고 평가하였다. 시스템은 선형 및 비선형 운동 동력을 포함한 4차원 상태 공간으로 모델링되었다. 총 1000 스텝의 시뮬레이션을 통해 다양한 조건에서 추적 효율성을 평가하였다. 시뮬레이션 과정에서 실제 표적의 움직임은 비선형 진동 요소를 포함한 상태 전이 행렬을 사용하여 생성되었으며, 이는 현실적인 운영 환경을 시뮬레이션하였다. 생성된 측정값은 가우시안 노이즈로 오염되었으며, 측정 노이즈 공분산 행렬로 특성화되었다. 칼만 필터의 성능은 평균 제곱 오차(MSE)와 Root Mean Square Error (RMSE)와 같은 주요 지표를 사용하여 정량화되었다. 결과적으로 X 축에서 평균 MSE는 0.5, Y축에서는 0.7로 나타났으며, 최종 RMSE는 0.85로 확인되었다. 이는 필터가 표적의 위치를 추정하는 데 효과적임을 나타낸다. 추가 분석 결과, 추정된 X 및 Y 속도는 각각 0.1과 0.09로, 실제 표적의 속도와 밀접하게 일치하였다. 위치 추적 정확성은 실제 위치, 측정 데이터 및 추정 상태의 경로를 나타내는 그래프를 통해 시각적으로 확인되었으며, 높은 상관관계를 보여주었다. 추적 정확성 외에도 본 연구는 칼만 필터의 계산 효율성을 평가하였으며, 각 측정값을 실시간으로 처리하여 스텝당 약 0.02초의 처리 시간을 유지하였다. 이러한 신속한 처리 능력은 해군 전투체계와 같은 실시간 응용 분야에서 시의적절하고 정확한 추적이 필수적임을 강조한다. 연구 결과는 복잡한 환경에서 표적 추적을 개선하기 위한 칼만 필터 기술의 잠재력을 부각시킨다. 본 연구는 해군 응용 분야에서 칼만 필터의 효과를 검증할 뿐만 아니라, 동적이고 불확실한 조건에서 작동할 수 있는 더욱 발전된 추적 시스템 개발을 위한 기초 자료를 제공한다. 향후 연구에서는 머신러닝 기술을 통합하여 추적 정확성과 적응성을 더욱 향상시키는 방향으로 나아갈 예정이다.

Key words : Kalman Filter, Target Tracking, Naval Combat Systems, Performance Evaluation, Nonlinear Motion

### 참고문헌

- [1] S. Kim, J. Lee, J. Jeong, S. Hwang and M. Han, "Design of a Kalman Filter Application System Based on Linearity Judgment for Improving Low-earth Satellite Positioning Precision," Journal of the KNST, vol. 7, No. 2, pp. 124-129, June 2024.
- [2] H. Hwang, S. Lee, S. Hong and M. Han, "Accuracy Improvement of Mine Detection Drone with Extended Kalman Filter Application," Journal of the KNST, vol. 7, No. 2, pp. 262-268, June 2024.
- [3] D. Sung, S. Kim, Y. Kim, Y. Jung and M. Han, "Prevention of Collision with Other Vessels Using Camera Sensors with Kalman Filter," Journal of the KIIECT, vol. 17, No. 3, pp. 130-140, June 2024.

\* 해군사관학교 전자제어공학과, mshan1024@navy.ac.kr, \*\*한화시스템 해양시스템1팀, \*\*\*해군사관학교 기계시스템공학과, hsk753@gmail.com(교신저자).

※ 이 논문은 2023년 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임(KRIT-CT-23-030).

# 군 업무 목적의 제로 트러스터 무선랜 기술의 적용과 활용방안에 관한 연구

황재룡\*

## A Study on the Application of Zero Trusted Wireless LAN Technology for Military Services

Jaeryong Hwang\*

### 요약

무선랜 기술의 보편성과 편의성에 따라 군 업무에서도 무선랜의 사용하고자 하는 시도가 많아지고 있다. 그러나 무선랜을 구축할 경우 비인가자의 네트워크 무단 접근을 차단하는 보안 취약점 극복이 요구된다. 제로트러스터 보안 아키텍처는 군 내부의 사용자라고 할지라도 다중 인증과 다계층 검증을 통해 내부 자원에 접근할 수 있도록 허용함으로써 내부 보안 위협의 원천을 차단한다. 이에 본 논문에서는 제로트러스터 아키텍처의 적용 방안과 교육, 군수, 작전 분야에서 군 내부 활용 방안을 소개한다.

### Abstract

The ubiquity and convenience of wireless LAN technology has led to a growing trend toward its use in military services. However, deploying wireless LANs requires overcoming security vulnerabilities to prevent unauthorized access to the network. The zero-trusted security architecture blocks the source of internal security threats by allowing users, even those inside the military, to access internal resources through multi-factor authentication and multi-layered verification. In this paper, we introduce the application of the zero-trusted Wireless LAN architecture and its use within the military in the areas of education, logistics, and operations.

### Key words

wireless LAN, security, zero-trusted architecture, military services

### 1. 서론

지능 정보기술의 발전과 병력 자원의 감소는 최신 ICT 기술의 군 내부 활용을 가속화 하고 있다. 군 부대와 같은 중요 기관에서 무선 네트워크 사용

이 증가하면서 보안 위협 또한 함께 증가하고 있다. 전통적인 네트워크 보안 모델은 내부의 모든 장치와 사용자를 신뢰하는 구조로, 비인가 접근이나 네트워크 침투가 발생할 경우 큰 위험을 초래할 수 있다.

\*해군사관학교, jhwang@navy.ac.kr

무선 네트워크 특성상 외부 침입 시도가 용이하고, 통신 중간자 공격(Man-in-the-Middle), 데이터 스니핑 등의 보안 위협이 발생할 가능성이 높아, 더욱 강력한 보안 모델의 도입이 요구된다.

따라서, 본 논문에서는 제로트러스터 기반의 무선랜 구축시 적용해야할 보안 정책에 대하여 살펴보고, 군에서 활용 가능한 방안에 대해서 살펴본다.

## II. 제로 트러스터 무선랜 보안 정책

제로 트러스트 기반의 무선랜은 무선 네트워크 환경에서 누구도 신뢰하지 않고 항상 검증하는 제로 트러스트 원칙을 적용하여 보안을 강화하는 접근 방식이다. 전통적인 무선랜 보안 모델은 네트워크 접근 시 초기 인증을 거치면 내부 네트워크를 자유롭게 사용할 수 있는 방식이었으나, 제로 트러스트 무선랜은 모든 연결과 접근을 지속적으로 검증하여 비인가 장치나 사용자가 내부 자원에 접근하지 못하도록 제한한다. 이를 위해 다중 인증(MFA), 네트워크 세분화, 실시간 모니터링 등 다양한 보안 기술을 결합해 무선 네트워크를 보호한다.

제로 트러스트 무선랜 아키텍처는 특히 외부 침입이 용이한 무선 환경의 특성을 보완하며, 비인가 접근, 중간자 공격(MITM), 데이터 스니핑 등의 위협으로부터 네트워크와 데이터를 보호할 수 있다. 군사나 정부 기관 등 높은 보안이 요구되는 환경에서, 제로 트러스트 무선랜은 민감한 정보 보호를 위해 최적의 보안 솔루션으로 활용될 수 있으며, 내부 네트워크의 안전성을 높여 군사 작전의 보안을 강화하는 데 효과적이다.

## III. 군 업무 활용 방안

군에서 업무를 위해 제로 트러스터 무선랜을 활용할 경우 교육, 군수, 작전 분야에 활용이 가능하다.

첫째, 교육 분야에서는 보안이 강화된 교육환경을 제공하고, 다중 인증을 통한 접속 관리가 가능하다. 군사 교육과 훈련 과정에서는 민감한 정보와 자료가 다루어지기 때문에, 제로 트러스트 기반 무선랜을 통해 교육용 네트워크에 접근하는 모든 장치와 사용자를 인증하고 검증하여, 비인가자의 위협으

로부터 민감한 자료를 보호할 수 있다.

두 번째, 군수분야에서는 실시간 모니터링을 통한 물자를 보호할 수 있다. 물자 관리 시스템의 접근 기록을 실시간으로 모니터링하고, 비정상적인 접근이나 의심스러운 활동을 즉시 차단할 수 있다. 이를 통해 군수 물자의 이동 및 재고 관리 과정에서 발생할 수 있는 보안 위협을 줄이고, 외부 침입에 대응할 수 있는 유연성을 제공할 수 있다.

마지막으로 작전분야에서는 민감 정보를 보호하고, 최소 권한 접근 및 네트워크 세분화가 가능하다. 작전 중 민감한 정보에 접근하는 모든 장치와 사용자를 지속적으로 검증함으로써, 내부 및 외부의 비인가 접근을 차단할 수 있다. 이를 통해 통신 중간자 공격(MITM)과 같은 위협을 예방하고, 작전 정보의 기밀성을 확보할 수 있다.

그리고, 작전 네트워크에 대한 접근 권한을 세분화하고, 필요한 인원만 최소한의 권한으로 특정 데이터에 접근할 수 있도록 설정할 수 있다. 작전 지역에서 각 부대원에게 부여된 권한 범위 내에서만 무선 네트워크에 접근하도록 하여, 정보의 불필요한 확산을 막고, 잠재적 위협을 줄이게 된다.

## III. 결 론

본 논문에서는 제로 트러스트 기반의 무선랜의 보안 정책과 군에서의 활용 방안을 살펴보았다. 무선랜의 활용의 필요성이 증가함에 따라 보안성이 강화된 무선랜의 정책 적용이 요구된다.

## 참 고 문 헌

- [1] Y. Liu, Z. Su, H. Peng, Y. Xiang, W. Wang and R. Li, "Zero Trust-Based Mobile Network Security Architecture," in IEEE Wireless Communications, vol. 31, no. 2, pp. 82-88, April 2024.
- [2] B. G. Jung, Y. -S. Yoo, K. Kim, B. -S. Kim, H. Lee and H. Park, "ZTA-based Federated Policy Control Paradigm for Enterprise Wireless Network Infrastructure," APCC, 2022, pp. 1-5.

# 실시간 센서 데이터 전송을 위한 SENT 네트워크 하드웨어 설계

오경목\*, 이재학, 송진한, 김병수\*\*

## SENT Network Hardware Design for Real-Time Sensor Data Transfer

Kyeongmook Oh\*, Jeahack Lee, Jinhan Song and Byung-Soo Kim\*\*

본 논문에서는 실시간 센서 데이터 처리를 위한 차량용 SENT(Single Edge Nibble Transmission) 네트워크 프로토콜 하드웨어 IP를 설계한 결과를 제시한다. 차량용 네트워크 시스템은 자율주행, 차량 안전 시스템 및 전기 자동차 등과 같은 차량 전장 기술이 발전하고 데이터양이 방대해짐에 따라 더욱 중요한 요소로 자리 잡았다. 특히, 차량 내 센서와 ECU(Electronic Control Unit) 간의 신뢰성 있는 데이터 전송은 차량 안전성에 중대한 영향을 미치며, 단일 엣지 기반의 프로토콜로써 데이터 전송 시에 간섭을 최소화할 수 있는 단방향 유선 차량용 프로토콜인 SENT는 차량용 센서 데이터 전송을 위한 최적화된 표준이다. 설계된 SENT 프로토콜 하드웨어 IP는 SAE-2716 표준을 전부 만족하며, 4bit 및 6bit CRC 연산이나 다양한 fast 채널 모드와 slow 채널 모드, pause enable 모드 등 모든 기능 및 옵션을 지원한다. 또한, synchronous 및 asynchronous 모드로 동작이 가능하도록 설계되었으며, fast 채널 데이터는 센서 데이터 전송을 주로 진행하고, slow 채널 데이터는 해당 fast 채널의 데이터에 오류가 발생했을 시에 이를 MCU로 알려줄 수 있는 역할을 할 수 있도록 설계되었다. 본 논문에서 제안하는 IP는 실시간으로 차량의 센서 데이터를 안전하게 내부의 ECU로 전달하기 위한 SENT 하드웨어 시스템으로 최적화되어 있다. CPU 코어와 AXI4 AMBA BUS 기반의 SoC(System on Chip) 구조로 구현되어 있으며, 제안하는 SENT 하드웨어 IP는 SENT 통신의 전송 속도를 고려하여 APB 인터페이스로 설계되었다. 제안하는 SENT 프로토콜 하드웨어 IP는 FPGA 보드와 오실로스코프를 이용하여 SENT 검증용 SW를 통한 검증을 진행하였으며, SAE-2716 표준의 지원하는 모든 기능을 검증하였으며, 최대 100MHz clock을 통해 SENT의 tick을 제어함으로써 데이터 및 CRC 연산의 유효성을 확인하였다. 또한, RTL 검증을 위한 SENT 디코더 모듈을 설계함으로써 다양한 시스템 환경에서 RTL 시뮬레이션을 통해 제안하는 SENT 하드웨어 IP를 검증하였으며, 추가적인 차량용 네트워크 하드웨어 구조 연구에 활용할 수 있는 기반을 마련하였다.

### Key words

Automotive network, SENT, Automotive sensor data processing, Real-time communication

### 참고문헌

- [1] SAE J2716 APR2016, "SENT - Single Edge Nibble Transmission for Automotive Applications," [https://www.sae.org/standards/content/j2716\\_201604/](https://www.sae.org/standards/content/j2716_201604/)

---

\*한국전자기술연구원, okm806@keti.re.kr \*\*한국전자기술연구원, bskim4k@keti.re.kr (교신저자)

※ 이 연구는 2024년도 산업통상자원부 및 산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임 ("20015705, 수소자동차 연료전지 Stack을 위한 300kpa 이하급 경로 환경용 센서 ECU 및 모듈 상용화 기술 개발")

## 무전원 테라헤르츠 신호검출기 설계

\*김영호, 박동우, 이의수

### A Design of Terahertz Signal Detector with Zero-Power Consumption

Youngho Kim, Dongwoo Park and Euisu Lee

본 논문에서는 전원 공급이 필요없는 InGaAs 쇼키베리어 다이오드(Schottky Barrier Diode, SBD)를 신호 검출의 핵심소자로 사용하여, 테라헤르츠(THz) 대역에서 신호 전파 강도를 정밀하게 측정할 수 있는 신호검출기 설계 방법을 다룬다. 본 설계는 기존 설계의 구조적 한계점을 살펴보고 이에 대한 해결 방법을 제시함으로써 향상된 신호 검출 성능을 제공하는 설계 방법을 소개한다.

보통 고주파 동작성능, 감도(Responsivity) 및 잡음특성 등을 고려하여 쇼키베리어 다이오드(SBD)를 사용한 신호검출기가 1THz 이하 주파수 대역에서 가장 선호되어 왔다[1,2]. 이때 신호를 효과적으로 검출하기 위해 필터의 높은 임피던스를 기반으로 하는 가상 접지점(Virtual Ground point)을 설계하여 쇼키베리어 다이오드(SBD)의 기준 동작점으로 삼는 방식이 테라헤르츠 신호검출기 모듈에 적용되어 왔다. 그러나 이 방법은 실제로 필터가 해당 동작 주파수에서 이론적으로 무한 임피던스를 갖지 못하기 때문에 검출 다이오드(SBD)에 효율적인 기준 동작점을 제공하지 못하는 단점이 있다. 또한 해당 구조는 가상 접지점을 설계하기 위하여 여러 번 개별블록으로 나누어 매칭 네트워크 설계와 병행하여 수행해야 하기 때문에 까다롭고, 많은 설계 시간과 설계 오차를 수반한다. 그래서 이 단점을 보완하여 좀 더 정확한 기준 동작점을 쇼키베리어 다이오드(SBD)에 제공하는 기술을 연구하게 되었고, 실제 접지(0V) 값에 기반한 가상 접지 방식이 본 논문에 제안되었다. 제안된 설계구조는 기존의 가상 접지점 설계 방식보다 설계 오차가 작으며 아울러, 성능 결과를 안정적으로 예측 및 유지할 수 있게 된다. 또한 신호 감도와 검출 효율은 극대화시키면서 회로구조는 단순해지게 만들어 신호의 유전체 손실을 최소화하는 성과를 보여준다.

#### Key words

terahertz signal detector, zero-bias/zero-power detector, InGaAs, Schottky Barrier Diode, terahertz rectifier, virtual ground

#### 참고문헌

- [1] J. L. Hesler and T. W. Crowe, "NEP and responsivity of THz zero-bias Schottky diode detectors," 2007 Joint 32nd International Conference on Infrared and Millimeter Waves and the 15th International Conference on Terahertz Electronics, Cardiff, UK, 2007, pp. 844-845, doi: 10.1109/ICIMW.2007.4516758.
- [2] A. Semenov, O. Cojocari, H. -W. Hubers, F. Song, A. Klushin and A. -S. Muller, "Application of Zero-Bias Quasi-Optical Schottky-Diode Detectors for Monitoring Short-Pulse and Weak Terahertz Radiation," in IEEE Electron Device Letters, vol. 31, no. 7, pp. 674-676, July 2010, doi: 10.1109/LED.2010.2048192.

---

\*한국전자통신연구원, youngho, dwpark, euisu@etri.re.kr

※ 본 연구 논문은 한국전자통신연구원 연구운영지원사업의 일환으로 수행되었음.  
[24ZH1320, 광전소자 기반 THz 근접장 센싱 및 필드 시각화 원천기술 개발]

# 디지털 라이프로그 데이터 기반 노쇠 증상 조기 감지 체계 프로토콜 개발

김연식\*, 염민선\*\*

## Development of a protocol for early detection of frailty symptoms from digital lifelog data

Kim Yeon Sik\* and Yeom Min Sun\*\*

본 연구는 스마트워치를 활용하여 수집된 라이프로그 데이터를 활용하여 노쇠(Frailty) 증상을 조기에 감지하는 체계를 개발하고자 한다. 노쇠는 신체의 내외부에서 발생하는 스트레스에 대항하는 생리적 여력이 줄어들었음을 의미하는 말로, 특히 고령 인구에서 신체적·정신적 기능 저하를 동반하며, 조기 발견 및 개입이 필수적이다. 그러나 기존 진단 방법은 정기적인 병원 방문과 같은 제한적인 접근을 요구한다. 본 연구에서는 웨어러블 기기와 스마트폰을 통해 수집된 라이프로그 데이터(활동량, 수면 패턴, 심박수 등)를 기반으로 노쇠의 징후를 자동으로 탐지할 수 있는 프로토콜을 개발하고자 한다. 이를 위해 머신러닝 알고리즘을 적용하여 데이터를 분석하고, 개인 맞춤형 조기 감지 시스템을 개발하고자 한다. 본 연구에서는 65세 이상의 고령자를 대상으로 하여, 서울아산병원, 연세 세브란스병원과 (주)디파이에서 임상적으로 진단된 노쇠 지표를 기반으로, 스마트워치를 통해 수집된 생체정보(심박수, 혈압, 혈중산소포화도, 체온 등)와 일상 활동 데이터(걷기, 수면 등)를 포함한 디지털 라이프로그 데이터를 비교 분석하였다. 이 과정을 통해 디지털 노쇠 지표를 추출하고, 이를 바탕으로 노쇠 상태에 대한 조기 감지와 대응이 가능한 프로토콜을 검증하고자 한다.

### Key words

frailty, healthcare, digital, intervention, lifelog

### 참고문헌

- [1] World Health Organization. Guidelines on Integrated Care for Older People (ICOPE) [Internet]. Geneva, Switzerland: World Health Organization; 2017 [cited 2021 Mar 25]. Available from: <https://www.who.int/ageing/publications/guidelines-icope/en/>.
- [2] Rose M Collard 1, Han Boter, Robert A Schoevers, Richard C Oude Voshaar, Prevalence of frailty in community-dwelling older persons: a systematic review, J Am Geriatr Soc. 2012 Aug;60(8):1487-92.
- [3] 원장원, 김희선, 김미지, 김문정, 김선영, 김성배, 김성욱, 노지훈, 백재욱, 신형은, 양지수, 오인환, 이대현, 전승연, 정은진, 정현경, 최지혜, 하은미, 일차의료에서 노쇠(frailty)의 의료적 비용 산출 및 통합 노쇠관리 의 비용효과분석, 국립보건연구원, 2021
- [4] Chang Won Won, Evaluation and management of frailty, J Korean Med Assoc. 2017 Apr;60(4):314-320.
- [5] Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and

\* 나무아이씨티 연구소, ys.kim@namuict.co.kr, \*\* 나무아이씨티 연구소, ms.yeom@namuict.co.kr(교신저자)

※ 본 연구는 질병관리청 국립보건연구원 헬스케어이종데이터활용체계및인공지능개발(R&D) 사업의 지원을 받아 수행된 연구임(1776000529)

- comorbidity: implications for improved targeting and care. *J Gerontol A Biol Sci Med Sci* 2004;59:255-263.
- [6] Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, McBurnie MA. Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001;56:M146-M156.
- [7] Abellan van Kan G, Rolland YM, Morley JE, Vellas B. Frailty: toward a clinical definition. *J Am Med Dir Assoc* 2008;9:71-72.
- [8] Jung HW, Yoo HJ, Park SY, Kim SW, Choi JY, Yoon SJ, Kim CH, Kim KI. The Korean version of the FRAIL scale: clinical feasibility and validity of assessing the frailty status of Korean elderly. *Korean J Intern Med* 2016;31:594-600.



# 고속 Chip 간 디지털 통신을 위한 인터페이스 회로 개발

이재학, 오경묵, 김병수

## Development of Interface Circuits for High-Speed Digital Communication

Jeahack Lee\*, Kyeongmook Oh\*, and Byung-Soo Kim\*\*

본 논문에서는 반도체 칩 간의 고속 디지털 통신을 위한 인터페이스 프로토콜 기술을 제안한다. 인공지능 등 다양한 응용에서 대량의 데이터를 저장하고 읽기 위해 DDR(Double Data Rate)[1, 2] 표준 및 PCIe(Peripheral Component Interconnect Express)[3], MIPI(Mobile Industry Processor Interface), SerDes(Serializer/Deserializer)와 같은 고속 인터페이스가 필수적으로 사용된다. 특히 DDR4, DDR5와 같은 고속 메모리 인터페이스는 높은 대역폭을 제공하지만, 고가의 IP 비용과 복잡한 설계로 인해 ASIC 설계 시 부담이 크다. 연구기관과 대학에서 IP 검증은 목적으로 고액의 메모리 인터페이스를 활용하는 것은 비용 효율적인 면에서 적합하지 않다. 기존의 다양한 고속 인터페이스 표준들은 또한 직렬/병렬 변환기 등 다양한 아날로그 및 제어 IP의 도입이 요구된다. 본 논문에서는 fully-digitalized 칩 간 통신 인터페이스를 제안하여 고속 데이터 통신이 필요한 반도체 핵심 IP 연구에 기여하고자 한다. 제안하는 칩 간 통신 인터페이스의 주요 특징은 다음과 같다. 첫째, PAD 수가 제한된 ASIC 설계 환경을 위해 데이터 라인을 입출력 포트로 설계하여 PAD 수를 최소화하였으며, 기존 입출력 구분이 있는 디지털 통신과 비교하여 약 50%의 PAD 수가 감소한다. 둘째, 연속 전송 길이 및 strobe 신호를 동일한 control 신호로 사용하여 PAD 수를 더욱 줄이고 대용량 데이터 전송의 효율성을 높였다. 최근 인공지능 관련 응용에서 연속적인 대용량 데이터 송수신을 지원하기 위해 burst 통신이 필수적이다. Burst 통신은 동일 방향으로의 연속적인 데이터 송수신으로 인해 output enable 신호 변화가 줄어들어, 입출력 data 포트를 활용한 본 프로토콜에 적합하다. Strobe 신호는 단일 전송 시 연속 통신 길이를 표현하는 포트와 함께 사용되며, Transfer 신호는 사용하여 현재 통신의 상태를 Idle, Busy, Burst, Single 상태로 통신 상태를 나타낸다. Idle은 어떠한 통신도 없으며, busy는 송신 단에서 다른 작업으로 통신을 못하는 상태, burst는 burst 통신 상태, single은 단일 통신 상태를 의미한다. 본 디지털 통신 프로토콜을 통해 추가적인 외부 통신 IP 없이 개별 연구 IP의 기능 검증을 가능하게 하였으며, FPGA 플랫폼에서 회로를 설계 및 검증하여 data 라인을 32, 64, 128, 256 비트로 검증 완료하였다.

### Key words

chip to chip interface, fully-digitalized interface, FPGA platform

### 참고문헌

- [1] JEDEC Solid State Technology Association. "DDR4 SDRAM Standard," JESD79-4, 2012.
- [2] JEDEC Solid State Technology Association. "DDR5 SDRAM Standard," JESD79-5, 2020.
- [3] PCI-SIG. "PCI Express Base Specification," Rev. 5.0, 2019.

\* 한국전자기술연구원, jhk507@keti.re.kr, \*한국전자기술연구원, okm806@keti.re.kr,

\*\* 한국전자기술연구원, bskim4k@keti.re.kr (교신저자)

※ 이 논문은 2022년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임(RS-2022-00155731, 융합 메모리와 고속 연산 엔진 제어를 위한 하드웨어/소프트웨어 융합 메모리 솔루션 기술 개발)

# 뇌졸중 심전도(ECG) 분류 모델별 성능 평가

임소윤<sup>†\*</sup>, 김철호<sup>\*\*</sup>, 김유섭<sup>\*</sup>

## Evaluate performance by stroke electrocardiogram (ECG) classification model

So-Yun Im<sup>\*</sup>, Chul-Ho Kim<sup>\*\*</sup>, and Yu-Seop Kim<sup>\*</sup>

### 요 약

뇌졸중은 고령자의 주요 사망 및 장애 원인 중 하나로, 조기 진단이 중요하다. 이에 따라 심전도(ECG) 신호를 활용한 뇌졸중 진단 연구가 활발히 진행되고 있다. 본 연구는 뇌졸중 환자의 ECG 분류 성능을 1D CNN, ResNet, DenseNet, LSTM, BiLSTM 모델을 통해 평가하였으며, 원본 데이터와 R-피크로 분할된 데이터 조건에서 실험을 진행하였다. 실험 결과, 원본 데이터에서는 DenseNet이 99.52%, 분할 데이터에서는 1D CNN이 98.30%로 가장 높은 정확도를 기록하였다. 반면, LSTM과 BiLSTM 모델은 낮은 성능을 보였다. 이는 CNN 기반 모델이 ECG의 국소적 특징을 효과적으로 포착함을 시사한다.

### Abstract

Stroke is one of the leading causes of death and disability among the elderly, making early diagnosis crucial. Accordingly, research on stroke diagnosis utilizing electrocardiogram (ECG) signals is actively underway. This study evaluated the ECG classification performance of stroke patients using 1D CNN, ResNet, DenseNet, LSTM, and BiLSTM models, conducting experiments under conditions of original data and data segmented by R-peaks. The experimental results showed that DenseNet achieved the highest accuracy of 99.52% with the original data, while 1D CNN recorded an accuracy of 98.30% with the segmented data. In contrast, the LSTM and BiLSTM models exhibited lower performance. This suggests that CNN-based models effectively capture the local features of ECG signals.

### Key words

ECG classification, stroke detection, 1D CNN, ResNet, DenseNet, LSTM, BiLSTM

## 1. 서 론

뇌졸중은 전 세계적으로 고령자 사망과 장애의

주요 원인 중 하나로 꼽힌다. 따라서 이 질병의 조기 진단은 매우 중요하다[1]. 뇌졸중 환자에게서 비정상적인 심전도(ECG) 패턴을 발견한 사례[2]가 보

\* 한림대학교 정보과학대학 융합소프트웨어학과, soyun012399@gmail.com<sup>†</sup>, yskim01@hallym.ac.kr,

\*\* 한림대학교 춘천성심병원 신경과, gumdol52@hallym.or.kr

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원[RS-2021-II212068, 인공지능 혁신 허브 연구 개발 (서울대학교, AI연구원)]과 2022년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2022R1A5A8019303).

고됨에 따라 심전도 신호를 뇌졸중 진단에 사용하는 시도가 활발히 이루어지고 있다. 다양한 방법으로 뇌졸중 환자의 ECG를 분류하는 연구 중에서도 딥러닝 모델을 사용하여 긍정적인 결과를 얻은 여러 사례를 찾아볼 수 있다. 한 연구에서는 1D Convolutional Neural Network(1D-CNN) 모델을 사용하여 약 90%의 정확도를 기록했으며, Gradient-weighted Class Activation Mapping (GRAD-CAM)을 통해 ECG 패턴을 분석하여 모델의 해석 가능성을 높였다[3]. 또 다른 접근법으로는 Densely Connected Convolutional Networks(DenseNet) 기반의 심층 신경망 모델을 활용하여 약 85.82%의 정확도를 달성한 경우가 있으며[4], Long Short-Term Memory(LSTM) 모델을 사용한 연구에서는 뇌졸중 환자의 ECG 데이터를 분석하여 93.78%의 정확도를 기록했다고 보고했다[5]. 본 연구에서는 뇌졸중에 대한 다양한 모델의 분류 성능이 연구진이 수집한 실제 뇌졸중 환자의 ECG 데이터에 적용하였을 때 어떠한 결과를 나타내는지 알아보려고 했다.

표 1. 실험 데이터 세트  
Table 1. Experimental data sets

원본 데이터				
	train	val	test	total
normal	480	103	104	687
stroke	480	103	104	687
분할 데이터				
	train	val	test	total
normal	4,500	1,000	1,000	6,500
stroke	4,500	1,000	1,000	6,500

## II. ECG 데이터 셋

ECG 분류는 연구마다 서로 다른 길이의 데이터 조건에서 실험되었다. 따라서 본 연구에서는 데이터 길이에 따른 모델별 성능 차이를 확인하기 위해 데이터 조건을 두 가지로 나누어 실험했다. 첫 번째는 원본 데이터 자체를 분류하였고 두 번째는 원본 데이터를 R-피크 기준으로 분할 하여 개별 파형으로 세분화하여 분류했다. R-피크는 심전도 신호에서 가장 두드러진 특징으로, 심실이 수축할 때 심장의 전기 활동을 나타내는 지점이다[6]. 원본 데이터는 모두 5000의 타임스텝을 가지고 있으며 분할 데이터

는 모두 256의 타임스텝으로 리샘플링하여 사용했다. 또한 모든 데이터에는 Min-Max 정규화를 적용했으며 동일 환자의 심전도가 학습과 검증/테스트에 함께 사용되지 않도록 데이터를 환자별로 분리했다. 학습에 사용된 데이터 셋 구성은 (표 1)에 제시되어 있다.

원본 심전도 데이터는 정상 심전도와 뇌졸중 환자의 심전도로 구분되며 각각 다음과 같은 경로로 수집했다. 정상 심전도 데이터는 공개된 데이터 셋 PTB-XL[7]에서 정상 심전도를 의미하는 NORM 클래스 데이터를 추출하여 사용했다. 뇌졸중 환자의 심전도 데이터는 한림대학교 춘천성심병원 (<https://chuncheon.hallym.or.kr/>)에서 제공받은 비공개 데이터를 사용했다. 개별 파형 심전도는 원본 심전도에 적응형 윈도우 분할 방법[8]을 적용하여 수집했다. 각 심박의 R-피크를 중심으로 이전 및 다음 R-피크 간의 거리의 75%를 사용하여 경계를 설정했다. 심전도 신호에서 피크 간 간격이 일정하지 않은 경우에도 유연하게 대응할 수 있도록 설계된 방법으로서 불규칙한 심전도의 불안정한 피크 탐지를 보완하는 방법으로 사용된다. (그림 1)은 이 두 종류의 데이터를 시각화한 것이다.

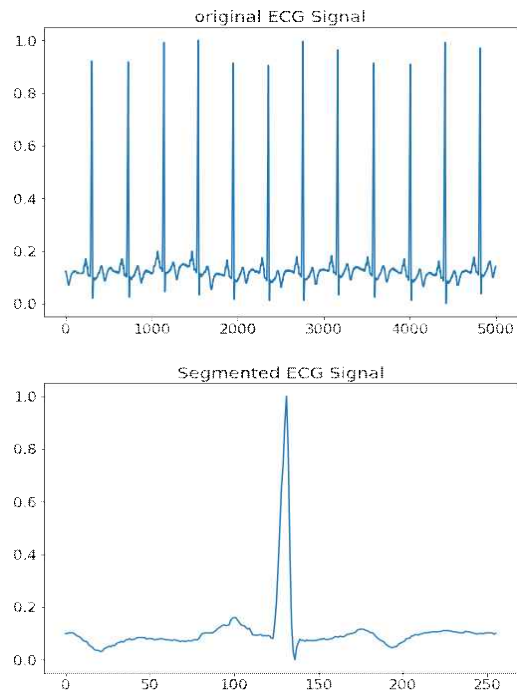


그림 1.. ECG 데이터 시각화  
Fig. 1. ECG data visualization

### III. 실험

실험은 정상 심전도와 뇌졸중 환자의 심전도 분류를 목적으로 하며 12-lead ECG 데이터에서 한 개의 리드만을 사용하여 1D 입력을 통한 단일 리드 분류를 수행했다. 뇌졸중과 관련된 가장 흔한 심전도 변화는 심장의 하벽 리드(II, III, aVF)에서 감지된다[2]. 따라서 단일 리드 분류를 수행하기 위해 세 개의 리드를 시각화하여 비교한 후, 피크 구별이 비교적 뚜렷하게 보인 II-lead를 선택해 사용했다.

실험에 사용된 모델은 1D CNN, ResNet, DenseNet, LSTM, BiLSTM이다. 다양한 모델의 분류 성능을 비교 분석하기 위해 서론에서 언급한 연구에 사용된 모델 외에 ResNet과 BiLSTM 모델을 추가로 실험하여 비교했다. 1D CNN[9]은 입력 신호의 국소적 특징을 잡아내는데 효과적이며[10] 이미지 처리에 중점을 두는 2D CNN과 비교해 신호 데이터 분석에 주로 사용된다.

Residual Network (ResNet)[11]과 DenseNet[12]은 모두 CNN을 기반으로 한 모델로서, 각각 잔차 연결과 층 간 연결 개념을 도입하여 더 깊고 효율적인 학습을 가능하게 한다. 이 모델들 역시 본 연구에서는 모두 1D CNN을 기반으로 설계되었다. LSTM[13]은 긴 시퀀스의 장기 의존성을 처리하는데 적합하며, Bidirectional Long Short-Term Memory(BiLSTM)[14]는 양방향에서 정보를 처리해 더 정교한 예측을 가능하게 한다. 각 모델은 가장 좋은 성능을 기록한 하이퍼파라미터와 모델 구조가 개별적으로 적용되었다. 성능 평가는 정확도와 F1

점수를 사용하여 평가했다.

### IV. 결과

원본 데이터에 대한 모델의 분류 성능을 확인한 결과 DenseNet 모델의 정확도가 99.52%로 다른 모델들과 비교하여 압도적인 분류 성능을 보였다. 그 다음은 ResNet 모델로 93.75%의 분류 정확도를 보였다. 반면 분할 데이터의 경우 1D CNN이 98.30%로 가장 높은 정확도를, DenseNet이 96.70%로 두번째로 높은 정확도를 보였다. 이를 통해 원본 데이터와 분할 데이터 간에 모델별 분명한 성능 차이가 존재한다는 것을 확인했다. 또한 본 연구에서 CNN 기반 모델들이 기존 연구보다 높은 성능을 기록하였는데 그 이유는 두 가지로 추측할 수 있다. 첫째, 적용된 모델 구조와 학습 데이터 조건의 차이가 성능 향상에 기여했을 가능성이 있다. 둘째, 뇌졸중과 구별되는 레이블로 정상 ECG로 진단된 안정적인 패턴의 데이터를 사용한 것이 분류 성능을 높였을 가능성이 있다. LSTM과 BiLSTM 모델의 경우 정확도가 현저히 낮았다. 이는 실험에 사용한 ECG 데이터가 상대적으로 짧은 구간에서 중요한 특징을 가지기 때문에 장기 의존성을 기반으로 한 학습 방식이 해당 특징을 포착하는 데 어려움을 겪은 것으로 해석할 수 있다. 전체 결과는 (표 2)에서 확인할 수 있으며 최고성능은 밑줄로 표시하였다. 이러한 결과는 ECG 데이터의 특성에 따라 적합한 모델이 다를 수 있으며, 연구진이 수집한 데이터의 경우에는 CNN 기반 모델이 더 적합하다는 것을 보여준다.

표 2. 모델별 분류 성능

Table 2. Classification performance by model

모델명	원본 데이터 분류 성능		분할 데이터 분류 성능	
	Accuracy	F1	Accuracy	F1
1D CNN[3]	87.50%	87.50%	98.30%	98.30%
ResNet	93.75%	93.74%	91.75%	91.74%
DenseNet[4]	99.52%	99.52%	96.70%	96.70%
LSTM[5]	62.98%	62.94%	58.10%	58.09%
BiLSTM	68.27%	67.68 %	58.65%	58.60%

## V. 결 론

본 연구에서는 뇌졸중 환자의 심전도 신호를 분류하는 데 있어 CNN 기반 모델이 적합함을 확인했다. 또한 원본 데이터와 분할 데이터로 나누어 성능을 평가한 결과 ECG 데이터 특성에 따라 모델별 성능이 차이를 보이는 것을 확인했다.

향후 연구에서는 환자의 임상 정보나 MRI와 같은 영상 데이터를 추가적으로 활용하여 연구를 진행할 계획이다. 이를 통해 진단의 신뢰도를 높이고 환자 맞춤형 뇌졸중 예측 모델을 개발하는 데 기여하고자 한다.

## 참 고 문 헌

- [1] P. Kunwar and P. Choudhary, "A stacked ensemble model for automatic stroke prediction using only raw electrocardiogram", *Intelligent Systems with Applications*, vol. 17, 200165, 2023.
- [2] X. Du and Y. Zhang, "Transient J wave-like ST-segment elevation in intracerebral hemorrhage: a case report", *BMC Cardiovascular Disorders*, vol. 22, no. 255, pp. 1-5, 2022.
- [3] E. S. Ho and Z. Ding, "Electrocardiogram analysis of post-stroke elderly people using one-dimensional convolutional neural network model with gradient-weighted class activation mapping", *Artificial Intelligence in Medicine*, vol. 130, 102342, 2022.
- [4] Y. Xie, H. Yang, X. Yuan, Q. He, R. Zhang, Q. Zhu, Z. Chu, C. Yang, P. Qin, and C. Yan, "Stroke prediction from electrocardiograms by deep neural network", *Multimedia Tools and Applications*, vol. 80, pp. 17291-17297, 2021.
- [5] M. A. Kumar, M. Kumar, and G. P. M. S, "Stroke disease prediction based on ECG signals using deep learning techniques", in *2022 1st International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*, Dehradun, India, pp. 1-5, 2022.
- [6] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm", *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230-236, Mar. 1985.
- [7] P. Wagner, N. Strodthoff, R.-D. Bousseljot, W. Samek, and T. Schaeffter, "PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3)", *PhysioNet*, 2022.
- [8] E. Adib, F. Afghah, and J. J. Prevost, "Arrhythmia classification using CGAN-augmented ECG signals", in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Las Vegas, NV, USA, pp. 1-6, 2022.
- [9] F. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller, "1D Convolutional Neural Networks for Time Series Classification", *Data Mining and Knowledge Discovery*, vol. 33, no. 4, pp. 917-963, 2019.
- [10] C.-H. Hsieh, Y.-S. Li, B.-J. Hwang, and C.-H. Hsiao, "Detection of atrial fibrillation using 1D convolutional neural network", *Sensors*, vol. 20, no. 7, p. 2136, Apr. 2020.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.
- [12] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700-4708, 2017.
- [13] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [14] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks", *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, 1997.

# 딥러닝(LSTM)과 통계적 기법(SARIMA)을 이용한 독거노인 생활반응 데이터의 예측 비교 분석 연구

김지윤, 강성범, 손언호, 최기석, 오세미  
(주)범일정보 기술연구소

## Predicting Behavioral Patterns of Elderly Living Alone Using Deep Learning (LSTM) and Statistical Methods (SARIMA): A Comparative Study

JiYun Kim, SungBum Kang, EonHo Son, KiSeok Choi, and SeMi Oh  
Bumil Information Co., Ltd Technology Research Institute, Daegu, Rep. of Korea

### 요 약

본 연구는 독거노인 가구의 생활반응 데이터를 활용하여 차월 데이터를 예측하는 두 가지 접근 방식인 LSTM(장단기 메모리) 예측 모델과 SARIMA(계절성 자기회귀 누적 이동평균) 통계 알고리즘의 예측 정확도를 분석하는 것을 목표로 한다. 두 가지 방법을 적용해 예측한 데이터와 실제 데이터를 비교하여 두 모델의 정확도를 평가하고 대조한다. 이 연구는 어떤 접근 방식이 더 정확한 예측을 제공하는지 식별하고, 딥러닝 모델을 전통적인 통계 방법에 비해 사용하는 것의 잠재적 이점을 탐구하고자 한다.

### Abstract

This research aims to analyze the prediction accuracy of next-month data using two different approaches: LSTM (Long Short-Term Memory) forecasting and SARIMA (Seasonal Auto-Regressive Integrated Moving Average) statistical algorithms. The study focuses on living response data collected from single-person elderly households. By applying both methods, the research compares the predicted data with actual data to evaluate and contrast the accuracy of the two models. The study seeks to identify which approach delivers more precise predictions and explore the potential benefits of using deep learning models over traditional statistical methods for forecasting future data trends in this demographic.

### Key words

Long Short-Term Memory, SARIMA, single-person elderly households, living response data

### 1. 서 론

현대 사회의 고령화는 독거노인의 증가를 초래하고 있으며, 이에 따른 사회적 문제가 대두되고 있

다. 특히, 독거노인 가구의 생활패턴과 건강 상태를 모니터링하고 예측하는 것은 이들의 복지를 향상시키고 적절한 지원을 제공하기 위해 중요한 과제로 대두되고 있다. 최근 들어 사물인터넷(IoT) 기술과

빅데이터 분석 기법의 발전으로 생활반응 데이터를 수집하고 분석하는 것이 가능해지면서, 이를 기반으로 독거노인 가구의 생활 패턴을 예측하는 연구가 활발히 이루어지고 있다.

본 연구에서는 독거노인 가구의 생활반응 데이터를 활용하여 차월 데이터를 예측하기 위한 두 가지 접근 방식을 활용했다. 첫 번째는 딥러닝 기반의 LSTM 모델을 사용하는 방법이며, 두 번째는 전통적인 통계 방법인 SARIMA 알고리즘을 사용하는 방법이다. 독거노인의 생활반응 데이터를 분석하여 그들의 생활 패턴을 이해하고, 이를 통해 필요한 서비스 제공의 시기를 조절하는 데 활용할 수 있는 방법을 제시한다. 또한, 각 모델의 예측된 데이터와 실제 데이터 간의 비교를 통해 모델의 정확성을 비교하였다.

## II. 배경 지식

### 2.1 독거노인 생활반응 데이터의 중요성

독거노인 가구의 생활 데이터를 활용한 연구는 노인의 안전과 돌봄 서비스의 효과성을 높이기 위해 꾸준히 진행되고 있다. 기존의 연구에서는 데이터 분석을 통해 특정 위험 패턴을 탐지하거나 긴급 상황을 예측하는 데 중점을 두었다.

### 2.2 LSTM을 활용한 시계열 예측

LSTM은 장기 의존성을 학습할 수 있는 능력 덕분에 시계열 데이터 분석에 적합한 모델로 평가받고 있다. 특히, LSTM은 데이터의 흐름에서 중요한 정보를 효율적으로 기억하고, 이를 바탕으로 미래 데이터를 예측하는 데 강점을 가지고 있다.

### 2.3 SARIMA를 활용한 계절성 데이터 예측

SARIMA(Seasonal ARIMA)는 시계열 데이터에서 계절성을 포함한 변화를 모델링할 수 있는 알고리즘으로, 시계열 데이터의 계절적 패턴과 경향을 반영하여 예측 정확성을 높이는 데 유용하다. 특히, 계절적 주기를 가진 생활반응 데이터에서 SARIMA 모델은 단기 예측과 계절적 변동을 설명하는 데 효

과적이며, ARIMA 모델보다 계절적 특성이 뚜렷한 데이터에 적합하다.

## III. 연구 방법

### 3.1 데이터 수집 및 전처리

독거노인 가구에서 수집된 수도 데이터(6시간 주기의 2개월치 데이터)를 사용하여 분석을 진행하였다. 데이터는 특정 시간대의 반응 패턴을 포함하고 있으며, 이상치 제거 및 정규화를 통해 모델 학습에 적합한 형태로 전처리하였다.

### 3.2 예측모델 설계

시계열 예측의 성능을 극대화하기 위해 LSTM과 GRU 모델을 함께 사용하는 앙상블 모델을 설계하였다. 모델의 학습을 위해 데이터셋을 학습용 데이터와 테스트용 데이터로 분리하였으며, 예측 결과의 정확도를 높이기 위한 다양한 하이퍼파라미터 튜닝을 진행하였다. 이를 통해 LSTM과 GRU의 장점을 결합하여 예측 성능을 더욱 향상시킬 수 있었다. `early_stopping`을 활용해 과적합을 방지해 정확도와 속도를 높였다.

표 1. 예측모델의 하이퍼파라미터

파라미터	값
LSTM 활성화함수	relu
LSTM 뉴런 수	128
GRU 활성화함수	relu
GRU 뉴런 수	80
epochs	1000
batch size	32

### 3.3 통계 알고리즘 설계

통계 알고리즘인 SARIMA를 활용하여 시계열 데이터를 예측하고자 하였다. SARIMA에서 계절 주기는 4로 설정하였고, 데이터의 특성에 맞게 각 가구별로 최적의 예측 성능을 확보하기 위해 `auto_arima`를 이용하여 최적의  $(p, d, q)$ 값을 자동으로 탐색하였다. 이러한 방법을 통해 각 가구별 시계열 데이터의 특성을 반영하고, SARIMA 알고리즘의 예측 정

확도를 최대화하고자 하였다.

표 2. SARIMA 파라미터

파라미터	값
자기회귀 차수(p)	4
미분 차수(d)	0
이동 평균 차수(q)	2
계절 주기(s)	4

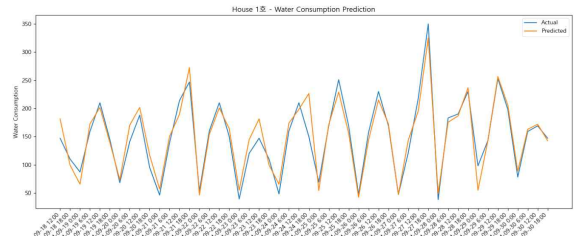


그림 2. LSTM 예측치

### 3.4 평가 지표

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

그림 1. 평가지표

예측된 데이터와 실제 데이터 간의 오차를 평가하기 위해 평균 절대 오차(MAE), 평균 제곱 오차(MSE), 그리고 평균 제곱근 오차(RMSE)를 사용하여 모델의 예측 정확성을 분석하고, 예측의 신뢰도를 검토하였다. 또한, 평균 절대 퍼센트 오차(MAPE)를 통해 예측 정확도를 퍼센트로 나타내어 모델의 예측 정확성과 신뢰도를 더욱 쉽게 해석할 수 있도록 하였다.

### IV. 연구 결과

생활반응 데이터 2개월치를 학습하여 다음 달의 데이터를 예측한 후 실제 데이터와 비교한 결과다.

표 3. 모델간 정확도 비교

	LSTM	SARIMA
MAE	12.21	33.14
MSE	259.39	2590.97
RMSE	16.11	50.90
MAPE	11.95%	28.81%
정확도	88.05%	71.19%

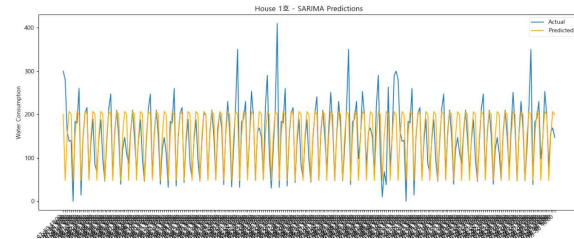


그림 3. SARIMA 예측치

결과를 보면 LSTM 모델은 SARIMA 알고리즘에 비해 독거노인의 생활반응 데이터를 기반으로 차월 데이터를 예측하는 데 있어 높은 정확도(16.86%)를 보였다.

### V. 결론

독거노인의 생활반응 데이터를 이용해 LSTM 모델과 SARIMA 알고리즘을 통해 다음 달 데이터를 예측하고, 두 모델의 예측 정확도를 비교했다. 분석 결과, LSTM 모델이 실제 데이터와의 오차가 더 작아, 차월 데이터 예측에 있어 더 높은 정확성을 보였다. 이러한 성능은 독거노인의 생활 패턴을 효과적으로 예측할 수 있는 가능성을 시사하며, 조기 대응을 통해 독거노인의 생활 상태 변화에 필요한 서비스를 적시에 제공할 수 있는 기초 자료로 활용될 수 있다.

향후 연구에서는 더 다양한 생활 데이터(전기, 가스 검침정보 및 각종 IoT 센싱값)와 많은 양의 시계열 데이터를 활용하여 모델의 예측 정확성을 높이는 연구개발이 필요하다.

이를 통해 독거노인 가구의 고독사를 조기에 발견하고 지원하는 서비스 플랫폼 개발에 기여할 수 있을 거라고 기대된다.



## 참 고 문 헌

- [1] 오진영. "독거노인의 고독사 방지를 위한 안전 시스템 방안 연구." 국내석사학위논문 전북대학교 행정대학원, 2018. 전북특별자치도
- [2] 김호현. "LSTM/GRU 순환신경망을 이용한 시계열데이터 예측." 국내석사학위논문 한국방송통신대학교, 2017. 서울

# FPGA 기반 원전 제어시스템의 사이버보안 특성

김태희\*, 이수봉\*\*, 이현주\*\*\*

## Cybersecurity Characteristics of FPGA-based Nuclear Power Plant Control Systems

Taehee Kim\*, Subong Lee\*\*, and Hyunjoo Lee\*\*\*

### 요 약

원자력발전소의 제어시스템은 최근 구체적이고 실제적인 사이버 위협에 노출되고 있다. 사이버 위협으로부터 비교적 안전한 것으로 알려진 FPGA 기반 제어시스템이 개발 및 도입되고 있으나, 이 역시 사이버 공격으로부터 완벽하게 안전한 것은 아니다. 본 논문에서는 FPGA 기반 제어시스템의 소프트웨어적, 하드웨어적 사이버보안 특성을 살펴보고, 앞으로의 연구 과제를 제시한다.

### Abstract

The I&C systems of nuclear power plants have recently been exposed to concrete and real cyber threats. Although FPGA-based systems, known for their relative safety against cyber threats, are being developed and implemented, they are not completely immune to cyber attacks. This paper examines the software and hardware cybersecurity characteristics of FPGA-based systems and presents future research directions.

### Key words

nuclear power plant, FPGA, I&C systems, cybersecurity

### 1. 원자력발전소의 사이버 위협

원자력발전소의 제어시스템은, 대체로 원자력 발전소의 안전한 운영을 목적으로 설계 및 제작되어 왔다. 이러한 목적의 원자력발전소 제어시스템은 고성능을 요구하지 않으며, 다양한 기능, 예컨대 보안 기능을 요구하기보다는 안전을 위해 신뢰성을 높일

수 있는 방향으로 발전해왔다.

최근 몇 년 사이에 원자력발전소에 대한 사이버 위협은 증가하고 있으며, 발전소는 아니지만 이란의 우라늄 농축시설을 목표로 한 2010년의 스텝스넷[1] 사례와 같이 원자력시설을 향한 구체적인 위협이 드러나고 있다.

반면 사이버 위협에 대한 대응 역시 점점 더 진

\* 한국원자력통제기술원, kimtaehee@kinac.re.kr(교신저자), \*\* 한국원자력통제기술원, sblee@kinac.re.kr, \*\*\* 한국원자력통제기술원, sn220865@kinac.re.kr

※ This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety(KoFONS), granted financial resource from the Nuclear Safety and Security Commission(NSSC), Republic of Korea. (No. 2106012)

화해가고 있다. 미국을 필두로, 세계 여러 나라들은 원자력발전소로 하여금 사이버보안 체계를 갖추도록 요구해 왔으며, 이에 대한 규제[2][3][4][5]역시 적용 중에 있다.

## II. 원자력발전소의 FPGA 기반 제어시스템

최근에는 원자력발전소의 제어시스템에서는 일반적인 중앙처리 장치 대신 FPGA(Field-Programmable Gate Array)[6]를 채택한 제어시스템이 등장하고 있다. FPGA는 내부에 많은 수의 논리게이트를 갖추고 있으며, 통합 개발툴(IDE)을 통해 이러한 논리게이트를 사용자의 의도대로 배열하여 여러 가지 소프트웨어적 연산과 하드웨어적 기능을 수행할 수 있다.

즉, 프로그램 가능한 하드웨어라고 정의할 수 있다. FPGA는 소프트웨어처럼 사용자가 원하는 대로 작성할 수 있지만, 하드웨어처럼 높은 신뢰도로 작동하며, 각 연산에 대해 병렬적으로 동시에 작동할 수 있는 장점이 있다.

그러나 FPGA의 가장 중요한 장점 중 하나는 사이버보안 관점에서 상대적으로 안전하다는 점이다.

## III. FPGA의 사이버보안 특성 및 공격 사례

### III-1. FPGA의 소프트웨어적/하드웨어적 사이버보안 특성

소프트웨어적 관점에서 볼 때, 일반적인 중앙처리 장치 기반 제어시스템의 소프트웨어는 운영체제, 미들웨어, 제어로직 등으로 구성되어 상대적으로 넓은 공격벡터를 갖게 된다.

그에 비하여 FPGA는 하드웨어와 소프트웨어가 결합된 고유의 특성으로 인해, 설계 파일인 비트스트림[6]을 단일한 공격벡터로 갖게 된다. 더구나 이러한 비트스트림이 FPGA의 내부 메모리에 암호화 저장될 경우, 이러한 비트스트림을 추출 및 해독하는 것은 일반 소프트웨어를 리버스 엔지니어링하는 것보다 훨씬 어렵게 된다. 즉, 보안 측면에서 큰 장점을 갖는다.

또한 하드웨어적인 관점에서, 비트스트림의 추출 및 해독하거나 공격을 위해서는 FPGA가 비트스트림을 사용할 때의 전력소비, 전자기방사, 시간 차이 등을 분석하거나, 열 레이지 등을 조사하여 암호화 키를 획득하여야 한다. 이는 동작 중인 제어시스템의 하드웨어에 물리적으로 근접하여야 하며, 분석에 수 시간 이상이 소요되는 어려움이 있다.

즉, FPGA는 소프트웨어적/하드웨어적으로 사이버보안 관점에서 상당한 강점을 갖는다.

### III-2. FPGA에 대한 사이버공격 사례

그러나 FPGA 또한 사이버공격에 완벽하게 안전한 것은 아니다. “III-1”에서 언급한 바와 같이 비트스트림의 추출 및 해독에 성공한다면 FPGA로 하여금 의도하지 않은 동작을 수행하게 하는 것이 불가능한 것은 아니다.

FPGA에 대한 사이버공격으로 잘 알려진 공격 기법 중 하나는 글리칭[7]이다. 글리칭은 FPGA의 비트스트림을 분석하여, 특정 시점에 전원 공급 또는 클럭 신호를 짧은 순간에 변조함으로써, FPGA로 하여금 의도된 동작을 수행하지 못하게 하는 공격 기법이다.

만약 FPGA의 전체 비트스트림에 대한 완벽한 분석이 가능하다면 글리칭 공격을 통해 FPGA의 동작을 정밀하게 제어할 수도 있을 것이다. 또한 비트스트림의 일부분의 분석에만 성공하더라도, 해당 부분의 전원이나 클럭 신호를 짧은 순간에 변조하게 되면 FPGA는 내부에서 처리 중인 데이터나 연산이 타이밍 오류를 일으키고, 잘못된 연산을 수행하거나 특정 동작을 생략하는 결과를 초래할 수 있다.

가장 잘 알려진 사례로는 마이크로소프트의 가정용 게임기인 xbox360[8]에 대한 클럭 글리칭 공격이 있다. xbox360은 불법 복제된 게임 소프트웨어가 실행될 수 없도록 부팅 시 보안을 인증을 수행한다. 해커들은 이러한 보안 인증을 우회하기 위하여 보안 인증이 마치 성공한 것처럼 보이는 결과 값이 나올 수 있는 시점에 대해 소프트웨어적으로 분석 후, 하드웨어적으로 글리칭 공격을 수행하는 칩을 xbox360의 기관에 이식하

였다. 이러한 공격은 하드웨어적이고 전기적인 수준에서 이루어지기 때문에 짧은 수초에서, 길면 십 수초 내에 보안 인증을 우회하는 결과에 도달하였다.

만약 숙련된 엔지니어가 동작이 정지된 제어시스템에 물리적으로 근접할 수만 있다면, 불과 수 분 이내에 글리치 칩의 이식에 성공할 수도 있다. 실제로 유튜브 등을 통해 공개된 글리치 칩의 이식 과정을 살펴보면 납땀 등 실제로 시간이 소요되는 작업량은 얼마 되지 않는 것으로 확인되었다.

#### IV. 결론

우리는 본 논문에서 FPGA의 소프트웨어적/하드웨어적 사이버보안 특성과 함께 이를 이용한 글리칭 공격 사례를 소개하였다. 글리칭 공격과 같이 FPGA의 사이버보안 특성에 대한 공격 시도는 모두 FPGA에 대한 물리적 접근에 시작할 수 있다.

비트스트림 분석을 위해서는 FPGA에 접근하여 전용 장비를 사용하여 메모리를 덤프하거나 전자기 방사, 열 레이저나 포커스 이온 빔을 조사하여야 한다. 또한 글리칭 공격을 위해서는 특정한 칩을 기관에 부착하여야 하며 역시 물리적인 접근이 필요하다.

이러한 접근은 공급망 통제 과정 중에 발생할 수도 있으며, 제어시스템의 운영 중 혹은 유지/보수 중에도 발생할 수 있다.

향후에는 FPGA의 보안을 위한 특성화된 접근 통제, 유지보수 장비의 반입 통제, FPGA의 불법적인 분석을 어렵게 만들 수 있는 설계 사례 등에 대한 수집과 연구가 필요할 것으로 판단된다.

#### 참 고 문 헌

[1] K. Zetter, *Countdown to zero day: Stuxnet and the launch of the world's first digital weapon*, in *Proc. IEEE Symp. Security and Privacy*, 2014.

[2] 대한민국, *원자력시설 등의 방호 및 방사능 방재 대책법*, 법률 제18664호, 2021. 12월 28일 개정.

[3] 한국원자력통제기술원, *심·검사 기준 KINAC/RS-015 원자력시설의 컴퓨터 및 정보시스템 보안*, 대한민국, 2023.

[4] U.S. Nuclear Regulatory Commission, *Protection of digital computer and communication systems and networks*, Code of Federal Regulations, Title 10, Part 73.54, 2009.

[5] U.S. Nuclear Regulatory Commission, *Cyber security programs for nuclear facilities*, Regulatory Guide 5.71, Rev. 1, 2023.

[6] A. Duncan et al., "FPGA Bitstream Security: A Day in the Life", in *Proc. 2019 IEEE Int. Test Conf. (ITC)*, Washington, DC, USA, 2019.

[7] Z. Liu et al., "Clock-glitch fault injection attacks on an FPGA AES implementation", *J. Electrotech. Electr. Eng. Manage.*, vol. 1, no. 1, pp. 23-27, 2017.

[8] Xbox360, "Xbox360," Xbox Official Website. [Online]. Available: <https://www.xbox.com/ko-KR/games>. [Accessed: Oct. 28, 2024].

# 하드웨어 효율적인 양자화 Vision Transformer 설계를 위한 Softmax 최적화 기법

김현성\*, 전석훈, 김병수\*\*

## Softmax Optimization Method for Hardware-Efficient Quantized Vision Transformer Design

Hyeonseong Kim, Seokhun Jeon, and Byung-Soo Kim

본 논문은 Vision Transformer (ViT) 모델의 Attention 메커니즘의 연산 효율을 향상시키기 위한 알고리즘의 하드웨어 구현 방법을 제안하고, 해당 알고리즘의 실효성을 분석한다. ViT는 뛰어난 이미지 분류 성능을 보여주지만, Transformer 모델의 막대한 파라미터 수와 Attention 연산의 높은 복잡도는 ViT의 실제 하드웨어 구현을 어렵게 한다. 이러한 문제를 해결하기 위해, 하드웨어 효율성을 고려한 ViT 설계를 목표로 일부 Softmax 연산을 생략할 수 있는 양자화된 Hierarchical Transformer 기반의 알고리즘[1]이 제안되었다. 해당 알고리즘은 산술 및 기하 평균과 양자화 레벨을 기존 수식에 적용하여 Softmax에서 확률이 낮은 요소를 저복잡도 연산으로 예측하고, 복잡한 연산의 일부를 생략하는 방법을 제시한다. 이 알고리즘은 소프트웨어적으로 Softmax 연산이 일부 생략될 수 있음을 보여주었으나, 하드웨어적 구현 가능성이나 Softmax 연산에서 어떤 부분이 구체적으로 생략되는지에 대한 분석이 미흡하다. 본 논문에서는 양자화 비트와 데이터 버퍼를 고려하여 해당 알고리즘을 적용한 Softmax 아키텍처를 제안한다. 제안된 아키텍처는 실제 연산 과정에서 나눗셈 연산의 생략을 통해 Softmax 연산의 효율을 높이며, 낮은 비트의 가산기와 Shift 연산기를 사용하여 하드웨어 오버헤드를 최소화한다. 결과적으로 본 논문은 Softmax 저확률 예측 알고리즘의 하드웨어 데이터플로우를 제시하였으며, 하드웨어 효율을 극대화하기 위해서는 양자화 레벨의 적절한 설정이 중요함을 보였다.

### Key words

Vision Transformer, Attention Mechanism, Hardware Efficient Transformer, Quantization, Softmax Optimization

### 참고문헌

- [1] Hyeonseong Kim and Byung-Soo Kim, "Hardware-aware Hierarchical Transformer using Softmax Low Probability Prediction", Proceedings of the International SoC Design Conference (ISOCC), August 2024

---

\*한국전자기술연구원, hyeonseong.kim@keti.re.kr, \*\*한국전자기술연구원, bskim4k@keti.re.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2024-00402898, 시뮬레이션 기반 고속/고정확도 데이터센터 워크로드/시스템 분석 플랫폼 개발)

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술기획평가원의 지원을 받아 수행된 연구임 (RS-2022-00155731, 융합 메모리와 고속 연산 엔진 제어를 위한 하드웨어/소프트웨어 융합 메모리 솔루션 기술 개발)

# 인공지능 모델 기반 리튬인산철 배터리 상태추정 연구

이은지\*, 김은주\*, 김민석\*, 박형욱\*\*

## A Study on Lithium Iron Phosphate Battery State Estimation Based on Artificial Intelligence Models

Eunji Lee\*, Eunju Kim\*, Minseok Kim\*, and Hyungwook Park\*\*

본 연구는 리튬인산철 배터리의 충방전 데이터를 이용하여 배터리의 상태를 추정하는 인공지능 모델 개발을 목적으로 한다. 배터리는 충방전 사이클 증가에 따라 열화가 발생하며, 초기 사이클의 방전 전압 특성이 리튬인산철 배터리의 최종 수명과 상관성이 있다는 선행 연구[1]를 바탕으로 배터리 수명 예측 모델을 설계하였다. 본 연구에서는 CNN(Convolutional Neural Network) 모델과 MLP(Multi Layer Perceptron) 모델을 활용하였으며, 입력 데이터로는 각 배터리의 사이클별 전압 데이터를 각 사이클당 1,000개씩 샘플링하여 사용하고, 출력 데이터로는 배터리의 최종 사이클 수를 설정하였다. 데이터는 공개된 리튬인산철 충방전 데이터셋[2]을 활용하였으며, 예측의 정확도를 높이고 모델의 과적합을 방지하기 위해 하이퍼파라미터 최적화 과정을 수행하였다. 또한, 배터리의 최종 사이클 수는 배터리의 상태에 따라 100에서 2,000사이클로 범위가 매우 넓기 때문에, 예측값과 실제값의 오차를 계산하여 임계값 범위 내에 있는 경우 정확한 예측으로 간주하여 정확도 평가를 진행하였다. 평가는 두 가지 조건에서 진행되었다. 첫 번째 테스트는 학습 데이터와 동일한 조건에서 수집된 43개의 배터리 데이터로 수행하였으며, 두 번째 테스트는 동일한 배터리 모델이지만 충방전 정책이 상이한 환경에서 수집된 40개의 데이터를 이용하여 모델의 일반화 성능을 검증하였다. 그 결과, MLP 모델은 첫 번째 테스트에서 93.02%, 두 번째 테스트에서 92.5%의 정확도를 보였으며, CNN 모델은 첫 번째 테스트에서 97.67%, 두 번째 테스트에서 90.00%로 상대적으로 높은 성능을 보였다. 본 연구는 리튬인산철 배터리의 성능을 추정하기 위해 방전 전압을 특징으로 활용한 선행 연구를 바탕으로 검증을 수행하였으며, 향후 연구에서는 시퀀스 데이터의 특성을 보다 정밀하게 반영할 수 있는 Transformer 모델[3]을 적용하여, 배터리 사용 패턴과 충방전 데이터에 기반한 재사용 리튬인산철 배터리의 잔여 수명 예측을 위한 연구를 수행할 계획이다.

### Key words

Battery State Estimation, Lithium Iron Phosphate Battery, Artificial Intelligence, Data Analysis

### 참고문헌

- [1] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, and P. K. Herring, "Data-driven prediction of battery cycle life before capacity degradation," Matroid, 2019.
- [2] [Data set] Available: <https://data.matr.io/1/projects/5c48dd2bc625d700019f3204>
- [3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I., "Attention Is All You Need," in Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, Dec. 2017, pp. 6000-6010.

\*한국과학기술정보연구원 데이터분석본부 호남지원, eunji\_lee@kisti.re.kr, ejkim@kisti.re.kr, kimminseok@kisti.re.kr,

\*\*hyungwook@kisti.re.kr(교신저자)

※ 본 연구는 2024년 한국과학기술정보연구원 주요사업에 의해 수행된 연구결과입니다(K24L3M3C5-04).

# ResNet 지원 Binarized 뉴럴 네트워크 프로세서 구조 연구

이재학\*

## Binarized Neural Network Processor Architecture for Residual Network Support

Jeahack Lee\*

본 논문에서는 residual network(ResNet)를 지원하는 이진화 신경망(Binarized Neural Network; BNN) 프로세서 구조를 제안하며, 이는 메모리 효율성과 처리 속도가 중요한 모바일 및 IoT 엣지 디바이스에 적합한 솔루션으로 기대된다. BNN은 가중치와 활성화 매개변수를 1 비트로 줄여 메모리 사용량을 크게 절감할 수 있지만, 역전파(backpropagation) 알고리즘 과정에서 기울기 소실 문제가 발생하기 쉬워 학습 효율이 낮아지는 단점이 있다. ResNetE는 ResNet에서 블록의 크기를 줄이고 더 많은 residual 경로를 추가하여 데이터 흐름을 원활하게 하는 구조적 변화를 통해 신경망의 성능을 향상시킨 변형 모델로, 높은 정확도와 함께 복잡도를 줄이는 장점이 있다. 본 논문은 ResNetE를 지원하는 높은 에너지 효율을 가지는 BNN 프로세서 구조를 제안하였다. 이진 데이터 처리 및 에너지 효율을 높이기 위해 두가지 특징을 가진다. 첫째, 제안된 BNN 프로세서는 외부 메모리 접근을 최소화하는 데이터 플로우를 통해 높은 에너지 효율을 보인다. 가중치와 활성화 매개변수를 내부 메모리에 저장하여 외부 메모리 접근 횟수를 줄임으로써 데이터 병목 현상을 해결하고, 이를 통해 실시간 데이터 처리가 가능하다. 또한, 각 레이어의 입력 데이터를 한 번만 로드하여 재사용하는 구조를 채택함으로써 메모리 사용의 효율성을 높였다. 이러한 구조는 기존의 BNN 가속기[2]들이 주로 VGG와 같은 단순 네트워크를 대상으로 했던 것과 달리, 복잡한 ResNet 경로를 효과적으로 지원한다. 둘째, 높은 데이터 처리 속도를 달성하기 위해 3x3 컨볼루션을 수행할 수 있는 processing element(PE) 어레이 구조로 설계하였다. XNOR 연산과 popcounter 기능을 통해 이진화된 입력 데이터를 빠르게 처리할 수 있으며, 이를 통해 연산 지연을 최소화한다. 9개의 BNN 코어로 구성된 프로세서는 각 코어가 ResNetE 네트워크 구조를 수행하도록 최적화되어 있어 높은 에너지 효율을 보인다. 높은 에너지 효율로 인해 전력과 공간이 제한된 모바일 및 IoT 엣지 디바이스에서도 CNN 모델을 효율적으로 구현할 수 있으며, 향후 인공지능 엣지 컴퓨팅 환경에서 폭넓은 응용이 기대된다.

Key words

Binarized neural network, Residual network, Neural network processor

### 참고문헌

- [1] J. Bethge, et al.: Back to simplicity: How to train accurate bnns from scratch?." arXiv preprint arXiv:1906.08637 (2019).
- [2] P. Guo, H. Ma, R. Chen, P. Li, S. Xie, and D. Wang, "FBNA: A fully binarized neural network accelerator," in Proc. IEEE Int. Conf. on Field Program. Logic and Applications (FPL), Aug. 2018, pp. 51-513.

---

\*한국전자기술연구원 SoC플랫폼연구센터, email : jhk507@keti.re.kr (교신저자)

※ 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.RS-2023-00277060, 개방형 엣지 AI 반도체 설계 및 SW 플랫폼 기술개발)

# 차량용 센서 오동작 방지를 위한 적응형 삼중 중복 모듈 설계

오경목\*, 이재학, 최성민, 김병수\*\*

## Adaptive Triple Module Redundancy Design for Fault Prevention in Automotive Sensors

Kyeongmook Oh\*, Jeahack Lee, Sungmin Choi, and Byung-Soo Kim\*\*

본 논문에서는 고신뢰성 및 고안전성을 갖춘 차량용 센서 데이터 처리를 위한 TMR(Triple Modular Redundancy) 알고리즘 및 설계된 하드웨어 IP를 제안한다. 자동차 전장 시스템에서 센서 데이터를 처리하는데 차량의 신뢰성과 안전성은 가장 우선시되어야 하는 요소이며, 특히 자율주행 및 첨단 운전자 보조 시스템(ADAS)과 같은 차량 내 센서가 증가함에 따라 오동작 방지 및 시스템 모니터링 기능 등이 필수적으로 요구되는 추세이다. 제안하는 TMR IP는 총 5가지 모드를 지원하며, 각 모드는 3개의 센서 입력 데이터를 받아 신뢰성 높은 1개의 보정된 데이터를 출력하도록 설계되어 있다 [1]. 각 채널에서 에러가 발생하는 경우에는 에러가 발생한 채널을 사용자에게 알려주며, 그 정보를 이용해 사용자는 전체 시스템의 불량 원인 파악 및 시스템 안전모드 구현 등을 통해 사고를 방지할 수 있으며, 자동차 전장 시스템에서 높은 신뢰성 및 안전성을 확보할 수 있다. 개발된 TMR 하드웨어 IP는 기본적인 평균 계산을 수행하는 기본동작 모드, 정규분포 형태의 데이터를 위한 가우시안 모드, 일정한 패턴의 데이터를 처리하는 평준화 모드, 사용자가 지정한 범위 내의 데이터를 처리하여 보정하는 범위 모드, 그리고 입력 데이터의 최소값과 최대값을 비교하여 오차범위를 고려한 최적의 출력을 제공하는 적응형 모드를 지원한다. 설계된 모든 모드에 대하여 C-modeling을 진행하여 reliability를 비교함으로써 실험 결과를 얻었으며, 제안하는 TMR 하드웨어 IP는 기본 TMR 대비 더 높은 reliability를 보여줌을 확인하였다 [2]. 특히, 본 연구에서 메인 타겟으로 설정한 실제 센서 데이터를 처리하는 응용 환경에서 우수한 성능을 실험 결과로 입증하였으며, 이는 제안하는 TMR 알고리즘 및 하드웨어 IP가 차량용 센서 데이터 처리에 최적화되었음을 보여준다. 제안하는 TMR 하드웨어 IP는 verilog로 설계 및 검증되었으며, vivado를 통해 FPGA 검증을 진행하였다.

### Key words

Triple Modular Redundancy(TMR), Automotive sensor signal processing, ADAS, Fault tolerant system

### 참고문헌

- [1] T. Arifeen, A. S. Hassan and J. A. Lee, "Approximate Triple Modular Redundancy: A Survey," in IEEE Access, vol. 8, 2020.
- [2] Hyunki Kim, Hyung-Joon Jeon, Keyseo Lee and Hyuntae Lee, "The design and evaluation of all voting triple modular redundancy system," Annual Reliability and Maintainability Symposium. 2002 Proceedings, Seattle, WA, USA, 2002.

---

\*한국전자기술연구원, okm806@keti.re.kr \*\*한국전자기술연구원, bskim4k@keti.re.kr (교신저자)

※ 이 연구는 2024년도 산업통상자원부 및 산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임 (RS-2024-00459534, 고신뢰성 대형승용차급 EMB(Electro-Mechanical Brake) 시스템용 AP 개발)



# 미생물을 활용한 건강한 농산물 생산을 위한 스마트팜 연구

김민희\*, 김희선\*\*

## Study on Smart Farm Technology for Healthy Agricultural Products Using Microorganisms

Min-Hee Kim\*, Heesun Kim\*\*

### 요 약

현대 농업은 식량 생산의 효율성 향상과 품질 개선을 위한 기술 혁신을 요구받고 있으며, 기후 변화와 환경 오염 문제로 인해 지속 가능한 농업의 필요성이 증가하고 있다. 스마트팜 기술은 데이터 기반의 정밀 농업 관리를 통해 자원 효율성을 극대화하고 환경 영향을 줄이는 대안으로 떠오르고 있다. 특히, 스마트팜의 중요한 요소로 미생물 활용이 주목받고 있다. 미생물은 토양 건강을 개선하고 작물의 성장과 면역력을 증진시키며, 병해충 억제 등의 기능을 수행한다. 이를 통해 화학 비료와 농약 사용을 줄여 환경 친화적이면서도 생산성을 높일 수 있다. 본 연구는 미생물을 활용한 스마트팜 기술과 적용 사례를 분석하여 건강한 농산물 생산을 위한 방법을 제안한다.

### Abstract

Modern agriculture demands technological innovations to improve efficiency and quality of food production, while the need for sustainable agriculture is increasing due to climate change and environmental pollution issues. Smart farm technology is emerging as an alternative that maximizes resource efficiency and reduces environmental impact through data-driven precision agricultural management. In particular, the utilization of microorganisms is gaining attention as a key element of smart farms. Microorganisms improve soil health, enhance crop growth and immunity, and help suppress pests and diseases. This enables a reduction in the use of chemical fertilizers and pesticides, resulting in eco-friendly and productive agriculture. This study analyzes smart farm technologies utilizing microorganisms and presents methods for producing healthy agricultural products.

### Key words

Modern agriculture, Smart farm technology, Data-based precision farming

---

\* 국립안동대학교 대학원 소프트웨어융합학과, dandy292@naver.com

\*\* 국립안동대학교 소프트웨어융합학과, hskim@anu.ac.kr

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

## I. 서론

현대 농업은 식량 생산의 효율성과 품질을 향상시키기 위한 다양한 기술 혁신을 필요로 하고 있다. 4차 혁명 이후 빅데이터 기반의 통합 환경 제어 시스템의 고도화, 자동제어가 가능한 시스템의 필요성이 요구되어지고 있다[1]. 특히 기후 변화와 환경오염 문제로 인해 지속 가능한 농업의 필요성이 증대되고 있으며, 이에 따라 농업 분야에서는 자원 활용을 최적화하고 환경에 미치는 부정적 영향을 줄이기 위한 연구가 활발히 이루어지고 있다. 이러한 배경 속에서 스마트팜 기술은 농생명 분야의 ICT 융복합 기술을 산업 기술과 융합하여 이를 통해 데이터를 수집하고 분석하여 농산물 생산에 효과적으로 적용하고 운용하고 있다[2]. 특히 데이터 기반의 정밀 농업은 농업 관리와 자원 효율성을 통해 새로운 대안으로 떠오르고 있다.

현대 농업에서 농산물 생산을 증대하기 위해 화학 비료와 농약의 사용이 보편화되었다. 그러나 이러한 화학 비료의 과도한 사용은 토양의 건강 악화, 수질 오염, 생태계 교란 등 여러 환경적 문제를 발생시킨다. 특히 장기적인 관점에서 화학 비료에 의존한 농업 방식은 토양의 비옥도를 저하시켜 작물의 생육을 저해하며, 소비자에게는 잔류 농약 및 화학 물질로 인한 건강 위험을 초래할 수 있다. 이에 환경 친화적이고 지속 가능한 농업 방식에 대한 요구가 증가하고 있으며, 이러한 필요성은 스마트팜 기술과 미생물 활용의 중요성을 부각시키고 있다.

미생물은 토양의 건강을 개선하고, 작물의 생장과 면역력을 증진시키며, 병해충 억제 등 다양한 기능을 수행한다. 이는 화학 비료와 농약 사용을 줄이고, 환경 친화적이면서도 높은 생산성을 유지할 수 있는 농업을 가능하게 한다. 이에 미생물을 활용한 스마트팜 기술은 지속 가능한 농업을 실현하기 위한 필수적인 접근법으로 평가되고 있다.

본 논문에서는 미생물 활용이 현재 적용되고 있는 스마트팜 기술을 분석하여 건강한 농산물 생산을 위한 효과적인 방법을 연구한다. 미생물은 농업 생산성 향상에 기여할 수 있는 중요한 요소로, 본 연구를 통해 미생물을 활용한 스마트팜 기술의 전반적인 효과와 그 적용 가능성을 연구하고자 한다.

## II. 관련 연구

### 2.1 애그테크(AgTech, Agricultural Technology)

애그테크는 농업(Agritech)과 기술(Tech)가 합쳐진 농업기술로 정보통신기술(ICT), 인공지능(AI), 사물인터넷(IoT) 등 첨단 기술을 융합하여 농업 생산성을 높이고, 자원 효율성을 증대시키며, 환경 친화적인 농업을 실현하는 것을 목표로 하는 분야이다. 그림 1은 애그테크는 농업의 스마트화를 가능하게 하는 핵심 기술로, 그린 바이오, 하드웨어, 소프트웨어 등의 다양한 분야에서 혁신적인 발전을 나타낸 그림이다[3].

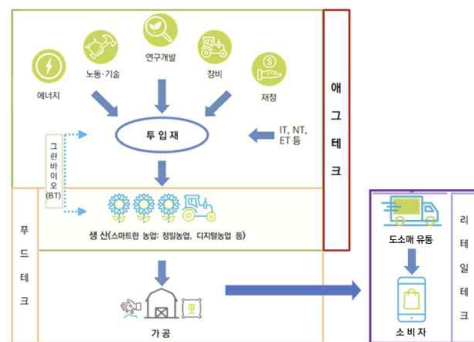


그림 1. 애그테크 범위  
Fig. 1. Scope of Agricultural Technology

애그테크는 농축산업의 생산활동에서 자원 투입의 효율화를 도모하고, 생산성 증대와 고부가가치 창출을 목표로 활용되는 첨단기술(IT, ET, NT, BT 등)을 의미한다. 이는 첨단 기술을 적용한 농업 투입재(input)와 이를 활용한 농업 생산 활동을 포괄하며, 다양한 기술 간의 융복합을 통해 시너지 효과를 극대화하는 특징을 지닌다. 기존의 농업 생산 방식의 한계를 극복하고, 지속 가능한 농업 체계를 구축하는 데 있다. 이러한 기술적 접근은 기후 변화, 환경 오염, 자원 고갈 등 문제를 해결하기 위해 더욱 주목받고 있다. 다양한 센서와 데이터를 기반으로 작물의 생장 상태를 지속적으로 모니터링하고 분석하는 기술은 빅데이터를 활용한 모니터링 시스템에 있다.

## 2.2 식물이 필요로 하는 원소의 종류

작물의 생육은 종자에서 시작하여 성장과 발달 과정을 거쳐 수확 가능한 상태에 이르는 일련의 과정으로 정의된다. 이 과정에서 식물은 줄기, 잎, 뿌리, 열매, 꽃 등 다양한 기관을 형성하며, 이러한 기관의 크기가 커지고 기능이 점차 강화되는 단계적 성장을 이룬다. 아래의 표 1은 작물 생장에 필요한 16가지의 무기 양분을 나타낸 표이다.

표 1. 작물 생장에 필요한 원소의 종류  
Table 1. Types of Elements Essential for Plant Growth

필수 원소 구분	필수 원소
자연에서 흡수	탄소(Carbon)
	수소(hydrogen)
	산소(oxygen)
다량원소	질소(nitrogen)
	칼륨(potassium)
	인(phosphorus)
소량원소	황(sulfur)
	칼슘(calcium)
	마그네슘(magnesium)
미량원소	붕소(boron)
	염소(chlorine)
	구리(copper)
	몰리브덴(molybdenum)
	아연(zinc)
	망간(manganese)
	철(iron)

식물이 흡수하는 양분이 무기질이라는 사실은 19세기 독일의 저명한 화학자 리비히(Justus von Liebig)가 1841년에 식물의 무기 양분설을 발표하였다. 무기 양분설은 작물의 생장에 필요한 16가지 무기 양분을 균형 있게 공급하여, 작물의 건강한 성장과 생산력을 유지하는 것을 목표로 한다. 이 이론에 따르면, 토양과 작물의 무기 양분 상태를 분석한 후, 부족한 무기 양분을 적절한 비율로 보충함으로써 작물에 필요한 영양분을 최적화하여 공급할 수 있다. 이를 통해 작물의 건강한 성장과 생산성을 향상시키는 것이 궁극적인 목적이다.

## 2.3 농업용 미생물

화학 농약의 잔류 독성 문제는 인체에 대한 위해성을 야기할 수 있어, 친환경 농산물에 대한 선호가 증가하고 있다. 이러한 상황에서 농업 분야는 화학 농약과 비료의 사용을 줄이는 방향으로 나아가고 있으며, 이는 농업 환경과 인체 건강에 긍정적인 영향을 미친다. 특히, 비료의 과다 사용은 토양 염류의 직접적인 축적과 환경 오염을 초래할 수 있다. 화학 농약의 장기적인 사용은 내성을 가진 저항성 균의 출현과 병충해 발생의 위험성을 높일 수 있다. 이러한 이유로, 지속 가능한 농업을 위해서는 토양 개량 및 작물의 생육을 촉진하는 친환경적 접근이 요구되고 있다. 이는 농업 환경의 건강을 유지하고, 생산성을 증대하는 데 중요한 요소로 작용한다. 아래의 표 2는 농업에서 사용하는 농업 미생물의 종류와 용도에 대해 정리한 표이다.

표 2. 대표적인 농업 유용 미생물  
Table 2. Representative Microorganisms in Agriculture

농업 유용 미생물	
고초균	작물 체질 향상, 작물 내병성 증가, 축사 악취 제거
유산균	유해균 억제, 유기물 가용화, 수확량, 토양 개선
광합성균	작물 생육 촉진, 토양 비옥화, 연작장애 개선
효모	미량원소 공급, 가축면역력 증가, 축사 악취제거
클로렐라	작물 생육 촉진, 품질, 저장성 증대, 유해 미생물 조절

농업 유용 미생물은 작물의 뿌리 주위에서 영양분의 공급을 증진시키며, 식물의 성장과 건강에 중요한 역할을 한다. 이러한 미생물은 식물의 뿌리 주변에서 서식하며, 영양분의 공급을 도와 식물의 전반적인 생육을 촉진한다. 또한, 식물호르몬의 대사활동을 유도하고, 호르몬의 균형을 유지하여 작물의 성장을 강화한다. 식물의 방어기작을 유도하여 병원균에 대한 저항성을 강화하는 역할을 수행한다. 따라서 유용 미생물은 작물의 성장 촉진, 종자 발

아 촉진, 작물 생산성 증대, 병원균 억제 등의 효과가 있다. 특정 유용 미생물이 환경과 상호작용하면서 식물 성장과 방어에 미치는 영향은 스마트팜 환경에서의 미생물 활용 가능성을 높이는 중요한 요소가 된다.

### III. 연구 동향

#### 3.1 생육환경 모니터링 시스템

모니터링 시스템은 스마트팜의 핵심 기술로, 시설재배지나 노지에서 농작물의 생육 환경을 감시하고 제어하기 위한 주요 장치로 활용된다. 이러한 시스템은 작물의 성장과 수확량을 예측할 수 있는 데이터를 제공함으로써, 효율적인 농업 생산을 위한 중요한 정보를 제공한다. 그림 2는 노지에 설치된 각종 센서로 생육환경 정보와 농작물이 생체정보를 측정할 수 있는 모니터링 시스템의 구성도이다.

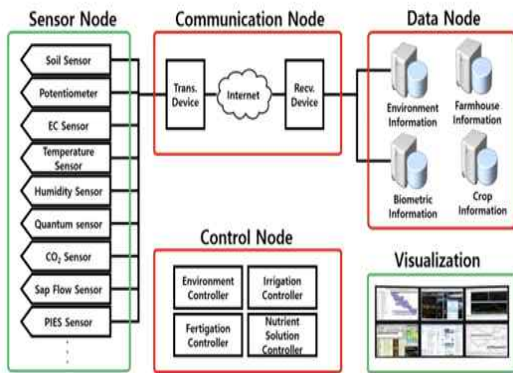


그림 2. 모니터링 시스템 구성  
Fig. 2. Configuration of Monitoring System

모니터링 시스템의 주요 구성 요소는 다양한 센서를 포함하며, 이 센서들은 농업 환경의 변수를 실시간으로 측정한다. 센서에서 측정된 아날로그 신호는 디지털 신호로 변환된 후, 데이터베이스에 빅데이터 형태로 저장된다. 이러한 데이터는 작물의 생육 상태를 분석하고, 농업 생산성 향상에 필요한 의사결정을 지원하는 데 중요한 역할을 한다[4]. 인공 신경망의 학습 시에 깊이 맵을 FIS로 정확하게 근사시키면서 학습이 되기에 기존의 방법보다 개선된 복원 결과를 보여준다.

#### 3.2 복합 미생물 배양기 딥러닝 구조 개발

비료와 분뇨와 같은 물질은 농가에서 많이 사용되며, 환경오염의 주요 원인으로, 악취로 인한 주민들의 민원과 무단 투기로 인한 수질오염 문제가 지속적으로 발생하고 있다. 이러한 문제를 해결하기 위해 세균의 성장을 억제하는 복합 미생물이 주목받고 있다[5]. 그림 3은 딥러닝 기반의 예측 결과를 활용한 복합 미생물 배양기를 개발을 위한 딥러닝 네트워크의 학습과정을 나타낸 설계도이다.

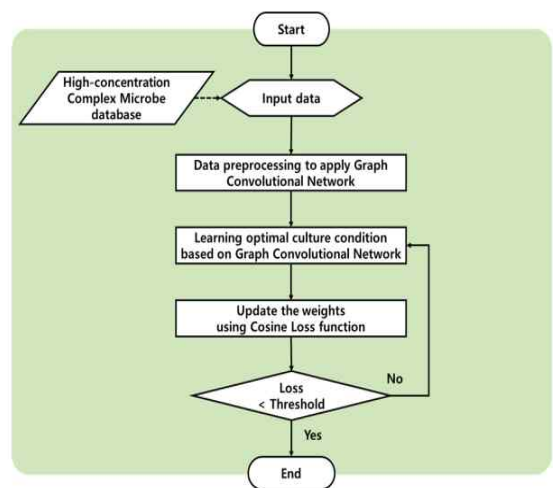


그림 3. 딥러닝 네트워크의 학습과정  
Fig. 3. Process of Deep Learning Network

개발된 배양기는 배양 실패 시 발생할 수 있는 경제적 손실을 최소화할 수 있도록 설계되었다. 전처리된 데이터의 인접 행렬과 속성 정보를 그래프 합성곱 네트워크(Graph Convolutional Network, GCN)의 입력으로 활용하여 학습을 진행하고, 네트워크의 출력으로 생성된 확률 정보를 바탕으로 최적의 배양 조건을 찾기 위한 학습을 수행한다[5]. 소규모 데이터 학습에 유용한 코사인(Cosine) 손실 함수를 사용하여 손실(Loss)이 임계값보다 작아질 때까지 반복 학습을 진행한다.

#### 3.3 복합환경 제어시스템

복합 환경 제어 시스템은 온실 내외부에 설치된 다양한 센서를 통해 환경 요소들을 측정하고, 측정

된 데이터를 바탕으로 환경 요소 간의 상호 관련성을 분석하여 작물의 최적 생육 환경을 조성한다[6]. 그림 4는 복합 환경 제어 시스템의 프로세스를 보여주고 있다.

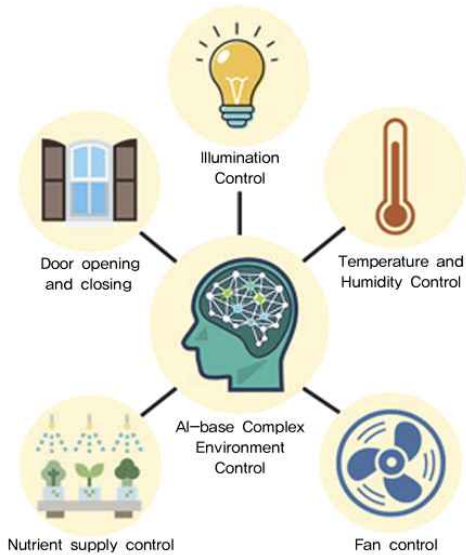


그림 4. AI기반 통합환경 제어시스템  
Fig. 4. AI-based Integrated Control System

이 시스템은 온도, 습도, 광량 등 주요 환경 변수를 실시간으로 모니터링하고, 분석된 결과에 따라 구동 기기를 정밀하게 제어함으로써 작물의 성장 조건을 최적화한다. 수집된 데이터를 저장하고 분석하여 재배 환경 관리의 효율성을 지속적으로 개선할 수 있도록 지원한다[6]. 복합 환경 제어 시스템은 농업 생산성을 높이고 재배 환경 관리를 개선할 수 있는 시스템이다.

#### IV. 결 론

본 논문에서는 미생물 활용이 스마트팜 기술에 미치는 영향을 분석하고, 건강한 농산물 생산을 위한 효과적인 방안을 제시하였다. 농업용 유용 미생물의 활용은 농작물의 생육 촉진과 병충해 저항성 강화에 중요한 역할을 한다. 유용 미생물은 작물의 뿌리 주변에서 영양분 공급을 증진시키고, 호르몬 균형을 유지하여 작물의 전반적인 성장을 촉진한다. 미생물의 방어 기작을 통해 병원균에 대한 저항성을 강화함으로써 화학 농약 사용을 줄이고, 농업

환경의 건강을 유지할 수 있다. 복합 미생물 배양기를 위한 딥러닝 구조를 개발을 통하여, 딥러닝 기반의 예측 결과를 통해 최적의 배양 조건을 학습하고, 배양 실패 시 발생할 수 있는 경제적 손실을 최소화하는 할 수 있는 연구 방법을 알아보았다. 이를 통해 농업 현장에서 복합 미생물의 효율적인 활용과 성공적인 배양을 지원할 수 있다. 복합 환경 제어 시스템은 온실 내외부의 센서를 통해 환경 요소를 측정하고 분석하여, 작물의 최적 생육 환경을 조성하는 기술이다. 이 시스템은 데이터를 기반으로 구동 기기를 정밀 제어하여 재배 환경을 최적화하고, 지속적인 개선을 지원한다.

향후 연구에서는 다양한 농업 환경에서의 미생물 적용과 함께 딥러닝 기반 배양 시스템의 정교화와 실효성 검증을 통해 농업 현장에서의 활용도를 높일 수 있는 시스템을 설계하고 구현할 계획이다.

#### 참 고 문 헌

- [1] 이창훈, 황승국, “스마트팜 재배를 위한 복합환경 제어시스템 설계”, 한국지능시스템학회 논문지, 제32권, 제2호, pp. 110-116, 2024. 4.
- [2] 주원균, 송유진 외3명, “빅데이터와 인공지능을 활용한 미래농업 활성화 방안에 대한 연구”, 한국콘텐츠학회지, pp. 379-380, 2022. 7.
- [3] 이정민, 김용렬, 최재현, “에그테크(AgTech)의 성장가능성과 활성화 방안, 한국농촌경제연구원, pp. 205-234, 2022. 1.
- [4] 이종열, 김병만, “빅데이터 모델을 활용한 노지 스마트팜 생육환경 모니터링 시스템”, 한국정보보통신학회, 제28권 제 1호, pp. 47-53. 2024. 1.
- [5] 김홍직, 이원복, 이승호, “딥러닝 예측 결과 정보를 적용하는 복합 미생물 배양기를 위한 딥러닝 구조 개발”, 한국정보통신학회, 제 27권, 제1호, pp. 116-121, 2023. 5.
- [6] 이창훈, 황승국, “스마트팜 재배를 위한 복합환경 제어시스템 설계”, 한국지능시스템학회지, 제 32권, 제2호, pp. 110-116, 2022. 4.



# 생성형 AI(가상 디지털 휴먼)와 사용자의 커뮤니케이션이 사용자 경험 및 이용 의도에 미치는 영향을 분석

관가우\*, 비팅팅\*\*, 순 샤오위\*\*\*, 왕사요\*\*\*\*

## Analyzing the Impact of Communication Between Generative AI (Virtual Digital Human) and Users on User Experience and Usage Intention

GUAN JIAYU\*, BI TINGTING\*\*, SUN XIAOYU\*\*\*, and WANG SIYAO\*\*\*\*

### 요 약

3D 모델링 기술과 대화형 AI 알고리즘의 발전으로 인해, 디지털 휴먼은 다양한 분야에서 활용되고 있으며, 신한은행, 농협은행 등 주요 은행에서는 가상 은행원이 등장하고 있다. 그러나, 디지털 휴먼과 관련된 연구는 주로 외형적 측면에 집중되어 있으며, 로봇의 의인화 과정에서 중요한 퍼소나 설정에 관한 연구는 부족한 상황이다. 본 연구에서는 AI 디지털 휴먼에 대한 사용자 경험을 조사하기 위해 실험을 수행하였다. 연구 결과, 사회적 실재감과 유용성은 시나리오와 의인화 수준 간에 상호작용 효과를 보였다. 친밀감, 신뢰도, 인지된 지능에서는 명확한 상호작용 효과가 나타나지 않았지만, 일정한 경향성을 확인할 수 있었다.

### Abstract

With the advancement of 3D modeling technology and conversational AI algorithms, digital humans are being utilized in various fields, and virtual bank tellers have begun to appear in major banks such as Shinhan Bank and Nonghyup Bank. However, research related to digital humans has mainly focused on their appearance, and there is a lack of studies on persona design, which is important in the process of anthropomorphizing robots. This study conducted experiments to examine user experiences with AI digital humans. The results showed that social presence and perceived usefulness had interaction effects with scenario and levels of anthropomorphism. Although there were no significant interaction effects on intimacy, trust, and perceived intelligence, certain trends were observed.

### Key words

AI,, Digital Human, Anthropomorphism

---

\* 국립부경대학교, 1886469@naver.com(주저자),

\*\* 국립 부경대학교, katherine0518@naver.com(교신저자),

\*\*\* 국립 부경대학교, sunxiaoyu981017@naver.com(교신저자),

\*\*\*\* 국립 부경대학교, siyao920629@naver.com(교신저자)

※ 이 논문은 2024학년도 부경대학교 국립대학육성사업 지원비(PhiNX 보호학문 차세대육성)에 의하여 연구되었음.

## I. 서 론

AIGC는 인공지능 기술을 이용해 자동으로 콘텐츠를 생성하는 방식이다. 이를 통해 텍스트, 이미지, 비디오, 오디오 등 다양한 형태의 콘텐츠를 분석하고 이해하며 생성할 수 있어, 콘텐츠 창작의 효율성과 품질을 크게 향상시킬 수 있다. AIGC의 발전에 따라, 기업들은 AI 디지털 휴먼 기술을 도입하여 사용자에게 보다 편리하고 몰입감 있는 서비스를 제공하기 시작했다.

AI 디지털 휴먼(Virtual/digital human/avatar)은 디지털화된 외형을 가진 가상 인물을 말하며 일반적으로 특정 콘텐츠의 대화를 위해 훈련된 인공지능 캐릭터로 특정 이미지 능력, 인식 능력, 표현 능력 및 엔터테인먼트 상호 작용 능력이 있다. 또한 디지털 휴먼은 가상 인간, 비추얼 인플루언서, 메타 휴먼 등 분야에 따라 다른 명칭으로 사용되고 있다. 이머전 리서치(EmergenResearch)에 따르면 디지털 휴먼 아바타의 세계 시장성장도는 2021년 이후 연평균 36.4%로 성장하여 2030년에는 약 730조 원까지도 커질 것으로 예상된다(한상열, 2022).

디지털 휴먼이 이전의 로봇 에이전트들과 동일한 발화 방식을 사용한다면 사람의 외형뿐 아니라 행동양식까지 모방하는 디지털 휴먼의 장점을 제대로 살리기 어렵다. 실제로, 사람들은 AI 에이전트와의 대화에서 선후 맥락을 이해하여 대화를 진행할 때 인공지능을 더 매력적으로 느낀다. 사람처럼 대화의 맥락을 이해하여 내재되어있는 의미, 즉 감성적인 부분을 읽어낼 수 있는 에이전트를 선호하는 것이다. 따라서, 본 연구에서 AI 디지털 휴먼이 사용자 감정을 정확하게 인식하고 반응할 수 있는 감정(신뢰도, 만족도)이 사용자 경험 및 이용 의도에 무슨 영향을 미치는지 탐구하겠다.

## II. 인공지능 기술 생산이 주도하는 AIGC 시대

AIGC는 텍스트, 오디오, 이미지 등의 기존 콘텐츠를 활용하여 유사한 콘텐츠를 새로 만들어내는 인공지능(AI) 기술이다. 생성형 인공지능은 단순히 콘텐츠의 패턴을 학습하여 추론 결과로 새로운 콘

텐츠를 만들어내는 것을 넘어 콘텐츠의 생성자와 만들어진 콘텐츠를 평가하는 판별자가 끊임없이 서로 대립하고 경쟁하며 새로운 콘텐츠를 생성해내는 기술이다. 특히, 이미지 분야에서는 특정 작가의 화풍을 모사한 그림으로 사진을 재생성하거나 가짜 인간 얼굴을 무제한으로 생성하여 쇼핑, 영화 등의 산업에서 활용한다. 음성 분야에서는 특정 장르의 음악을 작곡하거나 특정 노래를 원하는 가수의 음색으로 재생성하는 등으로 활용한다. 상업적 혹은 사회적인 쟁점으로 가장 대중적으로 알려진 생성형 인공지능으로는 ‘인물 합성 기술(deepfake)’ 이 있다(출처: 네이버 지식백과). AI 디지털 휴먼이 사람과 유사한 속성을 지닐수록 더욱 효과적인 상호작용이 가능할 것으로 예상된다. 이는 정서적 차원에서 카치오포 외(Cacioppo et al., 2006) 연구처럼 사용자가 AI 디지털 휴먼과 지속적인 상호작용을 통해 사회적 유대관계를 형성하여 직·간접적으로 AI 디지털 휴먼 이용에 중요한 역할을 할 것이다. 즉 AI 디지털 휴먼 사용자는 비인간적 대상물에게 감정, 의지 등 인간 고유의 속성을 부여하는 의인화, 준사회적 관계 형성으로 현실적 존재감을 인식하는 사회적 실재감, 그리고 AI 기술 기반의 개인화를 통해 AI 디지털 휴먼에 대한 인식을 형성할 것이다. 이는 AI 디지털 휴먼 사용자와 커뮤니케이션을 통해 이용에 중요한 역할 수행을 기대할 수 있다. 따라서 본 연구는 AI 디지털 휴먼과 사용자 간의 커뮤니케이션이 실제 AI 디지털 휴먼과 어떠한 관계가 있는지 살펴보고자 한다.

## III. AI 디지털 휴먼의 활용 사례

최근의 디지털 휴먼에 관한 기술적 및 활용 중심의 논문들은 인간을 모방하고 인간과 유사하게 보이도록 하는 기술에 중점을 두고 있다. 이 분야의 중심 키워드로는 ‘사람 같은 외형’, ‘3D 모델’, ‘인간 언어’, ‘사람 행동’ 등이 있다.

AI 디지털 휴먼은 다양한 분야에서 활용되고 있다. 중국 방송사 신화사는 2019년에 ‘AI 합성아나운서-신소호’을 공개하여 뉴스 방송에 AI를 도입했다. 교육 분야에서는 인기 있는 유튜버 ‘도티’가 AI튜터



로 활용되어 아이들에게 친절하게 학습 지원을 하고 있다(이태수, 2021). 대기업들도 AI 디지털 휴먼을 관광업에 도입하였다.

디지털 휴먼 정의	중심 키워드	저자
아바타 개념에서 발전된 형태로 사실적인 움직임이 가능한 3D 휴먼 모델	사실적 움직임, 3D 모델	김세영·허정윤 (2021)
사람의 외형을 모방하고 실제 사람의 행동 양식을 묘사하여 사람의 역할 대체	사람 외형 모방, 사람의 행동 묘사	서영호·오문석·한규훈(2021)
인간의 모습과 언어, 행동을 사실적으로 묘사한 3D 인간 모델	인간 모습, 인간 언어, 인간 행동, 3D 모델	한상열 (2021)
실제 인간의 모든 면을 컴퓨터가 모방하여 만들어낸 것	인간의 모든 면을 모방	구지연·강명진·이상준(2022)
인간처럼 보이고 행동하는 컴퓨터 생성 3차원 AI 캐릭터	인간같은 외형, 행동, 3D AI 캐릭터	서영호(2023)

그림 1. 디지털 휴먼 정의 (출처: 박성수·강태임 (2024))



출처: 중국 국가 박물관

이러한 AI 디지털 휴먼들은 사용자의 요청에 따라 언어와 비언어적인 대화를 포함한 다양한 서비스를 제공하며 발전하고 있다. 각 분야에서 AI 기술의 활용으로 인해 사람과의 상호작용에서 실질적인 변화와 혁신이 이루어지고 있다.

#### IV. 결 론

생성형 AI와 상호작용한 사용자들이 기존의 비생성형 AI와 상호작용한 사용자들보다 더 긍정적인 사용자 경험을 보고하였으며, 생성형 AI와의 상호작용은 사용자 만족도를 높일 것으로 나타났다. 이는 더 자연스럽게 인간적인 상호작용이 가능한 생성형 AI의 특징 때문이다. 생성형 AI와의 긍정적인 상호

작용은 사용자들의 향후 이용 의도를 높일 것으로 나타났다. 사용자들은 더 유익하고 만족스러운 경험을 통해 해당 시스템을 계속 사용하려는 경향을 보일 것이다.

#### V. 연구의의

생성형 AI가 기존의 비생성형 AI보다 사용자 경험을 향상시킬 수 있는지에 대한 실증적 증거를 제공한다. 이는 AI 기술 개발에 있어 중요한 피드백이 될 수 있고 사용자 특성에 따른 차별화된 경험을 분석함으로써, 다양한 사용자 그룹을 고려한 더 포괄적이고 사용자 중심적인 AI 시스템 디자인에 기여할 수 있다. 또한 생성형 AI와의 상호작용이 사용자 만족도와 이용 의도를 높이는 데 기여한다는 점에서, 기업들은 이 기술을 활용한 상업적 전략을 세울 수 있다

#### 참 고 문 헌

- [1] 네이버 지식백과: 생성형 인공지능 [generative AI, 生成型-] (IT용어사전, 한국정보통신기술협회).
- [2] 구지연, 강명진, 이상준. (2022). 디지털 휴먼의 특성 및 체험 경험이 이용자 체험 만족에 미치는 영향. <한국EA학회 학술발표논문집>, 2022(11), 97-106.
- [3] 김세영, 허정윤. (2021). 디지털 휴먼의 페르소나에 영향을 미치는 요소로서 패션 스타일 분석. <한국디자인학회 학술발표대회 논문집>, 5, 108-109.
- [4] 서영호. (2023). 디지털 휴먼의 제작 기술과 현황. 방송과 미디어, 28(3), 55-66.
- [5] 오문석, 한규훈, 서영호. (2021). 메타버스를 위한 디지털 휴먼과 메타휴먼의 제작기법의 분석. <한국디자인리서치>, 6(3), 133-142.
- [6] 윤영주. (2022). “AI 뱅커입니다” 은행 모바일 앱에 디지털 휴먼 최초 등장. (2022.8.4.), <https://www.aitimes.com/news/articleView.html?idxno=146177>.

# AI 부활 기술에 대한 대중의 반응: 감정적 분석 연구

비팅팅\*, 순 샤오위\*\*, 관가우

## Public Reactions to AI Resurrection Technology: An Emotional Analysis Study

BI TINGTING\*, SUN XIAOYU\*\*, and GUAN JIAYU

### 요약

본 연구는 현대 사회에서 인간과 기계의 감정적 상호작용이 확산됨에 따라 AI 부활 기술이 대중의 감정적 반응에 미치는 영향을 탐구하고자 한다. AI 부활 기술은 사망한 사람의 성격, 말투, 행동 등을 재현하여 유족이 디지털 아바타와 지속적으로 소통할 수 있도록 함으로써 정서적 위안을 제공할 수 있는 잠재력을 지니고 있다. 그러나, 이 기술은 윤리적, 철학적 논쟁을 불러일으키며 인간 본질과 기억의 의미에 대한 질문을 제기한다. 본 연구는 영화 <원더랜드>와 현실에서의 AI 부활 기술이 대중에게 미치는 감정적 반응을 비교하고, 이들 반응의 차이가 기술 수용도에 미치는 영향을 분석하고자 한다. 이를 통해 AI 부활 기술에 대한 정서적 요인의 중요성을 강조하고, 기술 발전 및 정책 수립에 유용한 통찰을 제공할 것이다.

### Abstract

This study aims to explore the impact of AI resurrection technology on public emotional responses as emotional interactions between humans and machines continue to expand in modern society. AI resurrection technology allows for the recreation of a deceased person's personality, speech patterns, and behaviors, enabling bereaved individuals to maintain communication with a digital avatar, thereby holding the potential to provide emotional comfort. However, this technology also raises ethical and philosophical debates, prompting questions about the essence of humanity and the meaning of memory. This study seeks to compare the emotional responses of the public toward AI resurrection technology as depicted in the film \*Wonderland\* and in real-life applications and to analyze how these differing responses influence technology acceptance. Through this, the study emphasizes the importance of emotional factors in AI resurrection technology acceptance and offers valuable insights for the development of technology and policy-making.

### Key words

AI resurrection technology, Public emotional response, Human-machine interaction

\*부경대학교, katherine0518@naver.com, \*\*부경대학교, sunxiaoyu981017@naver.com, 1886469@naver.com

※ 이 논문은 2024학년도 부경대학교 국립대학육성사업 지원비(PhiNX 보호학문 차세대육성)에 의하여 연구되었음

## I. 서론

인간 - 기계 소통(Human-Machine Communication, HMC)은 인간과 기계 간의 의사 소통을 의미한다 (Guzman, 2018). 특히, 현대 사회에서 사람들은 기계와 대면하여 소통하는 현상이 점점 더 일반화되고 있다. 이러한 추세는 사람들의 소통 방식을 변화시킬 뿐만 아니라 인간과 AI 간의 감정적 상호작용에 대한 폭넓은 논의를 촉발시키고 있다. 인간과 기계의 상호작용은 단순한 명령 수행을 넘어선 정서적 교류를 포함하게 되었고, 이는 사람들에게 새로운 형태의 관계를 형성할 가능성을 열어주고 있다.

기술의 발전과 함께, 인공지능(AI)은 점점 더 인간의 특성을 모방하고 이해할 수 있게 되었다. 이러한 발전은 AI와의 감정적 상호작용을 통해 사람들이 위로를 받거나, 사회적 고립감을 해소하는 등의 긍정적 영향을 가져올 수 있는 가능성을 제시한다. 예를 들어, 백상기, 장병희, 김현(2020)의 연구에 따르면, AI 스피커 이용에 있어 기능적 매력과 사회적 매력은 의사사회적 상호작용과 만족에 긍정적 영향을 미치며, 만족은 이용의도에 정적인 영향을 준다<sup>2)</sup>. 또한, 최지혜와 노기영(2022)의 연구에 따르면, 인공지능 챗봇의 의인화가 준사회적 상호작용에 긍정적 영향을 미치며, 특히 사람 외모의 챗봇이 로봇 외모의 챗봇보다 인지된 유사성과 사회적 현존감을 더 높게 느끼게 한다. 이로써 인간 친화적인 인공지능 챗봇 설계를 위한 의미있는 통찰을 제공하고 있다는 점에서 의미가 있다.<sup>3)</sup>

이와 같은 변화 속에서 특히 주목할 만한 분야는 'AI 부활 기술'이다. AI 부활 기술은 사망한 사람의 데이터와 기억을 기반으로 AI가 그들의 성격, 말투, 행동 방식을 재현하는 기술을 말한다. 이 기술은 단순히 과거 데이터를 활용하는 것을 넘어, 고인의 디지털 아바타를 통해 가족 및 친구들이 지속

적으로 소통할 수 있도록 돕는다. 이는 사랑하는 사람을 잃은 이들에게 정서적 위안을 제공할 수 있는 잠재력을 지니고 있다. 그러나 동시에, 이러한 기술은 윤리적, 철학적 논쟁을 불러일으키며, 인간의 본질과 기억의 의미에 대한 깊은 질문을 제기한다. 예를 들어, 최근 주목받고 있는 최신 영화 '원더랜드'에서 묘사된 유사한 상황들이 있다. 또한, 최근 중국에서 핫이슈는 대만 유명 음악가 포소백(Tino Bao·包小柏)이 인공지능(AI)을 이용해 죽은 딸의 목소리를 '부활시켜 화제다.

이러한 배경을 바탕으로, 본 연구는 다음과 같은 연구 문제를 탐구하고자 한다:

1. 영화 '원더랜드'에서 묘사된 AI 부활 기술은 대중에게 주로 어떤 감정적 반응(긍정적, 부정적)을 일으키는가?
2. 현실에서 AI를 통해 고인을 '부활'시키는 사례는 대중에게 주로 어떤 감정적 반응(긍정적, 부정적)을 일으키는가?
3. 영화와 현실에서의 AI 부활 기술이 대중의 감정적 반응에 있어 어떤 차이가 있는가?

본 연구는 영화와 현실에서의 AI 부활 기술이 대중의 감정적 반응에 미치는 영향을 비교하고, 이러한 차이가 AI 부활 기술에 대한 수용도에 어떤 영향을 미치는지 살펴보고자 한다. 이를 통해 AI 부활 기술에 대한 이해를 심화하고, 기술 수용에 있어 정서적 요인의 중요성을 강조하며, 나아가 관련 기술 발전과 정책 마련에 유용한 통찰을 제공할 것이다.

## II. 이론적 배경

감정 분석은 사람들의 심리 상태를 이해할 수 있으며, 이는 다양한 분야에 영향을 미칠 수 있어 유용하게 활용될 수 있다. 감정 분석은 주로 긍정과 부정을 구분하는 감성 분석과 감정을 여러 범주로 세분화하는 감정 분석으로 나눈다. 감성 분석은 "행복하다", "즐겁다", "신나다" 등의 모든 긍정적인 감정을 하나로 취급하고, "화나다", "피곤하다", "우울하다" 등의 모든 부정적인 감정을 하나로 분류한다. 산업이나 조직에서는 주로 사용자의 피드백에 대한 긍정적 또는 부정적인 반응을 파악하는 데 사용된

2) 백상기, 장병희, 김현. (2020). 대학생의 AI 스피커 이용에 영향을 미치는 요인에 대한 탐색적 연구 : 의사사회적 상호작용이론과 기대일치이론을 중심으로. 한국언론정보학보, 101, 305-340, 10.46407/kjci.2020.06.101.305

3) 최지혜, 노기영. (2022). 인공지능 챗봇의 의인화가 챗봇과의 준사회적 상호작용에 미치는 영향 : 인지된 유사성과 사회적 현존감의 매개효과. 한국광고홍보학보, 24(4), 521-549, 10.16914/kjapr.2022.24.4.521

다. 그러나 감정을 긍정과 부정으로만 분류하면 사용자의 정확한 감정을 이해하기 어려울 수 있다. 따라서 감정 분석은 설정된 다양한 범주에 따라 감정을 더 세부적으로 구분하여 분석한다. 연구자들은 주로 "즐겁다", "화나다", "힘오하다", "슬프다" 등과 같은 다양한 감정을 6~8가지로 분류하여 개인의 다양한 기분과 감정을 파악할 수 있다<sup>4)</sup>. 따라서 원하는 여러 감정 범주를 텍스트에서 추출하려면 감정 분석이 아니라 감정 분석을 사용해야 한다.

최근 몇 년간 감정 분석 연구는 주로 핫한 토픽을 중심으로 이루어졌다. 예를 들어 금희조 등의 연구에서는 이태원 사건을 예시로 들어 감정 섬네일과 댓글을 분석하는 연구가 진행되었다<sup>5)</sup>. 또한, 양혜승의 연구에서는 이 전장연과 관련된 지하철 시위에 대한 네이버 뉴스 댓글도 분석되었다<sup>6)</sup>. 그러나 AI 부활 기술에 대한 감정 분석 연구는 아직 시행되지 않은 상태다.

이에 따라 본 연구는 AI 부활 기술에 대한 대중의 감정적 반응을 분석하기 위해 다양한 데이터를 수집할 계획이다. 연구는 2024년 1월부터 시작하여 2024년 6월 30일까지의 데이터를 대상으로 한다. 이 기간은 다음과 같은 이유로 선택되었다:

한국: 영화 '원더랜드'의 개봉으로 AI 부활 기술에 대한 관심이 급증하고 있다.

중국: 유명 연예인 포소백이 AI를 이용해 사망한 딸의 목소리를 부활시킨 사건이 화제가 되면서 AI 부활 기술에 대한 논의가 활발하다.

데이터 수집은 AI 부활 기술과 관련된 특정 키워드를 사용하여 이루어질 것이다. 검색 키워드는 'AI 부활', 'AI resurrection', 'AI reincarnation', 'artificial

intelligence revival' 등에서 영화 '원더랜드', 포소백 사건 등 관련 사례를 포함하여 설정할 계획이다.

뉴스 기사: 주요 뉴스 포털 사이트에서 AI 부활 기술과 관련된 기사의 댓글을 수집하여 대중의 반응을 분석한다.

소셜 미디어: 트위터와 같은 주요 소셜 미디어 플랫폼에서 관련 게시물과 댓글을 수집한다. 소셜 미디어는 대중의 즉각적이고 진솔한 반응을 파악할 수 있는 중요한 데이터 소스이다.

게다가 텍스트 기반 감정분석을 하기 위해서는 먼저 자연어 처리를 수행해야 한다. 자연어 처리에는 토큰화, 불용어 제거, 어근 동일화 등을 포함한다. 토큰화는 전체 문서나 단락 또는 한 문장을 단어로 분리하는 작업이고, 불용어는 감정분석에 기여하지 않는 불필요한 단어를 제거하는 과정이다. 어근 추출과 표제어 추출을 위해 접미사를 잘라 어근으로 변환, 과거형을 현재형으로 변환하는 절차가 있다. 그리고 감정 사전을 통해 정제된 텍스트 데이터의 감정을 라벨링하고 감정 값을 내어 텍스트의 감정을 파악한다. 수집된 데이터는 전처리 과정을 거쳐 분석에 적합한 형태로 변환될 것이다.

### III. 결 론

본 연구는 현대 사회에서 인간과 기계의 감정적 상호작용이 확산됨에 따라, AI 부활 기술이 대중의 감정적 반응에 미치는 영향을 심층적으로 탐구하였다. AI 부활 기술은 사망한 이의 성격, 말투, 행동을 재현하여 유족이 디지털 아바타를 통해 지속적으로 소통할 수 있게 함으로써 정서적 위안을 제공할 잠재력을 지니고 있다. 하지만, 이와 같은 기술은 윤리적·철학적 논쟁을 불러일으키며 인간 본질, 기억의 의미, 사후 존엄성 등에 대한 깊은 질문을 제기한다.

연구 결과, AI 부활 기술은 대중에게 긍정적인 감정과 부정적인 감정을 모두 유발할 수 있음을 확인하였으며, 영화 \*원더랜드\*와 실제 사례에서의 대중의 감정적 반응에 차이가 있음을 발견하였다. 영화에서는 AI 부활 기술이 주로 긍정적 반응을 유발하며 감동과 위안을 주는 것으로 나타났으나, 현실

4) Nandwani, P., Verma, R. (2021). A review on sentiment analysis and emotion detection from text, *Social Network Analysis and Mining*, 11, 81. DOI : 10.1007/s13278-021-00776-6

5) 금희조, 박윤정, 윤준섭, 박정민, 이세영, 임인재. (2023). 유튜브 감정 섬네일과 댓글의 반시민성 : 이태원 참사 이슈를 중심으로 기계학습 식별과 전통적 내용분석의 융합. *한국언론정보학보*, 122, 7-47, 10.46407/kjci.2023.12.122.7

6) 양혜승. (2023). 장에인을 향한 시선 : 전장연 지하철 시위 관련 네이버 뉴스 댓글에 대한 텍스트 마이닝 분석. *한국방송학보*, 37(6), 197-241, 10.22876/kab.2023.37.6.006

에서는 기술 수용도가 낮고 불안감이나 거부감 등 부정적 반응이 더 두드러졌다. 이는 영화 속 AI 부활 기술이 이상화된 형태로 표현되는 반면, 현실에서는 기술에 대한 윤리적 우려와 프라이버시, 인격의 재현 정확성 등의 문제가 부각되기 때문으로 볼 수 있다.

이러한 연구 결과는 AI 부활 기술에 대한 수용도에 감정적 요인이 중요한 영향을 미친다는 점을 시사하며, 정서적 상호작용을 촉진하는 방향으로 기술을 설계하고 발전시켜야 함을 의미한다. 특히, 정책 수립에 있어 대중의 감정적 반응을 고려하여 윤리적 기준과 법적 규제를 마련하는 것이 필수적이다. 본 연구는 AI 부활 기술과 같은 감정적으로 민감한 기술의 발전과 도입에 있어 대중의 정서적 요구와 윤리적 기준을 충족하는 것이 중요하다는 점을 강조하며, 이를 통해 사회적 논의를 촉진하는 계기가 될 것이다.

인지된 유사성과 사회적 현존감의 매개효과. 한국광고홍보학보, 24(4), 521-549, 10.16914/kjapr.2022.24.4.521

## 참 고 문 헌

[1] Nandwani, P., Verma, R. (2021). A review on sentiment analysis and emotion detection from text, *Social Network Analysis and Mining*, 11, 81. DOI : 10.1007/s13278-021-00776-6

[2] 김희조, 박윤정, 윤준섭, 박정민, 이세영, 임인재. (2023). 유튜브 감정 섬네일과 댓글의 반시민성 : 이태원 참사 이슈를 중심으로 기계학습 식별과 전통적 내용분석의 융합. *한국언론정보학보*, 122, 7-47, 10.46407/kjci.2023.12.122.7

[3] 백상기, 장병희, 김현. (2020). 대학생의 AI 스피커 이용에 영향을 미치는 요인에 대한 탐색적 연구 : 의사사회적 상호작용이론과 기대일치이론을 중심으로. *한국언론정보학보*, 101, 305-340, 10.46407/kjci.2020.06.101.305

[4] 양혜승. (2023). 장애인을 향한 시선 : 전장연 지하철 시위 관련 네이버 뉴스 댓글에 대한 텍스트 마이닝 분석. *한국방송학보*, 37(6), 197-241, 10.22876/kab.2023.37.6.006

[5] 최지혜, 노기영. (2022). 인공지능 챗봇의 의인화가 챗봇과의 준사회적 상호작용에 미치는 영향 :

# 가상 현실 (VR) 미디어 콘텐츠가 사용자 경험에 미치는 영향: 몰입감과 상호작용을 중심으로

순 샤오위\*, 비 텅팅\*\*, 관가우\*\*

## The Impact of Virtual Reality (VR) Media Content on User Experience: Focusing on Immersion and Interaction

SUN XIAOYU\*, BI TINGTING\*\*, and GUAN JIAYU\*\*

### 요 약

가상 현실은 체험적 매체로서 다양한 역할과 정체성을 구현할 수 있는 능력을 통해 개인과 가상의 장면을 연결하며, 다양한 장면 경험과 피드백을 제공하는 기능과 역할을 수행한다. 이를 통해 사용자는 체험 과정에서 새로운 학습과 사회화 과정을 완성할 수 있다. 본 연구는 가상 현실의 체험적 속성에 기초하여 그 활용 과정에서 나타나는 몰입감과 상호작용성을 탐구하고, 가상 현실이 직면하고 있는 주요 과제를 분석하고자 한다. 또한, 인간 상호작용 메커니즘이 인간과 가상 인물 간 상호작용 방식에 미치는 효과를 고찰함으로써, 가상 현실 연구와 실질적 응용에 필요한 이론적 근거를 제공하고자 한다.

### Abstract

Virtual reality, as an experiential medium, plays a role in connecting individuals to virtual scenes by enabling various roles and identities and providing diverse scene experiences and feedback. Through this process, users can achieve new learning and socialization during the experience. This study explores the immersive and interactive qualities that emerge during the utilization of virtual reality based on its experiential characteristics, and it aims to analyze the key issues currently facing virtual reality. Additionally, it examines the effects of human interaction mechanisms on the modes of interaction between humans and virtual avatars, aiming to provide a theoretical basis for future virtual reality research and practical applications.

### Key words

Virtual reality, Media, Virtual Human, Flow, Interactiveness

---

\* 국립부경대학교, sunxiaoyu981017@naver.com,

\*\* 국립부경대학교, katherine0518@naver.com, 1886469@naver.com

※ 이 논문은 2024학년도 부경대학교 국립대학육성사업 지원비(PhiNX 보호학문 차세대육성)에 의하여 연구되었음

## I. 서 론

2020년, 코로나-19로 인해 비대면 생활 방식이 확산되면서 온라인 쇼핑이 급격히 성장하였다. 또한, 4차 산업혁명의 부상과 함께 사물인터넷, 인공지능, 빅데이터 등 5G 통신 기술의 발전과 모바일 기기의 보급이 이루어지면서 온라인 거래와 실시간 방송을 결합한 라이브 커머스 방식이 주목받기 시작하였다 [1]. 라이브 커머스는 브랜드 또는 운영자가 온라인 커머스 플랫폼 또는 라이브 방송 플랫폼을 통해 소비자에게 제품의 정보와 기능을 소개하고 소비자는 생방송 과정에서 구매하는 전자상거래의 일종이다 [2]. 라이브 커머스가 주목받으면서 최근에는 VR 기술을 활용한 온라인 쇼핑과 라이브 커머스가 새로운 트렌드로 자리 잡고 있다. VR을 통해 소비자들은 더욱 생생하고 몰입감 있는 쇼핑 경험을 제공받으며, 실시간 상호작용으로 제품 정보를 더욱 자세히 확인할 수 있어 이러한 형태의 쇼핑 방식이 빠르게 확산되고 있다. 가상현실 쇼핑(VR Shopping)이라는 VR 기술을 통해 가상의 구매공간을 만들고 그 공간 내에서 판매자와 소비자 간 거래가 진행된다. VR 쇼핑 플랫폼은 온라인 쇼핑과 오프라인 쇼핑을 결합한 형태라는 점에서 사용이 편리하다는 장점을 지닌다[3]. VR 피팅을 통해 의류나 가구를 가상으로 체험함으로써 더욱 향상된 구매 경험을 제공받을 수 있는 장점이 있다. VR 피팅 기술은 소비자가 실제로 제품을 착용하거나 배치해 보는 것과 유사한 체험을 가능하게 하여, 구매 결정 과정에서의 만족도와 몰입감을 높이는 데 기여한다. 이 외에도 주목받고 있는 가상 현실 콘텐츠에는 가상 인물이 있다. 가상 인물(Virtual/Digital Human/Avatar)은 디지털화된 외형을 지닌 가상 캐릭터로, 특정 콘텐츠와의 상호작용을 위해 훈련된 인공지능 캐릭터를 의미하며, 일정 수준의 이미지 표현 능력, 인지 능력, 표현 능력 및 엔터테인먼트 상호작용 능력을 갖추고 있다. 인물 이미지 생성, 음성 생성, 애니메이션 생성, 음성·영상 합성 및 상호작용 기술 등의 5가지 모듈로 이루어진 일반 시스템 프레임워크 내에서 현재 가상 인물은 비상호작용형과 상호작용형으로 구분되며, 상호작용형 가상 인물은 다시 실제

인물 기반 구동형과 인공지능 기반 구동형으로 세분화된다. 시장 응용 측면에서 가상 디지털 인물은 주로 정체성형과 서비스형으로 나뉘며, 실제 인간을 대체하는 가상 진행자와 가상 아이돌이 현재 시장의 주요 관심사이자 개발의 중심이 되고 있다. 현재 매체는 그 특성인 언제 어디서나 존재할 수 있는 특성을 통해 다양한 사회 분야와 사회 작업에 영향을 미치고 있으며, 이는 개인과 조직, 개인과 매체, 사회와 매체 간의 사회적 교류와 연결의 변화를 초래하고 있다. 가상 현실(Virtual Reality, VR)은 종합 정보 기술의 발전으로서 매체 환경에 투입되며, 사회 생활의 최적화를 이룰 뿐만 아니라 인류에게 더욱 직접적이고 편리한 기술 서비스를 제공하여 사용자 경험을 향상시키고 있다. VR 기술은 인간이 세상을 인식하는 방식을 변화시키고 있다. 본 연구는 가상 현실의 체험적 속성에 기초하여 그 활용 과정에서 나타나는 몰입감과 상호작용성을 탐구하고, 가상 현실이 직면하고 있는 주요 과제를 분석하고자 한다. 또한, 인간 상호작용 메커니즘이 인간과 가상 인물 간 상호작용 방식에 미치는 효과를 고찰함으로써, 가상 현실 연구와 실질적 응용에 필요한 이론적 근거를 제공하고자 한다.

## II. VR 개념 및 특성

VR(Virtual Reality)은 컴퓨터 인터페이스 기능을 활용하여 가상의 공간을 만들고 사용자로 하여금 그 공간이 실제 존재하는 것처럼 느끼게 만들어 주는 기술이다. 사용자는 그 공간에서 보고 듣고 느끼는 것을 거의 실존인 것처럼 생각하게 된다[4]. 가상의 3D 환경에서의 실존적인 느낌은 그래픽의 생동감(Graphic Vividness)과 색깔의 생동감(Color Vividness) 및 3D의 진실성(Auth-Enticity)을 통하여 만들어진다. 이러한 시각적 표현을 토대로 제공된 현실감은 사용자로 하여금 가상공간에서의 정보를 더욱 쉽고 재미있게 이용할 수 있도록 돕는다[5]. 상호작용성(interactiveness)은 가상현실 콘텐츠를 사용자가 변경할 수 있도록 하는 수준을 의미한다. 객체를 변경한다거나 주변 환경을 변경하는 것을 예로 들 수 있다. 미디어가 현실 세계와 가상세계를 연결하

는 역할을 할 때 사용자들이 체험하게 되는 꿈 같은 현실이 콘텐츠 제작업체에 의해 일방적으로 주어지는 것보다는 사용자 역시 이러한 체험을 변경하거나 취향에 따라 만들어낼 경우 사용자 체험은 더욱 강화될 수 있을 것이다[6]. 가상 현실 기술의 특성으로서 상호작용성은 사용자의 몰입을 강화시킬 수 있으며 가상 현실을 사용자 의식의 확장이 될 수 있다고 볼 수 있다[7]. 몰입(flow)은 심리학에서 개발되어 스포츠, 쇼핑, 댄싱, 게임 등에서 광범위하게 적용되는 개념이다. 몰입은 컴퓨터 매개 커뮤니케이션, 컴퓨터 매개 환경 관련 분야에서 연구가 많이 되었으며 온라인 소비자들의 경험에 영향을 주는 것으로 고려되고 있으며 실제 세계에서의 경험과도 유사하다고 주장된다[8]. 사용자들은 몰입이 되면 시간의 흐름과 자기 주변의 변화를 알지 못한다. 몰입상태에서는 행위가 의식적 개입이 없이 내면적 논리에 따라 자연적으로 발생하게 된다[8].

### III. 가상 현실이 가져오는 몰입감과 상호작용성

가상 현실에서의 경험 과정에서 사람들은 상대적으로 정지 상태와 움직임 상태 두 가지 방식으로 존재한다. 전자는 가상 현실이 장치와 기술을 통해 신호를 인간의 감각에 전달하여 사람들이 가상 세계를 느낄 수 있게 하는 방식이다. 후자는 가상 장치의 기술을 이용하여 자신의 움직임을 포착하고 이를 가상 세계에 입력하는 방식으로, 이후 대규모 데이터 연산을 통해 출력 기술을 사용하여 변화된 가상 세계를 인식하게 된다. 이러한 기술적 측면은 사람들의 몰입감을 높이는 데 기여한다. 이러한 몰입감은 감각적으로 반영되며, 동일한 콘텐츠의 영상에서 관객들은 2D보다 3D를 선호하는 경향을 보인다. 이러한 생리적 몰입은 매체의 전파 효율성을 극대화하며, 예술 작품이 관객에게 내용을 더 빠르고 효과적으로 이해시키는 데 도움을 줄 수 있어 전파 효과를 증대시킨다. 가상 현실의 상호작용성은 여러 측면에서 나타난다. 사람들은 가상 현실 기술을 통해 온라인 공간에서 의류를 착용해 볼 수 있으며, 이는 온라인 쇼핑에서 경험할 수 없는 제한을 극복할 수 있게 해준다. VR 기술에서 상호작용적 시각

경험의 사실성과 직관성은 특히 중요하다. 눈은 인간이 외부 환경과 소통하는 주요 매개체이며, 정보 획득의 중요한 경로이기도 하다. 가상 현실은 사용자에게 가상 기술로 구축된 실제 공간을 제공함으로써, 사용자가 상품 및 기타 정보를 직관적으로 획득할 수 있도록 한다. 이러한 경험은 소비자에게 더욱 향상된 구매 결정을 가능하게 하며, 가상 현실의 상호작용성이 실제 쇼핑 경험을 한층 더 풍부하게 만든다. 한편, 가상 인물의 도입은 인간과 함께할 수 있는 역할을 창출하였으며, 인간과의 상호작용 과정에서 심리적 만족감과 행복감을 제공하게 된다. 가상 인물은 외형, 표정, 자세 등을 통해 사용자 요구를 충족시키며, 가상 세계와 현실 세계를 연결하는 다리 역할을 수행한다. 이를 통해 가상 인물과 실제 인간 간의 상호작용 목표가 이루어지게 된다. 이러한 상호작용은 사용자의 정서적 경험을 풍부하게 하고, 가상 인물의 존재가 인간의 사회적 상호작용을 확대하는 데 기여하고 있다.

### IV. 문제점

산업 발전의 관점에서 볼 때, 가상 현실과 가상 인물의 생산은 모두 인간을 중심으로 이루어진다. 고도의 현실감과 의인화된 외형을 통해 상호작용 경험을 제공하고 있다. 그러나 일부 사람들은 이러한 고강도의 의인화 및 인공지능화가 일자리 감소를 초래할 것이라고 우려하고 있다. 현재 인공지능 기술로 개발된 가상 인물은 사회적 정체성을 부여받아 진행자, 방송인, 아이돌 등으로 활동하고 있다. 그러나 취업이 어려운 시점에서 과학 기술이 점령한 직업이 매우 많아지면서 구직의 어려움이 더욱 심화되고 있다. 한편, 윤리적 측면에서도 논란이 존재한다. 인간은 가상 세계와 가상 인물에 대한 수용 정도에 대해 고민하기 시작하였다. 예를 들어, 고유사한 의인화는 "불쾌한 골짜기" 효과를 초래한다. 불쾌한 골짜기 이론에 따르면, 로봇의 외형이 실제 인간과 유사할수록 관객의 감정 반응이 긍정적이지만, 유사성이 특정 임계점에 도달하면 사람들은 오히려 불쾌감과 거부감을 느끼게 된다. 따라서 가상 인물이 사용자에게 호감을 주되 불쾌감을 주지 않



는 수준에 도달하는 것이 현재 고려해야 할 문제이다. 몰입감과 관련해서도 많은 사람들이 지나치게 몰입할 경우 가상 현실과 현실 세계를 구분하지 못하게 되어 인간의 신체적 및 정신적 건강에 부정적인 영향을 미칠 수 있다는 우려를 하고 있다.

## V. 결 론

인공지능 기술의 발전과 응용은 피할 수 없는 시대적 흐름이다. 그러나 이러한 기술의 활용 과정에서 인간은 기술의 본질로 돌아가, 기술이 인류에 미치는 영향을 먼저 숙고해야 한다. 기술은 지속적으로 발전하고 있지만, 인간은 항상 기술의 수혜자이자 피드백 제공자로 남아야 한다. 맥 루한 미디어는 인간의 확장이라는 말처럼 매체 기술의 발전에 있어 핵심은 매체를 어떻게 이해하는가가 아니라, 매체를 사용하는 사람과 사람과 매체 간의 상호작용을 어떻게 이해하는가에 있다. 가상 현실 및 관련 기술에 대해서는 그 핵심 기능과 역할의 본질을 인식하는 것이 필수적이다. 인간과 가상 세계 간의 윤리를 신중하게 고려해야 한다. 인간과 가상 세계 간의 상호작용이 지속적으로 심화됨에 따라, 이러한 상호작용성과 몰입성이 긍정적이거나 부정적인 기준으로 평가될 수 있는 합리성에 대한 반성이 필요하다. 새로운 기술은 융합과 도전을 가져오며, 이를 어떻게 응용할지는 미래의 인류가 지속적으로 혁신과 개인 정보 보호 및 합리성에 대한 균형을 맞춰야 하는 과제이다. 인간 중심의 핵심 이념을 통해 도전에 더 잘 대응하고 사회에 봉사하며, 미래의 정보 기술 혁신을 맞이해야 한다. 이러한 혁신은 개인과 사회의 발전에 긍정적인 영향을 미칠 것이다.

## 참 고 문 헌

- [1] ZHU WENYU 쇼 호스트 특성이 가상현실, 몰입 및 구매의도에 미치는 영향에 관한 연구 : 중국 라이브 커머스 소비자를 대상으로 2024
- [2] 강박 쇼 호스트 및 라이브 커머스의 특성이 사용자의 행동 의도에 미치는 영향에 관한 연구 : 틱톡 이용자를 중심으로. 가천대학교 박사 학위

논문.2022.

- [3] 나운규, 정혜연 패션 VR 쇼핑 서비스의 가치기반 채택모델 개발 연구. e-비즈니스연구, 23(5), 157-175, 2022.
- [4] 정연승, 안수진 VR 쇼핑의 소비자 수용도 분석 연구. 국제융합학술지, 33(1), 79-97. 019.
- [5] 윤남희, 이하경, 이윤정 VR 점포의 생동감이 점포 몰입과 쇼핑경험에 대한 만족에 미치는 영향. 한국의류학회지, 43(5), 559-572 2021.
- [6] 김태하, 박찬희, 차훈상, 위정현. 가상 현실 몰입도 및 지속적 활용의도에 대한 실증 연구. 한국게임학회 논문지, 22(5),2022.
- [7] Animesh, Animesh, et al. "An odyssey into virtual worlds: exploring the impacts of technological anspatial environments on intention to purchase virtual products", *Mis Quarterly*, pp.789-810, 2011.
- [8] Cheon, E. "Energizing business transactions in virtual worlds: an empirical study of consumers' purchasing behaviors", *Information Technology and Management*, 14(4), pp.315-330, 2013.
- [9] Wang. Y. W. "Research on the factors influenching mobile game immersion and intention of continuous use", Graduate School Chonnam National University, 2021.

# 일회성 시스템의 무기체계 신뢰도 확보를 위한 실사격 수량에 대한 연구

이형철\*, 김현수\*\*

## Study on the number of live fires to ensure the reliability of the weapon system of a one-time system

Hyung-Chul Lee\*, Hyun-Su Kim\*\*

본 논문에서는 무기체계 중 유도탄의 특성인 한번 사용을 목적으로 설계하는 일회성 시스템과, 장기간의 저장기간이라는 특성을 고려하여 신뢰도 확보를 위한 실사격 수량에 대한 연구를 하였다. 유도탄의 신뢰도 평가방법 중 임무 성공확률의 신뢰도는 주로 신뢰도 평가를 위해 수행하는 작업으로 표본의 개수 및 신뢰구간의 범위에 따라 유도탄의 신뢰도 수준을 결정한다.[1] 유도탄의 신뢰성 평가방법은 n회의 실험 중 r회의 고장 또는 기능적 결함을 관찰하는 방식의 이항분포 확률을 적용하는 것이 일반적이고, 본 논문에서는 이러한 방법으로 대표 유도탄 5개 탄종을 선정하여 각 탄종별로 연간 생산수량 및 총 생산수량을 고려하여 탄종별 신뢰성 평가를 위한 적정 시료수를 산정하여 제시하였다.[2] 이러한 방법을 적용하여 기대수명 도래한 유도탄의 부품단위 신뢰성 평가에서 합격한 유도탄과 동일한 로트에서 실사격 시험 표본을 획득하여 유도탄 신뢰성 평가 시 동일 수준의 최소 신뢰도 및 신뢰수준을 바탕으로 최소한의 샘플을 적용할 수 있는 가장 효과적이며 경제적인 시험방법이라 할 수 있다.[3] 즉 고가의 유도탄 비용을 고려 시 최소한의 수량으로 최적의 신뢰도를 확보할 수 있는 방법을 제시하였다.

Key words

guided missiles, live fires, reliability, one-time system

### 참고문헌

- [1] Lee. D. J, No K. H, Lee S. B, "A Preliminary Investigation Analysis Study for the Live Fire Test of the Guided Missile", The Report of Research Service for Defence Agency for Technology and Quality, pp. 1-273, 2022.
- [2] Lee. H, C, Lee M. H, Shin S. B, Jo K. J, "A Preliminary Investigation Analysis Study for the Live Fire Test of the Guided Missile", The Technical Report of Defence Agency for Technology and Quality, Vol. 22, No 7438, pp. 1-8, 2022.
- [3] Sherwin, E.R., "Analysis of one-shot devices", START(Selected Topics in Assurance Related Technologies), Vol. 7, No 4, 2000.

---

\*국방기술품질원, lh10237@gmail.com, \*\*국방기술품질원, hyun2005@naver.com

※ 본 연구는 국방기술품질원 자체과제로 수행 되었음

# 사무직 근로자의 업무 수행 내역 자동 기록 방법 연구

황교찬\*, 최규성\*\*, 김노을\*\*\*

## Study on the Automated Recording Method for Office Workers' Task Performance Process

KyoChang Hwang\*, GyuSeong Choe\*\*, NoEul Kim\*\*\*

### 요 약

본 논문은 사무직 근로자의 업무 수행 내역을 자동으로 기록하는 방법에 대한 연구이다. 근로자의 대부분은 업무 활동을 기록하고 보고한다. 이러한 기록은 업무 관리에 있어서 근로자의 업무 수행 과정을 확인하기 위한 척도로 사용된다. 다만, 과도하게 세분된 업무 보고는 해당 문서를 작성하는 데 필요한 시간 소모를 초래하게 된다. 또한 가시적인 결과물이 없는 업무이거나 결과물 도출까지 기간 길다면 해당 기간의 업무 활동 보고는 주관적일 수밖에 없다. 업무 활동 보고는 코로나19 이후 세계적으로 재택근무가 일반화되어 가고 있는 지금 더욱 필요성과 중요도가 올라가고 있다. 이에 본 연구에서는 객관적인 데이터를 기반으로 근로자의 업무 상태를 검토할 수 있는 수단을 제공함으로써 공정하고 효과적인 인사관리를 지원하고 근로자들의 업무 효율성을 높이는 데 기여하고자 한다.

### Abstract

This paper is a study of how to automatically record the work performance of office workers. Most workers record and report their work activities. These records are used as a measure to check the worker's progress in managing their work. However, overly granular work reporting can lead to time-consuming documentation. Also, if a task has no tangible deliverables or a long lead time, the reporting of work activity during that period is subjective. Business activity reporting has become increasingly necessary and important in the post-COVID-19 world, where working from home is becoming more common. This study aims to contribute to fair and effective human resource management and improve worker efficiency by providing a means to review the status of workers' work based on objective data.

### Key words

Performance recording, Activity reporting, Remote work, Task Process

---

\* (주)위크스타일 대표이사, eekdro@gmail.com

\*\* (주)위크스타일 선임연구원, dev.tigermap@gmail.com

\*\*\* (주)위크스타일 연구원, commcencer@gmail.com

※ 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획 평가원의 지원을 받아 수행된 연구임(No.2022-0-00664, EX(Employee eXperience) 기반의 확장성이 보장된 비대면 공동업무 플랫폼 기술 개발

## 1. 서 론

코로나19 이후에도 재택근무는 지속적으로 확대되어 재택과 상주 근무의 장점을 혼합한 하이브리드(hybrid) 형태의 근무가 활성화 될 것으로 예상된다.[1] 이로 인해 비대면 상황에서도 업무 활동을 관리하는 중요성이 더욱 강조되면서 협업시스템을 적극적으로 도입하게 되었다.[2] 그 중에는 근로자의 PC에 특정 프로그램을 설치하여 업무 활동을 기록하는 방식도 존재한다. 광범위한 데이터 수집은 근로자의 프라이버시를 침해할 수 있는 문제가 있다.[3] 이러한 추세에 맞춰 프라이버시를 보호하면서 업무 활동을 자동으로 기록하는 시스템에 대한 연구는 지속적으로 발전하고 있다.[4]

본 논문에서는 기업에서 지정한 업무만을 근로자의 PC에서 모니터링하고 기록하여 프라이버시 문제를 해결하는 한편 업무 수행 과정을 자동으로 기록하여 업무 효율성을 높이는 방법을 제안한다.

## II. 본 론

### 1. 객관적 정보 수집 방법

사무직 근로자의 객관적인 데이터 수집을 위해 업무 중 PC에서 업무로 사용한 프로그램과 파일의 정보를 기록하고, 기록된 정보들을 기반으로 근무 수행 과정을 보고할 수 있도록 처리하고자 한다.

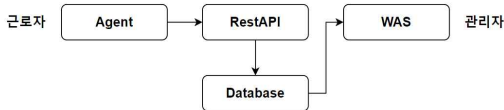


그림 1. 정보 수집 및 관리 시스템 구성도

그림 1은 객관적 데이터를 얻기 위한 시스템 구성으로 PC에 데이터를 수집하는 Application을(이후 Agent) 설치하게 된다. Agent는 OS(Operating System) 상에 구동되는 프로그램의 이름과 타이틀을 수집하고 사용 시간을 측정하게 된다. Agent에 의해 수집된 데이터는 Rest API 서버와 통신을 통해 Database에 기록되게 된다. Database에는 프로그램의 카테고리 및 업무 관련 여부를 구분하여 통계 처리가 가능

한 형태로 기록된다. 수집된 데이터를 기반으로 통계 데이터를 추출하여 가시화하는 WAS (Web Application Server)를 제공하여 근무자 및 관리자가 수집된 데이터를 검토할 수 있게 하였다.

### 2. 업무의 정의

업무는 기업마다 다를 수 있으며 디자이너, 프로그래머, 기획자 등 업무 직군별로도 구분될 수 있다. PC에서 사용하는 특정 프로그램을 업무용으로 지정했을 경우, 해당 프로그램을 업무 용도로 사용하고 있지 않아도 업무 수행으로 기록되는 결과가 나올 수 있다. 웹 브라우저의 경우 이메일, 원격회의, 정보 수집, 업무 관리 등 업무로도 사용할 수 있지만 개인적인 용도로도 사용이 가능하다. 이러한 상황에서 업무를 구분하기 위한 방법으로 본 논문에서는 PC에서 구동 중인 프로그램의 타이틀에서 특정 키워드를 추출하는 방식을 도입했다. 즉, 프로그램의 타이틀에 있는 문자열 중 업무로 지정된 키워드가 포함되어 있다면 업무로 기록하는 방식이다.

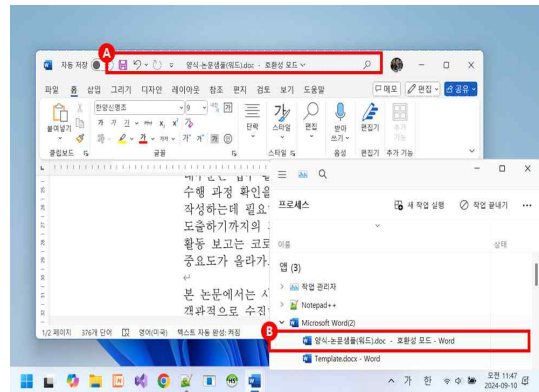


그림 2. 특정 프로그램의 타이틀 정보

그림 2. 'A'는 윈도우에서 MS Office의 Word가 구동 중인 상황이며 'B'는 윈도우 작업관리자 (Task Manager) 가 구동 중인 모습이다. 사용자가 가시적으로 볼 수 있는 프로그램은 모두 타이틀을 보유하고 있다.

특히 파일을 읽거나 편집할 수 있는 기능을 갖고 있는 프로그램의 경우 통상적으로 타이틀에 해당 파일명이 표기된다. 이를 통해 어떤 문서를 작업했는지도 확인이 가능하다.

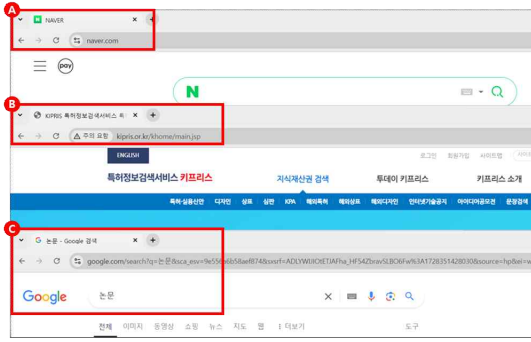


그림 3. 웹 브라우저의 타이틀 정보

그림 3. 을 보면 웹 브라우저의 경우 방문한 사이트의 정보가 타이틀에 표기된다. 그림 2. 'A'를 보면 naver.com 방문 시 타이틀도 NAVER로 변경되는 것을 확인할 수 있다. 'B'는 특허 검색 사이트를 방문했을 때의 타이틀이며 사이트에 대한 설명까지 타이틀에 포함되어 있다. 'C'는 Google에서 검색했을 때 해당 검색어가 타이틀에 표기되고 있는 것을 확인할 수 있다. 이와 같이 웹 브라우저의 타이틀 정보를 통해 웹 브라우저 내에서 어떤 업무를 수행하는지 판별이 가능하다.

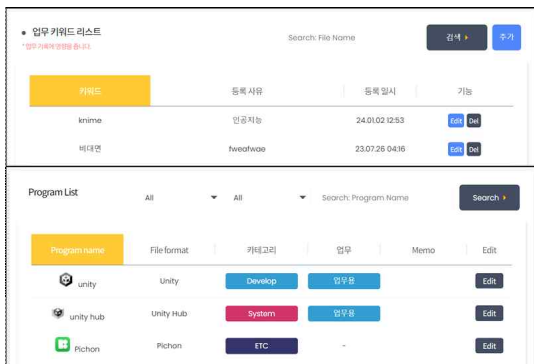


그림 4. WAS에서 키워드 및 프로그램 업무 지정

그림 4. 는 WAS 내에서 특정 키워드를 업무 지정하는 과정이다. WAS에서 키워드 및 프로그램을 업무로 등록하면 근로자 PC에서 업무 수행 중 발생하는 데이터를 선별할 수 있게 된다. Program List 에서는 특정 프로그램을 업무용으로 지정하고 카테고리를 지정할 수 있다. 이를 통해 업무용 프로그램 중에서도 어떤 카테고리에 속하는 프로그램을 주로 사용했는지도 확인이 가능하다.

### 3. 업무 정보 수집

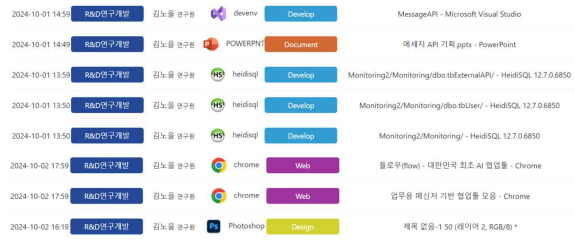


그림 5. WAS 에서 수집된 정보 확인

그림 5. 는 Agent를 통해 수집된 정보를 WAS에서 출력하여 제공하고 있다. 목록을 보면 사용된 프로그램 이름과 타이틀을 제공함으로써 어떤 프로그램을 언제 어떻게 사용했는지 확인할 수 있다.

### 4. 결과



그림 6. 업무 보고 일정 기록

그림 6. 는 그림 4에서 취합된 정보를 기반으로 일정을 기록하고 있다. 해당 일정을 효과적으로 표기하기 위해서 업무 내용을 간추릴 필요가 있다. 파일을 열기만 했는데 일정에 기록한다면 상당히 많은 내용이 일정에 기록될 수 있기 때문이다. 최소 10분 이상 실행되어 있는 프로그램 또는 파일을 기준으로 기록하여 불필요한 출력을 줄여 가독성을 높였다. 동일 프로그램이 반복적으로 열리거나 동일 문서를 반복적으로 작업하는 경우는 그림 6.의 A,B,C처럼 묶어서 표현하였다. 또한 프로그램의 카테고리별로 색상을 입혀 어떤 업무를 주로 했는지 한눈에 확인이 가능하게 하였다. 그림 5의 A는 오전 10시부터 11시 30분 사이에 업무를 나타내고 있다. 해당 시간에 근로자는 Microsoft Visual Studio로

HDF5\_Test 라는 프로젝트 파일을 열어서 작업한 것을 확인할 수 있다. B는 Chrome 웹 브라우저로 API에 관련된 내용을 검색한 것을 알 수 있다. 그림 5의 C는 PowerPoint로 ‘메시지 API 기획.pptx’ 파일을 열어서 작업을 한 내역이 자동으로 기록되어 있다.

이러한 자동 업무 기록의 객관적인 데이터를 활용한 업무 보고 체계는 비대면 상황에서도 고용주와 근로자 간의 상호 신뢰를 이어갈 수 있다.

### III. 결 론

본 연구에서 사무직 근로자의 업무 수행 과정을 자동으로 기록하는 방법을 제안하였다. 이러한 데이터는 근로자가 결과를 도출할 때까지의 여정을 객관적으로 판단할 수 있는 자료가 될 수 있다. 객관적인 데이터는 핵심 성과 지표(Key Performance Indicator, KPI)를 달성하지 못했을 때 빛을 발하게 된다. 실패의 원인이 업무의 과중인지 수행 방법의 실수 인지 객관적인 데이터를 기반으로 판단할 수 있게 된다. 업무에 있어서 성실은 결과만큼이나 중요하다. 국가 R&D에서도 연구 실패를 일반 실패와 성실 실패로 구분하고 있다. 도전적인 연구나 정량적 지표의 요구 수준이 상당히 높을 때, 성실하게 연구를 수행했음에도 정량적 지표를 완료하지 못했을 때 구제 방안으로 삼고 있다. 국가 R&D에서 연구노트를 요구하는 이유도 연구에 성실히 임하였는지 확인하기 위한 근거가 필요하기 때문이다. 차후 연구에는 사용한 파일의 내용을 AI로 분석하여 업무 수행을 과정을 기록하는 방법을 진행할 예정이다.

### 참 고 문 헌

[1] G. T. Oh, Changes in the way of working in the non-face-to-face era. Monthly Labor Review, November 2022 (Vol. 212, No. 11), 2022

[2] H. J. Seo, et al. The impact of information system quality of non-face-to-face collaboration system on

organizational performance. Global Management Journal, vol. 19, no.2. pp. 26-49, 2022

[3] K. C. Hwang, et al 3. A study on privacy protection measures during screen capture. Journal of Next Generation Convergence Technology, vol. 6, no. 7, pp. 1166-1174, 2022

[4] K. C. Hwang, et al 2. A study on Intelligent Collaboration Tools for Increasing Work Efficiency. Proceedings of the Korean Communications Society, 1193-1195, 2023

# 선택적 다중 시간 정보를 활용한 Mamba 기반 rPPG 추정

이준성\*, 박순찬\*, 류문욱\*

## rPPG Estimation Using Mamba with Selective Multi-Temporal Features

Jun Seong Lee\*, Soonchan Park\*, and Moonwook Ryu\*

### 요약

원격 광혈류측정법(rPPG)은 얼굴 영상으로부터 맥박 신호를 검출하는 비접촉 방식의 기술이다. rPPG는 얼굴 영상에서 광혈류 신호의 미세한 진폭을 추출하고, 시계열 데이터에서 의미 있는 맥락 정보를 효과적으로 해석하는 두 가지 주요 과제를 해결해야 한다. 기존의 연구들은 이러한 과제를 해결하기 위해 주로 CNN과 Transformer 기반의 접근법을 사용해 왔으나, CNN은 장기적인 의존성을 포착하는 데 한계가 있으며 Transformer는 높은 계산 복잡성을 요구하는 문제가 있었다. 본 연구에서는 이러한 한계를 극복하고 장기 의존성과 계산 복잡성 측면에서 모두 우수한 성능을 보이는 Mamba 기반의 새로운 접근법을 제안한다. 제안된 방법은 맥락 정보를 효과적으로 검출하기 위해 Mamba 구조를 활용하여 선택적 다중 시간 정보를 추출하고, 이를 활용하여 rPPG 신호를 추정함으로써 보다 정확한 rPPG 검출 성능을 달성한다. 제안된 방법은 MMPD 공개 데이터셋에서 3.15 MAE, 6.93 RMSE, 0.85 Pearson correlation coefficient를 보이며, 최신 방법들과 비교하여 우수한 성능을 입증하였다.

### Abstract

Remote photoplethysmography (rPPG) is a non-contact method to detect pulse signals from facial videos, with challenges in extracting subtle signal amplitudes and interpreting time-series data. While CNNs and Transformers have been used, CNNs struggle with long-term dependencies, and Transformers are computationally demanding. This study introduces a new Mamba-based approach that selectively captures multi-temporal information for more accurate rPPG detection, achieving 3.15 MAE, 6.93 RMSE, and a 0.85 Pearson correlation on the MMPD dataset, showing results comparable to recent methods.

### Key words

rPPG, Video-based physiological measurement, Mamba, State space sequence modeling

---

\* 한국전자통신연구원 jslee0708@etri.re.kr(교신), parksc@etri.re.kr, moonwook@etri.re.kr

※ 이 성과는 정부(과학기술정보통신부 및 대통령경호처)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (RS-2024-00443310).

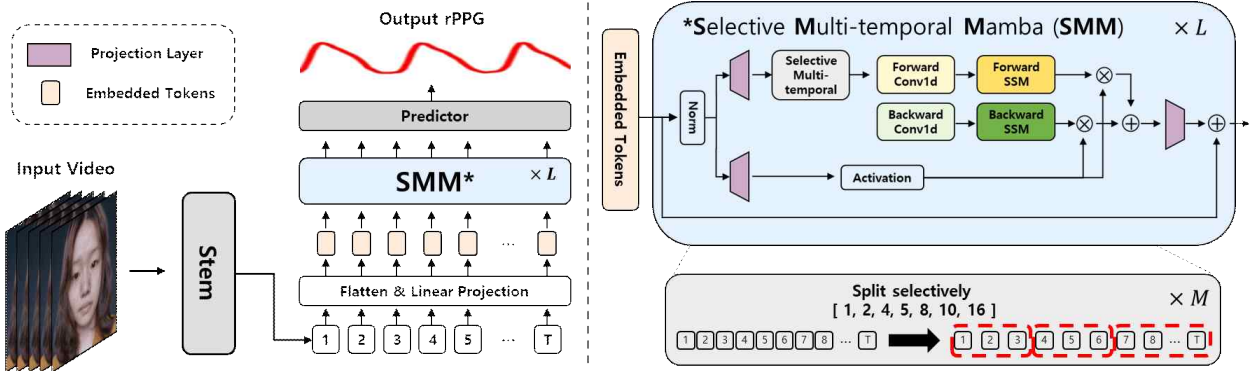


그림 1. 선택적 다중 시간 Mamba 기반 rPPG 검출 신경망 개요도

## I. 서 론

혈류량 맥박(BVP)은 심박수(HR)와 심박 변이성(HRV)과 같은 핵심 생체 신호를 추출할 수 있는 중요한 생리학적 신호로, 개인의 생리적 상태를 평가하는 데 중요한 역할을 한다. 광혈류측정법(PPG)은 비침습적인 방식으로 조직 내 혈류 변화를 측정하는 기술로, 전통적으로는 접촉 센서를 사용하여 BVP를 추출해 왔으나, 최근에는 얼굴 영상을 통해 비접촉 방식으로 BVP 정보를 추출할 수 있는 원격 광혈류측정법(rPPG)이 많은 주목을 받고 있다.

그러나 rPPG는 환경 변화와 잡음에 쉽게 영향을 받아, 신뢰할 수 있는 신호를 얻기 위해서는 얼굴 영상에서 광혈류 신호의 미세한 진폭을 정밀하게 추출하고, 시계열 데이터에서 의미 있는 맥락 정보를 효과적으로 해석해야 하는 두 가지 주요 과제를 해결해야 할 필요가 있다[1].

이러한 과제를 해결하기 위해 기존 연구들은 주로 CNN과 Transformer 기반의 접근법[2-7]을 사용해 왔으나, CNN은 장기적인 의존성을 효과적으로 포착하는 데 한계가 있으며 Transformer는 높은 계산 복잡성을 요구하는 문제가 있었다.

본 연구에서는 이러한 한계를 극복하고 장기 의존성과 계산 복잡성 측면에서 모두 우수한 성능을 보이는 Mamba 기반의 새로운 접근법을 제안한다. 제안된 방법은 Mamba 구조를 통해 선택적으로 다중 시간 정보를 활용함으로써 보다 효과적으로 전역 및 지역 특징들을 추출하고, 이를 활용하여 개선된 rPPG 추정을 달성한다.

## II. 선택적 다중 시간 Mamba 기반 rPPG 검출 인공지능망 프레임워크

제안하는 인공지능망의 프레임워크(그림 1)는 크게 3가지 부분으로 구성된다: (1) Frame Stem, (2) 선택적 다중 시간 Mamba(SMM), (3) rPPG 추정기이다.

Frame Stem은 입력 비디오를 임베딩하는 일련의 과정으로, 논문 [8]에서 제안된 방식을 차용했다. 이 과정에서는 7x7 크기의 2D 컨볼루션, 배치 정규화, ReLU 활성화 함수를 통해 주요 특징을 추출하고, 5x5 크기의 2D 컨볼루션으로 공간 정보를 통합한다. 또한, self-attention 메커니즘을 사용하여 피부 영역에서 더 강한 신호를 강조함으로써 rPPG 추정의 정확성을 높인다. Frame Stem의 마지막 단계에서는 글로벌 평균 풀링을 통해 임베딩된 신호  $x \in R^{N \times T \times C}$ 가 생성되며, 여기서  $N$ 은 배치 사이즈,  $T$ 는 입력 비디오의 길이,  $C$ 는 입력 비디오의 임베딩된 채널 크기이다.

SMM은 임베딩된 신호를 활용하여 전역 및 지역 정보를 추출한다. 이 과정에서 임베딩된 신호를 (1, 2, 4, 5, 8, 10, 16) 분할 선택지 중에서  $M$ 개의 분할을 선택적으로 활용한다. 예를 들어,  $M=4$ 일 때 선택적 분할 과정을 통해 (1, 2, 8, 16)이 선택된다면, 임베딩된 신호를 각각 1분할, 2분할, 8분할, 16분할로 나누어 Bidirectional Mamba를 통해 특징값을 추출하고, 이를 합산하여 최종 특징값을 도출한다. 1분할된 신호는 전역적인 정보를, 16분할된 신호는 지역적인 정보를 포함하며, 이러한 전역 및 지역 정



보를 통합함으로써 문맥 정보를 효과적으로 반영하는 특징을 생성할 수 있다. 중요한 점은 고정된 분할을 사용하는 대신 입력 임베딩에 따라 분할을 선택적으로 조정하여, 지역 정보가 더 중요한 경우 더 높은 분할을, 전역 정보가 더 중요한 경우 더 낮은 분할을 선택할 수 있도록 하는 것이다.

마지막으로, rPPG 추정기에서는 SMM으로부터 도출된 특징값을 바탕으로 rPPG 추정을 수행한다. rPPG 추정기는 간단한 MLP 레이어로 구성되며, 최종 출력은  $y \in R^{N \times T}$  이고, 여기서  $N$ 은 배치 사이즈,  $T$ 는 입력 비디오의 길이를 나타낸다.

### III. 실험 결과 및 결론

본 실험에서는 선택적 다중 시간 Mamba 기반 rPPG 검출 프레임워크의 구조적 변수  $L$ 과  $M$ 의 값으로  $L=24$ ,  $M=4$ 를 사용하였다. 학습 과정에서는 Adam 옵티마이저와 Negative Pearson Correlation Loss를 활용하여 모델을 최적화하였다.

제안된 방법의 성능 평가에는 MMPD 데이터셋을 사용하였다. MMPD 데이터셋은 Fitzpatrick 피부 유형 3-6을 가진 33명의 참가자가 네 가지 활동(정적 상태, 머리 회전, 말하기, 걷기)을 네 가지 조명 조건(LED-높음, LED-낮음, 백열등, 자연광)에서 수행한 데이터를 포함하며, 총 660개의 1분 길이 비디오(해상도 320x240, 프레임 레이트 30Hz)로 구성되어 있다. 실험에는 압축된 버전인 mini-MMPD를 사용하였다. 성능 평가는 평균 절대 오차(MAE), 평균 제곱근 오차(RMSE), 평균 절대 백분율 오차(MAPE), 피어슨 상관 계수, 신호 대 잡음비(SNR) 등 다섯 가지 평가지표를 사용하였다.

표 1은 MMPD 데이터셋에 대한 제안된 방법의 실험 결과를 나타낸다. 제안된 방법은 공개 데이터셋에서 3.15 MAE, 6.93 RMSE, 3.37 MAPE, 0.85 피어슨 상관 계수, 4.69 SNR을 기록하며, Transformer 기반 최신 접근법인 RhythmFormer와 유사한 성능을 보였다. 또한, 연산량 측면에서는 RhythmFormer 대비 약 1/3 수준으로 더 효율적인 성능을 달성하였다.

표 1. MMPD 데이터셋에 대한 제안된 방법과 최신 방법들의 결과 비교

Method	MMPD dataset				
	MAE	RMSE	MAPE	$\rho$	SNR
DeepPhys [2]	22.27	28.92	23.9	-0.03	-13.44
PhysNet [3]	4.80	11.80	4.74	0.60	1.77
TS-CAN [4]	9.71	17.22	10.36	0.44	-6.69
PhysFormer [5]	11.99	18.41	12.50	0.18	-9.51
EfficientPhys [6]	13.47	21.32	14.19	0.21	-9.25
RhythmFormer [7]	<b>3.07</b>	6.81	3.24	<b>0.86</b>	<b>5.46</b>
RhythmMamba [8]	3.16	<u>7.27</u>	<u>3.37</u>	0.84	4.74
Ours	3.15	<b>6.93</b>	<b>3.37</b>	<u>0.85</u>	4.69

결론적으로, 본 연구에서 제안한 선택적 다중 시간 Mamba 기반 rPPG 검출 프레임워크는 기존의 CNN 및 Transformer 기반 접근법의 한계를 극복하고, 연산 효율성과 성능 면에서 우수한 결과를 보여주었다. 특히, 다양한 시간 분할을 선택적으로 활용함으로써 전역 및 지역 정보를 효과적으로 통합하여 신뢰할 수 있는 rPPG 신호를 추정할 수 있었다. 이러한 결과는 제안하는 방법이 다양한 환경에서도 안정적인 생체 신호 추출에 기여할 수 있음을 시사하며, 향후 비접촉 생체 신호 측정 기술의 발전에 중요한 기여를 할 수 있을 것으로 기대된다.

### 참고 문헌

- [1] Lee, Jun Seong, et al. "Lstc-rppg: Long short-term convolutional network for remote photoplethysmography." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
- [2] Chen, Weixuan, and Daniel McDuff. "Deepphys: Video-based physiological measurement using convolutional attention networks." Proceedings of the european conference on computer vision (ECCV). 2018.
- [3] Yu, Zitong, Xiaobai Li, and Guoying Zhao. "Remote photoplethysmograph signal measurement

- from facial videos using spatio-temporal networks." arXiv preprint arXiv:1905.02419 (2019).
- [4] Liu, Xin, et al. "Multi-task temporal shift attention networks for on-device contactless vitals measurement." Advances in Neural Information Processing Systems 33 (2020): 19400-19411.
- [5] Yu, Zitong, et al. "Physformer: Facial video-based physiological measurement with temporal difference transformer." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
- [6] Liu, Xin, et al. "Efficientphys: Enabling simple, fast and accurate camera-based cardiac measurement." Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2023.
- [7] Zou, Bochao, et al. "Rhythmformer: Extracting rppg signals based on hierarchical temporal periodic transformer." arXiv preprint arXiv:2402.12788 (2024).
- [8] Zou, Bochao, et al. "Rhythmmamba: Fast remote physiological measurement with arbitrary length videos." arXiv preprint arXiv:2404.06483 (2024).

# 군중 객체 분석 및 이동 경로 기반 이상 행동 감지 시스템

이준성\*, 박순찬\*, 류문욱\*

## Crowd Object Analysis and Trajectory-based Abnormal Behavior Detection System

Jun Seong Lee\*, Soonchan Park\*, and Moonwook Ryu\*

본 연구에서는 군중 속 객체들을 추적하고 이동 경로를 기반으로 이상 행동을 감지하는 시스템을 제안한다. 군중 내 이상 행동 감지 기술은 공공 안전을 유지하고 대규모 인원 모임에서 발생할 수 있는 혼란을 예방하는데 중요한 요소이며, 이를 위해서는 개별 객체의 이동 패턴을 분석하고 다른 행동을 빠르게 감지하여 즉각적으로 대응하는 것이 필요하다[1]. 제안하는 시스템은 YOLO 모델[2]을 사용해 군중 내 객체를 탐지하고, ByteTrack[3] 알고리즘을 활용해 연속된 프레임에서 객체를 효과적으로 추적한다. 이후, 전체 객체의 평균 변위와 방위 값을 개별 객체의 변위 및 방위 값과 비교하여 군중 속에서 비정상적인 행동을 보이는 객체를 실시간으로 식별한다. 제안된 시스템은 UCSD Ped2 데이터셋[4]의 영상을 통해 성능을 검증하였으며, 군중 속에서 반대 방향으로 움직이거나 지나치게 빠르거나 느리게 움직이는 객체를 실시간으로 정확하게 검출할 수 있었다. 이를 통해 본 시스템은 대규모 군중 환경에서의 안전관리에 효과적으로 기여할 수 있음을 확인하였으며, 실시간으로 군중의 동태를 파악하고 이상 행동을 조기에 발견하여 공공 안전을 강화하고 잠재적인 위험 요소에 신속하게 대응할 수 있는 가능성을 보여준다.

### Key words

Crowd abnormal behavior detection, rPPG, Object detection, Object tracking, Movement trajectory analysis

### 참고문헌

- [1] Luo, Linbo, et al. "Detecting and Quantifying Crowd-level Abnormal Behaviors in Crowd Events." IEEE Transactions on Information Forensics and Security (2024).
- [2] Ultralytics, "YOLOv8: Ultralytics Object Detection Framework," <https://docs.ultralytics.com/>, 2023.
- [3] Zhang, Yifu, et al. "Bytetrack: Multi-object tracking by associating every detection box." European conference on computer vision. Cham: Springer Nature Switzerland, 2022.
- [4] M. S. Sabokrou, M. Fathy, M. Hoseini, and R. Klette, "Real-time anomaly detection and localization in crowded scenes," in 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 56-62, June 2015.

---

\*한국전자통신연구원 jslee0708@etri.re.kr(교신), parksc@etri.re.kr, moonwook@etri.re.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.RS-2024-00462874, 실외 혼잡 환경에서 위협상황 예측 및 선제대응을 위한 센티넬 AI (Sentinel AI) 시스템 기술 개발).

# 대규모 언어 모델 기반 프롬프트 엔지니어링을 활용한 OpenSCENARIO 자동 생성 연구

성기호\*, 조용현\*, 민경원\*

## A Study on Automated OpenSCENARIO Generation Using Large Language Model-based Prompt Engineering

Giho Sung\*, Yonghyeon Cho\*, and Kyungwon Min\*

자율주행 시뮬레이션 테스트를 위한 시나리오 생성은 안전하고 효율적인 자율주행 시스템 개발에 핵심적인 요소이다. 본 연구에서는 대규모 언어 모델(Large Language Model)인 ChatGPT와 Claude를 활용하여 OpenSCENARIO를 자동으로 생성하는 새로운 방법론을 제안한다. 프롬프트 엔지니어링 기법을 통해 자연어로 작성된 시나리오 설명을 OpenSCENARIO 형식의 XML 파일로 변환하는 과정의 자동화 가능성을 탐구하였다.

연구 방법으로는 차량의 Cut-in 시나리오를 대상으로 하여 프롬프트 템플릿을 설계하였다. 프롬프트 템플릿은 OpenSCENARIO 포맷의 구조화된 기술 방법을 고려하여, OpenSCENARIO의 주요 구성요소에 대한 설명과 함께 주요 구성 요소를 체계적으로 기술할 수 있도록 구성하였다. 생성된 OpenSCENARIO의 유효성은 esmini 시뮬레이터를 통해 검증하였으며, 이를 통해 생성된 시나리오의 실행 가능성을 확인하였다.

본 연구는 LLM과 프롬프트 엔지니어링을 활용한 OpenSCENARIO 자동 생성의 실현 가능성을 검증했다는 데 의의가 있다. 특히 자연어 설명만으로도 실행 가능한 OpenSCENARIO를 생성할 수 있음을 입증함으로써, 향후 다양한 자율주행 시나리오 생성 자동화 연구의 기반을 마련하였다. 이는 자율주행 테스트 시나리오 생성 과정의 효율성 향상에 기여할 것으로 기대된다.

### Key words

OpenSCENARIO, Large Language Model, Prompt engineering, Autonomous vehicle test

### 참고문헌

- [1] Vemprala, Sai H., et al. "Chatgpt for robotics: Design principles and model abilities." IEEE Access (2024).
- [2] Ding, Wenhao, et al. "Realgen: Retrieval augmented generation for controllable traffic scenarios." European Conference on Computer Vision (2024).

---

\* 한국전자기술연구원 모빌리티플랫폼연구센터, {gihosung@keti.re.kr, anorange7417@keti.re.kr, minkw@keti.re.kr}

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No. 2021-0-01415, 옛지연계 도심형 자율주행 서비스 검증을 위한 테스트 시나리오 생성 및 멀티에이전트기반 시뮬레이션 SW 기술개발)

# 동형암호 기반 이종 데이터 결합 및 분석에 관한 실증연구

최민호\*, 최호식\*\*, 신성철\*\*\*

## A Study of Integrated Analysis of Heterogeneous Data Using Homomorphic Encryption

Minho Choi\*, Hosik Choi\*\*, and Sung chul Shin\*\*\*

데이터 활용의 중요성이 증대됨에 따라 이종 데이터의 결합과 분석이 필수적으로 대두되고 있다. 그러나 현재의 가명처리 및 데이터 결합 신청 절차는 복잡하며, 민감한 개인정보 유출의 위험성으로 인해 실제 적용에 제약이 있다. 이러한 문제를 해결하기 위한 대안으로 동형암호(Homomorphic Encryption) 기술이 주목받고 있다. 동형암호를 활용하면 데이터의 원본 특성을 그대로 유지하면서도 암호화된 상태에서 데이터 결합과 분석이 가능하기 때문에, 데이터 유출의 가능성을 원천적으로 차단하여 보안성을 강화할 수 있다. 본 연구에서는 동형암호를 활용하여 택시 위치 데이터와 택시 미터기 요금 데이터를 암호화된 상태에서 결합하고, 이를 기반으로 한 통계분석을 수행하여, 동형암호를 활용함으로써 기존 절차보다 데이터 결합 과정이 간소화되고, 데이터 품질을 유지하면서 효율적인 분석이 가능함을 보여주고자 한다. 실증분석 결과, 동형암호 환경에서도 평문에서와 동일한 결과를 산출할 수 있음을 확인하였다. 이를 통해, 동형암호를 통한 이종 데이터 결합분석의 실용성과 그 장점을 제시함으로써 향후 개인정보 보호와 데이터 활용 간의 균형을 모색하는 데 기여할 것으로 기대된다.

### Key words

Homomorphic Encryption, Heterogeneous Data Integration, Combined Analysis

### 참고문헌

- [1] Kim, A., Song, Y., Lee, M., Cheon, J. H., & Park, K. (2020). Secure logistic regression based on homomorphic encryption: Design and evaluation. *IEEE Access*, 8, 226876-226888.
- [2] Khedr, A. M., Grotto, A., & Kuzu, M. (2023). Privacy-Preserving Data Integration Using Homomorphic Encryption. *IEEE Transactions on Dependable and Secure Computing*, 20(3), 1500-1512.
- [3] Zhang, Y., Yang, B., & Wang, X. (2023). Efficient Statistical Analysis over Encrypted Data in Cloud Computing. *IEEE Transactions on Cloud Computing*, 11(2), 345-358.

---

\* (주)크립토크 데이터사업개발실 연구원, minho8173@cryptolab.co.kr

\*\* 서울시립대학교 도시빅데이터융합학과 교수, choi.hosik@uos.ac.kr

\*\*\* (주)크립토크 데이터사업개발실 이사, scshin73@cryptolab.co.kr(교신저자)

※ 이 논문은 정부(국토교통부)의 국토교통 데이터산업 생태계 플랫폼 구축(R&D) 사업의 재원으로 국토교통과학기술진흥원의 지원을 받아 수행된 연구임 (NO.RS-2022-0144012).

# 반도체 공정 센서 제작 기술을 적용한 Closed Loop 제어 대전류 센서 제작 및 특성 분석 연구

이선하\*, 이형만\*

## Closed Loop Control Large Current Sensor Fabrication and Characteristics Analysis Research Using Semiconductor Process Sensor Manufacturing Technology

Sunha Lee\* and Hyungman Lee\*\*

대전류 센서는 최근 전기자동차 및 대전류 충전 시스템의 보급으로 인해 관심과 중요성이 증가하고 있다. 대전류 센서는 100A급 이상의 전류를 측정하기 위한 기술이며, Open Loop과 Closed Loop 제어 기술로 전류 센서를 제작하고 있다[1-2]. 본 연구에서는 고정확도를 구현할 수 있는 Closed Loop 제어 기반 대전류 센서 제작에 대한 연구를 수행하였다. 일반적인 대전류 센서는 홀센서를 적용한 정확도가 낮은 제품이 주를 이루고 있다. 이에 반해 고정확도를 갖는 전류 센서를 개발하기 위해서는 본 연구에서 제안하는 플럭스 게이트 방식을 적용한 Closed Loop 제어 기술을 적용한 센서 제작 기술 개발이 필요하다.

플럭스 게이트 방식을 적용한 대전류 센서를 제작하기 위해 반도체 공정을 적용하여 자성 박막으로 구성된 소자를 제작하였으며 이를 적용하여 플럭스 게이트 대전류 센서를 제작하고 특성을 평가하였다. 자성 박막을 사용한 미세 공정 센서를 제작하고 이를 적용하여 자성 코어로 제작된 대전류 센서 모듈을 제작하였다. 제작한 모듈을 적용하여 500A 대전류 센서 특성을 측정하고 분석하였다. 대전류 센서 정확도를 분석 결과 0.5% 이하의 고정확도를 갖는 센서 제작이 가능함을 확인하였다. 본 연구에서 개발한 고정확도 대전류 센서는 전기자동차와 관련 산업 분야에 파급효과가 클 것으로 기대된다.

### Key words

Current Sensor, Micro fluxgate, Semiconductor Fabrication, Magnetic, Thin Film

### 참고문헌

- [1] Costa, François & Poulichet, P. & Mazaleyrat, F. & Labouré, Eric. (2001). The Current Sensors in Power Electronics, a Review. EPE Journal. 11. 7-18. 10.1080/09398368.2001.11463473.
- [2] TOSHIKATSU S., UEDA R., KOGA K., "An AC and DC Current Sensor of High Accuracy" IEEE Transaction on Industry Applications 5 (1992) pp. 1087-1094.

---

\* Korea Electronics Technology Institute, lhm0703@keti.re.kr

\*\* This work was supported by the Technology development Program(S3207682) funded by the Ministry of SMEs and Startups(MSS, Korea).

# 드론촬영을 통한 태양광 발전소 태양전지 결함에 대한 AI진단 기술개발

허성필\*, 허철균\*\*

## Development of AI Diagnosis for Defects of PV in Solar Power Plants Using Drones

Sung-Phil Heo\*, Cheol Kyun Heo\*\*

### 요 약

국내 태양광 발전소에서 설치되어 운영중인 태양광 배터리는 시공 후 설비상태 점검과 주기적인 시설 점검에 따른 조기진단 대응보다는 발전량 모니터링에 의한 이벤트 발생사후 대응으로 유지보수가 운용되는 상황이다. 여러 개의 태양전지들이 붙어 있는 태양광 패널을 이용하여 전력을 생산하는 태양광 발전은 최근 신재생 에너지 기술로 빠르게 성장하고 있는 분야이다. 하지만 태양광 발전의 단점 중 하나인 불규칙한 전력 생산문제로 인해, 장비 및 패널 결함에 빠르게 대응하지 못하는 문제가 발생한다. 이에 본 논문에서는 드론 열화상 검사 AI 진단 체계로 전환하여 빠르고 정확하며 효율적인 시설물 관리 기술개발을 제안한다.

### Abstract

Solar PV module installed and operated in domestic solar power plants are maintained by monitoring power generation and responding to events after they occur, rather than responding to early diagnosis through periodic facility inspections and facility status inspections after construction. Solar power generation, which produces electricity using solar panels with multiple solar cells attached, is a rapidly growing field as a new renewable energy technology. However, due to the irregular power generation problem, which is one of the shortcomings of solar power generation, there is a problem in which it is difficult to respond quickly to equipment and panel defects. In this paper, we propose the development of a fast, accurate, and efficient facility management technology by switching to a drone thermal imaging inspection AI diagnosis system.

### Key words

Solar Power Plant, Defect, Drone, AI Diagnosis

## 1. 서 론

글로벌 태양광 시장은 2026년 기준 319조원에 달할 것으로 예상되고 있으며, 최근 기후변화에 따른 탄소제로, ESG, RE100 등의 영향으로 매년 평균 20% 이상 시장이 급격하게 성장하고 있으며, 국내

시장도 빠르게 성장하고 있다. 과거 국내 태양광에너지 발전비중(2017년 기준 설비용량 5.7GW)이 주요 선진국에 비해 낮은 상황이었지만, 최근 재생에너지 3020계획, 제5차 신재생에너지 기본계획 등 정부 보급정책에 따라 중대형 태양광발전소가 확대되고 있는 상황이다. 태양광 발전은 20~25년 이상 안

\* 국립강릉원주대학교, spheo@gwnu.ac.kr 주저자

\*\* (주)메타파스 calvinheo@metarpas.com 교신저자

정적인 전력을 생산하고 수익을 창출해야 하는 장기투자 사업으로, 태양광 발전설비가 확대되면서 태양광 PV시스템의 운영 유지보수(O&M: Operation & Maintenance)에 대한 수요도 증가하고 있다. 하지만 태양광 발전설비는 시간이 오래될수록 열화 및 환경 변화에 취약하기 때문에 모듈(Module), 스트링(String), 어레이(Array) 레벨의 결함이 자주 발생하여 발전효율이 떨어짐에도, 원격 관리가 어려우며 국내 태양광 O&M 업체는 유지보수 인력이나 문제 파악 또는 조기진단을 위한 지능형 시스템이 부족한 실정이다. 이에 본 논문에서는 태양광 발전소 시설물 점검에서 기존 수작업/현장중심으로 추진되는 체계를 항공(드론) 열화상 검사 AI 진단 체계로 전환하여 빠르고 정확하며 효율적인 시설물 관리 기술개발을 제안하고자 한다.

## II. 기존 태양광 발전소 배터리결함 진단방법 한계점

여러 개의 태양전지(PV: Photo Voltaic)들이 붙어 있는 태양광 패널을 이용하여 전력을 생산하는 태양광 발전은 최근 신재생 에너지 기술로 빠르게 성장하고 있는 분야이다. 하지만 태양광발전의 단점 중 하나인 날씨 등 주변환경의 영향을 받아 불규칙한 전력 생산문제로 인해, 장비 및 패널 결함을 빠르게 인식하지 못하고 즉각적으로 대응하지 못하는 문제가 발생한다. 태양전지(PV) 모듈은 태양광 발전에 있어서 가장 중요한 핵심부품으로 발전에 직접적인 영향을 미치고 있어 효과적으로 상태를 진단하고 파악하는 것이 매우 중요하다. 그린뉴딜, 탄소제로, RE100 충족 등을 위해 태양광 발전소의 국내 신규설비는 매년 4GW이상 급격하게 증가하고 있지만 발전소의 50%를 차지하고, 태양광 발전에서 가장 중요한 태양전지(PV) 모듈에 대한 대부분에 대한 점검은 아직까지도 전문인력이 계측기와 IR카메라를 사용하고 있다. 따라서 점검시 고전압 감전사고의 위험이 높고 많은 시간과 비용이 소요되고 있다.

그림 1.과 같은 인력기반의 점검 및 아날로그 방식의 결과 보고서는 전문 지식이 부족한 이해관계자들에게 난해하여 문제점을 적시에 해결하지 못함으로 인해 발전효율 저하로 이어진다.



그림 1. 기존 태양광 발전소 배터리결함 진단 방법

## III. 드론촬영 태양광 발전소 배터리결함 AI진단

드론을 활용하여 진단 대상 발전소를 정밀 항공 촬영 모델링을 수행하고 이를 기반으로 드론의 비행경로를 설정한다. IR이미지를 촬영 후 AI 머신러닝 기술을 이용하여 태양배터리 결함을 진단 및 검출하고 결함 위치를 파악하고 불량 태양전지모듈을 특정한다. 제안한 AI진단 방법은 발전소 진단/점검 시 감전사고와 같은 재해를 사전에 예방하고 점검 시간을 2배 이상 줄임으로써 유지보수 비용을 50% 이상 절감할 수 있다. 이러한 진단 및 유지보수에 대한 요구사항은 2017년부터 IEC-62446:3 표준으로 제정되어 미국, 유럽 등 태양광발전소에 도입되어 점검이 이루어지고 있다.



그림 2. 드론촬영 태양광 전지결함 AI진단 방법

태양광 발전소를 드론으로 진단시 기본적으로 IEC-62446:3 표준요구사항을 준수함으로써 검사결과 신뢰성을 높이고 점검시 판단에 필요한 적절한 데이터 및 이미지를 확보할 수 있다. 발전소를 촬영하기 위해서는 일사량이 최소 600W/m2 이상, PV모듈은 열화상카메라(IR: InfraRed Camera)로 촬영시 결함 등을 확인할 수 있다.

AI 진단방법은 기존 전문가가 드론촬영한 태양전지 이미지 점검시 소요되는 시간과 비용을 줄일 수 있고 보다 정확한 진단을 하는데 도움을 제공한다.



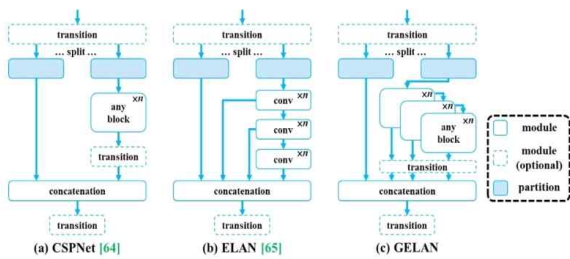


그림 3. YOLO v9 아키텍처

태양전지 및 결함을 학습하기 위해 PGI (Programmable Gradient Information) 기능을 사용하여 신경망의 심층레이어에서 발생할 수 있는 정보손실과 그래디언트 불안전성 문제로 손실되는 것을 최소화하고 학습 안전성을 확보하였다.

표 1. 플랫폼이 지원하는 태양전지 결함

결함 유형	결함 구분		결함 구분	
	세부 유형	결함 유형	세부 유형	결함 유형
Cell	Cell Low	Hot Spot Multi	Hot Spot Multi Low	
	Cell Medium		Hot Spot Multi Medium	
	Cell High		Hot Spot Multi High	
Cell Multi	Cell Multi Low	Module		
	Cell Multi Medium	Shading		
	Cell Multi High	Soiling		
Diode		String		
Diode Multi		Vegetation		
Hot Spot	Hot Spot Low			
	Hot Spot Medium			
	Hot Spot High			

플랫폼에서는 검출된 결함을 실제 촬영된 정사이미지를 사용하여 Digital Twin 방식으로 태양전지 모듈별로 결함의 종류를 표시하고 관리할 수 있다.

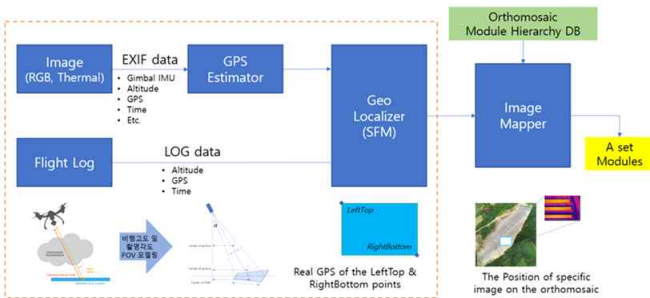


그림 4. 결함위치 Localization 기능

직관적으로 발전소의 정보와 결함의 정보를 알 수 있게 하려면 드론이 촬영한 이미지 파일에 포함되어 있는 메타데이터로부터 드론자세 정보와 카메라가 설치된 Gimbal 자세정보, GPS의 cordination 정보 등을 기반으로 결함의 위치좌표를 계산한다.

모듈위치에 결함을 표시하기 위해 촬영된 이미지에서 픽셀별 위치값을 계산하여 결함의 위치를 결정 후 결함정보를 표시함으로써 결함정보를 알 수 있도록 한다.

#### IV. 결 론

본 논문에서는 드론으로 발전소를 정밀 항공 촬영하여 모델링하는 방법을 제시하였다. 드론비행 경로를 설정하여 IR이미지를 촬영 후 AI 머신러닝 기술을 이용하여 태양광 배터리 결함을 진단 및 검출한다. 결함위치를 파악하고 불량 태양광 배터리 모듈을 특정하는 방법을 제시하였다. 기존 발전소의 점검방식에서 발생하는 진단, 점검시 감전사고 같은 재해를 사전에 예방하고 점검시간은 2배 이상 줄임으로써 유지보수 비용을 크게 절감할 수 있다. 또한, AI머신러닝 기술을 적용하여 빠르고 정확하게 태양광 배터리 결함을 진단하고 결함을 가시화시켜 관리자에게 유용한 정보를 신속하게 제공할 수 있었다.

#### 참 고 문 헌

- [1] Seungmin Lee, Woo Jin Lee, "Development of a System for Predicting Photovoltaic Power Generation and Detecting Defects Using Machine Learning", KIPS transactions on computer and communication systems, v.5 no.10, pp.353-360, 2016.
- [2] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting Solar Generation from Weather Forecasts Using Machine Learning," IEEE International Conference on Smart Grid Communications, 2011.
- [3] Younghoon Kim, Kyuseok Shim, "Computing the Efficiency of Solar Cell and Detecting Failures in the Solar Cells", Proceedings of the Korean Society of Computer Information Conference Vol.39, No.1(B), 2012.

# Wear OS의 AVD를 통한 스마트워치 포렌식 아티팩트 파악

김동영\*, 박창연\*\*, 유진우\*\*, 남성윤\*\*

## Identifying Forensic Artifacts in Smartwatches through Wear OS AVD

Dongyoung Kim\*, ChangYeon Park\*\*, Jinwoo YU\*\*, Seongyoun Nam\*\*

### 요 약

스마트워치의 이용률이 급증하고 성능이 향상됨에 따라, 스마트워치는 스마트폰을 대체할 수 있는 중요한 디지털 증거로 부각되고 있다. 특히 최근 출시되는 삼성 및 구글의 스마트워치는 Wear OS를 기반으로 하고 있지만, Wear OS에 관한 포렌식 연구는 다른 OS에 비해 상대적으로 부족하다. 본 연구에서는 Android Virtual Device(AVD)를 통해 Wear OS 스마트워치의 펌웨어 데이터를 수집하고 분석하여, 기존에 분석되지 않은 스마트워치 내 다양한 로그 파일을 제시한다. 기존 연구에서 밝혀진 통화 기록, 메시지, 연락처와 더불어 스마트워치 내부의 새로운 금융 관련 로그와 같은 정보를 분류하고 분석하여 디지털 포렌식에서 활용할 수 있는 데이터를 도출한다.

### Abstract

With the growing adoption and increasing performance of smartwatches, they are emerging as a significant source of digital evidence, potentially complementing or even replacing smartphones in forensic investigations. Although recent Samsung and Google smartwatches are based on Wear OS, forensic research on Wear OS remains limited compared to other operating systems. In this study, we utilize the Android Virtual Device (AVD) to acquire and analyze firmware data from Wear OS smartwatches, uncovering various log files that have not been previously examined. We categorize and analyze information such as call logs, messages, and contacts identified in prior studies, as well as newly discovered financial-related logs, thereby deriving data valuable for digital forensic investigations.

### Key words

Wear OS Forensics, Smartwatch Data Acquisition, AVD

### 1. 서 론

스마트워치의 이용자 수가 급증하고 성능이 향상

되면서 일상 속 중요한 디지털 기기로 자리 잡고 있다. 이에 따라 스마트폰 못지않게 디지털 포렌식에서도 중요한 증거 원천으로 주목받고 있다[1]. 특

\* 경북소프트웨어고등학교 소프트웨어개발과, dongyoung0912@gmail.com(교신저자),

\*\* 경북소프트웨어고등학교 소프트웨어개발과, changyeon0702@gmail.com

※ 이 논문은 2024년 학생 SW-AI 인재 양성 프로젝트의 일환으로 경상북도교육청의 지원을 받아 수행된 연구임.

히 구글의 스마트워치는 Wear OS를 기반으로 하며, 삼성의 갤럭시워치도 Wear OS를 기반으로 한 Wear OS powered by Samsung을 사용한다. 이러한 스마트워치는 통화 기록, 메시지, 연락처와 같은 개인 정보뿐만 아니라 다양한 활동 기록을 포함하고 있어 스마트폰을 보완하거나 대체할 수 있는 디지털 증거로 활용될 가능성이 크다. 그러나 Wear OS에 관한 포렌식 연구는 Tizen OS를 비롯한 다른 운영체제에 비해 상대적으로 부족하며, 다양한 데이터 유형과 그 활용 방안에 대한 연구가 필요한 상황이다.

본 연구는 Android Virtual Device (AVD)를 활용하여 Wear OS 스마트워치의 펌웨어 데이터를 수집하고 분석함으로써, 기존 비슷한 연구[2]에서 다루지 않았던 로그 파일들을 제시하는 것을 목적으로 한다. 이를 통해 통화 기록, 메시지, 연락처 외에도 금융 관련 로그 등 스마트워치 내부의 새로운 데이터를 발굴하여 디지털 포렌식 수사에서 활용할 수 있는 정보를 제공하고자 한다.

## II. AVD를 통한 Wear OS 펌웨어 분석

### 2.1 Wear OS

Wear OS 펌웨어는 Google사에서 제작한 Android 기반의 스마트워치 운영체제이다. Android 기반의 운영체제로써 기존 Android에서 제공되는 Android Debug Bridge(ADB) 등의 기능을 그대로 사용 가능하며 그림1 Android 계층 구조와 동일하다.

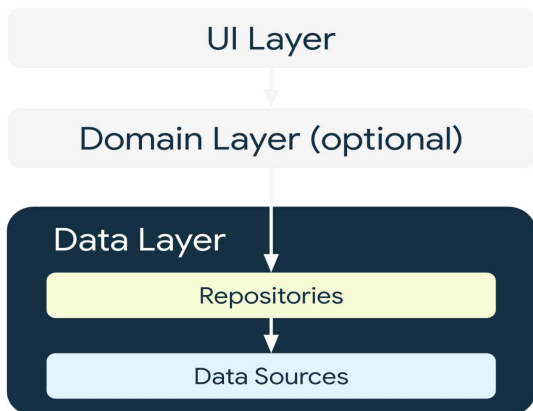


그림 1. 데이터 레이어 아키텍처  
Fig. 1. Data Layer Architecture

### 2.2 AVD를 통한 데이터 수집 방법

Wear OS가 사용된 실제 스마트워치를 가지고 연구를 진행하는 경우, 관리자 권한을 얻는 행위인 루팅을 차단해 두었기에 포렌식에서 얻고자 하는 이용자의 로그 정보까지 접근이 불가능하다. 이를 해결하고자 가상의 기기에 펌웨어를 구동시키는 Android Virtual Device(AVD)를 사용하여 환경을 구축하였다. AVD에서 작동되는 Wear OS는 자동으로 관리자 권한을 얻은 상태로 진행되기에 이용자의 로그 정보에 접근이 가능하다.

AVD에서 구동되는 Wear OS에 Android Debug Bridge(ADB)를 연결하여 가상 기기와 통신이 가능하도록 셸 환경과 내부 파일 분석 환경을 구축했다.

표 1. Wear OS 포렌식 실험 환경

Table 1. Wear OS Forensic Experiment Environment

Name	Type
Release Name	Tiramisu
API Level	33
Android	13.0 Android (Wear OS 4)
System Image	x86_64
Storage	6GB + 10GB (SDcard)

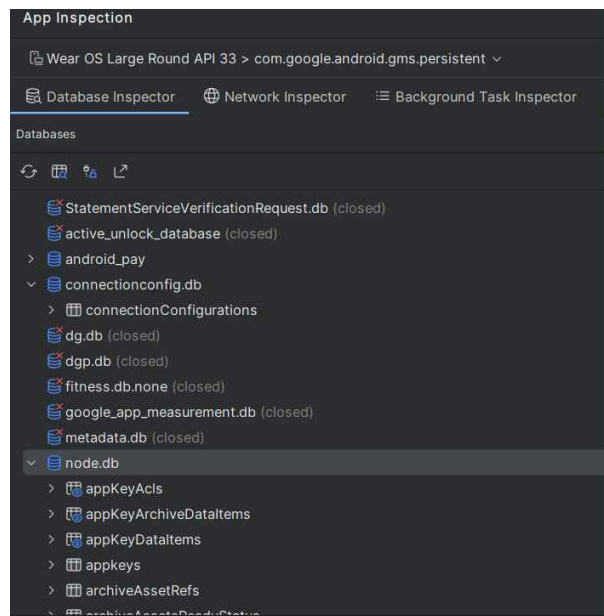


그림 2. 데이터 수집 환경 구축  
Fig. 2. Construct a data collection environment

### 2.3 AVD를 통한 Wear OS 분석 결과

본 논문에서는 AVD를 이용한 Wear OS 데이터 획득 방법을 사용하여 내부에 저장된 로그 파일을 기반으로 아래와 같이 분석할 수 있었다.

기존 연구[2]에서도 다룬 통화 기록, 메시지 기록 로그는 본 연구의 환경에서도 동일하게 파악할 수 있으며 추가로 스마트워치 이용자의 금융에 관한 정보, Google 스토어에 관련된 로그도 아래 표와 같이 파악이 가능하였다.

표 2. Wear OS 아티팩트  
Table 2. Wear OS Artifacts

Artifacts	Log Path
Call Log	com.android.providers.contact.databases\calllog.db
Message Log	com.android.providers.telephony.databases\mmsms.db
Contacts	com.android.providers.contact.databases\contacts2.db
Cards for Financial Payment	com.google.android.gms.persistent.PaymentCards
Account_id	com.google.android.gms.persistent.Account_id
Billing_id	com.google.android.gms.persistent.Billing_id
Other Wallet Data	com.google.android.gms.persistent.Additional wallet data
Recognized for the request	com.google.android.gms.persistent.Asknowledged_bundle
Unprocessed request	com.google.android.gms.persistent.Unasknowledged_bundle
Application install Information	com.android.vending.Install_data.db

### III. 결 론

본 연구에서는 Android Virtual Device (AVD)를 활용하여 Wear OS 스마트워치의 펌웨어 데이터를 효과적으로 수집하고 분석하는 방법을 제시하였다. AVD를 사용함으로써 루팅이 차단된 실제 스마트워치에서는 접근이 어려운 사용자 로그 정보에 접근할 수 있었다.

분석 결과, 기존 연구에서 다룬 통화 기록과 메시지 기록 로그 외에도 금융 관련 정보 및 Google

스토어와 관련된 로그들을 발견하였다. 이러한 결과는 Wear OS 스마트워치가 디지털 포렌식 수사에서 중요한 정보원을 제공할 수 있음을 입증하며, 사용자 개인의 다양한 활동을 기록하는 데이터들이 포렌식적 관점에서 매우 유용하다는 점을 강조한다.

앞으로의 연구에서는 Wear OS의 다양한 로그 파일을 더욱 심층적으로 분석하고, 이를 바탕으로 실질적인 포렌식 기법을 모색할 필요가 있다. 본 연구가 Wear OS 포렌식 분야의 기초 자료로 활용되기를 기대한다.

### 참 고 문 헌

- [1] Zhu, S., Xu, C., Fu, Y., Li, X., & Huang, C., "Smartwatch-based activity recognition: A pilot study", Proceedings of the 2015 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies, vol. 2015, pp. 1-6, June 2015.
- [2] Su-Bin Oh and Tae-Shik Shon, "Wear OS Forensics based on Android Emulator," Journal of Digital Forensics, vol. 18, pp. 110-121, January 2024.

# U-NET 신경망 모델을 활용한 1차원 신호에서 잡음 제거 방법

전단비\*, 유진호\*

## Noise Reduction Method for One-Dimensional Signals Using a U-NET Neural Network Model

Dan Bi Jeon\*, Jin-Ho You\*

### 요 약

본 논문에서는 U-NET 신경망 모델을 활용한 1차원 시계열 신호에서 가우시안 잡음을 제거하는 방법을 제안한다. 1차원 시계열 신호에는 점진적으로 완만하게 변하는 연속 신호와 펄스처럼 급격하게 변하는 불연속 신호가 있다. 본 논문에서는 완만한 연속신호와 불연속 펄스 신호, 그리고 이들이 혼합된 1차원 신호에 포함된 가우시안 잡음을 제거하는 U-NET 신경망 모델을 제안하고 성능을 평가한다. 제안한 신경망 학습기반 잡음제거 방법과 butterworth 디지털 필터를 비교한다. 제안한 방법은 높은 잡음 수준에서 원 신호 복원을 잘하고 특히 펄스 경계 부분의 원 신호 복원에서 좋은 성능을 보인다.

### Abstract

This paper proposes a method for removing Gaussian noise from one-dimensional time-series signals using a UNET neural network model. One-dimensional time-series signals include gradually varying continuous signals and abruptly changing discontinuous signals, such as pulses. In this study, we propose a UNET neural network model designed to remove Gaussian noise from continuous signals, discontinuous pulse signals, and mixed one-dimensional signals, and we evaluate its performance. The proposed noise reduction method, based on neural network training, is compared with the Butterworth digital filter. The proposed method demonstrates effective original signal recovery at high noise levels, particularly excelling in restoring the original signal at pulse boundaries.

### Key words

noise removal, time serial signal, neural network, UNET, deep learning

## 1. 서 론

1차원 시계열 신호의 주요 특징 중 하나는 연속적인 변화와 불연속적인 변화를 동시에 포함할 수

있다는 점이다. 연속 신호는 점진적으로 완만하게 변하는 반면, 불연속 신호는 펄스와 같은 급격한 변화를 특징으로 한다. 이러한 신호의 다양한 특성으로 인해 잡음 제거 작업이 복잡해지며, 전통적인

\* 국립목포대학교 융합소프트웨어학과

필터링 방법만으로는 연속 신호와 불연속 신호를 동시에 복원하는 데 한계가 존재한다[1]. 특히, 높은 수준의 잡음 환경에서는 신호의 복원이 더욱 어려워진다.

본 논문에서는 U-NET 신경망 모델[2, 3]을 활용하여 1차원 시계열 신호에서 발생하는 가우시안 잡음을 제거하는 방법을 제안하고, 이 방법의 성능을 평가하고자 한다. 특히, 본 연구에서는 연속적인 완만한 변화와 불연속적인 급격한 변화를 모두 포함한 1차원 신호의 잡음 제거를 목표로 하며, 제안한 U-NET 기반 신경망 모델의 성능을 Butterworth[4, 5] 디지털 필터와 비교 분석한다. 제안된 방법은 높은 잡음 수준에서도 원 신호를 효과적으로 복원하는 동시에, 펄스 신호의 경계 부분에서도 우수한 성능을 나타낸다.

## II. U-NET 신경망 모델기반 잡음 제거

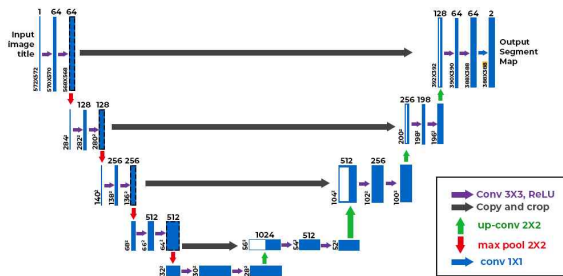


그림 1. U-NET 신경망 모델[1]  
Fig. 1. U-NET Architecture[1]

U-NET 신경망 모델은 이미지 분할 및 복원과 같은 고해상도 작업을 위해 설계된 인코더-디코더 구조의 심층 신경망 모델이다[2]. U-NET 모델은 2D 이미지의 노이즈 제거 문제에서도 효과적으로 활용되고 있다[2, 3]. U-NET이 노이즈 제거에 적절한 이유는 다음과 같다. 첫째, 인코더와 디코더의 스킵 연결 구조는 신호의 고주파 세부 정보를 보존하면서 노이즈 제거를 가능하게 한다. 이는 노이즈가 포함된 이미지의 경계선을 효과적으로 복원하는 데 기여하며, 경계 부분의 세밀한 정보가 중요한 경우 특히 유리하다. 둘째, 인코더 부분에서 입력 신호의 특징을 축약하여 중요한 정보를 추출한 후, 디코더에서 복원하는 구조는 노이즈와 신호를 구별하여 노이즈 제거 성능을 높이는 데 도움을 준다. 셋째,

딤러닝 기반으로 학습을 통해 노이즈 유형 및 패턴에 적응할 수 있으므로, U-NET 모델은 다양한 수준의 노이즈 상황에 유연하게 대응할 수 있다. 본 논문에서도 이를 바탕으로 1차원 시계열 신호에서의 잡음 제거를 수행하고자 한다.

## III. 1차원 신호 잡음 제거 U-NET 1D 모델

1차원 신호에서 잡음 제거를 위해 본 논문에서 제안한 U-NET 모델은 1차원 U-Net 아키텍처로서 다음과 같은 주요 구성 요소를 갖는다.

### Encoder (인코더)

기존 U-Net과 마찬가지로, 인코더 부분에서 입력 신호의 중요한 특성을 추출하고 점차적인 다운샘플링을 통해 정보를 압축한다. 이 과정은 반복적인 Conv1d와 MaxPool1d 연산으로 구성되며, 여기서 각 Conv1d 레이어는 BatchNorm1d와 ReLU 활성화 함수와 함께 적용되어 모델의 성능과 학습 안정성을 높인다(CBR1D). 첫 번째 CBR1D 레이어에서 채널 수가 1에서 64로 확장되고, 이후 단계별로 128, 256, 512, 1024 채널까지 증가되며, 각 단계마다 MaxPool1d를 통해 다운샘플링을 진행하여 인코더를 통과할수록 신호의 공간적 크기는 줄어들고, 대신 채널 수가 늘어나면서 정보가 점차 압축된다.

### Decoder (디코더)

디코더 부분에서는 인코더에서 압축된 정보를 점차 업샘플링하여 원래의 신호 크기로 복원하며 잡음을 제거한다. 이 과정에서는 ConvTranspose1d 레이어를 사용하여 업샘플링하며, 각 단계에서 대응되는 인코더의 출력을 연결하여 U-Net의 skip connection을 구현한다. 디코더의 첫 번째 레이어는 ConvTranspose1d로 시작해 인코더의 마지막 출력을 업샘플링하여 512 채널로 변환하고, 해당 단계의 인코더 출력(512 채널)과 결합하여 1024 채널로 만든다. 각 디코딩 단계마다 CBR1D 레이어를 통해 채널 수를 절반으로 줄이며 점진적으로 복원하며, 최종 레이어는 Conv1d를 사용하여 채널 수를 1로 줄여 최종 출력을 생성한다.

### 전체 Forward Pass

이 모델에서는 인코더와 디코더가 순차적으로 수행되며, 인코더 단계에서 다운샘플링된 특성들이 디코더의 각 단계에서 skip connection을 통해 결합되며 이러한 구조는 1차원 신호에서 중요한 정보를 보존하면서 점진적으로 복원하여, 잡음을 제거하는데 유리한 구조를 제공한다.

## IV. 실험 및 평가

본 논문에서는 점진적으로 변화하는 연속형 신호와 펄스형 신호들의 잡음을 효과적으로 제거하는 U-NET 1D 신경망을 학습하기 위하여 그림 2, 3과 같은 다양한 주파수와 크기를 갖는 사인과 신호와 펄스 신호를 생성하고 이에 가우시안 잡음을 여러 수준으로 주입한 잡음 신호와 원 신호기반 학습데이터를 생성하여 U-NET 신경망을 학습하였다.

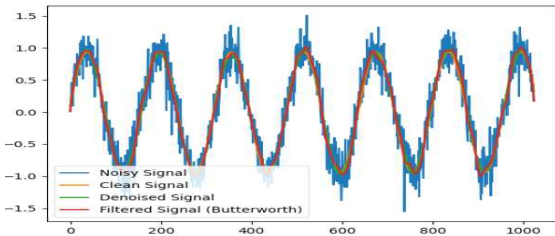


그림 2. 사인파형 신호에서 잡음 제거 학습  
Fig. 2. Noise removal training for sinusoidal signals

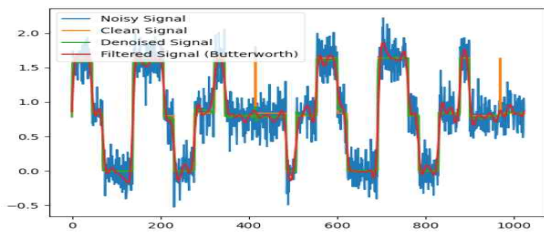


그림 3. 펄스형 신호에서 잡음 제거 학습  
Fig. 3. Noise removal training for pulse signals

또한 그림 2, 3와 같이 butterworth 디지털 필터는 노이즈 제거를 위해 차수를 높이거나 차단 주파수를 낮추면 특히 펄스형 신호의 경계 부분에서 진동이 심하게 나타남을 알 수 있다.

그림 4는 U-NET 1D 모델로 학습하지 않은 1차원 신호를 잡음을 제거한 결과를 보여준다. 대체적

으로 급격하게 계단처럼 변하는 신호에 대해서 원 신호에 가까운 잡음 제거 출력을 보여주지만, 높은 잡음 수준에서는 U-NET 1D의 출력이 원 신호와의 차이가 있음을 보여주고 있다.

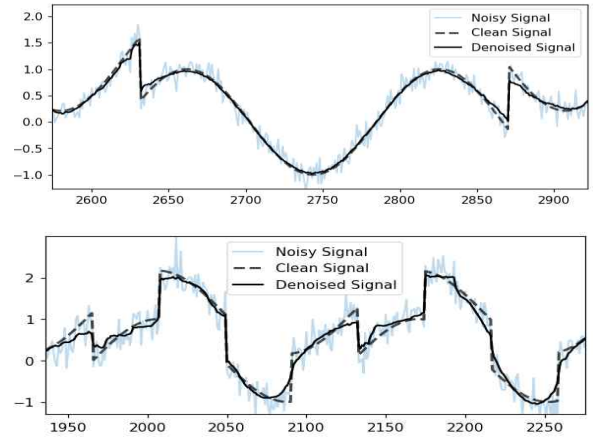


그림 4. 비학습된 신호에서 잡음 제거  
Fig. 4. Noise removal from untrained signals

## IV. 결 론

본 논문에서는 완만한 연속신호와 불연속 펄스 신호, 그리고 이들이 혼합된 1차원 신호에 포함된 가우시안 잡음을 제거하는 U-NET 1D 신경망 모델을 제안하고 성능을 평가하였다. 제안한 신경망 학습기반 잡음제거 방법은 범용적인 1차원 신호 잡음 제거에 사용되는 butterworth 디지털 필터와 비교하였으며, 높은 잡음 수준에서 원 신호 복원을 잘하고 특히 펄스 경계 부분의 원 신호 복원에서 좋은 성능을 보였다. 향후 연구방향은 다양한 1차원 신호 구조에서 가우시안과 다른 유형의 잡음 신호를 제거하는 U-NET 모델을 연구하고자 한다.

## 참 고 문 헌

- [1] Ifeachor, E.C. and Jervis, B.W., Digital Signal Processing: A Practical Approach, Prentice Hall, 2002.
- [2] Ronneberger, O., Fischer, P., Brox, T., "U-Net: Convolutional Networks for Biomedical Image Segmentation", MICCAI 2015
- [3] 한성현, 허지혜, 성연승, 임동훈, "CT 영상에서 U-Net 기반 변형가능 컨볼루션 GAN을 이용한

잡음제거,” 한국데이터정보과학회지, 34(4),  
567-585, 10.7465/jkdi.2023.34.4.567, 2023.

[4] \_, “Butterworth filter”, [https://en.wikipedia.org/wiki/Butterworth\\_filter](https://en.wikipedia.org/wiki/Butterworth_filter).

[5] \_, “Digital filter”, [https://en.wikipedia.org/wiki/Digital\\_filter](https://en.wikipedia.org/wiki/Digital_filter).



# 한글 디지털 문서 위조 데이터셋 자동 생성 방법에 대한 연구

배용열\*, 조대제\*\*, 정기현\*\*

## A Study on Auto Generation of Korean Digital Document Forgery Dataset

Yong Yeol Bae\*, Dae Jea Cho\*\*, and Ki Hyun Jung\*\*

### 요약

디지털 문서 위조 탐지 연구에서 한글로 된 위조 문서 데이터의 부족은 국내 디지털 문서 위조 탐지 알고리즘 개발에 대한 접근성을 떨어뜨릴 수 있다. 이를 해결하기 위해서 본 연구에서는 복사-이동(Copy-move) 방식과 삽입(Insertion) 방식의 위조 기법에 기반한 한글 디지털 위조 문서 데이터셋의 자동화 생성 방법을 다루고자 한다. 한글 문서 데이터는 대학교 성적증명서 양식으로 문서 안의 내용은 ChatGPT와 Dall-E2 API를 이용하여 무작위로 생성하였다. 생성된 데이터는 한글 디지털 문서 위조 탐지 연구에 필요한 실험 데이터셋으로 활용 가능하므로, 다양한 한글 문서 위변조 탐지 알고리즘 개발에 기여할 것으로 판단된다.

### Abstract

The lack of Korean forgery document data in digital document forgery detection can reduce the accessibility of developing domestic digital document forgery detection algorithms. To address this issue, this study proposes an automation method for generating a Korean digital forgery document dataset based on copy-move and insertion forgery techniques. The dataset consists of university transcript forms, and the contents in the documents are randomly generated by using the ChatGPT and Dall-E2 APIs. The generated dataset provides an experimental environment necessary for research in Korean digital document forgery detection and is expected to contribute to advancements in the related field.

### Key words

Document Forgery Detection, Dataset, Hangul, Artificial Intelligence

## 1. 서론

다양한 종류의 문서 및 인쇄물을 이미지 파일로

만들어 정보를 관리하는 방식은 정보 전달 및 관리 측면에서 다양한 이점을 제공할 수 있으며, 이러한 이유로 많은 기업과 정부 기관에서는 문서를 이미

\* 국립안동대학교 대학원 소프트웨어융합학과, baeyongyeol@gmail.com

\*\* 국립안동대학교 소프트웨어융합학과, djcho@anu.ac.kr, kingjung@anu.ac.kr

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

지 파일로 저장 및 관리하고 있다. 그러나 최근 이미지 편집 기술의 발전으로 인해 문서 이미지의 복잡한 위조 또는 변조 사례가 증가하고 있으며, 이는 경제적, 사회적 혼란을 초래할 수 있다. 따라서 이를 방지하기 위해 다양한 방법의 디지털 문서 조작 탐지 기법이 연구되고 있지만, 각 문서가 작성된 언어의 종류가 다른 경우 해당 언어의 형태적 특성[1]을 모두 반영한 디지털 문서 조작 탐지 기법을 개발하는 것은 매우 어렵다. 따라서 본 연구에서는 다양한 언어 중 한글로 작성된 문서에 대한 조작 영역 탐지 기법 개발에 활용할 수 있는 한글 디지털 위조 문서 데이터셋의 자동 생성 방법을 개발하였다. 개발한 방법은 다양한 디지털 문서 조작 기법 중 복사-이동(Copy-move)와 삽입(Insertion) 조작방법에 기반하였으며, 생성되는 문서는 ChatGPT와 Dall-E2에서 생성된 무작위 내용을 바탕으로 한 대학교 성적증명서 양식이다. 개발한 방법을 활용하면 복사-이동 또는 삽입 방식으로 조작된 디지털 문서 이미지를 쉽게 생성할 수 있으며, 이렇게 생성된 이미지를 활용하면 한글 디지털 문서 이미지의 조작 탐지 기법 개발의 발전에 기여할 수 있을 것이라 예상된다.

## II. 관련 연구

### 2.1 디지털 문서 조작 탐지 방법

디지털 문서의 조작 탐지 방법은 크게 능동적(Active) 방법과 수동적(Passive) 방법으로 구분할 수 있다. 능동적 방법은 이미지에 디지털 워터마킹(Digital Watermarking)이나 디지털 시그니처(Digital Signature)와 같은 정보를 미리 삽입한 후, 해당 이미지에서 워터마크나 시그니처 정보가 정상적으로 추출되는지 확인하여 조작 여부를 탐지하는 방법이고, 수동적 방법은 이미지의 통계적 특징, 픽셀의 일관성, 밝기 정보 등을 분석하여 위조된 부분을 찾아내는 방법이다. 수동적 방법은 추자 정보 없이 조작 탐지가 가능하기 때문에 더 광범위하게 사용할 수 있지만, 일반적으로 탐지가 더 어렵고 복잡한 경향이 있다[2].

### 2.2 디지털 문서 조작 방법

디지털 문서의 조작 방법에는 크게 복사-이동(Copy-move), 스플라이싱(Splicing), 삽입(Insertion) 등이 있다.

복사-이동 조작방법은 이미지의 한 영역을 복사하여 동일한 이미지 내 다른 위치에 붙여넣는 방식으로, 복사된 영역이 원본과 동일한 색상이나 텍스트를 가지기 때문에 탐지가 어렵다[3]. 스플라이싱 조작은 한 이미지의 일부를 다른 이미지에 삽입하여 새로운 합성 이미지를 만드는 방법으로, 이때 이미지의 경계나 색상, 조명의 일관성을 유지하는 것이 중요하다[4]. 마지막으로, 삽입 조작은 소프트웨어 도구를 사용해 텍스트를 수정하거나 문서 내 특정 위치에 텍스트를 추가하는 방식으로, 문서 내용의 변경이나 추가를 자연스럽게 보이도록 하는 것이 목표다[5].

## III. 한글 위조 문서 데이터셋 생성 방법

### 3.1 위조를 위한 데이터 생성

위조를 위한 데이터는 학생 정보와 성적 정보, 대학 명이 입력된 성적증명서이다. 데이터 생성을 위하여 파이썬의 docx 모듈과 ChatGPT API, Dall-E2 API를 활용하여 무작위 성적증명서를 MS word 파일로 생성한다. 이후 pdf2image 라이브러리를 활용하여 MS 워드 파일을 이미지 파일로 변환하여 성적증명서 이미지 파일을 생성한다.

### 3.2 복사-이동 조작 방식의 위조 문서 데이터

복사-이동 방식은 한 이미지 내에서 특정 영역을 복사한 후 조작하려는 위치에 복사된 이미지를 붙여넣는 방식으로 위조가 진행된다. 이러한 절차에 따라 복사-이동 방식으로 조작된 한글 디지털 위조 문서 데이터셋의 생성 순서는 아래와 같다.

단계 1: 성적증명서 이미지에서 OCR과 ChatGPT API를 활용하여 과목명에 해당하는 텍스트 좌표를 추출하고 해당 좌표 영역을 크롭핑하여 저장한다.

단계 2: 추출한 좌표 영역에 파이썬의 PIL 라이브러리를 이용하여 해당 영역의 크기와 같은 사각

형을 입력한다.

단계 3: 입력한 사각형 위에 단계 1에서 크롭핑한 이미지를 덮어씌운다.

단계 4: 이후 전체 이미지를 저장하고 추출했던 좌표 영역은 텍스트 파일에 입력하여 위조된 부분을 기록한다.

### 3.3 삽입 조작방식의 위조 문서 데이터

삽입 조작방식은 위조하려는 영역에 새로운 텍스트를 입력하여 문서를 조작하는 방법이다. 삽입 방식으로 조작된 한글 디지털 위조 문서 데이터셋의 생성 순서는 아래와 같다.

단계 1: 변환된 이미지에서 OCR과 ChatGPT API를 활용하여 학생명과 대학명, 학점에 해당하는 텍스트의 좌표를 추출하고 해당 좌표를 저장한다.

단계 2: 추출한 좌표 영역에 파이썬의 PIL 라이브러리를 이용하여 해당 영역의 크기와 같은 사각형을 입력한다.

단계 3: 입력한 사각형 위에 ChatGPT API를 활용하여 무작위로 생성된 대학명과 학생명, 학점을 같은 폰트 정보로 설정 후 입력한다.

단계 4: 이후 전체 이미지를 저장하고 추출했던 좌표 영역은 텍스트 파일에 입력하여 위조된 부분을 기록한다.

위 방법을 순서도로 표시하면 그림 1과 같다.

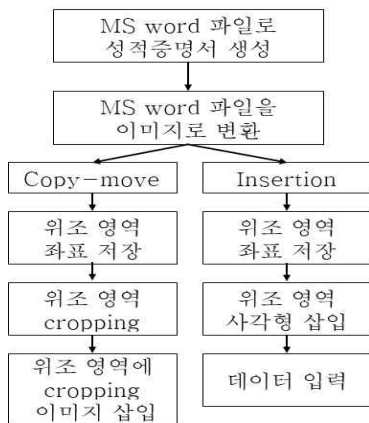


그림 1. 디지털 문서 조작 순서도  
Fig. 1. Flowchart of Digital Forgery Document

### 3.4 구현 결과

3.3에서 설명한 방식으로 디지털 문서를 조작한 결과는 그림 2와 그림 3과 같다. 성적증명서에서 붉은색 사각형으로 표시한 부분이 각각의 방법으로 위조한 영역이다. 성적증명서 안의 내용은 ChatGPT API를 활용하여 무작위로 구성하였다.

#### 성적증명서

학생 정보	
이름	김하늘
학번	202103475
생년월일	2002년 8월 15일
학년	2학년
전공	심리학

성적 정보					
학기	과목 코드	과목명	학점	성적	평점
2023년 봄학기	MATH101	미적분학	3	A	4.0
2023년 봄학기	PHY101	일반물리학	3	B+	3.5
2023년 봄학기	CHEM101	기초화학	3	A-	3.7
2023년 봄학기	CS101	프로그래밍입문	3	B	3.0
2023년 봄학기	ENG201	영문학개론	2	A	4.0

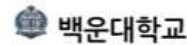


그림 2. Copy-move 방식 위조 결과  
Fig. 2. The Result of Copy-move Forgery Method

#### 성적증명서

학생 정보	
이름	김민준
학번	202103475
생년월일	2002년 8월 15일
학년	2학년
전공	심리학

성적 정보					
학기	과목 코드	과목명	학점	성적	평점
2023년 봄학기	MATH101	미적분학	3	A	4.0
2023년 봄학기	PHY101	일반물리학	3	B+	3.5
2023년 봄학기	CHEM101	기초화학	3	A-	3.7
2023년 봄학기	CS101	프로그래밍입문	3	B	3.0
2023년 봄학기	ENG201	영문학개론	2	A	4.0

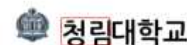


그림 3. 삽입 방식 위조 결과  
Fig. 3. The Result of Insertion Forgery Method

## IV. 결 론

본 연구에서는 한글 디지털 문서 위조 탐지 연구의 한계를 극복하기 위하여 복사-이동 조작과 삽입

조작 방법을 기반으로 한 자동화된 위조 데이터셋 생성 방법을 제안하였다. 제안된 방법은 성적증명서 양식을 활용해 ChatGPT와 Dall-E2 API로 무작위 데이터를 생성하고, 이를 바탕으로 조작된 문서 이미지 데이터를 손쉽게 만들 수 있도록 한다. 제안한 방법은 한글 디지털 문서 위조 탐지 기법 개발에 실질적인 기여를 할 것으로 기대되며, 추후 연구에서는 다양한 조작 기법과 해당 데이터셋을 활용한 디지털 문서 조작 탐지 방법에 대하여 연구해 보고자 한다.

### 참 고 문 헌

- [1] H. Miton and O. Morin, "Graphic complexity in writing systems," *Cognition*, Vol. 214, 104771, 2021.
- [2] A. H. Saber, M. Khan, and B. Mejbil, "A survey on image forgery detection using different forensic approaches," *Advances in Science, Technology and Engineering Systems Journal*, Vol. 5, pp. 361-370, Jan. 2020.
- [3] M. Zanardelli, F. Guerrini, R. Leonardi, and N. Adami, "Image forgery detection: a survey of recent deep-learning approaches," *Multimedia Tools and Applications*, Vol. 82, pp. 17521-17566, 2022.
- [4] X. Bi, Y. Wei, B. Xiao, and W. Li, "RRU-Net: The Ringed Residual U-Net for image splicing forgery detection," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA, pp. 30-39, 2019.
- [5] S. S. Gornale, G. Patil, and R. Benne, "Document image forgery detection using RGB color channel," *Transactions on Engineering and Computing Sciences*, Vol. 10, No. 5, pp. 1-14, 2022.

# 센서 장갑 기반 군집 제어가 가능한 군사용 다목적 무인 지상 로봇 기술 개발

최지훈\*, 박소은\*\*, 표기정\*\*, 박송희\*\*, 김영형\*\*

## Development of military multipurpose unmanned ground robot technology capable of swarm control based on sensor glove

Ji-Hoon, Choi\*, So-eun Park\*\*, Ki-jung Pyo\*\*, Song-Hee Park\*\*, and Young-hyeong Kim\*\*

### 요 약

전세계적인 인구감소로 병력 원이 감소함에 따라 인명 피해의 최소화 및 전투 효율성을 높이기 위해 미래 전장은 점차 무인 로봇을 이용한 무인·첨단화가 진행되고 있다. 이에 각 국방 관련 기업들은 사람과의 유·무인 복합 운영체계 구축을 위해 무인 지상 로봇들을 활발히 개발 중에 있지만 해당 표준 및 공용 플랫폼의 부재로 각 로봇 개체 간의 상호 호환이 어려운 실정이다. 본 논문에서는 1개의 임무 수행만 가능한 기존의 무인 지상 로봇의 기능적 요소를 통합하여 지상군이 요구하는 정찰, 공격, 수송, 탐지와 관련된 4가지 임무 수행이 모두 가능한 모듈형 센서 장갑 기반 군집 제어가 가능한 군사용 다목적 무인 지상 로봇에 대한 구현 방안을 제시하고자 한다.

### Abstract

As the number of troops decreases due to the worldwide population decline, future battlefields are gradually becoming unmanned and advanced using unmanned robots to minimize casualties and increase combat efficiency. Accordingly, each defense-related company is actively developing unmanned ground robots to establish a manned and unmanned composite operation system with people, but the lack of standards and common platforms makes it difficult for each robot to be compatible with each other. In this paper, we propose an implementation plan for a modular sensor armor-based swarm control-capable military multipurpose unmanned ground robot that can perform all four missions related to reconnaissance, attack, transportation, and detection required by ground forces by integrating the functional elements of existing unmanned ground robots that can perform only one mission.

### Key words

UGV, Military, Sensor Glove, Modular, Multipurpose, MUM-T (Manned and unmanned Team)

---

\* 아토즈, abcd7535@naver.com, \*\* 국립금오공과대학교, kic126@kumoh.ac.kr

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

## 1. 서 론

전 세계 병력자원 감소로 인명 피해를 최소화하고 전투 효율을 극대화하기 위해 군사 관련 무인 지상 로봇(UGV)의 첨단화가 급속히 진행되고 있지만, 국방 관련 방산 기업들은 각각의 프로토콜 및 플랫폼을 이용하여 1개의 임무 수행이 가능한 무인 지상 로봇 개발만 진행하는 실정이다. 이는 운용하기 위한 병력자원이 더 많이 필요한 상황을 초래하기 때문에 통합을 위한 공용 플랫폼 개발이 반드시 필요한 상황이다. [1] 따라서 본 논문에서는 유·무인 복합 운영체제 (MUM-T) 구축에 필수적인 군사용 무인 지상 로봇과 관련하여 지상군이 요구하는 4가지 임무 (정찰, 공격, 수송, 탐지)를 동시에 수행할 수 있도록 모듈화된 다목적 무인 지상 로봇을 제안하여 다양한 프로토콜을 사용하는 UGV와의 상호 운영성을 확보하고자 한다.

## II. 모듈화된 군사용 다목적 무인 지상 로봇

제안된 센서 장갑 기반 군집 제어가 가능한 군사용 다목적 무인 지상 로봇 (MUGR, Multipurpose Unmanned Ground Robot) 시스템 플랫폼의 경우 전원, 구동부, 제어부, 감지부를 가지는 Base Platform을 공용으로 사용하고 군에서 요구하는 군사용 무인 로봇의 4가지 임무(공격, 정찰, 탐지, 수송)에 관한 Top Module을 교체/장착함으로써 군사용 무인 로봇의 효율성을 극대화한 MUGR 제품과 LTE 기반의 TCP/IP 통신망과 GPS 좌표 정보를 기반으로 MUGR에 임무를 부여하고 각 로봇 개체에 대한 상태 정보를 확인할 수 있으며 MUGR의 군집 제어가 가능한 지상용 통합관제시스템 제품인 GCS (Ground Control System)으로 구성된다.

또한, 실질적인 전장에서 MUGR을 군인이 직접 운영할 수 있도록 손의 움직임과 손가락의 굽힘 정도를 Gyro Sensor와 Flexible Sensor로 감지하여 제어 명령어를 생성하는 Sensor Glove를 포함하고 있다.

MUGR의 군집 제어를 위해 포메이션 알고리즘과 코스트맵 기반의 경로 생성 알고리즘을 적용하였으며, LiDAR와 IMU (관성 측정 장치) Sensor를 통해

다중 SLAM 방식의 자율주행 기능을 구현하였다. 이때, 위치 추정 오차 누적에 대한 문제점을 해결하기 위해서는 특정 지점의 시각적 특징을 기억하여 위치 추정 오차를 보상하는 번들 조정을 통해 해결하고 로봇의 순간적 위치 상실에 대한 문제는 기억된 키프레임을 통해 상실된 위치 정보를 역추적하는 BoF (Bag of Features) 및 BoVW (Bag of Visual Words) 알고리즘을 활용하여 문제점 해결하였다.

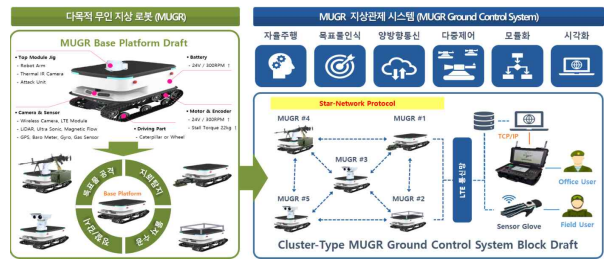


그림 1. 센서 장갑 기반 군집 제어가 가능한 군사용 다목적 무인 지상 로봇 (MUGR) 시스템 아키텍처  
Fig. 1. Military Multipurpose Unmanned Ground Robot (MUGR) System Architecture Capable of Sensor Glove-Based Swarm Control

## III. 결 론

본 논문에서는 각기 다른 프로토콜로 운용되는 무인 지상 로봇에 대한 시스템의 상호 호환성을 확보하기 위해 지상군이 요구하는 4가지 임무를 1개의 MUGR을 통해 수행할 수 있는 모듈화된 군사용 다목적 무인 지상 로봇에 대한 시스템 아키텍처를 제시하였다. 효율적이고 안정적인 로봇 운용을 위해 군집 제어, 자율주행에 대한 기능을 구현하였으며 군사 관련 소대급 소형 무인 지상 로봇에 대한 시스템 플랫폼을 고안하였다

## 참 고 문 헌

- [1] Dongbeom Kim · Hosung Kim : Development and Technological Strategy of Intelligent Ground Manned-Unmanned Composite Combat Systems : Focused on Future Ground Combat Systems, Journal of the KADIS, Vol. 30, No. 3, pp. 59-67, (2023)

# 납세자 인식 조사 기반의 투자 성향 의사 결정 요인 분석

최인수\*, 김우창\*\*

## An Analysis of Household Investment Preferences and Decision Factors Based on Taxpayer Perception Survey

Insu Choi\* and Woo Chang Kim\*\*

이 연구는 2017년부터 2022년까지 한국 가계의 위험 감수 태도의 변화를 분석하며, XGBoost의 예측력과 SHAP 값의 설명력을 활용한다. 대한민국의 재정패널조사 데이터의 납세자 인식 조사를 활용하여 가계의 금융 리스크 감수 성향에 영향을 미치는 요인을 식별하고자 하였다. 우리의 연구 결과는 경제 변동 기간 동안 주관적 인식과 금융 지식이 위험 회피 성향을 형성하는 것에 주안점을 둔다. 금융 투자에 대한 위험 회피 계수의 결정 요인을 패널 데이터를 바탕으로 분석함으로써, 본 연구는 금융 지식과 안정성을 향상시키기 위한 정보에 기반한 금융 및 경제 정책 개발에 유용한 통찰력을 제공하는 것을 목표로 한다.

### Key words

Financial Risk Aversion Coefficient; Machine Learning; Explainable Artificial Intelligence; Financial Behavior; Household Finance

---

\*한국과학기술원, jl.cheivly@kaist.ac.kr, \*\*한국과학기술원, wkim@kaist.ac.kr(교신저자)

# 한방 데이터 관리현황 및 품질관리 방안연구

김헌성\*, 이재훈\*\*

## A study on the current status of Korean medicine data management and quality control measures

Hun Sung Kim\*, Jaehoon Lee\*\*

### 요 약

한의학은 이병동치, 동병이치라는 고유한 특성을 가지고 있어 표준화 및 데이터화에 오랜 기간 어려움을 겪어왔다. 특히 한의업계의 국제적 진출과 내국민의 신뢰성 회복을 위해 이에 대한 필요성을 체감하고 이를 체계화하기 위한 연구가 진행되어 왔으며 관련 데이터를 제공하고 있다. 그러나 해당 데이터들은 한의약 진흥원, 오아시스 등 다양한 곳에 혼재 및 중복되어 있어 한의사, 환자, 연구자 입장에서 필요한 데이터에 대한 접근성이 현저히 떨어지는 상황이다. 또한 다수의 사이트들이 검색, 열람 기능으로 DB를 공개하고 있으며 실질적인 분석을 위해 필요한 파일 형태의 DB를 제공하지 않고 있다. 이에 본 연구를 통해 한방 데이터 관리 현황을 확인하고 앞에서 제시한 문제들을 개선할 수 있는 효율적인 품질관리 방안에 대해 연구해보고자 한다.

### Abstract

Traditional Korean medicine has unique characteristics, such as the principles of treating different diseases with the same method (異病同治) and treating the same disease with different methods (同病異治), which have long posed challenges in terms of standardization and data collection. As the need for international expansion of Korean medicine and the restoration of public trust has been recognized, research has been conducted to address these issues and provide related data. However, this data is dispersed and duplicated across various sources, such as the Korean Medicine Promotion Agency and OASIS, making access to necessary information difficult for practitioners, patients, and researchers. Moreover, many databases are made available only for search and viewing purposes, without offering downloadable files for practical analysis. This study aims to review the current state of Korean medicine data management and explore effective quality management measures to address these challenges.

### Key words

korean medicine data, data quality control, big data

## 1. 서 론

국내 고유의 의학인 한의학은 약 5000년간 한국의 역사와 함께 해온 유서깊은 학문이며, 양의학만

으로는 치료가 어려운 질환에 대해 긍정적인 치료 성과를 보여주는 사례들이 다수 보고될 정도로 효능 역시 입증되어 있다.

그러나 한의학은 이병동치(異病同治), 동병이치

\* 메디케이스ystem 기업부설연구소, oi9i@naver.com

\*\* 메디케이스ystem 기업부설연구소, herblinker@gmail.com

※ 본 연구는 보건복지부의 재원으로 한국보건산업진흥원의 보건의료기술연구개발사업 지원에 의하여 이루어진 것임 (과제고유번호 : HI23C149000)



(同病異治)라는 고유한 특성을 가지고 있어, 오랜 기간 동안 표준화 및 데이터화에 어려움을 겪어왔다. 이는 대체 의학시장의 해외 진출과 한의업계에 대한 국민들의 신뢰성 확보에 저해하는 주요 원인으로 지적되어 오고 있다. 이를 개선하기 위해 정부와 한국한의학연구원에서는 한약 국제표준 개발 연구, 한의기술 표준화 기반 구축 연구 등을 진행하고 있다. 그러나 수집 데이터의 양이 제한적일 뿐만 아니라 연구가 분산되어 진행되는 과정에서 하나의 통합된 한의학 데이터베이스를 확인할 수 없다는 문제를 가지고 있다. 이렇게 분산되어 수집, 관리가 진행되는 한의학 데이터들은 각기 다른 분류 기준을 적용하고 있어 데이터 통합이 어렵고 품질 관리 역시 균일하게 이루어지지 않고 있다. 이러한 문제점들은 한의학 데이터를 활용하여 추가적인 가치창출을 하는 데에 큰 장애 요인으로 지적되고 있다.[1][2][3]

따라서 본 연구에서는 현재 우리나라의 한방 데이터 관리 현황을 조사하고 향후 관련 데이터에 대한 품질 관리 방안을 제시하고자 한다.

## II. 한방 데이터 관리 현황 및 문제점

한방 데이터 관리 체계 및 품질 관리 현황에 대해 조사하였다. 국내 한방 데이터를 제공하고 있는 한의약진흥원 등 14개 기관(단체)에서 보유하고 있는 데이터를 데이터의 종류, 민간 공유 여부, 관리주체, DB제공 형태를 기준으로 정리하였다.

표 1과 같이 각 기관(단체)에 따라 한의학실태조사 결과, 한약재, 의료, 약재, 처방, 병증 등 다양한 한방 관련 데이터가 다양하게 제공되고 있으며 대부분의 데이터에 대해서 민간공유가 되고 있는 상황이다. 홈페이지 및 게시자에 따라 세부적인 저작권 방침에는 차이가 있으나 기업 연구 등과 같이 이해관계자 문제가 없는 경우라면 대부분의 데이터들이 무상/유상 형태고 공개되어 이용이 가능하다.

현재 국내에서 관리되고 있는 한방 데이터의 경우 다양한 웹사이트에 산발적으로 분산되어 있는 것이 가장 큰 문제다. 해당 웹사이트들이 연동되어 통합되지 않기 때문에 다른 사이트에서 동일한 데이터를 중복제공하거나 서로 다른 데이터의 명칭이 동일하게 관리되는 경우도 빈번하다.

조사한 14곳의 웹사이트 중 8곳이 DB 파일을 제공하는 대신 검색 기능을 중심으로 한 열람 방식으로 데이터를 공개하고 있는 것 역시도 데이터 접근성을 낮추는 원인으로 분석된다.

표 1. 한방 데이터 관리 현황

Table 1. Oriental medicine data management status

제공처	Data 종류	공유여부	관리주체	DB제공형태
한의약 진흥원	실태조사, 한약재, 의료 현황	공유(일부유료)	자체, 통계청	SPSS, XLS
한의약 소재은행	천연물, 한약재	공유	자체	X
국가한의학정보포털	임상연구, 진료지침	공유	자체	X
한의약융합연구정보센터	상호작용, 표준경혈, 생약	일부공유	자체	X
전통지식포털	논문, 약재, 처방, 병증 등	일부 O	자체, 통계청	XML
오아시스	논문, 보고서, 약재, 처방, 통계	멤버십 공유	자체	X
한약표준자원은행	한약, 문헌	일부공유	자체	X
한의약융합연구정보센터	상호작용, 표준경혈, 생약	일부공유	자체	X
한국한의학연구원	연구, 보고서	일부공유	자체	PDF
한의학 고전DB	원전, 진단치료, 종합의방, 분과, 의사실기	CC(BY-NC-ND)	자체	X
한의학정보 서비스	한의학, 한의기초, 경혈미디어	일부공유	자체	X
한의학정보은행 KDC	KDC, 체질건강인, 처방, 국외	공유(신청필요)	자체	XLS, SPSS, CSV
한의학맨티검색	약재, 처방, 효능, 주치, 병, 변증, 증상, 경락, 경혈, 침방	공유	자체	X
한의원라인정책 서비스	국내외 한의학, 한의약산업실태조사	공유	자체	PDF
한국생명공학 연구원	연구장비, 생명자원 연구성과물	공유	DATA	CSV

데이터 분석을 통해 추가적인 가치창출을 위해서는 SPSS, XLS 등의 DB 형태가 선호되고 있으나 조사한 14곳의 웹사이트 중 분석이 가능한 형태의 파일을 제공하는 곳은 한의학진흥원, 전통지식포털, 한의임상정보은행 KDC, 한국생명공학연구원으로 총 4곳에 불과했다. 특히 검색 및 열람 중심으로 제공되고 있는 데이터들의 경우 정제되지 않은 비정형 데이터들이 많은 것 역시 데이터에 대한 활용성을 감소시키는 요인으로 분석된다.

한의학진흥원의 경우 실태조사 3건, 한약재 6건, 의료 8건에 달하는 DB를 제공하고 있었으며 통계청을 중심으로 데이터 관리가 진행되는 상황이다. 또한 일부 데이터 제공에 있어서는 한의학진흥원의 승인 및 제공 여부를 결정하고 있는 구조이다. 해당 데이터들의 경우 통계청에서 품질 관리를 진행하고 있으며 통계 분석, 통계 공표, 관리 및 이용자 서비스, 사후 관리 등 품질 진단 매뉴얼을 만들어 관리하고 있다. 또한 해당 매뉴얼을 기준으로 매년 자체 통계품질진단을 시행하여 우수, 양호, 보통, 주의, 미흡 총 5단계로 평가를 진행하고 있다. [4][5][6][7][8][9][10][11][12][13][14][15][16][17]

### III. 품질관리 방안

이에 한방데이터 관리 현황 및 품질 관리를 위해 두 가지 방안을 제시하고자 한다.

첫째 한의학 데이터 표준화가 선행 되어야 한다. 여러 곳의 사이트에서 다양한 데이터가 관리되더라도 표준화된 명칭, 코드로 데이터가 관리된다면 향후 유의미한 데이터로 활용 될 수 있다.

둘째, 데이터 수집 및 제공 주체와 데이터 품질 관리 주체가 분리되어야 한다. 대부분 한의학 데이터는 데이터 수집 및 제공자가 데이터 관리까지 병행하고 있어 체계적인 관리 및 공유가 어려운 상황이다.

이에 그림 1과 같은 한의학 데이터 수집 및 제공과 데이터 관리자가 분리된 품질관리 방안을 제안한다. 각 대학, 병원, 제조기업 등에서 데이터를 수집 및 제공하고, 데이터 품질관리 플랫폼을 통해 표준화 DB에 의해서 데이터를 정제, 관리한다. 이렇

게 관리된 DB를 데이터 품질관리 플랫폼과 OpenAPI를 통해 기업, 병원, 대학 등과 공유할 수 있도록 한다.

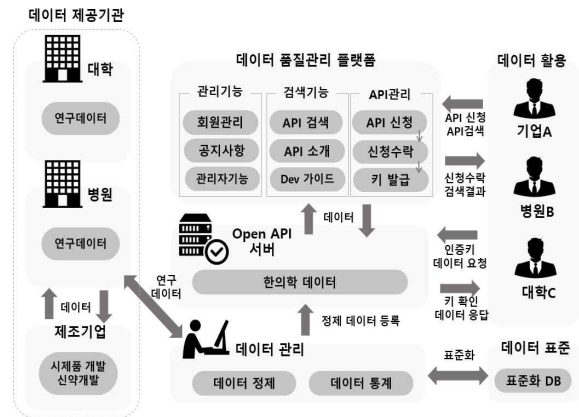


그림 1. 데이터 품질관리 플랫폼 개요도  
Fig. 1. Data quality management platform schematic

## IV. 결 론

본 논문에서는 현재 제공되고 있는 한방 데이터들의 현황을 분석하여 관리의 부실과 이로 인한 데이터 품질 유지에 대한 문제점을 도출하였다. 또한 해당 문제점을 보완할 수 있는 해결방안 두 가지를 제시하였다.

첫째 한의학 데이터 표준화한다. 둘째 데이터 품질관리 플랫폼을 통해 데이터 수집 및 제공 주체와 데이터 관리 주체를 분리하여 체계적이고 지속적으로 데이터 품질관리를 할 수 있게 하고자 한다. 해당 두 가지 방법을 통해 다양한 한방 데이터를 한 곳에서 확인하고 고품질 데이터에 대한 접근성을 높임으로써 다양한 산업 및 연구에서의 활용을 기대할 수 있다.

## 참 고 문 헌

[1] 구민석, 『輕寶新篇』의 醫案에 대한 연구, 경희대학교, pp.61, 2017  
[2] 신은주, 「세계시장 진출 위해 ‘표준’을 선점하겠다.」 『민족의학신문』, October 2016, <https://www.mjmedi.com/news/articleView.html?idxn>

o=31827

- [3] 최정희, 「한의학 표준화 어떻게 해야 하나」  
『민족의학신문』, June 2014
- [4] 한국한의약진흥원, [https://nikom.or.kr/koms/module/komsExtra/index.do?menu\\_idx=21](https://nikom.or.kr/koms/module/komsExtra/index.do?menu_idx=21), October 2024
- [5] 한의약소재은행, "Three-dimensional shape  
<https://nikom.or.kr/kmmmb/index.do>, October 2024
- [6] 국가한의임상정보포털, <https://nicom.or.kr/nckm/index.do>, October 2024
- [7] 한국전통지식포털,<https://koreantk.com/ktkp2014/>,  
October 2024
- [8] 오아시스, <https://oasis.kiom.re.kr/index.jsp>, October  
2024
- [9] 한약표준자원은행, <https://oasis.kiom.re.kr/herblib/>,  
October 2024
- [10] 한의약융합연구정보센터,<https://www.kmcric.com/>  
, October 2024
- [11] 한국한의약연구원,<https://kiom.re.kr/main/>,  
October 2024
- [12] 한의학고전DB, <https://mediclassics.kr/>, October  
2024
- [13] 한의임상정보서비스, <https://www.kmpedia.kr/>,  
October 2024
- [14] 한의임상정보은행, <https://www.kmpedia.kr/>,  
October 2024
- [15] 한의시맨틱검색, <https://cis.kiom.re.kr/semantic/>,  
October 2024
- [16] 한의온라인정책서비스,<https://policy.kiom.re.kr/>,  
October 2024
- [17] 한국생명공학연구원, <https://www.kribb.re.kr>,  
October 2024

# 다중 Agent와 반복적 답변 생성·평가 기법을 적용한 사규 검색용 Self-Corrective RAG 시스템 연구

정진철\*, 김대호\*\*, 정유철\*\*\*

## Research on a Self-Corrective RAG for Regulation Search Applying Multi-Agent and Iterative Generation and Evaluation

Jinchol Jeong\*, Daeho Kim\*\*, and Yuchul Jung\*\*\*

### 요 약

본 연구는 회사 내 복잡하고 빈번하게 개정되는 규정을 효율적으로 검색하고 신뢰성 있는 답변을 제공하기 위한 Retrieval Augmented Generation(RAG) 시스템을 제안한다. 전통적인 키워드 기반 검색의 한계를 극복하기 위해 LLM(Large Language Model)을 활용한 자연어 질의응답 방식이 주목받고 있으나, 외부의 상용 LLM을 활용하기에는 내부 민감 정보 유출, 망분리 문제, 높은 API 비용 등의 문제가 있다. 이에 따라 오픈소스 LLM을 활용한 Private RAG 시스템 구축이 필요하지만, 답변의 품질이나 성능의 한계가 존재한다. 본 연구에서는 다중 오픈소스 LLM을 활용하여 답변을 생성하고, 단답형 채점 방식으로 정량적으로 답변을 평가하여 결과에 따라 사용자 질의 재작성, 답변 생성을 반복하여 최적의 답변을 도출하는 Self-Corrective RAG 시스템을 구축하였다. 실험 결과, 제안 시스템은 회사 규정과 같은 복잡한 문서를 효과적으로 검색하고 정확한 답변을 제공함을 확인하였다.

### Abstract

This study proposes a Retrieval Augmented Generation (RAG) system to efficiently search for and provide reliable answers to complex and frequently updated company regulations. While natural language question-answering using Large Language Models (LLMs) is gaining attention as an alternative to keyword searches, using external commercial LLMs raises concerns such as sensitive data leakage, network separation, and high API costs. Thus, a private RAG system using open-source LLMs is needed, despite limitations in answer quality and performance. This study built a system that generates answers using multiple open-source LLMs (i.e., Multi-Agent) and evaluates them through short-answer grading. The system iteratively rewrites queries, regenerates answers, and evaluates them to find the best response. Experimental results show that it effectively searches complex documents and provides accurate answers.

### Key words

RAG, Self-Corrective RAG, Multi-Agent, Iterative, Open-source LLM, Evaluate, Regulation

---

\* 국립금오공과대학교 디지털융복합공학과 석사과정/한국전력기술, jinchol81@naver.com

\*\* 한국전력기술, kdaeho27@gmail.com

\*\*\* 국립금오공과대학교 인공지능공학과 부교수(교신저자), ivc@kumoh.ac.kr

## I. 서 론

회사 내에는 복잡하고 다양한 규정들이 존재하며, 이는 효율적인 업무 진행 및 직원들의 업무 수행을 보장하는 데 중요한 역할을 한다. 이러한 규정들은 분산되어 있고, 빈번히 개정됨에 따라 규정들에 따른 업무 수행을 위해 관련 규정을 수시로 검색해야만 한다. 그러나 전통적인 검색은 단순한 키워드 매칭 방식을 사용하고 있어 자연어 기반의 복잡한 질의로 원하는 정보를 찾는 데 한계가 있다[1].

최근 ChatGPT 같은 LLM(Large Language Model) 기반의 등장으로 자연어를 활용한 질의응답 형식의 검색은 높은 편리성과 유용성으로 주목받고 있다[3][4]. 그러나 LLM은 언어 이해와 생성에서 뛰어난 성능을 보이지만, 오래된 지식, 도메인 지식 부족, 환각(Hallucination) 현상, 답변 근거 제시 문제로 회사 내 적용에는 문제가 있다. 이러한 단점을 극복하기 위해 도메인 지식 및 근거를 참고하여 답변을 생성하고, 생성된 답변의 근거 제시함으로써 신뢰성 있는 답변을 제공할 수 있는 RAG(Retrieval Augmented Generation) 기술이 등장했다[5].

현재 OpenAI의 LLM인 ChatGPT 기반의 RAG 시스템이 가장 우수한 품질 및 성능을 보이는 것으로 알려져 있다. 그러나 ChatGPT의 회사의 외부에 위치함에 따른 내부 민감 정보의 유출 위험, 업무망과 망 분리 문제, 높은 API 비용 등 이유로 회사 내 인프라를 구축하고, 오픈소스 LLM을 사용하여 Private RAG 시스템을 도입하려는 추세이다.

## II. 관련 연구

오픈소스 LLM을 활용한 RAG 시스템은 ChatGPT를 활용한 RAG 시스템에 비해 품질이 떨어지는 문제가 존재한다. 이에 따라 답변 품질 향상을 위한 다양한 연구들이 진행되고 있다.

Mixture of Agents(MOA)는 여러 에이전트(LLM)가 협력하여 최적화된 답변을 생성하는 방식으로 제안되었으나, 여러 에이전트의 반복적인 활성화로 인해 계산량이 크게 증가하는 문제가 있다[6]. 또한 Self-RAG는 필요에 따라 정보를 추가로 검색하고 반영 토큰을 활용하여 스스로 평가하여 답변의 품질을 향상시키고자 하였다. 그러나 반영 토큰과 자

기 반영 과정이 추가되면서 계산량이 크게 증가하고 반복적인 평가로 인해 처리 시간이 길어지는 문제가 있다.[7].

본 연구에서는 이러한 문제를 해결하기 위해 다중의 오픈소스 LLM을 활용하여 후보 답변을 생성한 후, 이를 정량적으로 평가하는 Private RAG 시스템을 제안한다. 이 시스템은 답변 생성 및 평가를 반복적으로 수행하되, 경우에 따라서는 단 1회의 답변 생성으로도 최적의 답변을 도출할 수 있어 반복 횟수를 최소화하여 계산 비용을 줄일 수 있다. 이를 통해 기존 시스템 대비 성능 향상과 효율성을 동시에 추구하고자 한다.

## III. 연구 내용

### 3.1 제안하는 RAG 시스템 구성

제안하는 RAG 시스템은 그림 1과 같고, 주요 구성 요소 및 단계 다음과 같다. 각 단계에 품질 향상을 위한 방안을 적용하였다.

- Documents Chunks : 사규의 텍스트를 추출하여 임베딩을 위한 청크 단위로 나누는 단계이다. 사규는 장, 절, 조, 항, 목 등으로 구분되어 작성되어 있어, 단순한 일정한 크기 단위로 분할하지 않고 의미가 있는 조문 단위로 텍스트를 분할하여 품질 저하 문제를 개선하고자 하였다.
- Vector Embedding : 분할된 텍스트 데이터를 벡터 공간에 배치하는 단계이다.
- Retrieve : 주어진 사용자 질의(Query)에 대한 임베딩 모델(BGE\_M3)을 통해 Vector DB에서 관련 있는 컨텍스트를 선택하는 단계이다.
- Generator(LLM) : Retrieve 단계에서 선택된 컨텍스트를 바탕으로 주어진 질의에 대해 독립적인 후보 답변군을 생성하는 단계이다.
- Judge(LLM) : 생성된 후보 답변군의 품질을 평가하는 단계이다. 생성된 답변은 단답형 채점 방식으로 1~100점으로 정량적으로 평가한다.
- Query Rewrite(LLM): 품질평가 결과 95점 이상인 답변이 없을 경우, 사용자 질의를 분석하여 더 나은 품질의 답변을 얻을 수 있도록 사용자 질의를 재작성하는 단계이다.

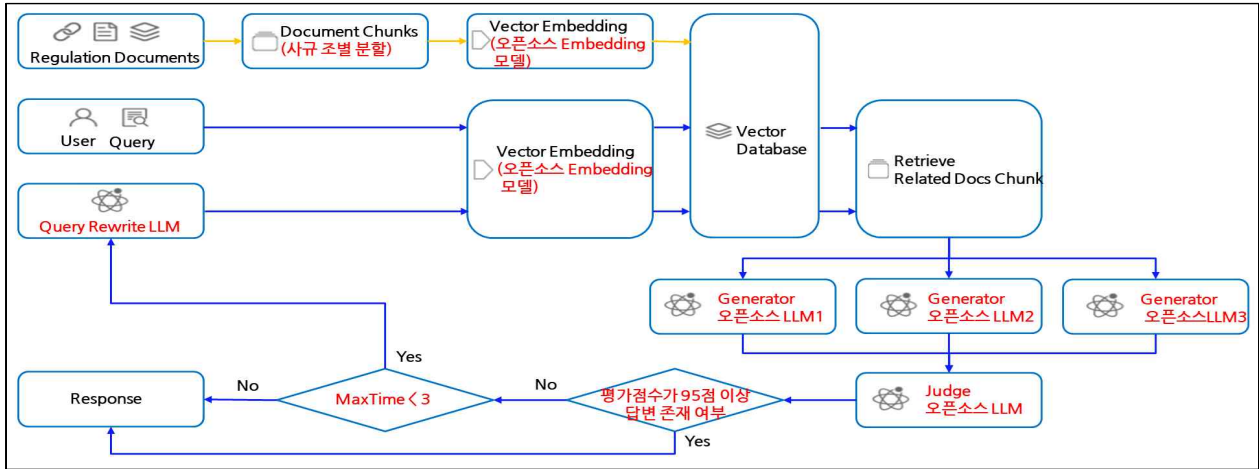


그림 1. 시스템 구성도  
Fig. 1. System Configuration

#### IV. 결론 및 향후 연구

#### 4.2 실험 결과

##### 4.1 실험 설정

본 실험에서는 아래와 같이 한국어에 우수한 오픈소스 LLM과 Embedding 모델을 중심으로 선정하여 RAG 시스템을 구성하여 진행하였다.

표 1. 실험에 적용한 LLM 및 Embedding 모델  
Table 1. LLMs and Embedding Model Applied in the Experiment

구분	명칭
Generator LLM(1~3)	Mixtral-8x22B
	Llama-3.1-70B
	Llama-3.2-90B
Judge LLM	Mixtral-8x22B
Query Rewrite LLM	Mixtral-8x22B
Embedding 모델	BGE-M3

RAG 시스템 평가는 gpt-4o를 활용하여 RAG 시스템 평가에 주로 사용되는 Metric인 Faithfulness, Answer Relevancy, Context Precision, Context Recall을 정량적으로 측정하여 수행하였으며, 이를 위해 50개의 사용자 질문 및 정답 Dataset을 구축하였다.

자산이 고장 나거나 분실된 경우, 전산 시스템에서 자산 상태 변경 절차는 어떻게 이루어지나요?
자산의 불용처리 이후 전산 시스템에 반영하는 과정은 무엇인가요?
자산 취득 시 해당 자산의 전산 기록 및 관리 절차는 어떻게 이루어지나요?
회사에 무단 접근한 직원에게는 어떤 징계가 있을 수 있나요?
직무급제가 무엇인지 알려주시고, 어떤 직무가 있고 각 직무별 직무급은 얼마 인가요?
한국전력기술에서 발전소를 건설하여 전기를 생산하여 팔수 있나요?

그림 2. 사용자 질문 예시  
Fig. 2. User Query Sample

제안한 RAG 시스템은 사규 질의에 대한 최종 답변의 품질평가 결과, 평가기준 4개의 평균점수가 0.957점(Faithfulness : 0.933점, Answer Relevancy : 0.972 점, Context Precision : 0.934점, Context Recall : 0.987 점)으로 비교 대상의 평균점수 대비 적게는 0.014점에서 많게는 0.259점까지 높은 품질을 보였다.

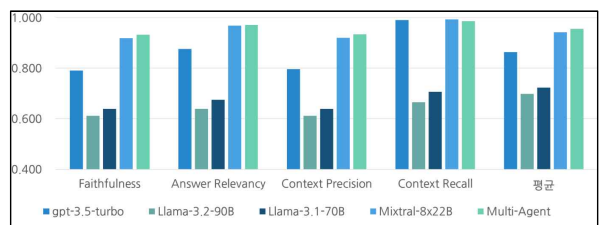


그림 3. 품질평가 결과  
Fig. 3. Quality Evaluation Result

Mixtral-8x22B도 모든 기준에서 준수한 성과를 보였고, gpt-3.5-turbo도 Context Recall에서 우수한 성과를 보였으나 제안한 RAG 시스템에는 미치지 못했다.

이에 따라 제안한 RAG 시스템이 사규 질의에 대한 보다 신뢰성 있는 답변을 생성한다는 것을 확인 할 수 있었다.

#### 4.3 향후 연구

향후 대규모 LLM을 대신하여 sLLM(Small-scale LLM)을 적용한 연구를 수행할 예정이다. sLLM은

경량화된 모델로, 더 빠르고 효율적인 응답을 제공하는 동시에, 시스템 처리 속도와 자원 효율성을 높여주는 장점이 있다. 또한, 대규모 LLM에 비해 인프라 구축 및 운영 비용을 크게 절감할 수 있을 것이다.

또한, 사규 외 다양한 도메인에서 제안하는 RAG 시스템의 성능을 비교하는 연구를 진행하여 본 시스템의 적용 가능성을 확장할 예정이다.

- A., & Gurevych, I. (2021). Beir: A heterogenous benchmark for zero-shot evaluation of information retrieval models. arXiv preprint arXiv:2104.08663.
- [10] Es, S., James, J., Espinosa-Anke, L., & Schockaert, S. (2023). Ragas: Automated evaluation of retrieval augmented generation. arXiv preprint arXiv:2309.15217.

## 참 고 문 헌

- [1] Singhal, A. (2001). Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.*, 24(4), 35-43.
- [2] Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. arXiv preprint arXiv:2303.18223.
- [3] Ai, Q., Bai, T., Cao, Z., Chang, Y., Chen, J., Chen, Z., ... & Zhu, X. (2023). Information retrieval meets large language models: a strategic report from chinese ir community. *AI Open*, 4, 80-90.
- [4] Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., ... & Ge, B. (2023). Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, 100017.
- [5] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. arXiv preprint arXiv:2312.10997.
- [6] Wang, J., Wang, J., Athiwaratkun, B., Zhang, C., & Zou, J. (2024). Mixture-of-Agents Enhances Large Language Model Capabilities. arXiv preprint arXiv:2406.04692.
- [7] Asai, A., Wu, Z., Wang, Y., Sil, A., & Hajishirzi, H. (2023). Self-rag: Learning to retrieve, generate, and critique through self-reflection. arXiv preprint arXiv:2310.11511.
- [8] Ma, X., Gong, Y., He, P., Zhao, H., & Duan, N. (2023). Query rewriting for retrieval-augmented large language models. arXiv preprint arXiv:2305.14283.
- [9] Thakur, N., Reimers, N., Rücklé, A., Srivastava,

# CNN 기반의 튜브 열화 등급 평가에 관한 연구

정용득\*, 오병진\*\*

## A study on the evaluation of tube degeneration based on CNN

Jung Youngdeug\*, and Oh Byungjin\*\*

### 요 약

발전소 보일러 튜브의 열화는 튜브의 수명을 감소시키고 고장정지의 원인이 된다. 현재는 손상된 튜브의 열화등급 평가를 위해 미세조직 이미지를 육안으로 평가하는 방식인데 이 방식은 평가자에 따라 주관적이며 비정량적인 방식으로 그 신뢰도가 낮다. 본 논문에서는 이러한 문제점을 최신 AI 알고리즘으로 해결하고자 연구를 진행하였다. 보일러 고온 고압부 설비에서 주로 사용되고 있는 재질로 시편을 만들어 온도/압력을 다르게 하여 가속열화 시험을 하였다. 이를 통해 얻은 이미지를 CNN 기반의 확장된 알고리즘에 실험한 결과, 이미지 판별(열화등급) 정확도가 92%로 높고 손실은 0.22의 결과를 얻을 수 있었다. 향후 CNN 기반 확장 모델을 활용하여 발전설비 보일러 튜브의 열화도 평가에 적용한다면, 설비진단 시 등급 평가의 정량화를 체계화하여 보일러 설비의 유지보수를 최적기에 할 수 있어 안정적인 설비운영에 기여하고, 신규 소재의 경우 추가 Data를 확보하여 활용한다면 동종업무 종사자들이 간편하고 신속하게 정량적인 잔존 수명 평가하는데 기여할 것이다.

### Abstract

The degeneration of the power plant boiler tube lowers the efficiency and causes the operation to stop. The method of visually evaluating the image for the evaluation of the corrosion grade of the damaged tube and facility part is a subjective and qualitative method according to the evaluator. In this paper, a study was conducted to solve this problem with the latest AI algorithm. As a result of experimenting with sample images collected from the degenerated facility in an extended algorithm based on CNN, it was possible to obtain results with high image discrimination accuracy and low loss. In the future, if the CNN-based extended model is used to evaluate the degeneration of the facility, it will be possible to maintain the facility and evaluate the expected life of the new material through the diagnosis, quantification and systematization of the degeneration facility.

### Key words

재질열화(Material Degradation), CNN(Convolution neural network), 발전 보일러 튜브(power plant boiler tube)

## 1. 서 론

화력발전소 보일러 튜브들은 높은 압력과 온도에

노출되어 열화가 발생되기 쉽다. 발전용 보일러 설비의 튜브 열화는 설비 수명을 떨어뜨리고 고장정지의 주된 원인이 된다[6]. 고온부에 장기간 노출된

\*정보기술품질연구원, ansonjung@naver.com, \*\*(주)비텍이엔씨, obj1222@naver.com.

\*한국데이터산업진흥원의 "2024년 데이터바우처사업" 지원의 성과임



튜브는 열화에 의하여 크리프 손상이 지속적으로 발생된다[6]. 크리프 손상은 시간에 따라 고온, 고압 환경에서 변형과 균열을 가져오며, 이러한 환경에서 장시간 노출되면 해당 부품의 강도가 또한 저하되어 이는 설비의 수명을 단축시키는 원인이 된다. 보일러 튜브 및 배관에서 크리프 손상이 발생하게 되면 미세조직에서 기공의 크기가 점점 커지면서 균열로 진전하게 되고, 탄화물의 크기가 증가해 탄화물 사이의 간격이 커지면서 경도 값이 감소하게 된다. 발전용 보일러 설비에 사용되는 재료를 가속 열화 시험하여 취득한 미세조직 이미지는 안정적인 발전소 운영을 위한 열화등급의 비교 데이터로 활용되기 때문에 빅데이터 처리와 AI알고리즘을 활용한 정량적이고 체계화된 분석이 필요하다, 하지만 튜브의 열화도나 크랙 손상 정도를 육안으로 평가하는 방식은 평가 전문가에 따라서 달라질 수밖에 없는 정성적 평가라는 점이다. 이에 따라 본 논문에서는 정량적이고, 체계화된 등급 평가 기준을 제시하고자 CNN 알고리즘을 적용한 이미지 판별기법을 방안을 제안하고자 한다.

### 1.1 연구의 범위 및 실험 데이터 확보

화력발전소에서 중요 설비인 보일러 튜브는 600°C 이상의 고온과 250kg/cm<sup>2</sup> 내외의 고압 환경에서 사용되는 설비로 고온에서 열화나 균열이 발생되며[2][4], 열화도의 정도를 튜브파열사고 전에 정량적으로 파악할 수 있다면 최적기에 튜브를 교체하거나 보수를 함으로써 발전소의 유지보수 비용을 획기적으로 절감할 수 있을 것이다.

본 논문의 연구 범위는 발전소 보일러 튜브를 가속열화 시험하여 그 미세조직을 광학현미경을 통해 이미지를 확보하였다. 이미지 데이터를 정교한 정제 과정을 거친 학습데이터로 만들어 CNN 기반 ResNet 모델에 적합하도록 훈련시킨 후 테스트, 이미지 데이터의 증강과 분석, 특징 추출을 반복하여 최적의 결과를 데이터간의 유사도를 분석하여 모델의 정확도와 손실함수 정도를 파악하였다.

고온 및 고압 조건에서 사용되는 발전소 보일러 튜브의 재료 특성을 반영하기 위해 다음과 같은 방

법을 사용하여 실험 데이터를 확보하였다.

- ① 저합금강인 Cr강 및 Cr-Mo강의 가속 열화 시킨 테스트용 시료
- ② SA213-T22, SA213-T23 시험용 시편은 가속시험 전에 용접부 비교용으로 제작한 시료
- ③ 가속열화시험에 따른 재료의 정보를 얻기 위한 인장 및 경도시험, 열화 재료의 모재 및 용접부의 모든 물성을 감안한 시료



그림 1. 튜브 조직(이미지) 사진 수집  
Fig. 1. tube image

### 1.2. Tensorflow 기반 CNN ResNet알고리즘

장기간 사용한 보일러 튜브의 경우 열화정도에 따른 미세조직의 등급분류가 필수적이나 진단업무를 수행하는 엔지니어마다 등급분류 판단이 다소 달라서 신뢰성이 떨어진다는 문제점이 존재한다.

이러한 문제점을 해결하고자 다양한 재질별로 열화 등급이 분류된 튜브의 미세조직 이미지를 수집, 분류하여 AI판별 모델에 훈련시키고, 또한 장기간 고온에서 노출된 튜브샘플이나 표면복제 시험을 통해 확보한 이미지를 Tensorflow에 기반한 CNN모델에 훈련시킨 후 CNN 이미지 분류 모델을 통해 이미지 등급을 평가할 수 있다면, 빠르고 정확하게 정량적인 판별을 할 수 있을 것이다.[9] 이미지 분류에 정확도가 높은 CNN ResNet 모델을 활용하여 향후 보일러 튜브의 미세조직의 열화등급을 분류하여 균열 진전여부와 잔존 수명을 예측할 수 있다는 장점이 있으며, 분류모델을 활용하면 고속연자가 아니더라도 업무수행이 가능하며 여러 유사 업무 관련회사에서 용이하게 활용할 수 있을 것이다.

본 논문의 구성은 다음과 같다. 제 2장에서는 이미지 분석을 위한 시스템 설계에 대해서 살펴본다. 제3장에서는 실험 및 결과, 제4장은 결론을 기술한다.

## II. 데이터 확보 비파괴 검사 및 실험환경

### 2.1 재질의 비파괴 검사 방법

고온, 고압 환경에 노출되는 보일러 설비 재료의 비파괴 시험(NDT: Non Destructive Testing)으로 조직분석법을 사용해왔으며 추가적으로 이 기술을 보완할 수 있는 금속자기 기억기술(Metal Magnetic Memory), 자기 카메라(Electro Magnetic Camera), 위상배열초음파탐상검사(Phased Array Ultrasonic Test) 기술이 적용되고 있다[4][5]. 배열 회수 보일러 Header Stub Tube 용접부 균열 검사기법인 MMM 비파괴진단법은 그 신뢰성이 향상되고 있는 단계이며, 튜브의 외면은 편이 부착되어 있어 사람이 접근하기 어려운 환경의 경우 내시경검사 기법으로 즉 육안으로 튜브 내부에 유동가속부식 등의 진전 여부를 확인하는 기법도 함께 사용하고 있다.

### 2.2 실험환경 구축 및 절차

표 1. 실험 환경

Table 1. Experimental environment

실험 컴퓨터 환경(H/W)	- Intel Core i7 - RAM : 32GB - Google Colab
프로그램 도구(S/W) 및 실행환경	- Tensorflow
프로그램 언어	- JAVA - Python

CNN기반 모델의 학습과 테스트, 검증, 최적화 모델 성능 개선의 절차는 다음과 같이 수립하였다.

- 가) 학습데이터(Training Data) 정의 및 DB 구축
  - 튜브 열화 상태의 식별 이미지 분류를 위한 데이터를 정의하기 위해서는 스타일들을 활용하여 클래스를 정의한다.
- 나) 입력-학습 데이터 셋 정제, 시각화, 표준화
- 다) 모델 만들어서 컴파일, 요약(Summary)
- 라) 학습데이터를 모델에 훈련, 훈련결과 시각화 (정확도, 손실율) 그래프 표출
- 마) 테스트 이미지를 모델에 적용, 테스트 실시
- 바) 모델을 개선하고 프로세스 반복
  - 데이터 증강 및 드롭아웃과 모델성능 개선

표 2. 실험 데이터 세트

Table 2. Experimental data sets

데이터 종류	데이터 원천	데이터 개수	데이터 유형
이미지(jpg)	튜브 조직사진	500	7종 106가지

본 논문에서는 [표2]와 같이 총 7종의 106가지 유형 500개 이미지를 확보하여 모델의 특징추출(feature extraction) 하는 방식과 데이터 증강방식을 사용하였다. 여기에 사전 학습된 모델을 미세 조정(fine tuning)하는 방법과 정확도, 손실함수를 계산하여 모델의 최적화를 위해 [그림 2]와 같은 CNN 기반 모델 흐름도(Workflow)를 적용하였다.

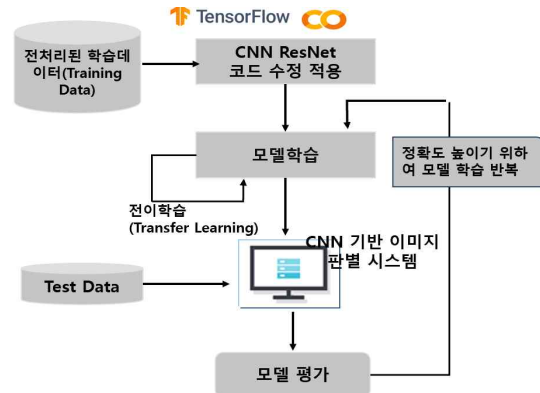


그림 2. CNN 기반 모델학습 흐름도

Fig. 2. Workflow based on CNN, Tensorflow

### 2.3. CNN 인공신경망 기법

CNN은 [그림 3]과 같이 컨벌루션층(Convolution Layer), 필터(Filter), 활성화함수(Activation function), 풀링(Pooling) 등으로 구성되어 결과를 예측할 수 있다. 컨벌루션 층에서는 필터를 통해 입력 데이터로부터 특징(Feature)을 추출하며 필터는 주요 특징이 데이터에 있는지 없는지를 검출해 주는 역할을 한다[1]. 활성화 함수는 결과 값을 참 또는 거짓으로 나타내지 않고 참에 가까워지면 0.5~1사이에서 1에 가까운 값을 부여하고, 거짓에 가까우면 0~0.5 사이의 값을 부여한다. 컨벌루션 합성곱(CNN)에서는 기본적으로 ReLu 함수를 활성화 함수로 사용한다. 샘플링 또는 풀링에서는 컨벌루션 층에서 어느 정도 추출된 특징을 기준으로 전체 데이터의 사이즈를 줄이고 오버피팅(Over Fitting)을 방지한다. 컨

별루션 필터와 활성화함수 (ReLU) 그리고 풀링층을 반복적으로 조합하고 특징을 추출하는 과정을 반복하여 결과 값에 대한 정확도를 높일 수 있다[1].

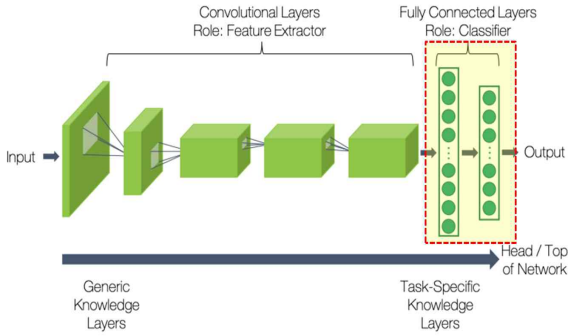


그림 3. CNN 알고리즘 처리 레이어[1]  
Fig. 3. CNN process Layer

### III. 실험 결과

#### 3.1 실험 준비

본 논문에서는 이미지 데이터의 분류 실험을 위한 테스트 데이터(test data)가 훈련데이터(Training data)와의 중복을 피하고자 현장과 똑같은 재질을 여러 온도조건에서 가속열화 시험하여 얻은 샘플을 광학현미경으로 분석하여 사용하였다. 이미지 분류 모델의 학습 반복 횟수(Training step)는 적절한 값을 찾기를 반복 수행하여 가장 높은 정확도를 보인 학습 반복 횟수(epoch)를 모델로 채택하였다.

#### 3.2 이미지의 데이터 증강(argumentation)

데이터 확보에 제한적인 환경에 따라 이 문제를 해결하기 위해 데이터 증강 기법을 활용하였으며, 기존 데이터의 다양한 변형을 통해 확보한 이미지 데이터 세트가 부족할 경우, 더 많은 데이터 세트를 확보하고, 모델이 더 다양한 특성을 반영할 수 있다[3]. 증강 데이터는 모델이 확보한 데이터를 효과적으로 일반화시킴으로서 실험 환경에서 성능을 개선하는데 도움이 된다[1][9].

또한 훈련을 위해 대량의 샘플을 만들어 데이터를 수집하고 준비하려면 비용과 시간이 많이 소요

됨에 따라, 데이터 증강 기법은 소규모 데이터 세트의 효율성을 높여 훈련 환경에서 대규모 데이터 세트에 대한 의존도를 줄여 줄 수 있다. 더불어 모델이 한정된 데이터 세트로만 훈련하면, 과적합되어 특정 데이터 유형에만 높은 예측력을 보여주는 단점을 완화할 수 있다[7][8].

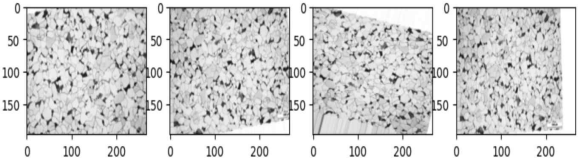


그림 4. 학습데이터 증강  
Fig. 4. Training Data Argumentation

#### 3.3 실험결과 - 이미지 유사도 분석

실험결과를 Tensorflow 기반 CNN 모델의 이미지 간 유사도를 통해 각각의 이미지를 고차원 벡터로 엠베딩(embedding) 시키고 벡터 간의 거리를 측정하는 방식을 사용하였다.

이미지 유사도 방식은 이미지 특징 추출(Feature Extraction) 후 이미지 간 비교를 위해 두 벡터(vector) 간의 코사인 유사도(Cosine Similarity)를 계산한다. 코사인 유사도(Cosine Similarity)는 같은 방향을 가지는 벡터(vector)는 1, 반대 방향을 가지는 벡터(vector)는 -1을 가지는 동작으로 테스트 이미지와 코사인 유사도(Cosine Similarity)가 높은 순으로 정렬함으로써 비슷한 이미지를 얻을 수 있다.[8] 여기서 얻은 코사인 유사도(Cosine Similarity)로 인덱싱을 통해 구현한 CNN 이미지 분류 시스템에 테스트 이미지로 유사 이미지를 추천하는 결과를 구하는 방식을 사용하였다.

이미지 간의 유사도 측정은 두 벡터의 내적(Dot product)을 구하여 코사인 유사도 방식으로 확인한 결과 -0.0072828로 서로 다른 이미지임을 확인할 수 있다. [그림 5]과 같이 훈련정확도와 검증정확도는 약 92%이며, 훈련손실과 검증 손실(loss)의 차이는 약 0.22로 측정되어 비교적 높은 정확도와 낮은 손실 함수 값을 도출하였다.

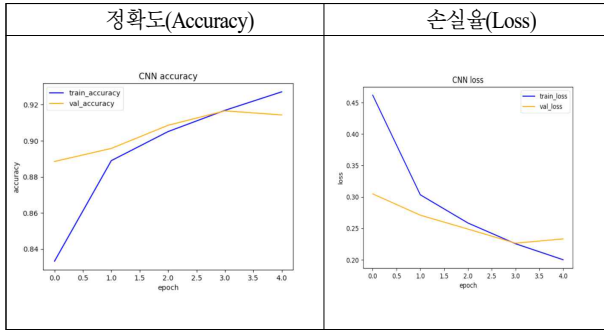


그림 5. 정확도 및 손실값 측정결과  
Fig. 5. Result for Accuracy and Loss

### 3.4 이미지 데이터 모델 성능 검증 결과

CNN기반의 모델을 사용한 실험 결과 이미지 데이터에 대한 성능과 검증 결과는 [표3]과 같다.

표 3. 이미지 성능 검증 결과  
Table 3. Result of Experimental Performance

Training Accuracy	Training Loss
0.9351458549499512	0.17346954345703125
Validation Accuracy	Validation Loss
0.9165833592414856	0.22636111080646515

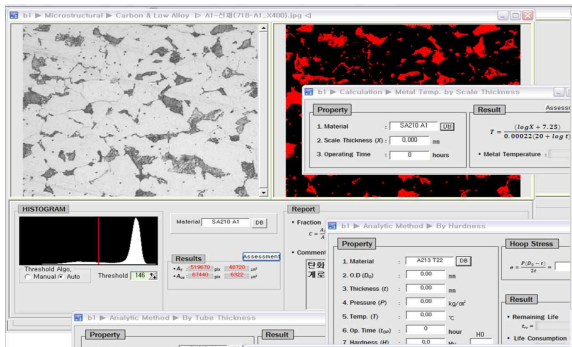


그림 4. 시스템 실험화면  
Fig. 4. Parameter Setting in DB system

## IV. 결 론

본 논문에서 실험한 시스템에서는 발전소 내부 설비에서 확보한 이미지를 광학현미경으로 확대 촬영한 이미지를 사용하였으며, 특징 추출을 통해 기존 크랙 이미지의 분석 평가 방법의 한계를 극복하기 위해 CNN기반 ResNet 모델을 활용하여 이미지 분류의 정확성과 손실함수 값을 측정하였다.

본 연구에서는 기존 열화 등급 평가 전문가의 개

입이나 노력이 필요 없는 이미지 학습 데이터를 이용하기 때문에, CNN 기반의 딥러닝 분석을 활용한 연구가 보다 보편화될 것이며, 개선된 CNN 알고리즘은 더욱 높아진 분석 정확도와 효율성으로 스마트팩토리와 같은 분야에서 활용도가 높아질 것으로 예상된다.

보일러 튜브 튜브의 열화도 검사 및 잔존수명 평가를 위해 CNN 기반 알고리즘을 이용하여 비교적 간편하게 적기에 튜브의 교체시기 등을 정량적으로 판단해 주는 비즈니스 모델로도 활용할 수 있을 것으로 기대한다.

## 참 고 문 헌

- [1] Anirudha Ghosh, AbuSu, Farhana, DebashisDe, "Fundamental Concepts of Convolutional Neural Network", pp.519-567, January 2020.
- [2] J.C. Vaillanta, B. Vandenberghe, B. Hahnc, H. Heuserd, C. Jochumd, "T/P23, 24, 911 and 92: New grades for advanced coal-fired power plants—Properties and experience" International Journal of Pressure Vessels and Piping 85 (2008) pp. 38-46, June 2007.
- [3] Journal of Knowledge Information Technology and Systems(JKITS), Vol. 13, No. 5, pp. 589-595, October 2018.
- [4] Takashi Watanabe, Masaaki Tabuchi, Masayoshi Yamazaki, Hiromichi Hongo, Tatsuhiro Tanabe "Creep damage evaluation of 9Cr-1Mo-V-Nb steel welded joints showing Type IV fracture", International Journal of Pressure Vessels and Piping 83 (2006) pp. 63-71, September 2005.
- [5] 박종서 외 2인, '고온설비 수명평가를 위한 표면복제 절차의 표준화', 대한기계학회논문집, Vol. 24, No. 9, pp. 2381-2386, 2000
- [6] 오병진, "발전설비 보일러 튜브 재료별 고온에서 모재와 용접부의 재질 열화 및 물성 변화에 대한 연구", 한양대학교 대학원 석사논문, 2011년 8월
- [7] 함성곱 신경망(CNN)기반 이미지 처리 시스템", 허이룬, 배제대학교 박사학위 논문, 2018년
- [8] <https://www.tensorflow.org/>
- [9] <https://wikidocs.net/28646>

# 가명 처리를 통한 법적 규제 및 인증 심사의 효율적 대응 전략 연구

함도윤\*, 유혜정\*\*

## An Efficient Response Strategy Study for Legal Regulations and Certification Audits Through Pseudonymization of Personal Information

Doyun Ham\*, and Hyejeong Yoo\*\*

### 요 약

오늘날 각국의 개인정보 보호 법률은 기업의 데이터 활용을 제한하는 동시에 강력한 보호 조치를 요구하고 있으며, 이에 따라 가명 정보는 법적 요구사항과 데이터 활용 간의 균형을 맞추는 핵심 수단으로 주목받고 있다. 본 연구에서는 유럽연합(GDPR), 미국(CCPA), 일본(APPI) 등 주요 법적 규제에서 요구하는 가명 처리 방법을 분석하여, 기업이 개인정보와 관련된 법적 리스크를 최소화하고 주요 인증 심사(ISMS-P, ISO/IEC 27701 등)에 효율적으로 대응함으로써 궁극적으로 개인정보 보호의 법적 테두리 내에서 데이터 활용을 극대화할 수 있는 전략을 제시하고자 한다.

### Abstract

Today, each country's privacy laws require strong protective measures while limiting the use of data by companies, and pseudonym information is attracting attention as a key means of balancing legal requirements and data utilization. This study analyzes pseudonymization methods required by major legal regulations such as GDPR (EU), CCPA (US), and APPI (Japan) to suggest strategies that can ultimately maximize data utilization within the legal boundaries of personal information protection by minimizing legal risks related to personal information and responding efficiently to major certification audits (ISMS-P, ISO/IEC 27701, etc.).

### Key words

pseudonymization, data protection, GDPR, CCPA, APPI, certification audits, legal requirements

### 1. 서 론

오늘날 디지털 환경이 급격하게 발전함에 따라 개인정보 보호에 대한 요구가 강화되고 있다. 각국

의 개인정보 보호 법률은 기업이 데이터를 안전하게 관리하면서도 법적 요구사항을 준수하도록 요구하고 있으며, 특히 가명 정보는 개인정보 보호와 데이터 활용의 균형을 맞출 수 있는 주요 기술로 주목받고 있다.

\* 세종사이버대학교 정보보호대학원, rhdaneh1957@naver.com,

\*\* 세종사이버대학교 정보보호학과, hjyoo@sjuc.ac.kr(교신저자)

가명 정보는 데이터의 일부 식별자를 제거하거나 변형하여 특정 개인을 식별할 수 없도록 한 정보를 의미하며, 이를 통해 데이터의 유출 위험을 줄이고 법적 요구사항을 준수할 수 있다. 유럽연합의 GDPR, 미국의 CCPA, 일본의 APPI와 같은 주요 법적 규제들은 각각 가명 처리에 대한 기준을 제시하며, 데이터 보안과 개인 프라이버시를 강화하기 위한 다양한 접근 방식을 요구한다.

따라서 기업은 이러한 법적 요구사항을 효과적으로 준수하기 위해 가명 처리 전략을 수립하고, 이를 통해 법적 규제 및 인증 심사(ISMS-P, ISO/IEC 27701 등)에 대응할 필요가 있다. 본 연구는 가명 처리 기술을 활용하여 법적 요구사항을 효율적으로 준수할 수 있는 구체적인 전략을 제시하고, 이를 통해 기업이 개인정보 보호의 법적 테두리 내에서 데이터 활용을 극대화할 수 있는 방안을 제시하는 데 그 목적이 있다.

표 1. 주요 국가의 개인정보 보호 법률과 가명 처리 요구사항 비교

Table 1. Comparison of Major Countries' Personal Information Protection Laws and Pseudonymization Requirements

법률	국가	가명 처리 요구사항	주요 요구사항
GDPR	유럽 연합	가명 처리된 데이터에 대한 재식별 위험 최소화	데이터 보호 조치 (암호화, 접근 통제 등), 재식별 평가
CCPA	미국	소비자 프라이버시 권리 보호를 위한 가명 처리	데이터 결합 시 재식별 위험 평가, 접근 제한 조치
APPI	일본	재식별 가능성 최소화 및 데이터 보관·폐기 관리	재식별 위험 평가, 접근 권한 제한

## II. 가명 처리를 통한 법적 규제 및 인증 심사 대응 전략

가명 처리에는 데이터의 일부 식별자를 제거하거나 변형하여 특정 개인을 식별할 수 없도록 하는 기술이 포함되며, 이를 통해 법적 요구사항 준수와

데이터 활용 효율성을 동시에 달성할 수 있다. 본 연구의 가명 처리 전략은 다음의 주요 사항들을 고려하여 수립되었다:

### ■ 법적 요구사항 분석

주요 국가의 개인정보 보호법에 명시된 가명 처리 규정은 데이터의 재식별 위험을 최소화하도록 요구한다. 예를 들어, GDPR 제4조와 CCPA Section 1798.140은 가명 정보를 데이터 보호의 핵심으로 간주하며, 가명 처리 시 재식별 가능성을 평가하고 데이터 접근을 제한하도록 권장하고 있다.

### ■ 인증 심사 기준과의 연계

ISMS-P와 ISO/IEC 27701 같은 인증 심사는 개인정보 처리 활동의 보안성을 평가하며, 가명 처리 전략이 이러한 심사 기준을 충족할 수 있도록 설정되었다. 가명 처리된 데이터는 재식별 위험 평가 및 접근 제한 조치를 통해 인증 심사 항목에 부합하는 방안으로 제안된다.

본 연구에서는 GDPR, CCPA, APPI 등 주요 법적 규제와 국내외 인증 기준(ISMS-P, ISO/IEC27701)에서 가명 처리 대응 가능성을 산정하기 위해 다음과 같은 기준을 적용하였다.

### ■ 항목 선정 근거

각 법적 규제에서 가명 처리를 통해 대응할 수 있는 조항을 선별하였다. 예를 들어, GDPR의 경우 데이터 보호와 재식별 가능성 최소화를 위해 제4조와 제32조 등에서 가명 처리를 권장하고 있으며, CCPA의 경우 소비자 프라이버시 보호를 위해 가명 정보의 사용을 지지하고 있다.

### ■ 대응도 산정 방법

가명 처리 대응 가능성은 각 법적 규제의 전체 조항 수 대비 가명 처리로 대응 가능한 항목 수의 비율을 계산하였다:

$$\text{대응도}(\%) = (\text{가명 처리 대응 가능 항목 수} / \text{전체 조항 수}) * 100$$

위 주요 사항을 고려하여 수립된 가명 처리 전략을 통해 글로벌 개인정보 보호법(GDPR, CCPA, APPI 등)과 국내외 인증 심사(ISMS-P, ISO/IEC 27701)에 얼마만큼 효과적으로 대응하는지는 다음과 같다.

1. GDPR (유럽연합) - 15% 대응 가능

GDPR은 가명 정보를 개인정보로 간주하며, 데이터를 처리할 때 재식별 가능성을 최소화할 것을 요구한다. 가명 정보 처리를 통해 GDPR의 약 15%에 해당하는 요구사항을 제한적으로 충족할 수 있으며, 이는 주로 기초적인 데이터 보호 조치와 일부 프라이버시 관리 요건을 개선하는 데 기여한다. 이를 통해 기업은 유럽연합의 일부 법적 요구에 대응하면서도 데이터 보호 수준을 최소한으로 유지할 수 있다.

2. CCPA (미국 캘리포니아주) - 10% 대응 가능

CCPA는 소비자의 프라이버시 권리를 보호하기 위해 가명 정보의 사용을 권장하며, 데이터 결합 시 재식별 가능성을 최소화할 것을 요구한다. 가명 처리를 통해 CCPA 요구사항의 약 10%에 해당하는 항목을 준수할 수 있으며, 이는 소비자 프라이버시 보호와 관련한 제한적인 수준의 기여를 할 수 있다. 결과적으로, 기업은 캘리포니아주의 최소한의 법적 요구사항에 대응하여 프라이버시 보호를 강화할 수 있다.

3. APPI (일본) - 12% 대응 가능

일본의 개인정보 보호법인 APPI는 데이터를 처리할 때 재식별 가능성을 최소화하고, 데이터 보관 시 보안을 강화할 것을 요구한다. 가명 처리는 APPI 요구사항 중 약 12%를 충족하는 데 기여할 수 있으며, 이는 주로 데이터 보관 및 접근 제어에서 최소 수준의 보안을 제공하는 역할을 한다. 이를 통해 기업은 일본 법적 요구사항에 일부 대응하여 데이터의 기초적인 보안을 유지할 수 있다.

4. 국내 인증 심사(ISMS-P) - 18% 대응 가능

국내의 정보보호 및 개인정보보호 관리체계

(ISMS-P)는 개인정보 처리 및 보호 조치의 적정성을 평가하며, 가명 처리는 ISMS-P 심사 기준의 약 18%에 대응할 수 있다. 이는 주로 데이터 보호와 접근 통제의 기초적 요구사항을 충족하는 데 기여하며, 기업의 법적 리스크를 일부 완화하고 데이터 보호의 최소한의 안전성을 제공할 수 있다.

5. ISO/IEC 27701 - 14% 대응 가능

ISO/IEC 27701은 개인정보 관리 시스템의 국제 표준으로, 가명 처리를 통해 재식별 위험 관리와 일부 보안 요구사항에 부분적으로 대응할 수 있다. 가명 처리는 ISO/IEC 27701 요구사항 중 약 14%를 충족하는 데 기여하며, 이는 데이터 접근 제한과 재식별 위험을 최소한의 수준에서 관리할 수 있도록 한다. 이를 통해 기업은 국제 표준의 기본 요건에 일부 대응하여 데이터 관리의 기초 효율성을 유지할 수 있다.

표 2. 가명 처리를 통한 법적 규제 및 인증 심사 대응도  
Table 2. The degree of response to legal regulation and certification audits through pseudonymization

분야	대응도
GDPR(EU)	15%
CCPA(미국)	10%
APPI(일본)	12%
ISMS-P(국내 인증 심사)	18%
ISO/IEC 27701	14%

### III. 결 론

본 연구에서는 가명 처리를 통해 법적 규제와 인증 심사에 효율적으로 대응할 수 있는 전략을 제안하였다. 제안된 전략은 각국의 개인정보 보호 법률과 인증 심사 기준을 기초로 하여, 가명 처리 기술을 적용함으로써 법적 요구사항을 부분적으로 준수하면서도 데이터 보호의 기초적 수준을 유지할 수 있게 한다. 본 연구의 결과, 가명 처리를 통해 제한적인 범위에서 법적 리스크를 줄이고 데이터 활용

의 안전성을 확보할 수 있는 가능성을 확인하였다. 이는 기존의 단순한 데이터 보호 방식보다 법적 대응 효율성을 개선하는 데 기여할 수 있음을 보여준다.

### 참 고 문 헌

- [1] 개인정보보호위원회, 가명정보 처리 가이드라인, 2023.1.
- [2] 국가법령정보센터, 개인정보 보호법, 2024.1.
- [3] 함현희, 가명 정보를 통한 개인 정보 활용에 대한 문제점 및 개선 방안, 2020.12.
- [4] 차훈수, 개인정보의 가명화를 고려한 ISMS-P 개선방안에 관한 연구, 2023.8.
- [5] 조재현, 개인정보 보호법상 가명정보 처리 특례에 관한 연구, 2022.8.
- [6] 오지민, 가명정보 활용 촉진을 위한 가이드라인 개선방안 연구, 2023.5.
- [7] 국가법령정보센터, 개인정보 보호법, 2024.
- [8] EU GDPR Portal, GDPR (General Data Protection Regulation).
- [9] 한국인터넷진흥원(KISA), ISMS-P 인증 기준.
- [10] ISO Official Website, ISO/IEC 27701.
- [11] 개인정보보호위원회, 개인정보보호위원회 보고서.
- [12] California Legislative Information, 캘리포니아 소비자 개인정보 보호법 (CCPA).
- [13] 개인정보보호위원회, 2021 개인정보보호 실태조사 보고서.
- [14] 개인정보보호위원회, 가명정보 처리 가이드라인, 2024.
- [15] 개인정보보호위원회, 2023 개인정보 주요 이슈 법령 해석 사례 30선, 2023.
- [16] 개인정보보호위원회, 개인정보보호 법령 및 지침 고시 해설서, 2020.
- [17] 개인정보보호위원회, 과학기술정보통신부, 보건복지부, 금융위원회, 국세청, 한국인터넷진흥원, 2023 가명정보 활용 우수사례집, 2023.
- [18] 교육부, 개인정보보호위원회, 교육분야 가명명처리 가이드라인, 2020.
- [19] 금융위원회, 금융감독원, 금융분야 가명·익명처리 안내서, 2022.
- [20] 행정안전부, 개인정보보호위원회, 공공분야 가명정보 제공 실무안내서, 2021.
- [21] 이한주, 가명정보 개념 도입을 통한 의료정보 활용 활성화 가능성의 법적 검토, 2020.6.
- [22] 이기호, 보건의료 데이터 활용 가이드라인의 내용과 한계에 관한 연구, 2020.11.



# 머신러닝 모델을 활용한 기후변화에 따른 한국의 주요 농산물 생산량 예측에 관한 연구

김윤섭\*, 함종수\*\*

## A Study on the Forecast of Major Agricultural Production in Korea by Climate Change Using Machine Learning Model

YunSub Kim\*, Jongsoo Ham\*\*

### 요 약

최근 기후변화로 인해 배추, 사과 등 주요 농산물의 생산량 불확실성이 증가하면서 가격 변동성이 크게 확대되고 있다. 본 연구는 머신러닝 기술을 활용하여 온도 상승, 강수 패턴 변화, 일조량 변화 등 기후변화 요인이 농작물 생산량에 미치는 영향을 분석하고, 이를 바탕으로 향후 10년간 주요 농산물의 생산량 변동을 예측하였다. 이를 통해 급격한 생산량 변동에 따른 가격 변동에 대한 대응 방안을 마련하기 위한 기초 자료로 활용될 수 있을 것으로 기대된다. 사용된 머신러닝 모델은 Random Forest, XGBoost, Ridge 회귀분석, 그리고 시계열 예측을 위한 Prophet 모델이며, 데이터는 각 농산물 주산지별 기후 데이터와 재배 면적을 결합하여 구축되었다. 분석 결과, Ridge 모델이 모든 농산물에서 가장 우수한 성능을 보였으며, 쌀과 배추는 생산량이 지속적으로 감소할 것으로 예측된 반면, 사과는 평년 수준의 생산량을 유지할 것으로 예상되었다.

### Abstract

Recently, the uncertainty in the production of major agricultural products such as cabbage and apples has increased due to climate change, leading to greater price volatility. This study analyzes the impact of climate change factors such as rising temperatures, changes in precipitation patterns, and variations in sunlight on crop production using machine learning techniques. Based on this analysis, we forecast the production fluctuations of major crops over the next 10 years. The findings of this study are expected to serve as foundational data for devising strategies to mitigate price fluctuations resulting from sudden changes in production. The machine learning models used in this study include Random Forest, XGBoost, Ridge regression, and the Prophet model for time series forecasting. The data were constructed by combining climate data and cultivated area information from the main production regions of each crop. The results showed that the Ridge model performed the best for all crops, predicting a continuous decline in the production of rice and cabbage, while apple production is expected to remain at average levels.

### Key words

machine learning, climate change, crop yield prediction, regression model, prophet model

---

#### \* 프리랜서

- ORCID : <https://orcid.org/0009-0009-6319>

- 이메일 : thgo34@naver.com

#### \*\* 한양대학교 인공지능융합학과(교신저자)

- ORCID : <https://orcid.org/0009-0006-8216-5248>

- 이메일 : hamionsoo0724@gmail.com

## I. 서 론

기후변화는 21세기 인류가 직면한 가장 큰 도전 과제 중 하나로, 그 영향은 지구상의 모든 생태계와 인간 활동에 걸쳐 나타나고 있다. 특히 농업 부문에서 기후변화로 인한 온도 상승, 강수 패턴 변화, 일조량 변화는 농산물 생산량에 큰 영향을 미치며, 이는 농산물의 가격 변동성과 인플레이션을 초래할 수 있다.[1][2] 한국은 전통적으로 농업에 의존하는 국가로, 쌀, 가을배추, 사과와 같은 주요 농산물의 생산량 변화는 국가 경제와 식량 안보에 중요한 역할을 한다. 이러한 농작물들의 생산량 변동은 기후변화로 인해 점차 불규칙해지고 있으며, 이는 농업 종사자와 소비자 모두에게 큰 영향을 미치고 있다. 예를 들어, 2023년에는 냉해 등 기후변화로 인해 사과 생산량이 30% 감소하고, 가격이 55% 이상 상승하는 현상이 발생했다.[3] 본 연구의 목적은 머신러닝 모델을 활용하여 기후변화가 한국의 주요 농산물 생산량에 미치는 영향을 분석하고, 향후 10년간의 생산량 변동을 예측하는 것이다. 이를 통해 농업 인플레이션에 대한 선제적인 대응 방안을 마련하고자 한다. 연구는 쌀, 가을배추, 사과의 주산지인 경상북도와 전라남도 지역의 기후 데이터를 수집하고, 생산량에 영향을 미치는 주요 요소인 재배 면적을 고려하여 예측을 진행한다. 분석 방법으로는 기후 및 재배 면적 데이터와 농산물 생산량 데이터를 결합하여 Random Forest, XGBoost, Ridge와 같은 머신러닝 회귀 모델을 활용하며, 기후 변수 간의 복잡한 상호작용을 효과적으로 다루고 비선형적 관계를 예측하는 강점을 가진다. 또한, 시계열 예측을 위해 Prophet 모델을 사용하여 추세와 계절성을 동시에 고려한 예측 데이터를 도출할 것이다. 최종적으로 본 연구는 AI 기반의 생산량 예측을 통해 농산물 부족이나 과잉 생산으로 인한 문제를 사전에 예방하고, 정책 수립 시 중요한 기초 자료로 활용될 수 있을 것으로 기대된다.

## II. 선행연구

우선 기후변화로 인한 농작물 생산량 예측에 관한 논문은 많다. 특히, 국내에서 소비가 많이 되는

쌀, 양파, 사과 등에 대해서 머신러닝을 활용하여 기후에 따라 생산량이 어떻게 변화하는지에 대한 연구가 많이 진행되었으며, 기후변화 외에 농작물 생산에 영향을 미치는 요인을 가지고 AI 모델을 활용하여 예측한 연구를 면밀히 살펴보고 참고하였다. 작물 생산량 예측을 위한 머신러닝 기법 활용 연구에서는 공공데이터 정보를 활용하여 분석기준 3개월로 하여 작물과 생환경 상태에 따른 생산량을 예측하였으며, 머신러닝 알고리즘인 Ridge Regression, Random Forest, XGboost를 사용하여 작물 생산량 예측의 적합도를 평가 분석하였고, XGboost가 MAE 0.233, RMSE 0.817로 최적모델로 선정되었다. 변수의 중요도는 재식밀도, 생장길이, 잎의 수 요인들이 출하량 예측에 가장 중요한 요소임을 확인할 수 있었다. XGboost 모델의 예측력이 가장 높았으며 약 77% 설명력을 보였다.[4] 날씨 데이터를 활용한 쌀 생산량 예측모델의 특성 선택 및 최적시기 탐색에서는 2012년부터 2021년까지 10년간의 기상청 종관 기상관측(ASOS) 데이터 중 쌀 생산량에 영향을 주는 변수를 다중선형회귀 및 다중공선성 등을 통해 선정하였다. 기상관측 29개 변수 중 6개 변수가 생산량에 큰 영향을 미치는 것을 확인했고 특히 8월의 데이터가 가장 좋은 예측 결과를 보인다는 것을 확인할 수 있었다.[5] 기후변화 관련 자연재해가 우리나라의 쌀 생산에 미치는 영향에서는 2008년~2017년까지 기후, 재해, 재배면적 등 자료를 바탕으로 회귀분석을 수행했으며, Adjusted R square 0.616 및 표준오차 19562.047의 예측결과를 보였다.[6]

## III. 자료수집 및 연구방법

### 1. 연구 데이터 수집

본 연구는 벼, 가을배추, 사과를 대상으로 하였으며, 지난 20년간 가장 생산량이 많았던 지역을 선별하였다. 벼는 전라남도에서 42만 톤, 사과는 경상북도에서 680만 톤, 가을배추는 전라남도에서 1,500만 톤이 생산되었다. 이에 따라 벼와 가을배추는 전라남도, 사과는 경상북도를 주산지로 선정하였고, 각 시군 중 생산량이 높은 해남, 영암, 나주, 고흥(벼), 안동, 청송, 문경(사과), 해남(가을배추)을 선택하였다.

## 2. 분석자료 및 방법

본 연구는 2000년부터 2023년까지 24년간의 기후 데이터를 바탕으로 한국의 주요 농산물인 쌀, 배추, 사과와의 생산량을 예측하는 데 중점을 두고 있다. 연구에 사용된 기후 데이터는 기상청의 종관기상관측(ASOS) 자료를 활용하였으며, 이 과정에서 주산지 관측소의 데이터를 수집하여 분석하였다. 총 29개의 기후 변수가 수집되었으며, 이 중 농업에 중요한 변수들을 선정하기 위해 농촌진흥청에서 제공하는 ‘농업날씨 365’[7]를 참고하였다. 이를 통해 각 지역의 평균기온, 최고기온, 최저기온, 일조시간, 일사량, 평균풍속, 초상온도, 지면온도 등 농업에 밀접한 기후 변수들이 최종적으로 선정되었다. 또한, 각 농산물의 생산량 변동을 보다 정확하게 예측하기 위해 해당 농작물의 주요 농사 기간에 맞춘 월별 데이터를 분석에 포함시켰다. 농촌진흥청에서 제공하는 ‘농사로’ 농업백서와 통계청의 국가통계포털(KOSIS)을 통해 2000년부터 2023년까지의 농작물 재배면적 및 연도별 생산량 데이터를 수집하였다. 이러한 기후 및 생산량 데이터를 기반으로 머신러닝 회귀 모델을 통해 쌀, 배추, 사과와의 생산량을 예측하였다. 머신러닝 모델로는 Ridge 회귀, RandomForest Regressor, XGBoost가 사용되었다. Ridge 회귀 모델은 다중공선성 문제를 해결하고 과적합을 방지하기 위해 L2 정규화를 적용한 선형 회귀 모델로, 기후 변수 간 복잡한 상호작용을 다루는 데 적합하며 예측 안정성을 제공한다.[8] RandomForest Regressor는 비선형 데이터를 다룰 수 있는 강력한 모델로, 여러 결정 트리를 학습하여 예측하며 변수 중요도를 계산해 각 변수의 중요성을 파악할 수 있다.[9] XGBoost는 그래디언트 부스팅 기법을 통해 트리의 오차를 수정하면서 성능을 점진적으로 향상시키는 모델로, 복잡한 변수 간 상호작용을 효과적으로 처리하며 우수한 예측 성능을 제공한다.[10] 이러한 모델들의 특성은 기후 데이터와 같은 비선형적 관계를 다루는 데 매우 유용하다. 또한, 시계열 데이터 예측을 위해 Prophet 모델이 사용되었다. Prophet 모델은 시계열 데이터의 계절성, 추세, 이벤트 등의 다양한 패턴을 학습할 수 있는 장점을 가지고 있으며, 비정상적인 기후 변동에도 강력한 예측 성능을 보인다. 본 연구에서는 각 농작물의 지역별 기후 데이터가 연간 생산량에

미치는 영향을 분석하기 위해 월별 기후 변수를 세분화하여 회귀 분석에 활용하였다. 이를 통해 4월의 평균기온, 일조시간 등 월별 기후 변수를 컬럼화하여 농작물 생산량 예측 모델을 구축하였다. 데이터 전처리 과정에서는 각 변수의 상관관계와 다중공선성 문제를 해결하기 위해 세 가지 케이스로 나누어 회귀 모델을 훈련하였다. 첫 번째 케이스는 상관계수가 0.3 이상인 변수를 선택하여 모델을 학습시키고, 두 번째 케이스는 다중공선성 문제가 적은 변수를 선택하였다. 마지막 케이스는 단계적 선택법을 사용하여 AIC(Akaike Information Criterion) 값을 기준으로 변수를 선택하였다. 이러한 과정을 통해 각 모델의 성능을 최적화하기 위해 Scaler를 적용하여 데이터를 정규화하고, Pipeline 함수를 사용하여 일관된 데이터 처리를 수행하였다. 추가적으로, 하이퍼파라미터 최적화 및 모델 성능 검증을 위해 GridSearchCV와 TimeSeriesSplit을 활용하였으며, 시계열 데이터 특성에 맞는 교차 검증을 통해 과적합(overfitting) 및 사전 관찰 문제(look-ahead bias)를 방지하고 보다 일반화된 예측 성능을 확보하고자 하였다. Prophet 모델을 통한 향후 10년간의 기후 데이터 예측은 동일한 절차로 진행되었으며, 이를 바탕으로 장기적인 기후 변화가 농작물 생산량에 미치는 영향을 분석하였다.

마지막으로, 모델 검증 단계에서는 회귀 모델의 성능을 평가하기 위해 RMSE(Root Mean Square Error)와 R<sup>2</sup>(R-squared) 지표를 사용하였으며, 이들 지표를 통해 가장 우수한 성능을 보인 모델을 최종적으로 선택하였다.

## IV. 연구결과 및 논의

### 1. 회귀모델 구축 및 검증

본 연구는 기후변화가 농작물 생산량에 미치는 영향을 예측하기 위해 다양한 회귀모델을 적용하였다. 각 모델은 상관관계, 다중공선성, 변수 선택법을 통해 적절한 변수를 선택하고, GridSearchCV를 통해 하이퍼파라미터를 최적화하였다. Ridge 회귀모델은 다중공선성 문제를 해결하고 과적합을 방지하기 위해 L2 정규화를 적용하였으며, Alpha 값을 조정하여 규제 정도를 최적화를 하였다. RandomForest 모델은 트리 개수, 깊이, 노드 수를 조정하여 비선형

데이터를 처리하였고, XGBoost 모델은 그래디언트 부스팅 기법을 사용하여 트리 개수와 학습률을 조정하였다. 모델의 성능은 R<sup>2</sup> 스코어와 RMSE 지표로 평가되었으며, 표 1에서 볼 수 있듯이 Ridge 모델이 가장 좋은 성능을 보였다. 쌀의 경우 Ridge 모델의 R<sup>2</sup> 스코어는 0.92, RMSE는 24,605톤으로 매우 높은 정확도를 보였다. 사과와 배추의 경우 R<sup>2</sup> 스코어 0.73, RMSE는 21,922톤이었고, 배추는 R<sup>2</sup> 스코어 0.90, RMSE 25,440톤으로 Ridge 모델이 모두 우수한 성능을 보였다. 특히, 월 단위로 데이터를 분석함으로써 기존 계절 단위 분석보다 성능이 약 15% 이상 향상되었다.

### 2. 시계열 모델을 통한 데이터 예측 결과

시계열 데이터 예측에는 Prophet 모델이 사용되었으며, 이 모델은 ARIMA의 한계를 보완하여 계절성, 추세, 이상치 등 다양한 패턴을 효과적으로 반영할 수 있다. Prophet 모델의 최적화는 GridSearchCV를 통해 진행되었고, 주요 파라미터로는 계절성 모드, 변곡점의 스케일, 연간 계절성 여부가 포함되었다. Prophet 모델은 RMSE 지표를 통해 평가되었으며, 향후 10년간의 기후 데이터와 재배 면적 데이터를 기반으로 농작물 생산량을 예측하였다.

벼 생산량에 영향을 미치는 주요 변수는 5월 최대풍속, 8월 최저초상온도, 일조시간, 9월 평균 상대습도였다. 벼 재배면적은 2024년 15만 ha에서 2033년 14만 ha로 감소할 것으로 예측되었으며, 최저초상온도는 23℃에서 24℃로 증가할 것으로 보인다. 사과의 경우, 2024년부터 2033년까지 상대습도와 최고기온이 증가할 것으로 예상되며, 상대습도는 77%에서 90%, 최고기온은 26℃에서 29℃로 상승할 것으로 보인다. 사과 재배면적도 500ha 증가할 것으로 예측되었으나, 5월 최저초상온도는 8.8℃에서 6.5℃로 하락할 가능성이 있다. 가을배추는 8월 일조시간이 약 15% 증가하고, 재배면적은 11% 감소할 것으로 분석되었다.

### 3. 향후 10년간 농산물 생산량 예측 결과

향후 10년간의 기후 및 재배면적 데이터를 바탕으로, 벼, 사과, 배추의 생산량 변동을 예측하였다. 그림 1에서 빨간 선 이후 향후 10년간의 생산량 추세를 확인할 수 있다. 벼는 평균 78만 톤으로, 현재 평균 89만 톤 대비 11톤 정도 감소할 것으로 보인다. 재배 면적 감소와 기후변화로 인해 급격한 생산량 감소와 가격 급등이 발생할 가능성이 크며, 생산량 감소가 주요 원인으로 작용할 것이다.

표 1. 농작물 별 회귀모델, 검증 결과, 투입 변수

Table1. Regression models by crop, verification results, input variables

농작물	모델명	R <sup>2</sup> Score /RMSE	변수
벼	Ridge	0.92/24,605	5월 최대풍속, 8월 평균 최저초상온도, 8월 일조시간, 9월 평균상대습도, 재배면적
사과	Ridge	0.73/21,922	4월 평균풍속, 5월 평균 최저초상온도, 6월 평균상대습도, 6월 강수량, 8월 지면평균온도, 10월 최고기온, 재배면적
가을배추	Ridge	0.90/25,440	8월 합계 일조시간, 11월 평균최저기온, 11월 최대풍속, 재배면적

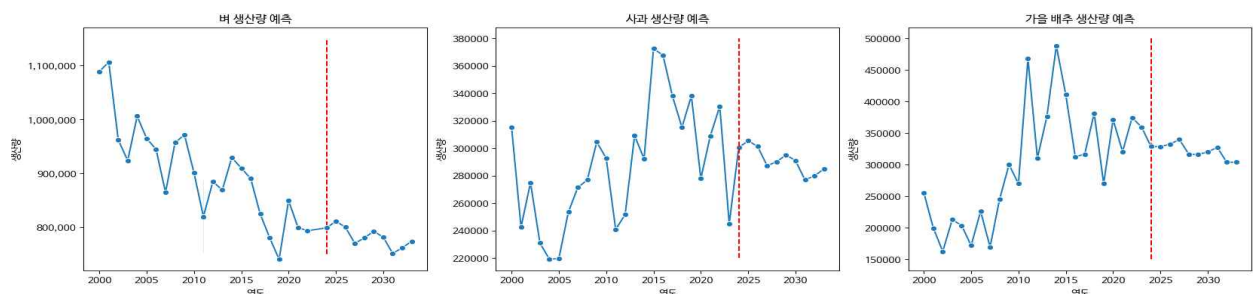


그림 1. 농산물 별 향후 10년(2024년~2033년)간 생산량 예측 그래프

Fig. 1. Graph of Production Forecast for the Next 10 Years (2024-2033) by Agricultural Products

사과는 평균 29만 톤으로, 현재 평균 28만 톤 대비 큰 변동 없이 기후변화의 영향을 받지 않을 것으로 보인다. 생산량이 안정적인 만큼 가격 변동도 크지 않을 가능성이 크다. 배추는 평균 32만 톤으로, 최고 생산량은 33만 톤, 최저는 30만 톤으로 예측되며, 최근 4년 평균인 35만 톤보다 낮은 수준이다. 이는 생산량 감소와 함께 가격이 크게 상승할 가능성이 있으며, 특히 배추 가격은 향후 10년간 50% 이상 상승할 수 있는 위험이 있다.

## V. 요약 및 결론

본 연구는 기후변화로 인한 농산물 생산량 변동과 그로 인한 인플레이션 문제를 해결하기 위해 머신러닝 모델을 활용한 예측 시스템을 제안하였다. 벼, 사과, 가을배추의 향후 10년간 생산량을 예측하기 위해 기후 데이터와 재배 면적 데이터를 통합하여 Random Forest, XGBoost, Ridge 회귀 모델 및 Prophet 시계열 모델을 적용하였다. 그 결과, 벼의 경우 재배 면적 감소와 기후변화에 따른 생산량 감소가 예상되었으며, 사과는 비교적 안정적인 생산량을 유지할 것으로 분석되었다. 그림 1에서 확인할 수 있듯이 가을배추는 2024년 배추 가격 상승 현상과 맞물려 향후 생산량 감소와 가격 상승 가능성이 높은 것으로 예측되었다. 본 연구에서 사용된 머신러닝 모델들은 기후 데이터의 복잡한 비선형 관계를 효과적으로 처리하였으며, 특히 Ridge 회귀 모델이 벼와 배추의 생산량 예측에서 가장 높은 정확도를 보였다. 시계열 데이터 분석에서는 Prophet 모델을 사용하여 계절성과 추세 변화를 동시에 반영한 예측 결과를 도출할 수 있었다. 이러한 예측 모델들은 다른 연구에서  $R^2$ 스코어 0.6~0.7[6][11]를 달성한 것과 비교해보면 기후 변수와 농작물 생산량 간의 상호작용을 더 정밀하게 반영함으로써 높은 예측 성능을 달성하였다. 본 연구는 기후변화가 농업 생산에 미치는 영향을 데이터 기반으로 정량화할 수 있는 머신러닝 모델의 가능성을 제시하였다. 향후 연구에서는 기후 외의 추가적인 경제적 변수들을 결합한 다변량 모델링 기법을 적용하여 예측 성능을 더욱 향상시킬 수 있을 것이다. 본 연구 결과

는 농업 인플레이션 대응을 위한 정책 수립에 기여할 수 있는 기술적 기반을 제공하며, 특히 기후변화에 따른 장기적인 농업 계획 수립에 있어 중요한 참조 자료가 될 것으로 기대된다.

## 참 고 문 헌

- [1] Lee, J.H., "Price Risks in Agricultural Products: Realities, Issues, and Tasks," GS&J Institute, 2020.
- [2] Isidro Juan Mirón, Cristina Linares, Julio Díaz, "The influence of climate change on food production and food safety", Environmental Research, Volume 216, Part 3, 2023
- [3] 이데일리, 금사과 된 이유가 있었네.. 올해 사과 생산량 30% 급감, 2023.12.22., <https://m.edaily.co.kr/News/Read?newsId=02079526635841144&mediaCodeNo=257>
- [4] Se-Won Kim, Younghee Kim, "A Study on the Application of Machine Learning Algorithm to Predict Crop Production", Journal of the Korea Academia-Industrial cooperation Society, Vol. 22, No. 7 pp. 403-408, 2021
- [5] Jin-Kyeong Seo, Da-Jeong Choi, Juryon Paik, "Research on Selecting Influential Climatic Factors and Optimal Timing Exploration for a Rice Production Forecast Model Using Weather Data", Journal of The Korea Society of Computer and Information, Vol. 28 No. 7, pp. 57-65, 2023
- [6] Soojeong Myeong "Impact of Climate Change Related Natural Disasters on Rice Production in South Korea", J. Korean Soc. Hazard Mitig 18(7): 53-60, 2018
- [7] 농업날씨 365, 주산지별 기상분석, <https://weather.rda.go.kr/farmProduce/frcPlpdAvg.do>
- [8] Marquardt, D. W., & Snee, R. D. (1975). "Ridge Regression in Practice." The American Statistician, 29(1), 3-20.
- [9] Mark R. Segal "Machine Learning Benchmarks and

Random Forest Regression", Division of Biostatistics,  
University of California, San Francisco, CA  
94143-0560, 2003

[10] Tianqi Chen and Carlos Guestrin "XGBoost: A Scalable Tree Boosting System", Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2003, 2016

[11] Oscar Barrero1, Sofiane Ouazaa, Camilo Ignacio Jaramillo-Barrios, Mauricio Quevedo, Nesrine Chaali, Sair Jaramillo, Isidro Beltran and Omar Montenegro, "Rice Yield Prediction Using On-Farm Data Sets and Machine Learning", SmartICT 2019, LNEE 684, pp. 422-430, 2020.

# 공중무인체계 상호운용성 시험 평가방법 연구

김다빈\*, 홍정희\*

## Research on the test and evaluation method for interoperability of unmanned aerial systems

Kim Dabin\*, and Hong Jeonghee\*

### 요 약

본 논문은 STANAG 4586 구현 적합성 검증 및 시험 가이드 AEP-84.2 문서와 기존 공중무인체계 운용시험 평가 항목 등을 고려하여 지상통제체계(GCS)와 무인기(UAV) 간 국방 상호운용성시험평가(JIT, Joint Interoperability Test) 방법에 관해 기술한다. 본 연구에서는 GCS와 UAV 간 연동성 측면을 중점적으로 고려하여 JIT 평가방법을 기술하였고, 향후 다양한 공중무인체계 운영개념과 상호운용성 수준을 함께 고려하여 JIT 평가방법론을 확대 연구할 필요가 있다.

### Abstract

This paper provides the interoperability test method between Ground Control System(GCS) and Unmanned Aerial Vehicles(UAVs) which is conducted in real operating environments. We described and interoperability test evaluation methods by referring to the STANAG 4586 Validation/Test Guideline for AEP-84.2 document and existing UAV operability test items. In the future, it is necessary to expand the test and evaluation methodology by considering various UAV operation concepts and the level of interoperability between GCS and UAV.

### Key words

Unmanned Aerial Vehicle System, STANAG 4586, AEP-84.2, Joint Interoperability Test

## 1. 서 론

과거 중대형 군용 공중무인체계는 지상통제체계(GCS, Ground Control System)와 무인기(UAV, Unmanned Aerial Vehicle)간 1:1 연통형(stovepipe) 구조로 개발되었으나 최근에는 무인기 개발 비용, 개발 기간 등의 비효율성으로 인해 미군과 NATO를 비롯한 주요국을 중심으로 GCS와 UAV 간 1:N 연동구조로 개발하는 추세로 변화하고 있다. 하나의

공통(범용) GCS로 여러 타 체계의 UAV를 통제할 수 있게 개발하기 위해서는 GCS와 UAV 간 상호운용성을 보장할 수 있는 연동 표준 프로토콜이 필요하다. NATO에서 연합국 간 무인기 상호운용성 확보를 위해 정립한 STANAG 4586이 대표적인 무인기 연동 표준 프로토콜에 해당한다. 최근 우리 군도 GCS와 UAV 간 1:N 연동구조로 개발하는 방향성에

\*한국국방연구원, dbkim@kida.re.kr(교신저자), jhhong@kida.re.kr.



그림 1. AEP-84.2 적합성 시험  
Fig. 1. AEP-84.2 Conformance Testing

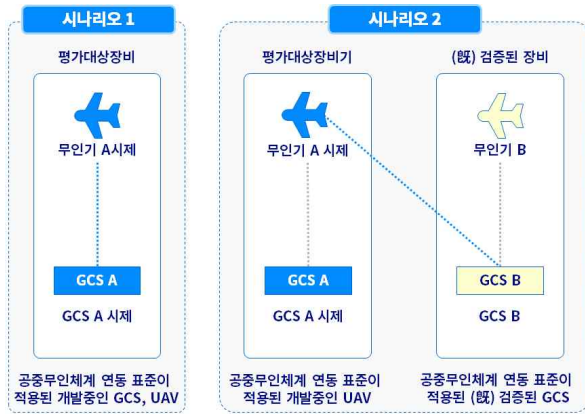


그림 2. 시험평가 시나리오  
Fig. 2. Operational scenarios

공감대를 형성하고, 향후 연구 개발되는 공중무인체계에 대해서는 STANAG 4586과 같은 연동 표준 프로토콜을 기반으로 개발하는 방안을 고려하고 있다. 이를 위해 GCS와 UAV 간 상호운용성 표준을 준수하여 개발되었는지를 검증하는 방안이 필요하지만, 아직 국방 무기체계 연구개발 단계 중 GCS와 UAV 간 상호운용성 평가에 관한 방안은 정립되어 있지 않다. 본고는 STANAG 4586 구현 적합성 검증 및 시험 가이드(AEP-84.2)와[1] 기존 공중무인체계 운용시험평가 항목 등을 고려하여 GCS와 UAV 간 국방 상호운용성시험평가(JIT, Joint Interoperability Test) 방법에 관해 기술한다.

## II. 상호운용성 평가방법

### 1. AEP-84.2

AEP-84.2는 STANAG 4586을 구현하는 시스템, 구성요소가 표준을 준수하여 구현되었는지 검증하는 방법에 관한 가이드라인을 제공한다. 시험방법은 그림 1과 같이 ① 컴포넌트 적합성 시험과 ② 시스템 적합성 시험으로 구분한다. 컴포넌트 적합성 시험을 통해 GCS와 UAV 간 연동에 필요한 각 컴포넌트에 대해 STANAG 4586 프로토콜 구현에 대한

표 1. AEP-84.2 적합성 시험

Table 1. AEP-84.2 Conformance Testing

① 컴포넌트 적합성 시험(Component Conformance Testing)	
시험 목적	시스템 적합성 시험 수행 전에 공중 무인체계 각 구성 컴포넌트에 대한 적합성 시험 수행
시험 단위	GCS를 구성하는 요소들의 기술적 구현 적합성을 각 컴포넌트 단위로 시험
시험 환경	각 컴포넌트 검증을 위한 추가 지원 도구 등을 이용한 시험 환경 구성(Stand-alone 시험 환경)
② 시스템 적합성 시험(System Conformance Testing)	
시험 목적	검증된 컴포넌트로 구성된 GCS ↔ UAV 간 체계 운용성에 대한 적합성 시험 수행
시험 단위	공중 무인체계를 이용한 GCS ↔ UAV (end-to-end) 간 운용시험
시험 환경	실 체계 운용환경 또는 그에 준하는 (인증된) GCS/UAV 모의 장비로 구성된 환경

표준 적합성을 검증하고, 이후 검증된 컴포넌트들로 구성된 전체 시스템 단위의 적합성 시험을 수행하는 방법론을 제공한다. 각 시험의 목적과 시험 환경 등에 관한 내용은 표 1과 같이 요약할 수 있다.

### 2. 상호운용성시험평가(JIT) 방법

국방상호운용성 평가는 무기체계 획득 단계별로 상호운용성 평가 대상을 정의하고 있다. 이 중 AEP-84.2에서 제시하는 적합성 시험에 해당하는 부분은 무기체계 획득 단계의 체계개발 단계 중 DT·OT에서 수행하는 표준적합성시험(SCT, Standard Conformance Test)과 상호운용성시험평가(JIT)에 대응 가능한 것으로 파악하였다. 본 논문은 AEP-84.2 시스템 적합성 시험 가이드라인과 연계하여 국방 공중무인체계 체계개발단계 중 OT 단계에서 실 체계 환경을 기반으로 구현 및 운용능력을 확인하는 상호운용성시험평가(JIT) 방법에 관해 기술한다. 시험 환경은 그림 2와 같이 두 가지 시나리오를 가정하였다. <시나리오 1>은 공중무인체계 연구개발 간 GCS 시제와 UAV 시제를 모두 개발하는 경우로 체계개발 사업 내에서 JIT를 수행하는 경우이다. <시나리오 2>는 신규 개발하는 UAV가 타 체계(무인기 연동 표준을 적용하여 이미 개발/검증된 GCS)와 상호운용성이 보장되는지를 검증하기 위해 타 체계 간 JIT를 수행하는 경우이다. 본 고에서는 기존 유사체계의 국방 상호운용성시험평가(JIT) 사례와 AEP-84.2의 시스템 적합성 시험 부분을 참고하여



공중무인체계 JIT 평가에 필요한 시험평가 항목과 방법을 다음과 같이 기술한다. <시나리오 1>에 대한 평가방법은 다음과 같다.

- ① **평가 대상 장비:** STANAG 4586 또는 군에서 정하는 무인기 연동 표준 프로토콜을 적용하여 체계 개발 중인 GCS 시제와 UAV 시제 각 1대
- ② **시험평가 방안:** 검사, 시연, 시험 중 선택
- ③ **시험평가 장소:** 비행(시험)장 또는 국방 차원에서 실 환경 모사 가능한 시험장소
- ④ **시험방법 (체계 내 시험)**
  - 시험 환경: GCS 시제 ↔ UAV 시제
  - 시험 항목: 무인기 연동 표준 프로토콜 및 본 연구에서 도출한 공중무인체계 JIT 평가항목
  - 결과확인: GCS 시제 ↔ UAV 시제기 간 데이터 링크 개설 후, GCS 시제 내부 운용자 컴퓨터(HCI)를 통해 UAV로 제어 명령(TC, Telecommand) 전송 및 화면에 도시되는 결과(TM, Telemetry) 확인
  - 체계 내 시험의 경우 국방 상호운용성시험평가 5대 항목 중 연동성 및 정보교환, 정보보호 일부 항목은 CT/OT 시험평가 결과 활용 가능.

<시나리오 2>에 대한 평가방법은 다음과 같다.

- ① **평가 대상 장비:** STANAG 4586 또는 군에서 정하는 무인기 연동 표준 프로토콜을 적용하여 체계 개발 중인 UAV 시제 1대
- ② **지원 장비:** 무인기 연동 표준 프로토콜을 적용하여 기(既) 검증된 GCS (또는 실 체계 수준의 평가가 가능한 검증된 GCS 모의 장비)
- ③ **시험평가 방안:** 검사, 시연, 시험 중 선택
- ④ **시험평가 장소:** 비행(시험)장 또는 국방 차원에서 실 환경 모사 가능한 시험장소
- ⑤ **시험방법 (타 체계 간 시험)**
  - 시험 환경: 검증된 GCS ↔ UAV 시제기
  - 시험 항목: 무인기 연동 표준 프로토콜 및 본 연구에서 도출한 공중무인체계 JIT 평가항목
  - 결과확인: GCS ↔ UAV 시제기 간 데이터링크 개설 후, GCS 내부 운용자 컴퓨터(HCI)를 통해 UAV로 제어 명령(TC) 전송 및 화면에 도시되는 결과(TM)확인

7) 검사(Inspection), 시연(Demonstration), 시험(Test)

표 2. 상호운용성시험(JIT) 평가항목(예)

Table 2. JIT Test cases (examples)

평가항목	평가 기준
네트워크 이중화/이원화 계획/구현의 적절성	<ul style="list-style-type: none"> <li>• GCS ↔ UAV 간 링크 다중화가 구현되어야 한다.</li> <li>• 주 링크 단절 시 보조링크로 자동 전환 여부를 확인한다.</li> <li>• 모든 링크 두절 상황에 대비한 비상 로직을 보유해야 한다.</li> </ul>
정보교환 내역 구현의 적절성	<ul style="list-style-type: none"> <li>• 실 체계 비행 간 GCS ↔ UAV 간 STANAG 4586 메시지 송수신이 가능해야 한다.</li> <li>• 실 체계 비행 간 GCS ↔ UAV 간 정보교환내역에 따라 정지영상/동영상 등의 메시지 송수신이 가능해야 한다.</li> </ul>
...	...

상기 언급한 시험방법 중 본 연구에서 도출한 공중무인체계 JIT 평가항목이란 국방상호운용성평가 기준지침서의 5대 평가항목<sup>8)</sup>에 대한 세부 평가항목에 대해 공중무인체계의 특성을 고려하여 GCS와 UAV 간 JIT 평가를 위해 도출한 항목을 말한다. 표 2는 본 연구에서 도출한 시험 평가항목 중 ‘연동성 및 정보교환’과 관련된 일부 항목으로 해당 평가항목에 대한 평가 기준 수립 사례를 제시하였다.

### III. 결 론

본 논문에서는 국방 공중무인체계의 GCS와 UAV 간 상호운용성 시험평가 방법 중 실 체계 환경을 기반으로 수행하는 JIT 평가방법에 관해 기술하였다. 본 연구에서는 GCS와 UAV 간 연동성 측면을 고려하여 JIT 평가방법을 기술하였으나, 향후 다양한 공중무인체계 운영개념을 함께 고려하여 JIT 평가방법론을 확대 연구할 필요가 있다.

### 참 고 문 헌

[1] Validation/Test Guideline for AEP-84, AEP-84.2 Edition A, Version 1, NATO Standardization Office(NSO), 2017.

8) 5대 평가항목: 운용개념 및 체계특성, 연동성 및 정보교환, 표준 및 아키텍처, 사이버보안, 주파수

# 신재생에너지 연구개발 동향 분석

김 윤 정\*

## Analyzing Trends of New & Renewable Energy Research and Development

Yun-Jeong Kim\*

### 요 약

기후 위기 대응과 탄소중립은 글로벌 의제로, 탄소 배출 저감을 위해 신재생에너지 기술 개발은 필수 요소이다. 본 논문에서는 국가R&D 과제 데이터를 기반으로 신재생에너지 연구개발 동향을 분석하였다. 주요 연구 수행주체인 중소기업의 신재생에너지 과제 분석을 통해 글로벌 에너지 전환 정책 및 수소 경제와 관련된 기술의 발전 등을 확인했다.

### Abstract

Climate crisis response and carbon neutrality are global agendas, and the development of new and renewable energy technologies is essential for reducing carbon emissions. In this paper, we analyzed the trends of new and renewable energy research and development based on national R&D project data. Through the analysis of new and renewable energy projects of small and medium-sized enterprises, which are major research performers, we confirmed the development of technologies related to global energy transition policies and the hydrogen economy.

### Key words

Net-Zero, renewable energy, R&D project, research trends, BERTopics

## I. 서 론

글로벌 탄소중립(Net-Zero)을 위해 재생에너지, 그린에너지, 청정에너지의 중요성이 점점 높아지고 있다. 2022년 러시아의 우크라이나 침공, 러시아의 핵심 광물 수출제한 조치 등으로 인한 국제 에너지 시장 위기 상황 속에서 에너지 안보 강화 필요성도 강조되고 있다.

탄소 배출 저감을 위해 신재생에너지 기술 개발은 필수이다. 본 연구는 국가 연구개발(R&D) 과제 데이터를 기반으로 신재생에너지 연구개발 동향을 다각도로 분석하여 다양한 연구수행주체에게 통찰을 제공하고자 한다.

## II. 국내외 동향

\* 한국과학기술정보연구원, miso@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원의 기본사업으로 수행된 연구입니다.(과제번호: J24JR042)

국제에너지기구(IEA)는 2023년 중국, 북미, 멕시코 등 주요 지역에 발생한 장기적 가뭄으로 수력발전 비중이 감소하고 화력발전이 증가한 것이 이산화탄소 배출량 최고치 기록의 원인이라 분석했다. 그러나 태양광, 풍력 등 주요 청정에너지 기술 보급으로 이산화탄소 배출량은 둔화하고 있다[1]. IEA의 보고서에 따르면 2023년 재생에너지 증가분의 약 75%를 태양광이 차지했다[2]. 태양광 발전은 현재 가장 저렴하고 안전한 전기 공급원이다.

KOSIS의 온실가스 배출량 통계에 의하면 배출량 상위 10개 국가(2021년 기준)는 미국, 러시아, 일본, 독일, 캐나다, 대한민국, 호주, 튀르키예, 영국, 프랑스, 이탈리아이다. 온실가스 배출량 1위 국가인 미국은 2022년 8월부터 시행된 인플레이션감축법(IRA)을 통해 청정에너지, 기후변화 대응 등에 대한 투자를 촉진하고 있으며 이는 재생에너지 기술 성장을 가속화할 것으로 전망된다.

유럽연합은 풍력발전 등 재생에너지 생산에 힘입어 2023년 이산화탄소 배출량 감소의 절반을 청정에너지 성장이 기여했다. 독일은 2023년 총 재생에너지 생산량이 약 260TWh로, 2022년에 비해 7% 증가했다. 영국의 재생에너지 전력 생산량은 2008년 23TWh에서 2023년 135TWh로 6배 증가했고, 2023년은 전력 생산의 47.3%가 재생에너지였다. 영국은 2030년까지 10GW의 그린수소 및 블루수소 발전 용량을 보유한다는 목표를 갖고 있다[3, 4].

전 세계적인 탈탄소화 흐름 속에서 우리나라 정부는 2021년 9월 탄소중립기본법을 제정했고, 전력 공급의무자에게 신재생에너지 의무공급량을 규정하는 등 신재생에너지 개발·이용·보급 촉진과 온실가스 감축을 추진 중이다. 제5차 신·재생에너지 기술개발 및 이용·보급 기본계획('20~'34), 탄소중립 중점기술 임무지향적 탄소중립 R&D 추진방안 탄소중립 녹색성장 기술혁신전략, 제1차 기후변화대응 기술개발 기본계획('23~'32) 등 정책도 연이어 발표되었다. 2024년 9월에는 '한·일 에너지대화'를 개최해 탄소중립 대응, 무탄소에너지의 활용 촉진을 위한 양국 협력을 약속했다.

### III. 신·재생에너지 국가R&D 현황

에너지 전환 및 탄소중립을 위해 신재생에너지 기술 개발은 필수이다. '신재생에너지' 연구분야는 국가과학기술표준분류체계의 '에너지/자원' 대분류의 중분류 중 하나이다. 신재생에너지의 세부영역(舊 소분류)은 신에너지와 재생에너지로 나뉜다.

국가R&D 과제 데이터를 기반으로 신재생에너지 R&D 동향을 분석하기 위해 국가R&D 사업 관련 정보를 공개·개방하는 국가과학기술지식정보서비스(NTIS)의 디렉토리검색을 이용하여 과제와 논문 데이터를 수집했다. 국가R&D 과제정보의 연구분야는 2015년까지 중분류, 2016년부터 소분류 입력이 필수이다. NTIS 과제의 조사·분석 항목 중 일부 오류 값은 정제 과정을 거쳤다.

기준년도	과제 고유번호	연구개발기관	사업등록번호	연구수행주체	연구수행주체(연장)
2021	9991008083	(사)수소융합알라이언스	278-82-00098	중소기업	기타
2018	1615010320	(주) 지지케이	215-81-53020	기타	중소기업
2023	1711182537	경북대학교	504-83-00558	국립연구소	대학
2021	1545024559	고등기술연구원	104-82-04565	중소기업	기타
2023	1711197006	재단법인 한국조선해양기자재연구원	602-82-06034	중소기업	기타
2023	1711198581	중앙대학교	108-82-01221	정부부처	대학
2023	1711198892	한국과학기술연구원	209-82-03522	기타	출연연구소
2015	1711028665	한국기계연구원	314-82-02693	대학	출연연구소
2017	1345266951	한국기초과학지원연구원	314-82-06097	대학	출연연구소
2018	1425127955	한국생산기술연구원	119-82-01008	중소기업	출연연구소
2023	1711186968	한국생산기술연구원	119-82-01008	국립연구소	출연연구소
2023	1711188642	한국세라믹기술원	119-82-06696	국립연구소	출연연구소
2019	1415166376	한국에너지기술연구원	314-82-02242	기타	출연연구소
2023	1711194952	한국화학연구원	314-82-00884	기타	출연연구소

그림 1. 과제 데이터 정제 예시  
Fig. 1. Example of project data refinement

신재생에너지 연구분야 2013~2024년 국가R&D 과제와 논문 현황은 그림 2와 같다. 조사·분석 확정 기준 최근 5년('18~'22) 동안 정부연구비와 과제가 연평균 9.7%, 4.5% 증가할 때 신재생에너지 연구분야는 34.3%, 14.3%로 증가해 신재생에너지 연구의 중요성을 보여준다. 정부연구비는 중소기업, 출연연구소, 대학 연구수행주체 순으로 집행되었다.

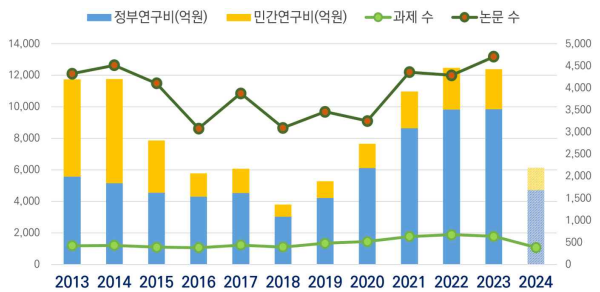


그림 2. 신재생에너지 국가R&D 현황(2013~2024)  
Fig. 2. Status of National R&D on New & Renewable Energy (2013~2024)

주 연구분야(연구분야1) 기준, 2023년에 수행 건수가 많은 세부영역은 ‘수소’, ‘태양광’, ‘연료전지’ 순이다. 국제공동연구는 국제협약(61%), 외국연구자 유치(32%), 정보교환(7%) 순이고, 미국, 독일과의 공동연구가 절반을 차지한다. 해외 기관이 수행한 위탁과제는 42건(‘13~’23)으로, 참여국가는 미국이 1위(83%)이며 위탁연구비의 절반(52.3%)이 3개 기관에서 집행되었다. 기관별로 태양광, 태양열 및 지열, 수소가 주 연구분야인 주관연구과제에 참여했다. 녹색기술분류별 연구개발 집행(‘13~’23)은 ‘풍력기술’, ‘차세대 고효율연료전지시스템기술’, ‘실리콘계 태양전지 기타 기술’ 순으로 나타났다.

신재생에너지와 다른 연구분야 간 융합 연구(동시 발생 기준, ‘16~’24)를 살펴보면, 나노화학공정기술(EC02), 에너지/환경기계시스템(EA07), 세라믹재료(EB02), 고분자재료(EB03), 금속재료(EB01), 화학공정(EC01), 폐기물 관리/자원순환(EH07), 전기화학(NC08)이 상위 연구분야이다.

정부연구비 기준으로 가장 많은 연구를 수행 중인 중소기업의 연구(‘16~’23) 주제를 BERTopic[7]으로 살펴보면 ‘바이오매스 및 폐기물 에너지화’, ‘풍력 발전기 부품 및 가공 기술’, ‘합성가스 생산 및 공정’이 주요 토픽이고, ‘수소 에너지와 연료전지’는 2020년 이후 급격한 상승세를 보인다.

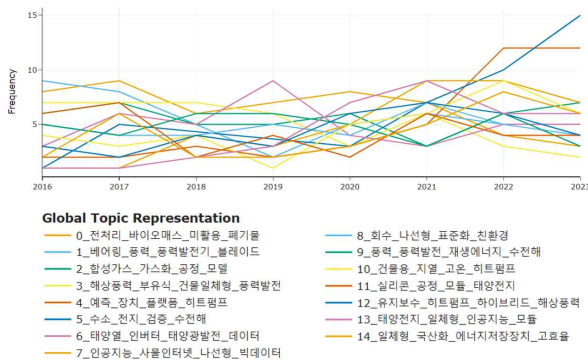


그림 3. 신재생에너지 연구분야 토픽모델링 결과(2016~2023, 중소기업)

Fig. 3. Topic Modeling Results of New & Renewable Energy Research (2016-2023, SME)

중소기업이 수행한 신재생에너지 연구과제의 국문 키워드 빈도수 기준으로 2023년과 2024년(2024년 10월 기준)을 비교하면, ‘태양광’, ‘수소’, ‘연료전

지’는 지속적으로 나타나고 있다. 특히 급부상한 ‘수소’는 수소 경제와 관련된 기술 개발이 가속화되고 있음을 의미한다. ‘연료전지’는 에너지 저장 관련 중요 기술로 자리 잡고 있음을 알 수 있다. 2023년 대비 2024년에 상승한 키워드는 ‘인공지능’, ‘암모니아’, ‘디지털트윈’, ‘액체수소’, ‘폐플라스틱’, ‘폐기물’, ‘바이오오일’, ‘고압수소’, ‘풍력터빈’, ‘로봇’, ‘자가발전’, ‘전통시장’, ‘탄소담지체’, ‘음식물’, ‘그리드포밍’, ‘청록수소’ 등이다.

#### IV. 결 론

국가R&D 과제 데이터 기반 연구개발 동향 분석을 통해 ‘태양광’이 중요한 연구 주제임을 확인했다. 에너지 예측 등에 활용하기 위한 ‘인공지능’이 급부상했고, ‘수소’, ‘연료전지’와 같은 연구 주제의 부상은 글로벌 에너지 전환 정책 및 수소 경제와 관련된 기술 발전이 가속화되고 있음을 의미한다. 신재생에너지와 다른 연구분야 간 융합연구 분석으로 에너지 변환 및 저장 기술, 배터리/연료전지/태양전지의 성능 향상, 폐기물 처리와 자원 재활용 연구의 중요성을 확인했다.

#### 참 고 문 헌

- [1] International Energy Agency, “CO<sub>2</sub> Emissions in 2023”, 2024.
- [2] International Energy Agency, “Renewables 2023 : Analysis and forecast to 2028”, 2024.
- [3] <https://www.carbonbrief.org>
- [4] <https://www.aitimes.com>
- [5] Renewable Energy Country Attractiveness Index, 63 edition, 2024
- [6] van Eck, Nees Jan and Waltman, Ludo, “VOSviewer: A Computer Program for Bibliometric Mapping”, *Scientometrics* 84, 523 - 538, 2010.
- [7] Grootendorst, Maarten., "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," arXiv preprint arXiv:2203.05794, 2022.

# 영화 추천 시스템(BERT-Mosen5)

정휘찬\*, 전준표\*\*, 이승엽\*\*\*

## Movie Recommender System(BERT-Mosen5)

Hwi chan Jeung\*, Jun pyo Jeon\*\*, and Seung yeop Lee\*\*\*

### 요약

기존의 영화 추천 모델들은 협업 필터링(Collaborative Filtering, CF), 콘텐츠 기반 필터링(Content-Based Filtering, CBF) 등의 방법을 통해 발전해 왔으나, 이러한 방식은 주로 데이터의 희소성(sparse) 문제와 콜드 스타트(cold start) 문제에서 한계를 드러냈다. 인공지능의 도입 이후 BERT 모델을 활용한 추천시스템들이 등장하며 성능이 개선되었으나, 대부분 임베딩만을 사용하거나 영화 리뷰의 특성 벡터만을 추출하여 추천에 적용하는 경우가 많아 여전히 데이터의 희소성 문제를 충분히 해결하지 못하였다. 본 논문에서는 BERT의 임베딩과 MLM(Masked Language Model) 방식을 함께 활용하여 영화 평점 예측의 정확도를 높인 새로운 모델을 제안한다.

### Abstract

Traditional movie recommendation models, such as Collaborative Filtering (CF), Content-Based Filtering (CBF) have developed extensively over time but often encounter limitations, particularly with data sparsity and the cold-start problem. With the advent of artificial intelligence, recommendation systems employing BERT have demonstrated performance improvements; however, these models frequently rely solely on embeddings or extract feature vectors from movie reviews for recommendation purposes, which falls short in fully addressing data sparsity issues. In this study, we propose an innovative model that leverages both BERT embeddings and the Masked Language Model (MLM) technique, aiming to enhance the accuracy of movie rating predictions.

### Key words

BERT, Sparsity, Sentiment score, Likert scale, Movie rating

## 1. Introduction

The advent of online streaming platforms and

diverse digital media services has liberated users from fixed times and locations, allowing them to access a vast array of content with ease[1]. However, the

\* Master's Course, Department of Media Communication, Pukyong National University, Busan 48513, Korea. wjdgnlcks@naver.com

\*\* Master's Course, Department of Media Communication, Pukyong National University, Busan 48513, Korea, jjp7380@naver.com

\*\*\*Assistant Professor, Department of Media and Communication, Pukyong National University, Busan 48513, Korea. yeopcp@pknu.ac.kr(Corresponding Author)

proliferation of media platforms has also led to a surge in content production, making it challenging for users to consume content efficiently[2]. In this context of information overload, recommender systems, which utilize users' preferences to provide personalized content recommendations, have gained increasing importance[3].

Traditional recommendation systems, such as Collaborative Filtering (CF) and Content-Based Filtering (CBF), rely on techniques like cosine similarity and Pearson correlation, which are limited in capturing the complex relationships between users and movies[4]. These methods also struggle to perform effectively in cases of sparse data. Models utilizing BERT have similarly relied solely on embeddings for rating prediction, failing to fully address existing limitations. Thus, this study proposes a novel model that leverages both BERT embeddings and the Masked Language Model (MLM) approach to enhance the accuracy of movie rating predictions.

## II. BERT-Mosen5

### II-1. CF

Collaborative filtering performs recommendations based on the similarity between users. The similarity between two users,  $u$  and  $v$ , is typically calculated using Pearson correlation after normalizing by the average rating. The similarity  $\text{sim}(u,v)$  between users  $u$  and  $v$  is defined by the following equation.

$$\text{sim}(u,v) = \frac{\sum_{i \in I_{uv}} (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in I_{uv}} (r_{ui} - \bar{r}_u)^2} \cdot \sqrt{\sum_{i \in I_{uv}} (r_{vi} - \bar{r}_v)^2}}$$

Here,  $r_{ui}$  represents the rating given by user  $u$  to movie  $i$ ,  $\bar{r}_u$  denotes the average rating of user  $u$ , and  $I_{uv}$  is the set of movies that both  $u$  and  $v$  have

rated. This similarity measure aids the CF model in selecting items to recommend[5].

### II-2. Sentiment Analysis

Sentiment Analysis is the process of extracting emotional information from text and classifying it as positive, neutral, or negative[6]. Using NLP models like BERT, emotional vectors from text are extracted and analyzed mathematically. A common approach involves averaging the embedding vectors of each word to derive an overall sentiment vector for the sentence, which is then passed to a classifier for categorization.

$$\text{Sentiment score} = \sigma\left(W \cdot \frac{1}{N} \sum_{i=1}^N h_i\right)$$

Here,  $h_i$  represents the embedding vector of word  $i$ ,  $W$  is the weight matrix, and  $\sigma$  denotes the sigmoid or softmax function. In sentiment analysis, the predicted results are used to determine the emotional tendency of the sentence, and the classifier's predictions quantify the sentiment of the text.

### II-3. BERT

BERT (Bidirectional Encoder Representations from Transformers) is a pre-trained language model proposed by Google that demonstrates excellent performance in comprehending the contextual meaning of text bidirectionally, thereby enhancing the accuracy of movie rating predictions. BERT uses the Transformer model to learn the interrelationships between words within a sentence, and it is trained through two main tasks: the Masked Language Model (MLM) and Next Sentence Prediction (NSP)[7].

$$P(w_i|w_{-i}) = \text{softmax}(W \cdot h_i)$$

## II-4. BERT-Mosen5

Studies that have utilized BERT for movie recommendations have only used userID and movieID as embeddings[8]. However, in this study, we additionally incorporate the average rating of the user and the average rating of the movie as supplementary information that aids recommendations in CF. This serves as a representative piece of information for a given user, especially in cases where a user profile is not available.

$$\mu_u = \frac{1}{|I_u|} \sum_{i \in I_u} r_{u,i}$$

$$\mu_i = \frac{1}{|U_i|} \sum_{u \in U_i} r_{u,i}$$

Here,  $r_{u,i}$  represents the rating given by user  $u$  to movie  $i$ .  $I_u$  is the set of movies rated by user  $u$ , and  $U_i$  is the set of users who have rated movie  $i$ .

The sentiment scores are derived from review data using BERT. Traditional sentiment analysis typically employs a binary scoring system, assigning a score of 1 for positive reviews and 0 for negative reviews [9]. However, in this study, to utilize user reviews more specifically for recommendations, we adopt a 5-point scale with five labels. If a review is very positive, it is assigned a score of 1; positive reviews receive a score of 0.75; neutral reviews are given a score of 0.5; negative reviews receive a score of 0.25; and very negative reviews are assigned a score of 0. The sentiment score scale of 5 is used as input for BERT.

$$s_{u,i} \in \{0, 0.25, 0.5, 0.75, 1\}$$

We propose a BERT-Movie Recommender System utilizing the 5-point sentiment score scale, which computes the similarity among user ID, average user rating, movie ID, average movie rating, sentiment score, and

actual rating through BERT's Masked Language Model (MLM) training methodology to predict ratings.

$$\text{BERT Input} = \text{UserID} \oplus \mu_u \oplus \text{MovieID} \oplus \mu_i \oplus s_{u,i} \oplus r$$

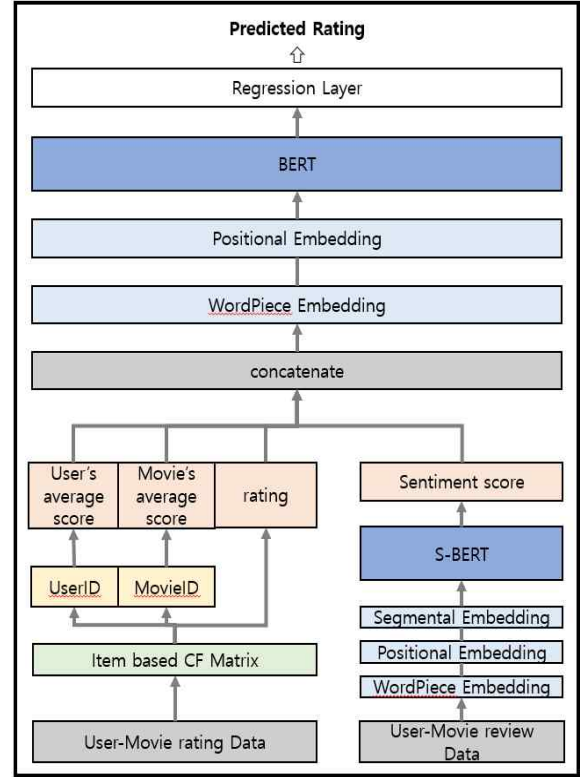


Fig. 1. BERT-Mosen5

## II-5. Experiment

In this study, experiments were conducted using MovieLens and Rotten Tomatoes datasets. The selection criteria included 610 users who had watched at least 30 movies and 618 movies that had been viewed more than 20 times. Among the user-movie matrix, only 8.4% of the ratings were actual ratings. Movie reviews were obtained from the ratings provided by film critics on Rotten Tomatoes.

Furthermore, this study compared the performance of several models for predicting movie ratings using BERT. The evaluation index employed were RMSE, M

SE, and MAE, and the performance comparisons were conducted five times through random sampling.

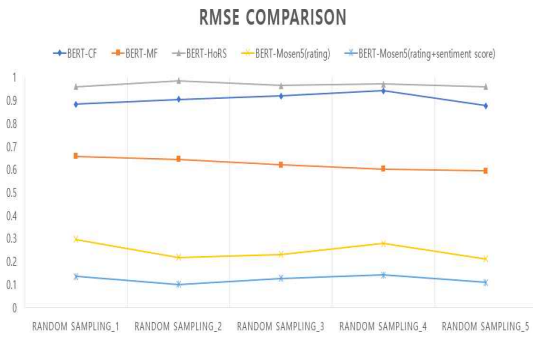


Fig. 2. rmse Performance Comparison

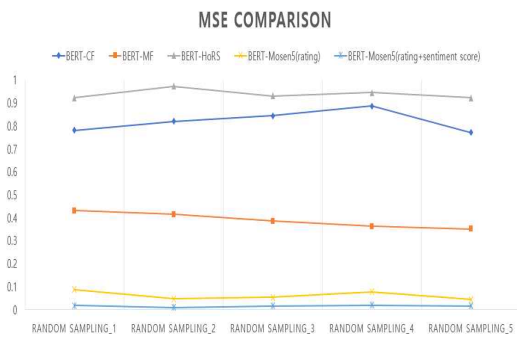


Fig. 3. mse Performance Comparison

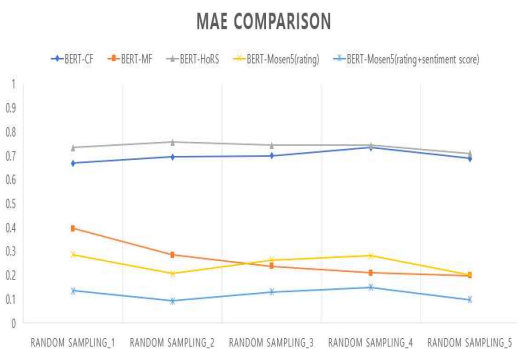


Fig. 4. mae Performance Comparison

### III. Conclusion

This study proposed the BERT-mosen5 model to enhance movie recommendation performance. This model utilized user ID and average rating, each movie's ID and average rating, and the sentiment scores derived from reviews on a 5-point scale,

employing BERT's Masked Language Model (MLM) for rating prediction. This approach demonstrated a significant improvement in movie recommendation performance, outperforming existing models, including those that applied BERT embeddings to Collaborative Filtering (CF) [10], methods that used MF latent vector initializations [11], and approaches that utilized MLP as input [12].

This study suggests that the integration of sentiment scores and rating data effectively enhances the accuracy and quality of personalized recommendations in recommendation systems. In particular, the use of 5-point sentiment scores opens up possibilities for more finely reflecting user preferences in reviews. This approach could also be applied to various media recommendation systems.

Future research suggestions include introducing methods to analyze specific types of emotions as well as the intensity of sentiment through a multidimensional interpretation of sentiment scores. Additionally, it would be meaningful to apply this model to various content recommendation scenarios to validate its generalizability.

### References

- [1] Menon, D., "Purchase and continuation intentions of over-the-top(OTT) video streaming platform subscriptions: A uses and gratification theory perspective", Telematics and Informatics Reports, 5, 100006, 2022
- [2] Villavicencio, C., Schiaffino, S., Diaz-Pace, J. A., & Monteserin, A., "Group recommender systems: A multi-agent solution", Knowledge-Based Systems, 164, 436-458, 2019.
- [3] G. Adomavicius, A. Tuzhilin., "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", IEEE Transactions on Knowledge and Data Engineering, 17, 734-749, 2005



- [4] A. Gokhale and M. Claypool, "Thresholds for More Accurate Collaborative Filtering", Proc. of the IASTED International Conference on Artificial Intelligence and Soft Computing, 1999.
- [5] THORAT, Poonam B.; GOUDAR, Rajeshwari M.; BARVE, Sunita. "Survey on collaborative filtering, content-based filtering and hybrid recommendation system", International Journal of Computer Applications, 110.4: 31-36, 2015.
- [6] B. Pang and L. Lee, "Opinion Mining and Sentiment analysis", Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, pp. 1-135, Jan 2008.
- [7] Kenton, Jacob Devlin Ming-Wei Chang, and Lee Kristina Toutanova. "Bert: Pre-training of deep bidirectional transformers for language understanding", Proceedings of naacL-HLT. Vol. 1. 2019.
- [8] Sun, Fei, et al. "BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer", Proceedings of the 28th ACM international conference on information and knowledge management. 2019.
- [9] S. Elbagir, and J. Yang, "Twitter Sentiment Analysis using Natural Language Toolkit and VADER Sentiment", Proceedings of the International MultiConference of Engineering and Computer Scientists, Hong Kong, pp. 216-225, 2019
- [10] Hoang, Bui Nguyen Minh, et al. "Using Bert Embedding to improve memory-based collaborative filtering recommender systems", 2021 RIVF International Conference on Computing and Communication Technologies (RIVF). IEEE, 2021.
- [11] Woo Seung Kang, Suwon Lee, Sang Min Choi, "A Matrix Factorization-based Recommendation Approach with SBERT Embedding", Journal of KIIT. Vol. 21, No. 11, pp. 203-211, Nov. 30, 2023
- [12] Seonu Park, Xinzhe Li, Jihyeon Kim, Jaekyung Kim, "Development of a BERT-based Hotel Recommender System Reflecting User's Specific Features", J Intell Inform Syst, vol. 30, No. 1, pp. 139-158, March, 2024.

## 온라인 스마트팜 교육과정 운영

이현진\*, 윤재현\*\*, 박찬수\*\*\*

### Operation of Online Smart Farm Education Curriculum

Hyunjin Lee\*, Jaehyun Yun\*\*, and Chan-Soo Park\*\*\*

최근 농업 분야에서는 인공지능(AI)과 정보통신기술(ICT)을 활용해 농업 생산성과 효율성을 극대화하는 스마트팜 기술이 주목받고 있어 이를 반영한 AI스마트팜학과를 신규 개설하여 운영하였다. 본 연구를 통해 학과의 개설 배경, 교육 과정 설계, 그리고 1년간의 운영 결과를 바탕으로 향후 발전 방안을 제안하고자 한다. 특히, 관련 산업체와의 연계를 통해 산업 요구에 맞춘 교육과정을 설계하고 현장실습을 도입하여 학생들이 실무 능력을 갖출 수 있도록 하였다. 또한, 학과 운영에 필요한 시설, 인력, 예산, 그리고 산학 협력 방안을 논의하였다. 본 연구는 AI 스마트팜 학과가 차세대 농업 인력 양성에 기여하며, 지속 가능한 농업 발전과 스마트 농업 혁신에 중요한 역할을 할 수 있음을 시사한다.

#### Key words

smartfarm, urban farm, online education, artificial intelligence, big data

최근 농업 분야에서 AI와 정보통신기술(ICT)을 활용한 스마트팜이 농업 혁신의 핵심 기술로 부상하고 있다. 이에 따라 2024년에 개설된 AI 스마트팜 학과는 이러한 기술적 요구에 부응하여 전문 인력을 양성하기 위해 1년간 시범 운영되었다. 교육과정은 AI와 ICT를 적용한 스마트팜의 기본 이론과 실제 농업 환경에서의 실무 교육을 중심으로 구성되었다. 특히 산업체와의 협력을 통해 산업 현장 요구에 맞춘 실습을 제공하며, 학생들이 현장에서 실제 스마트팜 기술을 체험할 수 있는 기회를 마련했다. 이를 위해 주요 과목은 스마트팜의 데이터 분석, 자동화 모델링, IoT 및 센서 기술 응용, 실습으로 구성되었으며, 이론과 실습을 연결하는 블렌디드 교육 방식이 도입되었다. 또한, 유관 지자체와 협력하여 지역 특성에 맞는 스마트팜 기술을 배우고 현장 실습을 할 수 있는 기회를 제공하였다. 1년간의 운영 결과, 학생들의 기술적 역량과 실무 적응력이 향상되었으며, 특히 현장실습과 블렌디드 교육 방식에 대해 높은 만족도가 나타났다. 이를 바탕으로 산업체와의 연계를 강화하고 블렌디드 교육의 비중을 확대하여, 교육 실효성을 더욱 높일 필요가 있음을 확인하였다.

#### 참고문헌

- [1] 임병욱, 황성일, 엄지범, “스마트팜 청년 창업농의 교육 요구 분석 - 교육과정 우선순위 도출을 중심으로”, 지역사회연구, 제29권, 1호, pp. 137-161, 2021.
- [2] 정중수, 정광욱, “스마트 팜 교육 시스템 설계 및 구현”, 한국컴퓨터정보학회 학술발표논문집, pp.145-148, 2022.

---

\* 서울사이버대학교 AI스마트팜학과 교수, Email: smartfarm@iscu.ac.kr

\*\* 서울사이버대학교 기계제어공학과 교수, Email: yoonjh@iscu.ac.kr

\*\*\* 서울사이버대학교 전기전자공학과 교수, Email: cspark@iscu.ac.kr

# 전력요금 관리를 위한 업무규칙 편집기 설계에 관한 연구

김예미\*, 송승미\*\*, 조윤지\*\*, 김효영\*\*, 고상원\*\*

## A Study on the Design of a Business Rule Editor for Electricity Billing Management

Ye-mi Kim\*, Seung-mi Song\*\*, Yoon-ji Cho\*\*, Hyo-young Kim\*\*, and Sang-won Koh\*\*

### 요 약

분산 에너지 활성화 특별법이 시행됨에 따라 지역별 차등 요금제 등 다양한 전력요금제도를 수용하는 관리 시스템이 필요하다. 본 논문에서는 비즈니스 규칙 엔진(BRE) 기반의 전력요금제도 관리 시스템을 구성하기 위한 업무규칙 편집기 설계를 제안한다. 전력요금관리에 특화된 업무규칙 유형, 업무규칙흐름 편집, 권한 관리 기능을 지원하여 전력판매 사업자가 정의한 요금규정 및 절차를 직관적이고 유연하게 관리한다.

### Abstract

With the implementation of the “Special Act on Activation of Distributed Energy”, it is necessary to design a system to manage various electricity pricing plans, including differential pricing plans by region. This paper presents a Business Rule Editor design for constructing a business rule engine(BRE) based electricity tariff management system. It supports business rule types, flow editing and user permissions specialized for electricity tariff management to intuitively and flexibly manage tariff rules and procedures defined by electricity sales operators.

### Key words

electricity tariff management, business rule engine(BRE), business rule

## 1. 서 론

전력요금은 정부에서 규제하는 공공요금으로, 에너지 정책 변화에 따라 요금체계가 개정되었으며 전국적으로 동일한 요금제도가 적용되었다. “분산 에너지 활성화 특별법”(2024.06 시행)이 시행됨에 따라 분산에너지 특화지역에 규제 특례가 적용되고, 지역별 차등 요금제가 도입될 예정이다[1]. 전력관

매 주체의 확대 및 요금제도 다양화로 전력요금에 대한 관리 리스크가 증가할 것으로 예상되며 요금제 분석검증부터 설계, 계산처리에 이르는 전반적 관리 기술에 대한 요구가 증가하고 있다. 복잡하고 변동성 높은 전력요금체계를 효율적으로 운영하기 위한 전력요금 관리 시스템의 설계가 필요하다.

업무규칙엔진(Business Rule Engine)은 시스템으로부터 비즈니스 로직을 분리하여 관리하는 소프트웨

\* 한전KDN, yemi\_0705@kdn.com, \*\*한전KDN, songsm\_8318@kdn.com, \*\*한전KDN, jamie\_yoonji313@kdn.com, \*\* 한전KDN, vitamin.2r@kdn.com, \*\*한전KDN, swkoh\_1995@kdn.com

어로, 운영자와 개발자의 업무를 분리하여 업무처리 규정을 직관적으로 관리한다. 비즈니스 로직이 복잡하거나 빈번하게 변경되는 경우 시스템 재배포 작업을 단축할 수 있다. 업무규칙 관리시스템에 입력된 업무규칙(Rules)과 모수 데이터(Facts)를 기반으로 업무규칙엔진이 업무로직을 실행 처리한다[2]. 특히 장문으로 작성된 전력요금 약관 및 세칙을 업무규칙으로 변환하는 편집기를 설계함으로써 전력판매 사업자가 정의한 요금규정을 명시적으로 관리하고 전력요금 다변화에 대비할 수 있다.

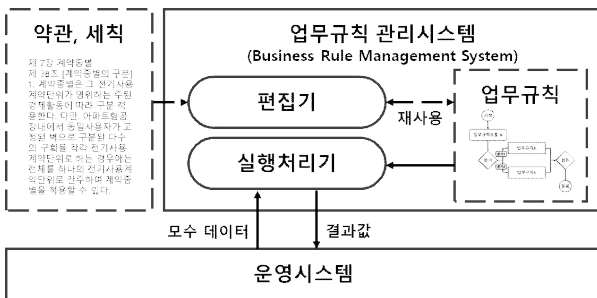


그림 1. 업무규칙 관리시스템 개요  
Fig. 1. Overview of Business Rule Management System

## II. 전력요금관리용 업무규칙 편집기

본 논문에서 제안하는 전력요금관리를 위한 업무규칙 편집기는 전력요금제도 규정을 업무규칙으로 편집하는 도구다. 전력판매사업자가 정의한 요금규정을 웹 UI로 제공되는 용어, 약관, 업무규칙 편집기를 통해 시스템에 등록한다. 용어는 Class 형태로 정의한 후 시스템 내부에서 요금규정에서 사용되는 자연어 형태의 용어와 매핑한다. 약관은 트리 형태로 관리되며 특정 구문을 검색하거나, 관련된 업무규칙을 조회할 수 있다. 사전 정의된 자연어 용어를 활용하여 업무규칙을 편집함으로써 담당자가 정책을 직접 반영할 수 있으며, 용어와 업무규칙을 재활용할 수 있어 업무 효율을 높인다.

업무규칙의 유형은 스크립트, 테이블, 스크립트 테이블, 단가표 형태를 지원한다. 각 유형은 정의 항목 및 방법이 상이하다. 먼저 스크립트 유형은 조건 및 행위를 정의한다. 조건을 충족할 경우 처리 행위를 명시하여 특정 변수의 데이터를 삽입, 변경, 삭제할 수 있다. 테이블 유형은 조건별 결과값을 명

시한다. 조건 인자 및 행위 인자를 선택하고 각 조건의 값을 여러 경우로 구분하여 일괄 처리하는 경우 유용하다. 스크립트 유형과 마찬가지로 행위 변수의 데이터를 삽입, 변경, 삭제할 수 있다. 스크립트 테이블 유형은 조건별 행위를 정의한다. 테이블 유형과 스크립트 유형이 결합된 형태로, 조건의 값을 여러 경우로 구분하여 특정 변수의 데이터를 처리한다. 단가표 유형은 전력요금 단가표 템플릿을 활용하여 요금제를 관리한다.

표 1. 업무규칙 유형

Table. 1. Type of Business Rule

업무규칙 유형	설명 및 예시															
스크립트	조건 및 행위 정의 조건 $대가족가구수 > 0 \text{ AND } 평균사용량 > 300$ 행위 $대가족할인대상여부 = "Y"$															
	조건별 결과값 명시 <table border="1"> <thead> <tr> <th>#</th> <th>대가족가구수 &gt;</th> <th>평균사용량 &gt;</th> <th>대가족할인대상여부</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>&gt; 0</td> <td>&gt; 300</td> <td>Y</td> </tr> </tbody> </table>	#	대가족가구수 >	평균사용량 >	대가족할인대상여부	1	> 0	> 300	Y							
#	대가족가구수 >	평균사용량 >	대가족할인대상여부													
1	> 0	> 300	Y													
스크립트 테이블	조건별 행위 정의 <table border="1"> <thead> <tr> <th>#</th> <th>대가족가구수 &gt;</th> <th>평균사용량 &gt;</th> <th>행위</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>&gt; 0</td> <td>&gt; 300</td> <td>대가족할인대상여부 = "Y"</td> </tr> </tbody> </table>	#	대가족가구수 >	평균사용량 >	행위	1	> 0	> 300	대가족할인대상여부 = "Y"							
#	대가족가구수 >	평균사용량 >	행위													
1	> 0	> 300	대가족할인대상여부 = "Y"													
단가표	전력요금 단가표 템플릿 활용 <table border="1"> <thead> <tr> <th rowspan="2">구분</th> <th rowspan="2">기본요금</th> <th rowspan="2">시간대</th> <th colspan="3">전력량 요금(원/kWh)</th> </tr> <tr> <th>여름철</th> <th>봄/가을철</th> <th>겨울철</th> </tr> </thead> <tbody> <tr> <td>고압 A 선택1</td> <td>7170</td> <td>경부하</td> <td>62.70</td> <td>62.70</td> <td>71.40</td> </tr> </tbody> </table>	구분	기본요금	시간대	전력량 요금(원/kWh)			여름철	봄/가을철	겨울철	고압 A 선택1	7170	경부하	62.70	62.70	71.40
구분	기본요금				시간대	전력량 요금(원/kWh)										
		여름철	봄/가을철	겨울철												
고압 A 선택1	7170	경부하	62.70	62.70	71.40											

복수의 업무규칙에 대한 조건별 처리 흐름을 정의하기 위해 업무규칙흐름 편집기를 활용한다. 시작, 종료, 분기, 합류 노드 및 연결선을 활용하여 전력요금 처리 흐름을 입력한다. 상위 업무규칙흐름은 하위 업무규칙 혹은 업무규칙흐름을 포함할 수 있어 복잡한 요금 처리 절차에 대한 분석 및 추적이 용이하다. 업무규칙 편집기로 작성한 업무규칙은 저장 시 문법 오류를 검증하며, 모수 데이터로 테스트 후 배포 승인을 획득해 파일 형태로 반영한다.

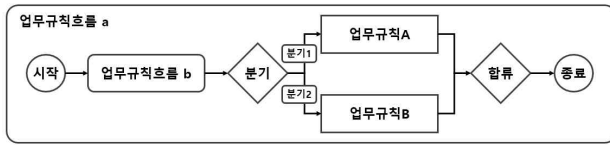


그림 2. 업무규칙흐름 설계 화면  
Fig. 2. Preview of Business Rule Flow

업무규칙은 사용자 그룹별 구분되는 업무분야와 업무 성격에 따른 단위업무의 2계층 구조로 관리된다. 업무규칙에 대한 조회 및 편집, 배포 권한은 시스템 관리자, 그룹 관리자, IT 담당자, 업무 관리자, 업무 담당자로 구분한다. 시스템 관리자는 모든 사용자 그룹의 모든 업무분야에 접근 가능하며, 사용자 그룹과 업무분야를 매핑한다. 그룹 관리자는 사용자 그룹 내 모든 단위업무에 접근 가능하며, 그룹에 포함되는 사용자 목록 및 권한을 관리한다. IT 담당자는 사용자 그룹별 용어등록 및 업무규칙 편집과 관련한 IT 작업을 담당한다. 업무 담당자가 업무규칙 편집기 및 업무규칙흐름 편집기를 활용해 업무규칙을 정의하면, 업무 관리자가 확인 후 배포 승인한다.

표 2. 사용자 권한 목록

Table. 2. List of User Permission

구분	권한			
	전체	그룹		
		사용자	등록	배포
시스템 관리자	O	O	O	O
그룹 관리자	X	O	O	O
IT 담당자	X	O	O	O
업무 관리자	X	X	X	O
업무 담당자	X	X	O	O

### III. 결 론

본 논문에서는 업무규칙엔진을 활용한 전력요금 관리 시스템을 구성하기 위해 업무규칙 편집기 설계를 제안하였다. 제안된 편집기는 전력요금관리에 특화된 업무규칙 유형을 제공하고, 업무규칙흐름 기능을 통해 조건별 처리 흐름을 정의하며, 사용자 권한에 따라 관리된다. 편집기를 통해 정의된 업무규칙은 다양한 고객 및 계약의 전력요금을 분석 및

검증하는데 활용되어 분산에너지 활성화에 따른 전력요금제도 다각화에 선제적 대응하는 핵심 역할을 수행할 것이다.

### 참 고 문 헌

- [1] 권경선, "「분산에너지 활성화 특별법(안)」에 대한 검토 - 지역에너지 분권의 관점에서 -", 지방자치법연구, vol.23, no.1, 통권 77호, pp. 139-180, 2023.
- [2] 주정은, 구상희, "BRE 도입을 위한 의사결정 모델", 지능정보연구, vol.11, no.3, pp. 103-115, 2005.

# Monolithic Architecture의 설계 유형에 따른 공공기관 형 Micro service Architecture 전환에 관한 연구

서주형\*, 이준호\*\*, 고상원\*\*\*

## Research on Transitioning Public Sector Monolithic Architectures to Micro Service Architectures

JuHyung Seo\*, JunHo Lee\*\*, and SangWon Koh\*\*\*

### 요 약

본 논문은 공공기관의 신규 및 기존 시스템의 Monolithic Architecture 적용 사례에 따라 향후 Micro Service Architecture로 전환 방법에 대해 서술한다. 서비스 지향 아키텍처(SOA, Service Oriented Architecture)와 데이터베이스 중심 Monolithic 시스템과 같은 전환의 유리한 사례와 전환에 불리한 사례를 분석하고 신규 시스템 개발 시 어떤 유형의 Monolithic 구조를 채택하는 것이 Micro Service Architecture로 전환에 비용, 기술, 효율성 확보를 목표로 서술한다.

### Abstract

This paper describes methods for transitioning public sector new and existing systems from monolithic architectures to microservice architectures. It analyzes favorable and unfavorable cases for transition, such as Service-Oriented Architecture (SOA) and database-centric monolithic systems. The study also discusses which type of monolithic structure should be adopted in the development of new systems to ensure cost and technical efficiency when transitioning to a microservice architecture.

### Key words

Micro Service Architecture, Monolithic Architecture, Service-Oriented Architecture, Architecture Transitioning

### I. 서 론

공공기관의 정보시스템은 전통적으로 Monolithic Architecture를 기반으로 운영됐다. 이러한 아키텍처는 초기 개발과 배포가 비교적 쉽다는 장점이 있으

나, 시간이 지남에 따라 확장성과 유지보수 측면에서 여러 한계를 드러내고 있다. 이로 인한 근래에 들어 민간사업자의 정보시스템이 마이크로서비스 아키텍처 전환은 활발히 추진 되고 있으며 정부 및 공공 부분에서도 주목받고 있는 것이 현재 상황이

\*한전KDN(주), juhyung\_0828@kdn.com, \*\*한전KDN(주), junho\_1121@kdn.com, \*\*\*한전KDN(주), swkoh\_1995@kdn.com

다. MSA(Micro Service Architecture)를 적용 시 시스템을 작은 독립적 서비스로 나누어 개발, 배포 및 확장을 독립적으로 수행할 수 있도록 하는 접근 방식으로, 이러한 변화는 특히 대한민국 정부의 디지털 정부 추진과 관련하여 중요한 이슈로 주목받고 있다. 정부는 디지털 플랫폼 정책을 통해 공공기관의 시스템을 보다 유연하고 효율적으로 운영하기 위해 클라우드 네이티브와 마이크로서비스 아키텍처로의 전환을 2030년까지 단계적으로 추진하고 있다. 본 논문에서는 공공기관의 신규 및 기존 시스템을 Monolithic Architecture에서 마이크로서비스 아키텍처로 전환하는 방법을 제안하고, 이 과정에서 발생할 수 있는 장단점을 분석하여, 이를 통해 디지털 정부 구현에 있어 더욱 효과적인 시스템 전환 전략을 모색하고자 한다.

표 1. MA와 MSA의 특징 비교

Table 1. Comparison of the features of MA and MSA.

구분	Monolithic	Micro Service
구조	단일 구조	서비스 단위 구조
확장성	서버 기능 재개발	마이크로서비스 추가
난이도	초기 개발 빠름	초기 개발 어려움
기술 스택	단일 기술 스택	서비스 단위 스택
장애 범위	시스템 전체	서비스 단위
배포	전체 재배포	서비스 단위 배포
통신	내부 프로세스	서비스 메시지 네트워크 영향

## II. 본론

### 2.1 MA 유형별 MSA 전환

이 논문에서는 공공기관에 적용된 3가지 MA 유형에 대한 MSA 전환에 관한 내용을 순차적으로 다루려 한다. 3가지 유형의 사례는 보편적으로 공공기관에 적용된 정보보안 수준에 맞추어 서술되며, 해당 내용은 국정원의 국가 정보보안 기본지침[1] 및 국가 클라우드 컴퓨팅 보안 가이드라인[2]을 준수하여 망 분리 정책 및 네트워크 보안장비 도입을 고려하여 작성된다.

### 2.2 3 Tier MA (Monolithic Architecture)

첫 번째로 일반적인 3 Tier 구조의 Monolithic 구조의 Architecture를 가진 서비스를 전환 시 고려해야 할 사항과 전환 방법에 관해 설명하고자 한다. 이와 같은 구조의 Architecture는 내부 WAS의 기능이 하나의 단일 구조와 단일 개발 코드로 구축되어 있어 서비스 단위의 분리가 최초개발자의 성향에 따라 달라진다. 기능단위의 종속성이 최대한 배제되어 개발되어 있다면, 기존 개발자의 리팩토링만으로도 기능단위의 컨테이너로 구성할 수 있다. 그렇지 않다면, 해당 서버는 재개발 이상의 공수가 투입되는 상황이 발생할 것이고 서비스의 규모에 따라 기능단위 분류보다 본 서버 자체를 컨테이너화하여 구성하는 방식을 고려해야 할 것이다.

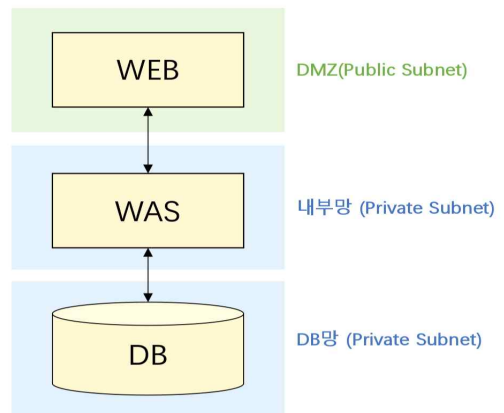


그림 1. 3 Tier MA (Monolithic Architecture)  
Fig. 1. 3 Tier MA (Monolithic Architecture)

### 2.3 복잡한 구조의 Monolithic Architecture

두 번째로 복잡한 구조의 Monolithic 구조의 Architecture를 가진 서비스를 전환 시 고려해야 할 사항이다. 오랜 시간 운영해 온 공공기관의 정보시스템은 보편적으로 그림2와 같은 형태를 가지게 된다. 최초 구성된 정보시스템에 추가적인 기능처리를 위해 API 서버가 추가 되거나 새로운 웹의 기능을 위해 WAS와 WEB 서버가 추가되는 형식이다. 이 같은 경우는 WAS와 API 서버의 규모에 분할 작업 없이 컨테이너로 전환할 수 있는 서버와 분할 작업이 필요한 서버를 분류하고 분할이 필요한 서버의

경우 재개발 규모를 산정하여 서비스 단위로 도메인을 분류해 기능을 묶어준다. 내부에서 RestAPI 혹은 내부 프로세스로 주고받는 데이터의 속성을 분류해 내외부 API Gateway 활용 및 ServiceMesh 구성을 위한 인터페이스 분류 작업을 수행해야 할 것이다.

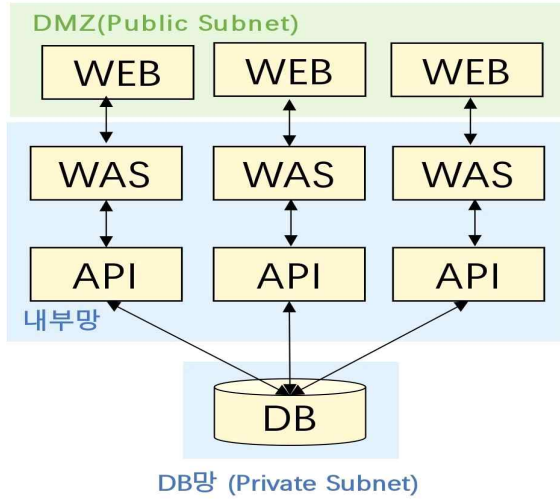


그림 2. 복합적인 구조의 MA (Monolithic Architecture)  
Fig. 2.. Complex structure MA (Monolithic Architecture)

#### 2.4 SOA(Service Oriented Architecture)

세 번째로 서비스 중심의 구조를 가진 Monolithic 구조의 Architecture를 가진 서비스를 전환 시 고려해야 할 사항에 관해 설명한다. 이미 서비스 단위로 분류되어 서버가 구성되어 있으므로 이 경우 서비스 도메인 단위의 분류 작업은 생략할 수 있다. 물론 개발자와 운영자의 설계철학에 따라 추가로 분류하여 나뉘어 질 수 있지만 시간적, 비용적 방면에서 이미 잘 나누어진 서비스를 추가로 나뉘어 관리 포인트를 확장하는 것에 대해서는 실무자의 실익 판단이 중요할 것이다. 이때는 Micro Service Achitecture 전환 시 기반 구성과 CI/CD, Service Mesh와 External Gateway 구성에 대한 부분을 중점적으로 고려하는 것이 중요할 것이다. 공공기관의 네트워크 보안 특성상, 내부망 운영이 필수이며, 게이트웨이와 통신하는 WEB 서버의 경우 DMZ존에서 Proxy 서버의 역할만 수행하게 되기에 내부망에서 자체적인 레포지터리와 레지스트리 저장소 구성과

내부 마이크로서비스를 관리하기 위한 컨테이너 관리 솔루션의 폐쇄망 환경 설치에 대한 부분도 고려되어야 할 것이다.

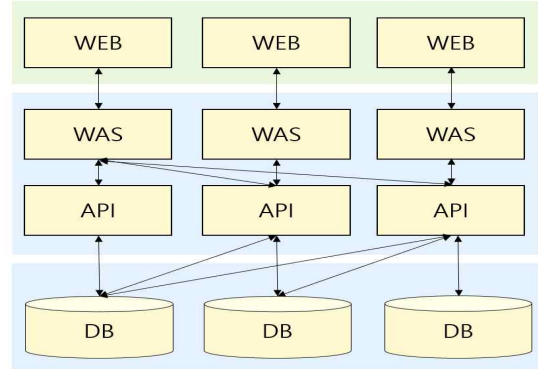


그림 3. SOA(Service Oriented Architecture)  
Fig. 3. SOA(Service Oriented Architecture)

#### 2.5 MSA(Micro Service Architecture)

마지막으로 Monolithic 구조의 Architecture가 최종적으로 구성되어야 할 MSA 구조를 설명하고 전환 작업에 대한 종합적인 내용을 기술하고자 한다.

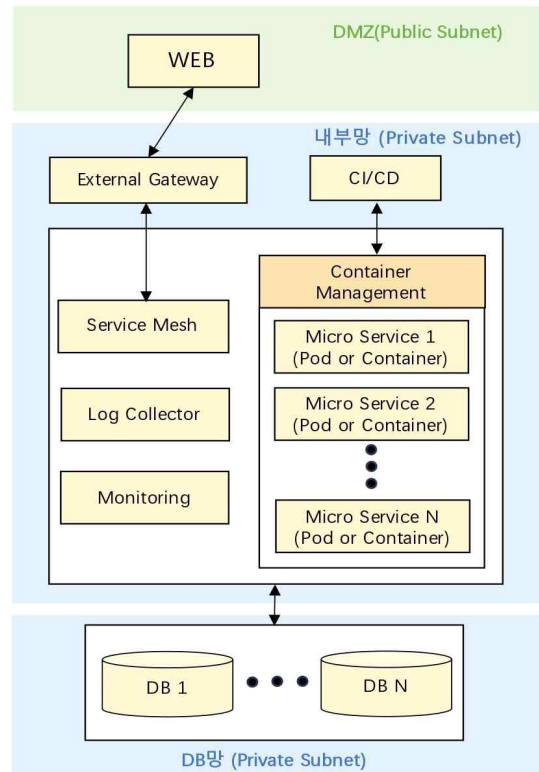


그림 4. MSA(Micro Service Architecture)  
Fig. 4. MSA(Micro Service Architecture)



상기 그림 4와 같이 기존의 Monolithic Architecture가 MSA로 전환됨에 따라 본래 내부 프로세스 레벨에서 동작하던 클래스 단위의 기능들이 하나의 Micro Service가 되고 이들의 트랜잭션 처리를 위해 Service Mesh와 API G/W가 기존의 내부 프로세스와 각각 서버 간의 RestAPI 연계 통신의 역할을 대신하게 된다. 이를 위해 컨테이너 관리 솔루션을 활용한 서비스 관리나 네트워킹 기능을 사용해야 하고, 기존에 수동으로 서버 기능을 업데이트할 때 적용하던 방식은 CI/CD 솔루션 연계를 통해 자동화하여 구성해야 한다. 수많은 기능으로 분산된 애플리케이션 단위의 컨테이너를 직접 빌드하고 배포한다는 것은 사실상 불가능에 가깝기 때문이다. 이 때문에 관리 포인트가 비대해짐에 따라 API 및 트랜잭션 처리에 대한 로그의 관리 또한 중요해진다. 과거 서버의 보안 트러블슈팅을 위해 기록되던 로그가 마이크로서비스의 관리적 측면에서 활용되는 등 그 역할이 확장되는 것이다. 외부 시스템 및 서비스와 통신 역시 기존의 RestAPI 등의 연계 방식에서 내부 마이크로서비스를 운영하는 클러스터와 격리되고 단일화된 하나의 External Gateway를 구축함으로써 API Gateway가 통합적으로 내부 서비스 접근 권한을 통제하고 인증 인가 및 라우팅, 서킷 브레이커와 같은 논리적 네트워킹 관리를 수행하여 안정성과 보안성을 높이게 된다.

## 2.6 공공기관의 MSA

공공기관 정보시스템 구축 가이드에 따르면 외부망과 내부망의 완충 지대인 DMZ 구간과 내부망 업무영역, 그리고 데이터를 전문적으로 관리하는 구간을 물리적으로 분리하여 운영하도록 지침을 내린다. 이는 기관의 Private Cloud 구성 시에도 그 기본 골자가 달라지지 않으며, 더욱 견고한 네트워크 보안 대책을 수립하도록 하고 있다. 이런 엄격한 보안정책은 향후 정부가 추진하는 디지털 정부의 정책을 따라가야 할 공공기관에는 커다란 부담으로 다가올 수 있다. 이를 생각하여 신규로 개발되는 정보시스템에 대한 Architecture를 어떻게 설계할지에 대한 기술적 고려가 필요한 현황이다.

## III. 결 론

본 논문에서는 3가지 Monolithic 구조에 대해 분석하고 MSA 전환에 필요한 사항에 대하여 고찰 하였다. 그림4와 같이 구성되는 것이 현재로서는 공공기관의 Micro Service Architecture 구성의 중간목표가 될 것이다. [3]

위 3단계의 Monolithic Architecture는 실제 3 Tier 구조에서 SOA 구조로 변환되어 가는 과정이기도 하다. 실제로 근래 MA 구조로 개발되는 공공기관의 정보시스템들은 SOA 구조를 채택하여 개발하는데 이는 향후 MSA로 정보시스템을 전환하고 나아가 클라우드 네이티브 환경에서의 운영을 고려한다면, 합리적인 설계 방안으로 보인다. [4] 3가지 예의 경우를 분석했듯 가장 MSA 전환에 시간적, 비용적 절감을 확보한 MA 구조가 SOA 구조의 정보시스템인 것으로 판단됨에 따라 향후 MSA로 구조 전환이 필요하나, 현재 기술 부채와 개발 비용 등 여러 조건을 고려 하여 MA로 우선 개발 후 MSA로 전환하는 정보시스템을 계획하고 있다면 SOA 구조로 개발하는 것이 가장 나은 선택지로 전망된다.

## 참 고 문 헌

- [1] 국가 정보보안 기본지침. 국가정보원. 2021. 11. 1
- [2] 국가 클라우드 컴퓨팅 보안 가이드라인. 국가정보원, 국가보안기술연구소. 2023. 1
- [3] 표준프레임워크 MSA 적용 개발 가이드. 한국지능정보사회진흥원(NIA). 2024. 4
- [4] 클라우드 네이티브 정보시스템 구축을 위한 개발자 안내서. 국지능정보사회진흥원(NIA). 2024. 2. 8



# Wireshark를 활용한 ISO 15118-20 프로토콜 분석에 대한 연구

김 동 완

## Study about analysis of ISO 15118-20 protocol by using wireshark

Dongwan Kim\*

ISO 15118-20은 차량 및 그리드 간 양방향 충·방전을 위한 차세대 V2G (Vehicle-to-grid) 통신 프로토콜이다. ISO 15118-20은 전기차 양방향 충·방전 통신을 포함하여 무선전력전송, 자동연결장치 팬터그래프를 위한 통신 프로토콜을 지원하며, TLS (Transport layer security) 1.3 기반 보안 인증 기능을 제공한다. 다만, TLS 1.3 인증은 강력한 보안 성능을 제공하는 동시에 인가된 사용자도 ISO 15118-20 프로토콜 분석을 방해하는 요소로 작용하기도 한다. 본 논문에서는 전기차 보급과 더불어 확산 중인 차세대 전기차 충전 통신 프로토콜인 ISO 15118-20 데이터 분석을 위한 패킷 스니핑 기술을 소개하고, 암호화된 ISO 15118-20 프로토콜 분석을 위한 요구 사항에 대해 살펴본다. 패킷 스니핑 기술은 네트워크를 통해 전송된 트래픽 패킷을 탐지하고 분석하는 기술로 인가된 사용자에게 네트워크 모니터링을 가능하게 하는 기능을 제공한다. 대표적 패킷 스니핑 프로그램은 GUI 기반 wireshark[1]와 명령어 프롬프트 기반 tcpdump[2]이며, 패킷 스니핑 프로그램을 통해 기존 통신 프로토콜에 대한 분석이 가능하다. 다만, wireshark 및 tcpdump와 같은 기존 패킷 스니핑 프로그램은 TLS 1.3 인증이 포함된 ISO 15118-20에 대한 통신 프로토콜 분석은 제공하지 않는다. 따라서, wireshark에서 ISO 15118-20 프로토콜 트래픽 분석을 위해서는 custom dissector를 통해 wireshark에 ISO 15118-20 프로토콜을 plugin 하여야 한다. 또한, wireshark에서 dissector를 사용하기 위해서는 proto\_register\_protocol 함수 호출을 통하여 dissector 핸들러를 생성하여야 한다. 마지막으로, TLS 1.3 보안 패킷은 Elliptic Curve Diffie Hellman Ephemeral 키 교환 알고리즘[3]을 사용하므로 wireshark를 통해 데이터 패킷 분석을 하기 위해서는 대칭키 정보에 대한 사전 학습이 요구되며, wireshark preference 기능에서 TLS 프로토콜에 대칭키 정보를 입력함으로써 데이터 패킷 분석이 가능하다.

### Key words

ISO 15118-20, packet sniffing, wireshark, TLS 1.3

### 참고문헌

- [1] <https://www.wireshark.org>, accessed on Sep. 07 2024
- [2] <https://www.tcpdump.org>, accessed on Sep. 07 2024
- [3] Haakegaard, R.; Lang, J. The Elliptic Curve Diffie-Hellman (ecdh). 2015. Available online: <https://koclab.cs.ucsb.edu/teaching/ecc/project/2015Projects/Haakegaard+Lang.pdf>, accessed on Sep. 07 2024

---

\* 동아대학교 전자공학과, dongwankim@dau.ac.kr(교신저자)

※ 이 연구는 2024년도 산업통상자원부 및 산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임(RS-2024-00409722)

# 연속수동운동 의료기기의 음성 모듈과 가동범위 측정 알고리즘구현

박민수\*, 손상희\*\*

## Implementation of the Voice Module and Ranging Algorithm of Continuous Manual Exercise Medical Devices

Min Soo Park\*, Sang-Hee Son\*\*

\*† Semiconductor Engineering of Cheongju University

### 요 약

현재 상용화된 CPM(Continuous Passive Motion) 의료기기들 대부분은 모터, 센서, 제어 시스템을 이용해 ROM(Range of Motion)측정을 하며, 의료기기의 데이터를 시각적 또는 디지털 형태로 사용자에게 제공하는 시스템을 채택하고 있다. 하지만 시각적 의료 데이터 제공 방식은 고령 또는 시각적 장애가 있는 환자에게는 불편함을 미칠 수가 있고, 현재 ROM 부하 측정 방식으로는 실시간 피드백 제공과 측정중 생기는 모터 부하 오류와 같은 한계들이 존재한다. 따라서 본 논문에서는 음성 모듈과 고정밀 ROM 측정 알고리즘을 활용해 환자가 ROM 측정 중에 생기는 부하 상태를 실시간으로 이해할 수 있게 지원하고, 의료기기가 즉각적인 피드백을 제공하는 시스템을 제안하고자 한다.

### Abstract

Most of the continuous passive motion (CPM) medical devices currently commercially available adopt a system that measures ROM (Range of Motion) using motors, sensors, and control systems and provides users with data from medical devices in a visual or digital form. However, the visual medical data provision method can be uncomfortable for patients with old age or visual impairment, and the current ROM load measurement method has limitations such as providing real-time feedback and motor load errors during measurement. Therefore, this paper proposes a system that supports patients to understand the load condition generated during ROM measurement in real time through voice modules and high-precision ROM measurement algorithms and provides immediate feedback for medical devices.

### Key words

Continuous Passive Motion, Range of Motion, immediate feedback

---

\* 청주대학교 반도체공학과 석사과정, pms137311@naver.com

\*\* 청주대학교 반도체공학과 교수, shson@cju.ac.kr

※ 이 논문은 2024~2025년도 청주대학교 연구장학 지원에 의한 것임.

## 1. 서 론

노인 인구가 증가함에 따라 관절 관련 질환 환자도 같이 증가하는 현대 사회에서 관절 질환을 가지고 있는 환자의 수와 관절 가동 범위를 점차 늘려주는 CPM 의료기기에 시장에서의 수요가 꾸준히 증가하고 있으며, ROM 측정을 위한 연구도 활발하게 진행중에 있다[1]. CPM은 관절의 강직됨을 방지하며, 혈액 순환을 원활하게 하는데 도움을 줘 재활 과정을 단축시키는데 도움이 된다[2]. 현재 ROM 측정 기능이 있는 CPM 의료기기는 실시간으로 기기 상태를 알 있게 해주는 음성 피드백 기능을 가지고 있지 않는다. 따라서 본 논문에서는 환자에게 이전보다 모터의 부하감지 능력이 뛰어난 고정밀 ROM 측정방식으로 측정된 각도 데이터와 음성모듈을 활용한 실시간 피드백을 통해 치료 과정을 개선할 수 있는 시스템을 제안한다[3, 4].

## II. ROM(Range of Motion) 측정 모드

### 2.1 ROM 측정 알고리즘

본 논문에서는 CPM 의료기기를 이용해 운동중인 환자의 관절 가동 범위를 측정하기 위해 모터에 흐르는 전류 감지 방식을 채택하였으며, 모터의 제어를 위해 쓰이는 모터드라이버 IC A4955의 전류 측정 기능을 사용하였으며, 해당 IC의 내부 구조도는 Fig. 1.이고 AIOUT 핀을 통해 출력된 전압 데이터가 MCU의 내부 ADC 핀으로 입력되어 전류를 측정한다. 입력된 ADC 신호를 활용한 전류 측정은 직류 전력공급기를 사용하여 200mA, 3000mA 두 전류값의 해당 AD 값을 이용해 두 지점의 포인트를 찍어 직선의 방정식으로 전류를 추적하는 방식으로 계산하였다. 다음으로 측정된 전류값을 활용하여 부하를 감지하는 알고리즘의 순서는 Fig 2 부하 측정 순서도와 같다. 입력된 ADC 신호 값을 0.1ms 카운터에 두고 100ms 동안 1000개의 ADC값을 더한 후 1000으로 나눠 평균을 내주는 방식으로 전류를 측정하였고 평균낸 값을 다시 100ms 주기로 Arr 배열에

저장하여 총 5개의 값을 다시 평균을 내었다. 평균값이 기준 Sensitivity(감도)보다 높다면 부하 감지가 되도록 설정했다. 여기서 기준 감도는 개발자가 LCD의 히든모드(개발자모드) 화면에서 총 10단계로 정해진 감도를 선택하는 방식으로 구현하여 의료기기의 부하 감지량을 다르게 설정하는 것이 가능해서 부하 감지 정확도가 높아졌다.

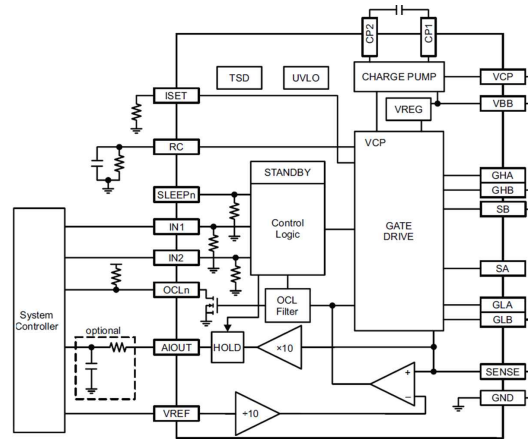


그림 1. A4955 드라이버 IC 내부 구조도  
Fig. 1. A4955 Driver IC internal structure

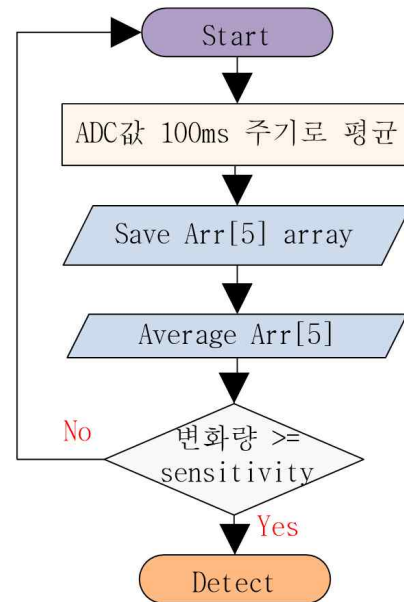


그림 2. ROM 측정 순서도  
Fig. 2. ROM measurement FLOWchart

하지만 높은 감지 정확도에도 불구하고 현재 의료기기의 문제점이 존재하는데 각도마다 부하량이 다르다는 점이다. 이러한 각도별 부하의 차이를 고려하여 각도마다 부하 감도의 설정을 다르게 설정

을 해렸으며, 90° ~ 170° 에서의 감도는 10° ~ 90° 감도 보다 대략 10% 정도 높게 설정해주었으며, 모터가 상한각으로 이동할 때는 하한각으로 이동할 때보다. 전류 부하 소모가 더 크다는 점을 고려하여 20%를 더욱 높게 설정을 하여 문제를 극복하였다.

## 2.2 음성모듈을 활용한 실시간 피드백

표 1. 음성 데이터 인덱스  
Table. 1. Voice data index

번호	인덱스	음성파일	번호	인덱스	음성파일
16	0x12	측정모드	19	0x15	측정이 완료되었습니다
17	0x13	측정을 시작합니다	21	0x17	비상정지 되었습니다
18	0x14	측정을 중지합니다	39	0x2A	부하가 감지되었습니다

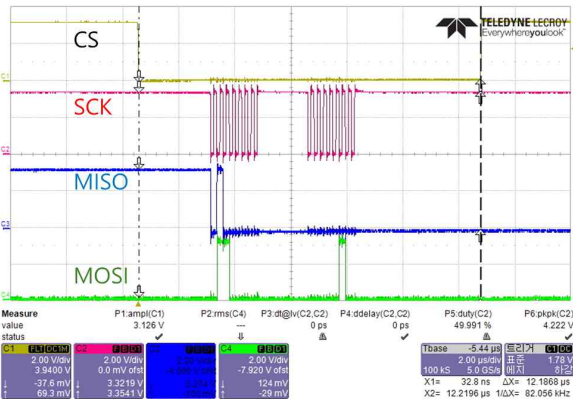


그림 3. 음성모듈 출력 신호 파형  
Fig. 3. Voice module output signal waveform

음성모듈을 사용하여 측정모드의 실시간 피드백 구현을 위해 Nuvoton Technology사의 ISD3800 음성 모듈을 사용하였으며, 128Mbit의 비휘발성 메모리를 가지고 있어 최대 16분 가량의 음성 파일 저장 가능 가능하다. MCU의 SPI의 MOSI에 표 1.의 인덱스 데이터를 보내면 해당 음성 데이터가 출력되는 구조를 가지고 있으며, Digital to Analog가 내장되어 있어 개별적인 신호처리를 하지 않아도 오디오 출력을 가능하도록 한다. 또한 인덱스 1개가 출력되는데 까지 소요되는 시간인 CS의 간격은 오실로스코프로

측정한 결과, 그림 3.과 같이 12 $\mu$ s의 소요시간이 확인되었다. 이를 통하여 고정밀 측정 알고리즘의 소요 시간이 짧은 것을 확인하였다. 따라서 CPM 의료기기의 상태를 실시간으로 확인 할 수 있게 되어 고정밀 측정 알고리즘의 효과를 더욱 극대화시킬 수 있었다.

## III. 결 론

본 논문에서는 고정밀 ROM(Range of Motion) 측정 알고리즘과 음성 모듈을 활용한 실시간 피드백을 제공 해주는 시스템을 개발하였다. 기존의 CPM 의료기 기들이 가진 한계점을 보완하고, 노인과 시각적 장애 를 가진 환자들에게 편리한 재활 환경을 제공하였다. 이를 통해 재활 운동에서의 효율성을 높였으며, 향후에는 다양한 환자들이 제공한 데이터를 활용하여 최 적의 ROM 측정 알고리즘 개선과 음성 데이터를 활용 한 사용자 인터페이스 최적화를 통해 CPM 의료기기의 재활 성능을 더욱 향상시킬 수 있을 것이다.

## 참 고 문 헌

- [1] Lee, K., Yang, O., Lee, C., "Design of Multifunctional Compound Joint Medical Equipment for Continuous Passive Motion," Journal of the Semiconductor & Display Technology, 21(4), pp. 126-131, (2022).
- [2] Jang, M., Shim, J., Lee, E., Choi, S., "Design of Continuous Passive Motion Medical Device System with Range of Motion Measurement Function," Journal of the Semiconductor & Display Technology, 22(4), pp. 87-92, (2023).
- [3] Kim, Y., "Study on the Integrated System of Potentiometer and Encoder for Motion Control," Journal of the Institute of Electronics Engineers of Korea, 45(3), pp. 354-362, (2011). Journal Paper
- [4] Jinwoo Park et al., "Design and Implementation of a Voice Feedback System in Medical Devices", Journal of the Korean Society of Medical & Biological Engineering, 2016.

# 센싱-통신 통합 시스템을 위한 프리코더 설계

서방원\*, 강민석\*\*, 공병용\*\*, 이상원\*\*

## Precoder design for ISAC systems

Bangwon Seo\*, Min-Seok Kang\*\*, Byeong Yong Kong\*\*, and Sang-Won Lee\*\*

### 요약

이 논문은 통신 기술과 레이더 센싱 기술을 통합하는 ISAC 시스템에서 통신 성능과 레이더 신호 센싱 성능을 동시에 고려하여 프리코딩 행렬을 설계하는 방법에 대해서 연구한다. 대부분의 기존 연구들은 통신 성능에 대한 요구 조건을 만족시키면서 레이더 신호의 SNR을 최대화시키도록 프리코딩 행렬을 설계하였다. 반면, 이 연구에서는 레이더 신호의 SNR에 대한 요구 조건을 만족시키면서 모든 사용자에게 대한 데이터 전송률 합이 최대가 되도록 프리코딩 행렬을 설계하는 방법은 제안하였다.

### Abstract

This paper studies a method for designing a precoding matrix in an integrated sensing and communications (ISAC) system that considers both communication performance and radar signal sensing performance. Most previous studies have focused on designing the precoding matrix to maximize the SNR of the radar signal while satisfying the requirements for communication performance. In contrast, this study proposes a method for designing the precoding matrix to maximize the sum data rate for all users while meeting the SNR requirements for radar signals.

### Key words

integrated sensing and communications (ISAC), precoder design, signal-to-noise ratio (SNR) constraints

## 1. 서론

통신과 센싱을 통합하는 기술(ISAC, Integrated Sensing and Communications)은 차세대 무선 통신 시스템에서 중요한 역할을 할 것으로 기대된다. 특히 자율주행, 드론, 스마트시티 등의 다양한 응용 분야에서 실시간으로 환경을 인식하고 데이터를 전송할 수 있는 능력을 제공하기 위해서는 ISAC 기술이 필수적으로 사용되어야 한다 [1]-[3]. ISAC 시스템은 동일한 무선 자원을 효율적으로 사용하여 통신과

레이더 센싱 기능을 동시에 수행할 수 있도록 설계된다 [4]-[5]. 이로 인해, 제한된 주파수 대역을 효율적으로 사용하고자 하는 요구가 증가하는 상황에서, 통신 성능과 센싱 성능을 동시에 고려하는 프리코딩 행렬 설계 기술이 필수적으로 요구된다.

현재 대부분의 연구들에서는 주로 통신 신호에 대한 신호-대-간섭 잡음비(SINR)가 특정 임계값 이상이 되도록 하는 제약 조건 하에서 레이더 신호의 신호-대-잡음비(SNR)가 최대가 되도록 프리코더를 설계하는 방식이 주를 이루고 있다. 이러한 접근 방

\*공주대학교 (교신저자, seobw@kongju.ac.kr), \*\*공주대학교

식은 통신 신호의 품질은 보장할 수 있지만, 레이더 신호 센싱 성능을 개선하는데는 어느 정도 한계가 존재하게 된다.

이 논문에서는 기존 연구와 차별화하여, 레이더 신호의 SNR이 특정 임계값을 초과해야 된다는 요구 조건 하에서, 통신 사용자들의 데이터 전송률 합이 최대가 되도록 프리코더를 설계한다. 이는 통신 신호의 품질을 최우선으로 고려하면서도 레이더 센싱 성능을 일정 수준 이상으로 유지할 수 있는 새로운 관점을 제시한다.

## II. 제안하는 프리코더 설계 기법

여기에서는 이동통신 시스템의 하향링크 통신과 레이더 신호 센싱을 동시에 수행하는 ISAC 시스템에서 기지국을 위한 프리코딩 행렬을 설계하는 기법을 제안한다.

기지국에서의 전송 신호는 다음과 같이 표현할 수 있다.

$$\mathbf{x} = \mathbf{W}\mathbf{s} = \sum_{m=1}^M \mathbf{w}_m s_m \quad (1)$$

여기에서  $\mathbf{w}_k$ 와  $s_k$ 는 각각 k-번째 사용자를 위한 프리코딩 벡터와 데이터 심볼을 나타낸다.

기지국에서 수신된 레이더 신호는 다음과 같이 표현할 수 있다.

$$\mathbf{y}_R = \alpha_t \mathbf{a}(\theta) \mathbf{a}^H(\theta) \mathbf{x} + \mathbf{z}_R \quad (2)$$

여기에서  $\alpha_t$ 는 타겟에서의 반사계수를 나타내고,  $\mathbf{a}(\theta)$ 는  $\theta$  방향에 대한 steering 벡터를 나타낸다. 또한 벡터  $\mathbf{z}_R$ 은 백색 가우시안 부가잡음 벡터를 나타낸다.

레이더 신호 센싱을 위한 검파기를  $\mathbf{u}$ 라고 하면, 검파기 출력 신호는 다음과 같이 표현할 수 있다.

$$\mathbf{u}^H \mathbf{y}_R = \alpha_t \mathbf{u}^H \mathbf{a}(\theta) \mathbf{a}^H(\theta) \mathbf{x} + \mathbf{u}^H \mathbf{z}_R \quad (3)$$

이제 타겟에 대한 레이더 신호의 SNR은 다음과 같이 주어진다.

$$\text{SNR}_R = \frac{\alpha_t^2 \|\mathbf{u}^H \mathbf{A} \mathbf{W}\|^2}{\sigma_z^2 \mathbf{u}^H \mathbf{u}} \quad (4)$$

여기에서  $\mathbf{A} = \mathbf{a}(\theta) \mathbf{a}^H(\theta)$ 이고,  $\sigma_z^2$ 는 부가잡음에 대한 분산 값을 나타낸다.

한편 하향링크 통신 시스템에서 M개의 사용자가 존재하고, 각 사용자는 수신 안테나 개수가 1개라고 가정한다. 이 때, k-번째 사용자에서의 수신 신호는 다음과 같이 표현할 수 있다.

$$y_k = \mathbf{h}_k^H \left( \sum_{m=1}^M \mathbf{w}_m s_m \right) + z_k \quad (5)$$

여기에서  $\mathbf{h}_k$ 는 기지국과 k-번째 사용자 간의 채널 벡터를 나타낸다.

k-번째 사용자에 대한 통신 신호의 SINR은 다음과 같이 주어진다.

$$\text{SINR}_k = \frac{\|\mathbf{h}_k^H \mathbf{w}_m\|^2}{\sum_{\substack{m=1 \\ m \neq k}}^M \|\mathbf{h}_k^H \mathbf{w}_m\|^2 + \sigma_z^2} \quad (5)$$

제안하는 프리코딩 설계 기법은 레이더 신호의 SNR에 최소 조건을 만족시키면서 모든 사용자에 대한 데이터 전송률의 합이 최대가 되도록 프리코딩 행렬을 설계한다. 이 문제는 다음과 같은 최적화 문제로 표현할 수 있다.

$$\begin{aligned} \max_{\mathbf{W}} \quad & \sum_{k=1}^K \log_2(1 + \text{SINR}_k) \\ \text{subject to} \quad & \text{SNR}_R \geq \Gamma_{th} \end{aligned} \quad (1)$$

최적화 기법을 사용하여 위 문제에 대한 해를 구할 수 있다.

## III. 결론

이 논문에서는 ISAC 시스템에서 통신 사용자들에 대한 데이터 전송률 합이 최대가 되도록 하면서도 레이더 신호의 SNR을 일정 임계값 이상으로 유



지하는 프리코더 설계 방법을 제안하였다. 기존 방식과 달리, 이 연구는 레이더 성능을 우선 보장하면서 통신 신호의 품질을 극대화하는 접근을 통해 통신과 센싱 성능 간의 균형을 효과적으로 조정하였다. 본 연구의 결과는 차세대 ISAC 시스템의 프리코딩 행렬 설계에 있어서 통신과 센싱 성능 향상에 실질적인 기여를 할 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1] Liu, F., et al. "Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond", *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728-1740, June 2022.
- [2] X. Fang, W. Feng, Y. Chen, N. Ge, and Y. Zhang, "Joint communication and sensing toward 6G: Models and potential of using MIMO", *IEEE Internet Things J.*, vol. 10, no. 5, pp. 4093-4116, March 2023.
- [3] H. Luo, R. Liu, M. Li, and Q. Liu, "RIS-Aided integrated sensing and communication: Joint beamforming and reflection design", *IEEE Trans. Vehic. Technol.*, vol. 72, no. 7, pp. 9626-9630, July 2023.
- [4] R. Zhang, Y. Wang, L. Mou, and Y. Zhang, "Research on beamforming design of ISAC system: An FP approach", *IEEE Wireless Commun. Lett.*, vol. 13, no. 4, pp. 964-968, April 2024.
- [5] Y. Liu, M. Jin, G. Guo, and J. Yao, "Secure beamforming for NOMA-ISAC with system imperfections", *IEEE Commun. Lett.*, vol. 28, no. 7, pp. 1559-1563, July 2024.

# 증분 학습과 객체 인식을 활용한 이미지 기반 자동 결제 시스템

배소이\*, 이준성\*\*

## Image-Based Automatic Payment System Utilizing Incremental Learning and Object Recognition

Soyee Bae\* and Jun Seong Lee\*\*

본 연구는 증분 학습과 객체 인식 기술을 활용한 이미지 기반 자동 결제 시스템을 제안한다. 기존의 바코드 시스템은 상품에서 바코드의 위치를 찾거나, 여러 상품을 구매할 경우 상품마다 바코드를 일일이 스캔해야 하는 불편함이 있다. 이에 반해, 본 연구에서 제안하는 시스템은 카메라로 촬영된 단일 이미지로 여러 상품을 한 번에 인식할 수 있고, 다양한 각도에도 강인하게 객체를 인식할 수 있다는 장점을 지닌다[1]. 제안하는 시스템은 YOLOv8[2] 기반의 객체 인식 모델을 사용하여 점포에서 판매하는 다양한 제품을 인식하도록 설계되었다. 점포에서 판매 중인 제품 이미지들을 다양한 각도에서 촬영하여 수집하고 레이블링함으로써 데이터셋을 구축하고, 구축된 데이터셋을 활용하여 YOLOv8 모델을 학습시킴으로써 점포에서 판매 중인 제품들을 인식할 수 있도록 한다. 또한, 점포에 새로운 제품이 추가될 경우 증분 학습(Incremental Learning)[3] 방법을 통해 모델을 지속적으로 업데이트할 수 있도록 설계하였으며, 기존 데이터를 다시 학습하지 않고도 새로운 제품을 효과적으로 인식할 수 있도록 하였다. 더 나아가, 시스템의 유연성을 강화하기 위해 알수없음(Unknown) 처리 메커니즘을 도입하여, 점포에서 판매하지 않는 제품들이나 학습되지 않은 물체들에 대해 잘못된 계산이 발생하지 않도록 하였다. 결과적으로, 제안된 시스템은 매장 내 상품 인식 및 결제 과정을 간소화하고, 지속적인 제품 업데이트를 통해 실시간으로 적용할 수 있는 고도화된 결제 솔루션을 제공함으로써, 사용자의 경험을 향상시킬 수 있을 것으로 기대된다.

### Key words

Automatic payment system, Object recognition, Incremental learning

### 참고문헌

- [1] 김희영. "빅데이터 분석에 기반한 AI 셀프서비스기술 자동결제 시스템 설계." 국내박사학위논문 광운대학교 대학원, 2022. 서울
- [2] Ultralytics, "YOLOv8: A state-of-the-art object detection model," GitHub repository, <https://github.com/ultralytics/ultralytics>, 2023.
- [3] Zhou, Da-Wei, et al. "Pycil: A python toolbox for class-incremental learning." (2023): 197101.

---

\* 니어블리스, ssoyeee@gmail.com(교신),

\*\* 한국전자통신연구원, jslee0708@etir.re.kr

# 1인 미디어 창작자용 AI 기반 인물 조명 알고리즘 설계

이용창\*

## Designing AI-Based Portrait Lighting Algorithm for One-Person Media Creators

YongChang, Lee\*

### 요 약

본 연구는 유튜브와 같은 비디오 플랫폼에서 1인 미디어 창작자를 위한 인공지능(AI) 기반 인물 조명 알고리즘을 설계하는 것이다. 전통적인 조명 설정의 복잡성과 비용 문제를 해결하기 위해, Convolutional Neural Network(CNN)와 Dlib의 얼굴 랜드마크 탐지 알고리즘을 사용하여 얼굴을 인식하고 주요 랜드마크를 추출하였다. HDR 이미징 기술로 주변 조명 조건을 분석하고, 실시간 데이터 분석 및 피드백 루프를 통해 최적의 조명 방향과 강도를 자동으로 설정하였다. 이를 통해 유튜브 개인 방송의 영상 품질 향상과 1인 미디어 창작자들이 보다 쉽게 고품질의 영상을 제작할 수 있도록 지원하며, AI 기반 인물 조명 기술은 다양한 분야에서 활용될 것으로 기대된다.

### Abstract

The purpose of this study was to design and implement an AI-based portrait lighting algorithm for solo content creators on video platforms such as YouTube. To avoid the complexity and cost issues of traditional lighting setups, we used a Convolutional Neural Network (CNN) and Dlib's facial landmark detection algorithm to recognize faces and extract key landmarks. We then analyzed surrounding lighting conditions using HDR imaging technology and automatically set the optimal lighting direction and intensity through real-time data analysis and feedback loops. This approach was able to enhance the video quality of YouTube solo broadcasts, making it easier for solo content creators to produce high-quality videos. This AI-based portrait lighting technology is expected to be utilized in various fields.

### Key words

YouTube, AI-Based, Portrait Lighting, Algorithms, One-Person Media

### 1. 서 론

유튜브와 같은 비디오 플랫폼은 콘텐츠 크리에이터들이 자신의 개성을 표현하고, 다양한 주제에 대

해 소통하는 공간을 제공한다. 특히, 개인 미디어 방송은 전문적인 장비 없이도 고품질의 콘텐츠를 제작할 수 있도록 한다. 그러나, 높은 영상 품질을 유지하기 위해서는 적절한 조명 설정이 필수적이다.

\*동아방송예술대학교 방송기술과, yclee@dima.ac.kr

전통적인 조명 설정은 시간과 비용이 많이 소요되며, 복잡한 조명 장비를 조작하는 데 높은 숙련도가 필요하다. 이러한 문제를 해결하기 위해 AI 기반의 인물 조명 시스템이 필요하다. AI 기술을 활용하면 효율적이고 최적화된 조명 설정을 자동화할 수 있어, 콘텐츠 크리에이터들이 조명 설정에 대한 부담을 덜고 콘텐츠 제작에 집중할 수 있다. 인물의 얼굴을 인식하고 주요 랜드마크를 추출하여 얼굴의 구조를 분석하며, 주변 조명 조건을 분석하여 최적의 조명 방향과 강도를 결정하고, 최적의 조명 조건을 계산하여 실시간으로 적용하여 실시간으로 조명 조건을 모니터링하고 최적화된 조명 설정을 유지하는 것이다[1][2][3].

## II. AI 기반 인물 조명 알고리즘 설계

본 연구에서는 AI 기반 인물 조명 알고리즘을 설계하고 구현하기 위해 여러 단계의 연구 방법을 적용하였다. 먼저 얼굴 인식과 랜드마크 추출 단계에서는 Convolutional Neural Network(CNN)를 사용하여 사진이나 영상에서 인물의 얼굴을 탐지한다[4]. 얼굴이 탐지되면, Dlib의 얼굴 랜드마크 탐지 알고리즘을 사용하여 얼굴의 주요 특징점(랜드마크)을 추출하여 시각화한다. 조명 환경 분석 단계에서는, 얼굴 랜드마크 데이터를 바탕으로 얼굴의 조명 프로필을 생성한다. 이를 위해 HDR(High Dynamic Range) 이미징 기술을 사용하여 주변 조명 조건을 분석하고, 얼굴의 입체감을 높이기 위한 최적의 조명 방향과 강도를 결정하였다[5]. 조명 최적화 및 적용 단계에서는, 얼굴 랜드마크와 조명 프로필 데이터를 기반으로 최적의 조명 방향과 강도를 계산하였다. 주요 조명(키 라이트)은 일반적으로 30° 각도로 설정되며, 보조 조명(필 라이트)과 배경 조명(립 라이트)은 각각 60°와 90° 각도로 설정하였다. 조명의 강도는 주변 환경의 조도와 색온도를 고려하여 자동으로 조정된다. 실시간 데이터 분석 및 피드백 루프 단계에서는, 알고리즘이 실시간으로 수집된 데이터를 분석하여 조명 설정을 최적화하도록 한다. 조도와 색온도 센서를 사용하여 주변 환경의 조명 조건을 모니터링 하며, 실시간 피드백 루프를

통해 조명 설정을 조정한다. 촬영 중에 조명의 각도나 강도가 변화할 경우, 알고리즘은 이를 감지하고 즉각적으로 최적의 조명 조건을 다시 계산하여 적용한다. 이와 같은 방법을 통해 AI 기반 인물 조명 시스템을 설계하고 구현함으로써, 유튜브 개인 방송의 영상 품질 향상 및 고품질의 영상을 제작할 수 있도록 지원할 수 있을 것이다.

## III. 결 론

AI 기반 인물 조명 알고리즘은 유튜브 등 개인 방송의 영상 품질을 향상할 수 있다. 이러한 알고리즘은 얼굴 인식, 조명 최적화, 실시간 데이터 분석 등을 통해 최적의 조명 조건을 자동으로 제공하며, 앞으로 더욱 다양하고 창의적인 활용 가능성을 열어줄 것이다. AI 기반 인물 조명 기술의 발전은 유튜브 개인 방송, 프로필 사진 촬영 등 다양한 분야에서 활용될 수 있을 것이다.

## 참 고 문 헌

- [1] Kazemi, Vahid, and Josephine Sullivan. "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions." *Computation*, 2023, 11(3), 52; <https://doi.org/10.3390/computation11030052>
- [2] Keiron O., Ryan N., "An Introduction to Convolutional Neural Networks", <https://ar5iv.labs.arxiv.org/html/1511.08458>
- [3] Zhang, Zhanpeng, et al. "Facial landmark detection by deep multi-task learning." *European conference on computer vision*. Springer, Cham, 2014.
- [4] Alzubaidi, L., Zhang, J., Humaidi, A.J. et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions", *J Big Data* 8, 53 (2021). <https://doi.org/10.1186/s40537-021-00444-8>
- [5] Google AI Blog. "Portrait Light: Enhancing Portrait Lighting with Machine Learning." Retrieved from Google AI Blog

# 한약 안전사용 및 한방데이터 품질관리 플랫폼 개발을 위한 설계

김헌성\*, 이재훈\*\*

## Korean medicine data quality management platform design

Hun Sung Kim\*, jaehoon Lee\*\*

### 요 약

한방 데이터는 한의약의 과학화와 체계화를 위해서 중요한 자원이다. 국내 한방 데이터 관리현황을 조사한 결과 14곳의 기관에서 한방 데이터를 제공하고 있었다. 그러나 대부분의 한방데이터가 혼재 및 중복으로 관리되고 있고, 데이터 제공에 있어서도 DB파일 제공이 아닌 검색을 중심으로 한 열람방식으로 데이터를 제공하고 있어서 데이터 활용 및 접근성이 낮은 상황이다. 이에 한의학 데이터 수집 및 제공파트와 데이터 관리파트가 분리되어 있고, OpenAPI를 통해 기업, 병원, 대학 등과 공유할 수 있도록 하는 한방데이터 품질관리 플랫폼 개발이 필요한 상황이다. 본 연구에서는 한방데이터 품질관리 플랫폼 개발을 위한 현재 시스템 분석, 개선 프로세스 설계, 요구사항 분석, Usecase Diagram을 연구하고자 한다.

### Abstract

Korean medicine data serves as a critical resource for the scientific and systematic advancement of traditional Korean medicine (TKM). A survey on the current state of TKM data management in South Korea reveals that 14 institutions provide access to TKM-related datasets. However, the majority of this data is inconsistently and redundantly managed, and the predominant access model relies on a search-based browsing interface rather than downloadable database files, thereby limiting data accessibility and usability. To address these issues, a quality management platform is required that separates data collection and provision from data management and facilitates data sharing with businesses, hospitals, and universities via OpenAPI. This study aims to examine the requirements for developing such a TKM data quality management platform, focusing on needs system analysis, process design, software requirements analysis, usecase design..

### Key words

Korean Medicine Data, Open API, Software Structure Design

---

\* 메디케이스스템 기업부설연구소, oioid@naver.com

\*\* 메디케이스스템 기업부설연구소, herblinker@gmail.com

※ 본 연구는 보건복지부의 재원으로 한국보건산업진흥원의 보건의료기술연구개발사업 지원에 의하여 이루어진 것임  
( 과제고유번호 : HI23C149000 )

## 1. 서 론

한의학은 고유한 치료 방식과 원리를 중심으로 발전해왔으나, 학문적 특성으로 인해 관련 데이터를 체계화, 표준화 하는 데에 어려움을 겪었다. 최근 정부와 한국한의학연구원을 중심으로 한약 국제표준화 기반 구축 연구를 진행하고 있으나, 데이터의 분산 관리와 통합 부재 등으로 인한 데이터 품질 관리 문제가 대두되고 있으며 이는 관련 데이터를 활용한 추가적인 가치창출 및 접근성에 부정적인 영향을 미치고 있는 상황이다. [8]

또한 환자들이 한약의 안정성에 대해 신뢰할 수 있는 자료를 찾기 어려운 실정이다. 현재 한의학 진흥원에서는 매년 한의학산업실태조사 및 한방의료 이용 및 한약소비 실태조사를 진행하여 한의학의 개선 사항과 이용 만족도에 대한 데이터를 도출 및 제공하고 있으나, 구체적인 한약의 안전성 정보, 약재별 작용과 부작용, 이에 대한 소비자 인식 수준을 다루는 실질적인 조사는 이루어지지 않고 있다. 이러한 신뢰성 데이터 부족은 한의학 소비자에게 중요한 한계점으로 나타나고 있다. [16]

이에 본 연구의 2장에서는 한약 안전 사용 플랫폼 개발을 위해 현재 다양한 한방 데이터 시스템에 대한 분석을 진행하였다. 또 이를 기반으로 3장에서는 본격적인 시스템의 개선프로세스를 설계하고 이에 대한 요구사항 분석, 시스템 구성도, 유스 케이스, 설계방법들을 연구하고자 한다.

## II. 현재 시스템 분석

그림1은 기존 한방데이터 검색 시스템의 기능분포도이다. 데이터 수집의 경우 모든 시스템에서 자체적으로 수집하고 있었고 검색기능은 제공하고 있었으나 OpenAPI 형태로 공유하는 곳은 없었다. 데이터 업데이트 이력이 없는 곳이 대부분으로 데이터 품질관리에도 어려움이 있는 것으로 파악되었다. 이에 본 논문에서 데이터 수집에서부터 검색기능, OpenAPI 기능, 데이터 업데이트 및 품질관리 기능이 있는 시스템을 설계하고자 한다.[1][2][3][4][5]

[6][7][8][9][10][11][12][13][14]

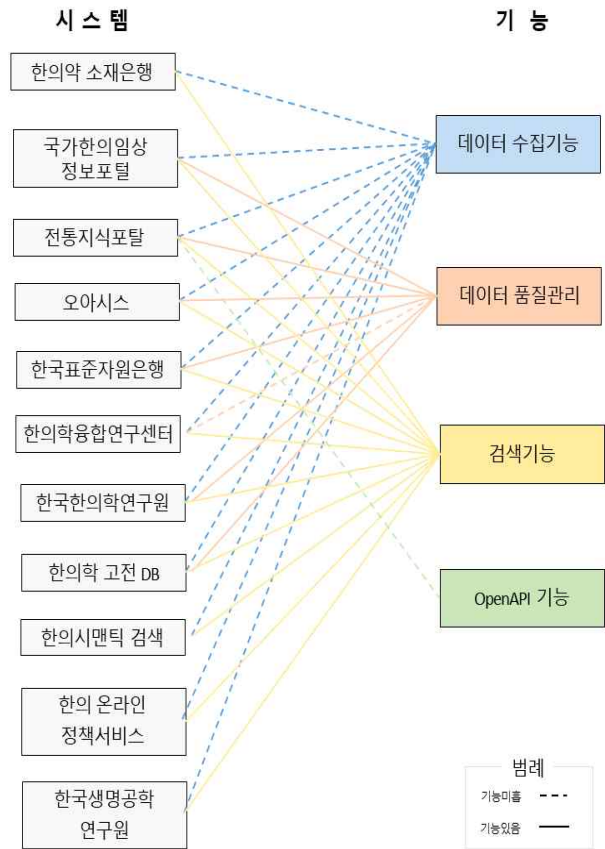


그림 1. 기존 한방데이터 검색 시스템 기능분포도  
Fig. 1. Function distribution diagram of existing Korean medicine data search system

## III. 개선 프로세스 설계

본 시스템의 목적은 첫째 데이터 수집파트 개발로 데이터 수집 및 제공자가 플랫폼을 통해 데이터를 업로드하면 플랫폼에서 데이터 정제 및 검수를 통해 정제된 데이터를 플랫폼에 적용시키고자 한다. 둘째 데이터 검색파트 개발로 플랫폼을 통해 일반 사용자들은 데이터를 검색할 수 있도록 한다. 마지막으로 데이터 공유파트 개발로 OpenAPI를 통해 데이터가 공유될 수 있도록 하고자 한다. 이는 기존에 14곳의 기관에서 한방데이터가 혼재 및 중복으로 관리되고 있어 데이터 활용 및 접근성이 낮았는데 통합관리를 통해 한방 데이터를 효율적으로 관리할 수 있게 한다.

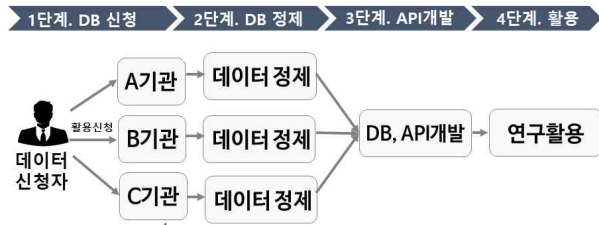


그림 2 기존 한방 데이터 연구 활용 프로세스 [4]

Fig. 2 Process of utilizing existing oriental medicine data research

그림2는 기존에 한방데이터를 활용하기 위한 프로세스를 도식화한 것이다. 데이터 신청자는 각 데이터 보유기관에게 데이터 활용을 신청하고, 데이터를 수신 받는다. 수신 받은 데이터를 연구데이터로 활용하기 위해서 데이터 정제작업을 거치고, 정제가 된 데이터를 연구에 활용하기 위해 데이터베이스화하여 API로 개발한다. 이렇게 연구데이터를 개발이 후에 연구에 직접적으로 활용할 수 있다. 이러한 방식은 한방 데이터를 연구에 활용하기 위해 데이터 신청자는 매번 데이터를 각 기관에 요청해야하고, 연구데이터로 활용하기 위해 데이터를 정제해야하여 비효율적이다.[4] 본 시스템을 활용한 개선된 프로세스는 그림3과 같다.

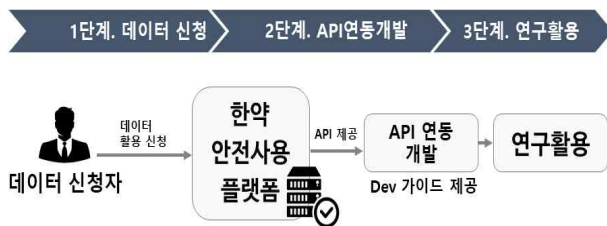


그림 3 개선된 프로세스

Fig. 3 Improved Process

그림3은 개선된 한방데이터 활용 프로세스이다. 데이터 신청자는 본 시스템을 통해서 원하는 한방 데이터를 선택하여 데이터 활용 신청을 할 수 있고, 신청이 승인되면 정제된 데이터로 구성된 API를 바로 연구에 활용할 수 있다. 데이터 신청자는 본 시스템을 통해서 데이터 제공 신청에서부터 API구축까지 한 번에 할 수 있어 기존에 여러 기관에 데이터 활용을 신청하여, 정제되지 않은 데이터를 직접

정제하고 API를 구축해서 연구에 활용해야했던 복잡하고 비효율적인 프로세스를 간소화 할 수 있다. 본 시스템의 API key발급 절차는 그림4와 같다.

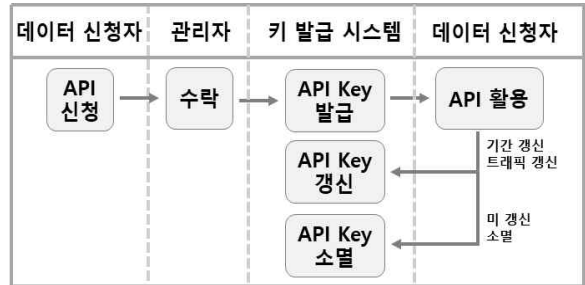


그림 4 API key 발급 프로세스

Fig. 4 API key issuance process

데이터 신청자는 본 시스템을 통해서 원하는 한방 데이터를 선택하여 API제공을 신청하고, 시스템의 관리자가 내용을 확인 후 수락버튼을 클릭하면 key 발급 시스템에 의해 자동으로 API key를 발급한다. API key는 유효기간, 트래픽에 대한 제한이 있으며 이를 API사용량에 따라 이를 갱신하여 사용할 수 있다. 기간 내에 API키를 갱신하지 않으면 API key는 소멸된다. 데이터 신청자는 발급받은 API key를 통해서 데이터를 바로 활용할 수 있다. 본 시스템의 데이터 제공자의 신청 프로세스는 그림5와 같다.

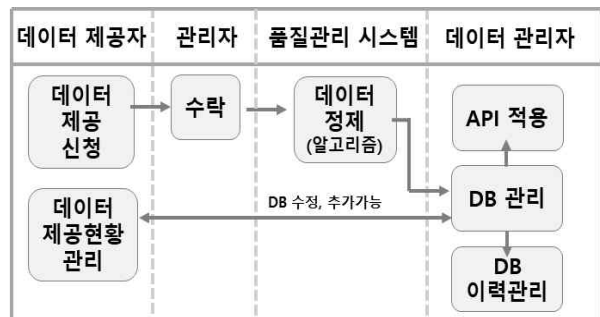


그림 5 데이터 제공 프로세스

Fig. 5 Data provision process

데이터 제공자는 본 시스템을 통해서 데이터 공개여부, API적용 여부 등을 선택하여 데이터 제공을 신청할 수 있다. 시스템 관리자는 신청내용을 확인하여 수락 할 수 있다. 데이터 관리자는 제공받은 데이터를 데이터 정제 알고리즘에 의해 데이터 오

타검사, DB화를 위한 띄어쓰기, 특수문자 검사 등을 실시한다. 정제작업이 끝난 데이터는 데이터베이스화하여 시스템에 적용하고, API적용을 신청한 데이터의 경우 API적용도 함께 진행한다. 또한 데이터 제공자는 데이터 제공현황 관리를 통해서 제공된 데이터를 직접 검색, 조회, 수정할 수 있다.

#### IV. 시스템 설계

본 시스템의 시스템구성도는 아래의 그림6과 같다. 먼저 공공 데이터 제공기관으로부터 데이터를 제공받아서 데이터 정제, API개발을 통해 시스템에 적용할 예정이다. 데이터사용자는 정리된 데이터를 한 번에 검색할 수 있고, API를 신청하여 활용할 수 있다. 데이터제공처에서 데이터를 개별로 관리할 수 있게 하여 따로 시스템을 제작하지 않고 본 시스템에서 연구데이터의 품질관리, 업데이트 관리 및 공유를 하도록 하여 시스템의 활용도를 높일 예정이다.

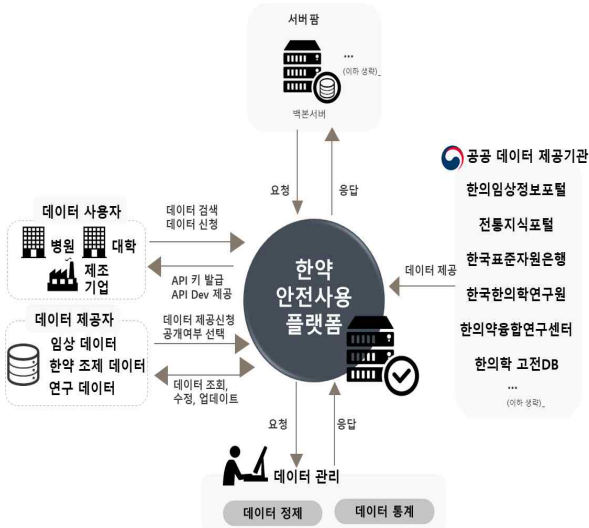


그림 6 시스템 구성도

Fig. 6 System configuration diagram

표 1은 여러 한방데이터 검색 시스템을 비교분석하고, 업계 관계자들과 인터뷰를 통해 도출된 요구사항을 요구사항 도출표로 작성하였고, 그림6은 Usecase Diagram으로 작성한 것이다.

표 1 요구사항 도출표

Table 1. Requirements Derivation Table

분류	기호	내용
데이터 수집	UC-001	데이터 제공 신청관리
	UC-002	데이터 신청 현황관리
	UC-003	데이터 제공처 관리
데이터 품질관리	UC-004	데이터 정제 관리
	UC-005	데이터 수정관리
	UC-006	데이터 검색통계 관리
검색	UC-007	데이터 검색관리
	UC-008	데이터 즐겨찾기 관리
	UC-009	데이터 수정요청 관리
	UC-010	수정요청 현황 관리
OpenAPI	UC-011	API 소개 관리
	UC-012	API Dev가이드 관리
	UC-013	API key 발급 신청관리
	UC-014	API key 발급 현황관리
회원관리	UC-015	회원권한 관리
	UC-016	로그인 관리
	UC-017	로그아웃 관리
	UC-018	회원가입 관리
	UC-019	회원 정보 관리
	UC-020	회원 탈퇴관리
	UC-021	회원 현황 관리
공지관리	UC-022	공지사항 관리
	UC-023	FAQ 게시판 관리
	UC-024	Dev 게시판 관리

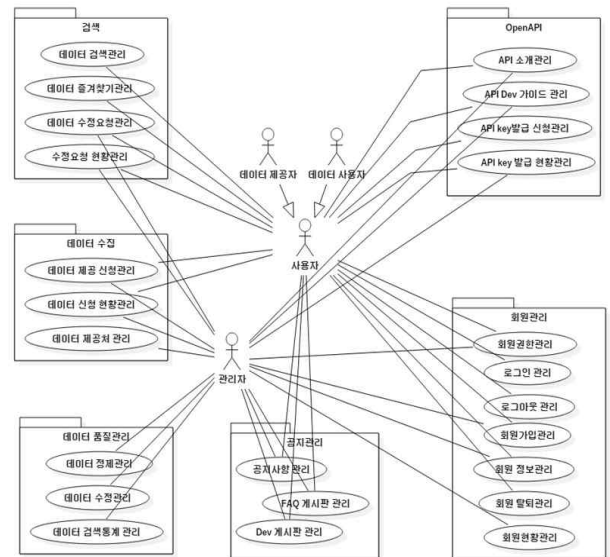


그림 7. 유스케이스 다이어그램

Fig. 7. Usecase Diagram



## V. 결 론

본 논문을 통해 기존에 혼재되어서 개별로 관리하는 한방 데이터의 접근성 및 연구 활용성을 높이고, 데이터 제공처에서 직접 시스템을 제작하지 않고도 데이터를 시각화(시스템)화하여 직접 관리할 수 있게 하여 데이터의 지속적인 관리와 데이터 품질을 향상시킬 수 있는 시스템을 제안하였다. 본 시스템은 한의약의 연구데이터뿐만 아니라 데이터 관리가 필요한 업종에서 활용될 수 있다고 생각한다.

[14] 한국생명공학연구원, <https://www.kribb.re.kr>,  
October 2024

[15] 공공데이터포털, <https://www.data.go.kr>,  
October 2024

[16] 한의약연구원, “한의약산업실태조사”, December  
2021

## 참 고 문 헌

- [1] 한국한의학진흥원, [https://nikom.or.kr/koms/module/komsExtra/index.do?menu\\_idx=21](https://nikom.or.kr/koms/module/komsExtra/index.do?menu_idx=21) , October 2024
- [2] 한의약소재은행, "Three-dimensional shape <https://nikom.or.kr/kmmb/index.do>, October 2024
- [3] 국가한의학정보포털, <https://nicom.or.kr/nckm/index.do>, October 2024
- [4] 한국전통지식포털,<https://koreantk.com/ktkp2014/>,  
October 2024
- [5] 오아시스, <https://oasis.kiom.re.kr/index.jsp>, October  
2024
- [6] 한약표준자료은행, <https://oasis.kiom.re.kr/herblib/>,  
October 2024
- [7] 한의약융합연구정보센터,<https://www.kmcric.com/>  
, October 2024
- [8] 한국한의학연구원,<https://kiom.re.kr/main/> ,  
October 2024
- [9] 한의학고전DB, <https://mediclassics.kr/>, October  
2024
- [10] 한의임상정보서비스, <https://www.kmpedia.kr/>,  
October 2024
- [11] 한의임상정보은행, <https://www.kmpedia.kr/>,  
October 2024
- [12] 한의시맨틱검색, <https://cis.kiom.re.kr/semantic/>,  
October 2024
- [13] 한의온라인정책서비스,<https://policy.kiom.re.kr/>,  
October 2024

# 보행 주기 특징 추출보행 주기 특징 추출

김유민\*, 최지우\*\*, 강태원\*, 최상일\*

## Gait cycle feature extraction

Yumin Kim\*, Jiwoo Choi\*\*, Taewon Kang\*, and Sangil Choi\*

### 요 약

본 논문은 LSTM 모델을 통해 추출한 보행 주기의 특징 벡터가 각 개인의 보행 고유성을 효과적으로 나타낼 수 있는지와 이를 통해 개인 식별이 가능한지를 연구한다. 연구 결과, 동일인의 보행 주기에서 추출된 특징 벡터는 일관된 상관관계를 보였으며, 반면 서로 다른 보행자 간의 특징 벡터는 상관 계수가 다르게 나타나 각 보행자에게 고유한 보행 양상이 존재함을 확인하였다. 이러한 결과는 본 연구에서 추출한 보행 특징 벡터가 개인별 고유성을 잘 반영하여 개인 식별에 유용하게 활용될 가능성을 시사한다.

### Abstract

This paper investigates whether feature vectors extracted from gait cycles using an LSTM model can effectively represent individual gait uniqueness and enable personal identification. The results show that feature vectors extracted from the same individual's gait cycles reveal consistent correlations. In contrast, those from different individuals show varying correlation coefficients, confirming unique gait characteristics for each person. These results present that the gait feature vectors extracted in this study accurately capture individual uniqueness, highlighting their potential applicability in personal identification.

### Key words

Deep Learning, Correlation, LSTM, Gait-Cycle

## I. 서 론

보행(Gait)은 행동 생체인식 기술의 하나로, 신체의 외형적 특징, 걷는 동작의 역학 등을 포함하는 걷는 방식으로 정의된다[1]. 보행은 의학, 보안, 스포츠 등 다양한 분야에서 주목받고 있다. 기존의 연구 중에서는 합성곱 신경망(CNN)을 이용하여 관성 데이터를 기반으로 개인을 식별할 수 있음을 확

인한 연구 결과가 있다[2].

본 연구에서는 이전 연구 [3]에서 관성 데이터로 얻은 보행 주기에서 특징(feature)을 추출하고, 추출한 특징이 개인을 식별하는데 유의미한 지표로 사용될 수 있는지를 검증하고자 한다.

보행 주기란 한 발의 발꿈치가 지면에 닿는 순간(HS: Heel Strike)부터 같은 발의 발가락이 지면에서 완전히 떨어지는 순간(TO: Toe Off)을 거쳐, 다시

\*강릉원주대학교, 20211315@gwnu.ac.kr (교신저자), twkang@gwnu.ac.kr, schoi@gwnu.ac.kr

\*\*인터포, cephular@gmail.com

같은 발의 HS가 발생할 때까지의 일련의 움직임은 의미한다[2]. 선행 연구[3]에서는 LSTM을 통해 관성 센서 데이터만을 이용하여 왼발 발꿈치 닿기(LHS), 왼발 발가락 떼기(LTO), 오른발 발꿈치 닿기(RHS), 오른발 발가락 떼기(RTO)의 각 순간으로 정의한 연구 결과가 있다[3].

보행 주기에서 추출한 특징이 유의하다면, 개인 식별, 보행 유형 분류 더 나아가 적절한 사용자 인증, 범죄자 추적, 의료 데이터 분석 등의 분야에 기여할 수 있을 것으로 기대된다.

## II. 보행 주기 습득

보행 데이터의 특징을 찾아내기 위해 이전 연구 [3]에서 수집한 관성 데이터를 사용하였다. 관성 데이터는 shimmer IMU 3를 통해 수집한 데이터를 선행보간법을 이용해 20ms 단위로 맞춘 관성 데이터를 전처리 후 사용하였다. 그림 1과 같이 전처리 과정을 끝낸 데이터를 LSTM 모델의 input data로 사용하였으며, output data로써 20ms당 보행 단계를 출력했다. 보행 단계란 LHS(class 1), LTO(class 2), RHS(class 3), RTO(class 4)의 각 순간을 의미한다.

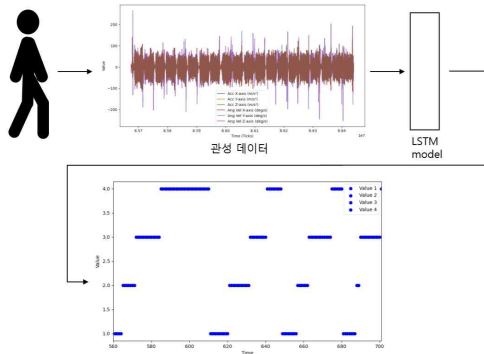


그림 1. LSTM 모델을 통한 보행 주기 결과  
Fig. 1. Gait cycle results with LSTM model

## III. 특징 추출

### 1) 특징 벡터(feature vector) 추출

본 연구는 위의 보행 단계를 바탕으로 그림2와 같은 보행 특징 15가지를 시간 단위를 기준으로 추

출하였다.

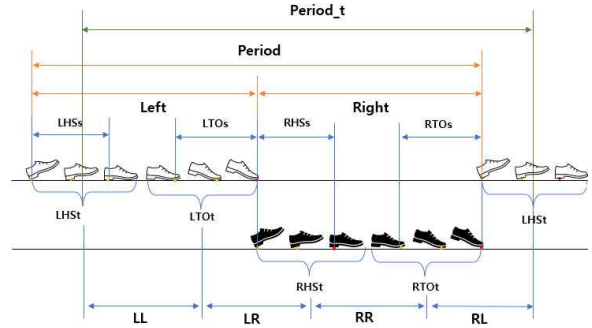


그림 2. 한 번의 보행 주기 내 특징 벡터 시각화  
Fig. 2.. Visualization of feature vectors in a single Gait cycle

한 사람의 전체 보행 주기에 대한 특징 벡터를 추출하면 식1과 같은 벡터를 얻을 수 있다.

$$X = \begin{bmatrix} Period_1 & Periodt_1 & \dots & RR_1 & RL_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Period_n & Periodt_n & \dots & RR_n & RL_n \end{bmatrix} \quad (식1)$$

추출한 보행 특징에 대한 자세한 설명은 표 1에 정리되어있다.

표 1. 특징 벡터 설명

Table 1. Description of Feature vector

name	설명
Period	LTO의 시작 시점부터 RTO의 종료 시점까지의 시간
Periodt	현재 보행 주기의 LTOt 시작 시점부터 다음 보행 주기의 LTOt 시작 전까지의 시간
LEFT	LTO 지속 시간과 LHS 지속 시간의 합
RIGHT	RTO 지속 시간과 RHS 지속 시간의 합
LHSs	LHS 지속 시간
LTOs	LTO 지속 시간
RHSs	RHS 지속 시간
RTOs	RTO 지속 시간
LHSt	LHS 구간 내 최고 확률 지점까지 소요시간
LTOt	LTO 구간 내 최고 확률 지점까지 소요시간
RHSt	RHS 구간 내 최고 확률 지점까지 소요시간
RTOt	RTO 구간 내 최고 확률 지점까지 소요시간
LL	LHSt부터 LTOt까지 걸린 시간
LR	LTOt부터 RHSt까지 걸린 시간
RR	RHSt부터 RTOt까지 걸린 시간
RL	RTOt부터 다음 보행 주기의 LHSt까지 걸린 시간

2) 특징 벡터의 상관관계 분석

선정한 특징 벡터가 보행 주기와 관련이 있는지, 각 특징 벡터가 유의미한 관계가 있는지 확인하기 위해 선정한 특징 벡터 간의 상관관계를 확인하였다. 동일인의 걸음걸이는 일관된 특성을 보이기 때문에, 동일인의 보행 데이터에 대한 특징 벡터의 상관관계는 높은 상관 계수를 나타낼 것으로 예상할 수 있다. 그림 4에서 확인할 수 있듯이, 본 연구에서 추출한 특징 벡터들이 동일인의 데이터에서 상관 계수가 최소 0.7, 최대 1.0으로 강한 상관관계를 가짐을 확인할 수 있다.

그림 4는 남녀 각 9명, 총 18명의 참가자의 보행 주기에서 추출한 특징 벡터 간 상관관계를 분석한 결과를 보여준다. 분석 결과, 상관 계수는 최대 0.95에서 최소 약 0.1까지의 범위를 가졌다. 이는 상관 계수가 높은 17번-12번, 12번-8번, 17번-8번, 16번-6번 간에는 유사한 보행 주기가 있음을 의미하며, 반대로 상관 계수가 낮은 7번-2번, 15번-2번, 16번-7번, 16번-15번 간에는 서로 매우 다른 보행 주기가 있음을 나타낸다.

V. 결론 및 향후 과제

본 연구에서는 LSTM 모델을 통해 추출한 보행 주기의 특징 벡터가 동일인의 데이터에서는 높은 상관관계를 보이는 반면, 서로 다른 보행자 간에는 상관관계의 차이를 보임을 확인하였다. 이는 본 연구에서 추출한 특징 벡터가 동일인 데이터에서는 일관된 양상을, 타인 데이터에서는 고유한 차이를 나타낸다는 점을 입증하며, 이러한 특징 벡터가 개인 식별에 유용한 지표로 활용될 수 있음을 시사한다. 또한, 본 연구 결과는 보행 패턴 분석을 통한 개인 식별뿐만 아니라, 향후 동적 환경에서의 보행 패턴 분석 및 실제 보행 인식 모델과의 비교 연구를 통해 보행 인식 기술의 정확성과 신뢰성을 향상시키기 위한 방향을 제시한다.

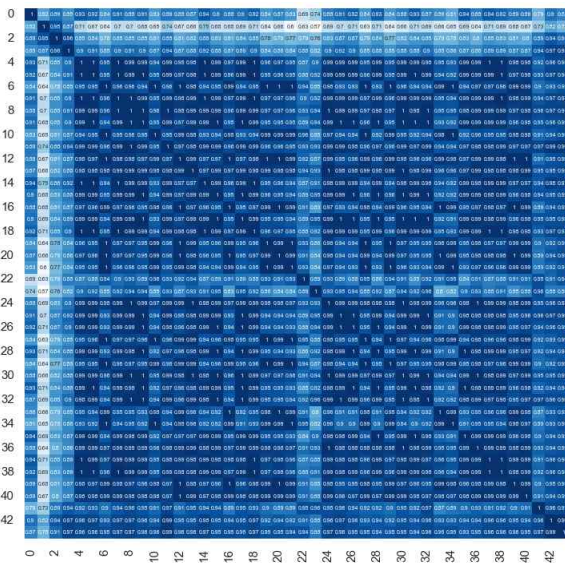


그림 4. 동일인의 특징 벡터 간 상관관계  
Fig. 4. Correlation between feature vectors of the same person

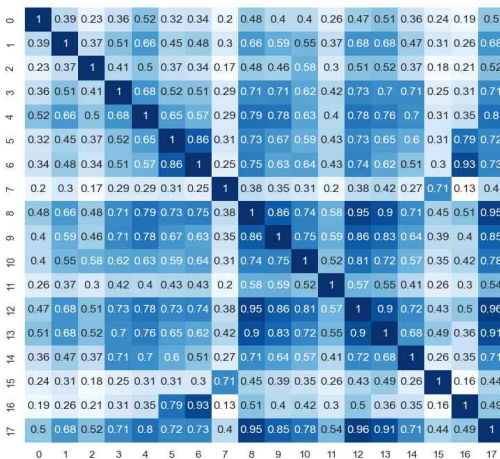


그림 5. 각 보행자 간의 상관관계  
Fig 5. Correlation between each pedestrian

참고 문헌

[1] A. Kharb, V. Saini, Y.K. Jain, and S. Dhiman, "A REVIEW OF GATE CYCLE AND ITS PARAMETERS", IJCEM, vol. 13, pp. 78, July 2011.

[2] 최지우, 최상일, 강태원, "보행주기를 이용한 개인식별 CNN 모델", JKIIIT, vol.20, No. 11, pp. 127-136, 2022.

[3] 최지우, 유형진, 최상일, 강태원, "LSTM을 사용한 보행주기 식별", JKIIIT, vol. 19, No. 8, pp. 27-39, 2021.

# 키워드 네트워크 분석 기반 감정인식 AI 분야 연구동향 분석

한유리\*

## Research Trends in Emotion Recognition AI: A Keyword Network Analysis Approach

Yuri Han\*

감정인식 AI 기술은 인간의 감정과 정서를 기계가 자동으로 인식하고 해석하는 기술로, 인공지능과 감정 분석이 융합된 첨단분야로 자리잡고 있다[1]. 감정인식 AI 시장 규모는 2022년 263억 달러 규모로 평가되었고, 2029년 748억 달러(CAGR 16.1%)에 이를 것으로 예상되는 만큼 다양한 산업에서 수요가 증가하고 있다 [2]. 또한 인공지능 기술의 급속 발전으로 그 가능성이 더욱 확장됨에 따라[3], 중점 연구주제를 파악하고 기술 응용범위와 글로벌 협력관계를 넓힐 수 있도록 최신 연구동향에 대한 이해가 필요하다. 이러한 배경에서 본 연구는 2000년대 이후 감정인식 AI 분야의 SCI급 논문 13,207건(Web of Science)의 서지정보를 활용하여 연구 발전추이와 연구동향을 분석하였고, 네트워크 분석을 통해 글로벌 협력관계와 시기별 주요 연구주제 및 응용분야를 고찰하였다. 연구 결과, 논문 성장 추이로는 2018년 트랜스포머 아키텍처 등장 이후 GPT 등 대규모 언어모델 출시 영향을 받아 최근 5년 이내 논문이 전체 비중의 83%를 차지하며 급속도로 발전해 온 것이 두드러진다. 국제협력 선도국가는 미국, 인도, 중국이 가장 영향력이 큰 것으로 나타났고, 한국은 글로벌 네트워크에서 높은 접근성과 여러 국가와의 다중 연결을 바탕으로 중요한 정보 교류 및 협력 위치에 있지만, 중개자로서의 역할은 상대적으로 제한적인 것으로 해석되었다. 따라서 다자간 협력보다는 개별 국가와의 양자 협력 강화를 통해 국제적 영향력을 넓혀가는 전략을 취해야 함이 적절하다. 연구분야 주요 시사점으로는 기존 텍스트 기반 자연어 처리 중심에서 최근 다양한 형태의 데이터(음성, 얼굴, 생체신호 등)를 기반으로 감정을 파악하는 Multi-Modal 접근법이 성장하고 있으며, 응용 분야로 챗봇, 정신건강, 헬스케어, 고객경험, 로봇, 주식, 관광산업 등 다양한 영역에서 응용연구가 수행되고 있는 것으로 파악되었다. 본 연구는 감정인식 AI 기술 발전 현황과 글로벌 협력 구조에 대한 실증 데이터를 제공함으로써 관련 연구자와 산업계의 지속적인 연구와 협력 전략 수립에 유용한 정보로 활용되길 기대한다.

### Key words

Emotion Recognition AI, Research Trend, Global Collaboration, Network Analysis, Keyword Analysis

### 참고문헌

- [1] Kalateh, S., Estrada-Jimenez, L. A., Nikghadam-Hojjati, S., & Barata, J. "A Systematic Review on Multimodal Emotion Recognition: Building Blocks, Current State, Applications, and Challenges." IEEE Access, vol. 12, pp. 103976-104019, 2024.
- [2] Fortune Business Insights. "Emotion Detection and Recognition Market Size." Fortune Business Insights, 2024.
- [3] Rai, M., & Pandey, J. K. "Using Machine Learning to Detect Emotions and Predict Human Psychology." IGI Global, 2024.

\*한국과학기술정보연구원 연구전략센터, yurihan@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: KISTI 기본사업 K24L5M1C1-01)

# AI 어플리케이션 지원을 위한 GPU IO 가속 연구 및 성능 분석

윤준원\*, 안도식\*\*, 흥태영\*\*\*

## GPU IO Acceleration Research and Performance Analysis to Support AI Applications

Junweon Yoon\*, Do-Sik An\*\*, and Taeyoung Hong\*\*\*

### 요 약

AI의 딥러닝 모델 학습은 수많은 데이터셋을 빠르게 GPU로 전송해야 하는 데이터 집중적 I/O 성능 향상이 중요하다. GDS는 스토리지에서 GPU 메모리로 데이터를 직접 전송하는 방법으로 CPU 메모리와 버스를 거치지 않아 데이터 전송이 속도를 가속하고 지연 시간을 줄일 수 있다.

본 논문에서는 GDS 성능을 측정하기 위한 벤치마크 도구를 사용하여 GDS 사용으로 인한 대역폭, 지연시간을 평가하고 CPU 리소스의 사용량을 확인한다. AI 어플리케이션 수행시 GDS는 데이터 병목 현상을 완화하여 전반적인 시스템 성능을 개선한다.

### Abstract

The performance of deep learning model training in AI relies heavily on the rapid transmission of large datasets to the GPU, making the enhancement of data-intensive I/O performance crucial. GPU Direct Storage (GDS) enables direct data transfer from storage to GPU memory, bypassing CPU memory and buses, thereby accelerating data transmission speeds and reducing latency.

This paper evaluates the performance of GDS using benchmark tools to assess bandwidth, latency, and CPU resource usage. During the execution of AI applications, GDS alleviates data bottlenecks and improves overall system performance.

### Key words

AI, application, GPU, GDS, data processing, parallel filesystem

---

\*한국과학기술정보연구원 슈퍼컴퓨팅인프라센터, jwyoona@kisti.re.kr, \*\*dsan@kisti.re.kr, \*\*\*tyhong@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원의 기본사업(과제명: 국가 플래그십 초고성능컴퓨터 인프라 구축 및 서비스, 과제번호: K24L2M1C1)으로 수행된 연구입니다.

## I. 서론

AI와 머신러닝 모델 학습, 특히 딥러닝 모델에서는 대규모 데이터셋을 GPU로 빠르게 전송해야 하는 요구사항이 증가하고 있다. 가령, 이미지나 텍스트 데이터 같은 방대한 양의 데이터를 실시간으로 GPU에 전달할 수 있어야만 효율적인 학습이 가능하다[1]. 기존의 데이터 전송 방식은 데이터가 저장 장치에서 CPU 메모리, GPU 메모리의 순서로 전송되어 CPU와 시스템 메모리가 병목 현상 발생하였다. 데이터 전송 속도가 모델의 처리 성능을 따라가지 못하는 경우 성능이 크게 저하되는 문제를 발생한다. 더욱이 SSD, NVMe와 같은 고성능 스토리지 장치들이 보편화되면서 데이터 전송 속도와 대역폭이 크게 증가했지만 GPU로 데이터를 전송하기 위해 CPU와 메모리를 거쳐야 하는 기존의 환경에서는 스토리지의 성능을 제대로 활용하기 어렵다. GPU Direct Storage(GDS)는 이러한 고속 스토리지 장치를 GPU와 직접 연결해 데이터를 빠르게 처리할 수 있게 하여 GPU와 스토리지 간의 병목 현상을 해소하고 고속 데이터 전송을 가능하게 한다[2]. 기존 시스템에서는 데이터 전송 경로가 복잡하여 데이터 병목 현상이 발생하고 특히 여러 GPU를 사용하는 환경에서는 더 큰 부하가 발생된다. 따라서 GDS는 GPU의 효율적 데이터 전송을 가능하게 하여 AI와 HPC 분야의 필수적인 기술로 자리 잡고 있다.

본 연구에서는 GDS 성능 벤치마크 도구를 사용하여 대역폭, 지연시간을 평가하고 CPU 리소스의 상태를 모니터링하여 성능 개선됨을 확인한다.

## II. 관련 연구

GPU Direct Storage(GDS)에는 효율적인 데이터 전송을 위해 필요한 세부 기술들이 포함된다. GDS는 주로 CUDA 환경에서 스토리지와 GPU 메모리(HBM) 간의 직접데이터 전송을 위해 RDMA(Remote Direct Memory Access) 기술을 사용한다. 일반적으로 CPU의 개입 없이 스토리지와 GPU 메모리 간 데이터를 직접 전송할 수 있다.

GDS는 CPU 메모리를 통과하지 않기 때문에 Direct Memory Access(DMA) 매핑이 필요하다. GPU 메모리와 스토리지 메모리 간에 매핑이 되면, 각 데이터 전송은 시스템 메모리를 거치지 않고 메모리 주소를 통해 직접 접근할 수 있다. 데이터 버퍼를 설정하고 pinning하여 GPU가 직접 접근 가능하도록 메모리 할당을 고정한다. <표 1>과 같이 메모리 할당 API를 사용하게 되는데 cudaHostRegister는 메모리를 고정(pinning)하여 GPU가 효율적으로 접근할 수 있게 하고 cudaHostGetDevicePointer는 버퍼 포인터를 GPU 메모리로 매핑하여 GPU에서 직접 접근할 수 있도록 한다[3].

표 1. GPU 메모리 직접 접근을 위한 GDS API  
Table 1. GDS API for direct GPU memory access

```
cudaHostRegister(buffer, bufferSize, cudaHostRegisterMapped);
cudaHostGetDevicePointer(&devicePtr, buffer, 0);
```

GDS는 비동기 I/O 처리를 지원하여 데이터 전송 중에도 GPU가 다른 작업을 수행할 수 있는데 cudaStream이나 nvmeQueue와 같은 비동기 처리 스트림을 사용하여 입출력 성능을 최적화할 수 있다.

## III. 성능 평가

CUUDA 툴킷에 포함된 GPU 프로그래밍 도구 중 하나인 GDSIO(GPU Data Streaming Input/Output)를 이용하여 성능을 검증하였다. GDSIO의 모드(-x 옵션) 선택을 통해 GPU 메모리에 직접 접근 또는 CPU를 통해 접근하는 성능을 비교하여 측정한다. <표 2>는 GDSIO 수행과 관련된 옵션을 보여준다.

표 2. GDSIO를 이용한 성능검증  
Table 2. Performance validation with GDSIO

```
$ /cuda/12.0/gds/tools/gdsio -T 120 -D directory/1 -d 0 -n 0 -w
64 -s 512M -i $iosize -x 0 -l 1

#mode=(0,2) xfer_type(0:Storage->GPU, 2:Storage->CPU->GPU)
#op=(0,2) -l <(read) 0/(write)1/ (randread) 2/ (randwrite) 3>
#d=0 -d <gpu_index (refer nvidia-smi)>
#T=120 -T <duration in seconds>
```

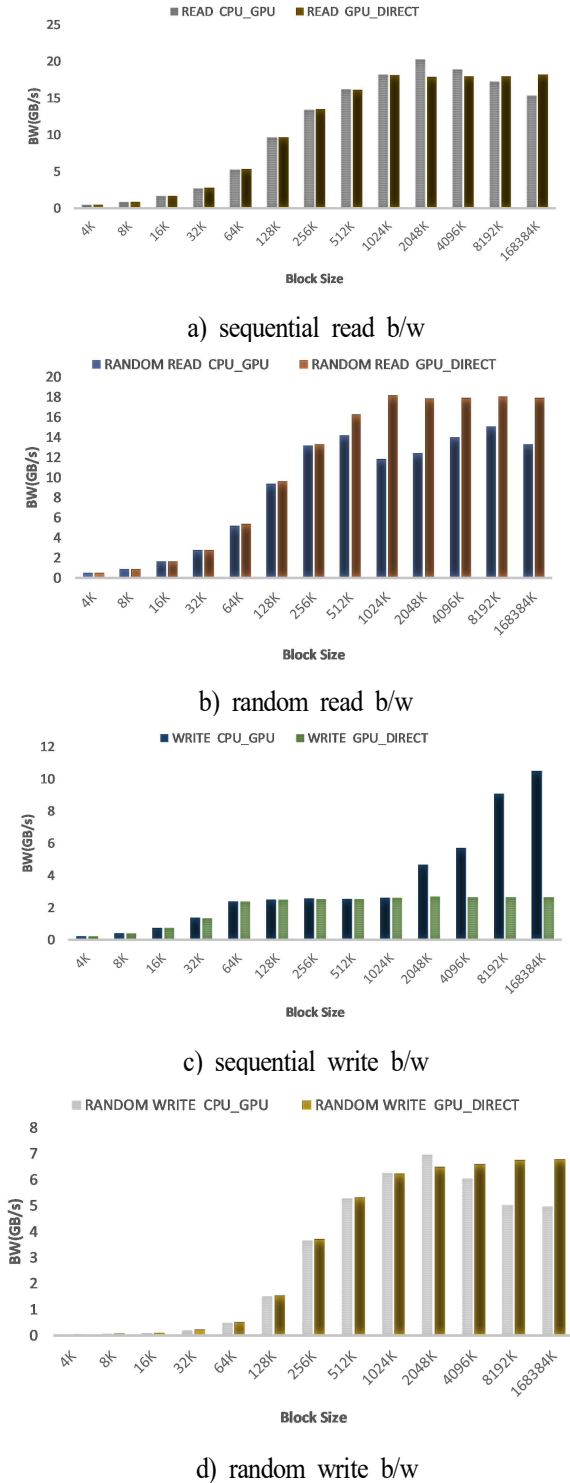


그림 1. GPU Direct IO와 GPU IO의 성능 비교  
 Fig. 1. GPU Direct IO vs. GPU IO Performance Comparison

(그림 1)은 GDSIO 벤치마크 도구를 이용하여 GPU Direct IO와 CPU IO의 성능을 IO 크기를 (4K~16M) 변경하며 비교한 그래프이다. GPU Direct

IO의 성능이 sequential/random, read/write에서 모두 향상된 성능을 얻었고 IO 크기가 클수록 일반적으로 성능이 차이가 더 나는 것을 확인할 수 있다.

#### IV. 결론

GDSIO는 스토리지와 GPU 간의 고속 데이터 전송을 가능하게 하며, CPU와 시스템 메모리를 우회하여 지연 시간을 줄이고 처리량을 증가시킨다. 또한 복수개의 GPU 작업을 지원하여 분산 컴퓨팅 환경에 적합하다. 따라서 AI의 머신 러닝, 딥 러닝, 대규모 데이터 처리와 같은 높은 데이터 처리량이 필요한 애플리케이션 수행에 있어 탁월한 성능 이점을 제공할 수 있다.

#### 참고 문헌

[1] M. Alian, Y. Yuan, J. Zhang, R. Wang, M. Jung, and N. S. Kim, "Data direct I/O characterization for future I/O system exploration", In 2020 IEEE International Symposium on Performance Analysis of Systems and Software(ISPASS), pp. 160-169, Aug 2020.

[2] A. Venkatesh, K. Hamidouche, S. Potluri, D. Rosetti, C. H. Chu, and D. K. Panda, "MPI-GDS: High performance MPI designs with GPUDirect-aSync for CPU-GPU control flow decoupling.", 46th International Conference on Parallel Processing(ICPP). IEEE, 2017.

[3] NVIDIA GPU Direct Storage. (<https://docs.nvidia.com/gpudirect-storage/index.html>).



# LLM 기반 한국어 추론적 읽기 문제 자동 생성 시스템 성능 평가

임성원\*, 류범모\*\*

## Performance Evaluation of an LLM-Based Automated Inferential Reading Question Generation System for Korean

Sungwon Lim\*, Pummo Ryu\*\*

본 연구는 한국어능력시험을 기반으로 한국어 학습자의 난이도에 맞춘 추론적 읽기 문제 자동 생성 시스템을 개발하고 그 성능을 평가하는 데 목적이 있다. 대형 언어 모델(LLM)을 활용하여 학습자 등급에 맞춘 문제를 자동 생성하고, 시스템의 유용성을 확인하기 위해 자동 평가, 학습자 평가, 전문가 평가[1]를 진행하였다. 문제 자동 생성 시스템을 구축하기 위하여 학습자 등급별 어휘, 문법 목록[2], 주제 목록을 구축하였고 실제 한국어능력 시험 추론적 읽기 문제를 바탕으로 사전 데이터를 구축하였다. 이를 통해 각 문제의 지문, 질문, 답변 선택지를 학습자 수준에 맞게 조정하였다. 그리고 LLM 중 GPT-4 모델을 사용하여 프롬프트를 제공하고 자동으로 문제를 생성하였다[3].

200개의 자동 생성 문제를 바탕으로 진행된 평가는 세 단계로 이루어졌다. 첫 번째로 자동 평가는 GPT-4를 이용하여 생성된 문제의 난이도별 어휘와 문법 적절성, 지문과 질문 간 일관성을 가졌는지에 대해 평가하였다. 그 결과, 중급 문제에서는 높은 성능을 보였으나 초급 문제에서는 중급 어휘와 문법이 자주 사용되는 경향이 나타났다. 두 번째로 학습자 평가는 실제 한국어능력시험 5급과 6급을 가진 학습자를 대상으로 진행하였고 문제 난이도의 적절성과 실제성을 평가하였다. 그 결과, 일부 초급 문제에서 중급 수준에 해당하는 어휘와 문법이 발견되었지만, 중급 문제와 실제성은 높은 결과로 나타났다. 마지막으로 전문가 평가는 현직 한국어 교육 10년 이상의 경력을 가진 전문가들이 평가를 진행하였고 평가 항목은 문제 난이도의 적절성, 문제의 일관성, 추론 문제의 적절성, 실제성으로 구성하였다. 그 결과, 초급 문제에서 어휘와 문법 난이도 조정, 선택지 복잡성이 적절하지 않다는 평가가 있었지만, 중급 문제에서는 적절한 난이도와 일관된 구성으로 높은 평가를 받았다. 세 단계의 평가를 통해 한국어 학습자 대상의 자동 문제 생성 시스템의 유용성을 확인하였고 향후 연구에서는 평가 결과를 바탕으로 학습자 난이도 조정을 위한 데이터의 양과 질을 향상하고, 세밀한 문제 생성이 가능하도록 프롬프트를 개선할 계획이다.

### Key words

Korean Language Learning Question Generation, LLM, Automated Question Generation

### 참고문헌

- [1] Aldabe, Itziar, et al. "Arikiturri: an automatic question generator based on corpora and nlp techniques." Intelligent Tutoring Systems: 8th International Conference, ITS 2006, Jhongli, Taiwan, June 26-30, 2006. Proceedings 8. Springer Berlin Heidelberg, 2006.
- [2] TOPIK, 학습자등급별어휘및문법목록, 한국어능력시험, 2024, <https://www.topik.go.kr/TWINFO/TWINFO0020.do>
- [3] 허동석외3, "프롬프트 개발을 통한 수능 국어 맞춤형 문제 생성 시스템 제안", 한국HCI학회 학술대회, 2024

\*부산외국어대학교 인공지능융합학과, wonirosoida33@gmail.com,

\*\*부산외국어대학교 인공지능융합학과, pmryu@bufs.ac.kr(교신저자)

# 오픈소스와 MSA 패턴을 적용한 클라우드 네이티브 환경 구축 방안

이준호\*<sup>1</sup>, 서주형\*, 고상원\*\*

## How to build a cloud-native environment with open source and MSA patterns

Lee Junho\*<sup>1</sup>, Seo Juhyung\*, and Koh Sangwon\*\*

### 요 약

디지털 전환이 가속됨에 따라 정부는 26년 신규 시스템에 대하여 클라우드 네이티브 환경을 70%까지 적용하기로 발표하였다. MSP 및 CSP를 통해 클라우드 네이티브 설계·구축 및 서비스 개발이 가능하지만, 자동화되는 클라우드 인프라 활용 및 SaaS 개발·운영을 위해서는 MSP 및 CSP에서 제공하는 기술에 대한 이해가 필수이다. 본 논문에서는 오픈소스 활용 및 MSA 패턴을 적용한 클라우드 네이티브 환경 구축 방안 및 MSA 시스템 구축 시 발생할 수 있는 주요 문제점과 대처 방안을 제시한다.

### Abstract

As digital transformation accelerates, the government announced that it will achieve up to 70% of cloud-native environments for new systems by 2026. Cloud-native construction and service development are possible through MSP and CSP, but understanding the technologies provided by MSP and CSP is essential to utilize automated cloud infrastructure and develop/operate SaaS. This paper presents a plan for building a cloud-native environment using open source and applying the MSA pattern, as well as major problems that may arise when building an MSA system and ways to deal with them.

### Key words

cloud-native, MSA, open source, SaaS

## 1. 서 론

디지털 전환이 가속됨에 따라 정부는 21개 행정·공공기관 정보시스템을 선정해 25년까지 클라우드 네이티브 전환을 추진하기로 결정하였다.

또한 2026년에는 신규 시스템에 대한 클라우드

네이티브 적용을 70%까지 달성한다는 것이 목표이다[1]. 이처럼 민간 및 정부에서 구축하고 있는 클라우드 네이티브는 서버나 저장소 등의 자원을 필요한 만큼 활용할 수 있는 클라우드 인프라의 장점을 극대화하는 정보시스템 구축 방식이다.

행안부에서 시행될 10개 기관 13개 시스템에 대한 클라우드 네이티브 전환이 완료될 시 최대 평균

\*한전KDN, junho\_1121@kdn.com<sup>1</sup>, \*한전KDN, juhyung\_0828@kdn.com

\*\*한전KDN, swkoh\_1995@kdn.com

시스템 중단시간 95% 감축, 서비스 요청처리 시간 26% 단축, 이용자 폭증 때 용량을 4.5배로 자동으로 확장시키는 등 이용자들이 서비스를 안정적으로 사용할 수 있을 것으로 예상된다[2].

이러한 클라우드 네이티브를 구축하기 위해서는 MSP(Managed Service Provider) 및 CSP(Cloud Service Provider)를 통한 설계, 구축 및 서비스 개발이 가능하다. 하지만 자동화되는 클라우드 인프라 활용 및 SaaS(Software as a Service) 개발, 운영을 위해서는 MSP 및 CSP에서 제공하는 클라우드 네이티브 기능 및 기술에 대한 이해가 필수이다. 본 논문에서는 Production 환경 및 PoC(Proof of Concept) 용으로 도입이 가능한 오픈소스 및 이를 활용한 클라우드 네이티브 환경 구축 방안을 제시한다.

## II. 클라우드 네이티브와 MSA 패턴을 적용한 시스템 구성

### 1. 클라우드 네이티브와 CNCF(Cloud Native Computing Foundation)

클라우드 네이티브란 IaaS(Infra as a Service) 환경 위에서 동작하는 PaaS(Platform as a Service)를 기반으로, MSA(Micro Service Architecture)로 설계된 어플리케이션들이 DevOps, CI/CD 등의 문화·기술 등과 융합되어 운영되는 환경이다.

위와 같은 클라우드 네이티브를 지원하고 도입을 가속화하기 위하여, 기술 정의 및 오픈소스를 관리하는 단체 CNCF가 발족되었다.

CNCF는 클라우드 네이티브의 주요 4가지 구성요소인 컨테이너, MSA, DevOps, CI/CD 기술과 관련된 기술 및 오픈소스를 지원하고 있다.

### 2. 오픈소스와 MSA 패턴을 적용한 시스템 구성

MSA 시스템은 하나의 유닛으로 구성된 Monolithic 구조에서 몇 개의 서비스 단위로 나누어서 시스템을 구성하는 SOA(Service Oriented Architecture)를 거쳐 일련의 서비스를 더욱 세밀하게 나누고 분리(Micro Service)하여 트랜잭션을 처리

하는 형태로 발전하였다.

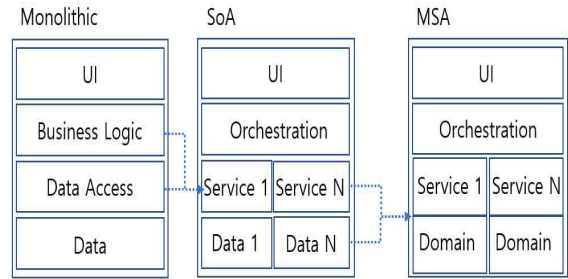


그림 1. Monolithic, SoA 및 MSA 아키텍처 간 비교

Monolithic 아키텍처가 하나의 유닛에 UI, 비즈니스 로직, 데이터 컴포넌트 및 데이터를 결합성이 높게 구성하는 시스템이라면, SOA 아키텍처의 경우 공통으로 사용되는 비즈니스 로직과 데이터 컴포넌트를 분리·독립시켜 반복적으로 동작시키는 형태이다. MSA 아키텍처는 도메인 바운더리를 정의하고 해당 영역에서만 비즈니스 로직을 처리하게 하며, 각 도메인 바운더리 및 컴포넌트, 데이터 영역을 느슨하게 결합하는 형태로 구성함으로써 시스템의 확장성 및 가용성을 극대화 할 수 있다[3].

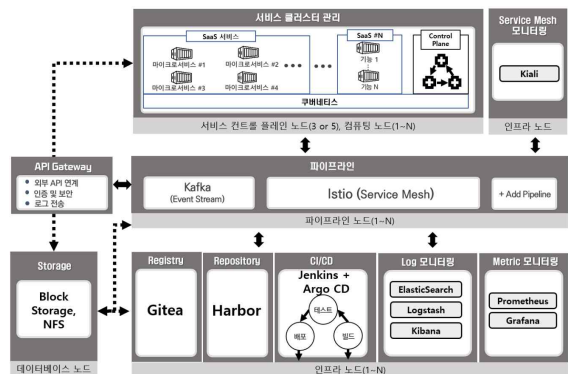


그림 2. 오픈소스로 구성된 클라우드 네이티브 환경

그림 2는 오픈소스만으로 구성된 클라우드 네이티브 환경이다. 관리 측면 및 시스템의 안정성을 위하여 클러스터의 정책을 관리하는 컨트롤 플레인 노드, 비즈니스 로직이 실행되는 컴퓨팅 노드, 컨테이너 삭제 시 데이터 유실의 위험을 방지하기 위한 외부 데이터베이스 노드, 이벤트 전달 및 마이크로서비스 간의 트래픽을 전송하고 제어하는 파이프라인 노드 및 형상관리, 컨테이너 이미지 관리, 분산

시스템 로그/메트릭 수집·모니터링을 담당하는 인프라 노드로 분리하여 구성하였다.

서비스 트래픽이 증가 시 컴퓨팅 노드는 자동으로 스케일링 되어 시스템의 가용성을 보장하며, 분산되어 있는 마이크로서비스들의 로그와 컨테이너의 메트릭 정보를 통합적으로 수집하고 모니터링을 함으로써 시스템의 안정성 및 서비스 품질을 제고할 수 있다.

또한, 서비스 개발 환경을 Registry, Repository 및 CI/CD 파이프라인과 연결함으로써 빠른 서비스 개발 및 배포가 가능해지며, Canary 배포 등을 통해 서비스의 중단 없이 업데이트가 가능해진다.

### 3. MSA 시스템에서의 트랜잭션 처리 및 데이터 동기화

MSA 시스템에서는 여러 개의 마이크로서비스 간 통신을 통해 1개의 트랜잭션을 처리하게 된다. 이와 같이 여러 마이크로서비스 간 통신 시의 완료성 및 통신 실패 시의 복원성을 보장하여야 한다.

또한, 마이크로서비스 마다 데이터베이스를 따로 구성하는 MSA 시스템의 경우 분산DB 간 데이터 동기화는 필수 불가결하다.

MSA 환경에서의 트랜잭션 처리 및 데이터베이스 동기화에는 3가지 방법이 있다.

서로 다른 DB간에 완벽한 정합성을 보장하는 2PC(Two Phase Commit), 보상 트랜잭션을 활용하여 최종적으로 일관성을 보장하는 CDC(Change Data Capture), 어플리케이션 개발자가 보상 트랜잭션을 정의하는 SAGA 패턴이 있다. 2PC 및 CDC는 시스템 소프트웨어가 담당하며, SAGA 패턴의 경우 어플리케이션 레벨에서 수행되고 MSA 사상에 가장 부합하는 방식이다.

시스템의 목적에 따라 강한 정합성이 필요한 경우에는 2PC, 반드시 실행되어야 하지만 개발자가 보상패턴을 정의하기 어려울 경우에는 CDC(Kafka 및 Debezium로 구성 가능), 시스템의 효율성 고려 및 개발자가 보상 패턴을 정의 가능할 경우에는 SAGA 패턴(Istio 활용)을 적용하는 것이 추천된다 [4]-[8].

### 4. MSA 시스템 성능, 보안성 및 취약점 관리

MSA 시스템의 내구성, 회복력 및 성능을 테스트하기 위한 방법으로 카오스 엔지니어링 기법 적용을 고려할 수 있다. 수십, 수백, 수천개의 마이크로서비스가 얽힌 MSA 시스템 환경에서 1개의 마이크로 서비스에서 발생한 문제가 전체 서비스에 영향을 미치는지 테스트 해 볼 수 있는 접근 방식으로서 어플리케이션 레벨에서의 error 발생, 호스트 레벨에서의 Service·Process 강제 중단, 리소스 어택(CPU, memory, IO) 및 네트워크 어택(packet loss 등) 등을 통하여 테스트 가능하며 이러한 카오스 엔지니어링을 지원하는 오픈소스로는 Chaos Monkey, Chaos ToolKit, Litmus, Istio, Gremlin 등이 존재한다.

마지막으로 클라우드 네이티브 환경에서 잘 알려진 취약점은 3가지 요소가 존재한다. 서비스 트래픽에 따라 자동으로 스케일링을 시행하는 오케스트레이션 도구의 취약점, 비즈니스 로직을 동작하게 하는 소스코드의 취약점 및 컨테이너 환경에서 소스코드와 여러 라이브러리 등을 포함하여 실제 동작을 수행하는 컨테이너 이미지 취약점이다.

오케스트레이션 도구로 잘 알려진 쿠버네티스의 보안 이슈 및 취약점은 CIS Kubernetes Benchmark와 kube-bench를 통해 검토가 가능하다[9]-[10].

비즈니스 로직을 구현하는 소스코드의 취약점은 SAST(Static Application Security Testing) 방식의 SonarQube를 통해 검토가 가능하며, 이미지 취약점의 경우 DAST(Dynamic Application Security Testing) 방식의 Trivy 엔진을 통해 검토가 가능하다[11].

## III. 결 론

본 논문에서는 오픈소스와 MSA 패턴을 적용한 클라우드 네이티브 환경 구축 방안 및 MSA 시스템 구축 시 발생할 수 있는 문제와 대처 방안을 제시하였다. 이러한 기술적인 측면 외에 클라우드 네이티브 환경 및 DevOps 문화가 정착하기 위해서는 조직원들의 기술 레벨, 조직 문화 및 정책 확립 등이 중요하다. 특히, 조직 및 구성원 간의 수직적인 직급 체계가 확립되어 있는 조직에서의 클라우드

네이티브 전환 및 DevOps 정착에는 많은 어려움이 따른다.

따라서, 클라우드 네이티브 기술 도입/교육과 함께 조직 구조의 변화, 정책 수립 및 DevOps 문화 교육 등을 통한 인문학적 노력이 같이 수반될 시 클라우드 네이티브 환경에서의 MSA 시스템 구축 및 원활한 서비스 제공이 가능할 것으로 기대된다.

## 참 고 문 헌

- [1] <https://www.etnews.com/20240410000027>
- [2] <https://www.korea.kr/news/policyNewsView.do?newsId=148935449>
- [3] <https://cazton.com/consulting/enterprise/software-architecture>
- [4] [https://en.wikipedia.org/wiki/X/Open\\_XA](https://en.wikipedia.org/wiki/X/Open_XA)
- [5] [https://en.wikipedia.org/wiki/Data\\_integration](https://en.wikipedia.org/wiki/Data_integration)
- [6] <https://microservices.io/patterns/data/saga.html>
- [7] <https://www.vinsguru.com/choreography-saga-pattern-with-spring-boot>
- [8] <https://www.vinsguru.com/architectural-pattern-orchestration-saga-pattern-implementation-using-kafka>
- [9] <https://tigera.io/learn/guides/kubernetes-security/kubernetes-cis-benchmark>
- [10] <https://www.armosec.io/glossary/cis-kubernetes-benchmark>
- [11] <https://about.gitlab.com/topics/devsecops/sast-vs-dast>



# 정규화된 양자 전송을 통한 양자 연합학습 프레임워크

백 한 결\*

## Normalized Quantum Teleportation for Quantum Federated Learning Framework

Hankyul Baek\*

### 요약

연합학습(FL: Federated Learning)은 데이터 편향성을 없애고 동시에, 보안성이 우수한 분산학습 알고리즘으로써 주목받고 있다. 하지만, 매 통신 단계마다 파라미터를 송/수신 해야 하기에, 이에 따른 통신비용 문제가 존재한다. 본 연구에서는 이를 해결하기 위하여 양자 정보이론과 양자 전송 알고리즘을 기반으로 한 차세대 양자 연합학습 알고리즘을 제안한다.

### Abstract

Federated Learning (FL) is a distributed learning algorithm that effectively mitigates data bias while ensuring user privacy. However, each aggregation requires the transmission and reception of parameters, leading to significant communication costs. To address this issue, this study proposes a next-generation quantum federated learning algorithm based on quantum information theory and quantum transmission algorithms.

### Key words

Federated Learning, Quantum Machine Learning

## 1. 서론

연합학습(FL: Federated Learning)을 비롯한 분산학습 알고리즘의 보안적 성능이 큰 관심을 받음에 따라, 연합학습 알고리즘을 최적화하기 위한 다양한 연구들이 진행되고 있다[1]. 하지만, 연합학습 알고리즘을 위해서는 학습 파라미터 및 그래디언트의 통

신이 필수적으로 요구됨에 따라, 통신비용이 매 통합단계(Aggregation Period)마다 소모되게 된다. 이를 해결하기 위하여 다양한 통신 최적화 알고리즘들이 연합학습에 추가되어 통신효율성을 극대화하였다. 많은 연구들이 통신효율성을 극대화하였지만, 더 적은 통신비용을 지출함과 동시에, 보안성을 유지하는 것은 큰 숙제로 남아있다. 대표적으로, 토큰

\*고려대학교, 67back@korea.ac.kr

※ This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00347180)

기반의 패스워드 시스템을 위해서는 매 통합 단계마다 모든 로컬 유저들이 토큰을 비교하는 방식으로 보안이 이루어지지만, 이를 생략함으로써 보안성을 잃을 수 있다. 이에 반해, 토큰의 숫자를 늘리고 보안성을 강화하기 위한 절차들을 추가할 때마다, 연합학습 시스템은 추가적인 통신비용을 감수해야 하는 문제점을 가지고 있다.

양자 전송(Quantum Teleportation)은 이를 해결할 수 있는 차세대 통신 기법으로 급부상하고있다. 하나의 양자비트(Qubit)의 상태변화가 다른 양자비트 상변화에 영향을 줌으로써, 통신 비용을 극적으로 줄임과 동시에 보안성을 해결할 수 있는 방안을 제시하였다. 하지만, 연합학습에 직접 적용하기에는 이는 매우 큰 학습성의 제약을 가지고있다. 양자 통신을 통하여 상변화가 일어났을 때, 이를 기반으로 하여 글로벌모델을 학습시키는 것은 높은 어려움과 낮은 학습성공률을 나타낼 수 있다. 본 논문에서는, 이러한 문제를 해결하기 위하여 연합학습 프레임워크와 학습가능한 양자회로 및 양자 전송을 기반으로 한 양자 연합학습 프레임워크를 제안한다.

## II. 양자 전송기반 양자 연합학습

본 연구에서는 양자통신 및 정보이론의 Fidelity 개념을 활용하여 각 로컬 모델의 파라미터를정규화 및 통합하였다 [2]. 글로벌 모델을 획득함에있어, 각  $i$ 번째 유저의 양자상태  $\pi_i, i \in \{1, 2, 3, \dots, N\}$  및 전체 양자상태의 정규화된 상태벡터

$\tilde{\phi} = \sum_{i=0}^N \pi_i / \|\sum_{i=0}^N \pi_i\|$  를 전송하여 적용한 Fidelity

알고리즘은 다음과 같이 표현할 수 있다.

$$F = \frac{1}{B} \sum_{j=1}^B \|\langle \phi_j^G, \tilde{\phi}_j \rangle\|^2$$

B는 총 배치사이즈, J는 각각의 데이터,  $\phi_j^G$ 는J번째 데이터의 데이터에대한 Global Model의 정규화된 양자상태를 나타낸다. 이후, 손실함수를 Inverse Fidelity, 1-F 로 설정하여 모델을 학습시켰다.

## III. 실험

### 3.1 실험 세팅

제안 양자 연합학습 알고리즘의 실효성확인을 위하여 전체유저 수  $N=4$ , 배치사이즈  $B=256$ 의 연합학습 프레임워크를 설계하였다. MNIST 데이터셋 중 Binary Classification을위한  $\{0,1\}$  데이터를 추출하여 실험데이터를 만들었고, 2-qubit 양자회로를 사용하였다.

### 3.2 실험 결과

그림 1은 각 Local User 및 Global Model의 Top-1

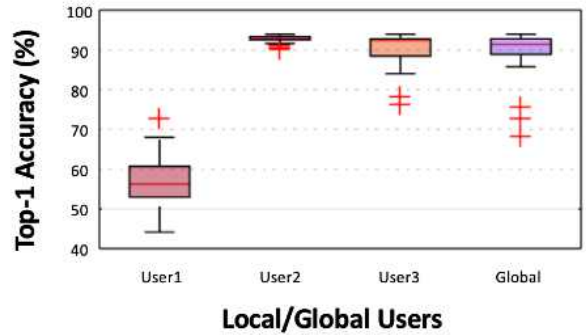


그림 1. Local/Global User의 Top-1Accuracy

를 나타낸다. Binary Classification에있어 User1은 낮은 학습률을 보였다. 반면, User2가 가장 안정된 성능을 보였다. 하지만, 정규화된 양자통신을 기반으로 학습된 Global 모델은 직접적인 파라미터 통신 및 Aggregation 과정없이, 평균 92%의 안정된 성능을 획득하였다. 이는 Fidelity만을 통하여 모델을 학습할때, 수렴하는 양자 연합학습 프레임워크를 만들 수 있다는 실험적 근거가 될 수 있다. 더불어, 각각의 모델들과의 정규화된 상태비교를 통하여 학습됨으로써, 보안성을 획득할 수 있을 것으로 기대된다.

## IV. 결론 및 향후 연구방향

본 연구에서는 양자 전송이론과 양자 정보이론 중 Fidelity개념을 차용하여, 차세대 양자 연합학습의 학습 가능성에 대한 실험을 진행하였다. 단일 양자회로 뿐만아니라, 분산 학습 시스템의 양자 인공지능의 학습 가능성을 실험적으로 증명하였다. 하지만, 매 Aggregation Period동안 제공되는 데이터셋이 동일하다는 제약이 존재하였다.



## 참 고 문 헌

- [1] H. Back, W. J. Yun, Y. Kwak, S. Jung, M. Ji, M. Bennis, J. Park, and J. Kim, "Joint superposition coding and training for federated learning over multi-width neural networks," in Proc. of the IEEE Conference on Computer Communications (INFOCOM), London, United Kingdom, May 2022, pp. 1729-1738.
- [2] J.R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," CoRR, vol. abs/1803.11173, 2018.

# 태양광 패널 기반 배터리 충전 전압의 실시간 모니터링을 위한 원격 시스템 구현

허민규\*, 양 오\*\*

## Implementation of Remote System for Real-Time Monitoring of Battery Charging Voltage for Solar Panels

MinKyu Heo\*, and Oh Yang\*\*

\*\* Semiconductor Engineering of Cheongju University

### 요 약

본 논문에서는 아날로그-디지털 변환기(ADC)와 LoRa 통신을 활용하여 배터리의 충전 전압을 실시간으로 모니터링하는 시스템을 제안한다. 태양광 패널과 같은 재생 가능 에너지원은 시간대나 날씨에 따라 충전 전압이 크게 변화할 수 있다. 이러한 전압 변동은 배터리의 충전 효율과 수명에 직접적인 영향을 미치기 때문에, 실시간 모니터링이 필수적이다. 저전력 및 장거리 통신이 가능한 LoRa 기술을 통해, 본 시스템은 원격으로 충전 상태 모니터링을 구현할 수 있다. 실험 결과, 제안된 시스템은 전압 변동에 대한 원격 모니터링 시스템을 제공하였으며, 이로 인한 편의성과 효율성을 보여주었다. 특히, 날씨 변화에 민감한 태양광 패널 기반 충전 시스템에서 배터리의 상태를 실시간으로 확인할 수 있어, 다양한 IoT 응용에 적합함을 확인하였다. 이러한 모니터링 시스템은 배터리의 수명 관리와 에너지 최적화에 크게 기여할 수 있을 것이다.

### Abstract

This study proposes a system for real-time monitoring of battery charging voltage using an Analog-to-Digital Converter (ADC) and LoRa communication. Renewable energy sources, such as solar panels, can experience significant fluctuations in charging voltage due to variations in time and weather conditions. These voltage changes directly affect the charging efficiency and lifespan of batteries, making real-time monitoring essential. By leveraging the low-power and long-range capabilities of LoRa technology, the proposed system enables remote monitoring of the charging status. Experimental results demonstrate that the system provides effective monitoring of voltage fluctuations, showcasing convenience and efficiency. Notably, the ability to check battery status in real-time within solar panel-based charging systems sensitive to weather changes indicates its suitability for various IoT applications. This monitoring system could significantly contribute to battery lifespan management and energy optimization.

### Key words

Battery monitoring, ADC, LoRa communication, real-time monitoring, IoT, charging voltage

---

청주대학교 반도체공학과 석사과정, Email : [hmk000111@naver.com](mailto:hmk000111@naver.com)

청주대학교 반도체공학과 교수, Email : [ohyang@cju.ac.kr](mailto:ohyang@cju.ac.kr)

이 논문은 2024년 대한민국 산업통상자원부와 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0017011)

## 1. 서 론

재생 가능 에너지원의 활용이 증가함에 따라 태양광 패널이 전력 생산의 핵심으로 자리 잡고 있다. 그러나 태양광 패널은 시간대나 날씨에 따라 전압 변동이 크고, 이는 배터리의 충전 효율과 수명에 부정적인 영향을 미친다. 이러한 변동을 실시간으로 모니터링할 수 있는 시스템의 필요성이 대두되며, 이는 전압 변동에 대한 즉각적인 대응을 가능하게 하여 충전 효율을 극대화하는 중요한 요소로 작용한다[1]. 전통적인 유선 방식이나 단거리 통신 시스템은 원격 모니터링에 한계가 있으며, 이에 따라 장거리 저전력 통신 기술인 LoRa가 주목받고 있다. 본 논문에서는 아날로그-디지털 변환기(ADC)와 LoRa 통신을 활용하여 태양광 패널로 공급되는 배터리의 충전 전압을 실시간으로 모니터링하는 시스템을 제안한다. 이 시스템은 태양광 기반 충전 환경에서도 신뢰성 있게 배터리 상태를 감시할 수 있는 가능성을 제시하며, IoT 환경에서의 다양한 응용 가능성을 지니고 있다. 이러한 시스템의 도입은 재생 가능 에너지의 효율적인 활용과 지속 가능한 에너지 관리에 크게 기여할 것으로 기대된다.

### 2.1 충전 모니터링 시스템 구성

본 논문의 구성은 그림 1과 같이 솔라 패널, 전류 센서(T60404-N4646-X410), 마이크로컨트롤러 유닛(MCU), 장거리 통신 모듈, 충전용 배터리 및 MCU, 전원 공급용 충전식 납 배터리로 구성되어 있다. 솔라 패널에서 생성된 전류는 전류 센서를 거쳐 충전용 배터리에 공급된다. 공급된 전류에 따른 전압 값은 STM사의 STM32F103 MCU의 12비트 ADC 핀으로 전달되어 디지털 데이터로 변환된다.[2] 변환된 ADC 데이터는 MCU를 통해 실제 전압 값으로 계산되어 1s 간격으로 Master LoRa를 통해 Slave LoRa로 전송된다.[3] 이후 Slave LoRa가 수신한 데이터는 PC로 전송되어 CoolTerm 프로그램을 통해 실시간으로 태양광 패널의 충전 전압을 모니터링할 수 있다.

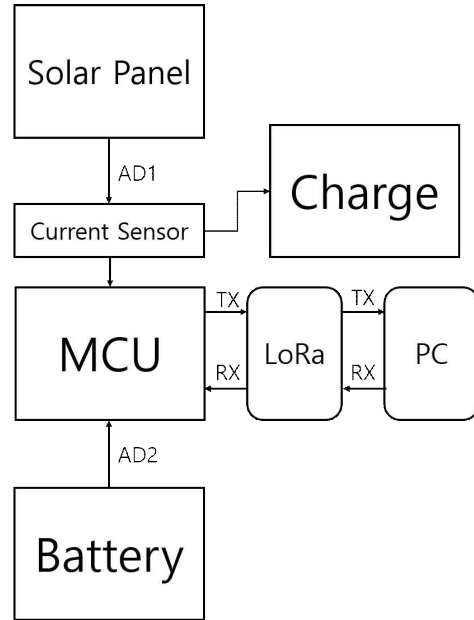


그림 1. 시스템 구성

Fig. 1. System configuration

### 2.2 실험 결과 및 보정 알고리즘

본 실험에서는 제안된 시스템을 통해 태양광 패널로부터 공급되는 전력을 측정하고, ADC 변환으로 수집된 데이터를 분석하여 원격으로 배터리의 충전 전압을 모니터링하였다. 초기 실험에서는 ADC 변환으로 수신한 값과 실제 전압 간에 약 10%의 오차가 발생했으며, 이는 실험에 사용된 MCU의 12 bit ADC 변환의 정밀도 한계에 기인한 것으로 나타났다. 이러한 오차는 배터리 충전 상태를 정확히 판단하는 데 있어 문제를 일으킬 수 있어, IAR 프로그램을 사용하여 간단한 보정 알고리즘을 설계하고 적용하였다. 보정 알고리즘은 다음과 같은 과정을 통해 구현되었다. 먼저 ADC로 수신된 데이터와 실제 전압 간의 차이를 분석하고, 이 오차를 보정하는 선형 보정 알고리즘을 추가하였다.[4] 이때 측정된 값을  $V_{adc}$ , 보정된 값을  $V_{cal}$ , 기울기 보정 계수를  $a$ , 절편 보정 계수를  $b$ 라고 할 때, 보정된 값은 아래의 수식으로 표현된다.

$$V_{cal} = a \cdot V_{adc} + b \quad (1)$$

이때, 계수  $a$ 는 실제 전압  $A_{act}$ 의 차이와 측정된 전압  $V_{adc}$ 의 차이의 비율로 계산한다.

$$a = \frac{V_{act_2} - V_{act_1}}{V_{adc_2} - V_{adc_1}} \quad (2)$$

계수  $b$ 는 보정된 선형 관계의 기준점을 설정하기 위해 사용되며, 기준점에서의 측정값과 실제 값 간 차이를 기반으로 한다.

$$b = V_{act_1} - a \cdot V_{adc_1} \quad (3)$$

위와 같은 보정 알고리즘 적용 후 재측정한 결과, 시스템의 오차율은 약 1% 이내로 감소하여 초기 오차율 대비 크게 개선되었다. 보정된 측정 데이터는 표 1에 제시하였다.

표 1. 보정 알고리즘이 적용된 원격 전압 측정  
Table 1. Remote voltage measurement with compensation algorithm

실제 전압(V)	LoRa 통신 수신값(V)	오차율(%)
4.11	4.09	0.49
8.03	8.00	0.37
11.98	12.07	0.75
16.17	16.12	0.31
19.85	19.86	0.05

그림 2과 같이, LoRa 모듈을 통해 1000ms 간격으로 송수신된 데이터는 PC의 CoolTerm 프로그램에서 실시간으로 확인할 수 있다.

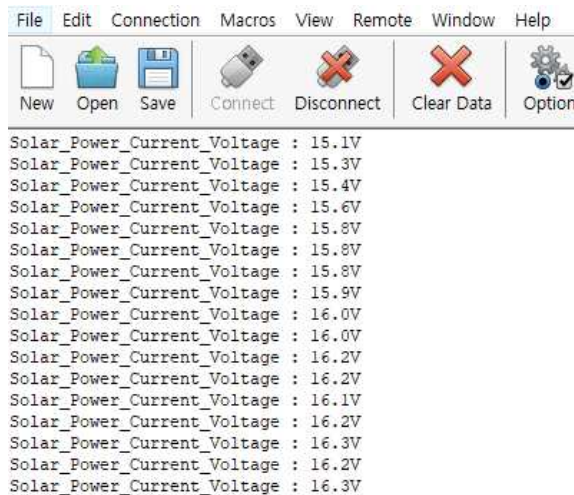


그림 2. CoolTerm 프로그램에서 확인한 전압값  
Fig. 2. Voltage value confirmed by CoolTerm program

### III. 결 론

본 논문에서는 배터리 충전 전압을 실시간으로 모니터링하기 위한 원격 시스템을 제안하고 구현하였다. 제안된 시스템은 12비트 ADC와 보정 알고리즘을 활용하여, 태양광 패널과 같이 전압 변동성이 큰 환경에서도 안정적으로 배터리 상태를 모니터링할 수 있도록 설계되었다. 또한, 저전력 및 장거리 통신이 가능한 LoRa 모듈을 채택하여 기존 유선 기반 시스템 혹은 단거리 통신 시스템보다 높은 유연성과 효율성을 제공하였다. 실험 결과, 다양한 충전 환경에서 안정적인 전압 모니터링이 가능함을 확인하였고, 이는 충전 전압 변동에 민감한 재생 가능 에너지 기반 시스템에 특히 유용함을 시사한다. 제안된 시스템은 IoT 기반 에너지 관리, 원격 모니터링 장치 등 다양한 응용 분야에서 배터리 수명 관리와 에너지 최적화에 기여할 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] Krishnan, R., Reddy, P., "An Efficient Data Communication Technique for Solar Panel-Based Energy Systems," International Journal of Renewable Energy Research, 9(2), pp. 637-645, (2019).
- [2] Suresh, P., & Ravi, K. "A 12-bit ADC with Low-Power Characteristics," International Journal of Electronics and Communications, vol. 78, pp. 15-25, 2021.
- [3] Haxhibeqiri, J., De Poorter, E., Moerman, I., & Hoebeke, J. (2018). A Survey of LoRaWAN for IoT: From Technology to Application. Sensors, 18(11), 3995. MDPI.
- [4] Hwang, H., & Lee, S., "A simple linearity correction algorithm for high-resolution ADCs," 2017 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1-4, (2017).

# 멀티스케일 디시전 트랜스포머에 관한 연구

이동욱\*, 홍송남\*\*

## Multi-scale Decision Transformer

Dongwook Lee\*, and Songnam Hong\*\*

오프라인 강화학습은 미리 수집된 고정된 데이터셋을 이용해 에이전트를 학습시키는 방법으로, 온라인 강화 학습과는 달리 학습 과정에서 환경과의 실시간 상호작용 없이 학습이 진행된다. 온라인 강화학습을 실생활에 적용하기 위해서는 실제 환경과 상호작용하는 과정에서 막대한 시간과 비용이 필요하며 실시간 탐색 중 에이전트가 택하는 잘못된 행동이 큰 피해를 초래할 수도 있다. 반면에 오프라인 강화학습은 기존에 가지고 있는 많은 양의 데이터셋을 적극 활용하고 실시간 탐색에서 발생할 수 있는 불안전성을 피할 수 있다는 측면에서 효과적이다. 대부분의 전통적 오프라인 강화학습 알고리즘은 정책 네트워크와 가치함수를 학습하지만, Decision Transformer는 트랜스포머 아키텍처를 사용해 오프라인 강화학습 문제를 시계열 예측 문제로 재해석하여 과거의 상태, 행동, 보상을 시간 순서로 연결된 시퀀스로 보고 목표 달성 경로를 예측한다. Decision Transformer는 새로운 측면에서 오프라인 강화학습 문제를 해결하였고 전통적인 오프라인 강화학습 알고리즘과 비교하여 좋은 성능을 내었다. 그러나 강화학습에서는 단기적인 보상과 장기적인 보상을 구별하는 것이 필요하고 이를 반영하기 위해 본 연구에서는 Decision Transformer에 멀티 스케일 계층적 표현을 도입하였다. 강화학습 환경에서는 현재의 행동이 미래의 보상에 즉각적 또는 장기적으로 영향을 줄 수 있는데 서로 다른 시간 단위의 멀티 스케일 표현을 통해 저해상도 스케일에서는 전체 시퀀스의 장기 의존성을 학습하고 고해상도 스케일에서는 개별 행동 및 상태의 세부 정보를 동시에 학습할 수 있다. 향후 연구에서는 멀티스케일 디시전 트랜스포머를 다양한 강화학습 환경에서 실험을 진행 할 예정이다.

### Key words

Machine Learning, Reinforcement Learning, Offline Reinforcement Learning

### 참고문헌

- [1] L. Chen, K. Lu, A. Rajeswaran, K. Lee, A. Grover, M. Laskin, P. Abbeel, A. Srinivas, and I. Mordatch, "Decision transformer: Reinforcement learning via sequence modeling", *Advances in Neural Information Processing Systems (NeurIPS)* 34, 2021.
- [2] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows", *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10012-10022, 2021.

---

\* 한양대학교 융합전자공학과 석사과정, dwlee2718@hanyang.ac.kr,

\*\* 한양대학교 융합전자공학부 교수, snhong@hanyang.ac.kr(교신저자)

# 인공지능을 이용한 신규 설진 지표 개발

김지혜\*

## Develop of New Tongue Diagnostic Parameter Using the Artificial Intelligence

Jihye Kim\*

설진은 한의학의 주요 진단 방법 중 하나로, 혀를 살펴 인체 건강 상태를 판단하는 데 사용된다[1]. 혀의 설질과 설태 색상, 설태의 양이 주요 지표로 활용되며, 이 지표의 변화를 관찰하여 병의 성질과 경중을 파악할 수 있다[1, 2]. 하지만 설진은 한의사의 육안 관찰에 의존하기 때문에 주관적 편향의 한계가 존재한다[3]. 이를 해결하기 위해 많은 연구자들이 설진의 정확도와 신뢰도 향상을 위한 연구를 수행하였으며, 최근 인공지능 발전은 한의학 진단에 새로운 지표를 제안하는 데 기여하고 있다[4]. 본 연구는 인공지능 방법론을 적용하여 기존 지표와 차별화된 신규 설진 지표를 제안하고자 한다. 2020년부터 2021년까지 5개 한방병원에서 모집된 건강한 성인의 혀 영상 1,810장을 분석에 활용하였다. 연구 경력 10년 이상의 진단학 박사 두 명이 열문 중증도에 따라 ‘열문 없음’, ‘열문 약간 있음’, ‘열문 있음’, ‘열문 심함’으로 혀 영상을 독립적으로 분류하고, 협의된 기준을 바탕으로 참값을 설정하였다. 이진화 마스크 이미지를 통해 혀와 열문 영역을 표기하고, U-Net 모델을 활용하여 의료용 이미지 분할을 수행하였다. 전체 데이터는 훈련용 데이터 셋(n=1,448)과 테스트용 데이터 셋(n=362)으로 8:2 비율로 나누었으며, 모델의 과적합을 방지하기 위해 매 에포크마다 검증 데이터 손실 변화를 모니터링하여, 30회 연속 검증 손실이 개선되지 않을 경우 학습을 중단하도록 설정하였다. 학습 후 저장된 weight matrix 파일을 사용하여 혀 영역을 예측하도록 테스트 데이터에 모델을 적용하였다. 열문 중증도는 전체 혀 영역 대비 열문 영역이 차지하는 비율로 정의하였다. 추정된 열문 중증도 지수와 실제 열문 중증도 지수 간의 상관 관계를 확인하기 위해 Pearson 상관계수를 측정된 결과, 상관계수는 0.8920, p-value  $\leq 0.0001$ 로 강한 상관관계를 보였다. 추정된 열문 중증도 지수를 0.348% 기준으로 열문설 유무를 분류했을 때, 전문가 판정을 참값 기준으로 정확도는 0.909, 민감도는 0.908, 특이도는 0.910, AUC는 0.958을 나타냈다. 이는 인공지능을 기반으로 계산된 열문 중증도 지수가 열문설 유무를 정확하게 분류하는 것을 의미한다. 본 연구는 인공지능 기반 열문 중증도 지수가 임상에서 열문 진단을 정량적·객관적으로 지원하며, 치료 전후의 호전도를 수치화하는 데 도움을 줄 수 있음을 보여주며, 한의학 진단의 과학화에 기여할 것으로 기대된다.

### Key words

tongue diagnosis, tongue feature, fissured tongue, AI, diagnostic parameter

### 참고문헌

[1] TIRYAKI, Burcu, et al. Artificial intelligence in tongue diagnosis: classification of tongue lesions and normal

\*한국한의학연구원 디지털임상연구부, wisdom00311@kiom.re.kr

※ This work was supported by the Korea Institute of Oriental Medicine (KSN1824130, KSN18314122).

- tongue images using deep convolutional neural network. *BMC Medical Imaging*, 2024, 24.1: 59.
- [2] CHIU, Chuang-Chien. A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue. *Computer methods and programs in biomedicine*, 2000, 61.2: 77-89.
- [3] LEE, Hyeon-Joo; KWEON, Na-Yeon; NAM, Dong-Hyun. A review on Clinical Trials by Using a Computerized Tongue Diagnosis System. *The Journal of the Society of Korean Medicine Diagnostics*, 2015, 19.1: 1-10.
- [4] LIU, Qi, et al. A survey of artificial intelligence in tongue image for disease diagnosis and syndrome differentiation. *Digital Health*, 2023, 9: 20552076231191044.

# NPMI-BERT4Rec: 카테고리 수준의 보완재 정보를 통합한 향상된 순차적 추천 시스템

원다정\*, 이의상\*\*, 김휘민\*\*\*, 함종수\*\*\*\*

## NPMI-BERT4Rec: Enhancing Sequential Recommendation Systems by Integrating Category-Level Complementary Information

Dajeong Won\*, Euisang Lee\*\*, Hwimin Kim\*\*\* and Jongsoo Ham\*\*\*\*

### 요 약

추천 시스템은 디지털 비즈니스에서 개인화된 경험을 제공하고 전환율을 높이는 데 중요한 역할을 한다. 기존 연구에서 주목받은 Transformer 기반의 BERT4Rec은 사용자의 상호작용을 양방향으로 학습해 기존 순차적 시스템보다 뛰어난 성능을 보였으나, 보완재와 같은 비순차적 관계를 반영하는데 한계가 있다. 본 연구에서는 이를 개선하기 위해 정규화된 점별 상호정보량(NPMI)을 BERT4Rec에 통합한 NPMI-BERT4Rec을 제안한다. NPMI-BERT4Rec은 개별 사용자 행동과 전체 사용자 행동 패턴을 동시에 학습해 더욱 정교한 추천을 가능하게 하며, 실험 결과 BERT4Rec 대비 평균 약 388.65%의 성능 향상을 보였다.

### Abstract

Recommendation systems play a crucial role in digital business by providing personalized user experiences and increasing conversion rates. The Transformer-based BERT4Rec model, which has garnered attention in previous studies, learns user interactions bidirectionally and has demonstrated superior performance to traditional sequential systems. However, it has limitations in capturing non-sequential relationships such as those between complementary items. This study proposes NPMI-BERT4Rec, which integrates Normalized Pointwise Mutual Information(NPMI) into BERT4Rec to address this issue. NPMI-BERT4Rec learns both individual and global user behavior patterns, enabling more precise recommendations. Experimental results show an average performance improvement of approximately 388.65% over BERT4Rec.

### Key words

BERT4Rec, Recommendation System, Sequential Recommendation, Personalized Recommendation, NPMI, Complementary Item Relationships

---

\* 한신대학교 수학과, ORCID: <https://orcid.org/0009-0005-4270-4238>

\*\* 성균관대학교 정보통신대학원 빅데이터학 석사, ORCID: <https://orcid.org/0009-0007-6716-2812>(공동 1저자)

\*\*\* 한국외국어대학교 ELLT학과, ORCID: <https://orcid.org/0009-0001-3883-2254>(공동 1저자)

\*\*\*\* 한양대학교 인공지능융합학과, ORCID: <https://orcid.org/0009-0006-8216-5248>(교신저자)



## I. 서 론

전자상거래, 스트리밍 서비스, 광고 플랫폼 등 다양한 디지털 비즈니스에서 맞춤형 추천 시스템은 사용자의 경험을 향상시키고 구매 전환율을 높이는 필수 도구로 자리 잡았다. 특히 Transformer 기반의 모델들은 자연어 처리 분야에서의 성공을 바탕으로 추천 시스템 분야에서도 성능 향상을 보여주며 주목받고 있다[1]. 이러한 모델의 대표적인 예로 Transformer 기반의 BERT 아키텍처를 추천 시스템에 적용한 BERT4Rec[2]이 있으며, 이는 사용자의 과거 행동 시퀀스를 양방향으로 학습하여 기존 순차적 추천 모델보다 정교하게 맞춤형 예측을 수행할 수 있음을 입증하였다. 그러나 BERT4Rec은 시퀀스 내 상품 간 순차적 의존성에 주로 초점을 맞추기 때문에 한계가 있다[3]. 이에 본 연구에서는 BERT4Rec의 비순차적 상호작용을 충분히 반영하지 못하는 한계를 보완하기 위해 상품 카테고리 간의 보완재 관계를 정규화된 점별 상호정보량(Normalized Pointwise Mutual Information, NPMI)으로 정량화하고, 이를 BERT4Rec에 통합한 NPMI-BERT4Rec을 제안한다.

여기서 보완재란 사용자들이 상품을 소비하는 과정에서 특정 상품이 다른 상품과 함께 구매되는 패턴을 의미한다. 예를 들어, ‘우유’를 구매하는 사용자는 ‘빵’이나 ‘시리얼’을 함께 구매하는 경향을 보인다. 본 연구에서는 이러한 상품 간의 관계를 보완재로 정의하고, 보완성을 정량화된 수치로 변환하기 위한 방법으로 NPMI를 도입한다. 이를 통해 일반적으로 개인화 추천 시스템에서 중요시되는 개별 사용자의 행동 패턴에 글로벌하게 집계된 NPMI 정보를 적용함으로써 모델이 모든 사용자 간의 공통된 소비 패턴 및 보완재 관계를 포착하도록 한다. 이는 단순히 사용자의 과거 행동에 기반한 예측을 넘어, 개별 사용자에게도 유효한 전체 사용자 기반의 정보를 통합하여 보다 풍부하고 관련성 높은 추천을 가능하게 한다.

## II. 본 론

### 2.1 카테고리 수준 NPMI 생성

NPMI는 점별 상호정보량(Pointwise Mutual Information, PMI)를  $[-1, 1]$ 로 정규화한 지표로, 두 사건 간의 연관성을 정량화하는 데 사용되는 정보이론에서 파생된 측정 방식이다[4]. 본 연구에서는 이를 카테고리 간 보완재성을 나타내는 지표로 활용하였다.

$$PMI(x,y) = \log_2 \frac{P(x,y)}{P(x)P(y)} \quad (1)$$

PMI는 카테고리  $x$ 의 등장 확률  $P(x)$ , 카테고리  $y$ 의 등장 확률  $P(y)$ , 카테고리  $x$ 와  $y$ 가 동시에 등장할 확률  $P(x,y)$ 에 기반하여 특정 카테고리들이 독립적으로 혹은 함께 나타나는 경향이 있는지를 수치화한다. 이후 PMI 값을  $[-1, 1]$  범위로 정규화한 NPMI를 통해 두 카테고리 간의 결합 정도를 정량적인 수치로 직관적으로 해석할 수 있다.

$$NPMI(x,y) = \frac{PMI(x,y)}{-\log_2 P(x,y)} \quad (2)$$

NPMI는 1에 가까울수록 시퀀스 내에서 두 카테고리가 함께 등장할 가능성이 높음을, -1에 가까울수록 두 카테고리가 함께 등장할 가능성이 낮음을 의미한다.

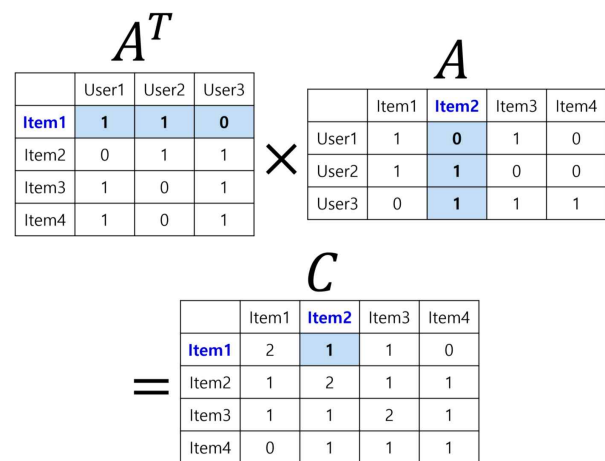


그림 1. 동시 발생 행렬 생성

Fig. 1. Co-occurrence matrix generation

이러한 확률 계산 시 동시 발생 행렬 (Co-occurrence matrix)이 선행적으로 생성되어야 하며, 본 연구에서는 연산의 효율성을 고려하여 그림 1과 같이 행렬 연산을 통해 구축하였다.

사용자-상품 상호작용 행렬  $A$ 에 대해, 동시 발생 행렬  $C$ 는  $C = A^T A$ 로 정의된다. 여기서  $C_{ij}$ 는 모든 사용자에게 대해 상품  $i$ 와  $j$ 가 함께 선택되는 강도를 나타낸다.

$$C_{12} = A_{11}^T A_{12} + A_{12}^T A_{22} + A_{13}^T A_{32} \quad (3)$$

예를 들어,  $C_{12}$ 는 식 (3)과 같은 방법으로 계산되며, 모든 사용자에게 대해 Item1과 Item2가 동시에 선택되는 강도를 나타낸다.

## 2.2 데이터셋

본 연구에서는 NPMI-BERT4Rec의 성능을 다양한 특성과 규모를 가진 실제 비즈니스 환경의 데이터셋을 활용해 평가하고자 했다. 카테고리 수준의 보완재 관계를 정의하고 학습시키기 위해 사용자, 상품, 카테고리 정보를 가진 데이터셋을 선정하였다.

표 1. 실험 데이터셋  
Table 1. Experimental datasets

데이터셋	사용자 수	상품 수	카테고리 수
Dunnhumby	2,500	92,339	308
ML-1m	6,040	3,706	301
ML-20m	138,493	26,744	1,329
Taobao	286,618	524,206	6,748

- Dunnhumby(The Complete Journey) [5]: Dunnhumby에서 제공하는 소매업체에서 자주 쇼핑하는 2,500명의 사용자에게 대한 2년 동안의 거래 내역을 가진 데이터이다.
- MovieLens [6]: 추천 알고리즘을 평가하기 위해 많이 사용되는 벤치마크 데이터셋이다. MovieLens 1m(ML-1m)과 MovieLens 20m(ML-20m)의 두 가지 버전을 사용했다. 영화 장르를 카테고리로 적용하였다.
- Taobao [7]: Alibaba에서 제공하는 사용자 행동

데이터셋이다. 사용자 행동 중 Buy(구매)에 대한 데이터를 채택하였다.

## 2.3 모델링

BERT4Rec은 Transformer의 Encoder 부분만 사용한 모델인 BERT에서 MLM(Masked Language Model)과 NSP(Next Sentence Prediction) 기법 중 MLM만을 채택하여 추천 시스템에 적합하게 변형시킨 모델이다. 본 연구에서는 기존의 BERT4Rec 재현을 위해 TensorFlow 2.5 버전에 맞게 자체적으로 구현하였다.

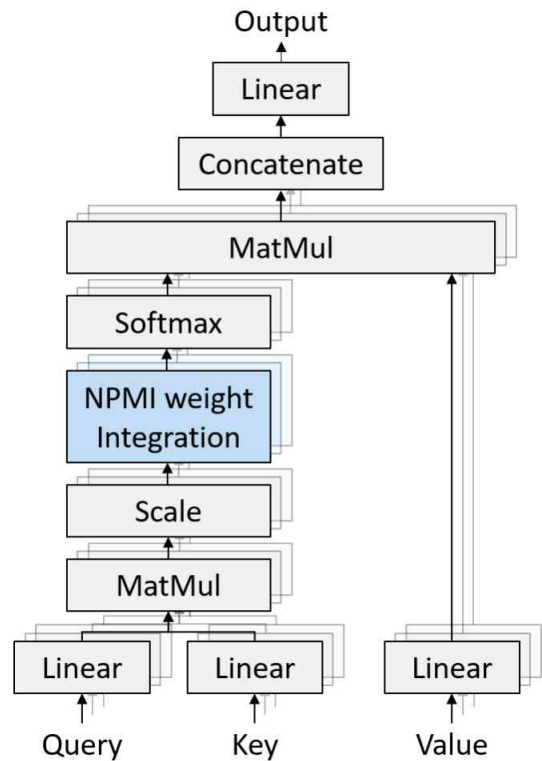


그림 2. NPMI-BERT4Rec 멀티헤드 어텐션  
Fig. 2. NPMI-BERT4Rec multi-head attention

NPMI-BERT4Rec의 아키텍처는 기존 BERT4Rec의 구조를 기반으로 하되, 어텐션 메커니즘에 NPMI 가중치를 통합하여 상품 간 보완재 관계를 반영한다. 훈련 데이터 전체를 이용하여 카테고리 간 NPMI 값을 계산하고, 이를 학습 과정에서 저장된 NPMI 가중치 행렬  $W_{NPMI}$ 를 불러와 모델에 적용한다.

NPMI 통합은 멀티헤드 어텐션에서 계산된 어텐션 스코어에 NPMI 가중치를 곱하여 조정된 어텐션

스코어를 사용함으로써 이루어진다.

$$S' = S \odot (1 + \beta \cdot W_{NPMI}) \quad (4)$$

여기서  $\beta$ 는 NPMI 가중치의 영향력을 조절하는 파라미터로, 목적에 따라 조정이 가능하나 기본적으로  $W_{NPMI}$ 가 [-1, 1]로 정규화되어 있으므로 본 연구에서는  $\beta = 1.0$ 으로 설정하여 모든 데이터셋에 대해 동일한 조건으로 비교하였다. 조정된 어텐션 스코어  $S'$ 은 이후 소프트맥스 함수를 적용하여, 컨텍스트 벡터를 계산하는 데 사용된다.

## 2.4 평가 지표

본 연구에서는 기존 연구의 평가 방식을 유지하였다. Leave-One-Out(LOO) 평가 방식을 적용하여, 각 사용자에게 대해 마지막으로 상호작용한 상품을 테스트 데이터로, 나머지 상품을 훈련 데이터로 사용하였다. 이를 통해 모델이 사용자의 다음 상품을 얼마나 잘 예측하는지를 평가한다.

또한, 평가의 효율성을 위한 부정 샘플링(Negative Sampling) 방식이 유지되었다. 각 사용자마다 실제 상호작용한 상품과 함께 100개의 상호작용이 없었던 부정 아이템을 무작위로 샘플링하여, 모델이 실제 상품을 부정 상품들 사이에서 얼마나 잘 순위화하는지를 평가했다.

평가 지표는 Hit Rate(HR), Normalized Discounted Cumulative Gain(NDCG), Mean Reciprocal Rank(MRR)을 사용하였다. 부정 샘플링 환경에서 HR@k는 Recall@k와 동일하게 해석되며, Precision@k에 비례한다. 따라서 k=1, 5, 10에 대한 HR와 NDCG, 그리고 MRR을 통해 기존 연구 결과와 동일한 평가 지표를 적용하여 NPMI-BERT4Rec의 성능을 BERT4Rec과 일관성 있게 평가한다.

## 2.5 성능 비교

본 연구에서는 제안한 NPMI-BERT4Rec과 기존 BERT4Rec을 네 가지 공개 데이터셋을 사용하여 비교 평가하였으며, 그 결과는 표 2에 요약되어 있다.

NPMI-BERT4Rec은 모든 데이터셋과 평가 지표에서 BERT4Rec보다 향상된 성능을 보였다. 특히

Dunnhumby와 ML-1m, ML-20m 데이터셋에서는 HR@1과 같은 주요 지표에서 80% 이상의 향상을 기록하였으며, 이는 NPMI 가중치 적용이 추천 정확도에 크게 기여했음을 시사한다. 반면 Taobao 데이터셋에서는 상대적으로 향상률이 낮았으나, 마찬가지로 모든 지표에서 성능이 개선되었다.

이러한 결과는 NPMI-BERT4Rec이 상품 간의 보완재 관계를 효과적으로 반영하여 사용자 행동 패턴을 보다 정교하게 예측할 수 있음을 보여준다. 특히, 카테고리 간의 NPMI 정보를 통합함으로써 모델이 개별 사용자의 행동뿐만 아니라 전체 사용자의 공통된 행동 패턴을 학습할 수 있어 추천의 정확성과 다양성이 향상된 것으로 해석된다.

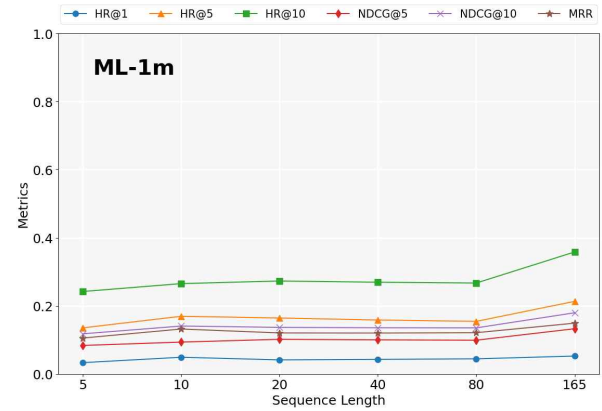


그림 3. ML-1m 데이터셋에서 시퀀스 길이에 따른 BERT4Rec의 평가 지표 변화

Fig. 3. Variation of BERT4Rec metrics with sequence length on the ML-1m dataset

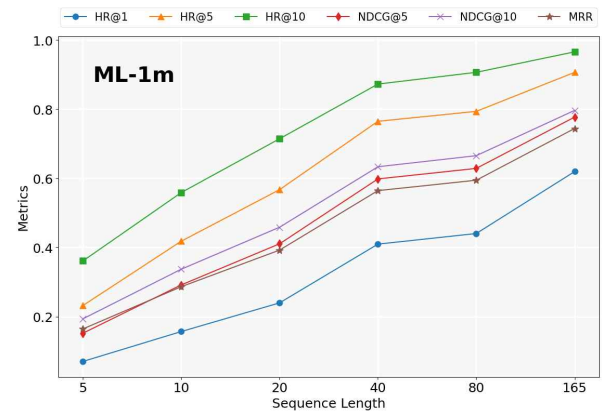


그림 408. ML-1m 데이터셋에서 시퀀스 길이에 따른 NPMI-BERT4Rec의 평가 지표 변화

Fig. 4. Variation of NPMI-BERT4Rec metrics with sequence length on the ML-1m dataset

표 2. BERT4Rec과 NPMI-BERT4Rec의 다음 상품 예측 성능 비교

Table 2. Performance comparison of BERT4Rec and NPMI-BERT4Rec on next-item prediction

데이터셋	평가 지표	BERT4Rec	NPMI-BERT4Rec	Improvement rate
Dunnhumby	HR@1	0.0656	0.6104	830.49%
	HR@5	0.1248	0.8736	600.00%
	HR@10	0.1692	0.9000	431.91%
	NDCG@5	0.0961	0.7568	687.51%
	NDCG@10	0.1102	0.7654	594.56%
	MRR	0.1106	0.7236	554.25%
	ML-1m	HR@1	0.0525	0.6194
HR@5		0.2131	0.9063	325.29%
HR@10		0.3581	0.9656	169.65%
NDCG@5		0.1327	0.7769	485.46%
NDCG@10		0.1796	0.7963	343.37%
MRR		0.1491	0.7439	398.93%
ML-20m		HR@1	0.0642	0.7098
	HR@5	0.2435	0.9218	278.56%
	HR@10	0.3925	0.9677	146.55%
	NDCG@5	0.1531	0.8268	440.04%
	NDCG@10	0.2010	0.8418	318.81%
	MRR	0.1671	0.8031	380.61%
	Taobao	HR@1	0.0128	0.0233
HR@5		0.0584	0.0810	38.70%
HR@10		0.1122	0.1393	24.15%
NDCG@5		0.0351	0.0519	47.86%
NDCG@10		0.0523	0.0706	34.99%
MRR		0.0554	0.0712	28.52%

또한, 데이터셋 규모와 특성에 따라 각 모델의 성능 지표 및 향상률에 차이가 나타났는데, 이는 NPMI-BERT4Rec이 대규모 데이터셋에서 더욱 효과적으로 작동함을 의미할 수 있다. Taobao 데이터셋에 대해서는 상대적으로 향상률이 낮았지만 결과적으로 모든 데이터셋에 대해서 성능이 개선되었다. 평균 시퀀스 길이가 5.1로 매우 짧았던 Taobao 데이터셋에서는 두 모델 모두 시퀀스 길이가 긴 데이터셋에 비해 상대적으로 저조한 결과를 보였는데, 이는 시퀀스 기반 Transformer 모델이 입력 시퀀스 길이에 큰 영향을 받을 수 있음을 유추하게 한다. 이러한 영향을 분석하기 위해 추가적인 실험을 진행하였다.

ML-1m 데이터셋을 선정하여, 충분한 사용자 수, 제품 수, 카테고리 수를 확보하고 데이터의 평균

시퀀스 길이를 기존 165에서 5, 10, 20, 40, 80으로 축소하여 테스트를 진행해 보았다. 그 결과, 그림 3의 BERT4Rec에서는 시퀀스 길이가 길어질수록 평가 지표가 소폭 증가하는 경향을 보였으나, 그 향상률은 미미했다. 반면, 그림 4의 NPMI-BERT4Rec에서는 시퀀스 길이가 길어짐에 따라 평가 지표가 상대적으로 크게 향상되는 결과를 확인하였다.

위 추가적인 실험을 통해 두 모델 모두 입력 시퀀스의 길이가 성능에 큰 영향을 미친다는 것을 나타냈지만, 여전히 NPMI-BERT4Rec이 BERT4Rec으로 충분한 성능을 발휘하지 못하는 데이터셋에서 역시 기본적인 성능을 보장하며 BERT4Rec보다 향상된 결과를 보여주었다.

종합적으로, 본 연구의 실험 결과는 NPMI-BERT4Rec이 기존의 BERT4Rec에 비해 추천 시스템의 성능을 전반적으로 향상시킬 수 있음을

입증하였으며, 이는 보완재 관계를 효과적으로 반영한 모델 설계의 중요성을 강조한다.

### III. 결 론

본 연구에서는 BERT4Rec의 한계를 보완하고자 카테고리 간 보완재 관계를 정량화한 NPMI 지표를 통합한 NPMI-BERT4Rec을 제안하였다. 이를 통해 추천 시스템이 단순히 개별 사용자의 과거 행동 패턴뿐만 아니라 전체 사용자 기반의 공통된 소비 패턴을 반영할 수 있도록 하였다.

네 가지 공개 데이터셋을 활용한 실험 결과, 제안된 NPMI-BERT4Rec은 기존의 BERT4Rec에 비해 전반적으로 향상된 성능을 나타냈다. 특히, 다양한 데이터셋에서 평가 지표를 통해 향상된 결과를 확인할 수 있었으며, 이는 학습 단계에서의 NPMI 가중치 통합이 추천 정확도에 긍정적인 영향을 미쳤음을 시사한다.

본 연구의 주요 시사점은 다음과 같다.

- 보완재 관계의 통합 효과: 상품 간의 보완재 관계를 모델에 통합함으로써 사용자의 소비 패턴을 보다 정교하게 반영할 수 있었으며, 이는 개인화 추천 시스템의 성능을 높이는 데 기여하였다.
- 글로벌 패턴의 활용: 개별 사용자의 행동 데이터뿐만 아니라 전체 사용자 기반의 공통된 소비 패턴을 학습함으로써 추천의 관련성과 다양성을 동시에 향상시킬 수 있었다.
- 모델의 확장 가능성: 본 연구에서 제안한 방법은 상품 분류와 같은 일부 카테고리에 한정되지 않고, 영화의 장르와 같은 다른 도메인의 카테고리 정보에도 적용 가능성을 확인하였다. 이는 다양한 도메인에서 추천 시스템의 응용 가능성을 넓혀준다.

향후 연구에서는 특정 도메인에 특화된 데이터셋에서의 성능 향상을 위해 도메인 특성에 맞춘 추가적인 최적화 방안을 모색할 필요가 있다. 또한, 보완재 관계 외에도 다양한 상품 간 관계를 통합함으로써 추천 시스템의 성능을 더욱 향상시킬 수 있을

것으로 기대된다.

종합적으로, 본 연구는 보완재 관계를 효과적으로 반영한 NPMI-BERT4Rec이 기존의 BERT4Rec보다 향상된 성능을 보임을 입증하였으며, 이는 추천 시스템 설계에 있어 상품 간의 상호관계를 정량화하여 통합하는 방법의 중요성을 강조한다. 이러한 접근은 다양한 디지털 비즈니스 환경에서 추천의 정확도를 향상시켜 서비스의 고도화를 이루는 데 기여할 수 있을 것이다.

### 참 고 문 헌

- [1] W.-C. Kang and J. McAuley, "Self-attentive sequential recommendation", Proc. IEEE Int. Conf. Data Mining (ICDM), pp. 197-206, Nov. 2018.
- [2] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, et al., "BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer", Proc. 28th ACM Int. Conf. Inf. Knowl. Manage., pp. 1441-1450, Nov. 2019.
- [3] Wang S, Hu L, Wang Y, Cao L, Sheng QZ, Orgun M. "Sequential recommender systems: challenges, progress and prospects." arXiv preprint arXiv:2001.04830. Dec. 2019.
- [4] J. Y. Lee, "A study on relative mutual information coefficients", Journal of the Korean Society for Library and Information Science, vol. 37, no. 4, pp. 178-198, Dec. 2003.
- [5] dunnhumby Source files - The Complete Journey, <https://www.dunnhumby.com/source-files/> [accessed : Oct. 30. 2024]
- [6] F. M. Harper and J. A. Konstan, "The MovieLens datasets: History and context", ACM Trans. Interact. Intell. Syst., vol. 5, no. 4, pp. 1-19, Dec. 2015.
- [7] Alibaba Cloud - User Behavior Data from Taobao for Recommendation <https://tianchi.aliyun.com/dataset/649?lang=en-us> [accessed : Oct. 30. 2024]

# 의존성 분석을 통한 산업제어시스템 보안조치 차등적용 방안 연구

이채창\*, 이현주\*\*

## A Study on Risk informed Graded Approach on Industrial Control Systems through Dependency Analysis

Chaechang Lee\*, and Hyunjoo Lee\*\*

### 요 약

산업제어시스템의 디지털화로 산업현장에 대한 사이버보안 위협이 증가하고 있다. 사이버공격에 따른 IT 환경의 보호를 위하여 다양한 보안조치를 적용하는 것과 같이, 산업제어시스템에도 사이버보안을 위한 다양한 보안조치들이 존재한다. 하지만 인력, 예산 등 한정된 자원으로 인해 모든 디지털 시스템들에 대한 일률적인 보안조치의 적용은 현실적이지 않고, 현장의 사이버보안 활동을 더욱 어렵게 만든다. 이 논문에서는 디지털 시스템의 의존성 분석을 통해 보안조치가 필요한 설비의 우선순위를 식별하여 보안조치를 차등적으로 적용하고, 사이버공격에 따른 영향성을 시각화할 수 있는 방법을 제안한다.

### Abstract

With the digitalization of industrial control systems, cybersecurity threats to industrial environments are increasing. Just as various security controls are applied to protect IT environments against cyber attacks, there are also diverse security controls for industrial control systems to enhance cybersecurity. However, due to limited resources such as personnel and budget, uniformly applying security controls to all systems is impractical and complicates cybersecurity efforts on-site. This paper proposes a method to identify the priority of systems requiring security controls through dependency analysis of systems, enabling differential application of security controls. Furthermore, we introduce a way to visualize the impact of cyber attacks.

### Key words

Industrial Control Systems, Cybersecurity Controls, Dependency Analysis, Risk informed graded approach

### 1. 서 론

최근 산업 현장의 산업제어시스템이 빠르게 디지

털화되면서 사이버위협도 크게 증가하고 있다. 산업 제어시스템의 사이버보안을 위해 다양한 표준과 보안조치들도 개발되어 공개되었다[1,2,3]. 하지만, 예

\*한국원자력통제기술원, chiching@kinac.re.kr(교신저자), \*\*한국원자력통제기술원, sn220865@kinac.re.kr

산이나 인력과 같은 한정된 자원으로 인해 모든 디지털시스템에 대해 가능한 모든 보안조치를 일률적으로 적용하는 것은 현실적으로 불가능하다. IAEA는 원자력시설에 대한 사이버 보안조치의 적용 시, 제어시스템의 침해로 인한 위협에 기반하여 보안조치를 차등적으로 적용할 것을 권고하고 있다[4].

이를 위하여 미국원자력협회(Nuclear Energy Institute)는 사이버공격으로 인해 원자력시설의 안전, 보안, 비상대응 기능(SSEP 기능) 등에 미칠 영향을 고려하여, 반드시 보호해야 하는 필수디지털자산(CDA)을 Direct CDA와 Non-Direct CDA로 구분하고 각각의 하드웨어 및 소프트웨어 특성에 맞게 적용 가능한 보안조치를 구분함으로써 위협기반 보안조치 차등적용 방법을 제안하였다[5].

미국 전력연구원(Electric Power Research Institute)은 디지털 시스템의 구성을 검토하여 시스템에 발생 가능한 사이버공격을 식별하고, 이에 따른 영향을 고려하여 보안조치를 적용함으로써 목표로 한 보안 수준을 만족하는지 여부를 정량화된 점수로 판단할 수 있는 Technical Assessment Methodology (TAM)을 제안하였다[6].

## II. 의존성 분석을 통한 보안조치 차등적용

우리나라는 미국과 함께 원자력시설의 컴퓨터 및 정보시스템이 설계기준위험에 포함된 사이버공격에 대해 보호됨을 보장하기 위해, SSEP 기능을 수행하는 디지털 시스템과, 침해 시 해당 기능들에 악영향을 미치는 지원시스템 및 기기들을 CDA로 정의하고 있다[8,9]. 원자력사업자는 CDA 식별을 위해 우선 필수시스템(CS)을 식별하여야 하며, 이 과정은 그림 1과 같다[7,8].

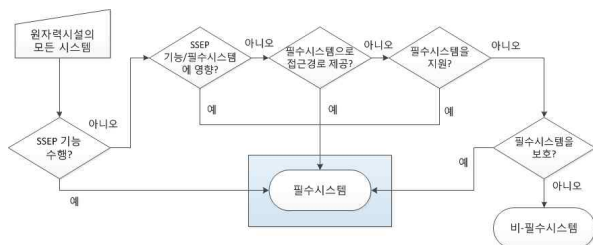


그림 1. 필수시스템 식별 절차  
Fig. 1. Evaluation process for determining critical systems

그림1에 따르면, SSEP 기능을 수행하는 시스템 뿐만 아니라, 의존성 분석을 통해 해당 기능 및 CS에 영향을 미치거나, 접근경로를 제공하거나, 지원하는 시스템, 또는 CS를 보호하는 시스템에 대해서도 CS로 식별하도록 요구하고 있다.

이에 따라, 우리나라에서 디지털시스템이 많은 원자력발전소를 대상으로 시스템의 의존성 분석 결과를 유향그래프(Directed Graph)를 활용해 시각화한 결과는 그림 2 및 그림 3과 같다.

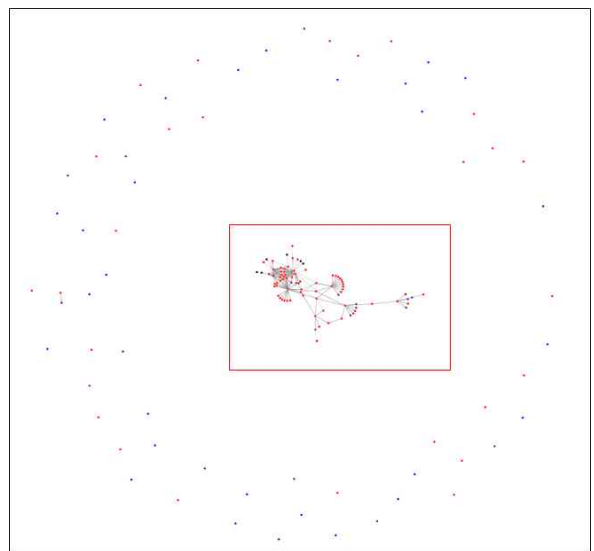


그림 2. 원자력발전소 모든 시스템에 대한 의존성 분석 시각화 결과

Fig. 2. Result on the visualization of dependency analysis of all critical systems at a nuclear power plant

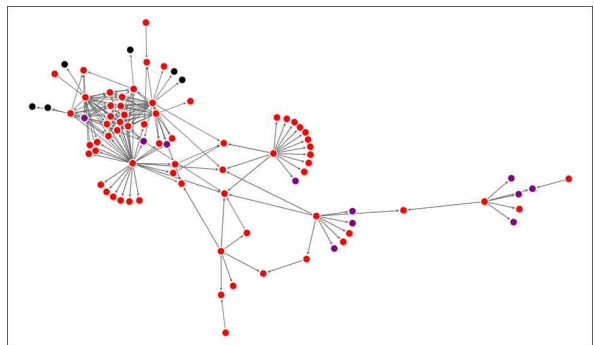


그림 3. 그림 2의 붉은색 사각형을 확대한 결과  
Fig. 3. A zoomed-in view of the red square in Fig. 2

그림 2 및 그림 3의 각 점(Vertex)들은 원자력발전소의 시스템을 의미하며, 붉은색은 그림 1에 따른

CS 식별 과정을 통해 CS로 식별된 시스템을, 파란색은 비 필수시스템(Non-CS)을, 보라색은 CS와 Non-CS 두 가지 특성으로 구성된 시스템을 나타낸다. 파란색으로 식별된 Non-CS는 그림 2에서 붉은색 사각형 바깥에 점으로 표시되었다. 검은색 점은 CS 식별 과정에서 오류로 인하여 시스템 이름이 잘못 작성되었거나 중복 작성된 사례를 구분하기 위해 별도로 표시하였다.

간선(Edge)들은 시스템들 간의 의존성 유무를 나타내며 A 시스템이 B 시스템에 영향을 주는 경우,  $A \rightarrow B$  형식으로 방향을 표시되었다. 간선을 가지고 있는 즉, 의존성을 가지는 시스템들은 그림 3에서와 같이 CS(붉은색), 또는 CS를 포함하고 있는 시스템(보라색)으로 식별되었다.

의존성 분석 결과를 진입 차수(In-degree)와 진출 차수(Out-degree)로 구분하여 상위 결과 일부를 내림차순으로 정리하면 표 1 및 표 2와 같다.

표 1. 진입 차수가 높은 필수시스템  
Table 1. CS with high in-degree order

필수시스템(CS)	진입 차수
발전소보호계통	10
정보처리계통	10
공학적안전설비 기기제어계통	9
공정기기제어계통	7
주요번수지시 및 경보계통-P	7

표 2. 진출 차수가 높은 필수시스템  
Table 1. CS with high out-degree order

필수시스템(CS)	진출 차수
전기 및 계측제어기기실 공기조화계통	54
터빈건물 공기조화계통	29
화재방호계통	26
필수전력공급계통	23
복합건물공기조화계통	18

진입 차수는 특정 점으로 들어오는 간선의 개수를 의미하며, 이에 따라 표 1은 특정 시스템에 영향을 주는 시스템의 개수를 보여준다. 즉, 여러 시스템에서 데이터를 받거나, 고장 및 침해에 따른 영향을 받을 수 있음을 의미한다.

진출 차수는 특정 점에서 나가는 간선의 개수를 나타내며, 표 2는 특정 시스템이 얼마나 많은 다른 시스템에 영향을 미치는지를 보여준다. 즉, 진출 차수가 높은 시스템에 대한 고장이나 침해는 그 만큼 많은 시스템에 악영향을 미칠 수 있음을 의미한다.

이와 같은 디지털시스템에 대한 의존성 분석은 사이버 보안조치가 필요한 시스템에 대한 우선순위를 고려하는데 도움이 될 수 있으며, 해당 시스템에 대한 사이버공격 및 침해 시 어떤 시스템에 영향이 전파될 수 있는지 손쉽게 파악할 수 있도록 도움을 준다.

다만, 이와 같은 의존성 분석을 통한 시스템의 식별은 사이버공격으로부터 보호가 필요한 시스템의 우선순위를 도출하고, 사이버공격에 따른 영향의 가시화에 도움 될 수 있지만, 어떤 보안조치를 적용해야 하는지를 설명해주진 않는다. 따라서, 의존성 분석 결과는 위험평가 시 고려되어야 하는 여러 요소 중 하나로 이해되어야 한다. 또한, 앞서 소개한 위험기반 보안조치 차등적용 방법론들을 함께 고려함으로써, 보안조치 적용이 필요한 시스템의 우선순위 식별과 함께 필요한 보안조치 목록들을 효과적으로 선별할 수 있을 것으로 기대된다[5,6].

### III. 결 론

본 논문에서는 디지털 시스템의 의존성 분석을 통해 한 시스템에 대한 사이버공격이 얼마나 많은 타 시스템에 악영향을 미칠 수 있는지 수치화함으로써, 강한 보안조치가 요구되는 중요한 시스템과 상대적으로 중요도가 낮은 시스템을 구분하는 위험기반 차등적 보안조치 방안을 제안하였다. 제안된 방법은 사이버공격으로부터 보호가 필요한 시스템의 우선순위를 도출하고, 사이버공격에 따른 영향을 가시적으로 보여줌으로써 사이버보안 침해 대응에 도움이 될 수 있다.

### 참 고 문 헌

[1] National Institute of Standards and Technology, "NIST Special Publication 800-53: Security and



- Privacy Controls for Information Systems and Organizations", USA, 2020.
- [2] National Institute of Standards and Technology, "NIST Special Publication 800-82: Guide to Industrial Control Systems (ICS) Security", USA, 2015.
  - [3] International Electrotechnical Commission, "IEC 62443: Industrial Communication Networks - Network and System Security for Industrial Automation and Control Systems", 2021.
  - [4] International Atomic Energy Agency, "IAEA Nuclear Security Series No. 33-T: Computer Security of Instrumentation and Control Systems at Nuclear Facilities", 2018.
  - [5] Nuclear Energy Institute, "NEI 13-10 Rev. 6: Cyber Security Plan for Nuclear Power Reactors", USA, 2016.
  - [6] Electric Power Research Institute, "Technical Assessment Methodology (TAM) for Cyber Security of Industrial Control Systems", USA, 2019.
  - [7] Korea Institute of Nuclear Nonproliferation and Control, "KINAC/RS-015: Regulatory Standard for Cybersecurity in Nuclear Facilities", South Korea, 2019.
  - [8] U.S. Nuclear Regulatory Commission, "Regulatory Guide 5.71: Cyber Security Programs for Nuclear Facilities", USA, 2010.

# 영역 분할 및 조밀 대응 관계를 활용한 비학습 기반 객체 단위 장면 변화 감지 기술

박한무\*

## Non-learning-based scene change detection utilizing region segmentation and dense correspondence

Han-Mu Park\*

### 요 약

본 논문에서는 임의의 시점에서 촬영된 두 장면 간의 변화를 감지하기 위한 비학습 기반 장면 변화 탐지 알고리즘을 제안한다. 제안된 알고리즘은 먼저 두 장면 사이의 조밀 대응 관계를 찾고, 영역 분할 결과에 따라 이러한 대응 관계를 그룹화하여 객체 단위의 변화 탐지가 가능하도록 한다. 제안된 알고리즘의 성능을 평가하기 위해 다수의 객체들이 포함된 여러 실내 장면을 수집하였으며, 제안된 알고리즘은 변경된 장면과 그 위치를 성공적으로 식별할 수 있었다.

### Abstract

In this paper, I propose a non-learning-based scene change detection algorithm for identifying changes between two scenes captured from arbitrary viewpoints. The proposed algorithm first establishes dense correspondences between the two scenes and then groups these correspondences based on area segmentation results, enabling object-level change detection. To evaluate the proposed algorithm, I collected several indoor scenes containing multiple objects, and the algorithm successfully identifies both the objects that have changed and their locations.

### Key words

change detection, segmentation, dense correspondence

## 1. 서 론

두 장면 사이의 변화된 환경을 찾는 기술은 컴퓨터 비전 분야에서 전통적인 연구 주제로 재난 상황

및 이상 상황 감지 등 감시 시스템을 위한 핵심 원천 기술이다. 또한, 최근에는 자율 주행, VR/AR 등의 분야에서도 능동적인 환경 인지를 위해 주변 환경에 대한 변화 감지 기술의 중요성이 높아지는 추

---

\*한국전자기술연구원, hanmu@keti.re.kr

※ This research was financially supported by the Ministry of Trade, Industry and Energy(MOTIE) and Korea Institute for Advancement of Technology(KIAT) through the International Cooperative R&D program. (No. P0019804, Digital twin based intelligent unmanned facility inspection solutions)

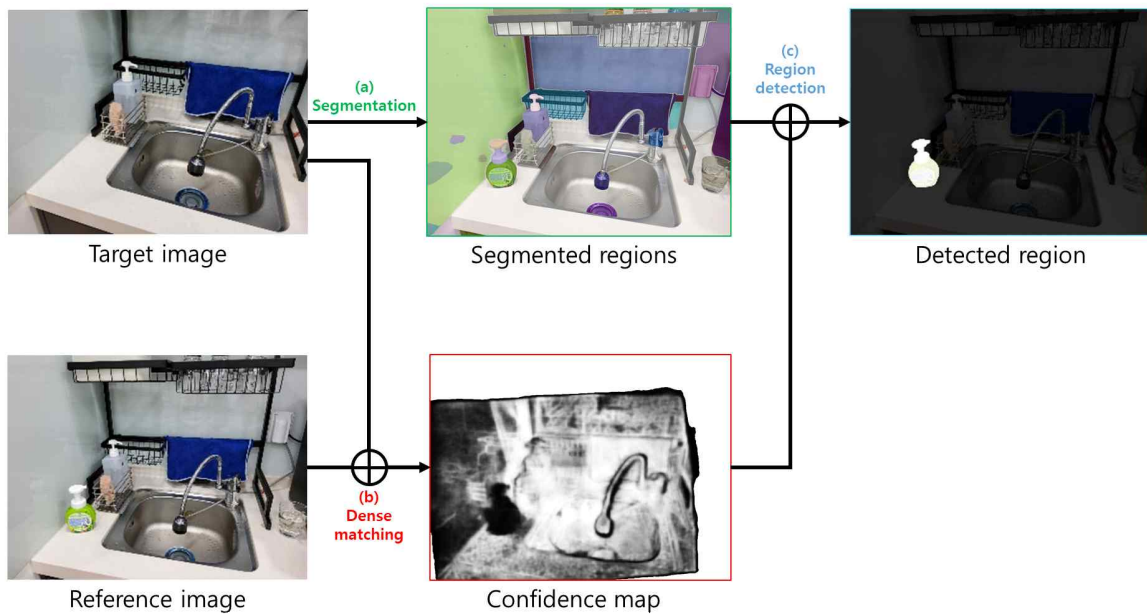


그림 1. 비학습 기반 환경 변화 검출 알고리즘  
Fig. 1. Non-learning-based scene change detection algorithm

세이다. 환경 변화 감지 기술이 적용되어야 하는 응용 시스템들은 조명, 촬영 시점 등이 변화하거나 나무 등 비강체 객체들이 혼재되어 있는 환경이 많아 이러한 미세한 변화에는 강건하면서 변화가 발생한 영역만을 정확하게 찾아내는 것 방법을 고려하는 것이 필수적이다. 이를 위해 기존 연구에서는 객체 검출 기술을 활용하거나[1], 다중 시점 영상들을 취합하여 3차원 공간을 복원하고 변화 영역을 검출하는 등의 방법을 사용하였다[2]. 그러나 이러한 방식은 사전에 학습된 객체들에 대해서만 검출이 가능하거나 너무 많은 시점의 입력 영상을 필요로 하기 때문에 실제 응용 시스템에 활용하기에 어려운 부분이 있었다.

본 논문에서는 이러한 문제점을 해결하기 위해 비학습 기반의 환경 변화 감지 알고리즘을 제안한다. 제안하는 방법은 두 장의 이미지만으로 변화가 발생한 영역을 검출할 수 있으며, 학습 데이터에 독립적으로 변화 영역을 검출하는 것이 가능하다.

## II. 비학습 기반 환경 변화 감지 기술

본 논문에서 제안하는 알고리즘은 그림 1과 같이 조밀 대응 관계 검출과 영역 분할 파트의 2단계로

구성되어 있다.

조밀 대응 관계 검출 기술은 두 이미지 사이의 정합 관계를 픽셀 단위로 검출하는 기술로 픽셀 단위의 신뢰성을 검출할 수 있다. 본 연구에서는 딥러닝 기반의 Dense Kernelized Feature Matching (DKM) [3] 알고리즘을 활용하였다. DKM 알고리즘은 그림 1(a)와 같이 시점 변화에 강건하면서 객체나 장면에 구애받지 않고 정합 관계를 구하는 것이 가능하다. 다음으로 영역 분할 파트에서는 각 장면을 딥러닝에 기반하여 동일 객체 또는 파트로 판단되는 영역들로 분할한다. 이를 위해 본 논문에서는 Segment Anything (SAM)[4] 알고리즘을 활용하였으며, 그림 1(b)와 같이 각 파트별로 영역이 분할되는 것을 확인할 수 있다. 마지막으로 변화가 발생한 영역을 검출하기 위해 SAM을 통해 검출된 각 영역의 신뢰도 값 통계를 분석함으로써 정합되지 않는 영역을 그림 1(c)와 같이 검출하게 된다. 이 때 두 장면 사이에 겹쳐지지 않는 경계 영역들이 존재할 수 있는데 해당 영역들에 대해서는 전체 영역에 대한 신뢰도 값을 정규화한 후 임계값을 설정하여 처리하였다.

## III. 실험

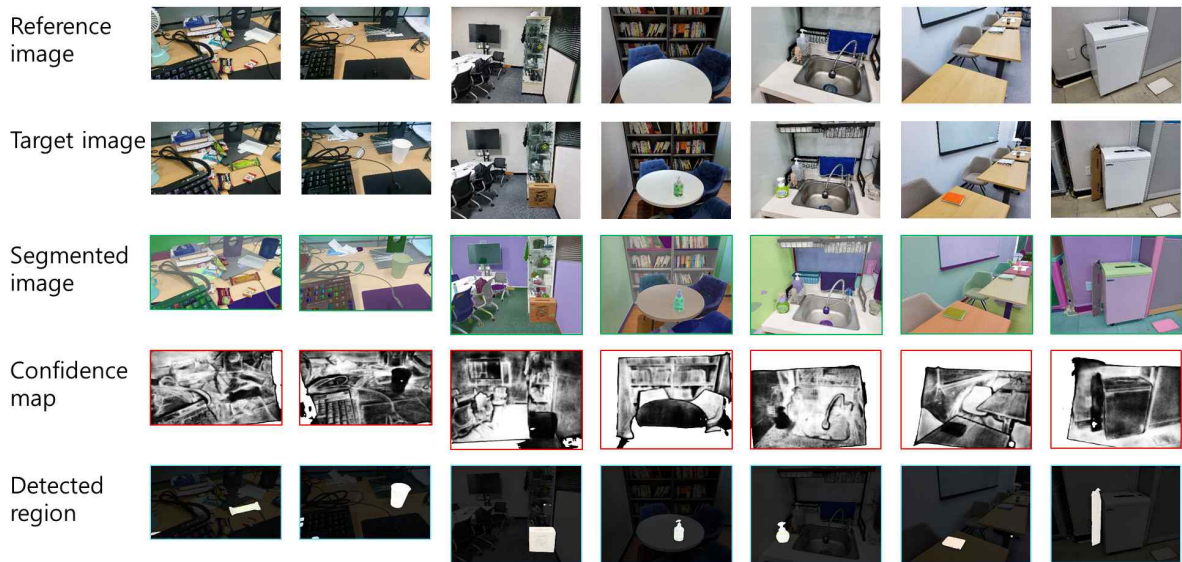


그림 2. 환경 변화 검출 결과  
Fig. 2. Environmental change detection results

본 논문에서 제안한 알고리즘의 성능을 분석하기 위해 그림 2와 같이 다수의 객체들이 등장하는 실내 환경을 촬영하였으며, 참조 이미지와 다른 시점에서 목적 이미지를 촬영하되 새로운 객체가 한 개 이상 등장하도록 장면을 구성하였다. 실험 결과 그림 2와 같이 시점 변화가 크고 형체가 다양한 객체들에 대해서도 검출하는 것에 성공하였다.

#### IV. 결 론

본 논문에서는 비학습 기반의 환경 변화 검출 알고리즘을 제안하였다. 영역 분할 기술과 조밀 대응 관계 탐색 기술을 활용함으로써 임의의 형체를 가진 객체를 시점 변화가 크더라도 검출하는데 성공하였다. 제안하는 알고리즘은 향후 자율주행, VR/AR 등 능동적인 환경 인지가 필요한 분야에 활용이 가능할 것으로 기대된다.

#### 참 고 문 헌

[1] K. Doi, R. Hamaguchi, S. Iwase, R. Yokota, Y. Matsuo, and K. Sakurada, "Epipolar-Guided Deep Object Matching for Scene Change Detection", arXiv, 2007.

[2] R. Huang, B. Jiang, Q. Zhao, W. Wang, Y. Zhang, and Q. Guo, "C-NeRF: Representing Scene Changes as Directional Consistency Difference-based NeRF", arXiv, 2023.

[3] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson, E. Mintun, J. Pan, K. V. Alwala, N. Carion, C.-Y. Wu, R. Girshick, P. Dollár, and C. Feichtenhofer, "SAM 2: Segment Anything in Images and Videos", arXiv, 2024.

[4] J. Edstedt, I. Athanasiadis, M. Wadenbäck, and M. Felsberg, "DKM: Dense Kernelized Feature Matching for Geometry Estimation", CVPR, 2023.

# 식물공장 양액공급시스템의 모니터링에 관한 연구

배 장 식\*

## A Study on the Nutrient Solution System Monitoring for Plant Factories

Bae Jang Sik\*

### 요 약

본 논문은 임베디드 시스템 기반의 식물공장 양액공급 시스템 모니터링 기능개발을 보여준다. 기존의 시스템은 시스템 자체의 디스플레이 혹은 전용 프로그램을 통하여 시스템의 상태를 모니터링 할 수 있었으며 이러한 부분을 해결하기 위하여 본 논문에서는 웹기반의 모니터링 기능을 구현하였다. 구현된 모니터링 기능은 실제 식물공장에 적용하여 기능 및 안정성을 확인하였다.

### Abstract

This paper demonstrates the development of a monitoring function for a nutrient solution supply system in a plant factory based on an embedded system. In the existing system, the status could be monitored through the system's own display or a dedicated program. To address this limitation, this paper implements a web-based monitoring function. The developed monitoring function was applied to an actual plant factory to verify its functionality and reliability.

### Key words

smartfarm nutrient solution supply system, embedded system, monitoring, plant factory

### 1. 서 론

본 논문에서는 식물공장에 필요한 양액공급시스템의 모니터링 기능 구현에 관하여 보여주고 있다. 시설농업 분야에서 양액공급시스템은 필수적인 요소로 이를 효율적으로 활용하기 위하여 외부에서의 모니터링 기능이 요구되고 있으며 이를 해결하기 위하여 기존의 시스템들은 전용 프로그램을 이용한 PC기반의 모니터링 기능을 지원하고 있다[1][2]. 본

논문의 모니터링 기능은 라즈베리파이를 이용하여 상용 양액공급시스템과의 통신을 위하여 RS-485기반의 modbus 통신을 사용하였다[3]. 수집된 양액공급시스템의 정보는 플로우 기반의 웹 UI도구인 Node-RED를 이용하여 개발된 웹 기반의 모니터링 기능을 통해 사용자에게 제공하였다[4]. 개발결과물은 실제 운영중인 식물공장에 적용하여 기능 및 신뢰성을 테스트하였으며 4개월 이상의 기간 동안 큰 문제없이 기능이 제공되는 것을 확인하였다. 본 논

\* 한국폴리텍대학, jsbae@kopo.ac.kr

문의 구성은 2장에서 시스템 구현의 세부적인 사항에 대하여 기술하고 3장의 결론에서 테스트결과 및 향후 연구방향에 대하여 기술한다.

## II. 본 론

본 논문에서는 구현한 모니터링 시스템의 구성은 아래 그림1에서 보여주고 있다. RS-485통신은 시리얼통신 기반의 멀티드롭 방식으로 다수의 제어시스템의 연결이 가능하여 스마트팜 분야에서 널리 사용되고 있는 통신방식이며, modbus통신은 PLC를 제어하기 위한 통신 프로토콜로 다양한 제어시스템에서 사용되고 있다[5][6]. 본 논문에서는 상용 양액공급 시스템의 터치디스플레이와 컨트롤보드 간의 RS-485 기반 modbus통신 데이터를 라즈베리파이 장치를 통해 캡처 후 분석하여 해당 정보를 웹 기반의 모니터링 서비스를 통해 사용자에게 제공하도록 하였다.

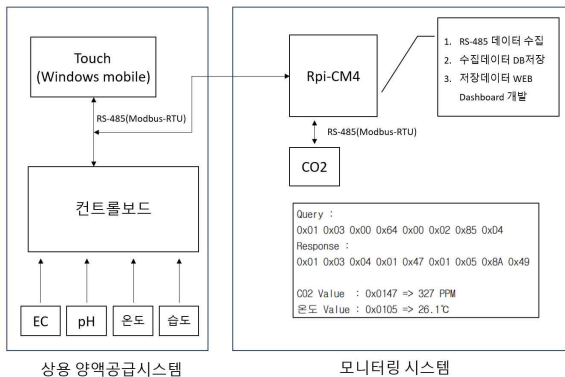


그림 1. 시스템 구성도  
Fig. 1. System diagram

모니터링 대상인 EC, pH, 온도, 습도 데이터는 양액공급시스템의 컨트롤보드에서 터치디스플레이로 modbus 통신을 통해 전송된다. 외부의 라즈베리파이 장치를 이용하여 양액공급시스템의 modbus 통신을 캡처 후 해당 데이터를 Node-RED기반의 웹 기반의 모니터링 시스템을 통해 사용자에게 제공하도록 한다. 데이터는 modbus function 3을 이용하여 전송되며 아래 그림 2는 modbus 데이터의 메시지 구조 및 양액공급시스템 구성요소 사이의 데이터 전송구조를 나타내고 있다.

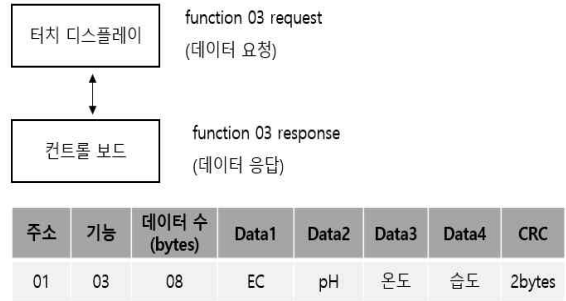


그림 2. modbus 메시지 구조  
Fig. 2. modbus message structure

Node-RED는 웹기반 시각 프로그래밍을 위한 플로우 기반 개발 도구로 노드로 표현되는 기능들을 와이어로 연결하여 동작 하도록 되어 있으며 개발되어 있는 다양한 노드를 활용하여 개발에 활용할 수 있다. 아래 그림3은 양액공급시스템의 modbus 데이터를 분석하여 EC, pH, 온도, 습도 데이터를 분리 후 Mysql 데이터베이스로의 저장 기능과 대시보드를 구성하기 위한 Node-RED 플로우를 보여준다.

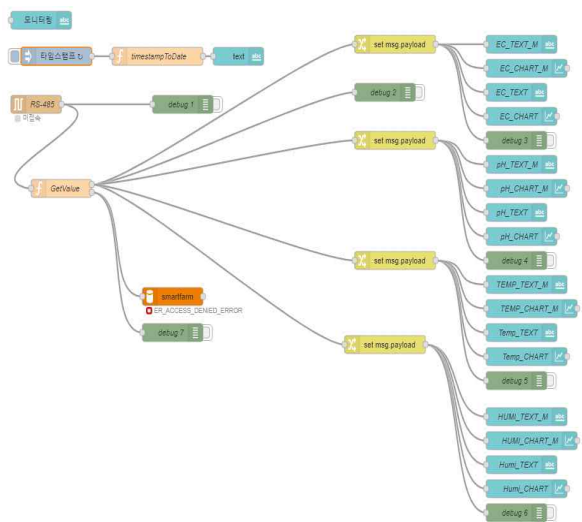


그림 3. 개발된 Node-RED flow 기능  
Fig. 3. Node-RED flow function

그림 4는 사용자에게 제공되는 웹기반의 모니터링 UI로 Node-RED의 대시보드 노드들을 활용하여 개발되었으며 PC와 모바일로 구분되어 구성하였다. PC는 데이터 조회 등의 모든 기능을 지원하며 모바일은 그래프 형태의 데이터 추이를 확인 할 수 있도록 하였다.

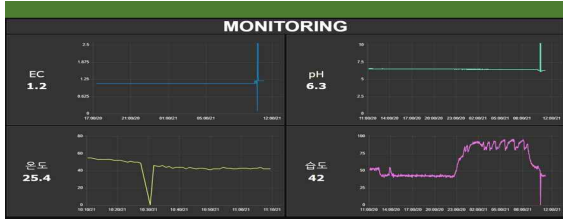


그림 4. 양액공급시스템 모니터링 대시보드  
Fig. 4. Node-RED based dashboard

개발 결과물은 현재 운영 중인 컨테이너 형태의 식물공장에 적용하여 테스트를 진행 중이며 1분간격으로 데이터를 수집하여 데이터베이스에 저장하였다. 표 1은 4개월 동안의 데이터 전송결과 분석을 나타낸다. 4개월 테스트 결과 1% 미만의 데이터 누락이 발생한 것으로 확인 되었다.

표 1. 데이터 전송결과 분석  
Table 1. Data transmission result

	3월	4월	5월	6월
전송데이터	43,200	44,640	43,200	44,640
저장데이터	43,187	44,499	43,101	44,512
전송률(%)	99.9	99.6	99.7	99.7

아래 그림 5는 개발 결과물의 테스트를 위한 식물공장 및 상용 양액공급시스템을 나타낸다.



그림 5. 테스트베드 환경  
Fig. 5. Testbed environment

### III. 결 론

본 논문에서는 식물공장에 필수적인 양액 공급 시스템의 모니터링 기능을 구현하였다. 이 기능은 상용 시스템의 MODBUS 통신을 캡처하고, 플로우 기반 웹 개발 도구인 Node-RED를 활용하여 모니터링 대상인 EC, pH, 온도, 습도 데이터를 분리한 후, 이를 데이터베이스에 저장하고 사용자에게 대시보

드를 통해 정보를 제공하였다. 결과물의 검증을 위하여 컨테이너 형태의 식물공장에 적용 후 4개월 이상의 테스트를 진행 중이었으며 시리얼 통신의 오류로 인한 데이터 누락 등의 문제가 발생하지만 기능적으로 문제가 없음을 확인하였다. 향후 한 대 이상의 양액공급시스템에 적용에 필요한 기능에 대한 연구를 진행할 예정이다.

### 참 고 문 헌

- [1] 황성일, 주종문, and 주성용. "수경인삼 식물공장 사례를 통한 ICT 기반 스마트 팜 팩토리 시스템." 한국통신학회논문지 40.4 (2015): 780-790.
- [2] 김관형. "스마트팜 모니터링 시스템 개발." 한국정보통신학회 종합학술대회 논문집 25.1 (2021): 287-288.
- [3] Thomas, George. "Introduction to the modbus protocol." The Extension 9.4 (2008): 1-4.
- [4] Tabaa, Mohamed, et al. "Industrial communication based on modbus and node-RED." Procedia computer science 130 (2018): 583-588.
- [5] 이명훈, and 여현. "빅데이터 기술을 활용한 스마트 팜 단계별 분석." 한국통신학회 학술대회 논문집 (2015): 71-72.
- [6] Jeong, Yunju, et al. "PLC and arduino interaction based on modbus protocol." Journal of Korea Multimedia Society 20.3 (2017): 511-519.





# 음성 정보 기반의 감정 추적을 통한 상담 만족도 예측 기법

이건수\*, 김찬호\*\*, 좌성훈\*\*\*

## A Method for Predicting Counseling Satisfaction through Emotion Tracking Based on Vocal Information

Keonsoo Lee\*, Chanho Kim\*\*, and Seonghoon Jwa\*\*\*

### 요 약

의사소통은 언어적 요소, 비언어적 요소, 그리고 청각적 요소로 구성된다. 본 연구에서는 대화 중 발화자들의 감정 상태를 각 발화의 청각적 요소만을 이용해 추적하고 이를 통해 대화 종료 후 참가자들의 만족도를 예측할 수 있는 방법을 제안한다. 청각적 요소를 통해 감정 요소 중 각성 속성의 변화를 추적하고, 종료 후 각성 상태의 변화 패턴을 통해 만족도를 예측하였다. 만족도에 대한 척도에 따라 다른 정확도를 보이지만, 만족/불만족에 대한 이진 만족도의 경우, 95.3% 분류가 가능했다.

### Abstract

Communication consists of verbal elements, nonverbal elements, and auditory elements. This study proposes a method for tracking the emotional states of participants during conversations using only the auditory elements of each speaker and predicting participants' satisfaction after the conversation. By tracking changes in arousal attributes through auditory elements, we were able to predict satisfaction based on the patterns of change in arousal states after the conversation. Different accuracy levels were observed depending on the satisfaction scale, but in the case of binary satisfaction (satisfied/dissatisfied), 95.3% classification accuracy was achieved.

### Keywords

Communication, Emotional Tracking, Auditory Elements, Satisfaction Prediction, Arousal Attributes

## 1. 서 론

커뮤니케이션은 비록 요소별 중요도 등에 대한 이견은 있지만, 크게 언어적 요소, 비언어적 요소, 그리고 청각적 요소로 구분할 수 있다[1]. 대화를 올바르게 이해하기 위해서는 모든 요소들 종합적으로 참조해야 하지만, 대화가 이루어지는 환경에 따

라, 일부 요소만 접근할 수 있는 상황이 발생한다. 가령, 채팅이나 편지의 경우 언어적 요소만이 사용되고, 전화의 경우 언어와 청각 요소만 사용된다. 특히, 서로 다른 언어를 사용하는 사람들이 전화로 통화하는 경우, 오로지 청각 정보만 사용할 수 있지만, 그럼에도 불구하고 그 대화가 어떻게 흘러가고 있는지를 알 수 있는 경우가 존재한다[2].

\* C&AI Research Lab, keonsoo.lee@cnae.kr(교신저자),

\*\* C&AI Research Lab, chanho.kim@cnae.kr

\*\*\* SINSIWAY Human Resources Training Team, junsan12@sinsiway.com

콜센터의 경우, 상담이 진행된 이후, 상담에 대한 품질평가가 진행된다. 비용상의 문제로 샘플 평가를 수행되지만, 전수 평가의 필요성 및 우월성은 항상 존재한다. 이에 본 연구에서는 청각 정보만을 사용해 상담 대화에서의 대화 참가자들, 즉 상담사와 상담 고객의 감정 변화를 추적해 상담 만족도를 예측할 수 있는 방법을 제안한다.

## II. 음성 정보로부터 발화자 감정 추출 방법

감정을 표현하기 위한 기본 모델은 긍/부정으로 표현되는 유쾌함(Valence)과 각성(Arousal)의 2차원 공간으로 표현할 수 있다[3].

표 1. 러셀의 감정 모델  
Table 1. Russell's Emotion Model

		각성	
		각성	비각성
유쾌함	유쾌	기쁨, 행복함	편안함, 졸림
	불쾌	놀람, 짜증남	슬픔, 지침

음성 정보로부터 추출할 수 있는 정보들은 크게 목소리 정보와, 발화 상황 정보로 구분할 수 있다. 목소리 정보는 피치(Pitch), 지터(Jitter), 심머(Shimmer), 에너지(Energy), 스펙트럼 평탄도(Spectral Flatness), 포먼트(Formant), 그리고 MFCC 등이 있다.

표 2. 음성 정보의 속성들과 감정 요소와의 관계  
Table 2. The relationship between the attributes of vocal information and emotional elements

속성	의미	각성도 상관성
피치	음성의 높낮이	+
지터	음성의 주기적 불규칙성	+
심머	음성의 진폭 변동성	+
에너지	음성 신호의 강도	+
스펙트럼 평탄도	음성신호의 하모닉스와 노이즈의 비율	+
포먼트	음질 정보로 F1, F2 포먼트의 주파수 변동의 방향성	+
MFCC	음색과 음질 속성으로 고주파수 대역의 에너지 분포 비율	+

목소리 정보는 최종적으로 각성도를 판별하기 위해 사용되며, 스펙트럼 평탄도의 경우 노이즈의 비율이 높을수록, 포먼트의 경우 상위 주파수로 이동할수록, MFCC의 경우, 고주파수 대역의 에너지 분포 비율이 높을수록 높은 각성도를 의미한다. 각각의 속성 결과를 통해 각성도가 높아졌는지, 혹은 낮아졌는지, 또는 유지되었는지를 평가하고 그 결과를 투표를 통해 취합한다.

발화 상황을 알 수 있는 정보로는 발화의 길이, 빈도, 발화 겹침, 무발화 구간, 그리고 발화간 간격 등이 있다. 이들 요소는 최종적으로 발화 교대(Turn-Taking)의 원활함을 측정하기 위한 요소로 작용한다. 즉 일방적으로 한 쪽만 말을 많이 했는지, 누군가가 말을 끊었는지, 서로가 아무말도 하지 않고 있는 상황인지를 확인하며, 두 대화 참여자가 동등한 분량의 대화를 순서대로 겹침과 공백 없이 대화하는 이상적인 상황에서 몇 개의 기준이 침해되었는지에 대한 침해 비율로 표기한다. 이때 사용하는 침해 기준은 기준이 되는 데이터를 토대로 사전 학습을 통해 결정한다.

## III. 감정 기반 대화 흐름 판별 방법

목소리 정보는 발화 교대가 일어나는 순간 계산되어 개별 발화자가 자신의 발화 순서에서의 각성도의 변화를 볼 수 있으며, 또한 상담 대화의 경우 대화는 크게 3부분으로 구분할 수 있다. 서론부의 경우 인사말 등의 교환단계로 10초 내외가 소요된다. 전개부의 경우, 문제를 설명하고 그 해결 방법을 찾는 부분으로 대화의 대부분을 포함한다. 결론부의 경우, 추가 질문의 여부를 확인하고 종료하는 부분으로, 상담이 잘 해결됐다면, 10초 내외로 진행된다. 상담이 실패한 경우, 존재 자체가 없을 수도 있다. 이를 통해 대화에서의 감정 변화 상태는 총 4 단계(서론, 전개, 결론, 종합)에서, 각 단계별로 각성도는 상승, 유지, 하강, 진동의 4가지 값을 갖는다. 이때, 발화의 언어적 의미를 알 수 없기 때문에, 각 단계는 고정 시간으로 사용하며, 고정 시간은 추후 변경할 수 있다.

발화 상황는 전체 대화를 대상으로 계산되며, 발

화 상황을 구성하는 조건들의 이진 상태 조합으로 계산한 뒤, 목소리 정보와 발화 상황에 대한 가중치 결합으로 최종 값을 산정한 뒤, 기준 값으로 만족/불만족을 평가한다.

이상의 과정은 대화의 도메인에 따라 각기 다른 기준값을 갖게 되기 때문에, 일반적인 학습과 추론의 과정을 거치며, 학습과정에서는 도메인 데이터를 기반으로 감정 변화 패턴의 종류와 각 패턴별 만족도의 확률표와 발화 상황 정보의 프레임 구조를 변경해야 한다. 이후, 두 요소 사이의 가중치를 계산하고, 이 가중평균의 결과를 최종 판별할 수 있는 기준 값을 생성해야 한다. 상기한 프로세스의 흐름은 다음 그림과 같다.

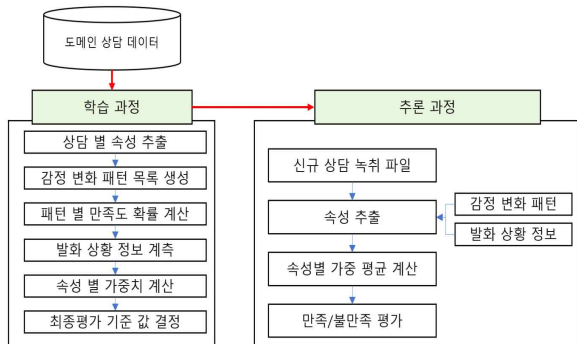


그림 1. 감정 기반 대화 만족도 판별 프로세스  
Fig. 1. Emotion-Based Conversation Satisfaction Determination Process

#### IV. 실험

제안 방법을 검증하기 위해 평균 3~5분의 시간으로 진행된 상담 녹취록 150개와 상담 후 상담 고객의 만족도를 협력사를 통해 확보했다. 만족도는 7점 리커트 척도로 수집되었고 본 연구에서는 만족, 불만족, 중립의 3단계로 병합하여 사용하였다. 이들 데이터 중, 중립 만조도 값을 받은 7개 데이터를 제외한 143개 중 70%인 100개를 선택해 감정 변화 패턴 및 가중치를 추출하였고, 이 추출 결과를 대상으로 43개에 대한 만족도 예측을 실험하였다.

학습 과정에서, 불만족에 대해 2개, 만족에 대해 5개로 구성된 총 7건의 패턴이 발견되었으며 요소별 가중치는 목소리 정보를 70%, 발화 상황을 30%로 계산한 뒤 0.65를 기준으로 만족과 불만족을

분류하였다. 그 결과 최종적으로 43개 중 41건에 대해 분류에 성공하였다. 오분류된 2건의 경우, 발화 상황에서 높은 각성도를 보였음에도 만족도가 높았던 예외 상황이었다는 것으로 분석되었다.

#### V. 결 론

콜센터 환경에서 상담 고객의 대화 만족도를 예측함에 있어서, 대화의 음성 정보만 사용하는 경우, 언어 정보를 해석하기 위한 계산 복잡도와 비용을 없앨 수 있기에 보다 효율적인 상담 품질 평가를 수행할 수 있다. 음성 정보만을 사용하는 경우, 감정을 정의하기 위한 유쾌함과 각성도를 모두 찾아낼 수는 없지만, 발화의 안정성과 발화 패턴을 통해 각성의 정도를 인지할 수 있고, 그 결과 상담 대화가 종료되었을 때, 상담 고객의 만족도를 예측할 수 있었다.

비록 만족/불만족의 정도를 예측하지는 못하지만, 만족/불만족의 판별은 실험 상황에서 96%의 정확도를 보였다. 음성 정보 이외의 대화 요소를 추가로 분석한다면, 보다 정확하고 상세한 분류가 가능할 것으로 예상된다.

마지막으로, 콜센터에서의 상담 만족도는 상담 고객을 대상으로 수집되기에, 콜센터의 상담 품질의 결과가 상담 고객의 만족에만 한정할 것이 아니라, 같이 대화에 참여한 상담사의 만족도 역시 포함할 수 있어야 할 것으로 사료된다.

#### 참 고 문 헌

- [1] D. Lapakko, "Communication is 93% nonverbal: An urban legend proliferates", *Communication and Theater Association of Minnesota Journal*, 34(1), 2, 2007
- [2] J. M. Dewaele, P. Moxsom-Turnbull, "Visual cues and perception of emotional intensity among L1 and LX users of English", *International Journal of Multilingualism*, 17(4), pp. 499-515, 2020.
- [3] J. A. Russell, "A circumplex model of affect", *Journal of personality and social psychology*, 39(6), 1161, 1980.

# 언어 정보를 통한 상담 대화 내 감정 변화 추적 방법

이건수\*, 김찬호\*\*, 좌성훈\*\*\*

## A Method of Tracking Emotional Changes in Counseling Conversations Through Textual Information

Keonsoo Lee\*, Chanho Kim\*\*, and Seonghoon Jwa\*\*\*

### 요 약

채팅을 통해 이루어지는 상담 대화의 경우, 일반적인 의사소통 과정에서 사용되는 비언어적 표현과 목소리에 대한 정보를 사용하지 못하고 오롯히 언어 정보만을 통해 대화가 이루어진다. 본 연구에서는 채팅 문장의 어휘, 맞춤법, 표현의 논리, 그리고 문장 발생 시간의 시간 및 문장 길이를 토대로 대화 참여자들의 감정 변화를 추적하기 위한 방법을 제안한다. 감정 변화의 추적은 상담 도메인에 독립적으로, 참여자들의 대화 만족도를 평가하는데 활용될 수 있다.

### Abstract

In chat-based counseling conversations, communication relies solely on textual information, as non-verbal expressions and vocal nuances typically used in general communication are unavailable. This study proposes a method to track the emotional changes of conversation participants based on the vocabulary, spelling, logical structure of expressions, the timing of sentence generation, and sentence length in chat messages. The tracking of emotional changes can be utilized independently of the counseling domain to assess participants' satisfaction with the conversation.

### Keywords

Chat-based Counseling, Emotional Tracking, Textual Information, Participant Satisfaction

## 1. 서 론

대화는 다양한 상황에서 다양한 목적을 위해, 2명 이상의 참가자가 자신의 생각이나 의견을 언어적 요소와 비언어적 요소를 포함해 상호작용하는 행위를 의미한다[1]. 컨택센터에서 이루어지는 채팅

을 통한 상담 대화의 경우, 그 대화의 범위는 좀 더 축소되는데, 문제를 갖고 있는 상담고객과 그 문제를 해결하려고 하는 상담사가 대화의 참여자이며, 상담 고객의 만족과 상담사의 성취감이 대화의 목적이며, 채팅의 경우, 비언어적 요소가 배제된다는 것이 그것이다[2]. 이처럼, 명확한 목적이 존재하는

\* C&AI Research Lab, keonsoo.lee@cnaikr(교신저자),

\*\* C&AI Research Lab, chanho.kim@cnaikr

\*\*\* SINSIWAY Human Resources Training Team, junsan12@sinsiway.com

상대 대화의 경우, 그 대화의 성공 여부를 대화 종료시의 대화 참여자들의 감정 상태를 통해 판별할 수 있고, 대화 중간에도 대화의 진행 상황이 올바른 방향으로 진행되고 있는지를 평가할 수 있다. 이에 본 연구에서는 채팅 문장을 토대로 대화 컨텍스트 안에서의 감정 변화를 평가하고 추적할 수 있는 방법을 제안한다.

## II. 언어 정보로부터 감정 정보 변경 방법

본 연구에서는 대화 참여자의 감정을 긍정(Positive), 중립(Neutral), 그리고 부정(Negative)의 3단계로 정의하였다. 그리고 대화 참여자의 감정은 자신 혹은 상대의 발화가 종료되는 순간을 기준으로 감정을 평가하였다.

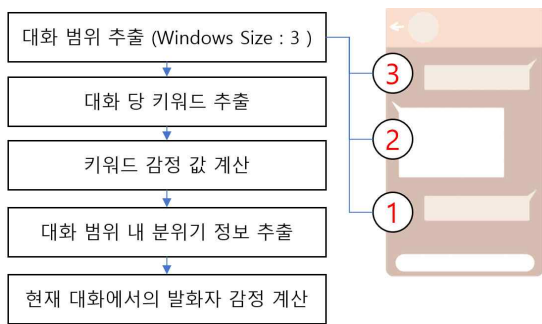


그림 1. 현재 발화자의 감정 추출 프로세스  
Fig. 1. Current Speaker's Emotion Extraction Process

우선 감정을 평가하기 위한 대화 윈도우로 현재 대화에서 상대의 순서 이전에 자신이 대화까지 총 3개의 발화 범위를 정한다. 즉 현재 대화에서의 발화자의 감정은 자신이 이전에 발화했던 내용, 상대방의 발화 내용, 그리고 지금 자신이 발화하는 내용을 토대로 만들어진다. 자신의 순서에서 발화의 길이는 중요하지 않고, 상대에게로 대화의 순서가 넘어가는 것을 발화의 종료로 한다. 이때, 감정 평가를 위한 윈도우를 이전 단계로 한정하는 까닭은 대화를 하는 순간의 감정은 과거 감정으로만 영향을 받기 때문이다.

감정은 분위기 속에서 영향을 받으며, 분위기에 대한 정의는 명확하게 정의되어 있지는 않지만, 일반적으로 명확하고, 즉각적이고, 강렬한 정서 반응

을 감정이라고 하고, 모호하고, 장기적이며, 일반적인 정서 반응을 분위기라고 한다[3]. 본 연구에서는 발화 안에서 보여지는 형태(Morphological), 구문(syntactic), 의미(semantic)의 언어 정보를 감정으로, 또 화용(pragmatic) 정보를 분위기로 구분하여 사용한다.

형태 정보의 경우, 발화 안에서 주요 키워드를 추출한 뒤, 해당 키워드의 감정을 계산한다. 키워드는 형용사, 명사, 동사, 감탄사, 부사를 추출하여 사용한다. 이때, 감정 표현 부호도 특수 키워드로서 같이 사용한다. 감정 표현 부호의 경우, 기 수집된 채팅 상담 데이터 1993개에서 도출한 상위 발생 빈도 표현 10가지(ㅋ, ㅎㅎ, ^^, ㅋㅋ, ㄱㄱ, --, ㅎㅎ, 0, ㄱㅅ, ㅅㄱ)와 그 변종(^^ → ^^, ^^, ^^ 등)에 대한 감정 상태를 부여했다. 또한 키워드에 대한 감정 상태의 경우, chatGPT, Gemini, Claude의 상용 LLM 모델을 사용해 도출된 값의 평균을 이용하였다. 이는 일반 공개모델의 경우, 한국어 모델에 대한 추가 학습을 수행하지 않는 이상, 결과에 대한 신뢰가 부족하기에 상용 LLM 모델을 결합하여 사용하였다.

구문 정보의 경우, 문장 구조의 변화와 구조를 통해 강조되는 키워드의 존재 유무를 확인하고, 강조된 키워드의 감정 상태에 가중치를 부여하는 역할을 한다. 이때, 일반적인 구분 분석 과정에서 도출할 수 있는 문법 오류나 오타의 경우, 분위기 분석을 위해 사용한다. 즉, 문법 오류가 발생했을 때, 그것을 알고 틀리는 것인지 몰라서 틀리는 것이냐가 중요하며, 오타의 경우도 알고 틀리는 것인지 몰라서 틀리는 것이냐가 중요하다. 몰라서 틀리는 경우는 그것이 발화자의 감정 상태와 연결되어 있지 않기 때문이며, 알면서 틀리는 경우는 감정적으로 격해진 경우로 해석할 수 있기 때문이다. 감정적인 흥분 상태는 화용 정보를 통한 분석 대상이며, 분석 범위 역시 화용 정보 분석 범위에서 수행되어야 하기에, 본 정보는 화용 정보 분석에서 담당한다.

의미 정보는 발화의 논리적 일관성을 판별하는데 사용한다. 표현의 논리는 화용 정보를 통해서 확인할 수도 있지만, 화용 정보를 통한 논리 판별을 위해서는 보다 복잡한 언어 이해 능력을 필요로 하기 때문에, 본 연구에서는 의미 정보에 기반한 논리적

일관성을 판단한다.

이 세가지 정보를 통해 만들어지는 감정은 형태 정보를 통해 대상 키워드를 추출하고, 구문 정보를 통해 가중치를 변경하며, 의미 정보를 통해 부정 감정의 강도를 보정한다. 의미 정보가 부정 감정에서만 사용되는 까닭은, 수집된 채팅 대화 상황을 토대로 확인했을 때, 부정 이외의 감정 상태에서 논리성이 결여되는 경우가 발생하지 않고, 논리성이 결여되는 경우, 반드시 감정은 부정으로 향하게 되기 때문이다.

화용 정보의 경우, 분위기를 생성하기 위해 사용되며, 발화 텍스트 이외의 정보를 주로 사용한다. 문장 발화의 시간, 문장 전송 간 시간, 발화 교대(Turn-taking), 그리고 구문 정보의 예외사항이었던 문법 오류와 오타 정보를 사용한다. 이전의 형태, 구문, 의미 정보를 사용할 때는 해당 발화 순서의 텍스트 정보만 사용했지만, 화용 정보를 추출하기 위해서는 대화 윈도우의 모든 발화를 사용한다.

문장 발화의 시간은 문장이 발화되는 시간의 변화를 기준으로 판단한다. 문장 사이의 시간은 대화 참여자가 자신의 발화를 얼마나 짧게 혹은 길게 보내는지를 판별한다. 가령, 자신의 발화 순서에 5개의 문장을 연속으로 하나씩 보냈는지, 혹은 하나의 장문으로 보냈는지에 대한 차이를 구분하기 위함이다. 발화 교대의 경우, 한 쪽이 일방적으로 말을 많이 하든지, 혹은 말끊기나 끼어들기 등의 발화 방해가 존재하는지를 확인하기 위함이다. 문법 오류 및 오타의 경우 해당 오류에 대한 지식을 알고 있는지가 중요하며, 이를 판단하기 위해, 대화 내내 동일 오류가 반복되는지를 판단한다. 전에는 오류나 오타가 없었지만, 현재 발생했다면, 흥분 상태임을 인지할 수 있다.

이들 네가지 종류의 정보는 발화자의 각성(Arousal)과 차분함(Calmness)을 파악하기 위한 정보이며, 이 정보를 통해서, 발화자의 감정 변화의 변화도를 변경한다. 이것은 높은 각성도 만큼 높은 차분함은 감정 변화의 정도를 동일한 수준으로 변경하기 때문이다.

이렇게 획득한 모든 정보를 결합하여 현재 발화 상황에서의 발화자의 감정상태를 결정하기 위한 규

칙은 다음과 같다.

$$Emotion_t = Emotion_{t-1} + \alpha(\sum_{k=0}^n Keyword_k Structure_k + Logic)$$

현재 감정은 현재 발화에 존재하는 개별 키워드들의 감정 값과 구문에 의한 강조 가중치의 합에 논리적 일관성에 대한 상수를 더해 화용 정보로부터 얻은 흥분 상태에 대한 가중치  $\alpha$ 를 곱한 뒤, 이전 감정의 값에 결합한 뒤, 이를 긍정, 중립, 부정의 범주 값으로 변환하여 도출한다.

### III. 결 론

본 연구에서는 컨텍센터에서 이루어지는 채팅 상담 상황을 기준으로 대화 참여자들의 감정 상태를 추적해 상담이 성공적으로 진행되었는지를 실시간으로 확인할 수 있는 감정 추적 모델을 제안하였다. 채팅 상담을 대상으로 하는 만큼, 대화는 텍스트 데이터를 통해 이루어지기 때문에, 텍스트 데이터에서 확보할 수 있는 언어 정보로, 형태, 구문, 의미, 화용 정보를 통해 감정의 변화를 추적하였다. 이 방법을 통해, 채팅 대화 안에서, 대화 도메인에 상관 없이 실시간으로 대화 참여자인 상담 고객과 상담사의 감정 변화를 추적할 수 있으며, 상담 종료시 두 참여자의 감정 상태를 통해 대화의 성공 여부를 자동으로 판단할 수 있다.

### 참 고 문 헌

- [1] M. McCarthy, "Discourse analysis for language teachers", Child language teaching and therapy, 8(1), pp. 96-100, 1992..
- [2] M. M. Kazmer, G. Burnett, M. H. Dickey, "Identity in customer service chat interaction: Implications for virtual reference", Library & information science research, 29(1), pp. 5-29, 2007.
- [3] D. Zillmann, "Theory of affective dynamics: Emotions and moods", Communication and emotion. Routledge, pp. 533-567, 2003.

# 대화 내 감정 변화 추적을 통한 상담 평가 기법

이건수\*, 김찬호\*\*, 좌성훈\*\*\*

## A Method of Evaluating Counseling through Tracking Emotional Changes in Dialogue

Keonsoo Lee\*, Chanho Kim\*\*, and Seonghoon Jwa\*\*\*

### 요 약

상담 대화는 역할이 명확히 부여된 2명 이상의 참여자가 하나의 목적을 위해 수행하는 대화 형태를 의미한다. 이러한 상담 대화의 성공과 실패는 상담의 품질을 보증하기 위해 사후 평가가 진행된다. 이러한 평가 과정을 상담 대화 전체에 대해 수행하기 위해, 본 연구에서는 대화 참여자의 감정 변화를 추적하여 평가를 수행하는 방법을 제안한다. 감정 변화에 기반한 상담 대화 평가는 상담 내용에 대한 언어적 평가 없이도 진행될 수 있기 때문에, 상담 도메인에 독립적으로 적용 가능하며, 대화 종료시의 참여자들의 감정 상태를 평가하기 때문에, 고객 만족도를 보다 현실적으로 평가할 수 있다.

### Abstract

Counseling dialogue refers to a form of conversation conducted by two or more participants, each with clearly defined roles, to achieve a common purpose. The success or failure of such counseling dialogues is assessed post-session to ensure the quality of the counseling. To conduct this evaluation process for the entire counseling dialogue, this study proposes a method that tracks the emotional changes of the dialogue participants. Evaluation based on emotional changes can be conducted without verbal assessment of the counseling content, allowing for application independent of the counseling domain. Furthermore, by assessing the emotional states of participants at the end of the conversation, customer satisfaction can be evaluated more realistically.

### Keywords

Counseling Dialogue, Emotion Extraction, Emotion Tracking, Quality Assurance

### 1. 서 론

대화는 다양한 상황에서 다양한 목적을 갖고 진

행된다 [1]. 그 중 컨텍센터에서 이루어지는 대화는 2명의 참여자가 ‘상담 고객’과 ‘상담사’라는 명확한 역할을 갖고 ‘상담 고객’이 요청한 문제 상황에 대

\* C&AI Research Lab, keonsoo.lee@cnaikr(교신저자),

\*\* C&AI Research Lab, chanho.kim@cnaikr

\*\*\* SINSIWAY Human Resources Training Team, junsan12@sinsiway.com

해 ‘상담사’가 해결책을 제시하는 목적을 위해 진행된다. 그리고 이러한 상담 대화는 이후 컨텍센터 관리자가 품질보증(Quality Assurance)과정을 통해 검수하고, 그 결과에 기반하여 상담사의 근무 수행 능력을 평가한다. 대부분의 컨텍센터에서는 관리자의 수가 상대적으로 부족하기 때문에, 전체 상담 대화 중 일부분에 대해서만 QA를 수행하는 경우가 일반적이다. 이러한 문제를 해결하기 위해, 자동화된 QA 방법을 도입하려는 시도가 존재하지만, 대부분의 경우, 성공적인 상담 대화의 수행 보다는, 상담 중 상담사가 정해진 규칙대로의 응대를 수행했는지를 검사할 뿐이다. 이에 본 연구에서는 컨텍센터에서 상담 대화의 QA를 처리하기 위한 기준으로서, 대화 참여자들의 감정 변화를 추적하고, 이를 통해 해당 대화를 평가하는 방법을 제안한다.

## II. 감정 추적 모델

상담 대화가 성공적으로 종료되었다는 것의 의미는 대화를 신청한 상담 고객의 문제가 올바르게 해결된 것만을 의미하지는 않는다. 설령 문제가 해결되지 않더라도, 빠른 응대, 상황 공감, 또는 대안 제시 같은 것으로도 상담은 성공적으로 종료될 수 있다[2]. 즉 상담이 종료되는 순간 상담 고객의 감정 상태를 통해서, 상담의 성공 여부를 판별할 수 있게 된다. 또한 이 조건은 상담사 자신에게도 적용된다. 감정 노동자로서의 상담사의 근무 만족도 역시 중요한 요소이기 때문에, 상담이 종료되는 순간 상담사의 기분 역시 좋아야 한다[3]. 그렇기 때문에, 본 연구에서 제안하는 대화의 상담 대화의 감정 모델은 상담 고객과 상담사의 2명의 감정 상태로 이루어진다.

감정은 긍정(Positive), 중립(Neutral), 그리고 부정(Negative)로 표현되며, 다시 긍정과 부정은 ‘상’, ‘중’, ‘하’의 3단계로 구분한다. 최종적으로는 7점 리커트 척도의 모양을 갖고 있으나, 긍정과 부정의 3점 상태는 각 상태의 내부 상태로, 모델은 중립 상태 기계의 모양을 갖는다. 감정의 내부상태를 갖는 까닭은, 동일한 외부 이벤트에 대해 감정이 반응하는 정도가 다르기 때문이다.

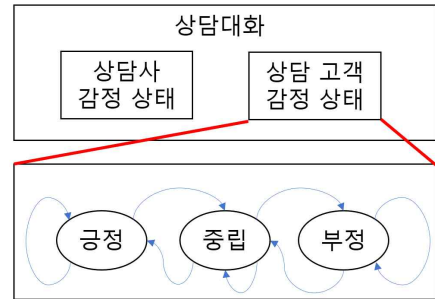


그림 1. 감정 상태 변환 모델  
Fig. 1. State Machine of Emotions

감정 상태를 변화시키는 이벤트의 종류는 일반적인 커뮤니케이션을 구성하는 언어적(Verbal) 요소, 음성적(Vocal) 요소, 그리고 비언어적(Non-verbal) 요소를 중에서 언어적 요소와 음성적 요소를 대상으로 자신과 상대의 발화로부터 추출해 사용한다. 이러한 제약은 컨텍센터에서 상담이 이루어지는 주요 채널이 전화이기 때문이며, 대면 상담이나 화상 상담의 경우라면, 비언어적 요소를 추가할 수 있다.

발화 정보에는 자신의 감정 상태를 변화시키는 자신의 발화와 상대의 감정 상태를 변화시키는 상대의 발화로 구분되며, 전자의 경우 자기비난, 자책, 자기비하 등의 표현을 포함하고, 후자의 경우 비꼼, 조롱, 폄하, 헐박, 무시 등의 표현을 포함한다.

발화 정보 이외에 대화의 진행 상태를 확인하기 위한 요소로 문제 해결 상태를 감정 변화를 야기하는 이벤트 종류에 포함한다. 문제 해결 상태는 대화의 현재 상태를 표현하기 위한 것으로 상담 대화가 본질적으로 내포하고 있는 문제 해결이라는 대화 목표의 진행도를 확인하기 위함이다.

본 연구에서는 문제 해결 상태를 표현하기 위해 PDCA(Plan-Do-Check-Action) 과정을 기반으로 한다. 이 과정에서 대화 참여자가 문제 해결을 위한 자신의 역할 수행에 실패하는 경우, 상대 혹은 자신의 감정을 변화시키는 발화로 이어지게 된다. 가령, PDCA에서의 특정 단계가 진행되지 못하고 정체되어 있는 경우, 특정 단계가 계속 반복되는 경우, 문제 해결의 실마리가 보이지 않는 경우, 참여자들의 발화는 서로에게 공격적으로 변화하게 되며, 문제가 원활하게 해결되는 과정에서는 서로에 대한 관용의 정도가 높아지게 된다.



표 1. 감정 상태 변환 이벤트 및 기분 구성 속성 종류  
Table 1. Types of Emotion States' Changing Events and Attributes for Determining Mood

그룹	이벤트	내용
감정	자신의 발화	자신의 감정을 변화시키는 자신의 발화 정보로 언어적, 음성적 요소로 구성됨
	상대의 발화	자신의 감정을 변화시키는 상대방의 발화 정보로 언어적, 음성적 요소로 구성됨
기분	문제 해결 상태	상담 대화의 목적을 달성하는 과정에서 현재 발화가 어떤 단계에서 진행되는지를 평가함
	발화 교대	대화 안에서 상호 원활한 대화의 흐름이 진행되는지 확인하기 위한 정보로, 말 끊기, 대화 독점 등을 통해 판별
	시간 흐름	대화 시작 후 문제 해결 상황 및 발화 내용 상, 요구되는 표준 시간을 토대로, 발화 속도/발화 공백/발화 순서 판별에 사용

발화 교대(Turn-taking)는 대화의 흐름속에서 참여자들이 자신들의 발화 차례를 주고 받는 것을 의미하며, 이를 통해 원활한 대화가 이루어지고 있는지를 확인할 수 있다. 말 끊기가 이루어지고 있는지, 한쪽이 혼자서 너무 많이 이야기하고 있는지 등을 확인한다. 대화의 목적에 따른 발화 교대 패턴이 달라지지만, 컨텍센터 내 상담 대화의 경우, 그림 2에서 정의된 대화 흐름 과정을 기준으로 할 때, 상담사와 고객의 발화 교대 패턴은 순서대로, 대화공백/말걸침 없이 5:5의 대화 점유를 해야 한다.

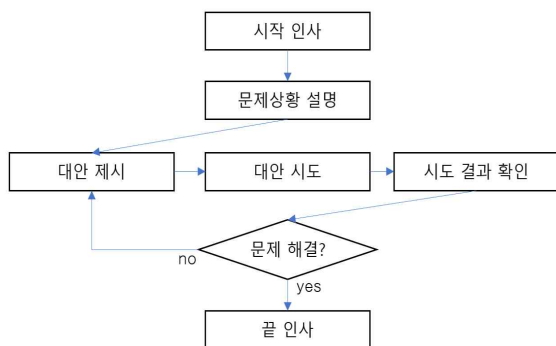


그림 2. 문제 해결 상태 변경 플로우  
Fig. 2. Problem Solving State Flow

시간 흐름이 대화 상황 정보를 구성하는 까닭은 앞에서 정의한 것처럼, 감정의 긍정과 부정은 세부 상태를 갖고 있고, 해당 상태에서 감정이 중립 상태로 되돌아가려는 항상성을 표현하기 위함이다. 즉 긍정/부정 상태에서 새로 발생한 발화가 중립 성향을 갖고 있다면, '중' 상태로의 지향성을 갖게 되며, '중'이 유지된다면, 그 감정은 다시 중립 상태로 변이하게 된다.

요컨대, 문제 해결 상태는 기분(Mood)으로 동작하고, 각각의 발화는 감정(Emotion)을 표현되며, 기분에 따라 발화가 감정을 변화시키는 강도를 조절시키는 방향으로 두 요소는 조합된다.

### III. 결 론

본 논문에서는 컨텍센터에서 진행되는 상담 대화에서 모든 대화 참여자의 감정 상태를 추적함으로써, 해당 상담의 성과를 평가하기 위한 모델을 제안하였다. 이를 위해 상태 변화 모델을 기반으로, 상담 상황에서 발생 가능한 이벤트를 정의하고, 각 감정 상태는 다시 리커트척도에 기반한 중첩 상태 기계로 모델링 하였다. 이 모델을 통해, 컨텍센터에서는 모든 발생 상담에 대한 QA테스트를 상호 비교 가능한 결과를 도출하며 수행할 수 있다. 또한 기존의 자동화된 QA 방법과는 다르게 상담 고객 뿐만 아니라, 상담사의 감정까지 고려하기 때문에, 직원의 근무 만족도를 동시에 개선시킬 수 있다.

### 참 고 문 헌

[1] D. Fishelov, "Types of dialogue: Echo, deaf, and dialectical", *Semiotica* 2013.195 pp. 249-275, 2013.  
 [2] R. Feinberg, I. Kim, L. Hokama, "Operational determinants of caller satisfaction in the call center", *International Journal of Service Industry Management* 11.2, pp. 131-141, 2000.  
 [3] K. Pugliesi, "The consequences of emotional labor: Effects on work stress, job satisfaction, and well-being", *Motivation and emotion* 23 pp. 125-154, 1999.

# 자가 검증을 통한 대화 분류 체계 개선 방법

이건수\*, 김찬호\*\*, 이남미\*\*\*

## A Method of Improving a Dialogue Classification Structure with Self-Validation

Keonsoo Lee\*, Chanho Kim\*\*, and Nammi Lee\*\*\*

### 요 약

상담 분야에서 상담 내용에 대한 분류는 상담 이후의 처리를 위한 중요한 과정이 된다. 기존의 상담 분류 체계는 도메인 담당자가 사전에 정의한 체계에 맞춰서 이루어지지만, 과연 그 기 정의된 분류 체계가 도메인을 정확하게 반영하고 있는지에 대한 검증은 체계적으로 진행되지 않고 있다. 본 연구에서는 담당자가 개인의 암묵지를 기준으로 만든 체계를 대신하여, 해당 도메인의 개별 목적에 맞춰 재분류하기 위한 방법을 제안한다. 제안된 분류 체계는 MECE의 원칙을 준수하며, 체계의 균형을 지향하며, 이를 통해, 기존 체계의 문제점을 탐색하고 그것의 개선안을 제안한다.

### Abstract

For counseling, classifying the content of counseling sessions is crucial for determining the subsequent actions. Existing classification systems are typically based on predefined structures established by domain experts. However, there is a lack of systematic validation to ensure that these predefined structures accurately reflect the counseling domain. In the paper, we propose a method for validating and advancing the classification structure, replacing existing systems rely on the implicit knowledge of experts. The resulting classification structure adheres to the principles of MECE and aims for balanced categorization. With this proposed method, limitations in existing classifications are identified and can be improved.

### Keywords

Counseling Type Classification, Classification Structure Validation, MECE, Category Analysis

## 1. 서 론

컨텍센터에서 진행되는 고객 상담 과정에서 상담사는 상담이 종료될 때마다, 정해진 분류 체계에 맞춰 그 상담의 유형을 분류한다. 이렇게 유형이 분류

되면, 그에 따른 후속 처리가 수행되기 때문에, 유형 분류는 고객 요청을 처리하기 위한 중요한 기준이 된다. 상담 유형의 분류 체계는 센터 도메인에 따라 도메인 전문가들이 자신들의 암묵지를 활용해서 작성된 후, 상담 센터로 내려보내진다. 하지만, 탐다

\* C&AI Research Lab, keonsoo.lee@cnae.kr(교신저자),

\*\* C&AI Research Lab, chanho.kim@cnae.kr

\*\*\* Strategic Planning Department, nammi.lee@ksjob.co.kr

운 방식으로 생성되는 이러한 분류 체계는 실제 상담 상황을 정확하게 반영하지 못하는 경우가 종종 발생한다. 가령, “기타”로 분류되는 상담의 수가 많아진다든지, 특정 분류 유형의 발생 빈도가 급격하게 낮아진다든지, 몇몇의 특정 유형에만 편중되어 있다든지 하는 경우가 그러하다.

이에 본 연구에서는 기존의 분류 체계를 검증하고, 개선하는 방법을 제안한다. 제안된 방법을 통해 도출된 상담 유형 분류 체계는 다음의 두 가지 측면에서 기존 체계를 개선한다. 첫째, MECE 원칙을 준수한다. MECE 원칙은 "Mutually Exclusive, Collectively Exhaustive"의 약자로, 이는 논리적 사고와 문제 해결 과정에서 중복되지 않고, 모든 가능성을 빠짐없이 포괄하는 분류 기준을 의미한다[1]. 둘째, 분류 체계의 균형을 개선한다. 분류 체계의 균형은 각각의 요소로 분류된 문서들의 수가 일정해야 함을 의미한다[2].

## II. 분류 체계의 검증 및 개선 방법

본 연구에서 MECE 원칙의 준수 여부를 검증하기 위해 제안하는 방법은 다음과 같다. 우선 상담 유형 체계에서의 MECE의 원칙은 다음과 같이 정의할 수 있다. (1)상위 유형과 그것의 자식 유형들 사이에는 부분-전체 관계를 갖고 있어야 한다. (2)형제 유형들 사이에는 상호배타적 관계를 갖고 있어야 한다. 이러한 관계를 도출하기 위해서 사용할 수 있는 중요한 정보는 유형의 이름이 되지만, 유형 이름은 자연어 특유의 모호함이 존재한다.

이 문제를 해결하기 위해, 본 연구에서는 요소 이름을 기준으로 기호화 변경 작업을 수행하고, 이를 토대로 관계를 유추하였다. 가령, “사용 내역”, “소득 공제”, “정보 찾기/변경” “배달\_가맹점”이라는 요소 이름이 있다고 할 때, 우선 단어 목록을 작성한다. [사용:1, 내역:2, 소득:3, 공제:4, 정보:5, 찾기:6, 변경:7, 배달:8, 가맹점:9]가 된다. 다음으로, 특수 기호에 대한 의미 확장을 수행한다. 슬래시는 디스정선(Disjunction)으로 변환하고, 언더바는 세분화(Specification)으로 변환하고 하이픈은 단순 공백으로 처리한다. 디스정선의 경우 후술할 유형 분할의 기준으로 사용할 수 있고 언더바의 경우 유형의 자

식 유형을 생성하기 위한 기준으로 사용할 수 있다. 이를 통하면, “12”, “34”, “56|57”, “8(9)”로 변경할 수 있다. 이 과정을 통해 추후 새로운 “수입 내역”이라는 이름의 요소가 들어왔다고 할 때, “수입”은 “소득:3”과 동의어라서, “32”로 변경되고, 이 요소는 추후 “2”로 인코딩된 다른 요소들과 결합될 수 있게 된다. 그 결과 “12”와 “32”는 “2”로 결합되고, 그 하위 유형으로 “12”와 “32”가 된다.

이러한 과정을 통해 앞에서 본 “주문확인”과 “주문내역확인”은 [주문:1, 확인:2, 내역:3] 이라고 할 때, “12”와 “132”가 형제 요소로 존재하는 경우, 상하위 관계를 가진 두 요소가 상호배타적 관계를 가져야 하는 형제 요소로 존재하기 때문에 MECE 원칙을 지키기 못했다고 판별하게 된다.

이때, 부모 요소와 자식 요소 사이의 관계가 상하 관계가 아닌 경우가 존재할 수 있다. 이것은 클래스와 인스턴스의 혼용으로 인해 생기는 문제라고 할 수 있다[3]. 그림 1은 상하 관계 변경이 수행되는 과정에 대한 개념도이다.

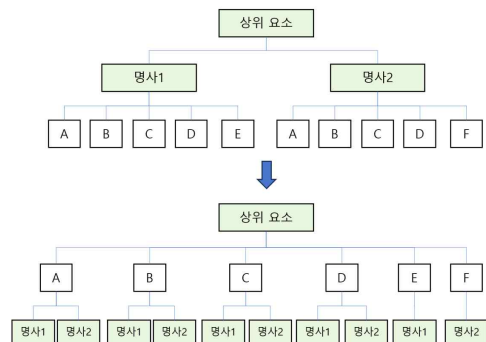


그림 1. 상하 관계 변경에 대한 개념도  
Fig. 1. Concept of Exchanging Hierarchy of Elements

MECE 원칙의 준수 여부를 확인 한 후, 체계의 균형을 확인한다. 분류 체계의 균형은 최종 단말 유형으로 분류된 상담 내용의 수가 비슷하다는 것을 의미하며, 이를 지키기 위해서는 요소들 사이의 결합 및 분리를 수행해야 한다. 즉 너무 세분화된 유형들은 서로 결합시키고, 너무 추상화된 유형은 더 세분화시킨다는 것이다.

결합의 경우, 결합되는 요소들끼리의 의미는 서로 묶일 수 있어야 하며, 묶일 수 없는 경우는 디스

정선 관계로 묶이게 된다. 또한 결합의 경우는 신규 자식 요소를 갖게 되며, 이 신규 요소는 모든 자식 요소의 합집합의 여집합, 즉 부모요소에 속하지만, 자식요소로 할당 될 수 없는 것을 할당시키는 “기타”의 역할을 수행한다.

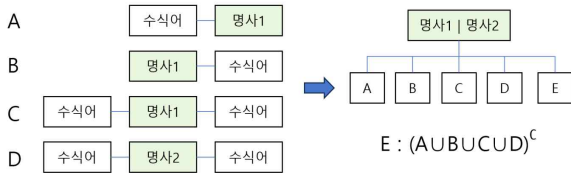


그림 2. 요소 결합에 대한 개념도  
Fig. 2. Concept of Combining Elements

분할의 경우, 해당 요소에 할당된 문서의 수가 평균을 초과하였을 때, 분할하게 되며, (1)사용자의 명시적인 구분 기준을 요청하는 방법과 (2)자동화된 구분 기준을 생성하는 방법으로 나뉘며, (1)의 경우에는 분할된 요소의 할당 균형에 대한 결과는 고려하지 않는다. (2)의 경우, 다음의 과정을 통해 단위 요소를 분할한다.

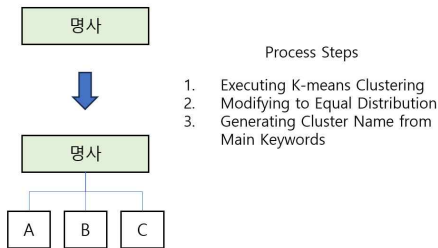


그림 3. 요소 분할에 대한 개념도  
Fig. 3. Concept of Separating Elements

부모 요소의 할당 문서 수를 기준으로 k개의 자식 요소를 결정하여 k-means 방법을 통해 클러스터를 생성하고, 클러스터에 따른 문서 수가 편중된 경우, 편중된 클러스터의 가장 외각의 문서들을 다른 클러스터와의 거리를 기준으로 재분배한다. 이 때 원 클러스터와의 거리가 타 클러스터와의 최단 거리보다 70% 이내로 짧다면, 재분배를 하지 않는다. 분배가 완료된 후 각각의 클러스터에 할당된 문서들의 대표 키워드를 추출한 뒤, 해당 키워드를 수식어로, 부모 요소의 이름을 주요 명사로 하는 클러스터의 이름을 생성한다.

### III. 결 론

본 연구에서는 상담 분류 체계를 검증하기 위해, 4개 업체의 상담 분류 체계를 대상으로 본 방법을 적용해, 기존 분류 체계의 단점을 찾아내고 개선안을 도출하였다. 개선안은 각 업체의 컨택센터를 운영하는 센터장과 매니저들과의 미팅을 통해 정성적 평가를 진행했고, 그 결과 기존 분류 체계의 문제점을 공유하고, 개선 결과에 대한 유의미성을 확인하였다.

하지만, 기존 체계에 대한 업체 담당자의 단점을 공식적으로 인정하고 개선한다는 점에서 담당자의 심적 거부감의 문제가 있었고, 또한 이미 사용되고 있는 시스템을 변경하는 점에서, 변경 시사점이 큰 경우 감정적인 거부감이 야기되는 경우가 발생하였다. 이것은 설령 명확한 단점이 존재하는 체계라고 해도 일단 조직 안에서 운영되고 사람들이 익숙해 지게 되면, 그것을 변경 혹은 개선하는 것에 반발이 생긴다는 것을 다시 한번 확인할 수 있었다.

본 방법을 이용하는 경우, 분류 체계를 도입하기 직전에 효과적으로 검증하고 수정할 수 있기 때문에, 기존에 운영중인 컨택센터보다 컨택센터를 운영하는 초기부터 도입하여 사용하는 것이 기술 외적인 부분에서의 활용에 더 효과적일 것으로 예상할 수 있다.

### 참 고 문 헌

[1] B. Minto “The Pyramid Principle:Present Your Thinking So Clearly That the Ideas Jump Off the Page and into the Reader’s Mind“, Financial Times/Prentice Hall, 2005

[2] Y. Lin, H. Tang, Y. Li, C. Fang, Z. Xu, Y. Zhou, A. Zhou, “Generating clusters of similar sizes by constrained balanced clustering”, Applied Intelligence. 52(3). pp. 5237-5289. March 2022

[3] M. Renate, J. Mylopoulos, "Classes and instances", International Journal of Intelligent and Cooperative Information Systems 1.01 pp. 61-92. 1992

# HPC & AI 환경의 대규모 데이터 처리를 위한 파일시스템 최적화 및 성능 분석

윤준원\*, 안도식\*\*, 흥태영\*\*\*

## Filesystem Optimization and Performance Analysis for Large-scale Data Processing in HPC & AI Environments

Junweon Yoon\*, Do-Sik An\*\*, and Taeyoung Hong\*\*\*

### 요 약

데이터의 특성 관점에서 HPC는 주로 과학적 시뮬레이션과 계산을 위해 최적화되었다면, AI의 ML/DL은 텍스트, 이미지, 음성과 같은 여러가지 형태의 방대한 데이터를 활용하여 학습을 통해 패턴을 찾고 추론하는데 중점을 둔다. 이러한 데이터를 효율적으로 처리하기 위해서는 대량의 연산 처리 능력과 더불어 대규모 데이터 병렬처리가 필요하다.

본 논문에서는 HPC와 AI 시스템에서 요구되는 데이터 처리의 특성을 분석하고 메타데이터 처리 성능 향상과 확장성을 높이기 위한 대표적 기법들을 적용한다. 이로써 대규모의 파일시스템에서 데이터 액세스 속도와 확장성, 처리 효율성을 지원함을 확인한다.

### Abstract

In terms of data characteristics, HPC is primarily optimized for scientific simulations and computations, while AI's ML/DL focuses on learning to find patterns and make inferences by utilizing massive amounts of data in various forms, such as text, images, and speech. Efficient processing of such data requires not only high computational power but also large-scale parallel data processing. This paper analyzes the data processing characteristics required in HPC and AI systems and applies representative techniques to improve metadata processing performance and scalability. As a result, it is confirmed that these techniques support data access speed, scalability, and processing efficiency in large-scale file systems.

### Key words

HPC, AI, parallel data processing, parallel filesystem, metadata

---

\* 한국과학기술정보연구원 슈퍼컴퓨팅인프라센터, jwyoona@kisti.re.kr, \*\*dsan@kisti.re.kr, \*\*\*tyhong@kisti.re.kr

※ 이 논문은 2024년도 한국과학기술정보연구원의 기본사업(과제명: 국가 플래그십 초고성능컴퓨터 인프라 구축 및 서비스, 과제번호: K24L2M1C1)으로 수행된 연구입니다.

## 1. 서 론

고성능 컴퓨팅(High Performance Computing, HPC)과 인공지능(AI) 모두 대규모 데이터를 처리하지만 방식과 특성에 차이가 있다. HPC는 주로 과학적 시뮬레이션과 계산을 위해 최적화된 반면에 AI는 데이터 기반의 모델 학습과 추론에 중점을 둔다[1]. HPC는 대규모의 과학적 계산이나 시뮬레이션을 처리할 때 수천 개에서 수백만 개의 CPU 또는 GPU 코어를 병렬로 활용하여 각 코어가 독립적으로 계산을 수행하고, 계산 결과를 병합하는 방식으로 처리된다. AI는 대규모 데이터를 처리하기 위해 GPU나 TPU를 사용하여 대량의 연산을 동시에 처리한다. 딥러닝 모델의 경우, 수백에서 수천 개의 레이어가 포함될 수 있고, 각각마다 대규모 행렬 연산이 발생한다. AI는 대부분 확률적 추론을 사용하며 완벽한 정확성보다는 패턴 인식을 기반으로 한 예측에 중점을 두어 HPC와 달리 데이터의 일부 손실이나 오류가 결과에 크게 영향을 미치지 않는다. 학습 과정에서는 손실 함수의 최소화를 목표로 하며 가장 가능성이 높은 답을 찾는다. 이렇게 AI는 대규모 데이터 세트를 처리하기 때문에, 데이터 저장장치와 네트워크 대역폭에 따른 입출력 속도가 연산 성능에 큰 영향을 끼친다[2].

본 논문에서는 HPC 시스템에 적용된 스토리지에 AI 데이터 처리에 요구되는 I/O 특성을 지원하기 위한 메타데이터 확장, 작은 파일 처리 향상 기법을 적용하고 BMT를 통해 성능의 향상을 검증한다.

## II. 메타데이터 활용 확장 기법

본 연구를 위해 HPC 환경에서 주로 적용되는 Lustre 병렬 파일시스템(Parallel File System)을 사용한다. Lustre는 고성능, 확장성, 신뢰성을 갖춘 파일 시스템으로 HPC 및 대규모 데이터 처리를 위한 과학 연구분야에 쓰이고 있다. Lustre의 DNE(Distributed Name Space)와 DoM(Data on Metadata) 기술은 대규모 데이터 처리 환경에서 중요한 역할로 HPC 및 AI 시스템에 데이터 액세스 속도와 확장성, 처리 효율성을 충족시킨다[3].

### 1) DNE(Distributed Name Space)

전통적인 파일시스템에서는 모든 디렉터리 이름과 메타데이터가 한곳의 서버에 저장된다. 이는 데이터 액세스가 많아질수록 병목 현상이 발생할 수 있다. DNE는 이를 해결하기 메타데이터를 여러 서버에 분산해 저장하여 디렉터리 생성, 삭제, 검색 등의 작업 속도를 개선한다.

### 2) DoM (Data on Metadata)

DoM은 파일의 소형 데이터를 메타데이터 서버에 직접 저장하는 기술로 파일의 데이터는 스토리지 서버(OSS, Object Storage Server)에 저장되지만, DoM 기술은 데이터가 아주 작을 때 메타데이터 서버에 저장하여 별도의 스토리지 서버를 통하지 않고 데이터에 직접 접근이 가능하다. 메타데이터 서버에 데이터가 함께 저장되므로, 작은 파일에 대한 메타데이터와 데이터 요청이 동시에 이루어져 기존 시스템과의 호환성을 유지하면서 성능이 향상된다.

### 3) MDTEST 벤치마크

MDTEST는 병렬 파일시스템 성능 평가도구로 메타데이터 성능을 측정한다. 메타데이터 작업에 대해 초당 수행 가능한 작업 수(IOPS: I/O operations per second)를 측정하며, 파일 생성 속도(초당 생성된 파일 개수), 파일 삭제 속도(초당 삭제된 파일 개수), 디렉터리 생성 및 삭제 속도(디렉터리 작업의 초당 처리량), 파일 읽기 및 열기 속도(파일에 접근하는데 걸리는 시간)의 성능 지표를 보여준다.

## III. 성능 평가

실험 환경에 적용된 Lustre 파일시스템은 DNE를 위해 복수개의 Metadata Target(MDT)이 구성되어 있다. Lustre 2.X 버전은 최대 256개까지 메타데이터를 분산하여 구성 가능하나 메타데이터 정보가 여러 곳에 분산되는 경우 파일의 정보를 수집해야 하는 부하가 증가할 수 있다. 따라서 스토리지 용량, 파일의 크기, 개수에 따라 메타데이터 어떻게 분산해야 할지 결정해야 한다. <표 1>과 같이 두 개의 디렉터리를 생성 후 하나의 디렉터리는 단일, 다른 하나는 두 개의 메타데이터로 구성하여 성능을 비교하였다.

표 592. DNE/DoM 설정 및 디렉터리 속성  
Table 592. DNE/DoM configuration and directory properties

```
[ DNE configurations ]
# lfs getdirstripe /scratch/jwoon/MDTEST/tmp/b
lmv_stripe_count:2 lmv_stripe_offset:1 lmv_hash_type: fnv_1a_64
mdtidx          FID[seq:oid:ver]
  1              [0x240040cd4:0x14a62:0x0]
  0              [0x200042688:0x1229c:0x0]

[ DoM Configuration ]
# lfs getstripe -d /scratch/jwoon/MDTEST/tmp/b
lcm_layout_gen: 2
lcm_mirror_count: 1
lcm_entry_count: 2
  lcme_id: 1
  lcme_mirror_id: 0
  lcme_flags: init
  lcme_extent.e_start: 0
  lcme_extent.e_end: 65536
  lmm_stripe_count: 0
  lmm_stripe_size: 65536
  lmm_pattern: mdt
  lmm_layout_gen: 0
  lmm_stripe_offset: 0

  lcme_id: 2
  lcme_mirror_id: 0
  lcme_flags: 0
  lcme_extent.e_start: 65536
  lcme_extent.e_end: EOF
  lmm_stripe_count: 1
  lmm_stripe_size: 1048576
  lmm_pattern: raid0
  lmm_layout_gen: 0
  lmm_stripe_offset: -1
```

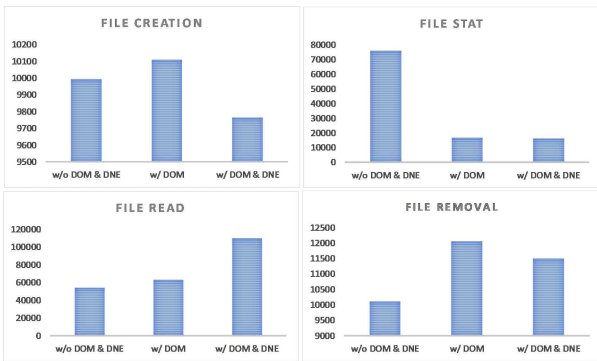


그림 1. DNE, DoM 설정에 따른 메타데이터(IOPS) 성능  
Fig. 1. Metadata (IOPS) performance based on DNE, DoM Configurations

MDTEST는 벤치마크는 메타데이터 작업 성능 측정이 목적으로 파일의 데이터 크기를 0으로 설정하여 파일 생성, 삭제, 상태 조회 등을 실행한다. File Creation(파일생성)에서 DoM 옵션을 사용한 경

우가 가장 높은 성능을 보였으나 File Stat(파일상태 조회)의 경우 DoM과 DNE 모두를 사용하지 않을 때 가장 좋은 성능을 얻었다. File Read(파일읽기)는 DoM과 DNE를 동시에 사용할 때 성능이 가장 높게 File Removal(파일삭제)는 DoM과 DNE를 동시에 사용할 경우 좋은 성능을 얻었다. 이 차이는 초기 생성시 양쪽에 메타데이터의 정보를 분산하여 저장하고 정보를 찾아 가져오는 부하로 볼 수 있다. 파일을 읽거나 제거할때는 작은 파일이 양쪽의 메타데이터에 분산되어 처리되므로 IOPS가 향상 됨을 확인할 수 있다. 즉, 파일 읽기와 삭제 작업에 중점을 두는 워크로드라면 DOM과 DNE 동시 적용은 향상된 I/O 성능을 가져온다.

#### IV. 결 론

HPC와 AI 환경에 GPU와 같은 가속기 기반 시스템 적용이 일반화되고 있으며 I/O 처리 요구 성능도 급격히 증가한다. 본 논문에서는 AI 환경에서 반복적으로 발생하는 대규모 데이터 처리를 위한 메타데이터 관련 설정을 적용하고 성능을 검증하였다. 다양한 I/O 패턴에 따라 서로 다른 성능 차이가 있음을 확인하였고 사용자 작업에 따른 메타데이터의 동적 처리 방안을 고려할 필요가 있다.

#### 참 고 문 헌

- [1] Paul, Arnab. K, Karimi, A. M and Wang, F, "Characterizing machine learning i/o workloads on leadership scale hpc systems", MASCOTS, pp. 1-8, 2021.
- [2] W. Lv, Y. Lu, Y. Zhang, P. Duan, and J. Shu, "An efficient metadata service for Large-Scale distributed filesystems", 20th USENIX Conference on FAST 22, pp. 313-328, 2022.
- [3] J. Fragalla, B. Loewe, and T. Kling Petersen, "New Lustre features to improve Lustre metadata and small-file performance. Concurrency and Computation", Practice and Experience, 32(20), e5649, 2020.

# IOT를 이용한 재활 운동 의료기기의 모니터링 시스템 설계

유도우\*, 양오\*\*

## Design of Monitoring System for Medical Devices for Rehabilitation Exercise using IOT

Do Woo Yu\* and Oh Yang\*\*

### 요 약

고령화 사회로 접어들게 되면서, 관절에 문제가 있는 환자들은 병원에서 수술 후 대부분 재활 운동을 한다. 최근에는 가정에서 재활 의료기기를 대여하여 병원을 방문할 필요 없이 사용자가 직접 치료를 하는 경우가 증가하고 있으며, 환자가 더욱 편리하게 재활치료를 할 수 있는 디지털 의료기기 개발의 필요성이 확대되고 있다. 본 논문에서는 기존 의료기기에 없는 블루투스 모듈을 이용한 무선 통신을 통해 현재 운동 상황을 모니터링하고 운동 정보를 기록하고 저장함으로써 사용자의 편의성을 도모하고자 한다.

### Abstract

As we enter an aging society, most patients with joint problems undergo rehabilitation exercises after surgery at a hospital. Recently, cases of patients renting rehabilitation medical devices at home and receiving treatment directly without visiting a hospital are increasing, and the need for developing digital medical devices that allow patients to receive rehabilitation treatment more conveniently is increasing. In this paper, we aim to improve user convenience by monitoring the current exercise status and recording and storing exercise information through wireless communication using a Bluetooth module that is not available in existing medical devices.

### Key words

Digital Medical Device, Bluetooth, Monitoring, Rehabilitation Exercises

## 1. 서 론

어깨, 팔꿈치, 무릎, 손목, 발목과 같은 관절에 문제가 생겨 수술을 받았던 환자들은 병원에서 관절 가동범위(ROM : Range of Motion)를 회복시키기 위해 다양한 재활 의료기기를 사용해왔다[1].

Fig. 1과 같이 통계청의 2023년 고령자 통계자료를 보면, 2023년에 대한민국의 65세 이상 고령인구는 950만 명으로 총인구 대비 18.4%이고, 점차 고령인구 비중이 늘어날 것으로 전망했다. 고령 사회로 진입함에 따라 관절 관련 질병 환자의 수가 증가하고 있고, 병원을 가지 않고 가정에서도 환자 스

\* 청주대학교 반도체공학과 석사과정, Email : ehdn4358@naver.com

\*\* 청주대학교 반도체공학과 교수, Email : ohyang@cju.ac.kr

※ 이 논문은 2024~2025년도 청주대학교 연구장학 지원에 의한 것임.



스로 재활 치료를 할 수 있는 디지털 의료기기 개발의 필요성이 확대되고 있다[2]. 따라서 본 논문에서는 현재 국내 재활 의료기기에 없는 블루투스 모듈을 이용하여 애플리케이션을 통해 현재 운동 상황을 모니터링하고 운동 정보를 기록하고 저장함으로써 환자의 편리성을 도모하고자 한다.

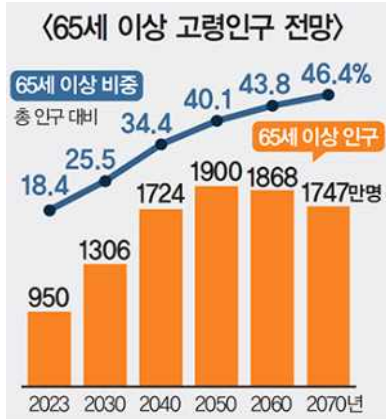


그림 1. 65세 이상 고령인구 전망  
Fig. 1. Outlook for the elderly population aged 65 and over

## II. 재활 운동 의료기기의 모니터링 시스템 설계

### 2.1 재활 운동 의료기기 모니터링 시스템 설계

본 논문에서는 재활 운동 의료기기의 모니터링 시스템을 구현하기 위해 Fig. 2와 같이 설계하였다. 의료기기 내부의 센서들로부터 얻은 환자 운동 정보 데이터를 ROM 측정 기구의 MCU(ST사의 STM32F103)에 저장한다. MCU와 블루투스 모듈 간 UART 통신으로 환자 운동 데이터를 전송하고 블루투스 모듈과 애플리케이션 간 무선 통신을 통해 현재 운동 정보를 스마트폰으로 확인할 수 있다.

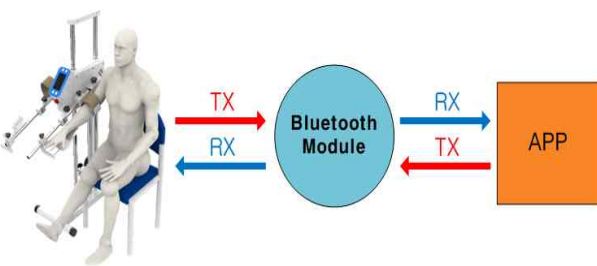


그림 2. 재활 운동 의료기기 모니터링 시스템 구성도  
Fig. 2. Rehabilitation Exercise Medical Device Monitoring System Configuration Diagram

### 2.2 모니터링 시스템

애플리케이션에서 블루투스를 거쳐 무선 통신으로 MCU에 100ms마다 데이터 요청 신호를 보내면 MCU에서는 이를 받아, 현재 환자가 진행 중인 운동 데이터를 전송한다. 이러한 운동 데이터를 통해 Fig. 3과 같이 5가지 관절 부위(어깨, 팔꿈치, 무릎, 손목, 발목)중, 어깨 부위에 대한 운동 정보를 애플리케이션으로 실시간으로 확인할 수 있다[3].

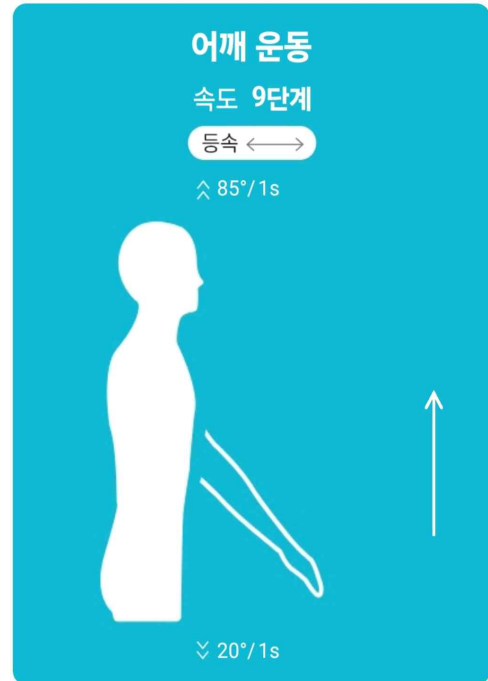


그림 3. 어깨 운동에 대한 실시간 모니터링 화면  
Fig. 3. Real-time monitoring screen for shoulder exercise

### 2.3 재활 운동 데이터 기록 및 저장 기능 구현

Fig. 4와 같이 재활 운동 의료기기에서 운동 종료 시 MCU에서 애플리케이션으로 운동 완료 신호(어깨)를 보내게 되는데 이를 통해 사용자가 스마트폰을 사용 중이 아니더라도 운동 완료 알림과 동시에 애플리케이션의 운동 기록 화면에서 현재 운동 정보 데이터를 자동으로 기록하고 저장할 수 있다. Fig. 5(a)의 운동 기록 화면(어깨)에서는 달력 하단의 표시를 통해 운동을 진행한 날짜를 확인할 수 있으며, 환자의 운동 정보에 대한 데이터를 저장하여 해당 달의 30일간의 운동 횟수를 x축은 날짜, y축은 운동 횟수로 설정한 막대형 그래프로 나타내었다. 또한, Fig. 5(b)에서 달력의 특정 날짜를 2초 동안 터치 유지 시 해당 날짜의 운동 기록 데이터를 팝업창을 통해 확인할 수 있다.

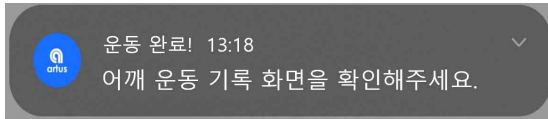


그림 4. 어깨 운동에 대한 운동 완료 신호

Fig. 4. Exercise Completion Signal for Shoulder Exercises



(a) 어깨 운동 기록 화면



(b) 날짜별 어깨 운동 기록 팝업창

그림 5. 어깨 운동 데이터 기록 및 저장

Fig. 5. Record and save shoulder exercise data

Shoulder exercise record screen (a)

Date-based shoulder exercise record pop-up window (b)

### III. 결 론

초고령화 사회로 진입함에 따라 병원에서 관절 수술 후 회복을 위해 재활 운동을 하는 환자의 수는 증가하고 있으며, 치료사 없이 가정에서 환자 스스로 치료가 가능한 디지털 의료기기 개발의 필요성이 확대되고 있다. 본 논문에서는 국내 재활 운동 의료기기에 없는 블루투스 모듈을 이용하여 환자가 스마트폰 애플리케이션으로 현재 운동 상황을 실시간 모니터링 할 수 있도록 구현하였다. 또한, 운동 종료와 동시에 애플리케이션으로 운동 완료 신호를 보내 현재 진행했던 운동 정보를 자동으로 저장하였으며, 애플리케이션의 운동 기록 화면에서 해당 날짜에 운동했던 과거의 운동 기록을 확인할 수 있도록 하여 국내 재활 운동 의료기기의 성능과 환자의 재활 운동의 편리성을 향상시켰다.

### 참 고 문 헌

- [1] Ki Mai Um, Yoon Kwon Yang, Soo Kyung Chang, "A Study on Hip Joint ROM of the Elderly", Journal of the Korean Society of Physical Therapists, Vol. 9, No. 2, pp.67-75, 2002.
- [2] Tae Ho Kim, Do Bong Oh, Da Yeon Kim, "Usability Testing of Digital Pressure Bio-feedback for Spinal Rehabilitation Exercise", Journal of Rehabilitation Welfare Engineering & Assistive Technology, Vol. 11, No. 3, pp.199-206, 2017.
- [3] Heung Ki Kim, Jin Soo Cho, "Design and Implementation of Real-time ECG Monitoring System for Personal Health Records", Journal of the Semiconductor & Display Technology, Vol. 11, No. 3, pp.45-50 September 2012.
- [4] 강성운, "Do it! 강샘의 안드로이드 앱 프로그래밍 with 코틀린", 이지스퍼블리싱

# XR 환경을 결합한 원격 모빌리티 제어 시스템의 개발

이로빈\*, 김윤상\*\*

## Development of a Remote Mobility Control System Integrated with an XR Environment

Robin Lee\*, Yoon Sang Kim\*\*

원격 제어 시스템은 특정 장비나 시스템을 물리적으로 가까이 있지 않은 상태에서 제어할 수 있도록 지원하는 기술로, 자율 주행 기술과 원격 제어 분야에서 큰 잠재력을 지니고 있다. 이러한 원격 제어 시스템에 추가적인 콘텐츠와 기능을 제공하기 위해서는 실제 모빌리티와 XR 기술의 결합이 필수적이다.

XR 기술은 원격 제어자가 마치 현장에 있는 것처럼 생생한 시각적, 공간적 정보를 제공하여 보다 정밀하고 직관적인 원격 제어를 가능하게 한다. 실제로 해양 로봇과 XR을 결합하여 제어하는 연구가 증가하고 있으며[1], 세계 최대 이동통신 전시회인 MWC 2023에서는 프랑스의 통신 기업 '오렌지'에서 확장현실을 통해 로봇을 제어하는 사례가 발표되었다. 이러한 이유로 원격 제어 기술과 XR 기술의 결합은 제품의 필수적인 기능으로서 중요한 발전 단계로 예상된다.

이에 본 논문은 이러한 기술적 배경을 바탕으로 XR 환경을 결합한 모빌리티 시스템 'Acro\_V1'을 제안한다. 제안된 시스템은 모바일 로봇을 제어하며, 모바일 로봇의 시야와 동기화된 실시간 영상과 XR 콘텐츠를 결합하여 사용자에게 전송하는 시스템이다. 본 시스템의 개발 및 제안을 통해 자율 주행 기술이 상용화되기 전 원격 주행 기술을 보완하는 데 도움이 될 것으로 예상된다.

### Key words

XR, Simulator, Mobile Robot, Remote System

### 참고문헌

[1] [https://www.kitech.re.kr/webzine\\_re/subpage.php?set\\_page=202301&set\\_subpage=sub1\\_1](https://www.kitech.re.kr/webzine_re/subpage.php?set_page=202301&set_subpage=sub1_1)

[2] <https://www.etnews.com/20230301000015>

---

\* 주식회사 소프트웨어, sgrobin@naver.com,

\*\* 한국기술교육대학교 컴퓨터공학과, yoonsang@koreatech.ac.kr(교신저자)

※ 본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다.(2021RIS-004)

# 무음구간의 임계값을 기반으로 한 잡음 억제를 위한 음성구간의 검출 방법

최재승\*

## A method for detecting voice segments based on threshold in silent segment for noise reduction

Jae-Seung Choi\*

근래에 음성인식과 디지털신호처리 기술이 발전되고 있지만 현재까지도 잡음이 많은 환경에서는 음성의 구간이 명확하게 검출되지 못하는 문제가 존재하기 때문에 음성구간 검출이 중요한 문제로 남아있다[1, 2]. 본 논문에서는 깨끗한 음성신호를 강조하고 잡음신호를 억제하기 위하여[3, 4, 5], 임계값 기반의 음성구간 검출 방법을 적용하는 알고리즘을 제안한다. 그러므로 음성의 각 프레임 단위로 임계값 기반의 음성구간 검출 방법을 적용하여 음성신호의 유무를 확인한 후에 무음신호 구간에서만 잡음의 전력을 추정하는 알고리즘을 제안한다. 제안하는 알고리즘에서는 먼저 음성구간 검출 방법에 있어서 임계값을 이용하여, 각 프레임의 주파수의 스펙트럼 전력을 비교하여 잡음신호 구간 및 음성신호 구간을 분류한다. 본 논문에서는 각 프레임별로 음성신호에 대한 잡음신호의 영향이 서로 다르다는 관점에서, 음성구간 검출 방법에 의한 잡음신호 및 음성신호의 분류를 각 프레임별로 실험을 진행하였다. 본 논문에서는 다양한 음성신호와 배경 잡음신호를 사용하여 실험을 진행하였으며, 제안한 알고리즘은 평균 입력 SNR이 0(dB)에서 본 알고리즘의 음성구간 검출 효과가 우수하였으며, 입력 잡음 레벨이 상당히 높은 잡음신호에 대해서도 제안한 알고리즘이 유효성이 높다는 것을 실험으로 확인할 수 있었다.

Key words

voice activity detection, speech recognition, speech enhancement, noise reduction

### 참고문헌

- [1] J. J. Lee, H. T. Han, Y. U. Choi, and H. Y. Lee, "Voice Activity Segment Audio Deduction Method using MobileNet", Journal of KIIT, vol. 20, no. 3, pp. 1-8, Mar. 2022.
- [2] J. S. Choi, "Gender Recognition for Speaker in Colored Noise by Speech/Non-speech Discrimination", Journal of KIIT, vol. 10, no. 11, pp. 63-68, Nov. 2012.
- [3] T. Lotter and P. Vary, "Noise reduction by joint maximum a posteriori spectral amplitude and phase estimation with super-Gaussian speech modelling", 2004 12th European Signal Processing Conference, pp. 1457-1460, Sep. 2004.
- [4] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction", IEEE Trans. on Acoustic Speech Signal Processing, vol. ASSP-27, no. 2, pp. 113-120, April 1979.

---

\*신라대학교 전기전자공학과.

# 캐릭터 이미지 공정이용 플랫폼 개발

김성찬\*, 신진섭\*\*

## Development of Character Image Fair Use Platform

Seongchan Kim\* and Jinseop Shin\*\*

캐릭터 IP 산업의 급성장과 함께 지식재산권 분쟁이 빈번해지면서 캐릭터 저작물의 공정이용을 보장할 수 있는 시스템적 인프라의 필요성이 대두되고 있다. 본 연구는 이러한 필요에 대응하기 위해 인공지능 기술을 활용하여 캐릭터 이미지 공정이용을 지원하는 플랫폼 개발을 목표로 한다. 본 플랫폼은 캐릭터 이미지의 저작권 보호와 공정 이용을 지원하기 위해 인공지능 기반의 유사도 측정, 이미지 검색, 세그멘테이션 및 PID(Persistent Identifier) 부여 및 등록 시스템 등 다양한 기능을 제공한다. 구체적으로는, 캐릭터 이미지의 세부 객체 인식을 통한 이미지의 유사도를 분석하는 모델을 기반으로 한 웹 인터페이스를 개발하고, 이를 API로 연동하여 사용자가 캐릭터 이미지의 유사도를 손쉽게 측정할 수 있도록 한다. 또한, 캐릭터 이미지 검색 기능과 PID 부여 시스템을 통해 이미지의 고유 식별자를 등록·관리할 수 있도록 한다. 아울러, 플랫폼 웹서비스는 독립적인 보안 서버를 통해 운용되며, 사용자 접근 제어와 인증 기능을 제공하여 보안을 강화한다. 본 플랫폼의 개발로 캐릭터 이미지의 유사도 분석 서비스를 통한 저작권 보호 체계를 구축하고, 캐릭터 저작물의 공정이용 기반을 마련함으로써 캐릭터 IP 산업의 지속 가능한 성장을 도모할 수 있을 것으로 기대된다.

### Key words

Character Image, Fair Use, Platform, Intellectual Property

### 참고문헌

- [1] 김성목, "디지털 콘텐츠 저작물의 공정이용 확대에 대한 연구", 문화기술의 융합 (The Journal of the Convergence on Culture Technology), 제6권 2호, Pages. 217-222, 2020.
- [2] Bin Yang, "Perceptual similarity measurement based on generative adversarial neural networks in graphics design," Applied Soft Computing, Volume 110, 2021.

---

\*한국과학기술정보연구원 인공지능데이터연구팀, UST-KISTI 응용AI sckim@kisti.re.kr; \*\*한국과학기술정보연구원 디지털큐레이션센터 js.shin@kisti.re.kr

※ 본연구는 2024년도 문화체육관광부 재원으로 한국콘텐츠진흥원의 지원을 받아 수행된 연구임(RS-2024-00397382, 인공지능을 활용한 캐릭터 유사도 검증 및 IP 라이선스 추출 기술 개발)

# 대중교통 이동시간을 고려한 중간지점 탐색 알고리즘

박지혜\*, 엄성용\*\*

## A Midpoint Search Algorithm Considering Public Transportation Travel Time

Ji Hye Park\*, and Seong Yong Ohm\*\*

모임 장소를 결정하는 과정에서 참여자들의 이동 시간을 고려하는 것은 매우 중요한 문제이다. 특히 대중교통을 주로 이용하는 현대 도시 환경에서는 단순한 지리적 중심점이 실제 이동 시간 측면에서는 최적의 장소가 아닐 수 있다. 기존의 무게중심을 이용한 중간점 계산 방식[1]은 3차원 좌표계에서의 위치만을 고려하여 실제 대중교통 노선과 환승에 따른 이동 시간을 반영하지 못하는 한계가 있으며, 다익스트라 알고리즘을 활용한 최단 경로 기반의 접근 방법[2] 역시 역 사이의 물리적 거리만을 고려할 뿐 환승 시간과 배차 간격 등 실제 이동에 영향을 미치는 요소들을 고려하지 못하는 한계가 있었다. 본 연구에서는 이러한 한계점들을 극복하기 위해 실제 대중교통 이동시간을 기반으로 한 최적의 중간지점 탐색 알고리즘을 제안하고자 한다.

제안하는 알고리즘은 ODsay API를 활용하여 실시간 대중교통 정보를 수집하고, 이를 바탕으로 각 참여자들의 이동 시간을 계산한다. 알고리즘은 주어진 출발지들의 중심점을 기준으로 반경을 점진적으로 확장(1km → 2km → 3km → 5km)하며 후보 지하철역들을 검색한다. 각 경로는 총 이동 시간이 5분 이내로 차이 나는 경우 환승 횟수를 우선적으로 고려하고, 환승 횟수가 동일한 경우 도보 시간이 적은 경로를 선택하는 방식으로 평가된다. 또한 최적 역사 선정 시에는 최대 소요시간이 60분을 초과하는 역사를 제외하고, 시간 편차가 20분 이상 차이 나는 경우 편차가 작은 역사를 우선 선택하며, 환승역의 경우 환승 편의성 점수(환승노선 수 \* 2분)를 반영한다.

서울시 내 실제 데이터를 통한 실험에서 제안된 알고리즘은 참여자들의 실제 이동시간과 환승 횟수를 효과적으로 최적화하여 모든 참여자의 이동 편의성을 향상시켰다. 향후 연구에서는 더 정확한 중간 지점 계산을 위해 다양한 경로 탐색 알고리즘을 추가로 연구하고, 사용자의 이동 패턴과 지역별 특성을 고려한 개선된 추천 시스템을 개발할 예정이다. 또한 모바일 애플리케이션으로 개발하여 실제 사용자들의 피드백을 수집하고 이를 반영하여 알고리즘을 지속적으로 발전시켜 나갈 예정이다.

### Key words

public transportation, optimal meeting point, travel time optimization, subway station search

### 참고문헌

- [1] 배혜지, 송지나, 이유진, and 이종우, "사용자의 위치와 선호도에 기반한 약속 장소 추천 모바일 애플리케이션 구현", 정보과학회 컴퓨팅의 실제 논문지, Vol. 21, No. 6, pp. 403-411, 2015.
- [2] 박나은, 문지연, 정운아, 조서연, and 허원희, "다익스트라 알고리즘을 활용한 중간지점 추천 애플리케이션 개발", 멀티미디어학회논문지, Vol. 24, No. 2, pp. 312-319, 2021.

\*서울여자대학교 소프트웨어융합학과, modzivv@swu.ac.kr,

\*\*서울여자대학교 소프트웨어융합학과, osy@swu.ac.kr(교신저자)

※ 본 연구는 교육부 및 한국연구재단 대학혁신지원사업으로부터 지원받은 연구임(2024)

# 다중 이미지 정합성 검증을 활용한 위변조 이미지 탐지 방법론

방민서\*, 전유선\*, 민예림\*, 김민서\*, 김남규\*\*

## A methodology for detecting AI-generated images using multi-image consistency verification

Minseo Bang\*, Yuseon Jeon\*, Yerim Min\*, Minseo Kim\*, and Namgyu Kim\*\*

### 요 약

최근 실제 차량 사고를 기반으로 사고의 내용을 왜곡하거나 피해 정도를 과장하여 보험금을 청구하는 보험 사기가 급증하고 있다. 특히 AI 기반 이미지 생성 모델을 활용하여 피해를 과장하는 수법이 더욱 교묘해짐에 따라, 위변조 이미지를 정확하게 탐지하는 일이 더욱 어려워지고 있다. 이에 본 연구는 AI 생성 이미지를 탐지하는 기존 모델을 활용하되, 이와 동시에 다중 이미지를 입력받아 이들간 정합성을 검증하여 모순 여부를 탐지하는 다중 이미지 정합성 검증 방법론을 제안한다. 제안 방법론은 위변조 이미지 탐지 능력을 향상시켜, 궁극적으로 보험업계의 보험 사기 피해를 방지하는 데 중요한 역할을 할 것으로 기대한다.

### Abstract

Recently, there has been a sharp rise in insurance fraud based on real vehicle accidents, where the details of the incident are distorted, or the extent of damage is exaggerated to claim higher insurance payouts. With the increasing sophistication of techniques leveraging AI-generated image models to exaggerate damages, detecting tampered images has become more challenging. In response, this study proposes a methodology that not only utilizes existing models for detecting AI-generated images but also verifies the consistency of multiple images to identify contradictions. The proposed multi-image consistency verification approach enhances the ability to detect tampered images and is expected to play a significant role in preventing insurance fraud in the industry.

### Key words

deep-learning, AI-generated image detection, multi-image consistency verification

## 1. 서 론

보험 사기는 매년 증가하는 문제로, 그 규모와 수법이 점점 더 정교해지고 있다. 특히 실제 사고를 기반으로 사고의 내용을 왜곡하거나 피해 정도를 과장하여 보험금을 청구하는 사고 내용 조작의 문제가 커지는 중이다. 최근 AI를 활용한 이미지 생

성 모델의 발전으로 더 교묘하게 피해를 과장하는 경우가 잦아지고 있다. 기존 손해 사정 시스템은 AI 이미지 생성 모델이나 이미지 조작 기술을 이용한 위변조 이미지를 탐지하기에는 한계가 있다. 이러한 상황은 보험업계에 새로운 도전 과제가 되고 있으며, 위변조 이미지 탐지를 위한 정교한 기술의 필요성이 점차 증가하였다.

본 연구에서는 다중 이미지 정합성 검증 기법을

\* 국민대학교 경영대학 학사과정, Email: {lala0708, ys02289, myr2021, alstj08077}@kookmin.ac.kr

\*\* 국민대학교 경영대학 교수(교신저자), Email: ngkim@kookmin.ac.kr

활용한 위변조 이미지 탐지 방법을 제안한다. 제안된 방법은 사고 현장에서 촬영된 이미지 쌍을 사용하여, 서로 다른 각도에서 촬영된 이미지를 대조함으로써 이미지의 진위 여부를 파악하는 방식이다. 먼저, CLIP 방식으로 학습한 이미지 인코더를 통해 이미지의 정보와 패턴을 포착한 후, 단층 네트워크를 통해 각 이미지가 AI로 생성된 이미지인지 탐지한다. 추가적으로 다중-뷰 이상 탐지 기술을 기반으로, 두 개 이상의 각도에서 촬영된 사고 이미지 쌍을 활용하여 두 이미지의 정합성을 검증하는 다중 이미지 정합성 검증 방법론을 제안한다.

## II. 관련 연구

### 2.1 기존 생성 이미지 탐지 기술

GAN(Generative Adversarial Networks) 기반 생성 이미지는 주파수 대역에서 나타나는 노이즈 및 텍스처 불규칙성을 통해 탐지할 수 있으며, Diffusion 모델은 노이즈 제거 과정에서 발생하는 픽셀 분포의 불균형을 기반으로 탐지된다 [1, 2]. 이러한 기술들은 딥페이크 탐지에 주로 활용되지만, 생성 모델이 고도화되면서 탐지 성능의 한계가 드러나고 있다.

### 2.2 CNN 및 ViT 기반 이미지 탐지 연구

CNN(Convolutional Neural Networks)은 이미지의 지역적 특징을 분석하여 GAN이 생성한 이미지의 텍스처 불규칙성을 탐지하는 데 효과적이다[1]. 지역 단위의 패턴을 잘 탐지하는 장점이 있으나, 전체 이미지의 문맥이나 전역적 패턴을 학습하는 데에는 한계가 있다. 이를 보완하기 위해 도입된 ViT(Vision Transformer)는 이미지 전체를 패치로 나누어 전역적 특징을 학습하는 방법으로, 고해상도 이미지에서도 우수한 성능을 발휘한다[3].

### 2.3 Universal Fake Detect 연구

Universal Fake Detect는 사전 학습된 ViT 기반 모델을 활용하여 다양한 생성 모델(GAN, Diffusion 등)에서 생성된 이미지를 탐지하는 방식이다. CLIP(Contrastive Language-Image Pretraining) 기반의

ViT 모델을 사용해 이미지의 특성을 인코딩한 후, KNN 알고리즘 및 리니어 분류기를 통해 탐지를 수행한다[3]. 이 모델은 다양한 생성 모델에 대한 일반화된 탐지 성능을 제공하며, 새로운 데이터셋이나 이미지의 생성 모델에도 유연하게 적용될 수 있는 장점이 있다.

## III. 다중 이미지 정합성 검증

본 장에서는 다중의 사고 이미지를 비교하여 생성된 이미지의 모순을 판단하는 모델의 프로세스를 설명한다. 이 시스템은 AI-Generated Image Detection 모델과 Multi-View Anomaly Detection 모델을 활용하여 위변조 이미지를 탐지한다. 사용자가 입력한 복수의 사고 이미지가 AI에 의해 생성된 것인지 판별하고, 이미지 쌍 간의 정합성을 검증하는 방식이다.

<그림 1>의 Phase 1은 AI-Generated Image Detection 모델을 통해 사용자로부터 입력 받은 복수의 사고 이미지가 실제 이미지인지 AI가 생성한 이미지인지 탐지한다. 이 과정에서 실제 사고 이미지와 생성된 사고 이미지로 학습한 분류기를 활용하여 이미지를 탐지한다. 이후 <그림 2>와 같이 이미지 인코더인 CLIP/ViT를 통해 각 이미지를 벡터로 추출한다. 추출된 벡터는 단일 레이어로 구성된 분류기를 통해 해당 이미지가 생성된 이미지인지 실제 이미지인지 판별한다.

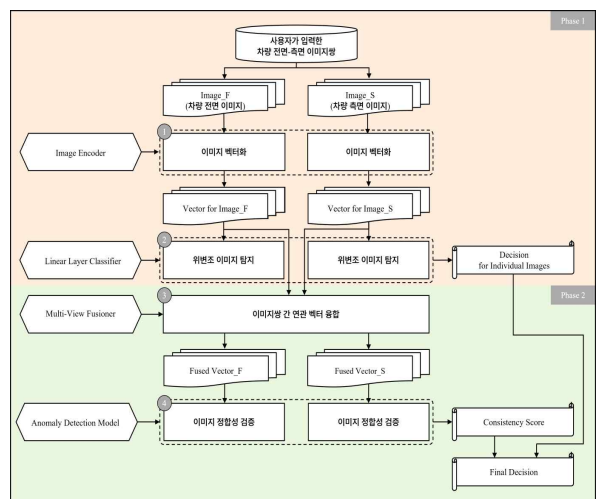


그림 1. 제안 방법론 개요  
Fig. 1. Overall Research Process



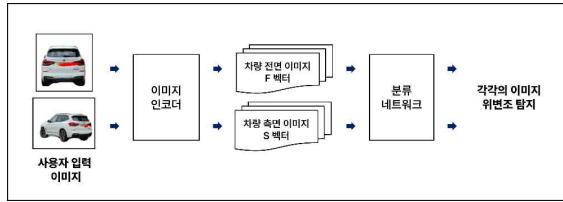


그림 2. AI 생성 이미지 탐지 모델 구조

Fig. 2. Structure of AI Generated Image Detection Model

이후 Phase 2에서는 <그림 3>과 같이 이전 단계에서 추출된 이미지 벡터를 이미지쌍 융합 모델에 입력한다. 이미지쌍 융합 모델은 두 이미지 사이의 상관관계를 포착하고, 이를 통해 벡터를 융합한다 [4]. 융합으로 인하여 각각의 이미지에 함께 입력된 이미지 정보가 추가된다. 이미지쌍 융합 모델은 ViT 모델에 두 이미지를 함께 입력하여 융합된 벡터를 구하는 방식과 Transformer 모델의 Cross-Attention 구조를 활용하여 두 이미지의 벡터를 융합하는 2가지 방식을 제안한다.

마지막으로, 융합된 이미지 벡터를 이미지 정합성 검증기에 입력하여 이미지 간에 서로 모순되는 상황, 즉 이상이 있는지 탐지한다. 정합성 검증기는 이상 탐지 모델을 기반으로 하여 융합된 각 벡터를 디코딩한다. 이 결과를 인코딩 과정의 벡터와 비교하며 기존 모순점에 대한 정합성 점수를 계산하고, 이를 통해 이미지의 정합성을 최종적으로 검증한다.

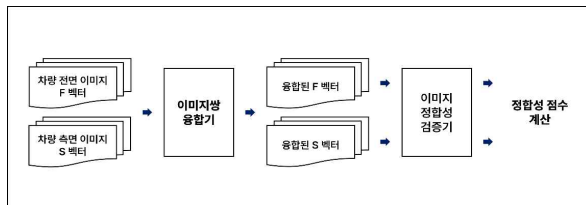


그림 3. 다중 이미지 정합성 검증 모델 구조

Fig. 3. Structure of Multi-Image Consistency Verification Model

#### IV. 결 론

본 연구에서는 다중의 사고 이미지를 비교하여 AI가 생성한 이미지의 모순 여부를 판별하는 시스템의 프로세스를 제안하였다. 제안 방법론은 보험 사기를 방지하는 데 중요한 역할을 할 것으로 기대되며, 보험업계의 위변조 이미지 탐지 능력을 향상

시킬 수 있을 것이다. 또한, AI 기반 자동화 탐지 시스템의 도입은 인적 자원의 효율성을 높이고, 보험 사기로 인한 손실을 줄이는 데 기여할 수 있다.

#### 참 고 문 헌

- [1] Wang, S. Y., Wang, O., Zhang, R., Owens, A., & Efros, A. A. (2020). CNN-Generated Images Are Surprisingly Easy to Spot... for Now, CVPR 2020, 8692-8701.
- [2] Wang, Z., Bao, J., Zhou, W., Wang, W., Hu, H., Chen, H., & Li, H. (2023). Dire for Diffusion-Generated Image Detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 22445-22455).
- [3] Ojha, U., Li, Y., & Lee, Y. J. (2023). Towards Universal Fake Image Detectors that Generalize Across Generative Models, CVPR 2023, 24480-24489.
- [4] Feng, C., Danier, D., Zhang, F., & Bull, D. (2024). MVAD: A Multiple Visual Artifact Detector for Video Streaming. arXiv preprint arXiv:2406.00212.

# 로봇 조립 성능 향상을 위한 심층 강화학습 기반 보상 스케일링 함수 비교 실험

박형근\*, 박원우\*, 홍승윤\*, 김병찬\*, 신지섭\*\*, 유승환\*\*, 윤종완\*\*\*, 고병진\*\*\*\*, 박태준\*

## Comparison of Reward Scaling Functions Based on Deep Reinforcement Learning for Improving Robotic Assembly Performance

Hyoung-Geun Park\*, Wonwoo Park\*, Seungyoun Hong\*, Byeongchan Kim\*, Jisub Shin\*\*,  
Seunghwan Yu\*\*, Jong-Wan Yoon\*\*\*, Byungjin Ko\*\*\*\*, and Taejoon Park\*

### 요 약

본 연구에서는 로봇 조립 작업을 위한 강화학습에서 조립 작업 성능을 개선하기 위하여 역수 보상 스케일링 함수를 제안한다. 기존의 보상 스케일링 방법, 특히 탄젠트 하이퍼볼릭 함수는 일반화와 성능에서 한계를 보였다. 우리는 역수 함수(Reciprocal)와 지수 감소 함수(Exponential Decay)를 포함한 5가지 비선형 스케일링 함수를 도입하고, 이들의 삽입 성공률에 미치는 영향을 평가하였다. 실험 결과, 역수 함수는 탄젠트 하이퍼볼릭 함수에 비해 삽입 성공률을 향상시켜, 접촉이 많은 작업에서 시뮬레이션과 현실 간의 간극을 줄이는 더 효과적인 방법임을 시사한다.

### Abstract

In this study, we propose a reciprocal reward scaling function to improve performance in reinforcement learning-based robotic assembly tasks. Traditional reward scaling methods, particularly the tangent hyperbolic function, showed limitations in generalization and performance. We introduce five non-linear scaling functions, including the reciprocal and exponential decay functions, and evaluate their effects on insertion success rates. Experimental results demonstrate that the reciprocal function outperforms the tangent hyperbolic function, suggesting that it is a more effective approach to narrowing the sim-to-real gap in contact-rich tasks.

### Key words

Robotic Reinforcement Learning, Robotic Assembly, Non-Linear Reward Scaling, Simulation-based Robot Learning

\* 한양대학교 로봇공학과, bborang3505@hanyang.ac.kr, Neal0907@hanyang.ac.kr, hong0807@hanyang.ac.kr, kchan124@hanyang.ac.kr, taejoon@hanyang.ac.kr(공동 교신저자)

\*\* 한양대학교 인공지능융합학과, mimak98@hanyang.ac.kr, sinanju06@hanyang.ac.kr

\*\*\* 한양대학교 지능형로봇학과, jongwanyoon@hanyang.ac.kr(공동 교신저자),

\*\*\*\* 한양대학교 스마트융합공학부, byungjinko@hanyang.ac.kr(공동 교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업(IITP-2024-RS-2020-II201741, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 인공지능융합혁신인재양성사업(IITP-2024-RS-2022-00155885, 50%) 연구임

## 1. 서 론

강화학습(Reinforcement Learning)은 복잡한 작업 환경에서 로봇이 자율적으로 효율적인 제어 정책을 학습하는 데 중요한 기술로 자리잡고 있다[1]. 최근에는 NVIDIA의 IsaacGym[2]과 IsaacLab[3]과 같은 고성능 시뮬레이션 플랫폼이 강화학습을 위한 가상 환경을 제공함으로써, 물리적 로봇 없이도 다양한 작업을 훈련할 수 있다. 이러한 시뮬레이션 기반 학습은 실제 로봇 시스템에서의 학습 비용을 줄이는데, 기여하지만, 여전히 현실과 시뮬레이션 간의 간극을 해결해야 할 과제가 남아 있다.

침투(interpenetration)는 시뮬레이션 상의 조립 작업에서 발생하는 비정상적인 현상으로, 두 물체가 실제로는 중첩되지 않아야 하지만, 시뮬레이터의 물리 모델링의 한계로 인해 서로 겹쳐지는 상황을 의미한다. 이러한 침투 현상은 시뮬레이션에서 실제 세계로 전이될 때 문제가 발생할 수 있다. 이러한 문제를 해결하기 위해 선행 연구인 IndustReal[3]에서는 SAPU(Simulation-Aware Policy Update)라는 기법을 제안했다. SAPU는 시뮬레이션에서 발생하는 비정상적인 침투를 감지하고, 이를 억제하는 정책을 학습함으로써 시뮬레이션에서 학습한 정책이 현실에서 성공적으로 전이될 수 있도록 한다. 이를 위해 SAPU는 침투 거리가 증가할수록 보상이 급격히 감소하도록 설계되었다. 특히, 선행 연구에서는 탄젠트 하이퍼볼릭 함수가 SAPU의 보상 스케일링에 사용되었다. SAPU에서 사용된 고정된 보상 스케일링 방식은 다양한 시나리오에서 84.9%에서 87.6%의 성공률을 기록하였지만, 이 방식에 대한 추가적인 개선 가능성은 여전히 존재한다. 특히, 다른 비선형 보상 스케일링 함수를 적용하는 연구가 필요하다.

본 연구에서는 IndustReal의 SAPU 보상 구조의 한계를 극복하기 위해 다섯 가지 비선형 스케일링 함수를 제안한다. 제안하는 스케일링 함수는 지수 감소 함수 (Exponential Decay), 역수 함수 (Reciprocal), 로지스틱 함수 (Logistic), 제곱 함수 (Quadratic), 그리고 선행 연구의 탄젠트 하이퍼볼릭 함수(Tangent Hyperbolic)로, 각각 침투 거리에 따른

보상을 조정한다. 본 연구의 목적은 이러한 다양한 비선형 스케일링 함수들이 조립 작업 성능에 미치는 영향을 비교 및 분석함으로써, 최적의 스케일링 함수를 찾는 것이다. 이를 통해 강화학습 기반 로봇 조립 제어 정책의 정확성이 한층 더 향상될 수 있을 것으로 기대된다.

## II. 방법론: 보상 스케일링 함수 설계

본 연구에서는 로봇 조립 작업에서 발생하는 침투를 억제하고 작업 성공률을 높이기 위해 다섯 가지 비선형 보상 스케일링 함수를 제안한다. 이 함수들은 침투 거리가 클수록 보상을 줄이는 방식으로 설계되었다. 각 함수의 수학적 정의는 다음과 같으며  $d_{ip}^{max}$ 는 최대 침투 거리,  $\epsilon_{ip}$ 는 침투 임계값이다. :

- 지수 감소 함수 (Exponential Decay):

$$reward_{exp} = e^{-\frac{d_{ip}^{max}}{\epsilon_{ip}}} \quad (1)$$

- 역수 함수 (Reciprocal):

$$reward_{reciprocal} = \frac{1}{1 + \frac{d_{ip}^{max}}{\epsilon_{ip}}} \quad (2)$$

- 로지스틱 함수 (Logistic):

$$reward_{log} = 1 - \frac{1}{1 + e^{-2\left(\frac{d_{ip}^{max}}{\epsilon_{ip}} - 0.5\right)}} \quad (3)$$

- 제곱 함수 (Quadratic):

$$reward_{quad} = 1 - \left(0.2 \cdot \frac{d_{ip}^{max}}{\epsilon_{ip}}\right)^2 \quad (4)$$

- 기준 함수 (tangent hyperbolic):

$$reward_{tanh} = 1 - \tanh\left(\frac{d_{ip}^{max}}{\epsilon_{ip}}\right) \quad (5)$$

이 다섯 가지 함수는 침투 거리를 기준으로 보상을 차등화하여 조립 작업의 성공률을 높이고, 비정상적인 침투를 최소화하기 위한 목적으로 설계되었다. 그림 1은 침투 거리에 따른 비선형 보상 스케일링 함수들의 그래프를 비교하여 각 함수의 특징을 시각적으로 나타낸다. 본 연구에서는 이 함수들이 실제로 로봇 조립 작업 성공률에 어떻게 영향을 미치는지 실험을 통해 정량적으로 분석하고자 한다.

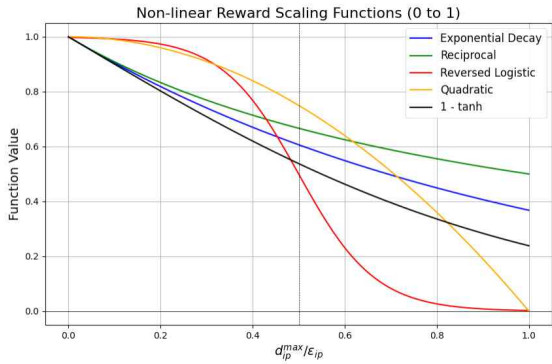


그림 1. 비선형 보상 스케일링 함수 비교 그래프  
Fig. 1. Comparison of non-linear reward scaling functions

### III. 실험

#### 3.1 실험 환경

본 연구에서는 IndustRealTaskPegsInsert라는 시뮬레이션 환경을 사용하여 비선형 보상 스케일링 함수들의 성능을 비교 평가하였다. 5가지 비선형 보상 스케일링 함수를 비교하기 위해 학습을 진행하였다.

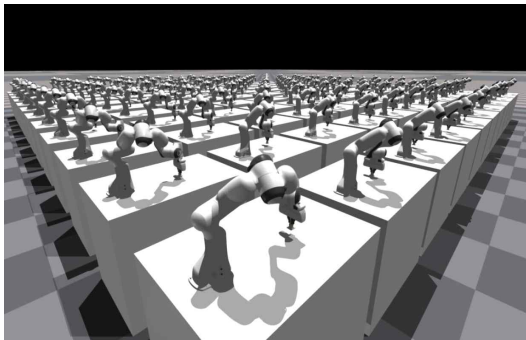


그림 2. IsaacGym에서 학습 중인 다수의 로봇 에이전트  
Fig. 2. Multiple robot agents learning in IsaacGym

학습 환경은 NVIDIA의 IsaacGym을 기반으로 구축되었으며, 로봇은 핀 삽입 작업을 수행하는 시나

리오에서 강화학습을 통해 학습되었다. 각 스케일링 함수는 침투 거리  $d_{ip}^{max}$ 에 따라 보상을 감소시키는 방식으로 설계되었으며, 이외의 실험 환경과 학습 조건은 동일하게 설정되었다.

#### 3.2 실험 설정

본 실험에서는 Proximal Policy Optimization (PPO) 알고리즘[5]을 사용하여 로봇의 학습을 진행하였으며, 각 비선형 보상 스케일링 함수는 150 에포크 동안 동일한 학습 시간을 가지며 충분한 훈련이 이루어졌다. 이는 각 함수별로 2000 스텝 이상의 학습을 포함하며, 학습이 안정화되는 시점을 고려한 것이다. 성능 평가를 위해 삽입 성공률, 평균 삽입 성공률, 보상을 주요 지표로 사용하였다.

#### 3.3 실험 결과

다섯 가지 비선형 스케일링 함수를 사용하여 학습한 결과는 그림 3과 같다. :

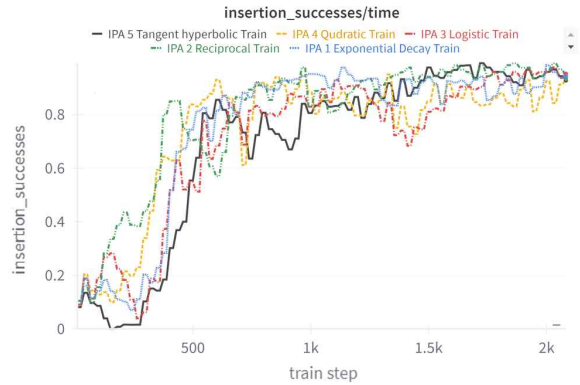


그림 3. 학습 중 삽입 성공률의 변화  
Fig. 3. Insertion success rate over time during training

표 1. 비선형 보상 스케일링 함수들의 성능 비교  
Table 1. Performance Comparison of Non-Linear Reward Scaling Functions

함수	삽입 성공률 (중앙값)	평균 삽입 성공률 (중앙값)
Exp Decay	0.851	0.857
Reciprocal	0.928	0.916
Logistic	0.768	0.792
Quadratic	0.718	0.704
Tangent Hyperbolic	0.920	0.895

각 스케일링 함수에 대해 학습한 모델을 토대로 16회 테스트를 통해 산출된 삽입 성공률, 평균 삽입 성공률, 그리고 보상의 중앙값은 표 1과 같다.

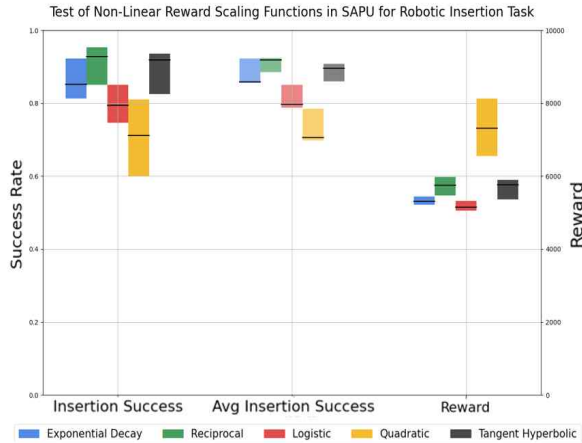


그림 4. 비선형 보상 스케일링 함수에 따른 실험 결과  
Fig. 4. Experimental results comparison based on non-linear reward scaling functions

Reciprocal 함수가 가장 높은 삽입 성공률(0.928)을 기록했으며, Tangent Hyperbolic 함수(0.920)도 유사한 성능을 보였다. Logistic 함수와 Quadratic 함수는 상대적으로 낮은 성공률을 기록하였다. Reciprocal 함수는 평균 삽입 성공률에서도 가장 높은 수치를 기록하였으며 (0.916637), Tangent Hyperbolic 함수가 그 뒤를 이었다. Logistic 및 Quadratic 함수는 성공률이 비교적 낮았다. Quadratic 함수는 삽입 성공률에서는 다소 낮은 성능을 보였지만, 높은 보상을 기록하였다. 이는 보상 구조 상 침투 거리에 따른 보상 감소율이 상대적으로 낮아 보상을 더 많이 받을 수 있음을 의미한다. 반면, Logistic 함수는 가장 낮은 보상을 기록하였다.

#### IV. 결론

본 연구에서는 비선형 보상 스케일링 함수들이 강화학습 기반 로봇 조립 작업에 미치는 영향을 분석하였다. 실험 결과 Reciprocal 함수와 Tangent Hyperbolic 함수가 삽입 성공률과 평균 삽입 성공률에서 우수한 성능을 보였다. 특히, Reciprocal 함수는 기존 연구의 Tangent Hyperbolic 함수보다 더 나

은 성능을 보여주었으며, Quadratic 함수는 높은 보상을 제공하였지만, 삽입 성공률 측면에서는 다소 아쉬운 성과를 보였다.

실험 결과, 역수 함수가 최적의 스케일링 함수임을 확인하였다. 이를 바탕으로, 로봇이 시뮬레이션 환경에서 학습한 정책을 현실 세계로 효과적으로 전이할 수 있음을 시사한다.

#### 참 고 문 헌

- [1] S. Levine et al., "Deep Reinforcement Learning for Robotics: A Survey of Real-World Successes," arXiv preprint, arXiv:2408.03539, 2024.
- [2] V. Makoviychuk et al., "Isaac Gym: High Performance GPU-Based Physics Simulation for Robot Learning," arXiv preprint, arXiv:2108.10470, 2021.
- [3] V. Prasad et al., "Orbit: A Unified Simulation Framework for Interactive Robot Learning Environments," in IEEE Robotics and Automation Letters, vol. 8, no. 2, pp. 698-705, April 2023.
- [4] B. Tang et al., "IndustReal: Transferring contact rich assembly tasks from simulation to reality," in Robotics: Science and Systems, 2023.
- [5] J. Schulman et al., "Proximal Policy Optimization Algorithms," arXiv preprint, arXiv:1707.06347, 2017.

# SAPU 가중치 함수와 필터-패널티를 사용한 심층 강화학습 기반 로봇 조립 시스템

박원우\*, 박형근\*, 홍승윤\*, 김병찬\*, 신지섭\*\*, 유승환\*\*, 윤종완\*\*\*, 고병진\*\*\*\*, 박태준\*

## SAPU Weight Function and Filter-Penalty Based Deep Reinforcement Learning Robot Assembly System

Wonwoo Park\*, Hyoung-Geun Park\*, Seungyoun Hong\*, Byeongchan Kim\*, Jisub Shin\*\*,  
Seunghwan Yu\*\*, Jong-Wan Yoon\*\*\*, Byungjin Ko\*\*\*\*, and Taejoon Park\*

### 요약

로봇 조립 작업은 높은 정밀도와 정확성을 요구하는 오랜 과제이다. 이를 해결하기 위해 시뮬레이션 기반 강화학습이 생산 환경의 유연성 향상과 비용 절감의 해결책으로 제시되었다. 본 연구에서는 시뮬레이션에서 로봇 조립 작업을 수행할 때 조립 대상 간 발생하는 상호 침투를 최소화하는 SAPU 알고리즘에서 기존 가중치 함수 이외에 두 가지 가중치 함수와 기존 Filter 대신 Filter-Penalty 방법론을 제안한다. 실험 결과, Gaussian 및 Sigmoid 가중치 함수와 Filter-Penalty 기법이 삽입 성공률을 크게 향상시키고 시드별 성공률 편차를 줄이는데 기여함을 확인하였다.

### Abstract

Robot assembly tasks require high precision and accuracy, making them a longstanding challenge. To address this, simulation-based reinforcement learning has been proposed as a solution to enhance flexibility and reduce costs in manufacturing environments. In this study, we propose two additional weight functions to complement the existing weight function in the SAPU algorithm, which is designed to minimize interpenetration between assembly parts during robotic assembly in simulation. Furthermore, instead of the conventional filtering method, we introduce the Filter-Penalty approach. Experimental results demonstrate that the Gaussian and Sigmoid weight functions, along with the Filter-Penalty method, significantly improve insertion success rates and reduce variance in success rates across different random seeds.

### Key words

Simulation, Robot Assemble, Insertion task, Deep Reinforcement Learning, Reward Shaping

\* 한양대학교 로봇공학과, neal0907@hanyang.ac.kr, bborang3505@hanyang.ac.kr, hong0807@hanyang.ac.kr, kchan124@hanyang.ac.kr, taejoon@hanyang.ac.kr(공동 교신저자)

\*\* 한양대학교 인공지능융합학과, mimak98@hanyang.ac.kr, sinanju06@hanyang.ac.kr

\*\*\* 한양대학교 지능형로봇학과, jongwanyoon@hanyang.ac.kr(공동 교신저자),

\*\*\*\* 한양대학교 스마트융합공학부, byungjinko@hanyang.ac.kr(공동 교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업(IITP-2024-RS-2020-II201741, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 인공지능융합혁신인재양성사업(IITP-2024-RS-2022-00155885, 50%) 연구임

## 1. 서 론

로봇 조립은 접촉이 많은 상호작용과 높은 정밀도 및 정확성이 요구되는 과제로, 이는 로봇 분야에서 해결해야 할 도전 과제이다. 로봇 조립 과제에서 높은 정밀도를 달성하기 위해 정밀도가 높은 하드웨어에 의존해왔다. 이와 같은 의존성은 로봇 설비 비용이 높아지는 문제로 이어진다. 또한 다품종 소량 생산의 필요성이 대두되면서, 다양한 부품, 자세, 환경에 대한 적응력이 요구되고 있다. 이러한 높은 설비 비용과 환경에 대한 적응력을 높이기 위해, 시뮬레이션과 심층 강화학습(Deep Reinforcement Learning)을 활용한 로봇 조립 정책의 학습이 해결 방안으로 제안되었다[1]. 이 방법은 고가의 센서(e.g. 토크 센서)가 없어도 조립 정책을 학습할 수 있기 때문에, 설비 비용을 절감하면서도 다양한 부품에 대한 적응력을 학습할 수 있다[2].

하지만, 시뮬레이션 환경에서는 조립 대상 물체 간의 상호 침투 현상 문제가 발생한다. 이 문제는 시뮬레이션에서 발생하는 비정상적인 동역학을 에이전트가 악용하여 비정상적인 정책을 학습하는 원인이 된다[3]. 이러한 문제를 해결하기 위해, 정확한 조립작업을 할 수 있는 IndsutReal[4]이 제안되었다. 특히, 물체 간의 상호 침투 거리에 따라 보상을 제공하는 Simulation-Aware Policy Update (SAPU) 알고리즘을 통해, 에이전트가 비정상적인 동역학을 악용하는 문제를 완화했다. SAPU는 물체 간의 상호 침투 거리에 따라 보상을 조정하여 침투 거리가 증가할수록 보상이 감소하도록 설계가 되었다. SAPU에서 사용된 스케일링 가중치 함수로는 Tanh 함수가 쓰였으며, 침투 거리가 일정 임계치를 넘어가면 보상값을 반환하지 않는 Filter가 존재한다. 이를 통해 86%의 삽입 성공률 성능을 달성하였다. 그러나 기존 스케일링 가중치 함수로 쓰인 Tanh 함수는 추가적인 파라미터 튜닝이 어려워 미세한 보상 조정이 힘들다. 또한 보상 값을 반환하지 않는 Filter는 에이전트가 잘못된 행동을 했을 경우 즉각적인 페널티가 없을 경우, 에이전트는 잘못된 행동을 반복할 수 있다. 따라서 파라미터 튜닝이 가능한 가중치 함수와 Filter에서의 확실한 페널티가 필요하다.

본 논문에서는 시뮬레이션 환경에서 발생하는 조립 물체 간의 상호 침투를 방지하는 SAPU 알고리즘에서 삽입 성공률을 높이기 위해 기존 SAPU에 가중치 함수로 쓰인 Tanh 함수 이외에도 추가적인 파라미터 튜닝이 가능한 Sigmoid 함수, 그리고 Gaussian 함수를 사용한다. 또한 기존 SAPU에서 쓰인 Filter 대신 침투 거리가 일정 임계치를 넘어가면 보상값에 Penalty를 부여하는 Filter-Penalty 방법을 제안한다.

## II. 방법론

선행연구[4]에서는 SAPU에 스케일링 벡터에서 Filter와 가중치 함수를 사용하였다. 사용된 Filter와 가중치 함수는 Step에서 최대 상호 침투거리  $d_{ip}^{max}$  보다 허용된 최대 상호 침투 거리  $\epsilon_{ip}$ (1mm)이 클 경우 정책 업데이트에서 반환값을 사용하지 않고, 반대인 경우에만 반환값을 가중치 함수  $1 - \tanh(d_{ip}^{max}/\epsilon_{ip})$ 를 사용하여 가중치 업데이트를 진행한다. 수식 1을 보면 기존 가중치 함수 Tanh는 일반적으로 함수의 형태를 조절하기 위해 추가적인 파라미터 튜닝이 어렵다. 즉, 주어진 함수의 경사를 변경하거나 보상의 감소 속도를 조절하는 데 한계가 있다. 이를 해결하기 위해, 파라미터 튜닝을 통해 기울기를 조절할 수 있는 Gaussian 및 Sigmoid 가중치 함수를 사용하였다. 수식 2에서 Gaussian 파라미터는  $\sigma = 0.001$ 를 수식 3에서 Sigmoid 파라미터는  $k = 0.0001$ 를 사용하였다. 이는 그림 1과 같은 그래프 형태로 확인할 수 있다.

$$\text{Tanh}(d_{ip}^{max}) = 1 - \tanh(d_{ip}^{max}/\epsilon_{ip}) \quad (1)$$

$$\text{Gaussian}(d_{ip}^{max}) = \exp\left(-\frac{(d_{ip}^{max})^2}{2\sigma^2}\right) \quad (2)$$

$$\text{Sigmoid}(d_{ip}^{max}) = \frac{1}{1 + \exp\left(\frac{d_{ip}^{max} - \epsilon_{ip}}{k}\right)}$$

(3)

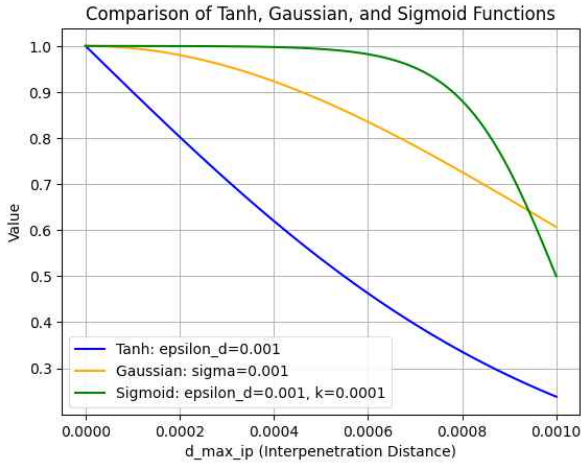


그림 443. 가중치 함수 Tanh, Gaussian, Sigmoid에 따른 보상

Fig. 1. Rewards based on Tanh, Gaussian, and Sigmoid weighting functions

기존 Filter는  $d_{ip}^{max} < \epsilon_{ip}$  인 경우에만 업데이트 반환값을 사용한다. 대신, 우리는 Filter-Penalty를 제안한다. 제안하는 Filter-Penalty란  $d_{ip}^{max} > \epsilon_{ip}$  인 경우에 보상이 적어질 수 있도록 이전 Step 보상값 (prev\_step\_reward)에 Penalty를 부과하여 보상을 업데이트 한다. 이를 통해 물체 간 침투와 반환된 보상값의 연관성을 강화시켜, 학습 진행 시 RL 에이전트가 침투를 최대한으로 피하도록 유도한다.

**Algorithm 1** Filter-Penalty Algorithm

```

for every step do
  checking interpenetration and compute reward
  if  $d_{max}^{ip} > \epsilon_{ip}$  then
    reward = penalty × prev_step_reward
  else
    reward = prev_step_reward
  end if
end for
    
```

그림 444. Filter-Penalty 알고리즘  
Fig. 2. Filter-Penalty Algorithm

III. 실험

실험 환경은 NVIDIA IsaacGym[5] 시뮬레이터를

이용하여 구축되었다. 그림 3과 같은 환경에서 선행 연구의 기존 파라미터를 이어 사용하였고 로봇 머니플레이터는 Franka Emika Panda가 사용되었으며 일정 범위에 랜덤으로 위치되는 막대를 소켓에 삽입시키는 작업이다. 학습은 Proximal Policy Optimization (PPO) 알고리즘[6]을 사용하여 로봇 128개가 150 epoch 동안 정책을 학습하고 이 중 가장 큰 보상 값을 갖는 에피소드를 체크포인트에 저장한 이후 테스트에서 삽입 성공률로 성능을 평가한다. 모든 실험 결과는 선행연구[4], Gaussian/Filter-Penalty, Sigmoid/Filter-Penalty 총 3종류의 학습을 각 3개의 시드로 학습하고 각 학습 종류마다 학습된 평균을 나타낸 것이다. 테스트는 학습된 시드당 로봇 128개가 16번의 에피소드에서 나온 결과를 나타낸 것이며 학습과 마찬가지로 3개의 테스트 결과의 평균이다.

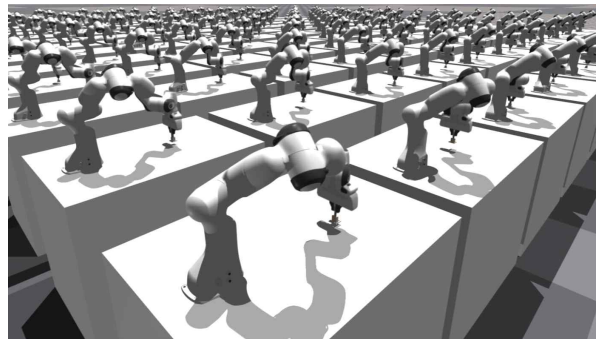


그림 445. IsaacGym 시뮬레이터 테스트 환경

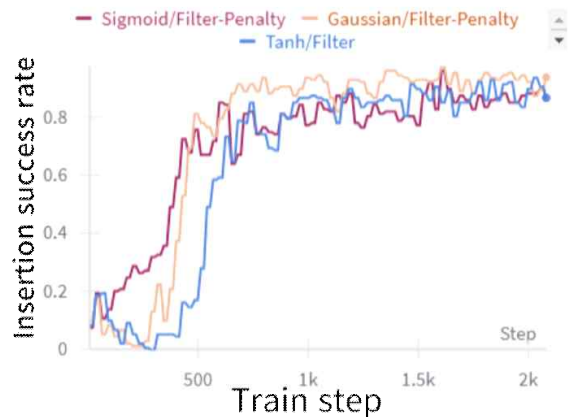


그림 446. 학습 중 삽입 성공률의 변화 그래프  
Fig. 4. Graph of training insertion success rate

그림 4를 보면 학습 시 삽입 성공률을 확인할 수



있는데 Filter-Penalty가 적용된 Gaussian이 전반적으로 가장 높고 그 뒤로 Filter-Penalty가 적용된 Sigmoid와 Tanh가 비슷하게 나왔다.

#### IV. 결 과

테스트 결과인 표 1과 그림 4를 보면 Gaussain 가중치 함수와 Filter-Penalty가 적용된 학습이 테스트 결과 평균 삽입률 92.3%로 Tanh 가중치 함수가 적용된 학습 테스트 결과인 81%보다 약 10% 증가한 것을 확인할 수 있다. 또한 기존 삽입 성공률의 편차 7% 평균 삽입 성공률의 편차 4%에 비해 삽입 평균 성공률의 편차 2%, 평균 삽입 성공률의 편차 0.3%로 적어짐을 확인할 수 있다. 이는 시드별 학습 편차가 제일 적음을 나타낸다.

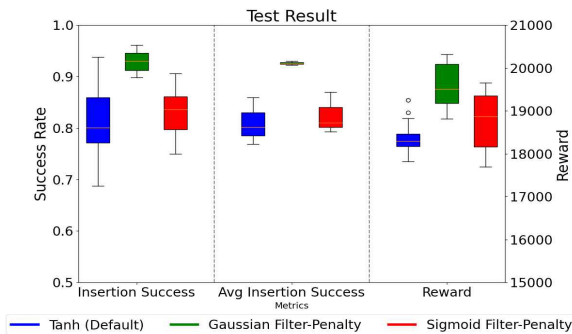


그림 447. 테스트 결과 그래프  
Fig. 5. Test Result Graph

결론적으로 변경한 가중치 함수와 Filter-Penalty는 삽입 성공률을 증가시켰으며, 그중에서도 Gaussian과 Filter-Penalty가 적용된 학습이 테스트 결과에서 92.3%로 가장 높은 삽입 성공률을 보여준다.

표 1. 테스트 결과 성능 지표  
Table 1. Test Result Performance Metrics

방법	삽입 성공률	평균 삽입 성공률	평균 보상
Tanh (Default)	0.8 ± 0.07	0.81 ± 0.04	18313 ± 523
Gaussian Filter/Penalty	0.923 ± 0.02	0.923 ± 0.003	19451 ± 405
Sigmoid Filter/Penalty	0.836 ± 0.04	0.8242 ± 0.03	18778 ± 496

#### V. 결 론

본 연구에서는 로봇 조립 작업 그 중에서도 접촉면이 많은 삽입 작업의 성공률을 높이고, 시뮬레이션 환경에서 조립 대상 물체 간의 상호 침투 현상을 방지하기 위한 Simulation-Aware Policy Update (SAPU) 기법을 개선하는 방법을 탐구하였다. 특히, 기존의 Tanh 함수 기반의 Filter와 달리 파라미터 튜닝이 가능한 Gaussian 및 Sigmoid 가중치 함수와 Filter-Penalty 기법을 도입하여, 삽입 성공률을 높이고 성공률 편차를 감소시켰다.

실험 결과, 제안한 방법론 중 Gaussian 가중치 함수와 Filter-Penalty가 함께 쓰였을 때 기존 방법론 삽입 성공률이 81%에서 92.3%로 크게 증가하였고, 시드별 삽입 성공률 편차도 7%에서 2%로 낮아졌다. 이로써 제안하는 방법론이 로봇 조립 작업의 정밀도 향상에 기여할 수 있음을 입증하였다.

#### 참 고 문 헌

- [1] S. Levine et al., "Deep Reinforcement Learning for Robotics: A Survey of Real-World Successes," arXiv preprint, arXiv:2408.03539, 2024.
- [2] J. Schulman et al., "Proximal Policy Optimization Algorithms," arXiv preprint, arXiv:1707.06347, 2017.
- [3] J. Skalse et al., "Defining and Characterizing Reward Hacking," arXiv preprint, arXiv:2209.13085, 2022.
- [4] B. Tang et al., "IndustReal: Transferring contact rich assembly tasks from simulation to reality," in Robotics: Science and Systems, 2023.
- [5] V. Makoviyshuk et al., "Isaac Gym: High Performance GPU-Based Physics Simulation for Robot Learning," arXiv preprint, arXiv:2108.10470, 2021.
- [6] V. Prasad et al., "Orbit: A Unified Simulation Framework for Interactive Robot Learning Environments," in IEEE Robotics and Automation Letters, vol. 8, no. 2, pp. 698-705, April 2023.

# 가중치 SDF 보상 함수를 이용한 로봇 조립을 위한 심층 강화 학습

김병찬\*, 박형근\*, 홍승윤\*, 박원우\*, 신지섭\*\*, 유승환\*\*, 윤종완\*\*\*, 고병진\*\*\*\*, 박태준\*

## Deep Reinforcement Learning for Robotic Assembly Using a Weighted SDF Reward Function

Byeongchan Kim\*, Hyoung-Geun Park\*, Seungyoun Hong\*, Wonwoo Park\*, Jisub Shin\*\*,  
Seunghwan Yu\*\*, Jong-Wan Yoon\*\*\*, Byungjin Ko\*\*\*\*, and Taejoon Park\*

### 요 약

본 연구는 로봇 삽입 작업에서 강화학습 정책의 삽입 정밀도를 증가시키기 위해, 가중치를 적용한 SDF 기반 보상 함수를 제안한다. 플러그와 소켓 간의 거리가 임계값 이하일 때 보상에 가중치를 적용하였고, 이를 통해 부품 삽입 성공률을 선행 연구보다 약 7% 향상시켰다. 본 실험 결과는 소켓 주변에서의 플러그를 삽입하지 못하고 주변에 충돌하는 문제에서 상당한 개선을 보인다. 여기서 제시된 작업은 향후 자동차, 항공우주를 포함한 로봇 조립 자동화가 필요한 공정에서 로봇 조립 작업의 정밀성을 높이는 데 기여한다.

### Abstract

This study proposes a weighted SDF-based reward function to increase insertion precision in reinforcement learning policies for robotic insertion tasks. Weighting was applied to the reward when the distance between the plug and socket was below a certain threshold, improving insertion success rates by approximately 7% over previous studies. Experimental results show significant improvements in reducing collisions around the socket when the plug fails to insert. The task presented here contributes to enhancing the precision of robotic assembly tasks in automation processes, including automotive and aerospace applications.

### Key words

Robotic insertion task, SDF (Signed Distance Field), Weight application

---

\* 한양대학교 로봇공학과, neal0907@hanyang.ac.kr, bborang3505@hanyang.ac.kr,  
hong0807@hanyang.ac.kr, kchan124@hanyang.ac.kr, taejoon@hanyang.ac.kr(공동 교신저자)

\*\* 한양대학교 인공지능융합학과, mimak98@hanyang.ac.kr, sinanju06@hanyang.ac.kr

\*\*\* 한양대학교 지능형로봇학과, jongwanyoon@hanyang.ac.kr(공동 교신저자),

\*\*\*\* 한양대학교 스마트융합공학부, byungjinko@hanyang.ac.kr(공동 교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업(IITP-2024-RS-2020-II201741, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 인공지능융합혁신인재양성사업(IITP-2024-RS-2022-00155885, 50%) 연구임

## I. 서론

로봇 조립 작업은 자동차, 항공우주, 전자 산업 등 다양한 분야에서 중요한 역할을 수행하며, 두 부품(e.g. 플러그와 소켓) 간의 정밀한 삽입 능력을 요구한다. 삽입 작업에서 두 부품 간의 위치와 방향에 대한 미세한 오차는 작업의 성공 여부에 결정적인 영향을 미친다. 이러한 오차를 줄이기 위해, 시행착오를 통해 조립 제어 정책을 학습할 수 있는 강화 학습을 사용할 수 있다. 또한 시행착오를 수행하는 비용을 줄이기 위해서 시뮬레이션을 사용한다. 그러나 시뮬레이션에서 발생하는 불안정한 동역학(e.g. 물체 간의 침투 현상)을 에이전트가 보상으로 악용할 수 있다. 이에 따라 에이전트가 보상을 악용하는 것을 방지하기 위해, 보상 함수를 정확하게 설계하는 것이 요구된다.

SDF(Signed Distance Field) 기반의 보상 알고리즘은 두 부품 간의 위치와 방향 오차를 동시에 고려하여 보상을 제공하기 때문에, 기존의 Key-point based reward와 Distance based reward 방식들보다 정밀한 조립 작업을 수행할 수 있다. SDF는 플러그 표면과 소켓 표면 간의 SDF 거리를 계산한다. 플러그 표면 내에 소켓 표면이 있으면 SDF 거리가 양의 값을 갖고, 반대의 경우에는 음의 값을 갖는다. 이를 통해 부품이 표면에 얼마나 가까운지, 얼마나 정확하게 정렬되어 있는지를 연속적으로 평가하고, 특히, 부품이 목표 표면에 가까워질수록 보상이 기하급수적으로 증가한다. 선행 연구[1]에서는 주어진 플러그 표면에서 여러 개의 점을 샘플링한 후, 각 점에 대해 목표 위치에 있는 소켓 표면의 SDF 값을 계산하여, 평균값을 음의 로그를 통해 부품이 목표 표면에 가까워질수록 보상이 크게 증가하도록 설계되었다. 그러나, 플러그가 소켓 근처에 있을 때 받는 보상과 부분적으로 삽입되었을 때 받는 보상의 차이가 크지 않다. 이는 소켓 주변에서의 플러그를 삽입하지 못하고 주변에 충돌하는 문제로 야기될 수 있다.

이러한 문제를 해결하기 위해, 본 연구는 두 부품 간의 거리가 임계값 이하일 때 가중치를 적용하는 새로운 보상 함수를 제안한다. 이는 소켓 주변에서 플러그가 주변에서 충돌하지 않고 소켓 안으로

삽입할 수 있게 적극적으로 유도함으로써, 삽입 작업의 성공률 향상에 기여한다.

## II. 방법론

SDF 값은 플러그와 소켓 간의 서명된 거리로 정의되며, 플러그가 소켓에 가까워질수록 양의 값에서 음의 값으로 변하게 된다. 이를 통해 삽입 작업의 진행 정도를 정밀하게 측정할 수 있다.

SDF 값의 계산은 각 타임 스텝에서, 플러그의 표면에서 샘플링한 점들의 SDF 값을 계산한다. 이 값은 플러그와 소켓 간의 거리 정보를 제공하며, 이를 바탕으로 로봇의 조립 작업 정도를 평가할 수 있다. SDF 보상 함수  $R$ 에서  $x_i$ 는 샘플링된 점이고,  $\Phi(x_i)$ 은 점  $x_i$ 의 SDF 거리 값을 의미한다.  $N$ 은 샘플링된 점의 개수이고, 논문 및 실험에서는 1000개를 사용한다.

$$R = -\log\left(\frac{1}{N} \sum_1^N \frac{\Phi(x_i)}{N}\right) \quad (1)$$

가중치 적용은 플러그가 소켓 사이의 거리가 임계값 이하일 때 적용된다. 임계 값은 0.5cm와 1cm로 설정하였다. 구체적으로 점  $x_i$ 의 SDF 거리 값인  $\Phi(x_i)$ 가 임계 값 이하 일 때  $\Phi(x_i)$ 를 0.5배로 계산하여, 수식 (1)의 음의 로그 진수 부분을 작게 만든다. 이는 플러그가 소켓에 부분적으로 삽입되었을 때 SDF 보상 값( $R$ )을 증가시켜, 정확한 삽입 작업을 유도한다.

## III. 실험

실험 환경은 NVIDIA Isaac gym[2] 시뮬레이터에서 제공하는 IndustReal 벤치마크를 사용하였다. 로봇 매니플레이터는 Franka Emika Panda를 사용하였다. 128개의 plug-socket 환경을 복제하여 학습 및 테스트를 진행하였다.

학습 모델은 PPO(Proximal Policy Optimization) [5]를 사용하였다. 각각의 방법들을 PPO 학습 모델로 150epoch 학습을 하였다. 삽입 성능 지표로는 삽입 성공률을(insertion\_successes) 사용하였다. 삽입 성공

횃수/전체 삽입 시도 횃수로 계산하였다. 실험 결과는 그림1과 같은 실험 환경에서 총 16번 테스트를 진행하였고, 삽입 성공률의 평균을 계산 하였다.



그림 448. test 환경

베이스라인 모델은 3가지를 사용하였다. Base는 가중치를 적용하지 않은 SDF기반 리워드 방식이다. Weighted SDF\_1은 그림2 사진에서 플러그 및 소켓의 거리가 임계값인 1cm 이하일 때 가중치를 적용하는 SDF기반 리워드 방식이다. Weighted SDF\_2은 임계값이 0.5cm 일 때 가중치를 적용하는 SDF기반 리워드 방식이다.

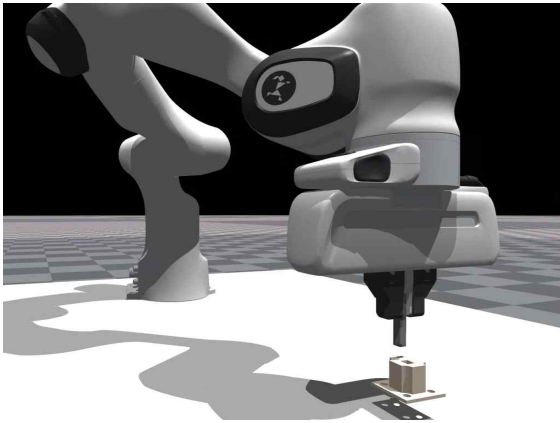


그림 449 로봇이 소켓에 플러그를 삽입하는 모습

#### IV. 결과

표 1은 Base, Weighted SDF\_1, Weighted SDF\_2의 16번의 test 결과의 평균이다. 실험 결과, Weighted SDF\_1이 삽입 성공률 94.6%, Weighted SDF\_2이 삽입 성공률 91.46%로 기존의 base 87.54% 보다 각각 약 7%, 4% 증가하였다.

표 1 테스트 결과(평균)

방법	학습 횃수 (epoch)	삽입 성공률 (%)
Base	150	87.54
Weighted SDF_1	150	94.60
Weighted SDF_2	150	91.46

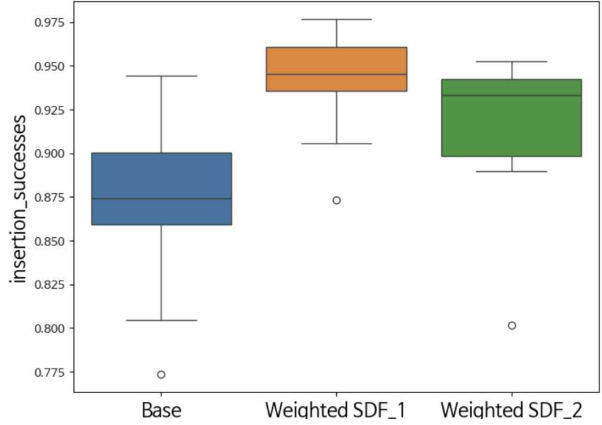


그림 450 테스트 결과 Boxplot 그래브

그림 3은 실험 결과를 Boxplot 그래프로 나타낸 것이다. Base의 성공률 편차가 약 7%, Weighted SDF\_1, Weighted SDF\_2의 성공률 편차가 약 5%로 성공률 편차도 적어졌다. 이러한 결과를 통해 선행 연구보다 로봇 삽입 작업의 정밀도가 향상되었음을 확인할 수 있다.

#### V. 결 론

본 논문에서는 시뮬레이션 환경에서 심층 강화학습 기반 로봇 조립 작업의 성능을 향상하기 위해 두 부품 간의 거리가 임계값 이하일 때 SDF 값에 가중치를 적용하는 방법을 제안하였다. 실험 결과 플러그와 소켓 간의 거리가 임계값 1cm 이하일 때 가중치 적용했을 때, 선행연구에서의 삽입 성공률이 87.54%에서 94.6% 크게 증가하였다. 이로써 제안하는 방법론이 로봇 조립 작업의 정밀도 향상에 기여할 수 있음을 입증하였다.

#### 참 고 문 헌

[1] IndustReal: Transferring Contact-Rich Assembly Tasks from Simulation to Reality, Bingjie Tang,

Michael A. Lin, Iretiayo Akinola, Ankur Handa,  
Gaurav S. Sukhatme, Fabio Ramos, Dieter Fox,  
Yashraj Narang

- [2] V. Makoviychuk et al., "Isaac Gym: High Performance GPU-Based Physics Simulation for Robot Learning," arXiv preprint, arXiv:2108.10470, 2021.
- [3] V. Prasad et al., "Orbit: A Unified Simulation Framework for Interactive Robot Learning Environments," in IEEE Robotics and Automation Letters, vol. 8, no. 2, pp. 698-705, April 2023.
- [4] B. Tang et al., "IndustReal: Transferring contact rich assembly tasks from simulation to reality," in Robotics: Science and Systems, 2023.
- [5] J. Schulman et al., "Proximal Policy Optimization Algorithms," arXiv preprint, arXiv:1707.06347, 2017.

# SDF 보상 스케일링을 활용한 시뮬레이션 환경에서의 심층 강화학습 기반 로봇 조립 작업

홍승윤\*, 박원우\*, 박형근\*, 김병찬\*, 신지섭\*\*, 유승환\*\*, 윤종완\*\*\*, 고병진\*\*\*\*, 박태준\*

## Deep Reinforcement Learning-Based Robotic Assembly in Simulation Environments Using SDF Reward Scaling

Seungyoun Hong\*, Wonwoo Park\*, Hyoung-Geun Park\*, Byeongchan Kim\*, Jisub Shin\*\*, Seunghwan Yu\*\*, Jong-Wan Yoon\*\*\*, Byungin Ko\*\*\*\*, Taejoon Park\*

### 요약

본 연구에서는 심층 강화학습 기반 로봇 조립 작업의 성능 향상을 위한 SDF 보상 스케일링을 활용한 방법론을 제안한다. 선행 연구의 시스템에서 사용된 SDF(Signed Distance Field)-기반의 Dense Reward를 개선하여 양수 영역에 대한 페널티를 도입하고, 보상 설계를 위해 SDF 보상의 로그 함수 수정과 보상 스케일링을 조정하였다. 이러한 개선을 통해, 플러그와 소켓 간의 접촉 시뮬레이션 환경에서 삽입 작업의 성공률이 최대 7%까지 향상되었다. 실험 결과는 제안된 방법론이 시뮬레이션 환경에서 로봇 조립 작업의 성공률을 효과적으로 높일 수 있음을 보여준다.

### Abstract

In this study, we propose a methodology that utilizes SDF (Signed Distance Field) reward scaling to improve the performance of robotic assembly tasks based on deep reinforcement learning. By enhancing the SDF-based dense reward system used in previous research, we introduce a penalty for positive regions and adjust the reward design by modifying the logarithmic function of the SDF reward and scaling the rewards. These improvements resulted in up to a 7% increase in the success rate of insertion tasks in a simulation environment involving plug-and-socket interactions. The experimental results demonstrate that the proposed methodology effectively enhances the success rate of robotic assembly tasks in a simulated environment.

### Key words

Simulation, Robot assemble, Insertion task, Deep reinforcement learning, Reward shaping

---

\* 한양대학교 로봇공학과, hong0807@hanyang.ac.kr, neal0907@hanyang.ac.kr, bborang3505@hanyang.ac.kr, kchan124@hanyang.ac.kr, taejoon@hanyang.ac.kr(공동 교신저자),

\*\* 한양대학교 인공지능융합학과, mimak98@hanyang.ac.kr, sinanju06@hanyang.ac.kr,

\*\*\* 한양대학교 지능형로봇학과, jongwanyoon@hanyang.ac.kr(공동 교신저자),

\*\*\*\* 한양대학교 스마트융합공학부, byungjinko@hanyang.ac.kr(공동 교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업(IITP-2024-RS-2020-II201741, 50%)과 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 인공지능융합혁신인재양성사업(IITP-2024-RS-2022-00155885, 50%) 연구임

## I. 서론

로봇 조립 작업은 높은 정밀도와 복잡한 상호작용을 요구하는 매우 도전적인 과제이다. 로봇이 부품을 정확하게 조립하기 위해서는 두 부품 간의 기구학과 접촉 동역학을 동시에 고려하는 모션 플래닝이 필요하다. 특히 산업 자동화의 발전과 함께 로봇 조립의 효율성을 높이고, 작업의 정확성을 극대화하는 기술적 요구가 점점 커지고 있다. 이와 같은 작업을 성공적으로 수행하기 위해서는 시뮬레이션을 기반으로 심층 강화학습을 통해 로봇 조립 정책을 학습하는 것이 중요하다[1].

시뮬레이션에서 로봇 조립 정책을 학습하기 위해, 접촉이 많은 조립 작업에서 강화학습을 사용하여 정책을 학습할 수 있도록 설계된 IndustReal이 제안되었다[2]. 해당 연구의 SDF(Signed Distance Field)-Based Dense Reward 함수는 플러그와 소켓 간의 정렬을 평가하기 위해 SDF 값을 사용하여 두 물체 사이의 거리를 계산하고, 이를 기반으로 보상을 제공한다. 그러나 SDF의 양수 영역 처리가 미흡하고, SDF 보상함수가 자연로그로 구현되어 있어 보상의 변화가 완만해지는 한계가 있다. 이러한 한계는 플러그와 소켓이 정밀하게 삽입되는 과정의 학습 성능을 저해하는 원인이 된다.

이러한 한계를 극복하기 위해, 본 논문에서는 세 가지 주요 개선 방안을 제시한다. 첫째, 학습 과정에서 음수 영역뿐만 아니라 양수 영역의 중요성도 반영하기 위해, SDF의 양수 영역에 페널티를 적용하는 방안을 제시한다. 둘째, SDF 거리에 따른 보상 값의 차이를 더욱 극대화하기 위해, SDF 보상함수를 자연로그에서 상용로그로 변환하는 방안을 제안한다. 마지막으로, 성공적인 삽입 작업 시 주어지는 보상을 극대화하기 위해, 보상 스케일링을 조정하는 방안을 제안한다.

이와 같은 세 가지 방안들을 통해, 시뮬레이션 환경의 로봇 조립 삽입 성공률이 선행 연구보다 최대 7% 향상되었다. 이러한 결과는 시뮬레이션 환경에서의 로봇 조립 실용성을 한층 더 높일 수 있을 것으로 기대한다.

## II. 방법론

본 연구에서는 로봇 조립 작업에서의 성능 향상을 위해 세 가지 주요 방법론을 제안한다. 각 방법론은 SDF 보상을 기반으로 경계 처리 개선, 보상 함수 수정, 보상 스케일링 조정에 초점을 맞추고 있다. 선행 연구[2]에서 제안하는 SDF의 보상 수식은 아래와 같다.

$$R = -\log\left(\frac{1}{N} \sum_1^N \frac{\Phi(x_i)}{N}\right) \quad (1)$$

첫 번째 방안은 SDF 경계 처리 개선이다. 이 방안은 SDF의 양수 영역에 특정 페널티를 부여하여 학습 과정에서 양수 영역의 중요성을 반영한다. 선행 연구에서는 음수 값만 절댓값으로 변환하고 양수는 0으로 처리하였다. 이를 개선하기 위해, 양수 값에 대한 positive penalty factor(양수 페널티 계수)를 도입하였다. 이 값은 1.0일 경우 양수 값을 음수처럼 동일하게 처리하고, 0.5로 설정하면 양수에 대해 절반의 중요도를 부여한다. 이러한 방식은 양수 영역의 보상을 강화하여 로봇의 조립 작업에서 양수 영역의 영향을 추가적으로 고려하게 한다.

SDF 보상의 로그 함수 수정은 보상의 민감도를 높이고 삽입 성공률을 극대화하는 방법이다. 기존에 사용된 자연로그 대신 상용로그를 사용함으로써, SDF 보상 값이 높아질 경우 더 큰 보상을 부여할 수 있도록 한다. 이 접근은 로봇이 더욱 정밀한 삽입 작업을 수행하도록 돕는다.



그림 1. NVIDIA Isaac Gym Peg and Hole 환경  
Fig. 1. NVIDIA Isaac Gym Peg and Hole Environment

SDF 보상 스케일링 조정은 engagement reward scale(접촉 보상 스케일)을 증가시켜 플러그가 소켓에 더 잘 맞을 때 더 많은 보상을 부여하는 방식이다. 이를 통해 성공적으로 삽입된 경우의 보상을 증가시켜 성공적인 삽입을 할 수 있도록 유도한다. 위와 같은 3가지 방법론을 본 논문의 보상함수 모델로 선정하였다.

### III. 실험

본 논문에서 제안하는 방법의 성능을 평가하기 위해 여러 실험을 진행하였다. 실험은 NVIDIA Isaac Gym[3] 시뮬레이션 환경에서 수행하였으며 시뮬레이션 환경에서 로봇은 Franka Emilka Panda를 사용하였다. 성공률을 선행 연구와 비교하여, 본 연구에서 제시한 세 가지 방법론의 보상함수들이 삽입 작업을 성공적으로 유도했는지를 분석하였다.

앞서 방법론에서 언급한 SDF의 양수 영역에서의 특정 페널티 부여(SDF boundary smoothing), SDF 보상 로그 함수 변경(SDF reward log function), SDF 보상 스케일링 조정(reward scale weight)들에 대한 학습 알고리즘으로 Proximal Policy Optimization (PPO)[5]를 사용하였다. 학습 횟수는 150 에포크로 설정하였다. 성능 검증은 학습 성능이 가장 높은 경우를 기반으로 15회 진행하여 성공률을 확인하였다. 학습 시에는 128대의 로봇을 하나의 프레임에 불러와 병렬로 학습을 진행하였고, 검증 시에는 1대의 로봇을 불러와서 검증하였다.

삽입 성공 지표로는 평균 삽입 성공률(Average Insertion success rate)과 보상(Reward) 두 가지를 사용한다. 평균 삽입 성공률은 전체 검증의 삽입 성공률에 대한 평균값을 의미하며, 보상은 삽입 성공에 대한 최종 보상값을 의미한다.

학습에서의 삽입 성공률은 마지막 학습 데이터 지표 기준으로 선행 연구에 대비하여 SDF boundary smoothing은 약 5%, SDF reward log function은 약 8%, Reward scale weight는 약 3% 증가하였다. 검증에서의 평균 삽입 성공률은 세 가지 방법론 모두 선행 연구에 비해 높은 지표를 보였다. 선행 연구의 삽입 성공률 대비 SDF boundary smoothing의 경우

약 2% 향상, SDF reward log function은 약 4% 향상, Reward scale weight는 약 7% 향상의 결과를 확인할 수 있다.

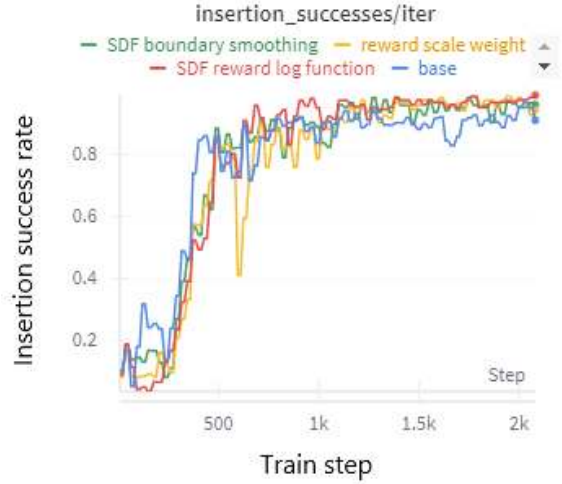


그림 2. 학습 결과 그래프  
Fig. 2. Training Result Graph

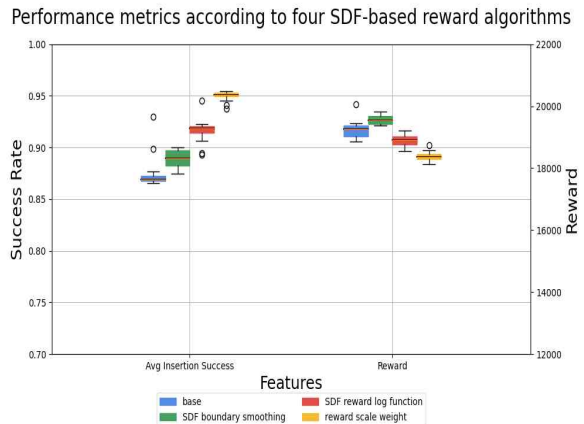


그림 3. 실험 결과 그래프  
Fig. 3. Experimental Result Graph

표 1. 선행 연구 및 SDF 보상 방법론 성능 결과  
Table 1. Performance Results of Prior Research and SDF Reward Methodologies

방법론	평균 삽입 성공률(%)
Base	87.49
SDF boundary smoothing	88.92
SDF reward log function	91.64
Reward scale weight	95.00



#### IV. 결 론

본 논문에서는 시뮬레이션 환경에서 심층 강화학습 기반 로봇 조립 작업의 성능을 향상하기 위한 SDF 보상 스케일링을 활용한 방법론을 제안하였다. 선행 연구에서 나타난 SDF의 양수 영역 처리 부족과 보상 구조의 한계를 극복하기 위해, 세 가지 주요 개선 방안을 도입하였다.

실험 결과, 제안된 방법론들은 모두 선행 연구와 비교하여 평균 삽입 성공률에서 최소 약 1.5%에서 최대 약 7.5%의 유의미한 향상을 보였다. 특히 보상 스케일링을 활용한 방법이 최대 향상으로 가장 높은 성능을 보였다.

향후 연구에서는 제안된 방법론을 다양한 시뮬레이션 환경에서의 로봇 플랫폼에 적용하여 그 범용성을 검증하고, 더욱 향상된 성능을 위한 추가적인 개선 방안을 모색할 것이다.

#### 참 고 문 헌

- [1] S. Levine et al., "Deep Reinforcement Learning for Robotics: A Survey of Real-World Successes," arXiv preprint, arXiv:2408.03539, 2024.
- [2] B. Tang et al., "IndustReal: Transferring contact rich assembly tasks from simulation to reality," in Robotics: Science and Systems, 2023.
- [3] V. Makoviychuk et al., "Isaac Gym: High Performance GPU-Based Physics Simulation for Robot Learning," arXiv preprint, arXiv:2108.10470, 2021.
- [4] V. Prasad et al., "Orbit: A Unified Simulation Framework for Interactive Robot Learning Environments," in IEEE Robotics and Automation Letters, vol. 8, no. 2, pp. 698-705, April 2023.
- [5] J. Schulman et al., "Proximal Policy Optimization Algorithms," arXiv preprint, arXiv:1707.06347, 2017.

# Auto EDA: SLM을 활용한 EDA 자동화 어플리케이션

최예림\*, 민유진\*\*, 오세종\*\*\*

## Auto EDA: EDA Automization application with SLM

Yerim Choi\*, Yujin Min\*\*, and Sejong Oh\*\*\*

### 요 약

데이터 분석의 중요성이 증가함에 따라 탐색적 자료 분석(EDA)은 데이터 이해와 해석을 위한 필수 도구로 자리 잡았다. 본 논문은 탐색적 자료 EDA를 자동화하는 Auto EDA 어플리케이션을 제안한다. 사용자가 데이터 셋을 입력하면 자동으로 EDA를 수행한 후 보고서와 해석을 제공한다. 특히, 기존 자동화 도구들의 한계점을 보완하고 SLM 모델을 활용하여 데이터 분석의 전반적인 과정을 지원한다. 이러한 기능을 바탕으로, 데이터에 대한 깊은 이해와 효율적인 데이터 분석이 가능할 것으로 기대된다.

### Abstract

As the importance of data analysis is growing, Exploratory Data Analysis (EDA) has become essential for understanding and interpreting data. This paper introduces an Auto EDA application that automates the EDA process. This application automatically performs EDA, generates a report, and provides an interpretation of the uploaded dataset. It addresses the limitations of existing automation tools by leveraging an SLM model to support the entire data analysis process. Based on these functions, it is expected to enable deeper insights and more efficient data analysis.

### Key words

eda; microsoft phi-3; prompt engineering; report generation;

## I. 서 론

1977년 John Tukey에 의해 도입된 탐색적 자료 분석(EDA, Exploratory Data Analysis)은 “숫자적 탐정 작업, 계산적 탐정 작업, 또는 그래픽적 탐정 작업”라고 정의되며, 데이터를 탐구하고 분석하는 중

요한 방법론으로 자리 잡았다[1]. 최근 4차 산업 혁명과 함께 데이터의 양과 복잡성이 급증함에 따라 데이터 분석의 중요성이 비즈니스와 연구 환경에서 강조되었으며, EDA는 데이터를 이해하고 해석할 때 사용하는 필수 과정이다.

이러한 흐름 속에서 본 연구에서는 Small

\* 단국대학교 sw융합대학 소프트웨어학과 학사과정, cyelim0621@gmail.com,

\*\* 단국대학교 sw융합대학 정보통계학과 학사과정, minyujin951225@gmail.com,

\*\*\* 단국대학교 sw융합대학 소프트웨어학과, sejongoh@dankook.ac.kr (교신저자)

※ 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2023-00222191, 분산된 데이터에 대한 논리적 데이터 통합과 복합분석을 지원하는 데이터 패브릭 기술개발)

Language Model(SLM)을 활용하여 사용자가 데이터를 쉽게 탐색하고 분석할 수 있는 Auto EDA를 제안한다. 사용자가 데이터셋을 입력하면 자동으로 EDA를 수행하고, EDA 보고서와 함께 언어 모델을 활용한 분석 결과 해석 기능을 제공한다. 이는 기존 어플리케이션과의 차별점으로, SLM을 활용해 직관적이고 명확한 데이터 이해를 돕는다. 이를 통해 사용자는 데이터를 더욱 효과적으로 활용할 수 있게 된다.

## II. 기본 개념

### 2.1 EDA

탐색적 자료 분석(EDA)는 데이터를 탐색하여 이해를 높이고 접근성을 개선하는 통계적인 방법론이다. 데이터 시각화 도구는 EDA의 핵심 요소로 데이터 내에 숨겨진 정보를 쉽게 파악할 수 있다. EDA는 데이터 자체에 초점을 맞추어 가설을 세우기 전 통계적 추론의 기초를 마련하고 새로운 가설을 제시하는 기틀이 된다[2].

### 2.2 Phi3

SLM(Small Language Model)은 대규모 언어 모델(LLM)의 핵심 기능을 유지하면서 모델 크기와 계산 비용을 최소화한 생성형 AI 모델이다. 매개변수가 1천억 개 이상인 경우 LLM으로 분류되며, 수십억 개의 매개변수로 연산 작업을 단축시킨 모델은 SLM으로 분류된다[3].

phi-3-mini-128k-instruct는 128K 길이의 문맥을 처리할 수 있는 phi-3-mini 모델의 변형이다. phi-3-mini 모델은 3.8억 개의 파라미터로 이루어진 SLM으로, 작은 규모에 더 불구하고 GPT-3.5와 유사한 성능을 보여주고 있다[4].

### 2.3 Prompt Engineering

프롬프트 엔지니어링은 LLM의 출력을 과제에 맞게 안내하는 프롬프트를 설계·개선·구현하는 것을 의미한다[5]. 여기서 프롬프트는 LLM을 프로그

래밍하기 위해 사용자가 제공하는 명령어 집합이다 [6]. 프롬프트가 작업을 지정하는 만큼, 적절한 프롬프트 템플릿을 선택하는 것은 정확도뿐만 아니라 모델이 처음에 수행할 작업에도 큰 영향을 미친다 [5].

## III. Auto EDA 어플리케이션

### 3.1 시스템 구조도

본 연구는 기존 EDA 자동화 도구들의 문제를 보완하고, SLM 모델을 통해 사용자가 데이터 분석에 익숙하지 않더라도 쉽게 해석할 수 있도록 한다. 사용자는 언어와 SLM 모델 분석 여부를 선택 후, Data 페이지에서 데이터를 업로드하고 편집한다. 분석 결과는 Overview, Visualization, Correlation 페이지에서 확인할 수 있으며, Modeling 페이지에서 학습 결과를 비교할 수 있다.

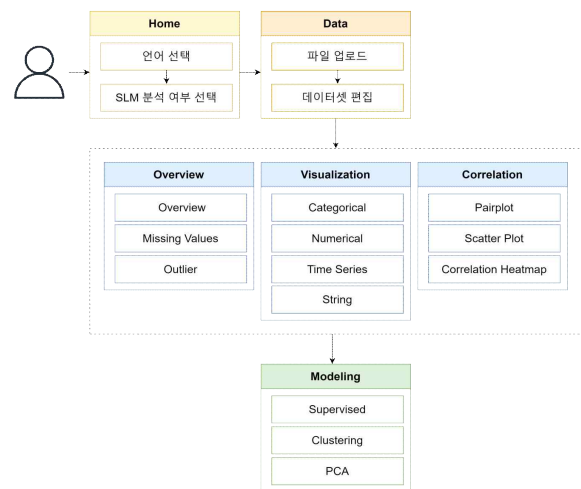


그림 1. 시스템 구조  
Fig. 1. System Structure

### 3.2 EDA 보고서

Auto EDA는 사용자가 업로드한 데이터의 분석 결과를 보고서로 제공한다. 이 보고서는 Overview, Visualization, Correlation 페이지로 구성된다. 보고서는 그림 2와 같이 출력된다. Overview 페이지는 데이터에 대한 전반적인 분석을 제공한다. 열 수,

행 수, 결측치, 중복된 행 등을 포함한 데이터의 개요와 결측치, 이상치에 대한 표와 시각화를 제공한다. Visualization은 페이지는 사용자가 편집한 데이터 타입을 기반으로 각 컬럼의 시각화 결과를 제공한다. 사용자는 Data 페이지에서 각 컬럼의 타입을 Categorical, Boolean, Discrete, Continuous, Time Series, String 중 원하는 것을 선택할 수 있다. Correlation은 pairplot, 산점도 그래프, 상관관계 히트맵을 제공한다. Modeling 페이지는 대표적인 지도 학습 및 비지도학습을 수행할 수 있는 인터페이스를 제공한다.

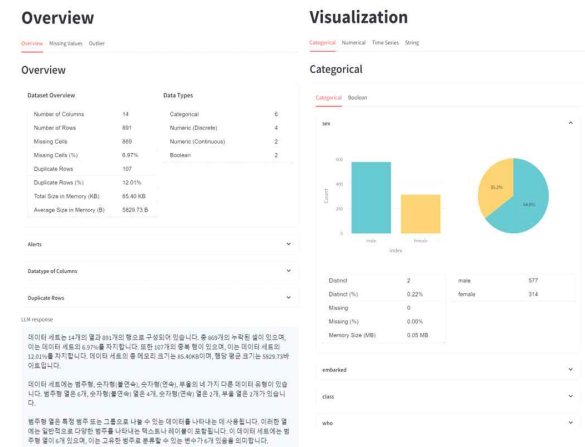


그림 2. Auto EDA 어플리케이션 구동 화면  
Fig. 2. Screen of running Auto EDA application

### 3.3 SLM 분석

이번 연구에서 SLM 분석을 위해 microsoft의 Phi3 모델을 사용하였다. SLM 분석은 Overview와 Correlation 페이지에 분석 보고서 내용을 요약하고 추가적인 해석을 제공한다. 해석은 영어와 한국어로 제공된다.

Overview 페이지는 데이터의 전반적인 내용을 해석한다. Overview 표는 데이터의 기본 특성을 요약한다. 결측치는 문제가 될 수 있는 결과 그 이유, 그리고 결측치 처리 방법을 제안한다. 마지막으로 이상치의 영향과 적절한 전처리 방법을 추천한다.

Correlation 페이지는 피쳐들의 상관관계를 해석한다. pairplot은 관련성이 높은 변수와 그 이유를 제시한다. 산점도 그래프는 두 변수의 상관관계의

방향성과 그 강도를 설명한다. 상관관계 히트맵은 정규성 검정의 결과를 바탕으로 spearman, pearson, kendall 모델 중 적합한 방법을 선택하고 변수 간의 관계를 해석한다.

## IV. 한계 및 향후 연구

Auto EDA 어플리케이션은 다양한 분석 결과를 제공하지만, 한계가 존재한다. SLM 모델을 사용할 때 토큰 제한 문제가 발생할 수 있다. SLM 모델은 입력력이 가능한 토큰 수가 적기 때문에 실험에 사용한 데이터보다 규모가 더 큰 데이터는 문제가 발생할 수 있다. 또한 SLM 모델은 이미지를 입력으로 사용하지 못하기 때문에 Visualization 페이지의 해석을 제공하지 못한다.

향후 연구는 Visualization 페이지를 위한 프롭트 엔지니어링을 진행할 예정이다. 또한 LLM을 사용한 챗봇 기능을 도입하여 어플리케이션과 사용자의 상호 작용을 더욱 개선할 계획이다. 이를 통해 사용자가 실시간으로 질문을 하거나 모델링 결과에 대한 추가 정보를 얻어 데이터 분석의 효율성을 더 높이고자 한다.

## V. 결론

본 논문은 SLM 모델인 Phi3를 활용한 Auto EDA 어플리케이션을 제안한다. Auto EDA는 보고서를 자동으로 생성 및 해석 기능을 제공한다. 기존 EDA 도구들의 문제를 개선하여 데이터 타입 설정, 시계열 분석, 머신러닝 모델링을 추가함으로써 데이터 분석을 전반적으로 지원한다. 또한 SLM 모델을 사용해 데이터를 해석하여 사용자에게 도움을 준다.

그러나 데이터 분석 지식이 없는 사용자는 SLM 모델의 오류를 판단하고 수정하는 데 한계가 있을 수 있다. 따라서 기본적인 데이터 분석 지식은 여전히 중요하다. 향후 연구에서는 더 정확한 결과를 위해 LLM 도입 및 실시간 질문-응답이 가능한 챗봇 기능 추가를 통해 데이터 분석에 대해 더 깊은 이해를 도모할 수 있을 것이다.

## 참 고 문 헌

- [1] Tukey, John Wilder. "Exploratory data analysis." Reading/Addison-Wesley (1977).
- [2] Reid, Machel, et al. "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context." arXiv preprint arXiv:2403.05530 (2024).
- [3] Meskó, Bertalan. "Prompt engineering as an important emerging skill for medical professionals: tutorial." Journal of medical Internet research 25 (2023): e50638.
- [4] White, Jules, et al. "A prompt pattern catalog to enhance prompt engineering with chatgpt." arXiv preprint arXiv:2302.11382 (2023).
- [5] Liu, Pengfei, et al. "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing." ACM Computing Surveys 55.9 (2023): 1-35.

# 머신러닝 모델 성능 향상을 위한 피처 합성 자동화 도구

최서영\*, 오세종\*\*

## Automated Feature Synthesis Tool for Improving Machine Learning Model Performance

Seoyoung Choi\* and Sejong Oh\*\*

### 요약

본 연구는 기존 피처 합성 방법의 한계를 극복하고, 머신러닝 과정의 부담을 줄이기 위해 피처 선택 과정까지 자동화하여 모델의 성능 최적화를 도모하였다. 특히, 사용자가 모델을 선택할 수 있는 유연성을 제공하며, 다양한 데이터 변환 기법을 도입하여 새로운 피처를 생성함으로써 머신러닝 모델의 성능을 유의미하게 향상시켰다.

### Abstract

This study attempted to optimize model performance by overcoming the limitations of existing feature synthesis methods and automating the feature selection process to reduce the burden on the machine learning process. In particular, it provides users with the flexibility to select models and significantly improves the performance of machine learning models by introducing various data conversion techniques to create new features

### Key words

AutoML, Machine Learning, Feature Engineering, Feature Synthesis, Feature Selection

## 1. 서론

최근 다양한 형태의 데이터의 증가와 데이터 양의 급격한 확대에 따라 머신러닝 및 빅데이터 애플리케이션에서 데이터 처리와 피처 엔지니어링의 중요성이 강조되고 있다. 이러한 배경 속에서 데이터 수집 및 처리 기능의 자동화 접근법은 최근에 상당

한 주목을 받고 있다[1]. 자동화된 머신러닝 접근법은 데이터 전문가에 대한 의존도를 줄이고 도메인 전문가가 머신러닝 모델을 구축할 수 있도록 지원하는 데 목적을 두고 있다[2]. 그러나, 자동화된 데이터 처리 방법의 중요성에도 불구하고, 이와 관련된 체계적인 연구는 여전히 미흡한 실정이다.

현재 몇몇 피처 합성 패키지는 이러한 문제를 부

\* 단국대학교 sw융합대학 소프트웨어학과, 2001609@naver.com,

\*\* 단국대학교 sw융합대학 소프트웨어학과, sejongoh@dankook.ac.kr (교신저자)

※ 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2023-00222191, 분산된 데이터에 대한 논리적 데이터 통합과 복합분석을 지원하는 데이터 패브릭 기술개발)

분적으로 해결하고 있다. 예를 들어, Featuretools는 ‘Deep Feature Synthesis(DFS)’ 기법을 통해 숫자 및 명목 데이터에서 피처를 자동 생성하는 기능을 제공한다[3]. Featuretools는 여러 데이터 소스의 관계성을 분석하여 새로운 피처를 생성함으로써 데이터 전처리를 효율화하고 있다. 또 다른 예로, AutoFeat 패키지는 AutoFeatRegressor 및 AutoFeatClassifier 모델을 제공하며, 원래 데이터를 기반으로 추가 비선형 입력 피처를 자동으로 생성하고 선택한 다음 이러한 피처로 선형 예측 모델을 학습한다[4].

그러나, 현재 존재하는 피처 합성 연구의 대부분은 사용자가 모델을 선택할 수 없거나, 피처 합성 단계까지만 자동화되고, 이후 중요한 피처 선택 및 최적화 과정은 사용자에게 맡기는 경향이 있다. 이로 인해 정해진 모델에 의존하게 되어 유연성이 감소하며, 사용자는 수작업으로 피처를 선택하고 성능을 최적화하는 데 많은 시간을 소모하게 된다.

이러한 문제를 해결하기 위해 본 연구는 자동화된 피처 합성 및 피처 선택을 통해 머신러닝 모델의 성능을 최적화하고, 이를 실질적인 데이터 분석 및 예측 문제에 적용하는 데 기여하고자 하였다. 이를 통해 데이터 과학자는 피처 엔지니어링에 대한 부담을 줄이고, 더 나은 데이터 품질과 모델 성능을 달성할 수 있을 것으로 기대된다.

## II. 피처 합성

본 연구에서는 머신러닝 모델의 성능 향상을 위한 피처 합성 접근법을 체계적으로 다루었다. 먼저 의미 없는 피처들을 포함하여 불필요한 계산을 증가시키기보다는, 중요한 피처 위주로 합성을 진행하여 모델 성능을 높이고 계산 효율성을 극대화하였다. 이후, 선택된 중요 피처에 대해 로그, 제곱근, 역수 등과 같은 12가지의 수학적 변환을 통해 새로운 피처를 생성하였다. 이렇게 생성된 피처들을 다시 산술 연산을 통해 확장되었으며, 이 과정에서는 두 개의 피처를 선택하여 합, 차, 곱, 나누기 등의 산술 연산이 적용되었다. 이러한 변환을 통해 모델이 더욱 풍부한 정보를 학습할 수 있도록 지원하였다.

피처 합성 과정에서는 산술 연산 외에도 다양한 비교 연산이 추가적으로 적용되었다. 이러한 비교 연산은 데이터 간의 관계를 더욱 심층적으로 반영할 수 있게 하며, 모델이 보다 풍부한 특징을 학습할 수 있도록 돕는다. 또한 누적 연산도 사용되어 각 피처의 누적 합, 누적 평균, 차분 등을 계산함으로써 시계열적 특성을 반영한 새로운 피처들을 생성하였다. 이러한 다중 연산 과정을 통해 데이터에서 중요한 패턴을 보다 효과적으로 학습할 수 있는 환경이 마련되었다.

본 연구에서 적용한 피처 합성 과정은 크게 세 단계로 요약될 수 있다. 첫 번째 단계에서는 원본 데이터에 다양한 변환(로그, 제곱근, 역수 등)을 적용하여 새로운 피처를 생성하고, 두 번째 단계에서는 생성된 피처들 간의 산술 및 비교 연산을 통해 데이터를 확장하였다. 마지막 단계에서는 각 피처에 대해 누적 연산을 적용하여 피처의 시계열적 특성을 반영한 새로운 피처들을 추가하였다.

## III. 피처 선택

본 연구에서는 피처 합성 이후, 피처 선택 절차를 통해 모델 성능을 더욱 향상시키는 과정을 거쳤다. 새로 생성된 피처들은 데이터의 복잡성을 증가시키고, 모델이 더 많은 정보를 학습할 수 있도록 돕는다. 그러나 피처가 많아질수록, 그 중 일부는 모델에 도움이 되지 않거나, 오히려 성능을 저해할 가능성이 있기 때문에 적절한 피처 선택이 필수적이다.

본 연구에서는 피처 선택 과정을 통해 모델 성능을 향상시키기 위한 데이터 전처리 단계를 체계적으로 진행하였다. 먼저, Null 비율이 높은 피처나 값의 분포가 매우 적은 피처를 제거함으로써 낮은 정보를 제공하는 피처들을 필터링하였다. 이 과정은 불필요한 데이터를 사전에 제거하여 계산 복잡성을 줄이고, 학습 속도를 향상시키는 역할을 하였다. 다음으로, 다중공선성 문제를 해결하기 위해 VIF(Variance Inflation Factor) 수치가 10 이상인 피처를 제거하였다. 이를 통해 피처 간의 상관관계를 낮추어 모델의 안정성과 성능을 보장할 수 있었다.

이후, 보다 정교한 Feature Selection의 Filter method를 적용하여 다수의 피처를 필터링하는 과정을 거쳤다. 이렇게 필터링된 피처는 보통 1000개 이상의 피처가 남게 되며, 이 상태에서 Permutation Importance 방식을 적용하여 원본 피처에서 중요 피처를 추출할 때 사용한 것과 같은 개수만큼의 피처를 다시 선택하였다.

다음으로는, 원본 피처와 새로 합성된 피처를 모두 결합하면 약 두 배 정도의 피처가 생성된다. 이 단계에서는 Feature Selection의 Wrapper 방식을 적용하여 최종적으로 선택할 피처들을 다시 한 번 선별하였다. 피처의 최종 개수는 사용자가 조정할 수 있도록 하였으나, 기본값으로는 처음 선택한 중요 피처의 개수와 동일하게 설정하였다. 이렇게 추출된 피처들을 바탕으로 최종적으로 성능을 비교하여 최적의 피처 조합을 결정하는 방식으로 연구를 진행하였다.

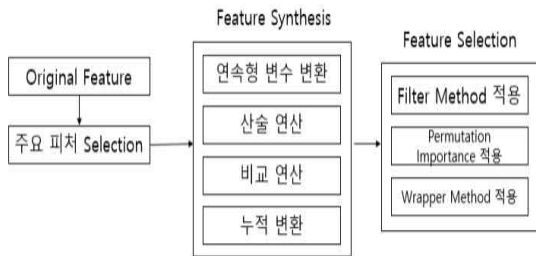


그림 1. 시스템 구조  
Fig.. 1. System Structure

#### IV. 실험 결과

본 연구에서는 다양한 데이터셋을 활용하여 피처 합성이 머신러닝 모델의 성능 향상에 미치는 영향을 분석하였다. 사용된 데이터셋은 모두 Kaggle에서 수집하였으며, 각 실험 결과에 대한 세부 사항은 다음과 같다. Heart Attack Analysis & Prediction Dataset에서는 총 13개의 피처 중 10개가 최종 선택되었으며, 원본 피처 5개와 합성 피처 5개로 구성되었다. 모델의 정확도는 원본 피처 사용 시 0.8150, 합성 피처 포함 시 0.8350으로, 약 0.02의 성능 향상이 나타났다. Wine Quality Dataset에서는 11개의 피처 중 10개를 선택하였고, 원본 피처 4개와 합성 피

처 6개가 포함되었다. 이 경우, 정확도는 0.6064에서 0.6122로 상승하였다. Water Quality Dataset 데이터셋에서는 20개의 피처 중 10개를 선택하였으며, 원본 피처 4개와 합성 피처 6개가 포함되었다. 원본 피처의 정확도는 0.8929에서 합성 피처 포함 시 0.8979로 증가하였다. Body Fat Prediction Dataset에서는 총 14개의 피처 중 10개가 선택되었으며, MSE는 원본 피처 2.0031에서 합성 피처 포함 시 1.6536으로 감소하였다. 또한, R<sup>2</sup> 값이 0.9673에서 0.9732로 증가하는 성능 개선으로 이어졌다.

표 1. 성능 향상 비교  
Table 1. Performance Improvement Comparison

	Heart Dataset	Wine Dataset	Water Dataset	BodyFat Dataset
원본 피처 개수	13개	11개	20개	14개
피처 조합	원본 5/ 합성 5	원본 4/ 합성 6	원본 4/ 합성 6	원본 1/ 합성 5
성능 개선	2.45% 향상	1.25% 향상	0.56% 향상	17.4% 향상

#### V. 결론

본 연구에서는 피처 합성을 통한 머신러닝 모델의 성능 향상 가능성을 체계적으로 검토하였으며, 다양한 데이터셋을 활용한 실험 결과 유의미한 성능 향상을 확인할 수 있었다.

기존 연구와 차별화되는 점은 사용자가 직접 모델을 선택할 수 있도록 하였으며, 여러 가지 피처 합성 과정을 한 번에 자동화하여 각 함수를 개별적으로 실행해야 하는 번거로움을 줄인 것이다. 또한, 피처 선택 과정까지 자동화하여 부담을 경감하고, 더 나은 결과를 도출할 수 있도록 하였다.

본 연구에서는 연속형 변수의 합성만을 다루고 있어, 시계열 데이터와 같은 다른 유형의 데이터에 적용하기 위한 추가적인 연구가 필요하다. 또한, 피처 선택 과정을 개선하여 더 우수한 피처를 선택할 수 있는 방법에 대한 연구도 필요하다.

#### 참고 문헌



- [1] Alhassan Mumuni, Fuseini Mumuni, "Automated data processing and feature engineering for deep learning and big data applications: A survey", *Journal of Information and Intelligence*, <https://doi.org/10.1016/j.jiixd.2024.01.002>.
- [2] Zöllner, Marc-André and Marco F. Huber. "Benchmark and Survey of Automated Machine Learning Frameworks." *J. Artif. Intell. Res.* 70
- [3] Bosch, Nigel. 2021. "AutoML Feature Engineering for Student Modeling Yields High Accuracy, But Limited Interpretability". *Journal of Educational Data Mining* 13 (2):55-79. <https://doi.org/10.5281/zenodo.5275314>.
- [4] Horn, F., Robert T. Pack and Michael Rieger. "The autofeat Python Library for Automated Feature Engineering and Selection." *PKDD/ECML Workshops* (2019).

# 위험 소리 식별 기술개발을 위한 효과적인 CNN 기반 학습모델의 실험적 비교

김도균\*, 이준영\*\*, 하옥균\*\*\*

## An Experimental Comparison of Effective CNN-based Learning Models for Developing Hazard Sound Identification Technology

DoGyun Kim\*, JunYeong Lee\*\*, and Ok-Kyoon Ha\*\*\*

### 요 약

본 논문에서는 딥러닝 기술을 이용하여 실내외에서 발생하는 위험소리를 식별하는 학습모델의 개발을 위해 10종의 소리 데이터를 2D-CNN 기법으로 학습하여 개발된 딥러닝 학습모델을 대상으로 소리 데이터의 두 가지 정규화 방법에 따른 정확성의 비교와 경량화를 위한 각 양자화에 따른 효과성을 실험적으로 비교한다. 실험 결과 MFCC 정규화 방법을 적용한 모델이 Mel-Spectrogram 방법에 비해 개선된 정확성을 보이고, 3가지 양자화 방법들 중 Float16 양자화 방법이 원본 모델에 비해 약 17%의 용량만으로 유사한 정확성을 보인다.

### Abstract

This paper addresses the development of a deep learning model for identifying hazardous sounds occurring in both indoor and outdoor environments using deep learning techniques. In this study, a deep learning model is trained on a dataset comprising 10 types of sound data using a 2D-CNN technique. The study experimentally compares the accuracy of the model based on two different sound normalization methods for data preprocessing and evaluates the effectiveness of various quantization techniques for model compression. The experimental results demonstrate that the model utilizing the MFCC normalization method achieves improved accuracy compared to the Mel-Spectrogram method. Among the three quantization methods tested, the Float16 quantization reduces the model size to approximately 17% of the original while maintaining similar accuracy.

### Key words

shape recovery, shape from focus, artificial neural network, depth estimation, focus measure

## 1. 서 론

현대사회에서 다양한 요인으로 인해 청각장애인과 난청인이 증가하고 있고 이들은 일상 및 직업

생활에서 비장애인 및 비난청인에 비해 수용할 수 있는 정보의 제한으로 인해 안전사고 등의 위험 정보 획득에 어려움을 겪고 있다[1]. 이러한 문제를 해결하기 위해 주변의 위험소리를 딥러닝 기술로

\*경운대학교, 201905060@ikw.ac.kr, \*\*경운대학교, dlwnsdud1206@naver.com, \*\*\*경운대학교 okha@ikw.ac.kr(교신저자)

식별하여 사용자에게 시각화 또는 진동으로 알려주는 기술이 제시되고 있다.

스마트 폰, 스마트 워치, 스마트 Hub 등의 임베디드 장치에 탑재되는 기존의 방식들을 위해서 본 연구에서는 딥러닝 학습에 필요한 소리 데이터의 전처리를 위한 2가지 정규화 방법과 더불어 정확성의 손실을 최소화하는 경량화 방법을 실험적으로 비교하였다. 실험 결과 MFCC 정규화 방법을 적용한 학습모델이 Mel-Spectrogram 방법을 적용한 모델에 비해 평균 약 2.5% 개선된 정확도를 보인다. 또한, 3가지 양자화 방법을 적용하여 학습모델을 경량화한 결과의 비교에서는 Float16 양자화를 적용한 경우 원본 모델에 비해 약 17%의 파일 크기만으로 약 7%만 감소된 정확도를 보인다. 이러한 결과는 스마트 폰과 같은 저사양 장치에 본 학습모델을 적용하여 청각장애인 및 난청인을 위한 서비스 개발에 유용하게 활용할 수 있을 것으로 기대한다.

## II. 연구 방법

CNN과 같은 딥러닝을 통해 소리 신호의 특징을 추출하는 기법에는 소리 신호를 일정 구간으로 나누어 구간별 스펙트럼을 분석하여 특징점을 추출하는 Mel-Frequency Cepstral Coefficients (MFCC) 방법이 대표적이다 [2, 3]. MFCC는 Mel Spectrum에서 Cepstral 분석을 통해 추출된 값을 의미하며, 소리 신호를 프레임별로 고속 푸리에 변환(FFT)를 적용해 생성된 배음 구조의 차이를 포함하는 Spectrum에 물리적인 주파수와 실제 사람이 인식하는 주파수의 관계에 기반한 Filter Bank를 적용하여 도출한 Mel Spectrum을 대상으로 Cepstral 분석을 적용해 추출된다. 이때 Cepstral 분석은 소리 주파수의 피크 값들의 연결 곡선과 Spectrum을 분리하는 것으로 Discrete Cosine Transform (DCT)를 취한 값에서 계수를 구한 결과이다. Mel-Spectrogram은 음성 신호를 시각적으로 변환하는 방법으로, 음성 인식이나 음향 분석과 같은 기계 학습 작업에 유용하게 사용된다.

이 방법은 음성 신호를 단시간 푸리에 변환(STFT)을 적용해 짧은 프레임으로 나누어 시간에

따른 주파수 정보를 인간의 청각을 반영한 멜(Mel) 스케일로 변환된다. 이를 통해 결과적으로, x축은 시간을, y축은 주파수를, 색 강도는 진폭을 나타내는 스펙트로그램(Spectrogram)이 생성된다. 멜 스케일을 사용함으로써 스펙트로그램은 인간의 음성이거나 음향 분석 작업에 더 적합해 지기 때문에 이를 통한 최종 출력은 소리를 2D 이미지로 표현한 형태로 변형되어 CNN과 같은 모델에 입력 데이터로 쉽게 활용할 수 있다.

본 연구에서는 딥러닝 기반 CNN 모델을 활용하여 위험 소리를 식별하기 위한 소리 데이터를 전처리한 후, 학습 모델을 개발하고 경량화를 통해 효율성을 개선하는 과정을 실험적으로 검토한다. 실험은 두 가지 정규화 방법(MFCC, Mel-Spectrogram)과 세 가지 양자화 방법을 적용한 모델을 통해 진행하며, 이를 바탕으로 모델의 정확성과 파일 크기, 추론 시간을 비교한다. 그리고, 실험에 사용할 소리 데이터는 표 1과 같이 10종(경적, 유리깨짐, 비명, 경보, 도난, 화재, 응급, 가스누출, 폭발, 충격)의 소리를 AI 허브 및 국내 빅데이터 사이트 등의 인터넷 수집을 통해 총 3,512개를 확보하여 실험에 사용한다.

표 1. 실험 데이터 세트 단위(개)  
Table 1. Experimental data sets

경적	유리 깨짐	비명	경보	도난	화재	응급	가스 누출	폭발	충격
400	400	358	400	277	300	300	320	400	357

## III. 비교 실험

실험을 위해 학습할 소리 신호 데이터를 모노타입으로 통일하고 다른 Sampling Rate를 Re-Sampling하여 22,050Hz로 통일하여 Bit Depth를 정규화하는 과정을 거쳤다. 이러한 MFCC와 Mel-Spectrogram으로 추출된 각 데이터를 구분하고 Conv2D 4-layer, Max\_pooling2D 3-layer, Global\_average\_pooling2D 1-layer, Dense 3-layer로 구성된 단일 CNN 모델을 적용하여 학습에 사용하였다.

단일 CNN 모델을 적용하여 두 분류의 데이터 집합을 비교한 실험 결과 MFCC 정규화 기반의 데이터를 적용한 경우 학습 초기에 정확도가 빠르게

상승하였으며, 최종적으로 훈련 정확도 99.57%, 검증 정확도 97.65%를 기록하였다. 반면, Mel-Spectrogram 정규화 모델은 상대적으로 느린 학습 곡선을 보였으며, 최종적으로 훈련 정확도 98.36%, 검증 정확도 96.30%를 기록하였다. 최종적으로, MFCC 정규화 데이터를 사용한 CNN 모델이 평균적으로 약 2.5% 더 높은 정확도를 보였으며, 이는 MFCC가 소리 특징을 더 효과적으로 추출함을 시사한다.

한편, 스마트 폰, 스마트 워치, 스마트 Hub 등의 임베디드 장치에 탑재되는 기존의 방식들을 위해서 학습 모델의 경량화의 필요성이 증가하는데, 이를 위해 MFCC 기반 데이터를 학습한 모델을 대상으로 양자화(Quantization) 기반의 경량화를 적용하여 파일 크기, 응답시간, F1-Score 기반의 정확도를 실험하여 비교하였다.

표 2. 양자화에 따른 실험 결과

Table 2. Experimental results according to quantization

Model name	File size (MB)	Inference Time (sec)	F1-Score
original (HDF5)	285.3	1.15	0.98
Full int quantized	24.6	0.63	0.003
Float16 quantized	47.4	0.58	0.91
Dynamic Range quantized	24.4	0.63	0.84

실험 결과는 표 2에서 보인것과 같이 원본 모델 (HDF5)은 285.3MB 크기를 차지하며, 추론 시간은 1.15초로 나타났다. 양자화를 적용한 학습모델은 평균 32.1MB 크기로 줄어들었고, 응답시간도 평균 0.61초로 단축된 반면, 정확도는 평균 58.4%로 크게 감소된 결과를 보였다. 이들 양자화를 통한 경량화 방법 중 Full int quantized 방법을 적용한 모델은 파일 크기가 24.6MB로 줄었으며, 추론 시간도 0.630초로 감소하였으나 정확도가 0.003을 기록하여 정확도의 손실이 매우 심각한 결과를 보였다. Float16 quantized 방법을 적용한 경량화 모델은 파일 크기 47.4MB, 추론 시간 0.585초로 가장 빨랐고, F1-Score는 0.91로 원본 대비 약 7%의 정확도 손실을 보였다. 마지막으로 Dynamic Range quantized 모델은 Full int 양자화와 유사한 성능을 보였으며, F1-Score

가 0.84로 원본 대비 14%의 정확도 손실을 보였다.

결론적으로 모든 양자화 방식에서 원본 모델 대비 파일 크기와 추론 시간이 크게 감소했음을 확인할 수 있었으며, Full int quantized 방법을 제외한 두 가지 방법에서 낮은 정확도 손실로 경량화할 수 있는 것으로 확인되었다. 특히 Float16 양자화 모델의 경우 약 7%의 정확도 감소만으로 파일 크기를 약 83%로 줄이면서도 추론 시간을 50% 이상 단축시켜 스마트폰 등의 임베디드 시스템을 위한 경량화 방법으로 가장 우수한 결과를 보였다.

### III. 결 론

본 연구에서는 위험 소리를 식별하기 위한 딥러닝 모델을 개발하고, 두 가지 정규화 방법(MFCC, Mel-Spectrogram)과 세 가지 양자화 방법의 성능을 실험적으로 비교하였다. 실험 결과, MFCC 정규화 방법을 적용한 모델이 Mel-Spectrogram 방식에 비해 더 높은 정확도를 보였으며, 경량화 실험에서는 Float16 양자화가 파일 크기를 약 83% 줄이면서도 원본 모델과 유사한 정확성을 유지하였다. 이러한 결과는 소리 식별 모델이 스마트폰 및 스마트워치와 같은 저사양 임베디드 장치에서도 효율적으로 적용될 수 있음을 시사한다. 특히, 청각장애인 및 난청인을 위한 위험 소리 식별 서비스의 실시간 제공에 큰 기여를 할 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] 이현아, 이준우, "청각장애인의 위험 판단을 위한 소리 정보, 인지 조건과 행동 반응에 관한 연구", 한국장애인복지학, 38(38), 149-180, 2017.
- [2] Abdelmajid H. Mansour, Gafar Zen Alabdeen Salh, Khalid A. Mohammed, "Voice Recognition using Dynamic Time Warping and Mel-Frequency Cepstral Coefficients Algorithms," *Inter. Journal of Computer Applications*, 116(2), 34-41, April, 2015.
- [3] 윤원정, "MFCC를 활용한 딥러닝 기반 음성 인식 모델에 관한 연구." 국내석사학위논문 한세대학교 대학원, 2022.

# VR환경에서 특수교육 전문가 AI를 활용한 예비교사의 인식 조사

김가은\*, 김혜정\*\*,

## Investigating Pre-Service Teachers' Perceptions of Expert AI in Special Education within a Virtual Reality Environment

Ga Eun Kim\*, and Hye Jeong Kim\*\*

### 요 약

본 연구는 예비 특수교사인 대학생의 장애아동의 문제행동에 대한 대처 자신감을 향상시키기 위한 VR환경에서의 ChatGPT기반 전문가AI의 활용 후 인식을 조사하였다. 본 연구에서는 예비 특수교사(n=17)의 장애학생의 문제행동 대처 방안을 “(1) 검색엔진”과 “(2) 텍스트 기반 ChatGPT” 및 “(3) VR환경에서의 음성기반 ChatGPT”로 구현한 전문가AI를 통해 학습하고 인식을 조사하였다. 분석 결과, 검색엔진으로 학습한 그룹보다 텍스트와 VR환경에서의 ChatGPT 전문가AI를 활용한 학생들이 장애아동이 보이는 문제행동의 대처에 보다 높은 자신감을 나타내었다.

### Abstract

This study investigated the perceptions of college students who are pre-service teachers after using ChatGPT-based expert AI in VR environment to improve their confidence in dealing with problem behaviors of children with special needs. In this study, pre-service teachers (n=17) learned how to deal with problem behaviors of students with special needs with “(1) search engine,” “(2) text-based ChatGPT,” and “(3) voice-based ChatGPT in VR environment” and investigated their perceptions. As a result of the analysis, students who used ChatGPT in text and VR environment were more confident in dealing with problematic behaviors of children with special needs than those who used search engine.

### Key words

special education, problem behaviors, ChatGPT, Unity

---

\* 국립창원대학교 특수교육과, 20230185@gs.cwnu.ac.kr,

\*\* 국립창원대학교 특수교육과 교수, hjkimcw@changwon.ac.kr (교신저자)

※ 이 연구는 「과학기술정보통신부 방송통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.

## 1. 서 론

특수교육 전공 과정에서 실시하는 교육실습과 교육 봉사활동을 거치면서 특수교육과 학생들은 장애아동의 장애의 유형과 정도별로 다양하고 복합적인 문제행동으로 인해 어려움을 경험한다. 현장의 특수교사들도 장애아동이 일으키는 문제행동에 대한 중재 시 여러 난관을 만나게 된다[1]. 장애아동의 문제행동은 때리기, 물기와 같은 신체 공격부터 반항, 규칙 위반과 같은 사회적 기술상의 문제까지 넓은 범위로 나타난다. 예비교사인 학생들이 교육현장에서 겪는 장애학생의 문제행동에 대해 문제해결 능력 습득을 위한 자기주도 학습의 필요성이 제기된다.

한편, 2023년 말 ChatGPT라는 생성형 AI의 등장으로 사전학습된 지식 자원을 바탕으로 프롬프팅을 거쳐 시공간적 한계 없이 전문가 수준의 지식을 얻을 수 있게 되었다[2]. 이러한 이점을 활용하여 ChatGPT로 가상현실 공간에 특수교육 분야의 전문가AI를 도입한 환경을 구현하였고 이는 예비 특수교사가 특수교육 분야의 전문가와 대화하듯 장애아동의 문제행동에 관한 자문을 받고 대처 방안을 학습하게 할 수 있다.

따라서 본 연구에서는 예비교사인 특수교육과 학생들에게 ‘검색엔진 검색’, ‘텍스트 기반 ChatGPT’, ‘VR환경에서의 음성기반 ChatGPT’에서의 장애아동이 보이는 문제행동의 대처 방안에 대한 학습을 하고, 학습이 도움이 되는지에 대해 학생들의 자신감을 포함한 인식을 조사하고자 한다.

## II. 연구 방법

본 연구에서는 ‘특수교육 전문가AI’가 예비 특수교사의 장애학생의 문제행동 대처에 도움이 되는지 조사하기 위해 특수교육과 학생 17명을 대상으로 3개 그룹을 나누어 조사를 실시하였다. ‘특수교육 전문가AI’의 효과를 검증하기 위해 실험집단을 ‘검색엔진 (n=5)’, ‘텍스트기반 ChatGPT (n=6)’, ‘VR환경에서의 음성기반 ChatGPT (n=6)’ 사용 집단으로 구분하였다. 실험 수행 전 사전 설문을 실시하여 문제

행동 대처에 대한 인식을 조사하였다. 다음으로는 장애아동의 문제행동에 대한 대처 방안에 대한 정보 수집과 학습을 각 30분간 매체를 사용하여 실시하였다. 학습이 끝난 후, 전체 실험집단을 대상으로 사후 설문을 진행하였다.

본 연구에 사용된 음성형 전문가AI는 Unity를 이용해 구현되었다. Unity는 3D 프로그램 또는 VR 콘텐츠 제작 도구로, 몰입감 있는 VR 콘텐츠 구현이 가능한 프로그램이다[3].



그림 457. Unity로 구현한 음성형 전문가 AI  
Fig 1. Voice-type expert AI implemented in Unity

예비 특수교사들이 전문적인 지식을 가진 사람과 대화하며 배우는 느낌을 받을 수 있도록 [그림 1]에서 보이는 것처럼 강의실 환경을 구현하였으며, 교수 NPC(Non-player character)를 배치하여 실제 교육현장과 유사한 경험을 제공하고자 하였다. 이 NPC는 ChatGPT API[4]와 연결되어 자연어 처리가 가능하며, TTS와 STT 기능이 추가되어 사용자가 음성으로 대화할 수 있게 구현하였다. 또한, NPC위에 대화 UI를 배치하여 사용자가 NPC와의 상호작용 내용을 시각적으로 확인할 수 있게 하였다.

UI는 음성 녹음, GPT 답변, 대화 기록으로 구성되어 있으며, 사용자는 녹음 버튼을 클릭해 질문을 음성으로 입력할 수 있다. NPC의 응답은 음성으로 출력되고 동시에 UI 창에 텍스트로 표시된다. 이를 통해 사용자는 대화 내용을 시각적, 청각적으로 모두 확인할 수 있어 직관적이고 몰입감 있는 상호작용이 가능하다.

특수교육 전문가AI를 구현하기 위해 ‘텍스트 및 VR환경에서의 ChatGPT’는 사전에 장애아동에 대한 대처 방안과 특수교육학적 지식을 전달하기 위해

다음과 같은 프롬프트를 구성하였다.

‘당신은 특수학교 근무 경력 20년 차 특수교사입니다. 예비 특수교사들이 지금부터 당신에게 질문을 할 것입니다. 당신은 장애아동의 장애 유형과 정도에 따라 나타나는 문제행동에 대해 전문적인 지식과 경험이 있는 전문가입니다. 그러므로 학교 현장 경험이 없는 예비 특수교사에게 구체적인 자문을 해 주어야 합니다. 답변은 전체적으로 자연스러운 대화 형식을 갖추고 교사가 해야 하는 대응에 대해 구체적인 예시를 제시하여야 합니다. 예비교사의 질문과 응답이 완료되면 다시 질문을 받는 형태로 반복할 것입니다.’

### III. 연구 결과

본 연구조사 분석 결과, ‘검색엔진’, ‘텍스트 기반 ChatGPT’, ‘VR환경에서의 음성기반 ChatGPT’ 그룹별로 장애아동이 보이는 문제행동의 대처 방안에 대해 학습을 하고, 학습이 도움이 되는지에 대해 학생들의 자신감을 포함한 인식을 조사하였다. 그룹별 사전·사후 자신감의 평균 차이는 ‘텍스트 기반 ChatGPT(1.00)’, ‘VR환경에서의 음성기반 ChatGPT(0.95)’, ‘검색엔진(0.53)’순으로 높게 나타났다.

표 1. 그룹별 사전사후 비교

Table 1. Pre- and Post-test comparison by groups

	① 검색엔진	② Text ChatGPT	③ Voice ChatGPT
사전	3.40	2.17	2.94
사후	3.93	3.17	3.89
증가	<b>0.53</b>	<b>1.00</b>	<b>0.95</b>

또한, 참여한 학생들이 생각하는 자기주도 학습에 도움이 되는 방식은 ‘VR환경에서의 음성기반 ChatGPT(4.50)’, ‘텍스트 기반 ChatGPT(4.33)’, ‘검색엔진 검색(3.80)’순으로 높게 나타났다.

표 2. 그룹별 자기주도학습 도움 정도

Table 2. Perceived Level of Helpfulness for Self-Directed Learning by Group

	① 검색엔진	② Text ChatGPT	③ Voice ChatGPT
자기주도 학습 도움	3.80	4.33	4.50

## IV. 결 론

본 연구는 ‘검색엔진’, ‘텍스트기반 ChatGPT’, ‘VR환경기반 음성형 ChatGPT’를 활용하고 장애아동이 보이는 문제행동의 대처 방안에 대한 예비 특수교사들의 자신감을 포함한 인식을 조사하였다. 설문문을 통해 진행된 인식 조사에서 예비 특수교사들은 검색엔진을 사용하여 학습한 그룹보다 텍스트 및 VR기반 음성형 ChatGPT 사용 그룹이 문제행동을 대처하는 자신감 향상에 긍정적으로 도움이 된다고 인식하고 있음을 알 수 있다.

본 연구에서는 VR환경 및 텍스트 기반의 대화형 인공지능인 ChatGPT를 적용한 ‘특수교사 전문가AI’의 사용은 예비교사들이 교육현장에서 경험하는 장애아동의 문제행동에 대해 적절한 대처 방안을 수립하는 데 도움이 될 수 있다는 가능성을 확인하였다.

## 참 고 문 헌

- [1] S. Y. Hwang, K. R. Lee and H. H. Lee, “The Problem Behavior Types of Students with Disabilities and the Recognition of Teachers' Intervention Practices”. The Journal of Special Children Education, vol. 16, no. 2, pp. 69-98, 2014.
- [2] S. E. Park and J. Y. Kang, “Analysis of Prompt Engineering Methodologies and Research Status to Improve Inference Capability of ChatGPT and Other Large Language Models”. Journal Intelligence and Information Systems, vol. 29, no. 4, pp. 287-308, 2023.
- [3] Unity Technologies, <https://unity.com/>
- [4] ChatGPT API, <https://openai.com/index/openai-api/>

# End-to-End 개인화 추천 지속 학습 시스템

장찬우\*, 서영덕\*\*

## End-to-End Continual Personalized Recommender System

ChanWoo Jang\*, and Young-Duk Seo\*\*

### 요약

지속 학습 기반 추천시스템은 사용자 선호 변화를 반영할 수 있게 되었지만, 단계적 학습 처리로 인한 성능 저하와 치명적 망각 현상의 한계가 있다. 본 논문에서는 이를 해결하기 위해 동적 샘플링과 정규화를 End-to-End 방식으로 학습하는 지속 학습 프레임워크를 제안한다. 제안된 모델은 사전 계산 없이 사용자 선호 변화를 반영하며, 정확하고 효과적인 추천을 제공한다.

### Abstract

Continual learning-based recommendation systems have enabled the reflection of user preference changes; however, they face limitations such as performance degradation due to stepwise learning processes and catastrophic forgetting. To address these issues, this paper proposes a continual learning framework that incorporates dynamic sampling and regularization in an end-to-end learning approach. The proposed model reflects user preference changes without requiring prior calculations, thereby providing accurate and effective recommendations

### Keywords

Continual Learning, Personalized Recommender System, End-to-End Learning

## 1. 서론

개인화 추천 시스템의 핵심은 사용자 선호의 변화를 포착하여 만족스러운 결과를 제공하는 데 있다 [1]. 최근 사용자의 선호를 빠르게 반영하기 위해 지속학습 기반 추천시스템이 제안되었으나 [3-4], 이를 안정적으로 구현하는 데는 여전히 한계가 있

다. 특히, 사용자 선호 변화에 유연하게 대응하면서 기존 선호도를 유지하기 위해서는 가소성과 안정성의 균형이 필수적이다 [2].

Recup [2]은 사용자 임베딩 벡터 분포의 비교하여 선호도 변화량을 계산한다. 이를 통해 적응형 정규화와 적응형 재사용을 통해 가소성과 안정성을 유지한다. 그러나, 선호도 변화량을 사전에 계산하

\*인하대학교 통계학과, J12202050@inha.edu

\*\*인하대학교 전기컴퓨터공학과, mysid88@inha.ac.kr(교신저자)

※ 본 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2022R1C1C1012408, 우수신진 연구)과 정보통신기획평가단의 지원(No.2022-0-00448/RS-2022-II220448, 인간처럼 회상이 가능한 인공 신경망 지속학습 플랫폼 개발, No.RS-2022-00155915, 인공지능융합혁신인재양성(인하대학교))을 받아 수행한 연구임.



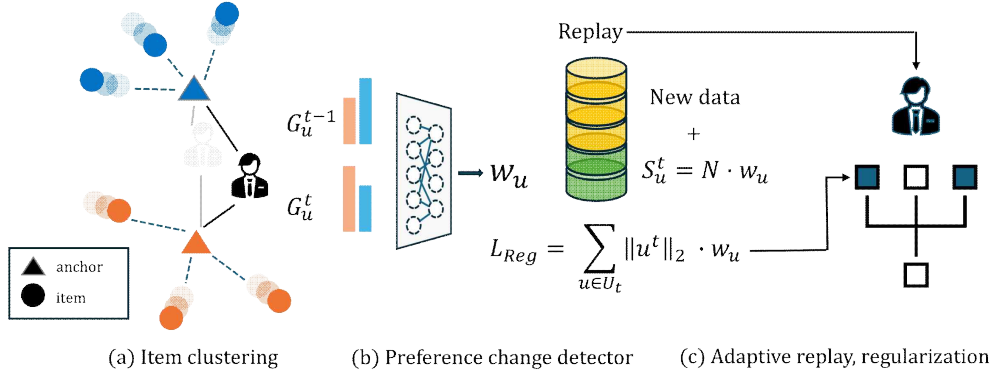


그림 1. 전체 개요

Fig. 1. Overview

기 때문에 동적으로 사용자 변화량의 반영이 어려우며, 중간 단계에서 정보 손실이 발생할 수 있다 [4]. 한편, PIW [3]는 사용자 선호 변화를 세밀하게 반영하기 위해 설계된 End-to-End 기반의 접근법이다. 사용자 임베딩과 아이템 클러스터 앵커 간의 거리를 측정해 사용자별 변화량을 조정한다. 이를 통해 변화가 빠른 사용자에 대한 가소성을 강화하지만, 안정성을 유지하는 데는 한계가 있다.

본 논문에서는 End-to-End 지속 학습을 통해 Recup과 PIW를 결합한 새로운 프레임워크를 제안한다. 이 프레임워크는 중간 정보 손실과 추가적인 계산 요구 없이 동적으로 사용자 선호 변화를 반영하며, 자동화된 개인화 추천을 제공한다.

## II. 제안 방법

본 논문에서 제안하는 프레임워크는 개인화 선호도 변화량을 학습 중 동적으로 학습하여 샘플링 크기와 정규화 강도를 동적으로 조절한다. 이를 통해 사용자 선호 변화에 즉각적으로 적응할 수 있다. 제안하는 프레임워크의 구조는 그림 1과 같다.

학습된 아이템 임베딩을 목표 분포에 근접시켜 정확한 아이템 분포를 학습한다 (a). 이후, 아이템 클러스터 중심과 사용자 간 거리 분포의 차이를 계산하여 선호도 변화를 탐지한다 (b). 마지막으로, 탐지된 사용자 선호도 변화량을 동적 샘플링 및 정규화 과정에 활용하여 추천을 제공한다 (c).

### 2.1 개인화 선호도 변화량 계산

사용자  $u$ 의  $t$ 와  $t-1$ 시점의 임베딩 차이를 2-layer MLP로 처리하여 개인화 선호도 변화량  $w_u^t$ 를 학습한다. 이  $w_u^t$ 는 사용자별 샘플링 크기와 정규화 강도를 결정하는 데 활용된다.

### 2.2 사용자별 샘플링 크기 조절

개인화 선호도 변화량  $w_u^t$ 을 기반으로 사용자  $u$ 의  $t$ 시점 샘플링 크기  $S_u^t$ 는 다음과 같이 계산한다.

$$S_u^t = \frac{N}{1 + kw_u^t} \quad (1)$$

해당 과정을 통해 선호도 변화의 차이가 큰 사용자에게 더 적은 과거 데이터를 반영한다. 반대로, 선호도 변화가 적은 사용자에게 더 많은 과거 데이터를 반영해 안정성을 보장한다. 여기서  $k$ 는 샘플링 비율을 조절하는 상수이다.

### 2.3 정규화 강도 조절

또한,  $w_u^t$ 로 정규화 강도를 동적 조절해 가소성과 안정성의 균형을 맞출 수 있고, 학습 중 치명적 망각 현상을 방지한다. 정규화 손실  $L_{Reg}$ 은 다음과 같이 정의된다.

$$L_{Reg} = \sum_{u \in U^t} \frac{1}{w_u^t + \epsilon} \cdot \|u_u^t\|_2 \quad (2)$$

표 1. 실험 결과 (Recall@20)  
Table 1. Experimental results (Recall@20)

Baseline	Block 1	Block 2	Block 3	Avg.
Finetune	0.0581	0.0635	<b>0.0684</b>	0.0633
LWC-KD	0.0522	0.0429	0.0397	0.0449
LWC-KD-PIW	0.0538	0.0392	0.0426	0.0452
<b>Ours</b>	<b>0.0597</b>	<b>0.0656</b>	0.0681	<b>0.0644</b>

## 2.4 최종 손실 함수

제안하는 프레임워크의 최종 손실 함수는 동적 정규화 손실  $L_{Reg}$ , 아이템 클러스터링 손실  $L_{kl}^{soft}$ , 그리고 추천시스템에서 주로 사용되는 BPR 손실  $L_{BPR}$ 을 포함한다.

$$L_{total} = L_{BPR} + L_{kl}^{soft} + \lambda_{Reg} L_{Reg} \quad (3)$$

여기서  $\lambda_{Reg}$ 는 정규화 손실 강도를 결정하기 위한 가중치이다.

## III. 실험 및 평가

### 3.1 실험 환경

본 연구에서는 제안된 방법의 성능을 검증하기 위해 영화 추천 데이터셋인 MovieLens-1M을 사용하였다. 백본 모델로는 Graph Collaborative Filtering 기반의 LightGCN [6]을 사용하였으며, 비교 모델로는 Fine-tune, LWC-KD, LWC-KD-PIW [3] 모델을 선정하였다. 데이터셋은 시간 순서에 따라 기본 블록과 4개의 재학습용 블록으로 나누어 학습을 진행하였다. 모델 성능 비교를 위해 Recall@20 지표를 사용하여 모델의 각 추천 성능을 평가하였다.

### 3.2 실험 결과

실험 결과 제안된 방법이 Baseline들보다 우수한 성능을 보임을 나타내며, Block 3를 제외하면 가장 개선된 성능을 보인다. 평균 점수는 시간에 따라 변

화하는 사용자 선호에 적응하는 제안 방법의 효과성을 반영한다.

## IV. 결론

본 논문에서는 사용자 선호도 변화를 End-to-End 구조로 지속 학습하는 추천 시스템을 제안한다. 제안된 방법은 중간 정보 손실 및 사전 계산 없이 사용자의 선호도를 동적으로 반영할 수 있다. 실험 결과, 제안 모델은 기존 비교 모델 대비 우수한 성능을 나타냈으며, 사전 계산 없이도 효율적인 개인화 추천이 가능함을 확인하였다.

향후 연구에서는 다양한 데이터셋 적용과 모델 경량화를 통해 제안 모델의 처리 성능을 더욱 향상시키는 방향으로 확장할 계획이다.

## 참 고 문 헌

- [1] S. Maneeroj, N. Sritrakool, "An End-to-End Personalized Preference Drift Aware Sequential Recommender System With Optimal Item Utilization", IEEE Access, vol. 10, pp. 62932-62952, 2022.
- [2] J. W. Jeong, H. J. Kim, Y. D. Seo, "Reflecting Changes in User Preference for Continual Learning in Recommender Systems", KIISE Transactions on Computing Practice, Vol. 30, no. 8, pp. 382-291, 2024
- [3] Y. Wang, Y. Zhang, A. Valkanas, R. Tang, C. Ma, J. Hao, and M. Coates, "Structure aware incremental learning with personalized imitation weights for recommender systems", AAAI Conference on Artificial Intelligence, pp. 4711-4719
- [4] Y. Li, K. Lui, R. Satapathy, S. Wang, E. Cambria, "Recent Developments in Recommender Systems: A Survey", IEEE Comput. Intell. Mag, 19, 2 (May 2024), 78-95.
- [5] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "LightGCN: Simplifying and powering graph convolution network for recommendation", Proc. ACM Int. Conference on Research and Development in Information Retrieval, Xi'an, China, 25-30 July 2020; pp.639-648

# Unity Engine 및 생성형 AI 기반 영어회화 학습 플랫폼 개발

박재성\*, 박민호\*, 안지훈\*, 김성영\*\*

## Development of a English Conversation Learning Platform Leveraging Unity Engine and Generative AI

JaesungPark\*, MinhoPark\*, JihunAn\*, and SungyoungKim\*\*

### 요 약

본 논문에서는 메타버스와 인공지능을 결합한 영어 회화 학습 플랫폼을 제안한다. 사용자는 유니티 엔진을 통해 구현한 오픈월드 가상 공간을 이동하면서 가상 공간에 배치된 NPC들과 자유롭게 상호작용 할 수 있다. 사용자의 발화는 STT를 통해 텍스트로 변환되어 GPT-4로 전달되고, GPT-4의 응답은 다시 TTS를 통해 음성으로 변환되어 NPC를 통해 재생된다. 각 NPC들은 고유한 이름과 역할을 지니고 있으며, 이는 프롬프트 엔지니어링을 통해 GPT-4에 함께 전달되어 각 NPC의 이름과 역할에 맞는 응답을 반환할 수 있도록 하였다. 이를 통해 사용자에게 현실감 있는 대화 경험을 제공한다. 대화가 종료되면 대화내용을 분석하여 피드백을 제공함으로써 사용자가 회화능력을 발전시키는 데 도움을 줄 수 있을 것으로 기대한다.

### Abstract

In this paper, we propose an English conversation learning platform that combines the metaverse with artificial intelligence. Users navigate an open-world virtual space implemented using the Unity Engine, where they can freely interact with NPCs placed throughout the environment. The user's speech is converted into text through STT technology and sent to GPT-4, which generates responses that are then converted back into speech via TTS and played by the NPCs. Each NPC is assigned a unique name and role, which are integrated into the prompts sent to GPT-4, ensuring that the responses align with each NPC's identity. This setup offers users a realistic conversational experience. Upon the conclusion of interactions, the system analyzes the conversation and provides feedback, aiding users in enhancing their conversational skills.

### Key words

Generative AI, Metaverse, English Conversation, GPT

---

\* 국립금오공과대학교 컴퓨터공학과, pjs1193@kumoh.ac.kr, alsgh4442@kumloh.ac.kr, ajh0129@kumoh.ac.kr

\*\* 국립 금오공과대학교 컴퓨터공학과, sykim@kumoh.ac.kr, 교신저자

## I. 서론

최근 메타버스와 인공지능 기술의 급격한 발전은 교육 분야, 특히 언어 학습에 새로운 가능성을 제시하고 있다[1][2][3]. 메타버스는 다양한 가상 환경을 제공하여 학습자에게 몰입감 있는 학습 경험을 선사한다. 동시에, 인공지능 기술, 특히 GPT-4[4]와 같은 대규모 언어 모델은 자연스러운 대화를 가능하게 하여 언어 학습에 혁신적인 도구로 자리매김하고 있다. 본 논문에서는 메타버스와 인공지능을 결합한 영어 회화 학습 플랫폼을 제안한다. 사용자는 유니티 엔진을 통해 구현된 오픈월드 가상 공간을 이동하며, 가상 공간에 배치된 NPC들과 자유롭게 상호작용할 수 있다. 사용자의 발화는 음성 인식(STT)을 통해 텍스트로 변환되어 GPT-4로 전달되고, GPT-4의 응답은 음성 합성(TTS)을 통해 NPC를 통해 재생된다. 각 NPC는 고유한 이름과 직업을 지니고 있으며, 이는 프롬프트 엔지니어링[5]을 통해 각 NPC의 역할에 맞는 응답을 반환하도록 설계되었다. 이를 통해 현실감 있는 대화 경험을 제공하며, 대화가 종료되면 대화 내용을 분석하여 피드백을 제공함으로써 사용자의 회화 능력 향상을 돕는다. 이 플랫폼은 학습자가 시간과 공간의 제약을 완화하여 효율적인 언어 학습을 할 수 있도록 지원하며, 인공지능과 메타버스의 결합을 통해 보다 효과적인 학습 환경을 제공한다.

## II. 시스템 아키텍처 설계

본 논문에서 소개하는 영어회화 학습 플랫폼의 전체적인 흐름은 그림1과 같다.

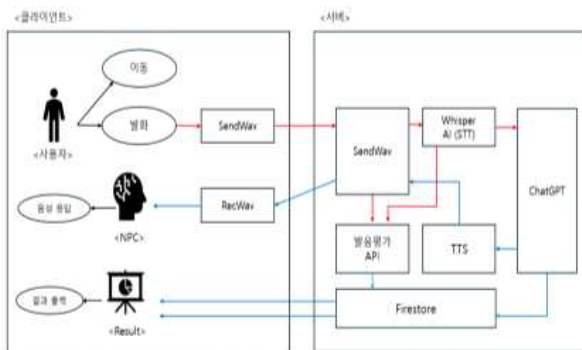


그림 1. 시스템 아키텍처  
Fig. 1. System Architecture

사용자는 유니티로 구현된 가상 공간에서 자유롭게 이동할 수 있으며, 가상 공간에 배치된 NPC들과 상호작용 할 수 있다. 사용자가 상호작용을 시작하면, 서버와 사용자는 Python 기반의 TCP 소켓 통신을 이용해 통신한다. 상호작용 시 사용자가 발화한 음성은 STT를 통해 텍스트로 변환되어 GPT-4로 전달되고, GPT-4는 이를 입력으로 받아 적절한 응답 텍스트를 출력한다. 출력된 텍스트는 TTS를 통해 다시 음성으로 변환되어 유니티로 전달된다. 대화가 종료되면 GPT-4 지금까지 진행한 대화를 분석하여 대화에 대한 피드백을 출력한다. 서버는 Python 소켓프로그래밍을 이용하여 TCP로 통신하도록 구현하였고 데이터베이스는 NoSQL 데이터베이스인 구글의 Firestore[6]를 이용하였다.

## III. 개발 내용

### 3.1 가상 공간 구현

사용자가 자유롭게 이동하면서 현실감 있는 대화를 경험할 수 있도록 유니티엔진을 이용해 오픈월드 그림 2와 같은 가상 공간을 구현하였다. 사용자가 여러 상황을 가정하여 영어를 사용할 수 있도록 마을, 공원, 상점 등 다양한 장소를 포함하며, 카툰풍의 에셋인 Toon City[7]로 제작하여 사용자에게 가볍고 친근하게 다가갈 수 있는 분위기를 조성하였다.



그림 2. 유니티 엔진으로 구현한 가상 공간  
Fig. 2. Virtual space implemented by Unity

그림 3처럼 가상 공간에 배치된 NPC들은 스크립트를 통해 각각의 직업과 이름을 부여받는다. NPC들의 이름과 직업은 상호작용 시 서버로 전달되고,

서버는 응답 생산 시 해당 정보들을 참조하여 생산되는 응답의 다양성을 높였다.



그림 3. 가상 공간 속 NPC  
Fig. 3. NPC in virtual space

사용자는 Tab키를 눌러 그림 4와 같은 미니맵을 펼쳐 가상공간의 구조를 확인할 수 있고, NPC와 플레이어의 위치는 미니맵 상에 마커로 표기하여 플레이어는 NPC와 자신의 위치를 확인할 수 있다.



그림 4. 미니맵 기능  
Fig. 4. Minimap

### 3.2 사용자 UI 구현

NPC와 상호작용 시 사용자 UI가 나타나고 대화 시작, 진행, 종료 등 과정이 원활하게 이루어지도록 그림 5와 같이 UI를 배치하였으며, 이를 통해 사용자가 직관적으로 플랫폼을 이용할 수 있도록 했다.



그림 5. 상호작용 시 유저 인터페이스  
Fig. 5. user interface in interact situation

사용자가 NPC에 접근하여 상호작용 키를 누르면 NPC와 대화를 진행할 수 있다. 대화 로그 창을 통해 대화 내용을 실시간으로 확인할 수 있으며, 사용자, 혹은 서버 응답 로그를 선택적으로 비활성화 하는 기능을 제공하여 회화 학습에 도움을 준다. NPC는 서버로부터 전달받은 감정에 따라 적절한 애니메이션을 실행시키도록 하여 행동을 풍부화하였다. 사용자는 녹음 버튼을 통해 대화를 시작하고 Send 버튼을 통해 음성을 전송할 수 있는데, 사용자가 직접 녹음 및 전송을 제어하도록 하여 원하지 않는 음성이 녹음 및 전송되는 것을 방지하였다. Exit 버튼을 통해 상호작용을 종료하고 결과를 확인할 수 있다.

### 3.3 서버 구현

사용자가 NPC와 상호작용 시 NPC의 이름과 직업 정보를 받아서 프롬프트를 설정한다. 그 후 사용자의 음성이 들어오면 WhisperAI[8]로 구현된 STT로 음성을 분석하여 텍스트 정보로 반환하여 ChatGPT로 보낸다. ChatGPT는 설정된 프롬프트를 바탕으로 텍스트를 생성하여 그림 6처럼 반환한다. 반환된 텍스트는 VITS[9]로 구현된 TTS에서 음성으로 변환하여 유니티로 송신한다.

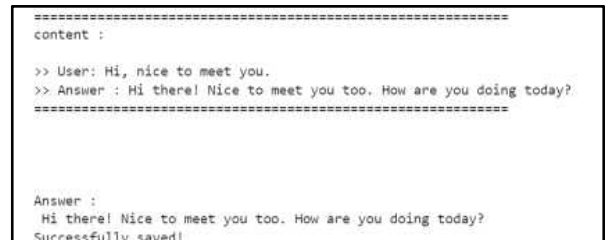


그림 6. 반환된 텍스트  
Fig. 6. Text output

사용자가 NPC와 대화를 하며 서버와 송수신하는 과정에서 Firestore로 대화내용은 저장된다. 그림 7은 대화기록이 저장되는 데이터베이스이다. 데이터베이스 구조는 User - [NPC이름] - log - [시퀀스번호] - dialogue\_log로 설계하였고 사용자가 대화하는 NPC의 이름별로 구분하여 저장한다. 또한 log 내부의 metadata는 Int형 정수값을 가지는데 이는 마지막 시퀀스번호를 가리키고 대화가 저장될 때 metadata의 정수값은 1씩 증가하게 된다. log는 새로 생성될

때마다 metadata의 값을 참조하여 생성된다.



그림 7. 데이터베이스  
Fig. 7. Database

dialogue\_log는 message, speaker\_id, timestamp 총 3 가지 값을 저장한다. message는 대화 내용이 기록되고 speaker\_id에는 해당 대화를 말하는 주체가 기록된다. 이 값들은 추후 피드백 결과를 생성하는 데 사용된다.

### 3.4 피드백 기능 구현

사용자가 NPC와 대화한 내용을 기반으로 ChatGPT가 분석하여 분석내용을 반환한다. 피드백은 사용자가 원하는 대화 기록을 선택하여 확인이 가능하다

```
content :
>> User: Hi, nice to meet you. What's your name?
>> Answer: Hi! Nice to meet you too. My name is Alice. What's your name?
>> User: My name is Jack. How are you?
>> Answer: I'm doing well, thanks for asking, Jack! How about you? How are you doing today?
>> User: I lost my wallet, so I'm so sad.
>> Answer: Oh no, I'm really sorry to hear that, Jack. Losing a wallet can be really upsetting. Is there anything I can do to help?
>> User: Thank you, but I have to go to school now. Sorry, see you later.
>> Answer: Of course, Jack. I hope things get better soon. Take care, and I'll see you next time!
emotion : sadness
```

그림 466. 대화 내용  
Fig 8. Conversation

```
평가 및 분석
- 대화의 흐름 및 응답 -
사용자의 대화는 자연스러운 인사와 상호 소개로 시작하여, 개인적인 상황 공유로 이어지고 마지막으로 급히 종료되는 흐름을 가지고 있습니다. 대화는 각 단계에서 연속성을 유지하며, 사용자는 감정을 표현하고 바쁜 일정을 고려하여 대화를 마무리합니다. 이러한 구성은 일상 대화에서 흔히 볼 수 있는 패턴을 따릅니다.
- 문맥 및 문법 수정 -
[감정 표현의 수정]
원문: "I lost my wallet, so I'm so sad."
수정 제안: "I've lost my wallet and I'm very sad about it."
설명: "so"의 반복을 피하고 "very"로 변경하여 문장을 좀 더 문법적으로 자연스럽게 합니다. 또한, 완료형 "I've lost"를 사용하여 사건의 완결성과 지속적인 영향을 강조합니다.
[대화 종료 문장의 수정]
원문: "Thank you, but I have to go to school now. Sorry, see you later."
수정 제안: "Thank you for your understanding, but I need to get to school now. See you later!"
설명: "for your understanding"을 추가하여 감사의 표현을 좀 더 구체적으로 하며, "need to get to"로 변경하여 일상적인 표현을 사용합니다.
- 개선점 및 제안 -
[다양한 감정 표현 방법]
원문: "I'm so sad."
개선된 예문:
"I'm really upset about losing my wallet."
"I feel quite disheartened after misplacing my wallet."
"I'm truly distressed because my wallet is gone."
설명: 이 예문들은 각각 "upset," "disheartened," "distressed"를 사용하여 같은 상황에 대해 다양한 감정을 표현함으로써 언어의 풍부함을 보여줍니다. 이는 상황에 따른 더 깊은 감정 이해와 공감을 촉진할 수 있습니다.
[보다 상세한 배경 제공]
원문: "I lost my wallet."
개선된 예문:
"I lost my wallet somewhere in the cafe this morning. It had all my credit cards and my ID!"
"My wallet went missing during my commute today, and it contained important documents and my driver's license."
설명: 이 예문들은 종료할 때 사용하는 표현을 더 활동적이고 긍정적으로 변경하여 대화를 기억에 남는 방식으로 마무리합니다. 이는 상대방과의 관계를 강화하고, 다음 대화를 기대하게 만드는 긍정적인 기대감을 조성합니다.
[적극적인 대화 종료 방법]
원문: "Sorry, see you later."
개선된 예문:
"I must dash off to class now, but let's catch up later this week!"
"I really have to run now, but I'm looking forward to our next chat. Catch you soon!"
설명: 이 예문들은 종료할 때 사용하는 표현을 더 활동적이고 긍정적으로 변경하여 대화를 기억에 남는 방식으로 마무리합니다. 이는 상대방과의 관계를 강화하고, 다음 대화를 기대하게 만드는 긍정적인 기대감을 조성합니다.
```

그림 467. 피드백 내용  
Fig 9. Details of the feedback

다. 그림 8의 대화 내용을 기반으로 평가 및 피드백 실행 시 피드백이 출력된다. 또한, 유니티 환경에서 사용자가 그림 5의 Result버튼을 눌렀을 때 결과화면이 출력되어 사용자는 각각의 문장에 대해서 ChatGPT가 분석한 내용을 확인할 수 있다. 문장마다 피드백 내용을 확인하는 것뿐만 아니라 전체적인 대화 내용의 문법, 단어 선택, 대답의 적합성을 분석하여 점수로 평가하고 이를 그래프로 시각화하여 확인할 수 있다.

#### IV. 결과

본 논문에서는 메타버스와 인공지능을 결합한 영어 회화 학습 플랫폼을 제안한다. 이 플랫폼은 사용자에게 효율적이고 몰입감 있는 언어 학습 경험을 제공하는 것을 목표로 한다. 사용자는 유니티 엔진으로 구현된 가상공간에서 다양한 NPC와 상호작용하며 영어 회화를 연습할 수 있다. 사용자의 발화는 음성 인식(STT)을 통해 텍스트로 변환되어 GPT-4로 전달되며, GPT-4의 응답은 음성 합성(TTS)을 통해 다시 음성으로 재생된다. 이 과정에서 각 NPC는 고유한 이름과 역할을 가지고 있어, 프롬프트 엔지니어링을 통해 NPC의 역할에 맞는 응답을 제공할 수 있다. 대화가 종료되면 시스템이 대화 내용을 분석하여 어휘, 문법, 발음 등의 요소에 대한 피드백을 제공하고 이를 통해 사용자는 자신의 회화 능력을 점검하고 지속적으로 발전시킬 수 있다. 추후, 피드백 화면 구현 및 개선과 더 몰입감있는 학습환경 제공 및 효과적인 학습을 위한 피드백을 위해서 NPC의 감정 추가 및 감정 TTS를 이용한 회화과정에서의 자연스러움 증가, 사용자 음성의 빠르기, 발음 등을 분석 기능을 추가하여 텍스트 분석뿐만 아닌 음성 분석 기능을 추가하여 피드백 기능을 강화할 예정이다.

#### References

[1] Eun-jeong Song. "A Study on the Effects and Application Cases of Education Using Metaverse in the Non-Face-To-Face Era." *Journal of practical engineering education*, vol. 14, no. 2, pp. 361-366,

2022.

- [2] A. Berns, A. Gonzalez-Pardo, and D. Camacho, "Game-like language learning in 3-D virtual environments," *Computers & Education*, vol. 60, no. 1, p. p. 210-220, Jan. 2013.
- [3] Shinjung Park and Eunjee Song, "A Proposal of Korean Education Program for Multi-cultural Families through Metaverse," in *한국정보통신학회 종합 학술대회 논문집*, 2022, pp. 581-583.
- [4] J. Achiam, B. McGrew *et al.*, "Gpt-4 technical report," arXiv preprint arXiv:2303.08774.
- [5] J. White *et al.*, "A prompt pattern catalog to enhance prompt engineering with ChatGPT," arXiv preprint arXiv:2302.11382, 2023.
- [6] Firebase, <https://firebase.google.com/> [Accessed 25 October 2024].
- [7] Toon City, [https://assetstore.unity.com/packages/3d/environments/urban/toon-city-88379?locale=ko-KR&srsltid=AfmBOoqY9XWfCFQRWoBtUE\\_ASdFyJZyeLFxqTCA9XBZdlgd9-YgWdkN9](https://assetstore.unity.com/packages/3d/environments/urban/toon-city-88379?locale=ko-KR&srsltid=AfmBOoqY9XWfCFQRWoBtUE_ASdFyJZyeLFxqTCA9XBZdlgd9-YgWdkN9)
- [8] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," In *International Conference on Machine Learning*, pp. 28492-28518, Jul 2023.
- [9] J. Kim, J. Kong, and J. Son, "Conditional variational autoencoder with adversarial learning for end-to-end text-to-speech," In *International Conference on Machine Learning*, pp. 5530-5540, Jul 2021.

# Colla : 대학생들을 위한 팀 프로젝트 관리 플랫폼

박준하\*, 오세찬\*, 설진영\*, 박재민\*, 안형태\*\*

## Colla : Team Project Management Platform for College Students

Junha Park\*, Sechan Oh\*, Jinyeong Seol\*, Jaemin Park\*, and Hyeongtae Ahn\*\*

### 요약

대학생들은 팀 프로젝트 수행 중 일정 조율, 자료 관리, 소통 등을 위해 여러 개의 협업 도구를 병행 사용해야 하며, 이로 인해 정보의 산재와 관리 복잡성으로 인해 작업 효율성이 저하되고 있다. 본 논문은 대학생들이 팀 프로젝트를 진행하는 과정에서 발생하는 협업 도구의 파편화 문제를 해결하기 위한 Colla 플랫폼을 제안한다. 제안된 플랫폼은 팀 생성 및 참여, 채팅, 목적별 게시글 작성, 파일 공유 및 자료 저장소 등 다양한 협업 기능을 통합적으로 제공하는 시스템이다. 기존의 복잡한 협업 도구를 단순화하여 대학생들이 쉽게 사용할 수 있는 직관적인 인터페이스를 제공함으로써, 사용자가 느끼는 피로감을 줄이는 등 사용자 경험을 개선한다. 이를 통해 대학생들은 팀 프로젝트 관리를 보다 체계적이고 효율적으로 수행할 수 있다.

### Abstract

University students often use multiple collaboration tools to manage schedules, organize materials, and communicate during team projects. This results in information fragmentation and increased management complexity, leading to decreased work efficiency. This paper proposes the Colla platform to address the fragmentation of collaboration tools encountered by students during team projects. The proposed platform integrates various collaboration features, including team creation and participation, chat, purpose-specific posting, file sharing, and a resource repository. By simplifying existing complex collaboration tools and providing an intuitive interface, Colla aims to reduce user fatigue and enhance the overall user experience. As a result, students can manage team projects more systematically and efficiently.

### Key words

Team Project Management, Collaboration, University Students, React, Spring Boot, File Sharing

---

\*국립금오공과대학교 컴퓨터공학과 학부생 박준하 qkrwnsgk1624@icloud.com

\*국립금오공과대학교 컴퓨터공학과 학부생 오세찬 seb\_s19@naver.com

\*국립금오공과대학교 컴퓨터공학과 학부생 설진영 wlsdud5654@gmail.com

\*국립금오공과대학교 컴퓨터공학과 학부생 박재민 jaemin2167@naver.com

\*\*국립금오공과대학교 컴퓨터공학과 조교수 안형태 anten@kumoh.ac.kr

※ 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1G1A1005366).



## 1. 서론

오늘날 대학 강의에서 팀 프로젝트의 비중이 증가함에 따라, 학생들은 이를 원활하게 진행하기 위해 다양한 협업 도구를 사용하고 있다. 그러나 하나의 팀 프로젝트를 완료하기 위해 메신저, 드라이브, 일정 조율 서비스 등 여러 가지 도구를 동시에 사용해야 하는 경우가 많다. 이러한 방식은 정보가 분산되고, 여러 플랫폼을 이동하며 소통하거나 자료를 관리하는 과정에서 불필요한 시간 소모와 혼란을 초래한다.

대학내일연구소의 '2018년 20대 대학생, 조별 과제에 대한 생각'에 따르면 대학생들은 학기당 평균 3.34개의 조별 과제를 수행하고 있으며, 이는 2012년 한 학기 평균 2.99개에서 증가한 수치이다. 특히, '5개 이상의 조별 과제 수행'이라고 응답한 학생들의 비율은 23.8%로 가장 높았다[1]. 이는 조별 과제의 빈도와 중요성이 점차 증가하고 있음을 나타낸다.

또한, 코로나19 팬데믹으로 인해 디지털 전환이 가속화되면서, 대학생들도 조별 과제에서 비대면 협업 도구를 더욱 의존하게 되었다. 글로벌 협업 도구 시장은 2020년 402억 달러 규모였으며, 2024년에는 483억 달러에 이를 것으로 예상된다(그림 1)[2]. 이는 학생들이 사용하는 협업 도구의 선택 폭을 넓혔으나, 동시에 다양한 도구 간의 통합 부족과 사용의 번거로움도 증가시키고 있다. 국내에서도 협업 도구 사용이 증가하고 있으며, 다양한 서비스들이 시장에서 활성화되고 있다. 이러한 문제를 해결하고자, 본 연구에서는 대학생들이 보다 효율적으로 조별 과제를 수행할 수 있도록 돕는 통합 협업 플랫폼을 제안한다. 본 플랫폼은 파일 공유, 팀 생성, 커뮤니케이션 일정 조율 기능을 하나의 플랫폼에서 제공함으로써, 학생들이 여러 협업 도구를 번갈아 사용하는 번거로움에서 벗어나, 정보가 한 곳에 통합되어 관리될 수 있는 환경을 제공한다. 또한, 개인정보를 주고받지 않고도 안전하게 협업할 수 있는 기능을 갖추고 있어, 학생들이 보다 체계적이고 효율적으로 팀 프로젝트를 진행할 수 있을 것으로 기대된다.

글로벌 협업 툴 시장 전망

자료: IDC

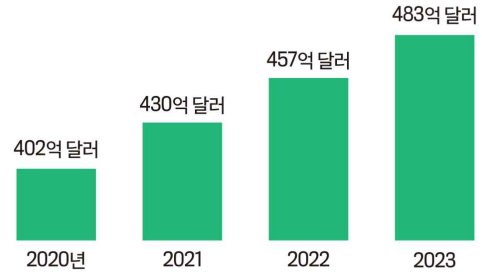


그림 1. 글로벌 협업 툴 시장 전망

Fig. 1. Global Collaboration Tool Market Outlook

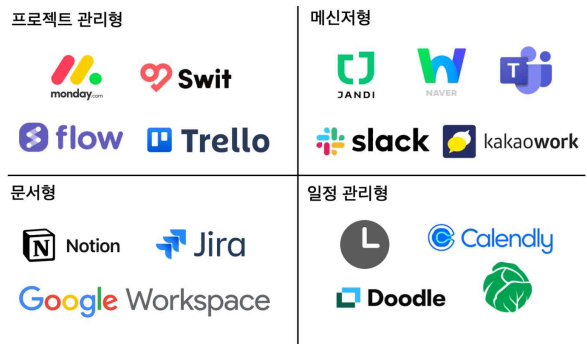


그림 2. 종류별 협업 툴

Fig. 2. Types of Collaboration Tools

## II. 관련사례

노션(Notion)[3]과 플로우(Flow)[4]와 같은 도구들은 팀프로젝트에서 협업을 지원하기 위해 자주 사용되는 플랫폼이다. 각 도구는 고유한 장점을 가지고 있지만, 대학 팀프로젝트 환경에서는 한계점이 존재한다. 노션은 메모 및 문서 기반의 협업에 매우 유리하며, 학생들에게 높은 자유도를 제공하는 장점이 있다. 하지만 실시간 채팅 기능이 부재하여 팀원들 간의 즉각적인 소통이 어렵고, 자유도가 높다는 점이 오히려 학업 환경에서는 러닝 커브가 발생할 수 있다. 플로우는 프로젝트 관리와 업무 분배에 특화된 기능을 제공하지만, 기업 중심으로 설계된 만큼 유료 서비스로 운영되며, 대학생들이 단기적인 팀프로젝트를 위해 사용하기에는 비용적인 부담이 크고, 러닝 커브가 있어 초기에 익히기 어렵다는 단점이 있다.

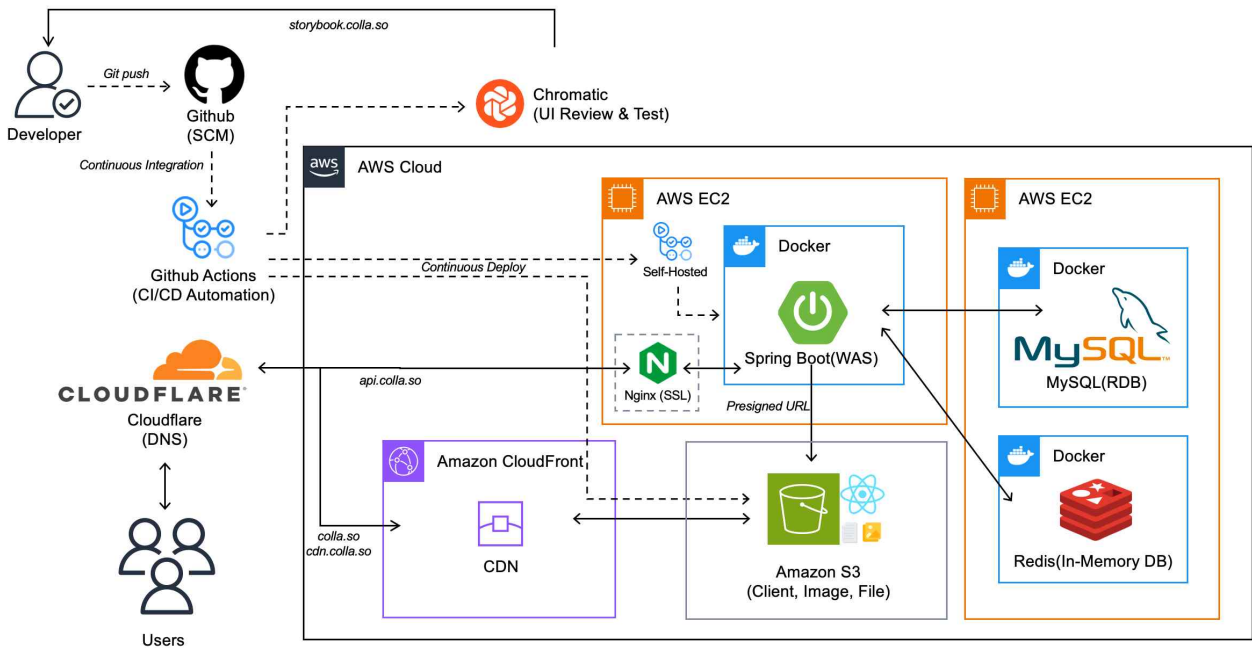


그림 3. 전체 시스템 아키텍처  
Fig. 3. Total System Architecture

본 플랫폼에서는 실시간 채팅과 직관적인 인터페이스 등을 모두 지원한다. 특히, 대학생들은 각기 다른 수업 시간표와 개인 일정 때문에 팀원 모두가 모일 수 있는 공통의 시간을 찾기 어렵다는 문제를 자주 겪는다. 일정 조율 기능을 통해 팀원들이 가능한 시간대를 입력하고, 공통의 회의 시간을 찾아주는 기능을 제공하여 팀프로젝트의 효율성을 높인다.

### III. Colla

#### 2.1 시스템 개요

본 논문에서는 기존의 기업 중심 협업 도구들의 복잡성과 불편화된 서비스 사용으로 인한 불편함을 해소한 새로운 플랫폼을 제안한다. 본 시스템은 팀 생성 및 참여, 팀원 관리, 팀원 초대와 같은 협업의 핵심 기능을 담당하는 팀스페이스를 제공한다. 또한 주제별 대화방을 생성해 기본 메시지 생성부터 사진과 파일 전송을 지원하는 채팅 기능과, 팀원 간 제목, 사진, 내용을 조회할 수 있는 게시글(이하 ‘피드’)을 제공한다. 팀원들의 일정을 효율적으로 조율할 수 있는 일정 조율 피드와 팀원 각각의 자료수집 현황을 체계적으로 관리할 수 있는 자료수집 피

드도 함께 제공한다. 더불어 팀스페이스의 피드와 채팅에서 업로드된 파일들을 한눈에 볼 수 있는 자료저장소를 통해 대학생들이 팀프로젝트를 수행하는 데 필요한 다양한 서비스와 편의성을 제공한다.

#### 2.2 시스템 구성

그림 3은 본 시스템의 전반적인 구조를 나타낸다. 본 시스템은 프론트엔드, 백엔드, 데이터베이스, 인프라 등 각 계층별로 최신 기술 스택을 활용하여 구축되었다. 각 계층의 기술들이 유기적으로 결합되어 안정적이고 확장 가능한 서비스를 제공한다.

프론트엔드 개발에는 정적 타입 기반의 TypeScript[5]를 채택하여 코드의 안정성과 개발 생산성을 향상시켰다. Virtual DOM[6] 기반 라이브러리인 React[7]를 사용하여 상태 변경에 따른 UI 업데이트를 효율적으로 관리했다. 서버 상태 관리를 위해 TanStack Query[8]를 도입하여 데이터를 효율적으로 요청하고 캐싱할 수 있도록 했다. 클라이언트 측 상태 관리는 Zustand[9]를 사용하여 React의 단방향 데이터 흐름을 손쉽게 유지하면서도 다양한 컴포넌트에서 상태를 간편하게 공유할 수 있게 했다. 스타일링은 Styled Component[10]를 활용해 공통 컴포넌트

스타일을 효과적으로 관리했다. UI 컴포넌트 개발 과정에서는 Storybook[11]을 활용하여 버튼, 아이콘 등 공통 컴포넌트의 독립적인 개발과 테스트가 가능하도록 하였으며, MSW[12]를 통해 백엔드 API가 완성되기 전에도 프론트엔드 개발이 가능한 환경을 구축했다.

백엔드는 Spring Boot와 Spring Data JPA[13]를 기반으로 구축하여 반복적인 코드를 줄이고 객체 중심의 직관적인 데이터 처리가 가능하도록 했다. 특히 팀스페이스 멤버 관리나 피드 CRUD와 같은 핵심 기능 구현에 있어 JPA의 연관관계 매핑과 트랜잭션 관리가 큰 역할을 했다. RESTful API 설계 및 문서화를 위해 Swagger[14]를 활용하였으며, JUnit 5 [15]를 통해 회원가입, 인증 등 중요 비즈니스 로직에 대한 단위 테스트를 구현하여 코드의 안정성을 확보했다. 데이터 저장소로는 MySQL[16]을 주 데이터베이스로 활용하여 회원정보, 팀스페이스, 피드 등 핵심 데이터를 관리하며, Redis[17]를 도입하여 팀스페이스 초대 코드와 소셜 로그인 세션 관리를 수행했다. MySQL에서는 핵심 데이터를 안정적으로 관리하고, Redis를 캐시 저장소로 활용하여 빈번한 접근이 필요한 데이터를 고속으로 처리하는 구조를 구축했다. 실시간 채팅 기능은 Spring의 WebSocket과 STOMP 프로토콜을 활용하여 구현했다. STOMP의 pub/sub 구조를 통해 팀스페이스 내 채팅방별 메시지 전송을 처리했으며, 채팅 메시지의 영구 저장은 MySQL에서 관리하여 채팅 이력 조회가 가능하도록 했다.

인프라 구축을 위해 Amazon Web Services(AWS) [18]를 활용하였으며, EC2 인스턴스를 통해 애플리케이션 서버를 운영한다. 프론트엔드 애플리케이션의 빌드 결과물과 피드에 업로드되는 이미지 등의 정적 자원들은 S3 버킷에 저장하고, CloudFront[19]를 통해 정적 콘텐츠를 효율적으로 전송하고 웹 사이트의 로딩 속도를 개선했다. Nginx[20]를 리버스 프록시로 활용하여 백엔드 애플리케이션에 대한 HTTPS 통신을 제공하고, SSL/TLS 인증서 관리와 보안 연결을 효율적으로 처리했다. 플랫폼의 모든 구성요소를 독립된 패키지로 묶기 위해 Docker[21]와 Docker Compose[22]를 도입하였으며, 이를 통해 일

관된 개발 및 운영 환경을 제공한다. GitHub Actions[23]를 활용하여 CI/CD 파이프라인을 구축함으로써 자동화된 빌드 및 배포 프로세스를 확립하였으며, 프론트엔드의 경우 빌드 결과물을 자동으로 S3에 업로드하는 과정까지 자동화했다.

현대적인 기술 스택의 조합을 통해 개발 생산성과 시스템 안정성을 극대화하였으며, 실시간 채팅과 피드 기능에서 보여지듯 사용자에게 원활하고 직관적인 서비스 경험을 제공하는 효율적인 시스템을 구축할 수 있었다.

### 2.3 시스템 동작

본 논문에서 제안하는 플랫폼은 크게 네 가지 핵심 기능으로 구성된다. 첫 번째는 회원 관리 기능이다(그림 4). 사용자는 해당 서비스를 사용하기 위해 먼저 회원가입을 한다. 이 과정에서 이메일 인증, 이메일 중복 여부 검사, 비밀번호 확인 검사가 필수적으로 요구된다. 사용자 편의성을 위해 카카오, 네이버, 구글 소셜 미디어 계정을 사용해 OAuth 2.0[24] 기반 간편 회원가입 옵션도 제공된다. 회원가입이 완료되면, 사용자는 로그인 과정을 거쳐 JWT(JSON Web Token)[25]를 이용한 인증 시스템을 통해 액세스(access) 토큰과 리프레시(refresh) 토큰을 발급받을 수 있다. 여기서 액세스 토큰은 사용자 인증 및 인가에 필수적인 역할을 하며, 리프레시 토큰은 사용자가 시스템에 지속해서 접근할 수 있도록 한다.



그림 4. 로그인 및 회원가입 화면  
Fig. 4. Login and Sign Up Screen

사용자의 로그인 절차는 그림 5와 같은 흐름을 따른다. 먼저 사용자가 플랫폼에 접근할 때 클라이언트는 저장된 토큰이 있는지 확인한다. 만약 저장된 토큰이 존재한다면, 서버에서 해당 토큰의 유효성을 검증한다. 이때, 토큰이 유효하다면 사용자에게 별도의 로그인 절차 없이 플랫폼에 접근할 수 있다. 반면, 토큰이 존재하지 않거나 유효하지 않다면 사용자는 로그인 절차를 거친다. 로그인에 성공하면 서버는 새로운 액세스 토큰을 발급하고, 클라이언트는 이를 저장해 이후 요청에 사용하게 된다. 로그인 후 액세스 토큰이 만료될 경우 클라이언트는 axios interceptor[26] 를 통해 리프레시 토큰으로 새로운 액세스 토큰을 발급받고 요청을 재시도하는 방식으로 사용자 경험을 개선한다. 이후 로그인이 성공적으로 이루어지면 사용자와 서버 간에 웹소켓(WebSocket)[27] 연결이 설정되어 실시간 양방향 통신이 가능해진다. 웹소켓 연결 시에도 JWT가 인증 수단으로 사용되며, 이를 통해 실시간 데이터 전송 및 상호작용이 이루어진다. 이러한 과정을 통해 본 플랫폼은 사용자 정보 보호와 편리한 접근을 보장한다.

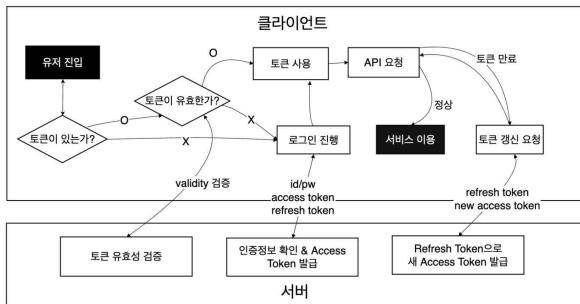


그림 5. 인증 및 인가 절차  
Fig. 5. Authentication and Authorization Process

두 번째는 팀스페이스 기능이다. 사용자는 그림 6과 같이 팀 스페이스를 생성하거나 참가할 수 있다. 팀스페이스 생성자는 팀장이 되고, 참여자는 팀원이 된다. 그림 7과 같이 페이지 상단의 내비게이션 바에서 현재 가입되어 있는 팀스페이스 목록을 볼 수 있으며 선택할 경우 해당 팀스페이스로 이동한다. 사용자는 현재 위치한 팀 스페이스의 팀원 목록을 볼 수 있다. 이메일을 입력해 팀스페이스 초대 메일을 전송하거나 초대 링크를 발급하여 공유함으로써 팀스페이스에 참가할 수 있다. 팀스페이스에

속한 사용자는 자신의 팀스페이스 내에서 피드, 채팅, 자료 저장소 기능을 사용할 수 있게 된다. 팀장은 그림 8과 같이 팀스페이스 이미지와 이름을 설정할 수 있으며, 팀스페이스에 소속된 사용자들의 팀스페이스 내 역할을 생성하고 지정할 수 있다.



그림 6. 팀스페이스 생성 또는 참가 화면  
Fig. 6. Team Space Creation or Join Screen

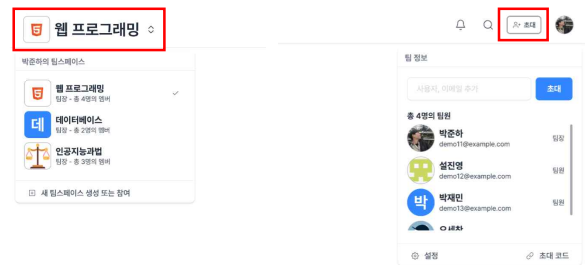


그림 7. 팀스페이스 목록 및 팀 정보 조회 화면  
Fig. 7. TeamSpace List and Team Information Screen

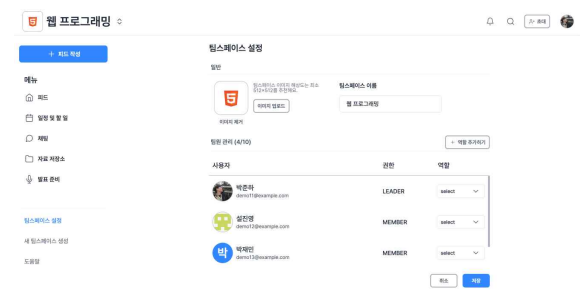


그림 8. 팀스페이스 설정 화면  
Fig. 8. TeamSpace Settings Screen

세 번째는 팀 채팅 기능이다. 그림 9와 같이 사용자가 채팅방에 입장하여 실시간으로 텍스트 메시지, 사진, 파일을 주고받을 수 있다. 실시간으로 활성화된 채팅방에 대한 정보들을 제공하며 팀 스페이스에 속한 사용자 모두 채팅방에 참가하여 대화를 나눌 수 있다. 채팅방 생성 버튼을 통해 원하는 채팅방 이름을 작성한 후 새로운 채팅방을 생성할 수 있다. 그림 10과 같이 팀 스페이스마다 현재 로그인한 사용자가 읽지 않은 전체 채팅 개수를 제공한다. 팀 스페이스, 사용자에게 따라 변경되는 전체 채팅 개수를 제공해 주며 이를 통해 현재 팀 스페이스에 자신이 읽지 않은 전체 채팅 개수를 파악할 수 있다.

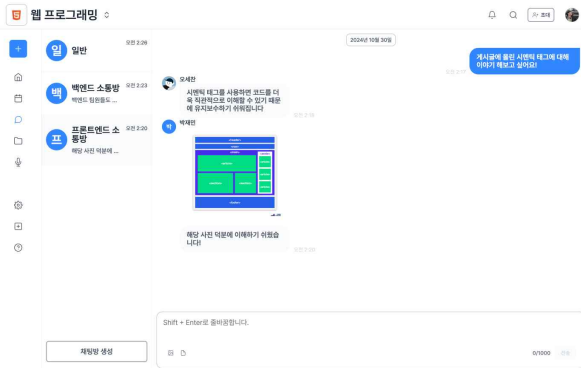


그림 9. 팀 채팅 화면  
Fig. 9. Team Chat Screen

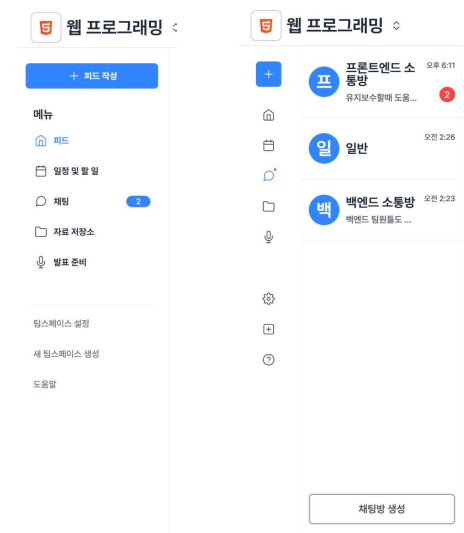


그림 10. 읽지 않은 채팅 수 표시 화면  
Fig. 10. Unread Chat Count Display Screen

네 번째는 목적별 피드 기능이다. 사용자는 그림 11의 피드 조회 화면에서 동일한 팀스페이스에 속한 팀원들이 작성한 피드들을 조회할 수 있다. 또한, 목적에 따라 필터링하여 특정 피드만 조회할 수 있다. 좌측 상단의 피드 작성 버튼을 클릭해 필요한 목적의 피드를 작성할 수 있다. 피드는 일반, 일정 조율, 자료 수집 세 가지로 나뉘며 목적에 따라 별도의 기능을 제공한다.

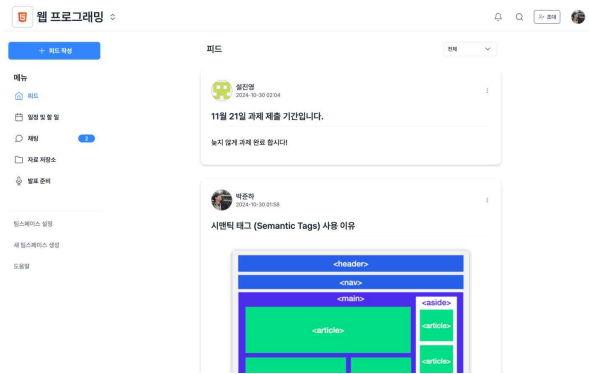


그림 11. 피드 조회 화면  
Fig. 11. Feed View Screen

일반 피드는 공지, 정보 공유 등을 위해 사용한다. 그림 12와 같이 제목, 글, 이미지를 WYSIWYG [28] 에디터를 통해 작성하고 파일을 첨부할 수 있다.

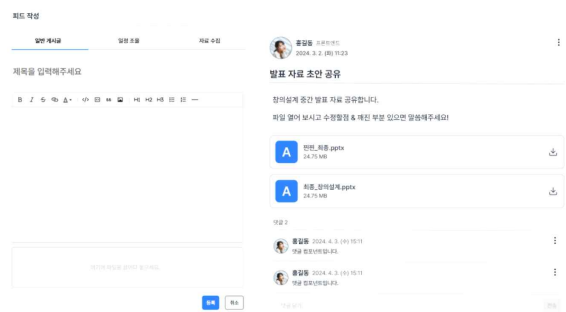


그림 12. 일반 피드 작성 및 조회 화면  
Fig. 12. Feed Creation and View Screen

일정 조율 피드는 팀원 간 일정 조율을 목적으로 사용한다. 작성자는 그림 13과 같이 일정 진행 날짜와 마감 일시 및 시간 범위를 지정할 수 있다. 팀원들은 그림 14와 같이 일정 추가 버튼을 클릭해 가용 시간을 작성할 수 있다. 각 팀원이 제출한 시간대는 차트에 시각적으로 표시되며 일정 조율에 참여한 팀원들을 확인할 수 있다. 시간대별로 가용 가능 팀원이 많을수록 강조되어 팀원들의 가용 시간대를 시각적으로 쉽게 파악하고 효율적인 일정 조율을 가능하게 한다.

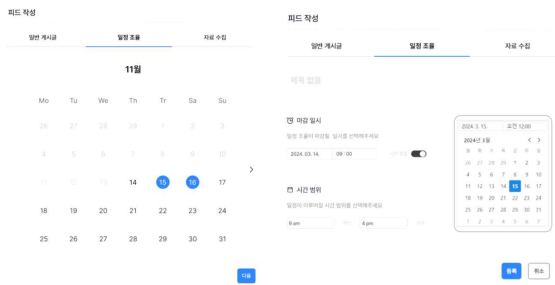


그림 13. 일정 조율 피드 작성 화면  
Fig. 13. Schedule Coordination Feed Creation Screen

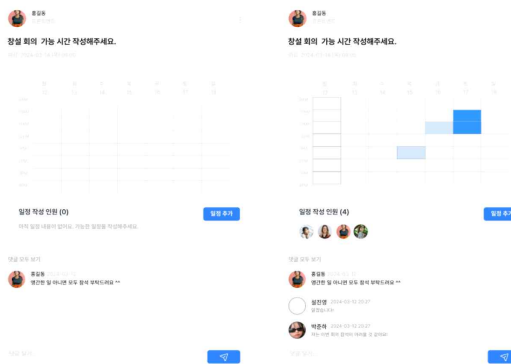


그림 14. 가용 시간 등록 및 확인 화면  
Fig. 14. Available Time Registration and Confirmation Screen

자료 수집 피드(그림 15)는 아이디어, 의견, 자료 수집을 목적으로 작성하는 피드이다. 작성자는 제목과 내용을 작성하고 자료 수집 참여 인원을 지정할 수 있다. 지정된 팀원은 자신에게 할당된 입력란을 통해 글을 작성하고 자료 수집 진행 상황 상태를 요청, 진행, 피드백, 완료 네 가지로 나눠 직관적으로 표시할 수 있다.

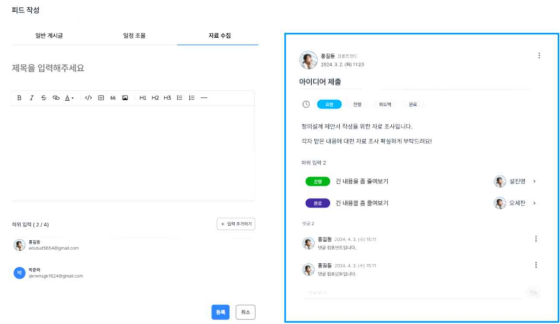


그림 15. 자료 수집 피드 작성 및 조회 화면  
Fig. 15. Data Collection Feed Creation and View Screen

모든 피드에는 댓글 작성 기능이 제공되어, 팀원들이 자유롭게 의견을 교환할 수 있다.

네 번째는 자료 저장소 기능이다. 팀 스페이스 내 분산되어 업로드된 모든 파일을 한 곳에서 효율적으로 관리하고 접근할 수 있다. 사용자는 그림 16과 같이 자료 저장소에서 팀 스페이스의 피드, 채팅을 통해 업로드 한 파일들을 조회할 수 있다. 또한, 파일 타입, 확장자, 등록자명으로 필터링해 특정 파일만을 조회할 수 있다. 사용자는 조회된 파일을 클릭해 다운로드 받을 수 있다.



그림 16. 자료 저장소 화면  
Fig. 16. Storage Screen

## IV. 결 론

본 논문은 대학생들이 조별 과제를 수행하는 과정에서 겪는 불편함을 해결하기 위해 개발된 협업 플랫폼 'Colla'를 제안한다. 본 플랫폼은 팀 프로젝트에 필요한 기능을 통합하여, 학생들이 더 효율적이고 체계적으로 프로젝트를 진행할 수 있도록 지원한다.

첫째, 본 플랫폼은 팀 생성, 파일 공유, 커뮤니케이션, 일정 조율 기능을 한 플랫폼에서 제공함으로써 여러 도구를 번갈아 사용하는 번거로움을 없앤다. 이를 통해 학생들은 한 곳에서 모든 팀 협업 작업을 효율적으로 처리할 수 있으며, 자료를 쉽게 공유하고 관리할 수 있다. 둘째, 자료 저장소 기능은 팀원들이 여러 채널을 통해 분산된 자료를 찾는 번거로움을 해결하고, 한곳에서 모든 자료를 관리할 수 있게 해준다. 이로 인해 파일을 중복 관리하거나 찾는 데 시간을 허비하지 않으며, 팀원들이 최신 자료에 쉽게 접근할 수 있어 협업 과정이 더욱 원활해진다. 셋째, 개인정보 보호 측면에서도 장점을 제공한다. 기존에는 조별 과제를 진행할 때 팀원 간 연락처를 주고받아 개인 정보를 공유해야 했으나, 본 플랫폼을 사용하면 이러한 절차가 필요 없다. 플랫폼 내에서 팀원들과 직접 소통할 수 있어, 불필요한 개인정보 노출을 방지하고 안전한 협업 환경을 제공한다.

향후 본 플랫폼 사용자 피드백을 바탕으로 지속적인 개선이 이루어질 것이며, 대학생들뿐만 아니라 다양한 협업 환경에서도 활용될 수 있는 폭넓은 플랫폼으로 발전할 가능성이 있다. 이를 통해 팀 프로젝트에 필수적인 도구로 자리매김할 수 있을 것이다. 따라서 본 플랫폼은 대학생들이 조별 과제를 효율적으로 수행할 수 있도록 여러 불편을 해소하고, 더 나은 협업 경험을 제공하는 데 크게 기여할 것으로 기대된다.

## 참 고 문 헌

[1] 20s Research Institute. "Thoughts of College Students in Their 20s on Group Projects." University Tomorrow Research Institute, 5 Dec. 2018. Available: <https://www.20slab.org/Archives/28766>.

[2] The Collaboration Tools Grown by Pandemic and Chosen by Digital, What's the Market Situation? Available: <https://jmagazine.joins.com/economist/view/333608>

[3] Notion. Notion Official Website [Internet]. Available: <https://www.notion.so>

[4] Flow. Flow Official Website [Internet]. Available: <https://flow.team>

[5] TypeScript. Documentation [Internet]. Available: <https://www.typescriptlang.org/docs/>

[6] VirtualDOM [Internet]. Available: <https://ko.legacyreactjs.org/docs/faq-internals.html>

[7] React. React Documentation [Internet]. Available: <https://react.dev/>

[8] TanStack. TanStack Query Documentation [Internet]. Available: <https://tanstack.com/query/latest/>

[9] Zustand [Internet]. Available: <https://zustand-demo.pmind.rs/>

[10] Styled Components. Documentation [Internet]. Available: <https://styled-components.com/docs>

[11] Storybook. Documentation [Internet]. Available: <https://storybook.js.org/docs/>

[12] Mock Service Worker. Documentation [Internet]. Available: <https://mswjs.io/docs/>

[13] Spring. Spring Data JPA Documentation [Internet]. Available: <https://docs.spring.io/spring-data/jpa/docs/current/reference/html/>

[14] Swagger. Documentation [Internet]. Available: <https://swagger.io/docs/>

[15] JUnit. User Guide [Internet]. Available: <https://junit.org/junit5/docs/current/user-guide/>

[16] MySQL. Documentation [Internet]. Available: <https://dev.mysql.com/doc/>

[17] Redis. Documentation [Internet]. Available: <https://redis.io/docs/>

[18] Amazon Web Services. AWS Documentation [Internet]. Available: <https://docs.aws.amazon.com/>

[19] Amazon CloudFront. Documentation [Internet]. Available: <https://docs.aws.amazon.com/cloudfront/>

[20] Nginx. Documentation [Internet]. Available: <https://nginx.org/en/docs/>

[21] Docker. Get Started [Internet]. Available: <https://docs.docker.com/get-started/>

[22] Docker Compose. Overview [Internet]. Available: <https://docs.docker.com/compose/>

[23] GitHub. GitHub Actions Documentation [Internet].

Available: <https://docs.github.com/en/actions>

[24] OAuth 2.0 [Internet]. Available: <https://oauth.net/2/>

[25] JSON Web Token [Internet]. Available: <https://jwt.io/>

[26] Axios Interceptor [Internet]. Available: <https://axios-http.com/kr/docs/interceptors>

[27] WebSocket [Internet]. Available: <https://developer.mozilla.org/ko/docs/Web/API/WebSocket>

[28] WYSIWYG [Internet]. Available: <https://ko.wikipedia.org/wiki/위지위그>



# 대형언어모델 기반 다중 에이전트 시스템을 사용한 자동 웹 크롤러 생성

곽지호\*, 김민근\*, 정유철\*\*

## Automated Web Crawler Generation Using LLM-Based Multi-Agent Systems

Jiho Gwak\*, Mingeun Kim\*, and Yuchul Jung\*\*

### 요약

데이터를 효율적으로 수집하기 위한 대표적인 도구로 크롤러가 있으며, 기존 크롤러의 구현을 위해서는 HTML 구조를 사람이 분석하는 과정이 필요하므로 이에 대한 추가 비용이 발생한다는 한계가 있다. 본 논문에서는 이러한 문제를 해결하기 위해 대형언어모델을 활용한 멀티 에이전트를 기반으로 하여 자동 크롤러 작성 기법을 제안하여, 사람의 개입 없이도 효율적인 데이터 수집의 가능성을 확인하였다.

### Abstract

Crawlers are typical tools for efficiently collecting data. Existing crawlers have the limitation that humans must analyze the HTML structure for implementation, which incurs additional costs. This paper proposes an automated crawler generation technology based on a multi-agent system utilizing LLM to solve this problem and confirms the possibility of efficiently collecting data without human intervention.

### Key words

multi agent system, automatic crawler, data collection, large language model

## 1. 서론

컴퓨터 시스템이 발전함에 따라 데이터는 시장 분석, 동향 분석 등에서 핵심 자원으로서 중요한 역할을 한다. 이러한 데이터를 효율적으로 수집하기 위한 도구로 웹 크롤러가 있으며, 이는 웹 페이지에서 데이터를 자동으로 수집하여 유용한 정보를 제

공한다. 그러나 웹의 데이터는 대부분 비정형 데이터로 존재[1]하여, 정형화된 정보를 얻기 위해 웹 페이지의 구조를 탐색하고 HTML 형식의 코드를 분석하는 등 사람의 개입이 필요하다는 한계가 있다. 대형 언어 모델 에이전트는 다양한 분야에서 활용되고 있으며[2], 최근에는 이를 통해 웹 콘텐츠를 분석하는 도구[3], 계층 구조를 이용해 점진적 이해

\* 국립금오공과대학교 컴퓨터공학과, wlg6709@kumoh.ac.kr, bbbababb@kumoh.ac.kr,,

\*\* 국립금오공과대학교 인공지능공학과, jyc@kumoh.ac.kr 교신저자

※ The work is supported by the Korea Basic Science Institute (Project No.A438200)

를 돕는 AutoCrawler 연구[4], 동적 및 정적 크롤러 병합 기술에 관한 연구[5]도 진행되었다. 본 연구에서는 정형 데이터를 효과적으로 수집하기 위해 대형언어모델 멀티 에이전트 시스템을 적용해 동적 및 정적을 결합한 웹 크롤러 개발 기법을 제안한다. 이 기법은 웹 페이지 구조 인식, 크롤러 코드 작성, 디버깅 작업을 사람의 개입 없이 에이전트가 반복적으로 수행하여, 적은 비용으로 정형화된 데이터를 추출하고, 수집할 수 있는 것을 보였다.

## II. 에이전트 기반 크롤러 코드 생성

본 논문에서는 실험을 위해 GPT-4o-mini API를 에이전트로 활용하였다. 에이전트의 상태는 State function을 통해 관리하며, 이는 대화 기록, 프로세스 상태, 요청 횟수 등의 정보가 포함한다. 에이전트는 크롤링 코드를 구현하는 development 에이전트와 해당 코드를 실행하는 executor 에이전트로 구성된다. development 에이전트는 동적 에이전트와 정적 에이전트로 나뉘며, 각각 크롤링 방식에 따라 구현된다. 그림 1은 전체 프로세스의 구성을 나타낸 것이다.

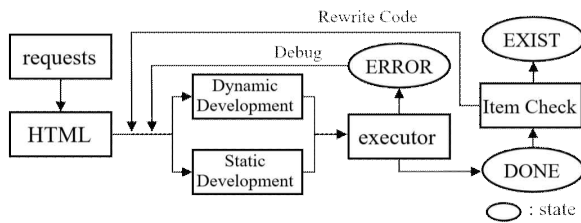


그림 1. 제안 기법 프로세스  
Fig. 1. Process of the proposed technique

사용자가 수집을 원하는 사이트의 URL을 입력하면, requests 요청을 통해 HTML 형식의 사이트 정보를 응답받고, 이를 각 Development 에이전트에 제공한다. 동적 에이전트는 Selenium을 기반으로 하여 동적으로 데이터가 있는 페이지를 탐색하며, 정적 에이전트는 BeautifulSoup를 활용하여 정적 페이지 분석을 통해 구조화된 데이터를 추출하는 역할을 한다. 에이전트는 실행될 수 있는 python code를 생성하여 해당 코드는 executor 에이전트를 통해 실행

된다. 실행 과정에서 오류가 발생하면 로그와 코드를 다시 제시하여 디버깅을 진행하고, 실행이 정상적으로 완료되면 작업을 마친다. 이후 수집된 Item을 확인한 후 Item이 부족하거나 잘못되면 추가 개발을 요구하고, 충분한 경우 프로세스를 종료하게 된다.

## III. 평가

평가는 nserc([https://www.nserc-crsng.gc.ca/ase-oro/index\\_eng.asp](https://www.nserc-crsng.gc.ca/ase-oro/index_eng.asp)), nsf(<https://www.nsf.gov/awardsearch/>), 금오공과대학교 게시판([https://www.kumoh.ac.kr/ko/sub06\\_01\\_01\\_01.do](https://www.kumoh.ac.kr/ko/sub06_01_01_01.do)), 네이버 지식인(<https://kin.naver.com/qna/list.naver>)네 개 사이트에 대해 항목(Item)을 올바르게 수집했는지에 대해 진행되었으며, 수작업으로 작성한 데이터를 정답으로 두고 두 결과를 비교하였다. 평가 항목은 수집된 Item의 수와 정보의 정확성 두 가지로 구성하였으며, 각각 1, 3회 반복 작업에 따른 결과 비교를 진행하였다.

표 1. 평가 결과  
Table 1. Evaluation results

사이트	제안 기법의 수집 Item 수		정확히 수집한 Item 수		실제 Item 수
	1	3	1	3	
반복	1	3	1	3	
nserc	7	16	6	15	16
nsf	12	24	12	24	28
본교 게시판	7	7	7	7	7
지식인	1	1	1	1	4

위 결과와 같이, 제안한 기법이 높은 수집률과 정확도를 보였으며, 반복 과정을 수행한 후 수집률과 정확도가 더욱 향상됨을 확인할 수 있었다. 사이트의 형태가 어느 정도 구조화된 경우는 정보 대부분을 수집할 수 있는 코드가 작성되었으나, 주요 정보의 판단이 어려운 지식인 같은 경우는 코드 작성에 한계를 확인할 수 있었다. 또한, 본 기법은 1회 평균 0.0009\$에서 3회 평균 0.0013\$의 매우 적은 비용으로 크롤러를 구축할 수 있음을 확인하였다.

#### IV. 결 론

본 논문에서는 멀티 에이전트 기반으로 한 크롤러 코드 작성 기법을 제안하였다. 제안된 기법은 에이전트를 기반으로 하여 사람의 개입 없이 높은 수집률을 가지는 크롤러를 효율적으로 구축할 수 있음을 보였다. 그러나 사이트 구조가 복잡한 경우 데이터 수집을 위한 코드 작성에 한계가 존재하며, 이러한 문제는 추후 연구에서 다룰 예정이다.

#### 참 고 문 헌

- [1] [https://en.wikipedia.org/wiki/Unstructured\\_data](https://en.wikipedia.org/wiki/Unstructured_data) [Accessed: Oct. 21, 2024]
- [2] A. Dorri, S. S. Kanhere, and R. Jurdak, "Multi-Agent Systems: A Survey," IEEE Access, vol. 6, pp. 28573-28593, 2018. doi: 10.1109/ACCESS.2018.2831228
- [3] <https://dev.to/jadouse5/building-an-openai-swarm-web-scraping-and-content-analysis-streamlit-web-app-with-multi-agent-systems-bl0> [Accessed: Oct 21, 2024]
- [4] W. Huang, Z. Gu, C. Peng, Z. Li, J. Liang, Y. Xiao, L. Wen, and Z. Chen, "AutoScraper: A Progressive Understanding Web Agent for Web Scraper Generation," arXiv preprint, 2024. Available: <http://arxiv.org/abs/2404.12753>
- [5] Q. Liu, R. Yahyapour, H. Liu, and Y. Hu, "A novel combining method of dynamic and static web crawler with parallel computing," Multimedia Tools and Applications, vol. 83, no. 21, pp. 60343-60364, 2024. doi: 10.1007/s11042-023-17925-y
- [6] <https://platform.openai.com/docs/models/gpt-4o-mini> [Accessed: Oct. 22, 2024]

# 다양한 학습자를 위한 몰입형 VR 기반 수학 교육 어플리케이션 개발

강다운\*, Azwanina Shahizan\*\*, 박지윤\*\*\*

## Development of an Immersive VR-Based Mathematics Education Application for Diverse Learners

D. E. Kang\*, A. Shahizan\*\* and J. Park\*\*\*

### 요 약

테크놀로지의 발전은 교육 현장에 많은 변화를 가져왔으며, 특히 교육용 어플리케이션은 수학 학습에 긍정적인 영향을 미치는 사례로 주목받고 있다. 기존 연구에서는 수학 교육용 어플리케이션을 활용했을 때 학생들의 수학 학업성취도와 태도에 긍정적인 효과가 나타났다고 보고된 바 있다. 그러나 장애학생의 경우, 교육용 어플리케이션의 접근성이 낮아 그 효과를 기대하기 어려운 실정이다. 이에 본 연구는 다양한 학습자의 요구를 반영하고, 효과적인 수학 학습 경험을 제공할 수 있는 새로운 어플리케이션을 제안하고자 한다. 본 연구에서 제안하는 어플리케이션은 수세기 개념을 학습할 수 있도록 설계되었으며, 스토리텔링을 통한 흥미 유발과 폭넓은 선택지를 제공함으로써 학습자의 참여를 높이는 데 중점을 두었다. 또한 VR 기술을 접목하여 현실감 있는 학습 환경을 구현함으로써 수학적 개념 학습에 효과적일 것으로 기대된다.

### Abstract

As technology use in education continues to increase, educational applications have emerged as a powerful tool for supporting math learning. Prior studies have shown that the usage of these applications resulted in improvement of academic performance and positive attitudes toward math, among users. However, the same cannot be said for learners with disabilities, as the benefits of educational applications is often limited to them due to accessibility challenges. Therefore, the current study aims to propose a new mathematics educational application that is specially designed to meet the diverse needs of learners, thus providing a more inclusive learning experience. The application proposed in the current study focuses on foundational math concepts like numbers and counting, and aims to increase learners engagement through methods such as storytelling and offering users a wide range of choices. By integrating virtual reality (VR) technology, this new application also seeks to create a realistic learning environment for the users which is expected to support effective mathematical learning.

### Key words

mathematics education, educational applications, diverse learners, VR

---

\* 국립창원대학교 특수교육과, 20220185@gs.cwnu.ac.kr

\*\* 국립창원대학교 특수교육과 박사 과정, azwanina@gmail.com

\*\*\* 국립창원대학교 특수교육과 교수, jpark@changwon.ac.kr(교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.

## I. 서 론

테크놀로지는 실생활과 융합되면서 교육 분야에서도 그 활용도가 높아지고 있다. 교육용 어플리케이션은 교육적 목적과 내용을 담아 국어, 수학, 영어 등 다양한 교과 영역에서 활용되고 있으며, 특히 수학은 학습자가 실생활 문제를 해결하고 지역사회에서 자립적으로 살아가는 데 필수적인 학문이다 [1]. 김태화, 최재호(2021)에 따르면, 수학 교육용 앱을 활용한 수업은 학생들의 수학 학업성취도와 수학적 태도에 긍정적인 영향을 미치는 것으로 나타났다[2]. 이는 교육용 어플리케이션이 학습 효과를 높이고 학생의 태도를 개선하는 효과적인 교육 도구임을 시사한다.

다만, 교육용 어플리케이션은 테크놀로지로의 접근이 낮은 학생들에게는 사용이 어려운 도구라는 단점이 있다. 즉, 교육용 어플리케이션의 활용도가 낮은 집단은 교육용 어플리케이션을 통한 교육 효과를 기대할 수 없다. 특히 장애학생의 경우, 개별적 특성으로 인해 이러한 교육 도구를 충분히 활용하지 못하는 경우가 있다. 대부분의 교육용 어플리케이션이 다양한 학습자의 요구를 반영하지 못하여 장애학생의 접근성이 제한되고 있다. 따라서 다양한 학습자의 요구를 고려한 교육용 어플리케이션을 개발하여 장애학생의 학습 기회를 확대한다면, 이들의 수학 학습에 실질적인 도움을 줄 수 있을 것으로 기대된다.

## II. 새로운 어플리케이션에 대한 논의

새로운 어플리케이션을 도출할 때의 주안점은 다음과 같다.

1. 다양한 학습자를 고려할 수 있는가?
2. 스토리텔링을 통한 사용자의 흥미 유발과 적절한 피드백, 사용자의 폭넓은 선택지를 제공할 수 있는가?

위의 두 가지 주안점을 바탕으로 새로운 수학 교육용 어플리케이션을 고안하였다. 본 어플리케이션의 가장 큰 특징은 VR(Virtual Reality) 기술을 접목하였다는 점이다. VR 기술을 접목한 이유는 다음과

같다. 첫째, VR 기술은 시각적, 청각적, 신체적 자극을 제공하여 다양한 학습 유형을 지원할 수 있기 때문이다. 이는 시각적 학습자, 청각적 학습자, 신체적 학습자를 모두 고려할 수 있어 학습자의 특성에 따른 맞춤형 자극 제공이 가능하다. 이를 통해 학습자는 자신에게 맞는 자극을 받아들여 학습 효과를 극대화 할 수 있다.

둘째, VR은 높은 현실감을 제공하기 때문이다. 생활연령은 높지만 정신연령이 낮은 장애학생을 교육할 때에는 생활연령에 맞는 교구가 요구된다. 예를 들어, 중학교 3학년인 장애학생이 초등학교 2학년 수준의 수 세기를 개념을 학습할 때 아동용 캐릭터가 그려진 숫자블록을 사용하는 것은 생활연령에 맞지 않아 부적절할 수 있다. 따라서 장애학생의 생활연령을 고려하여 현실성이 높은 자료를 제공해야 할 것이다. VR은 장애학생의 생활연령에 부합하는 현실성 높은 자료를 제공할 수 있어, 장애학생의 학습 요구에 적합한 환경을 조성하는 데 유리하다. 안희두, 서만호, 이순천, 정희경(2018)의 연구에서도 VR의 실재감이 학습자의 흥미와 관심을 높이고 교육 효과에 긍정적인 영향을 미쳤다고 보고된 바 있다[3].

셋째, VR 기술은 몰입감을 높일 수 있는 요소를 제공한다. 여러 선행연구에서 VR를 통한 높은 몰입도가 학습 효과를 증대시키는 것으로 나타났다. 예를 들어, 양성광, 김재웅(2022)은 VR 기반 교육 게임을 통해 학생의 몰입도를 높이고 학습 효과에 긍정적인 영향을 미칠 수 있었다고 보고하였다[4]. VR을 통한 몰입감 증대는 스토리텔링을 통한 흥미 유발을 지원하며, 학습자의 참여도를 더욱 향상시킬 수 있다.

본 어플리케이션의 목적은 다양한 학습자를 고려하여 수학 개념을 학습하고 문제를 해결하는 경험을 제공하는 데 있다. 수학 교과는 다양한 영역(예: 수와 연산, 도형, 측정 등)과 수준으로 구성되어 있으나, 본 어플리케이션은 초기 개발 단계에서 수세기 개념에 집중하고자 한다. 따라서 다양한 학습자를 고려한 수세기 교육용 어플리케이션을 목표로 하고 있으며, 이로써 학습자들이 수학의 기초 개념을 효과적으로 익힐 수 있을 것으로 기대된다. 전체

적인 어플리케이션의 시나리오는 다음과 같다.

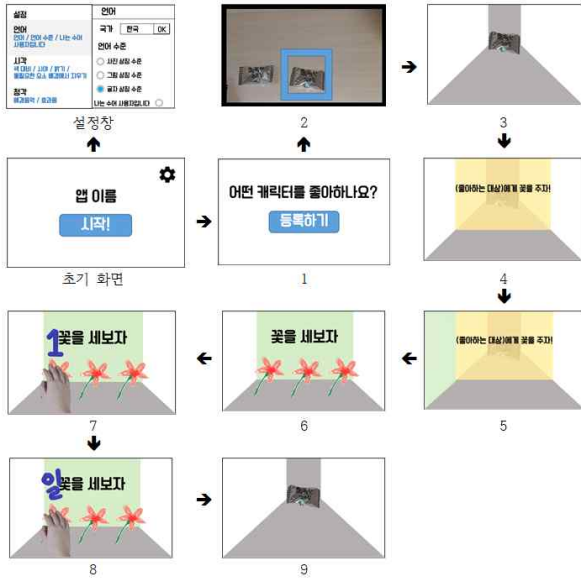


그림 490. 새로운 어플리케이션의 시나리오  
Fig. 1. Scenarios for new applications

그림 1의 설정창을 보면, 이 앱에서는 언어(국가, 언어 수준, 수어 사용 여부), 시각(색대비, 시야, 밝기, 불필요한 요소 배경에서 지우기), 청각(배경음악, 효과음)을 사용자의 특성에 맞게 변경할 수 있도록 하였다. 이를 통해 장애 유형과 정도가 다르더라도 이 어플리케이션을 통해 학습할 수 있도록 고려하였다.

다음으로 그림 1의 1에서 ‘등록하기’ 버튼을 누르면 사용자가 좋아하는 대상의 사진을 등록할 수 있게 된다. 등록이 완료되면 VR이 실행되고 그림 1의 3과 같이 등록된 좋아하는 대상이 길 끝에서 보인다. 이렇게 사용자가 선호하는 대상을 향해 이동해야 한다는 스토리텔링을 통해 사용자의 동기와 흥미를 유발하고자 하였다.

사용자는 자유롭게 이동할 수 있으며, 좋아하는 대상에게 가까이 다가갈 수 있다. 그러나 이동 중 그림 1의 4처럼 벽이 길을 막고 더 이상 이동할 수 없게 되면, 사용자를 막고 있는 벽에 적힌 지시사항과 음성 안내를 통해 사용자가 좌측을 보도록 유도한다.

사용자의 좌측에는 그림 1의 6처럼 지시가 나타난다. ‘꽃을 세 보자’라는 지시에 따라 꽃을 잡으면 그림 1의 7과 같이 ‘1’ 모형과 ‘일’이라는 음성이

출력된 후, 그림 1의 8과 ‘일’이라는 모형이 나타난다. 이러한 방식으로 꽃의 개수인 3까지 센다면 정답이 맞았음을 알리는 효과음이 발생한다. 이후, 사용자를 가로막았던 벽이 사라지며 다시 대상에게 접근할 수 있게 된다.

이러한 이동의 제한과 지시의 제공 과정(그림 1의 3~그림1의 9의 과정)을 5회 반복하며, 사용자는 조금씩 좋아하는 대상에게 접근할 수 있다. 결국 마지막에는 사용자가 좋아하는 대상에 완전히 도달하여 대상을 얻을 수 있다.

5회 반복되는 동안 지시는 매번 달라진다. 표 1에 지시와 사용자 반응에 대한 예시를 제시하였다.

표 1. 지시와 사용자 반응의 예시

Table 1. Examples of instructions and user responses

순서	지시 / 제시물	요구되는 사용자 반응
1	꽃을 세 보자 / 꽃 3개	꽃 하나씩 잡기
2	꽃을 세 보자 / 꽃 5개	꽃 하나씩 잡기
3	사탕을 세 보자 / 사탕 3개	사탕 하나씩 잡기
4	꽃을 세 보자, 직접 세보아요 / 꽃 3개	꽃 하나씩 잡으면서 수 세기
5	사탕을 세 보자, 직접 세보아요 / 사탕 5개	사탕 하나씩 잡으면서 수 세기
6	-	대상을 향해 이동하기

다음 예시와 같이 5회 반복하는 과정에서 지시는 다양하게 변화하여, 사용자는 여러 상황에서 수세기의 원리가 적용되는 방식을 학습할 수 있다. 이를 통해 개념의 일반화가 가능해진다.

### III. 어플리케이션 개발 결과

앞서 논의한 내용을 바탕으로 어플리케이션을 개발하였다. 새로운 어플리케이션을 통해 수세기를 학습하는 과정은 다음과 같다.

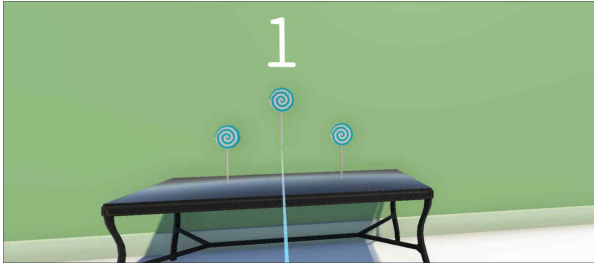


그림 491. 어플리케이션 실행 모습  
Fig. 2. App execution screen

그림 2는 어플리케이션의 실행 모습을 보여준다. 사용자가 물건을 잡으면, 그림 2와 같이 '1' 모양이 생성되고 동시에 '일'이라는 음성이 들린다. 이후 '일'이라는 모양이 화면에 나타나며, 사용자는 물건 1개에 숫자 '1'과 음성 '일'이 연결되는 과정을 통해 숫자 개념을 학습하게 된다. 이후 물건을 또 하나 잡으면 '2'라는 숫자와 음성 '이'가 등장하며, 이를 통해 '1' 다음이 '2'임을 학습할 수 있다. 이처럼 시각적, 청각적, 신체적 자극을 통해 수세기 개념을 효과적으로 학습할 수 있도록 설계하였다.

#### IV. 결 론

본 논문에서는 수학 교육용 어플리케이션이 수학 학업 성취도와 수학에 대한 긍정적인 태도 향상에 기여할 수 있음을 바탕으로, 장애학생의 활용도를 높이기 위한 새로운 어플리케이션을 개발하고자 하였다. 제안된 어플리케이션은 VR를 접목하여 시각적, 청각적, 신체적 자극을 제공하고, 높은 현실감과 몰입 효과를 통해 사용자의 동기를 유도할 수 있다는 장점이 있다.

개발된 어플리케이션은 사용자가 선호하는 대상을 등록하여 동기와 흥미를 유발하고, 다양한 학습자가 수세기를 흥미롭게 학습할 수 있도록 한다. 그러나 본 연구에서는 장애학생을 대상으로 어플리케이션을 직접 적용해 본 결과는 다루지 않았다는 한계가 있다. 후속 연구에서는 이러한 점을 보완하여 실제 적용 사례를 통해 효과를 검증할 필요가 있다. 또한 본 논문에서는 수학 교육 내용 중 수세기만 다루었다. 장애학생에게는 사칙연산과 같은 수세기 외의 기능도 중요하므로, 추후 연구에서는 다양한

수학적 개념을 포함한 어플리케이션으로 확장할 필요가 있다.

#### 참 고 문 헌

- [1] Ministry of Education, 2022 Revised Basic Education Curriculum, Ministry of Education Notice No. 2022-34. 135, 2022
- [2] T. H. Kim and J. H. Choi, The Effect of Mathematics Education App on Mathematics Academic Achievement and Mathematical Attitude, The Research of Science Mathematics Education, vol. 44, pp. 67-83, 2021
- [3] H. D. An, M. H. Seo, S. C. Lee and H. K. Jung, Study for the Presence and The Interaction in the VR Solar System Education Contents, Proceedings of HCI Korea 2018, pp. 903-906, 2018
- [4] S. G. Yang and J. W. Kim, A Study on the Impact Mechanism of Student Immersion on the Learning Effect of Virtual Reality Educational Games, The Journal of Humanities and Social Science, vol. 13, pp.3987-4002, 2022

# 우울증 완화를 위한 인터랙티브 무비 콘텐츠

오경문\*, 최현빈\*\*, 유선진\*\*\*

## Interactive Movie Contents for Relief Depression

Gyungmoon O\*, Hyunbin Choi\*\*, and Sunjin Yu\*\*\*

### 요 약

현대사회에서 우울증은 뇌졸중이나 공황장애 등 정신적 혹은 신체적 질환으로 이어질 수도 있는 큰 문제이다. 우울증은 실제로 많은 사람들이 겪는 질환이지만 정신 질환의 부정적인 인식으로 적극적으로 치료에 임하는 경우는 소수에 불과하다. 이 논문에서 다루는 콘텐츠는 현대인들의 모습을 투영한 가상의 '민수'라는 인물을 통해 주인공과 동질감을 느끼고 공감을 통하여 이런 어려움을 겪는 것이 나 혼자가 아니라는 생각을 통해 우울증 치료에 도움이 되는 인터랙티브 무비이다.

### Abstract

In modern society, depression is a big problem that can lead to mental or physical diseases such as stroke and panic disorder. Depression is actually a disease experienced by many people, but only a few people actively engage in treatment due to negative perceptions of mental illness. The content covered in this paper is an interactive movie that helps treat depression through the idea that I am not alone in experiencing these difficulties through empathy and feeling the same sense with the protagonist through "Min-soo" who projects the appearance of modern people.

### Key words

interactive movie , depression , empathy , mental disease , remedy

## 1. 서 론

우리나라의 우울감 경험률은 13.2%로 일본 9.3%, 미국 9.0%, 덴마크 7.8% 등 다른 나라에 비해 높다. 증가하는 속도 또한 매우 빨라서 2015년에는 62만 3,938명이었지만 2019년에는 82만6,834명으로 4년간 30만명이 넘게 증가하여 우울증이 우리나라의 문제

임을 알 수 있다[1].

한국 사회에서 많은 사람들이 지나친 경쟁 속에 살고 있다. 경쟁에서의 실패나 지나친 경쟁의 연속은 우울증이나 무기력증을 유발하게 되며 치료 또한 어렵다.

우울증은 실제로 많은 사람들이 겪는 질환이지만 정신 질환의 부정적인 인식으로 적극적으

\* 국립창원대학교 문화테크노학과, 20234621@gs.cwnu.ac.kr,

\*\* 국립창원대학교 문화융합기술협동과정 석사 과정, 20217292@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수 sjyu@changwon.ac.kr, (교신 저자)

※ "이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다."



로 치료에 임하는 경우는 소수에 불과하다. 우울증을 앓고 있더라도 반응하지 않은 환자의 비율이 30%에 달하는 정도로 우울증에 대한 사람들의 이해가 부족하고 치료를 위해 정신과 전문의와 상담을 통해 우울증을 치료하거나 비용대비 효과적인 측면을 고려하여 항우울제를 투약하며 우울증을 치료하는 방법도 고려되지만, 결국 우울증을 치료하기 위해서는 병원 방문에 따라 학생의 경우는 방학이 아닌 경우에는 학업을 중단하거나 결석을 할 수 밖에 없게 되고 직장인의 경우에는 결근일 수 발생, 일의 능률 또한 감소함에 따라서 적극적으로 치료에 임할 수 없는 환경이 된다.

이러한 경제적, 환경적 요소 이외에도 사회적 인식으로 인하여 본인의 정신 질환을 숨기는 경우가 많다. 또한, 우울감을 경험한 잠재적 우울증 환자들이 치료를 필요로 하는 의학적 상태라는 것에 대한 인식이 부족하다. 보통의 우울증은 적절한 치료나 도움이 있다면 비교적 쉽게 치료할 수 있지만 장기화 된다면 치료가 점점 어려워지는 문제점이 있다[2].

## II. 인터랙티브 무비

본 논문에서는 이러한 우울증 완화를 위한 인터랙티브 무비를 구현하기 위하여 Unity를 사용한다. 개발 환경은 PC이며. 콘솔 컨트롤러 혹은 마우스를 사용하여 주인공의 행동을 선택하는 방식이며 여러 편의 영상을 여러 가지의 스테이지로 나눠놓은 방식으로 구성되어있지만 플레이하는 사람이 느끼기에는 선택지의 선택을 통한 이야기의 전개로 한 편의 영화처럼 이어지는 느낌이며 플레이 하는 사람의 선택에 따라서 해피엔딩이 될 수도 배드엔딩이 될 수도 있다.

이 인터랙티브 무비는 대부분의 현대인들이 겪는 문제점을 겪고 있는 ‘민수’ 라는 주인공을 통하여 청년들이 자신의 모습을 투영하는 주인공에게서 자아를 인식하며 주인공과의 동감을 통하여 이런 어려움을 겪는 것이 나 혼자가 아니라는 생각을 하게 되면서 자연스럽게 아픔과 스트레스를 덜어내고 희망을 느끼고 소외감을 떨쳐낼 수 있는 기대효과를

찾는다.



그림 1. 인터랙티브 무비의 한 장면  
Fig. 1. A Scene of Interactive Movie

## III. 결 론

본 논문에서 제안한 인터랙티브 무비는 단순하게 영화를 감상하는 것이 전부인 일반적인 영화의 형태를 벗어나 게임을 통한 몰입감 증가, 주인공과 함께한다는 유대감 증가와 자신이 직접 할 수 있다는 행복감 증진과 자신감 또한 얻을 수 있다[4][5].

## 참 고 문 헌

- [1] 신민섭, 박광배, 오경자, “우울증과 충동성이 청소년들의 자살 행위에 미치는 영향,” Korean Journal of Clinical Psychology, Vol.10, No.1, pp.286-297, 1991.
- [2] A. Shrivastava, M. Johnston, and Y. Bureau, "Stigma of mental illness-I: Clinical reflections," Mens sana monographs, Vol.10, No.1, p.70, 2012
- [3] 김종로. (2010). 기획된 영화 감상이 행복감 증진에 미치는 영향. 인문과학연구, 25, 369-388.
- [4] 하승완, 김주섭. (2018-01-31). 정서적 유대감을 높여주는 인터랙티브 사진. 한국HCI학회 학술대회, 강원.
- [5] 신경아. (2018). 우울증 중년여성을 위한 영화치료 프로그램의 효과. 디지털융복합연구, 16(10), 511-522.

# RFID 모듈과 Unity 엔진을 통한 음악 정서 체험 콘텐츠

안은채\*, 부서희\*, 이세진\*\*, 유선진\*\*\*

## Music emotional experience content with RFID modules and the Unity engine

Eunchae An\*, Seohee Boo\*, Sejin Lee\*\*, and Sunjin Yu\*\*

### 요 약

우리는 종종 음악을 단순히 듣고 즐기는 청각적 경험으로만 인식하지만, 사실 음악은 우리 생활과 밀접하게 연관되어 있으며 많은 부분 정서적 반응을 동반한다. 본 연구에서는 시대별 대중가요 영상을 감상할 수 있도록 RFID와 윈도우 폼을 활용하였다. 또한 개인 성향 분석을 통한 맞춤형 음악 장르 추천 테스트를 개발하기 위해 Unity를 사용해 제작하였다. 그 결과 체험자는 음악을 통해 공감대를 형성하고 자기 표현의 기회를 얻을 수 있다.

### Abstract

We often perceive music as an auditory experience that we simply listen to and enjoy, but it is closely connected to our lives and is often accompanied by emotional responses. In this study, we utilized RFID and window forms to allow users to watch videos of popular songs from different eras. We also used Unity to develop a personalized music genre recommendation test based on personality analysis. As a result, users are able to connect with music and find opportunities for self-expression.

### Key words

ardunio, rfid, unity, windows forms, music

## 1. 서 론

우리는 종종 음악을 단순히 듣고 즐기는 청각적 경험으로만 인식하지만, 사실 음악은 우리 생활과 밀접하게 연관되어 있으며 많은 부분 정서적 반응을 동반한다[1]. 음악은 감정을 표현하며 미적인 즐

거움을 제공하고 기분 전환이나 여흥을 위해서 사용되기도 하며, 정서 및 감정을 전달한다고 한다[2]. 다시 말해, 음악은 예술의 경계를 넘어 우리의 감정과 정체성을 표현하는 강력한 수단이다. 이러한 음악의 정서적 역할을 이해하기 위해 본 연구에서는 아두이노 RFID(Radio-Frequency Identification) 기술과

\* 국립창원대학교 문화테크노학과 학부과정, 20214515@gs.cwnu.ac.kr, 20214513@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화테크노학과 석사과정, 20237203@gs.changwon.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr (교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.

Unity 엔진을 활용하여 음악의 기능 중 추억 소환, 공감 생성, 그리고 자기 표현을 조명하고자 한다.

## II. 아두이노 활용 영상 시청 콘텐츠

아두이노 RFID 모듈과 윈도우 폼(Windows Forms)을 활용하여, 관람객이 시대별 대중가요 영상을 감상할 수 있도록 체험형 인터랙티브 콘텐츠를 제작한다. 관람객이 특정 오브젝트를 리더기 위에 올려두면, 오브젝트 하단에 부착된 RFID 태그가 인식된다. 이때 시리얼 통신을 통해 윈도우 폼으로 신호를 보내면 UID에 할당된 영상이 자동으로 재생된다. 관람객은 시대별로 제작된 오브젝트들을 통해 시대별 영상을 감상할 수 있다. 영상은 총 5개로 구성되며, 각 영상은 특정 시대의 대중가요와 그 시대의 감성을 반영하는 요소들을 담고 있다.



그림 1. 아두이노 RFID 구현 장면  
Fig. 1. Arduino RFID Implementation Scene

그림 1은 상징물 하단에 부착된 RFID 태그가 인식되어 영상이 송출되는 과정을 보여준다. 이를 통해 관람객이 각 시대의 대중가요를 통해 추억을 떠올리며, 음악을 매개로 한 정서적 공감을 경험할 수 있다.

## III. Unity 활용 음악 취향 분석 테스트

본 콘텐츠는 Unity 엔진을 사용해 음악 취향 분석 테스트로 제작되었으며, 태블릿과 PC 환경에서도 실행할 수 있다. 그림 2는 Unity 엔진으로 개발된 음악 취향 분석 테스트 화면을 보여준다. 테스트는 음악 취향과 관련된 10개의 질문으로 구성되어 있으며, 각 질문에는 두 개의 선택지가 주어진다. 선택한 답변에 따라 특정 음악 장르의 점수가 올라가며, 모든 질문에 답하면 “당신의 취향은 [최고 점

수를 받은 장르]입니다”라는 결과가 표시된다.



그림 2. Unity 엔진을 활용한 음악 취향 분석 테스트  
Fig. 2. Testing music taste analysis with the Unity Engine

이후, ‘플레이리스트 확인하기’ 버튼을 통해 8개의 장르 중 개인 취향에 맞는 장르를 선택하면, 해당 장르의 랜덤 플레이리스트가 생성된다. 체험자가 생성된 플레이리스트를 소장할 수 있도록, 버튼에 ScreenCapture 기능을 연결하였다. 버튼을 누르면 화면이 저장되며, 이후 배치된 프린터를 통해 플레이리스트를 인쇄함으로써 콘텐츠 체험이 마무리된다.

## IV. 결 론

본 연구에서는 Arduino RFID 기술과 Unity 엔진을 활용하여 음악의 정서적 역할을 체험하는 콘텐츠를 제안하였다. 그 결과 체험자는 음악을 통해 공감대를 형성하고 자기 표현의 기회를 얻을 수 있다. 향후 연구에서는 더욱 사용자 맞춤형 요소를 강화하여 인터랙티브 콘텐츠를 기반으로 한 개인 맞춤형 음악 체험 콘텐츠로 발전시킬 계획이다.

## 참 고 문 헌

- [1] J. Y. Lee, "The emotional regulation function of music appreciation," Ph.D. dissertation, Graduate School, Seoul National University, Seoul, South Korea, 2015.
- [2] E. Y. Hwang, "The Study on the General Public Musical Life, Music Function and Emotional Regulation Strategies as Psychological Healing Using Music", Journal of Arts Psychotherapy, vol.17, no. 3, pp.237-265, 2021.



# 블록체인을 이용한 아파트 하자 관리 시스템

이용탁\*, 김건우\*\*, 손기민\*\*\*, 이창준\*\*\*\*, 신연순\*\*\*\*\*

## Blockchain-Based Defect Management System for Apartments

Yong-Tak Lee, Gun-Woo Kim, Ki-Min Son, Chang-Joon Lee and Youn-Soon Shin

### 요 약

본 논문은 온라인 시스템으로 아파트 시공사와 입주민 간의 하자를 관리하는 데 있어 데이터 무결성을 보장하기 위해 블록체인 기술을 적용한 시스템의 설계 및 구현 과정을 다룬다. 제안하는 시스템은 SHA-256 해시 함수와 머클 트리 구조를 통해 시공사와 입주민이 상호 인정한 하자 정보를 효율적으로 저장 및 검증하며, 모든 노드가 독립적으로 유효성을 확인함으로써 시스템의 신뢰성을 강화하고 분쟁을 최소화한다. 본 연구의 결과는 블록체인 기술이 아파트 하자 관리에 적용될 때 높은 수준의 데이터 무결성과 보안을 제공할 수 있음을 확인한다.

### Abstract

This paper addresses the design and implementation of a blockchain-based system to ensure data integrity in apartment defect management. The system leverages SHA-256 hashing and Merkle tree structures to efficiently verify defect information, while independent validation by each node enhances system reliability. Additionally, the P2P network architecture maintains data consistency and security in a distributed environment. The results of this study confirm that blockchain technology can provide a high level of data integrity and security when applied to apartment defect management.

### Key words

Blockchain, Apartment defect management, Validity verification

---

\* 동국대학교 공과대학 컴퓨터정보통신공학부 컴퓨터공학전공, ytleea98@gmail.com,

\*\* gun0005@naver.com

\*\*\* kimin0331@naver.com

\*\*\*\* 동국대학교 불교대학 불교학부, lchj1999@gmail.com

\*\*\*\*\* 동국대학교 AI소프트웨어융합학부, ysshin@dongguk.edu(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음”(2023-0-00049)

## I. 서 론

아파트 건설 및 관리에서 하자 관리는 중대한 문제이다. 기존의 아파트 하자 관리 시스템은 종이 기반 또는 중앙집중형 데이터베이스를 사용하는 경우가 많아 데이터 손실 또는 변조의 위험이 존재한다. “채들” [1]과 같은 솔루션은 기존에 지류로 관리되던 하자 정보를 전산화하고, 하자를 발견하고 등록하는 프로세스를 디지털화하여 관리 효율성을 높였으나, 하자 데이터를 엑셀 파일과 같은 형식으로 관리하기 때문에 데이터의 위변조 위험에서 벗어날 수 없다. 논문[2]는 다양한 건설 분야에서 데이터의 정확성, 투명성, 보안성, 안정성을 위한 목적으로 블록체인을 사용하고 있다고 한다.

본 연구에서도 아파트 하자 데이터를 관리하는데 있어 투명성을 보장하기 위해 블록체인 기술을 적용하고자 한다. 제안한 시스템은 SHA-256 해시 함수와 머클 트리 구조를 활용하여 하자 정보의 변조를 방지하고, 데이터의 무결성을 체계적으로 검증할 수 있도록 설계되었다.

## II. 이론적 배경

### 2.1 블록체인

블록체인은 분산 데이터 저장 기술로, 각 참여자가 네트워크상의 모든 거래 기록의 사본을 보유하고, 모든 노드가 거래의 유효성을 독립적으로 검증한다. 이 기술은 데이터의 무결성과 보안성을 보장하며, 중앙 집중식 관리가 필요 없다는 장점이 있다. 특히, 본 연구에서 적용된 블록체인은 합의 알고리즘 없이 각 노드가 독립적으로 데이터의 유효성을 검증하므로 처리 속도와 네트워크 효율성의 향상이 기대된다.

### 2.2 머클 트리

머클 트리는 블록체인 기술에서 널리 사용되는 데이터 구조로, 각 노드의 데이터 무결성과 일관성을 보장하는 데 핵심적인 역할을 한다. 해당 이진 트리 구조는 각 리프 노드가 개별 데이터 블록의

해시를 포함하고, 비리프 노드는 자식 노드의 해시값을 합쳐서 해시를 생성한다. 이 구조는 데이터나 트랜잭션이 변경될 때 해당 변경이 트리의 루트까지 전파되므로, 머클 루트의 해시값만 비교함으로써 전체 데이터 세트의 무결성을 효과적으로 검증할 수 있다. 본 논문의 블록체인 시스템에서는 논문[3]과 유사하게 같이 머클 트리를 활용하여 데이터 무결성 검증 절차를 간소화하고, 빠른 검증을 가능하게 한다.

### 2.3 P2P 네트워크

논문 [4]에서는 P2P 시스템은 네트워크 환경에서 집중화된 서비스 개념 없이 분산 자원의 공유를 목적으로 동등한 자격을 가진 자율적 객체로 이루어진 자율 구성 시스템으로 정의 된다고 한다. 해당 네트워크 구조는 중앙 서버의 필요성을 없애고, 각 노드가 직접 데이터를 교환하며 통신한다. 이러한 구조는 네트워크의 장애 내성을 향상시키고, 탈중앙화를 실현하는 데 기여한다. 블록체인에서 P2P 네트워크는 트랜잭션과 블록 데이터를 네트워크 전체에 효과적으로 전파하며, 각 노드가 독립적으로 트랜잭션의 유효성을 검증하고 블록을 추가할 수 있게 한다. 이는 네트워크의 확장성과 보안을 동시에 증진시키는 중요한 요소이다.[4]

## III. 시스템 설계 및 구현

### 3.1 시스템 아키텍처

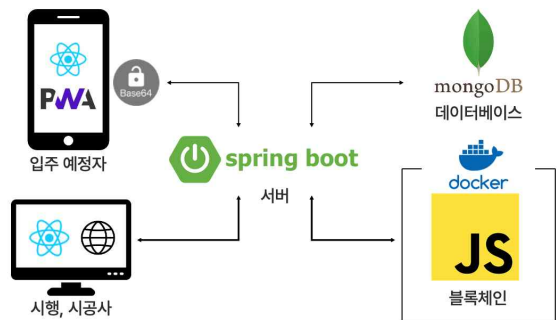


그림 1. 시스템 구조도

Fig 1. System Structure Diagram

그림 1은 제안하는 시스템의 구조도이다. 입주

예정자는 하자 등록 신청을 한다. 하자 등록 신청 데이터는 하자에 대한 정보들과 그것을 증빙하는 음성, 영상 또는 이미지 데이터이다. 하자를 증빙하는 데이터는 Base64를 통해 인코딩되어 서버를 통해 데이터베이스에 저장된다. 그 후 서버는 시행, 시공사로 하자 등록 신청을 전달한다. 시행, 시공사는 하자 등록 신청에 대해 승인 또는 거부를 할 수 있다. 거부를 선택할 경우 시행, 시공사는 거부 사유를 작성해야 하며 서버는 작성된 거부 사유와 함께 데이터베이스에 하자 정보가 거부되었음을 업데이트하고 입주 예정자에게 알린다. 시행, 시공사가 승인을 선택할 경우 서버는 블록체인을 저장하는 노드에 블록 생성을 요청해 블록을 생성하고 생성된 블록에 하자를 저장한다.

블록이 생성된 노드는 P2P 네트워크를 통해 다른 노드에 블록을 전파한다. 블록을 전파 받은 노드에서는 받은 블록에 대해 유효성 검사를 진행하고, 유효하지 않을 경우 다른 노드의 체인을 요청하여 유효성 검사를 수행한다. 다른 노드의 체인과 새로운 블록이 유효하다면, 기존의 체인을 해당 체인으로 교체하고 블록을 저장한다.

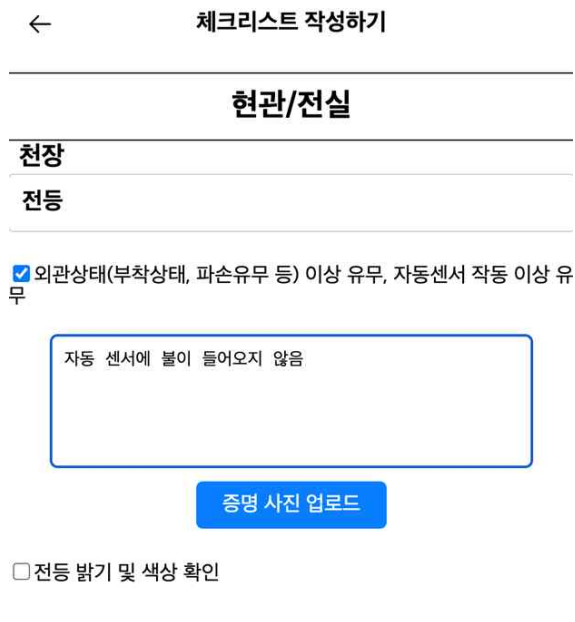


그림 2. 체크리스트 작성 페이지

Fig 2. write checklist page

시스템 구조를 조금 더 상세히 살펴보면, 그림 2는 발견한 하자 등록을 위해 입주 예정자가 체크리스트를 작성하는 페이지이다.



그림 3. 하자 승인 페이지

Fig 3. fault approve page

그림 3은 입주 예정자가 신청한 하자에 대해 시행, 시공사가 승인 혹은 거절을 하는 페이지이다.



그림 4. 하자 거부 페이지

Fig 4. fault reject page

그림 4는 그림 3에서 시행, 시공사가 거절을 선택하는 경우 나타나는 모달 창을 보여주는 것이다. 거절 사유를 시행, 시공사가 작성하고 저장을 누르면 입주 예정자에게 거절 사유가 전달된다.

### 3.2 데이터 관리

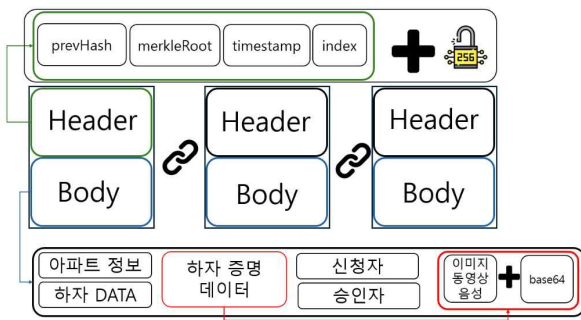


그림 5. 블록체인 구조도  
Fig 5. Blockchain Structure

그림 5는 제안한 시스템의 블록체인 구조도이다. 아파트 하자 정보를 블록 형태로 저장하며, 각 블록은 생성된 시간순으로 연결된다. 데이터의 무결성을 보장하기 위해 블록의 Header에는 이전 해시, 머클 루트, 생성 시간, 인덱스가 포함된다.

블록의 Body에는 아파트 하자에 대한 정보가 저장된다. 하자가 발생한 아파트, 하자에 관한 내용, 하자를 증명하는 이미지, 영상 또는 음성 데이터, 하자를 신청한 입주 예정자, 하자를 승인한 시행, 시공사 측 관리자의 정보가 포함된다.

블록이 서버의 메모리에 직접적으로 저장되기 때문에 하자를 증명하는 데이터인 이미지, 영상 또는 음성 데이터를 저장할 때 문제가 발생했다. 이를 해결하기 위해 제안한 시스템에서는 하자를 증명하는 데이터를 base64를 이용해 인코딩해 얻은 string을 저장한다.

#### IV. 분석

블록체인 기반 아파트 하자 관리 시스템에서 데이터 무결성은 신뢰성을 보장하는 핵심 요소이다. 본 시스템은 하자 정보가 블록체인에 정확히 기록되고 유지되도록 하여 무결성을 확보한다. 이를 위해 각 노드는 독립적으로 유효성을 검증하는 방식을 채택했으며, 이 과정에서 SHA-256 해시 함수와 머클 트리를 활용하여 효율적인 검증을 수행한다.

##### 4.1 무결성 검증 절차

새로운 하자 정보가 등록되면 각 노드는 해당 데이터를 독립적으로 검증한다. 시스템은 블록 생성

단계부터 무결성을 보장하며, 다음과 같은 절차를 따른다.

##### 4.1.1 블록 생성 및 전파

하자 데이터가 블록으로 생성되면 SHA-256 해시를 사용하여 블록의 고유 해시값을 생성하고 이를 블록 Header에 저장한다. 블록이 생성된 후, P2P 네트워크를 통해 다른 노드로 전파된다.

##### 4.1.2 노드 간 검증

전파된 블록을 받은 각 노드는 블록의 유효성을 검증한다. 불일치가 감지될 경우, 해당 노드는 다른 노드의 체인을 요청하여 데이터의 일관성을 유지한다.

##### 4.2 해시 함수의 역할

SHA-256 해시 함수는 블록 내 데이터의 무결성을 보장하는 데 중요한 역할을 한다. 블록이 생성될 때 SHA-256을 통해 생성된 해시값은 블록 간 연결을 유지하고, 데이터 변조 시 이를 즉각 탐지할 수 있도록 한다.

##### 4.3 체인 무결성 검증

블록체인의 연결 구조는 하자 관리 시스템의 신뢰성을 뒷받침한다. 각 블록은 이전 블록의 해시를 참조하고 있어 하나의 블록이 변경되면 이후 모든 블록이 무효화된다. 이로써 블록체인 전체의 데이터 일관성을 유지할 수 있다.

```

{
  "block": {
    "header": {
      "version": "1.0.0",
      "previousBlockHash": "0000000000000000000000000000000000000000000000000000000000000000",
      "merkleRoot": "03E49D1646C23D03822DE9F3F171655B87F6E6C4F406A1735F34D23D97C6E622",
      "hash": "5e3f1d8a7c77d64cdacd170f9607477ac5beb64e53937325dce26c40b7422db6",
      "merkleRoot": "03E49D1646C23D03822DE9F3F171655B87F6E6C4F406A1735F34D23D97C6E622",
      "merkleConstructionProcess": [
        {
          "block": {
            "header": {
              "version": "1.0.0",
              "previousBlockHash": "5e3f1d8a7c77d64cdacd170f9607477ac5beb64e53937325dce26c40b7422db6",
              "merkleRoot": "991013f83f6ecbf5401ffD5BFA1F15EBE4D8EF0EE11B25FC83852751ACD873F",
            }
          }
        }
      ]
    }
  }
}
    
```

그림 6. 생성된 블록

Fig 6. created blocks

그림 6은 생성된 블록에 대해 보여준다. 각 생성



된 블록은 자신의 머클 루트가 맞게 생성되었는지 확인 할 수 있다. 또한 다음의 생성된 블록이 가진 이전 블록의 해시값을 통해 무결성이 검증되었음을 알 수 있다.

<https://doi.org/10.22648/ETRI.2006.J.210501>

## V. 결론 및 향후 연구

본 연구에서는 블록체인 기술을 활용하여 아파트 하자 관리의 무결성을 보장하는 시스템을 설계하였다. 기존 시스템과 비교하여, 제안한 시스템은 SHA-256 해시 함수와 머클 트리, 체인 무결성 검증을 통해 데이터의 정확성과 보안을 강화함으로써 데이터 손실 또는 변조의 위험을 상당히 줄일 수 있음을 확인하였다. 또한, 분산된 P2P 네트워크 환경에서도 효율적으로 하자 정보를 관리할 수 있음을 보였다. 이는 데이터 관리의 투명성을 높이고, 전통적인 중앙집중식 시스템에 비해 더욱 안정적이며 확장 가능한 무결성 보장 방안을 제시한다.

향후 연구에서는 합의 알고리즘을 추가해 기존보다 더 안정적이고 확장 가능한 무결성 보장 방안을 고려할 수 있다. 예를 들어, PoS(Proof of Stake)와 같은 경량 합의 알고리즘을 통해 네트워크 전체의 신뢰성과 보안을 더욱 강화할 수 있을 것으로 기대한다. 이러한 개선은 블록체인 기반 하자 관리 시스템이 더욱 높은 신뢰성을 제공하고, 데이터의 안전성과 보안을 한층 강화하는 데 기여할 것이다.

## 참 고 문 헌

- [1] 서비스 채들, <https://eumservice.com/>
- [2] 김범수, 김성진, 김도영. (2023). 건설 분야에서 블록체인 활용 연구. 한국컴퓨터정보학회논문지, 28(1), 79-86.
- [3] 단관, 손이심, 고금명, 정희경. (2024). 블록체인 기반 클라우드 데이터의 안전한 저장에 관한 연구. 디지털콘텐츠학회논문지, 25(6), 1581-1588, 10.9728/dcs.2024.25.6.1581
- [4] 박호진, 박광로 . (2006). P2P 기술 동향 및 홈네트워크 응용. 한국전자통신연구원.

# 빅데이터 기반 다중 선형 회귀 모델을 통한 청년층의 지역별 문화·여가 선호도 분석

이민우\*, 채주혁\*, 안형태\*\*

## Analysis of Youth Regional Cultural and Leisure Preferences Using a Big Data-Based Multiple Linear Regression Model

MinWoo Lee\*, JuHyeok Chae\*, and HyeongTae Ahn\*\*

### 요 약

현대 한국 사회는 청년층의 수도권 집중화와 지방 청년 인구 감소라는 심각한 사회경제적 문제에 직면해 있다. 이러한 인구 이동 현상은 지방의 경제적 활력 저하와 사회적 불균형을 심화시키며, 장기적으로는 지역 소멸의 위기를 초래한다. 이에 본 연구에서는 청년층의 인구 이동 현상에서 지역별 문화·여가 선호도에 대한 영향력을 빅데이터 기반으로 분석하고 예측하는 모델을 제안한다. 지역별 인구 데이터와 문화향수 실태조사 데이터를 활용하여 청년 인구와 문화 활동 선호도를 결합한 데이터를 구축하고, 다중 선형 회귀 모델을 통해 청년층의 지역별 선호도를 예측한다. 본 연구의 결과는 향후 청년층을 위한 지역별 문화·여가 정책 수립에 활용될 것으로 기대된다.

### Abstract

Contemporary Korean society is facing serious socioeconomic problems such as the concentration of young people in the metropolitan area and the decrease in the youth population in local areas. This population movement phenomenon deepens the economic vitality and social imbalance of local areas, and in the long term, it leads to the crisis of regional extinction. Therefore, this study proposes a model that analyzes and predicts the influence of the population movement phenomenon of young people on regional cultural and leisure preferences based on big data. Using regional population data and cultural enjoyment survey data, data combining the youth population and cultural activity preferences are constructed, and the regional preferences of young people are predicted using a multiple linear regression model. The results of this study are expected to be utilized in the establishment of regional cultural and leisure policies for young people in the future.

### Key words

Big Data, Multiple Linear Regression, Regional Analysis, Cultural and Leisure, Youth, Migration

---

\* 금오공과대학교 컴퓨터공학과 학부생, minwoo1119@kumoh.ac.kr, kyoung0161@kumoh.ac.kr

\*\* 금오공과대학교 컴퓨터공학과 교수(교신저자), anten@kumoh.ac.kr

※ 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1G1A1005366).

## I. 서론

현대 한국 사회는 청년층의 수도권 집중화와 지방 청년 인구 감소라는 심각한 사회경제적 문제에 직면해 있다. 행정안전부의 보고서에 따르면, 청년 인구는 수도권으로 집중되는 경향을 보이며, 이에 따라 지방의 청년 인구는 지속적으로 감소하고 있다[1]. 이러한 현상은 지역 간의 사회경제적 불균형을 심화시키며, 지방의 지속 가능성을 위협하는 요인으로 작용하고 있다.

그림 1은 2015년부터 2023년까지의 수도권과 지방 청년 인구의 변화를 보여준다. 그래프를 보면, 수도권의 지방의 청년 인구 격차가 증가하는 것을 나타낸다. 이는 청년층이 수도권으로 이동하는 현상이 강화되고 있음을 나타낸다.

청년들이 대도시로 이동하는 이유는 다양하다. 교육 기회의 확대, 경제적 기회의 증가, 문화 및 여가 시설의 다양성, 활발한 사회적 네트워크, 편리한 인프라, 최신 기술 접근성 등이 주요 요인으로 작용한다[2]. 이러한 요인들은 서로 상호작용하며 청년층의 선택에 영향을 미친다. 특히, 문화와 여가적 요소는 청년층에게 중요한 의미를 가진다. 문화적 및 여가 활동은 청년들이 사회적 관계를 형성하고, 문화적 관심과 선호를 충족하는 데 중요한 역할을 한다[3].

표 1은 한국문화정보원의 검색기반 문화시설 트렌드 데이터를 활용하여 청년층이 자주 찾는 문화·여가 시설의 검색 데이터를 나타낸 자료이다. 수도권에 위치한 주요 문화 시설들은 높은 검색 횟수를 기록하며, 이는 수도권에 청년층의 문화적 취향을 만족시키고 있음을 나타낸다. 반면, 지방의 문화 시설들은 상대적으로 낮은 검색 횟수를 보이며, 지방의 문화적 인프라가 수도권에 비해 부족하다는 것을 보여준다.

기존 연구들은 주로 거시적 요인에 초점을 맞추고 단순 통계적 분석 방법에 의존하여, 청년층의 복합적인 선호도를 충분히 반영하지 못하는 한계가 있다. 이에 본 연구에서는 청년층의 복잡한 문화·여가 선호 요인을 분석하기 위해 다중 선형 회귀

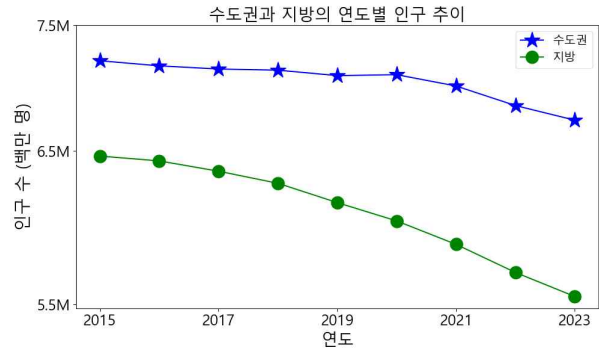


그림 1. 수도권과 지방의 연도별 청년 인구 (2015-2023)  
Figure 1. Youth population by year in the metropolitan area and local areas (2015-2023)

(Multiple Linear Regression) 모델을 활용하여, 문화와 여가가 청년 인구의 이동에 미치는 영향을 분석한다.

## II. 본론

본 연구에서는 청년층의 지역별 문화 선호도를 예측하기 위해 다중 선형 회귀 모델을 기반으로 분석했다. 분석 과정은 크게 데이터 준비, 모델 설계 및 학습, 그리고 모델 평가의 세 단계로 구성된다.

먼저, 청년층의 지역별 문화 선호도를 예측하기 위해 통계청에서 제공하는 인구 데이터를 사용했다. 지역별로 20~30대 청년 인구만을 필터링하여 분석에 사용함으로써, 청년층이 집중된 지역과 그렇지 않은 지역의 차이를 분류했다. 또한, 청년들의 문화 활동에 대한 선호도를 분석하기 위해 각 문화·여가

표 1. 청년층이 자주 찾는 문화·여가 시설의 검색 횟수 상위 10개

Table 1. Top 10 Cultural and Leisure Facilities Searched by Youth

지역	문화시설명	검색 횟수
서울	충무아트센터	4870.0
부산	영화의전당	3423.0
경기	성남아트센터	3121.0
서울	마포아트센터	2324.0
대전	대전예술의전당	2154.0
대구	대구콘서트하우스	1548.0
대구	수성아트피아	1533.0
충북	청주예술의전당	1460.0
대구	대구오페라하우스	1456.0
강원	강릉아트센터	1422.0

표 2. 주요 특성 분석

Table 2. Key Feature Analysis

순위	피처명	계수
1	문화 정보 지수	1.646259
2	전시시설 만족도	1.482777
3	인기 음악 공연 만족도	0.787311
4	전시관 상태	0.019832
5	문화여가 시간 충분지수	0.002827
6	극장 만족도	0.000032
7	종교 시설 만족도	0.000015
8	전통 문화 만족도	0.000007

활동별 선호도, 만족도, SNS 문화 행사 홍보를 나타내는 문화 정보 지수 등의 데이터를 사용했다. 입력 데이터는 인구수와 문화 관련 관람 수, 만족도, 참여 의향 등의 데이터를 포함하며, 지역별 문화·여가 만족도 점수인 total\_satisfaction 칼럼을 타겟 변수로 사용했다. 지역명은 범주형 데이터이므로, 모델에 포함시키지 않았으며, 대신 각 지역의 인구와 문화 관련 데이터를 활용하여 예측을 수행하였다.

학습된 모델의 성능은 MSE(Mean Square Error)를 사용하여 평가했다. 데이터는 train 데이터와 test 데이터를 8:2 비율로 분할하여, MSE를 통해 성능을 평가했다. 검증 결과 MSE가 0.087으로, 높은 정확도를 보였다.

모델을 기반으로 청년층이 높은 관심을 보이는 지역의 문화적 특징을 분석했다. 표 2에서 청년층의 지역 선택에 영향을 미치는 주요 변수로 문화 정보 지수, 전시시설 만족도, 인기 음악 공연 만족 지수 등이 도출되었으며, 문화·여가 시간 충분지수, 극장 만족도, 종교 시설 만족도, 전통 문화 만족도 등의 특성들은 청년들의 문화·여가 선호도에 거의 영향을 미치지 않는다.

표 3은 서울 수도권 및 주요 도시와 같은 대도시들이 높은 선호도를 보이는 반면, 강원도와 경상북도와 같은 광역시가 아닌 지역의 선호도가 상대적으로 낮게 예측되었다. 또한, 다중 선형 회귀 모델로 예측한 선호도가 실제 청년 인구 분포와 유사하게 나타났다. 이러한 결과는 본 연구에서 제안한 모델이 문화 및 여가 선호도에 영향을 미치는 주요 요인을 효과적으로 반영하고 있음을 검증한다.

본 연구의 결과는 비수도권 지역이 청년층의 문화·여가적 요구를 충분히 충족시키기 위해, 적극적

표 3. 지역별 선호도 예측 결과

Table 3. Regional Preference Prediction Results

순위	지역	예측 선호도 점수	청년 인구 비율 (순위)
1	서울특별시	6.987410	29.54(1)
2	부산광역시	6.558789	23.01(10)
3	세종특별자치시	6.345355	24.67(6)
4	인천광역시	5.971700	25.70(4)
5	대구광역시	5.826798	23.41(9)
6	경기도	5.155028	25.87(3)
7	광주광역시	5.110526	25.37(5)
8	대전광역시	5.000195	26.76(2)
9	충청북도	5.000192	22.90(7)
10	제주특별자치도	4.891984	22.14(11)
11	울산광역시	4.812505	23.18(8)
12	경상남도	4.59827	20.30(14)
13	충청남도	4.360571	22.16(12)
14	전라북도	4.255824	20.41(13)
15	전라남도	4.248611	18.98(17)
16	강원특별자치도	4.189673	20.80(15)
17	경상북도	3.566051	19.54(16)

으로 문화·여가 행사를 홍보하고 팝업스토어와 같은 전시시설을 확충하며 인기 음악 공연에 대한 지원을 확대해야 한다는 것을 시사한다.

### III. 결론

본 연구는 빅데이터 기반의 다중 선형 회귀 모델을 활용하여 청년층의 지역 선호도에 영향을 미치는 문화·여가적 요인을 분석하였다. 주요 연구 결과로, 비수도권 지역의 문화·여가에 대한 충분한 정보 제공이 가장 중요한 요인임을 확인했다. 또한, 청년층에 대한 지방의 매력을 높이기 위해 문화·여가 시설의 확충과 음악·팝업 이벤트와 같은 다양한 활동이 필요하다. 본 연구 결과를 바탕으로 청년층의 선호와 요구를 반영한 맞춤형 정책을 수립한다면, 지방의 청년 인구 감소 문제를 완화할 수 있을 것으로 기대된다. 추후 연구에서는 더 다양한 문화 데이터를 활용하여 청년층의 문화적 선호도를 보다 세부적으로 예측함으로써, 특정 지역에서 선호되는

문화 활동의 특성과 청년층의 요구를 정확하게 파악할 수 있을 것이다.

### 참 고 문 헌

- [1] 행정안전부, "2023년 지역별 인구 변화 보고서," 행정안전부, 2023.
- [2] 서울연구원, "청년층의 지역별 이동 요인과 유출 방지 정책 연구," 서울연구원, 2022.
- [3] 한국문화관광연구원, "문화·여가 활동과 국민 행복도 연구," 한국문화관광연구원, 2023.
- [4] 한국관광공사, "DATA LAB - 지역별 분석: 강원도 양양군," 한국관광공사, 2023.
- [5] 한경, "전통 문화자산과 첨단기술 융합: 경상북도의 청년층 유입을 위한 노력," 한국경제신문사 보고서, 2022.

# Graph-RAG를 활용한 추천 시스템 성능 개선: 여행 및 장소 POI 데이터셋 기반 분석

서희찬\*, 김세훈\*, 박익범\*, 신연순\*\*

## Enhancing Recommendation System Performance with Graph-RAG: An Analysis Using Travel and POI Datasets

Hui-Chan Seo\*, Se-Hun Kim\*, Ik-Bum Park\*, Youn-Soon Shin\*\*

### 요약

본 연구는 추천 시스템 환경에서 Graph-RAG 방식의 도입이 성능 개선에 미치는 영향을 분석하였다. 여행 및 장소 관련 64,614개의 POI(Point of Interest) 데이터셋을 사용하여 RAG, Multi-RAG, Graph-RAG 기법을 적용한 챗봇을 각각 개발하고, 여러 성능 지표를 기반으로 실험을 수행하였다. 실험 결과, Graph-RAG는 추천 정확도, 응답 일관성, 사용자 만족도에서 우수한 성능을 보이며, 추천 시스템의 성능을 향상시킬 새로운 가능성을 제시하였다. 특히, 그래프 기반 정보 연결을 통해 사용자의 개별 선호도를 더욱 정확하게 파악하고, 맞춤형 추천을 제공할 수 있음을 확인하였다.

### Abstract

This study focuses on the Retrieval-Augmented Generation (RAG) technique, applicable in the prompt engineering process of generative AI models, to analyze how the introduction of the Graph-RAG method contributes to performance improvement based on various performance metrics. Experiments were conducted in a recommendation system environment using a dataset comprising 64,614 Points of Interest (POI) related to travel and locations. In the experimental setup, chatbots applying RAG, Multi-RAG, and Graph-RAG techniques were each developed utilizing a Large Language Model (LLM). The experimental results demonstrated that the Graph-RAG method exhibited overall superior performance improvements compared to the existing RAG and Multi-RAG methods in terms of recommendation accuracy, response consistency, and user satisfaction.

### Key words

Retrieval-Augmented Generation (RAG), Graph-RAG, Multi-RAG, Large Language Model (LLM), Recommendation Systems, Chatbot, Performance Improvement, Prompt Engineering

---

\*동국대학교 공과대학 컴퓨터공학전공(gmlcks00513@gmail.com, pibox0824@gmail.com, pibum@naver.com),

\*\*동국대학교 AI융합대학 AI소프트웨어융합학부, ysshin@dongguk.edu(교신저자)

※ "본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음"(2023-0-00049)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음(IITP-2024-2020-0-01789)

## I. 서 론

최근 자연어 처리(Natural Language Processing, NLP) 분야에서 생성형 모델(Generative Model)의 발전은 언어 생성, 번역, 요약 등 다양한 응용 분야에서 중요한 성과를 거두고 있다. 특히, 대규모 사전 훈련 언어 모델(Pre-trained Language Models)은 문맥을 이해하고 자연스러운 문장을 생성하는 능력을 통해 인공지능(AI) 연구의 중요한 이정표로 여겨진다. 그러나 이러한 생성형 모델은 항상 정밀한 정보에 기반한 답변을 생성하지 못하는 한계가 있다. 이러한 한계를 극복하기 위한 접근 방식 중 하나로 Retrieval-Augmented Generation (RAG)이 제안되었다. RAG는 생성 과정에서 사전 훈련된 모델 외부의 추가 정보를 검색하여 결과를 생성하는 방식으로, 정보의 정확성과 다양성을 높이는 데 기여한다. 하지만 기존 RAG 모델은 검색된 정보를 활용하여 높은 정확도와 유창성을 보이지만, 적재된 지식을 재작업 없이 사용해 답변을 생성하는 과정에서 정확도가 떨어질 수 있다. 또한, RAG 구성 시점 이후의 실시간 데이터를 반영할 수 없어 맥락 이해 능력이 부족하고 편향된 정보 문제를 야기할 수 있는 한계가 존재한다. [1]

본 연구는 이러한 한계를 개선하고자 추천 시스템 환경에서 Graph-RAG 방식을 도입하였다. Graph-RAG는 데이터 간의 관계를 그래프 구조로 모델링함으로써 정보 검색과 생성 과정 간의 상호작용을 더욱 효율적으로 개선할 잠재력을 지니고 있다. 본 연구의 목적은 RAG와 Graph-RAG, Multi-RAG의 성능을 비교함으로써 추천 시스템에서의 활용 가능성을 평가하는 데 있다.

## II. 관련 연구

RAG[2,3]에 대한 연구는 자연어 처리(NLP)와 추천 시스템 분야에서 중요한 발전을 이루어왔다. Lewis, P., Perez,[2]는 RAG를 처음으로 제안하였으며, 대규모 비구조화 데이터를 기반으로 한 질의응답(Q&A) 시스템에서 우수한 성능을 입증하였다. RAG는 생성 모델의 약점을 보완하기 위해 외부에

서 관련 정보를 검색하고, 이를 모델의 출력에 포함시키는 방식으로 작동한다. 이 과정에서 검색된 정보의 품질과 다양성은 RAG 모델의 성능에 중요한 영향을 미친다. 추천 시스템 분야에서도 RAG의 적용이 활발히 연구되고 있다. 최근에는 LangChain 프레임워크를 활용하여 대규모 언어 모델(LLM)의 성능을 기반으로 한 프롬프트 엔지니어링 과정에서 RAG 방식을 적용하는 기술이 주목받고 있다. LangChain은 LLM의 프롬프트와 외부 검색 정보를 결합하여 보다 정확한 추천 결과를 생성할 수 있는 환경을 제공한다. 이 과정에서 RAG 방식은 검색 정보의 품질과 생성 모델의 응답 정확도를 향상시키는 데 중요한 역할을 한다.[4] 그러나 기존 연구에서는 검색된 정보 간의 관계를 명시적으로 고려하지 않는 한계가 있었다.

이러한 한계를 극복하기 위해 Graph-RAG 방식[1]이 제안되었다. Graph-RAG는 데이터 간의 관계를 그래프 구조로 모델링하여 정보 검색의 정확성을 높이고 생성 모델의 성능을 강화하는 접근 방식이다. 이를 통해 데이터 간의 상호 연관성을 반영하여 보다 정밀한 검색 결과를 생성할 수 있으며, 추천 시스템의 성능 또한 개선할 수 있다. Graph-RAG 방식에서는 다양한 피처를 그래프 노드로 표현하고, 이들 간의 관계를 그래프 에지로 나타내어 추천 항목 생성에 활용한다. Graph-RAG는 데이터의 상호 연관성을 명시적으로 모델링함으로써, 기존 RAG 방식의 한계를 극복하고자 한다. 기존 RAG 방식에서는 개별 문서들 간의 관계를 고려하지 않고 독립적으로 정보를 검색하고 생성에 반영하였다. 반면 Graph-RAG는 데이터 간의 관계를 그래프로 표현하여, 정보를 검색하고 이를 바탕으로 생성하는 과정에서 이러한 관계를 반영한다. 이를 통해 보다 정교하고 일관된 정보 생성이 가능하며, 사용자 맞춤형 추천 시스템의 성능도 크게 향상될 수 있다.

## III. 연구 방법

본 연구는 RAG, Multi-RAG, Graph-RAG방식의 성능을 추천 시스템 환경에서 비교하기 위해 약 64,614개의 POI 데이터를 기반으로 실험을 진행하

였다. 데이터셋은 공개 데이터와 직접 수집한 데이터를 결합하여 구성하였으며, RAG와 Multi-RAG는 MySQL 데이터베이스를 활용한 반면, Graph-RAG는 Neo4j그래프 데이터베이스를 기반으로 지식 그래프(Knowledge Graph)를 사용하였다. 이를 통해 서로 다른 구조에서의 성능 차이를 비교하였다.

### 1. 데이터셋 및 그래프 모델링

데이터 수집 및 전처리: 여행 및 장소 관련 POI 데이터를 수집하고 정제하였다. 정제 과정에서 데이터 컬럼은 Place Name, Category Group Name, Category Name, Latitude, Longitude, Road Address Name, Place URL로 구성되었다.

이후 추가적인 전처리 과정을 거쳐 Place노드와 Category노드로 저장, 장소 간 근접성은 NEARBY관계로 연결하였다.

Neo4j 모델링: 각 장소는 Place 노드로 저장, 카테고리는 Category 노드와 HAS\_CATEGORY관계로 연결, 장소 간 근접성은 NEARBY관계로 모델링하였다.

### 2. 관계 도출 및 수학적 모델링

그래프 구조: Neo4j의 데이터는  $G=(V,E)$ 로 표현되며,  $V$ 는 노드(장소,카테고리)의 집합,  $E$ 는 노드 간 관계(엣지)를 나타낸다.

카테고리 관계(HAS\_CATEGORY): 특정 장소와 카테고리 간의 관계를 수식 (1)과 같이 표현할 수 있다.  $R(Place, Category)$  는 장소 P가 카테고리 C에 속할 때 1, 그렇지 않을 때 0으로 정의된다.

$$R(P, C) = \begin{cases} 1 & \text{if } Place \text{ belongs to } Category \ C \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

근접성 관계(NEARBY): 두 장소 간의 거리  $d(P_1, P_2)$ 가 임계값  $\delta$  이하일 때, NEARBY 관계가 성립한다. 이는 수식 (2,3)과 같이 표현할 수 있다.

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2)$$

$$E(P_1, P_2) = \begin{cases} 1 & \text{if } d(P_1, P_2) \leq \delta \\ 0 & \text{if } d(P_1, P_2) > \delta \end{cases} \quad (3)$$

### 3. RAG, Multi-RAG, Graph-RAG 구현

RAG: 그림 1과 같이 MySQL 데이터베이스에서 관련 데이터를 검색하여 답변을 생성한다. 이 방식에서는 질문과 관련된 데이터를 검색한 후, 이를 그대로 참조하여 답변을 제공한다.

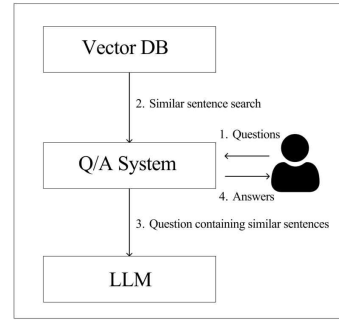


그림 1. RAG 구조도  
Fig. 1. RAG Structural Chart

Multi-RAG: 그림 2처럼 키워드 기반 벡터 DB를 사용하여 질문에서 추출된 키워드와 관련된 카테고리 데이터만 검색하는 방식으로 동작한다. 질문에서 특정 키워드를 추출한 후, 해당 키워드와 연관된 데이터만 검색하여 더 구체적인 답변을 제공한다.

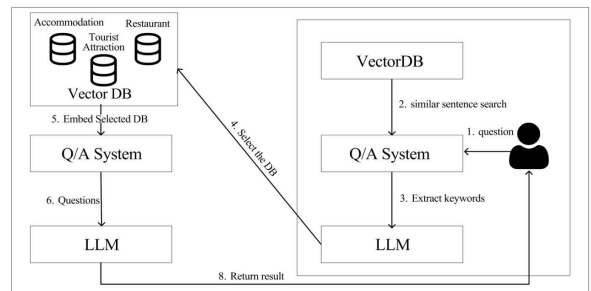


그림 2. Multi-RAG 구조도  
Fig. 2. Multi-RAG Structural Chart

Graph-RAG: Neo4j 그래프 데이터베이스를 활용하여 장소 간의 관계를 고려한 답변을 생성한다. 이 기법은 카테고리 관계

(HAS\_CATEGORY)와 근접성 관계(NEARBY)를 활용하여 장소간의 상호작용을 고려한 답변을 생성한다. Graph-RAG는 복잡한 질문에 대해 장소 간의 연관성과 근접성을 반영하여 더 정교하고 맞춤형 답변을 제공할 수 있다.



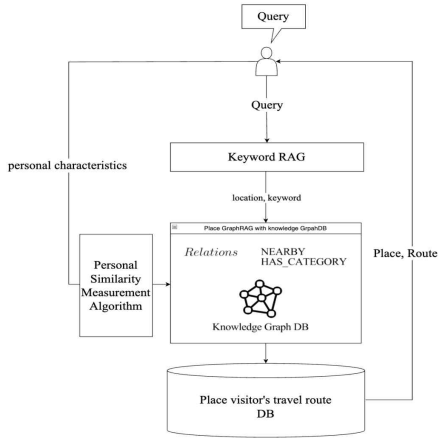


그림 3. Graph-RAG 구조도  
Fig. 3. Graph-RAG Structural Chart

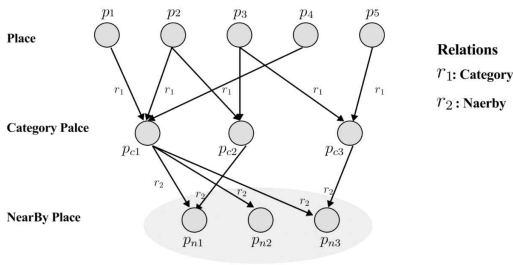


그림 4. 지식 그래프 노드간 관계  
Fig. 4. Knowledge graph node-to-node

#### 4. 실험 설계 및 성능 평가

##### 4-1. 실험 설계

본 연구에서는 RAG, Multi-RAG, Graph-RAG 세 가지 기법을 비교하기 위해 다양한 사용자 질문을 기반으로 실험을 설계하였다. 각 기법은 실시간으로 질문에 대한 응답을 생성하며, 이를 통해 챗봇의 응답을 성능 지표로 평가하였다.

##### 4-2. 성능 측정 지표

세 가지 RAG 방식은 다음 지표를 기준으로 평가되었다:

**추천 정확도:** 사용자의 질문에 대해 얼마나 적절하고 구체적인 장소를 추천하는지를 평가한다. 예를 들어, 사용자가 특정 지역의 맛집을 묻는 경우, 적절한 장소와 카테고리의 추천 여부를 기준으로 측정한다.

**응답 일관성:** 복합적인 질문에 대해 일관성 있게 답변을 제공하는지 평가한다. 연속적인 질문(예:

장소 추천 후 해당 장소 근처 명소 문의)에 대한 응답의 일관성을 분석한다.

**응답 시간:** 사용자가 질문한 후 챗봇이 응답하기까지 걸린 시간을 측정하여, 각 기법의 처리 속도를 비교한다.

**사용자 만족도:** 응답의 적절성, 속도, 일관성에 대해 사용자가 얼마나 만족했는지 평가한다. 이 항목은 설문조사를 통해 수집한 데이터를 기반으로 분석된다.

#### 4.3. 사용자 설문조사 및 통계 분석

본 실험에서는 50명의 장소 추천 서비스를 이용하는 사용자를 대상으로 설문조사를 실시하였다. 사용자들은 세 가지 RAG 방식으로 응답을 받으며, 각 기법의 추천 정확도, 응답 일관성, 응답 시간, 사용자 만족도를 평가하였다. 설문조사 데이터를 기반으로 세 가지 RAG 방식의 성능 차이를 통계적으로 분석하였다.

### IV. 연구 결과

본 절에서는 RAG, Multi-RAG, Graph-RAG 세 가지 기법을 적용한 5개의 챗봇의 대화 시나리오를 통해 성능을 비교하였다. 시나리오는 모두 다양한 조건(예: 장소 유형, 사용자 요구)과 주요 지역명이 포함된 대화식 추천요청으로 구성하였다.

아래 그림 5의 시나리오 예는 실제 사용자와 챗봇 간의 대화를 기반으로 설정되었으며, 각 기법에 대해 위 성능 지표 측면에서의 차이를 평가하였다.

Question	RAG	Multi-RAG	Graph-RAG
서울 중구 관광명소 추천해줘	서울에서 인기가 많은 곳으로 남산타워와 명동거리가 있습니다.	서울 중구에서 추천하는 관광지로는 덕수궁과 남산을 한옥마을이 있습니다.	서울 중구에 덕수궁이 있습니다. 특히 덕수궁 주변에는 남산을 한옥마을과 다양한 전통 문화 공간도 있어요. 방문해 보시는 걸 추천드려요.
근처에 식당 있어?	명동 근처에는 많은 음식점이 있어요. 한식, 일식, 양식 모두 즐길 수 있습니다.	중구에는 다양한 한식당이 있습니다. 가깝게는 한식 전문점 '한옥'을 추천합니다.	덕수궁 근처에 '맛찬들'이라는 한식당이 있습니다. 전통적인 분위기와 함께 한식을 즐길 수 있어요. 또한, 덕수궁에서 걸어갈 수 있는 거리에는 전통 찻집도 있습니다.
식당 근처에 숙소는 뭐 있어?	서울 시내에 다양한 숙소가 있어요. 명동이나 종로 주변에 호텔과 게스트하우스가 많이 위치해 있습니다.	주변 숙소 정보는 제공되지 않지만, 중구에는 다양한 숙박 시설이 있습니다.	맛찬들 근처에 위치한 '호텔 중구'를 추천드려요. 덕수궁과도 가깝고 식당과 관광지를 편리하게 이동할 수 있습니다. 조용한 숙소로 인기가 많습니다.

그림 5. 서울 중구 관광명소 추천에 관한 대화  
Fig. 5. Conversation on Recommendation of Tourist Attractions in Jung-gu, Seoul

표1. 서울 중구 관광명소 추천 대화 분석  
Table 1. Analysis of recommended conversation for tourist attractions in Jung-gu, Seoul

대화 내용	RAG	Multi-RAG	Graph-RAG
서울 중구 관광명소 추천해줘	명소 추천	명소 추천	명소 추천
근처에 식당 있어?	근처 정보 제공 불가	근처 정보 제공 불가 일반적인 추천	앞서 추천해준 명소 근처 식당 추천
식당 근처에 숙소는 뭐 있어?	근처 정보 제공 불가 일반적인 추천	근처 정보 제공 불가	앞서 추천해준 식당 근처 숙소 추천

실험 결과, 표1과 같이 Graph-RAG기법은 다른 두 방식에 비해 사용자 질문에 대한 세부적이고 맥락을 고려한 응답을 제공하며, 대화의 자연스러운 흐름을 유지하는 데 탁월한 성능을 보였다. 이를 통해 사용자 만족도가 가장 높게 평가되었으며, 복잡한 질문에도 일관된 응답을 제공할 수 있었다. 평가 항목별 결과는 표2와 같다.

표2. 평가 항목별 결과  
Table 2. Results by Evaluation Item

평가항목	RAG	Multi-RAG	Graph-RAG
추천 정확도	65%	80%	95%
응답 일관성	50%	70%	90%
응답 시간	0.5초	0.4초	0.6초
사용자 만족도	1.7/5	3.2/5	4.8/5

## V. 결 론

본 연구는 RAG, Multi-RAG, Graph-RAG 기법을 추천 시스템 환경에서 비교 분석하기 위해 여행 및 장소 관련 POI 데이터를 활용하여 각 기법을 적용한 챗봇을 개발하고 성능을 평가하였다. 실험 결과, Graph-RAG 방식이 RAG와 Multi-RAG 대비 추천 정확도, 응답 일관성, 사용자 만족도에서 전반적으로 우수한 성능을 보였다. 이는 Graph-RAG가 데이터 간 관계를 그래프로 모델링하여 정보 검색과 생성 과정에서 연관성을 효과적으로 반영했기 때문이다.

특히, Graph-RAG는 맥락 이해와 정확한 답변 생성 측면에서 할루시네이션을 감소시키며 일관된 정보를 제공할 수 있었다.

그러나 본 연구에는 몇 가지 한계가 있다. Graph-RAG는 정보 연관성에 기반해 성능을 개선했지만, RAG 모델들의 공통적인 단점인 처리 속도 문제를 완전히 해결하지 못했다. 데이터 규모가 커질수록 지연 시간이 증가하며, Graph-RAG 방식도 복잡한 그래프 구조 처리에서 추가적인 연산 비용이 발생해 실시간 응답에서 제약이 있다.

향후 연구에서는 Graph-RAG의 실시간 데이터 처리와 그래프 최적화 방안을 모색하고, 다양한 도메인에서의 적용 가능성을 검토하여 추천 시스템 성능 향상에 기여할 필요가 있다.

## 참 고 문 헌

- [1] 정천수. (2024). 지식 기반 QA개선을 위한 Advanced RAG 시스템 구현 방법: Graph Agent 활용. 지식경영연구, 25(3), 99-119.
- [2] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Riedel, S. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. \*Advances in Neural Information Processing Systems (NeurIPS)\*, 33, 9459-9474.
- [3] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877-1901.
- [4] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024). Retrieval-Augmented Generation for Large Language Models: A Survey. arXiv preprint arXiv:2312.10997v5.

# Malk:법령 데이터 기반의 지식 그래프를 활용한 멀티 에이전트 프레임워크 제안

박선화\*, 최승민\* 정유철\*\*†

## Malk:Proposal for a Multi-Agent Framework Utilizing Knowledge Graphs Based on Legal Data

Seonhwa Bak\*, Seungmin Choi\* and Yuchul Jung\*\*

### 요 약

본 논문은 법령 데이터를 기반으로 한 지식 그래프(Knowledge Graph)를 활용한 RAG(Retrieval-Augmented Generation) 기반 멀티 에이전트 프레임워크인 Malk를 제안한다. Malk는 한국 법령 데이터의 구조를 분석&정의하고 이를 기반으로 법령 지식 그래프를 구축한다. Malk의 파이프라인은 멀티 에이전트로 구성되는데, 각 에이전트는 정보 정제자, 작성자, 편집자, 판별자로 구성된다. 이렇게 구성된 Malk는 법률 정보 처리의 효율성과 정확성을 향상하고, 방대한 법령 데이터의 체계적인 관리와 연결성을 강화하는 것에 중점을 두고, 법률 문서 처리, 준법 감시, 법률 결정 지원 시스템 등 다양한 업무에 높은 활용도를 가진다.

### Abstract

This paper proposes Malk, a multi-agent framework based on Retrieval-Augmented Generation (RAG) utilizing a Knowledge Graph constructed from legal data. Malk analyzes and defines the structure of Korean legal data to build a legal knowledge graph. The pipeline of Malk is composed of multiple agents, each designated as an information purifier, generator, editor, and evaluator. This configuration of Malk aims to improve the efficiency and accuracy of legal information processing, strengthen the systematic management and connectivity of extensive legal data, and enhance applicability across various tasks, including legal document processing, compliance monitoring, and decision support systems in legal contexts.

### Key words

Malk, KG, RAG, legal data, multi-agent

---

\*국립금오공과대학교 컴퓨터공학과, 박선화 paksh0725@kumoh.ac.kr,

\*국립금오공과대학교 컴퓨터공학과, 최승민 csm1018csm@gmail.com,

\*\*국립금오공과대학교 인공지능공학과, 정유철 jyc@kumoh.ac.kr†

## I. 서론

법률 데이터를 활용한 챗봇 개발은 현재까지 꾸준히 연구 되어왔다. 그러나 법률 데이터 구조의 복잡성이나 문맥의 연결성 등의 문제로 인해 LLM 모델에 치명적인 환각 증상이 존재[1]했다. 해당 문제를 해결하기 위해 본 연구에서는 한국의 법률 데이터 및 판례 등을 활용한 지식 그래프(Knowledge Graph) 구축 및 RAG(Retrieval-Augmented Generation) 기반 멀티 에이전트 프레임워크를 제안한다.

본 연구에서 제안한 프레임워크를 위해 ‘법제처, 한국법령정보센터’에 존재하는 대한민국 법령 정보를 수집하고, 법령 정보를 판례와 사건 기준으로 정의한다. 이렇게 정의된 법령 정보를 활용하여 지식 그래프를 구축하고, 구축된 지식 그래프를 본 연구의 목적에 맞는 파이프라인을 통해 멀티 에이전트를 활용한 프레임워크인 Malk를 제안한다.

## II. 데이터 구조

본 연구에서는 해당 목적에 맞는 데이터를 실제 대한민국 법령 정보와 판례 등을 수집했다. 수집된 데이터는 상위 50개의 판례를 기준으로 수집 및 정제를 진행하고, 이렇게 정의된 구조는 다음과 같다.

표 1. 법령 데이터 판례 구조 정의  
Table 1. Definition of Legal Data and Case Law Structure

구조	설명
원고, 피고	소송 제기 하거나 받은 사람
(피)상고인	원심판결에 불복하여 대법원에 상고하거나 받은 사람
원심판결	대법원 상고 직전, 2심 법원판결
주문	대법원의 결론 1) 파기환송 2) 파기이송 3) 파기자판
이유	주문이 이뤄진 법리적 경과 설명
날짜	대법원 판결이 행해진 날짜
법관 서명	상고심 관여 대법관 서명 or 날인

## III. 파이프라인

본 논문에서는 지식 그래프를 활용한 RAG기반 멀티 에이전트 파이프라인을 구축했다.

### 1. 법령 지식 그래프 구축

법령 데이터 기반 지식 그래프 구축은 임베딩이 중요하다. 특히 법령 데이터의 경우, 판례나 특정 도메인에 따른 지역적 특성이 반영되어야 하기 때문에, 이를 통하여 본 연구에서는 임베딩 과정을 세분화하였다.

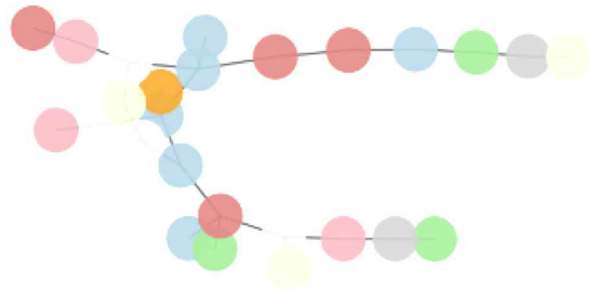


그림 507. 법령 데이터 지식 그래프 구축 예시

Fig. 1. Example of Building a Knowledge Graph from Legal Data

그래프 노드 임베딩을 통하여 각 노드를 저차원 공간으로 맵핑하는데, 핵심 노드가 될 핵심 정보를 추출하여 임베딩을 통한 재구성 작업을 수행한다. 이 과정에서 노드가 가지고 있는 핵심 정보는 유실되지 않도록 유사도를 최대화한다. 이 과정에서 학습은 귀납적으로 수행하여 지역적 구조를 통한 새로운 정보에도 대응할 수 있도록 한다[2]. 최종적인 산출은 관련 정보와 문제 해결을 위한 필요 정보를 병렬 산출하는데 부가적으로 KoBert가 검색 도구의 역할 수행한다.

### 2. Multi-Agent 판별자

판별기의 역할을 진행한다. 1번 지식 그래프를 통하여, 다음의 2가지 질문을 점수 및 정보로 반환한다. “사용자의 문제를 해결하기 위한 정보를 가지고 있는가?”, “사용자의 문제를 해결하기 위해, 필요한 정보는 무엇인가?”로 앞선 두 개의 질문으로

반환한다.

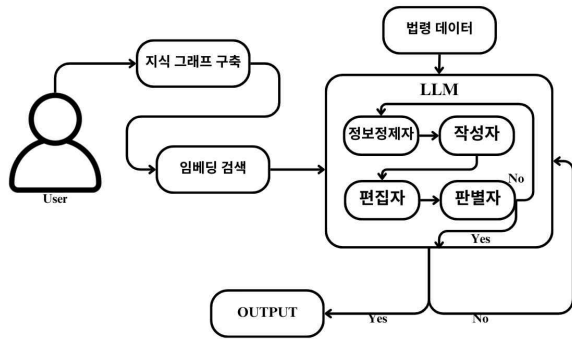


그림 508. Malk 전체 흐름도

Fig. 2. Overall Structure of the Malk Framework

1번 질문에서 긍정인 경우, 다음 작업으로 넘어가고, 부정이라면 다시 2번을 반복한다.

### 3. 멀티 에이전트

멀티 에이전트 기법은 다음 4개의 에이전트로 구현된다. 각 에이전트는 정보 정제자, 작성자, 편집자, 판별자의 역할을 진행한다.

각각 정제, 정보 바탕 답변 생성, 답변 수정 및 Markdown 변환, 질문에 대한 충분한 답변 제공의 기능을 한다. 이때 판별자의 경우, 부정의 답변이 도출된다면 초기 정보 정제자의 역할을 다시 수행한다.

이후로 랭 체인의 챗 메모리 관리를 위한 Multi turn을 사용하여 이전 대화의 내용을 함축적으로 모델에 전달하여, 이전 입력된 질문의 정보가 유실되지 않도록 한다[3].

### 4. QA 구조 데이터셋

사용자들이 자주 요청되는 query는 알고리즘 데이터 베이스에 저장하고, 데이터베이스의 노드와 사용자의 질문 일치도가 ROUGE-F1 score로 80% 이상인 경우, 그대로 답변하도록 한다.

## IV. 결론

최종적으로 본 연구는 법령 데이터를 통하여 지식 그래프를 구축하고, RAG를 기반한 멀티 에이전트를 구현했다. 제안된 기법을 활용한다면, 특정 도

메인인 법령 데이터에서 기존의 해결 불가능 문제였던 LLM의 환각 문제를 최소화할 수 있다.

이러한 모듈식 설계는 도메인의 확장성을 높이고, 신규 법령이나 수정되는 법령에 대한 적절한 대응이 가능하여 해결책을 제공할 수 있다.

## 참고 문헌

- [1] 박대민, and 이한중. "챗 GPT 등장 이후 인공지능 환각 연구의 문헌 검토: 아카이브 (arXiv) 의 논문을 중심으로." 정보화정책 31.2 (2024): 3-38.
- [2] Hamilton, Will, Zhitao Ying, and Jure Leskovec. "Inductive representation learning on large graphs." Advances in neural information processing systems 30 (2017).
- [3] Yu, Lingfan, Jinkun Lin, and Jinyang Li. "Stateful large language model serving with pensieve." arXiv preprint arXiv:2312.05516 (2023).

# 아두이노를 활용한 인터랙티브 엔딩 크레딧 콘텐츠 제작

김민경\*, 박민수\*, 변공규\*\*, 유선진\*\*\*

## Creating Interactive Ending Credits Contents using Arduino

Minkyoung Kim\*, Minsoo Park\*, Gongkyu Byeon\*\*, and Sunjin Yu\*\*\*

### 요약

최근 급변하는 사회·경제적 환경 속에서 청년들이 느끼는 불확실성과 혼란은 매우 크다. 이는 자아의 혼란을 초래하고, 무가치감을 느끼는 정서적장애로 이어질 수 있다. 이에 본 연구에서는 아두이노와 인터랙티브 기술을 활용하여 '엔딩 크레딧' 형태의 체험형 콘텐츠를 통해 청년들의 자아존중감 향상을 도모하고 미래에 대한 긍정적 사고를 촉진시킬 수 있는 인터랙티브 콘텐츠를 제안한다.

### Abstract

In the rapidly changing socio-economic environment, the uncertainty and confusion experienced by young people are significant. This can lead to a crisis of identity and emotional disorders characterized by feelings of worthlessness. In response to this, the present study proposes interactive content using Arduino and interactive technologies, designed as an 'ending credit' experience. This interactive content aims to enhance young people's self-esteem and promote positive thinking about the future.

### Key words

arduino, interactive technologies, python, web server framework, self-esteem

### 1. 서론

2024년 1월 통계청에 따르면 연령/교육정도별 실업률 동향의 15 - 29세 대졸이상의 실업률은 6.8%로 전체 대졸이상 실업률 2.7%보다 약 3배 더 높은 것으로 나타났다. 취업포털 잡코리아는 청년구직자

를 대상으로 스트레스 현황을 조사 한 결과 응답자의 93.8%가 '취업으로 인하여 스트레스를 받는다'고 하였다[1]. 이 같은 취업스트레스가 지속되면 우울, 불안, 낮은 자아존중감, 대인기피 등의 심리적·신체적 반응에 노출된다[2]. 이 때문에 청년들은 대학 생활을 회고하고 자아존중감을 향상시킬 기회가 필

\* 국립창원대학교 문화테크노학과, suuu0622@gs.cwnu.ac.kr, 3347girl@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화융합기술협동과정 박사과정, bgk9517@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr

※ 본 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구임.

요하다.

인터랙티브 기술은 관람자의 능동적 참여를 유도하며, 이를 통해 단순한 관람을 넘어, 자신의 생각과 감정을 표현하고, 새로운 시각을 얻을 수 있다 [3]. 이를 아두이노와 접목하여 관람자에게 자신만의 대학 시절 엔딩크레딧을 선사하여 더욱 깊은 몰입감과 새로운 경험을 전달한다.

## II. 아두이노 인터랙티브 콘텐츠 제작

인터랙티브 콘텐츠 ‘엔딩 크레딧’은 참여자가 직접 대학 생활을 돌아보고 정리할 수 있는 매개체로서, 자신의 대학 생활을 정리하며 성취감과 뿌듯함을 느낄 수 있도록 구성된다. 이 콘텐츠는 영상 기술을 결합하여 관람자가 콘텐츠와 상호작용하며 몰입감 있게 체험해볼 수 있다.

본 콘텐츠는 아두이노(Arduino)와 파이썬(Python) 서버구축을 기반으로 제작되었으며, 아두이노 초음파 센서를 활용해 참여자가 장치 앞에 서면 이를 인식하여 전시 의도를 설명하는 영상이 자동 재생된다. 이후 파이썬 웹사이트가 열리며, 참여자가 자신의 엔딩 크레딧을 작성해보는 인터랙티브 콘텐츠 경험을 제공한다.



그림 1. 아두이노 초음파 센서 작동 예시  
Fig. 1. Example of Arduino ultrasonic sensor operation

그림 1은 아두이노 초음파 센서를 사용하여 사람을 인식하는 그림이다. 실행 시 초음파 센서로부터 거리를 측정하고 거리가 5cm 이내이면 영상이 재생된다. 영상을 감상한 후 파이썬 웹사이트가 열리며, 화면에 입력할 수 있는 칸이 생성된다. 입력 항목은 '감독', '감사한 분들', '도움을 주신 분들', '친구 팀의 총 4개로 구성되어 있다. 참여자가 모든 항목을 입력하고 나면, 입력된 내용이 엔딩 크레딧 영상 형식으로 구현된다.

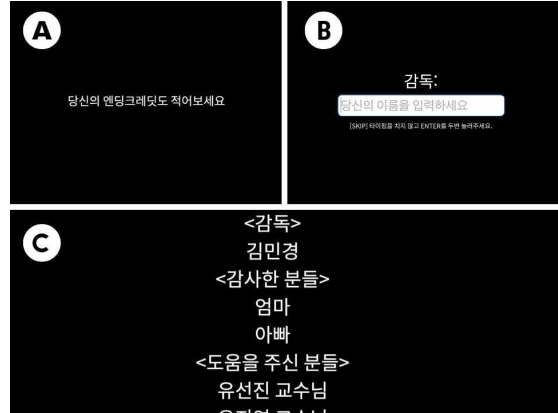


그림 2. 웹 서버 구축 콘텐츠 작동 예시  
Fig. 2. Example of Web Server Setup Content Operation

그림 2는 입력 항목과 엔딩 크레딧 영상이 구현되는 장면이다. 엔딩 크레딧 영상은 입력된 항목이 화면 밖으로 사라질 때까지 재생된다.

## III. 결 론

본 연구에서는 아두이노와 웹 서버 구축 콘텐츠를 활용하여 대학 생활을 회고하고 기록할 수 있는 인터랙티브 콘텐츠를 제작하였다. 관람자는 졸업에 대한 여러 감정을 정리하고 기록하면서 단순한 과거의 회상이 아닌, 자신을 더욱 깊이 이해하고 성장할 수 있는 기회를 얻었다. 향후 본 엔딩크레딧 인터랙티브 콘텐츠는 수집된 데이터를 통한 분석으로 청년들의 시작과 마무리를 응원하는 확장 콘텐츠 기술로 발전할 것이다.

## 참 고 문 헌

- [1] <https://www.sedaily.com/NewsView/1VPKIS3D9G>, Assessed, 2024.10.28.
- [2] B. S. KIM. , S. J. KIM. , and Y. J. AN. "Analysis of Major Syndromes of Psychological and Physical Responses of Dealers by Stages. Journal of Career Education Research.", 2009.
- [3] H. J. KIM. "Interdisciplinary Study for Analyzing Audience Experience in Interactive Media Art," Journal of Korean Society of Image Science and Technology, 2019.

# 아두이노와 프로세싱을 활용한 인터랙티브 감정 치유 콘텐츠

윤슬기\*, 추상아\*, 최현빈\*\*, 유선진\*\*\*

## Interactive Emotion Healing Contents using Arduino and Processing

Seulgi Yoon\*, Sanga Choo\*, Hyunbin Choi\*\*, and Sunjin Yu\*\*\*

### 요 약

현대 사회는 급격한 기술 발전과 함께 다양한 사회적 변화가 동반되면서 개인의 감정 관리가 이슈이다. 이에 따라 본 논문에서는 현대인의 억압된 감정을 완화할 수 있는 인터랙티브 콘텐츠를 제시한다. 아두이노를 활용해 체험자의 맥박 정보를 수집하고, 프로세싱으로 인터랙티브 콘텐츠와 음악을 출력한다. 이를 통해 체험자에게 자신의 감정을 직관적으로 이해하고, 정신 건강을 증진할 수 있는 경험을 제공한다.

### Abstract

With the rapid technological advancements and diverse societal changes of modern society, emotional management for individuals has become a critical issue. Accordingly, this paper presents an interactive content approach to alleviate the suppressed emotions of modern individuals. Utilizing Arduino, pulse information from the participant is collected, and Processing is used to display interactive content and output music. Through this experience, participants can intuitively understand their emotions, providing a beneficial experience that promotes mental well-being.

### Key words

arduino, processing, emotion suppression, media art, interactive art

## 1. 서 론

현대 사회는 급격한 기술 발전과 함께 다양한 사회적 변화가 동반되면서 개인의 감정 조절이 중요한 이슈이다. 특히 최근에는 업무 스트레스나 사회적 기대 및 압력, 대인 관계의 어려움 등으로 인해

많은 현대인이 자신의 감정을 제대로 표출하지 못하고 이를 억압하는 현상이 증가하고 있다. 그 결과 현대 사회의 구성원 대부분이 강박증이나 불면증, 식이장애, 감정 조절 장애와 같은 정신 질환에 시달린다.[1] 이에 따라 억제하려는 정보는 억제하면 할수록 커지고, 표현을 할 때 해소된다.[2] 따라서 본

\* 국립창원대학교 문화테크노학과 20200261@gs.cwnu.ac.kr, csa4545@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화융합기술협동과정 석사과정 20217292@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수 sjyu@changwon.ac.kr (교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스티원센터의 지원을 받아 수행된 연구입니다.



연구에서는 현대인이 자신의 감정을 더욱 효과적으로 이해하고, 이를 완화하기 위해 인터랙티브 콘텐츠와 음악을 활용하는 방법을 제시한다.

## II. 인터랙티브 콘텐츠 제작

본 콘텐츠는 아두이노(Arduino)와 프로세싱(Processing) 프로그램을 통해 개발한다. 아두이노 심박 센서를 통해 체험자의 맥박 정보 값을 수집한다. 이때, 아두이노의 PulseSensor Playground 라이브러리와 프로세싱의 Arduino(Filamata) 라이브러리를 설치한다. 이를 통해 수집된 맥박 정보 값이 프로세싱에 전달되도록 한다. 이후 프로세싱을 통해 인터랙티브 콘텐츠를 보여주고, 동시에 Sound 라이브러리를 활용하여 음악이 출력되도록 한다.

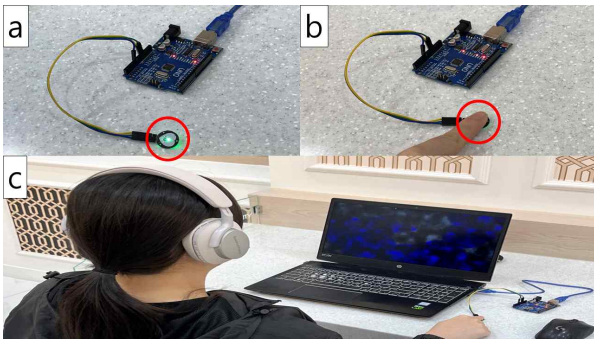


그림 1. 인터랙티브 콘텐츠 체험 방식  
Fig. 1. Interactive content experience method

그림 1은 인터랙티브 콘텐츠를 체험하는 과정이다. a에 표시된 맥박 센서 위에 b처럼 손을 올린 후, c와 같이 헤드셋을 착용하고 인터랙티브 콘텐츠를 감상한다. 체험자의 맥박에 따른 감정 상태는 우울/불안(분당 맥박수 70~80회), 기쁨/행복(분당 맥박수 81~90회), 흥분/분노(분당 맥박수 91~100회)의 3가지로 분류한다. 우울/불안을 완화하기 위해서는 빠르고 강한 템포의 록(Rock) 장르 음악을, 기쁨/행복을 유지하기 위해서는 우리가 흔히 듣는 팝(Pop) 장르 음악을, 흥분/분노를 완화하기 위해서는 느리고 차분한 템포의 클래식(Classic) 장르 음악을 선정하여 인터랙티브 콘텐츠 개발에 활용한다.

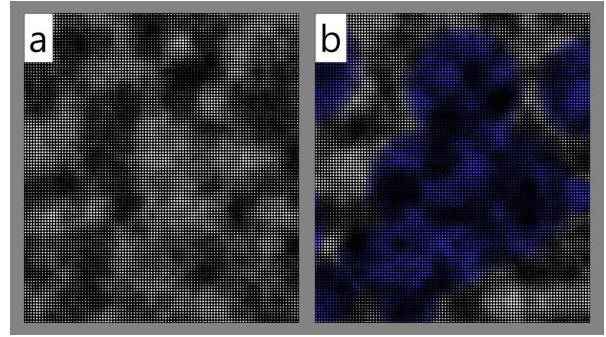


그림 2. 인터랙티브 콘텐츠 화면  
Fig. 2. Media art screen

그림 2는 인터랙티브 콘텐츠 화면 구성을 보여준다. 망점 효과를 사용하여 다양한 감정을 드러내며, 체험자가 맥박을 측정하기 전에는 a와 같이 현대인의 억압된 감정을 상징하는 무채색 화면이 나타난다. 이후 맥박 측정을 통해 드러난 체험자의 감정이 우울/불안일 때는 활기를 북돋는 빨간색을, 기쁨/행복일 때는 초록색을, 흥분/분노일 때는 마음을 안정시키는 파란색을 활용하여 화면을 b처럼 전환한다.

## III. 결 론

본 연구에서는 현대인의 억압된 감정을 효과적으로 조절할 수 있는 콘텐츠를 제작하였다. 시각적으로 제공된 인터랙티브 콘텐츠를 통해 체험자는 자신의 감정을 직관적으로 이해할 수 있고, 감정 상태를 치유하는 음악을 들으며 정신 건강을 증진할 수 있을 것이다. 향후 연구에서는 생성형 인공지능을 활용하여 다채로운 화면 구성이나 사운드를 추가할 것이다.

## 참 고 문 헌

- [1] <https://mdtoday.co.kr/news/view/179581555128280>  
[Assessed 2024-10-27]
- [2] Editorial team, “[Brain Traffic Light] The Waves of Emotions, Control Needed Rather than Suppression,” *Brain*, Vol. 59, Korea Institute of Brain Science, pp. 59, 2016.

# 유니티와 웹캠을 활용한 인터랙티브 미디어 리터러시 콘텐츠

김혜린\*, 안승현\*, 이세진\*\*, 유선진\*\*\*

## Interactive Media Literacy Contents using Unity and WebCam

Hyelin Kim\*, Seunghyun Ahn\*, Sejin Lee\*\*, and Sunjin Yu\*\*\*

### 요 약

최근 가짜뉴스와 다크패턴 등이 사회문제로 대두됨에 따라, 미디어 리터러시는 현대인의 필수 역량으로 요구된다. 이에 본 논문에서는 미디어 리터러시와 다크패턴의 개념을 접목한 인터랙티브 콘텐츠를 제작하였다. 유니티(Unity)를 활용해 사용자의 선택을 방해하도록 조작된 가상 웹페이지 환경을 구현하고 사용자의 무의식적인 반응을 고찰한다. 이후 웹캠을 통해 사용자의 모습을 기록하여 미디어로 인한 감시와 조종의 과정을 드러낸다. 그 결과 사용자는 미디어의 영향력을 경험하며 미디어 리터러시 능력의 중요성을 인지할 수 있다.

### Abstract

As fake news and dark patterns have recently emerged as social problems, media literacy is required as an essential competency for modern people. Accordingly, in this paper, interactive content was created by combining media literacy and the concept of dark patterns that deceive users. In a web page using Unity, a virtual environment is created to interfere with the user's choices and the user's unconscious reactions are considered. After, the user's appearance is recorded through webcam, revealing the process of surveillance and manipulation by the media. As a result, users can experience the influence of media and recognize the importance of media literacy skills.

### Key words

media literacy, dark patterns, unity, webcam, manipulation

## I. 서 론

현대사회에서는 정보과부하 속 가짜뉴스와 다크패턴 등이 문제로 떠오르며, 미디어 리터러시 교육의 중요성이 대두되고 있다. 미디어 리터러시란, 미

디어를 다양한 관점에서 분석하여 이해하는 능력을 의미한다[1]. 특히 미디어 리터러시가 부족한 사람들은 사용자의 의도와 상관없이 특정 행동을 유도하거나 실수를 유발하도록 디자인된 다크패턴에 더 쉽게 속을 수 있다[2].

\* 국립창원대학교 문화테크노학과, rinz20200593@gs.cwnu.ac.kr, 20200257@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화융합기술협동과정 석사과정, 20237203@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr (교신저자)

※ 본 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.

이러한 연관성을 활용해 본 논문에서는 사용자에게 의도적으로 다크패턴이 주는 불편함을 느끼게 하고 미디어 리터러시의 중요성을 인지할 수 있는 인터랙티브 콘텐츠를 제안한다.

## II. 인터랙티브 웹페이지 제작

본 콘텐츠의 콘셉트는 미디어 리터러시의 역할을 조망하는 가상의 사회모임 ‘미소클럽’이 운영하는 ‘프리덤 프로젝트’이다. 다크패턴 요소를 넣은 웹페이지와 웹캠 및 빔프로젝터 콘텐츠로 구성되며 개발 플랫폼으로는 유니티(Unity) 엔진을 활용한다.

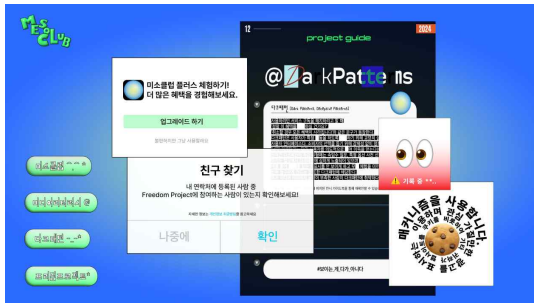


그림 1. 가상 유니티 웹페이지 화면  
Fig. 1. Virtual web page scene

유니티의 Canvas와 Button UI를 활용하여 웹페이지 화면과 팝업창을 구현한다. 사용자는 개인정보 수집에 동의한 후 ‘미소클럽’의 회원으로 가입하며 자유롭게 웹페이지를 탐색할 수 있다. 버튼을 클릭하면 각각의 폴더와 관련된 설명 팝업창이 등장한다. 5초 후 그림 1과 같은 4개의 새 팝업창이 화면에 생성된다. 이는 사용자의 주의를 분산시키고 시각적 방해물 불러일으키는 다크패턴 요소로 작용한다. 모든 팝업창이 표시되면 숨겨진 비상구 폴더 아이콘이 활성화된다. 이는 팝업창을 지울 수 있는 탈출구로 가장하여 사용자의 클릭을 유도한다.

클릭 시 기존 웹페이지 씬은 사라지고 데스크탑과 연결된 웹캠 씬으로 전환된다. WebCamTexture를 활용하여 실시간으로 웹캠을 사용할 수 있는 환경을 구성한다. 촬영 버튼을 누르면 3초 후 촬영이 이루어진다. 촬영된 이미지는 새 오브젝트로 생성되고 화면에 무작위로 배치된다. 이때 System.Date.No

w 클래스를 이용해 촬영 당시의 날짜 및 시간을 이미지 우측 상단에 표시한다.



그림 2. 웹캠 촬영 사진 누적 화면  
Fig. 2. Scene after webcam shooting

그림 2는 웹캠으로 촬영된 사진이 누적된 모습이다. 해당 화면은 외부 빔프로젝터를 통해 송출된다. 이는 사용자가 미디어를 벗어나더라도 개인과 타인의 기록이 계속 남아 있다는 것을 시각적으로 보여주는 화면이다. 또한, 미디어 속에서 한번 생성된 자료는 쉽게 삭제되지 않는다는 의미를 내포한다.

## III. 결 론

본 논문에서는 유니티와 웹캠을 기반으로 다크패턴 요소를 적용한 인터랙티브 미디어 리터러시 콘텐츠를 제안한다. 표면적인 자유 속에서 의도된 불편함을 경험함으로써 사용자는 미디어의 통제와 감시에 대한 경각심을 느끼고 미디어 리터러시의 중요성을 이해하게 된다.

향후 몰입감을 더하는 상호작용 요소를 추가해 미디어 리터러시 교육에 활용할 수 있는 콘텐츠로 발전시킬 것이다.

## 참 고 문 헌

- [1] J. Y. Yum and S. H. Jung, "Predictors of fake news exposure and sharing", Korean Journal of Journalism and Communication Studies, pp. 19, Feb 2019.
- [2] B. Harry, "Deceptive Patterns: Exposing the Tricks Tech Companies Use to Control You", across book, April 2024.

# Arduino와 Unity를 활용한 심리 콘텐츠 개발

이경민\*, 변공규\*\*, 유선진\*\*\*

## Development of Psychological Content using Arduino and Unity

GyeongMin Lee\*, Gongkyu Byeon\*\*, and SeonJin Yu\*\*\*

### 요 약

불안감은 현대인이 느끼는 요소 중 빠질 수 없는 요소이며 사회에서 꾸준히 나오는 주제이다. 그러나 자신의 감정을 제대로 다루지 못해 불안감에 시달리는 시민들이 늘어나고 있다. 따라서 본 논문에는 Unity 엔진과 Arduino를 활용해 불안감을 시각적으로 표현하여 각자의 불안감의 정도를 눈으로 보고 자신에 대해 생각해 보는 시간을 가질 수 있는 콘텐츠를 제안한다.

### Abstract

Anxiety is an essential element that modern individuals experience and is a consistently recurring theme in society. However, there is a growing number of citizens who suffer from anxiety due to their inability to properly manage their emotions. Therefore, this paper proposes content that visually expresses anxiety using the Unity engine and Arduino, allowing individuals to see the degree of their own anxiety and take time to reflect on themselves.

### Key words

arduino, unity, mentality, psychological test

## 1. 서 론

2021년 정신건강실태조사에 따르면 우리나라 국민의 정신질환 유병률은 27.8%로, 정신건강 문제를 겪는 사람이 10명 중 3명에 가깝다[1]. 극심한 고용 불안과 경제난이 이어지면서 우울증과 불안감에 시달리는 시민들이 늘어나고 있다. 통계청이 발표한 '2023년 4월 고용동향'에 따르면 청년층 취업자는

13만 7000명 줄어들며 6개월 연속 하락했다[2]. 앞으로 건강한 사회를 위해서라도 자신의 감정을 제대로 이해할 수 있어야 한다. 이를 위해선 불안감이라는 무형적인 요소를 시각적으로 표현해낼 수 있어야 한다. Arduino는 보드의 입력을 읽고 모터를 작동시키거나 LED를 켤 수 있다[3]는 점에서 사용자의 감정을 공간에 시각적으로 표현하고자 하는 목적과 알맞다. 본 논문에서는 이러한 Arduino의 특

\* 국립창원대학교 문화테크노학과, iy1615@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화융합기술협동과정 박사과정, bgk9517@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr (교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.

성을 활용하여 관람객들은 자신의 불안을 시각적으로 표현해보고 자신에게 필요한 말은 무엇인지 알아가는 경험을 가질 수 있도록 한다.

## II. 심리 테스트 아두이노 콘텐츠

본 콘텐츠는 Unity와 Arduino를 기반으로 개발하였다. Unity로 심리 테스트를 만들고 Arduino를 활용하여 결과를 시각적으로 표현하였다. 또한 Unity AR을 활용하여 결과 카드를 만들었다. 체험 진행 시, 사용자는 미리 준비해둔 작은 공간에 들어가 앉아 심리 테스트의 첫 화면에 시작하기 버튼을 누르며 체험이 시작된다.

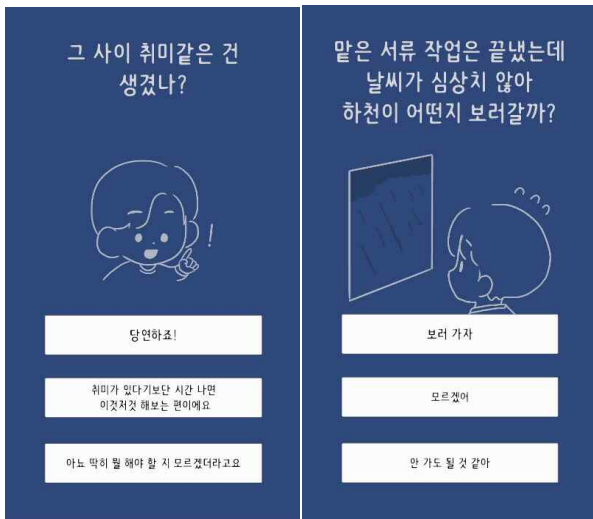


그림 1. 심리테스트  
Fig. 1. Psychological test

그림1은 사용자가 볼 심리 테스트 화면이다. 불안감이라는 말을 직접적으로 사용하지 않고 스토리텔링을 활용해 사용자의 자연스러운 선택을 이끌어 낼 수 있도록 하였다. 또한 사용자의 생생한 체험을 위해 콘텐츠를 진행하는 동안 배경 음악이 나오도록 하여 이야기에 집중할 수 있도록 하였다.

콘텐츠가 시작되면 사용자는 미리 만들어 놓은 공간으로 들어간 후 심리 테스트를 진행한다. 안에는 어떤 조명도 설치하지 않아 빛을 최대한 차단하고, 자신이 선택한 선택지에 따라 체험자 오른쪽에 위치한 수조에 물이 차오르도록 한다. 이를 통해 심

리 테스트의 결과를 시각적으로 볼 수 있게 하여 콘텐츠에 대한 몰입도를 증가시킨다. ‘당신의 번호는 5번입니다!’와 같이 숫자 1부터 5까지의 결과가 화면에 뜨는 것으로 심리 테스트가 종료된다.

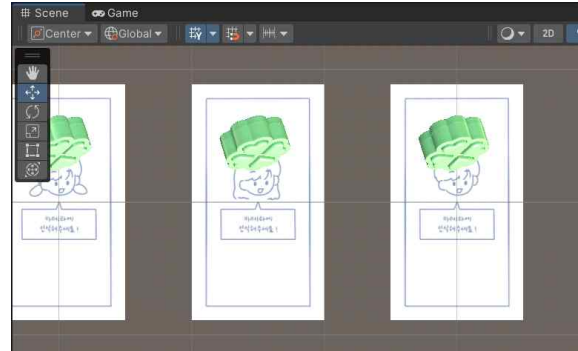


그림 2. 이미지 타겟  
Fig. 2. Image target

이후 공간 바깥에 준비되어있는 카드 중 결과로 본 숫자가 적혀있는 카드를 가져간다. 준비된 패드를 카드에 인식하면 클로버와 각 결과 단계에 필요한 말들이 그림 2처럼 떠오른 모습을 볼 수 있다.

## III. 결 론

본 논문에서는 Arduino와 Unity를 활용한 심리 콘텐츠를 제안하였다. 제안된 방법은 제한된 공간에서 체험자 혼자 체험하며 직접 눈으로 불안감을 볼 수 있도록 한다. 이를 통해 자신의 불안감이 어느 정도인지 알아보며 자신을 돌아보는 데 도움을 줄 수 있다. 향후 이 연구에서는 AI를 이용하여 사용자가 체험한 내용을 참고해 불안감에 대한 데이터를 수집하여 콘텐츠의 다양성을 발전시킬 것이다.

## 참 고 문 헌

- [1] <http://www.ujeil.com/news/articleView.html?idxno=355744>  
[Accessed 2024.10.28.]
- [2] <https://www.1conomynews.co.kr/news/articleView.html?idxno=23562>  
[Accessed 2024.10.28.]
- [3] <https://www.arduino.cc/en/Guide/Introduction>  
[Accessed 2024.10.27.]

# 핵심 어구 추출을 위한 언어 모델의 프롬프트 정제 기법 제안

박선화\*, 권수민\*, 정유철\*\*

## Proposal of a Prompt Refinement Technique for Language Models for Keyphrase Extraction

Seonhwa Bak\*, Sumin Kwon\* and Yuchul Jung\*\*

### 요약

핵심 어구 추출 문제는 뉴스나 논문 등의 텍스트에서 오랜 기간 연구되어 왔지만, 다른 자연어 처리 분야에 비해 상대적으로 낮은 성능을 보이고 있다. 본 연구에서는 기존 UKE(Unsupervised Keyword Extraction) 접근법에서 벗어나, 비지도 학습 기반의 LLM을 활용하여 과학 및 컴퓨터 분야 논문에서 핵심 어구를 효과적으로 추출하는 프롬프트 정제 기법을 제안한다. 이를 위해, 다양한 모델에 개별 프롬프트를 적용하여 추출된 핵심 어구의 F1 Score를 측정함으로써 Prompt와 LLM 간의 관계를 파악하는 실험을 진행한다. 실험을 통해 각 모델을 대상으로 핵심 어구 추출 성능이 가장 높은 LLM-Prompt 쌍을 선정하였으며, 평균적으로 20% 이상의 성능 향상을 확인하였다. 이러한 연구 결과를 토대로, 본 연구에서는 언어 모델의 핵심 어구 추출의 정확도를 높이기 위한 프롬프트 정제 기법을 제안한다.

### Abstract

Although the problem of keyphrase extraction has been studied for a long time in texts such as news articles and academic papers, it has shown relatively lower performance compared to other natural language processing fields. In this study, we propose a prompt refinement technique that effectively extracts keyphrases from scientific and computer science papers by utilizing unsupervised learning-based large language models (LLMs), moving away from existing unsupervised keyword extraction (UKE) approaches. To this end, we conduct experiments to understand the relationship between prompts and LLMs by applying individual prompts to various models and measuring the F1 scores of the extracted keyphrases. Through these experiments, we selected the LLM-prompt pairs with the highest keyphrase extraction performance for each model and confirmed an average performance improvement of over 20%. Based on these research findings, we propose a prompt refinement technique to enhance the accuracy of keyphrase extraction using language models.

### Key words

sLM, LLM, keyword extraction, prompt, f1 score, UKE

---

\*국립금오공과대학교 컴퓨터공학과 박선화 paksh0725@kumoh.ac.kr

\*국립금오공과대학교 인공지능공학과 권수민 20230056@kumoh.ac.kr

\*\*국립금오공과대학교 인공지능공학과 정교수 정유철 jyc@kumoh.ac.kr

※ The work is supported by the Korea Basic Science Institute (Project No.A438200)

## 1. 서 론

생산되는 문서 자원의 양이 지속적으로 증가함에 따라, 그로부터 유용한 정보를 효과적으로 추출하는 것이 중요한 과제로 부상하고 있다. 특히 핵심 어구는 문서의 의미를 간결하게 나타내는 핵심적인 역할을 하며, 정보 검색, 요약, 문서 분류, 주제 탐지, 지식 그래프 등 다양한 자연어 처리 분야에서 유용한 메타 정보로 활용되고 있다. 기존의 핵심 어구 추출 기법은 비교사 학습과 교사 학습으로 구분되었다. 비교사 학습에서는 TF-IDF와 같은 통계 모델이나 TextRank, SingleRank, TopicRank 등의 그래프 모델이 주로 사용되었다. 교사 학습에서는 은닉 마르코프 모델, 최대 엔트로피 마르코프 모델, 조건부 랜덤 필드 등의 알고리즘이 적용되었으며, 딥러닝 기술의 발전으로 양방향 LSTM 네트워크 등이 도입되어 성능 향상을 이루었다.

그러나 딥러닝 기법 중에서 대형 언어 모델을 활용한 핵심 어구 추출에 대한 연구는 그 잠재력이 충분히 탐구되지 않았다. 기존의 키워드 추출 모델과 달리, LLM을 활용한 핵심 어구 추출은 Prompt 설계와 모델에 따라 결과가 크게 달라질 수 있어, 이에 대한 체계적인 연구가 필요하다.

본 연구에서는 이러한 배경을 바탕으로, 기존의 UKE(Unsupervised Keyword Extraction) 접근법에서 벗어나 비지도 학습 기반의 LLM을 활용하여 과학 및 컴퓨터 분야 논문에서 핵심 어구를 효과적으로 추출하는 프롬프트 정제 기법을 제안한다. 본 연구를 통해, 프롬프트 정제만으로도 다양한 데이터에서 LLM의 핵심 어구 추출 능력을 높일 수 있음을 확인한다.

## II. 실 험

### (1) 실험 개요

본 연구에서는 언어 모델의 핵심 어구 추출 성능에 영향을 미치는 주요 요소로 언어 모델의 사전 학습 지식과 프롬프트 설계를 선정하였다. 언어 모델의 사전 학습 데이터는 모델의 답변 내용과 지식

수준에 직접적인 영향을 미치며, 프롬프트는 모델의 응답 품질과 정확성에 큰 영향을 준다. 따라서, 이 두 가지 요소에 따른 성능 개선을 확인하기 위한 실험을 설계하였다. 실험 대상의 모델로는 Llama2, Solar, Gemma 모델을 선정하였다. 각 모델은 사전 학습 데이터의 양과 특성이 상이하며, 이는 모델의 응답 일관성과 정확성에 영향을 준다. 각 모델의 특성은 아래 [표 1] 과 같다.

표 1. 모델 특성

Table 1. Model Characteristics

모델	특성
Llama2	대규모 웹 크롤링 데이터와 학술 논문/도서를 학습한 모델
Solar	GPT 모델을 활용하여, 정제된 데이터로 학습한 DPO 모델
Gemma	다양한 언어 스타일과 광범위한 웹 데이터로 학습된 모델

프롬프트 설계 측면에서는 기본 프롬프트에 Content Control, CoT(chain-of-Thought), Step-by-Step, Output Formatting, Structured Prompting 기법을 다중 구조로 추가하여 모델의 응답을 유도하였다. 특히 각 모델의 성능을 높이기 위해, 대상 모델에 맞춰 3 가지 프롬프트를 기획하였다.

### (2) 실험 방법

실험에는 핵심 어구 추출 분야의 공개된 벤치 마크 데이터인 KP20K, Karpivin, Inspec, Semeval를 사용하여 성능을 측정하였다.

표 2. 실험 데이터 세트

Table 2. Experimental data sets

데이터명	분야	데이터 개수
KP20K	과학	500
Karpivin	과학	500
Inspec	과학	500
Semeval	학술 논문	500

이후, 5가지 방식의 Prompt를 적절하게 배열하여 핵심 어구 추출을 진행한다. 5가지 방식의 Prompt는 아래 [표 3]과 같다.

표 3. 프롬프트 기법  
Table 3. Prompt Technique

데이터명	데이터 개수
Basic	목적과 결과 예시로만 구성된 Zero-shot Prompt
Content Control	핵심 어구 선택 시 유의할 점을 명시하여 모델의 출력 내용을 통제
Output Formatting	응답 형식을 제시하여 모델의 출력을 제한
CoT(chain-of-Thought Prompting)	작업 수행을 위한 단계별 사고 과정을 안내하여 모델의 논리적 흐름을 제어
Structured Prompting	모델의 태그를 이용하여 체계적으로 사고 과정을 지정
Step-by-Step	단계별로 응답 생성을 제시

위 기법을 다중 구조로 작성하여, 복합 Prompt를 통해 모델을 제어하며 실험을 진행하였다. 최종 작성된 Prompt와 기법은 아래와 같다.

표 4. 실험 프롬프트  
Table 4. Experimental Prompt

	Prompt	목적	사용 기법
1	prompt = """ [Detailed prompt text for prompt 1]"""	논문의 내용 및 주제 파악	1. Step-by-Step 2. Output Formatting 3. Content Control
2	prompt = """ [Detailed prompt text for prompt 2]"""	결과 형식 및 사고 과정 지정	1. Structured Prompting 2. Output Formatting
3	prompt = """ [Detailed prompt text for prompt 3]"""	결과 내용을 분석하고 순위 지정	1. Chain-of-Thought 2. Content Control 3. Output Formatting

실험은 아래의 [그림 1]과 같이 데이터 세트를 Alpaca-lora의 instruction 형태로 구조화한 이후, 논문의 초록과 Prompt를 조합하여 모델의 입력으로 사용한다. 이후 모델은 초록에서 핵심 어구를 추출하고, 이를 자동 정제하여 평가 과정으로 넘긴다. 이

후 원본 데이터의 정답 핵심 어구와 추출 핵심 어구를 F1 Score와 Rouge를 사용하여 평가한다.

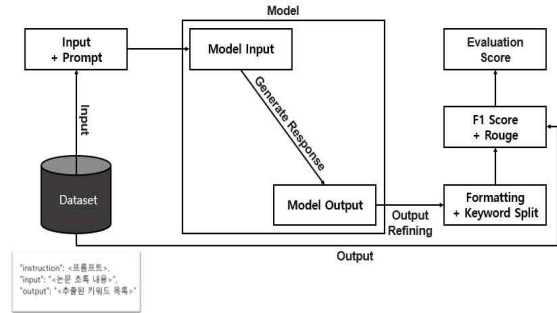


그림 1. 실험 방법  
Fig. 1. Experiment method

Rouge와 F1 Score를 통해, 하나의 핵심 어구 당 70% 이상의 일치도를 가지고 있으면 정답으로 인정하는 방식의 평가를 통해, 평가 모델이 정답에 대해 유연하게 평가할 수 있도록 한다. 이후 위와 같은 과정을 반복하여, 평균 성능을 확인한다.

(3) 실험 결과

실험 시, 결과 형식과 사고 과정을 지정해주었던 Prompt-2의 경우, 다량의 보고서/논문으로 학습된 Llama2와 정제된 데이터를 통해 학습된 Solar 모델에서 가장 높은 정확도를 기록하였다. 논문의 내용 및 주제 파악에 초점을 두었던 Prompt-1의 경우, 다양한 언어 스타일과 문맥 이해에 초점을 둔 Gemma가 높은 정확도를 보였으나, Solar 모델은 가장 낮은 평균 정확도를 기록하였다. 결과 내용을 분석하고 순위를 지정하는 Prompt-3의 경우, Gemma와 Llama2 모델에서 가장 낮은 정확도를 기록하였다. 각 Prompt에 대한 평균 정확도 결과는 아래 [표 5]와 같다.

표 5. 실험 결과  
Table 5. Experimental Result

모델 명 / 평균 정확도	1번 Prompt	2번 Prompt	3번 Prompt
Llama2	12.54%	58.52%	9.59%
Solar	45.80%	53.17%	46.80%
Gemma	43.45%	39.46%	9.77%



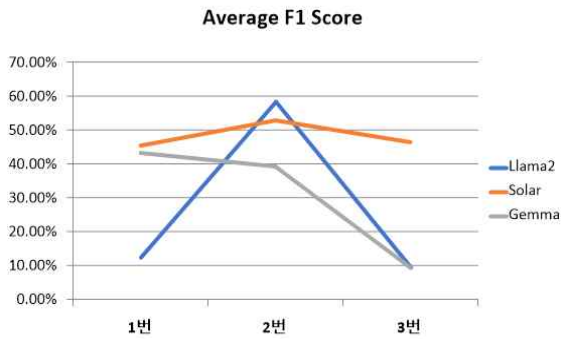


그림 2. 실험 결과  
Fig. 1. Experimental Result

이러한 실험 결과를 통해, 프롬프트와 언어 모델의 조합이 핵심 어구 추출 성능에 큰 영향을 미침을 확인하였다. 또한, 각 모델의 특성을 반영하여 Prompt를 정제하는 것만으로도 높은 성능을 기대할 수 있다는 사실을 확인하였다.

### III. 결 론

실험을 통해 프롬프트 설계가 언어 모델의 핵심 어구 추출 성능에 중요한 영향을 미친다는 것을 확인하였다. 모델별로 최적화된 프롬프트를 적용함으로써 평균적으로 20% 이상의 F1 Score 향상을 달성하였다. 또한, 언어 모델의 사전 학습 데이터의 특성과 프롬프트의 조합이 모델의 응답 품질에 큰 영향을 미친다는 것을 발견하였다. 이는 모델의 사전 학습 지식이 프롬프트에 의해 효과적으로 활용될 수 있음을 시사한다.

결론적으로, 본 연구를 통해 언어 모델의 핵심 어구 추출 정확도를 높이기 위해서는 모델의 특성에 맞는 프롬프트 정제가 필수적임을 확인하였다. 이는 향후 자연어 처리 분야에서 대형 언어 모델을 활용한 다양한 응용에 있어 중요한 시사점을 제공할 것으로 기대된다.

### 참 고 문 헌

[1] Rui Meng, Sanqiang Zhao, Shuguang Han, Daqing He, Peter Brusilovsky, Yu Chi, "Deep Keyphrase Generation", May 2021

[2] Rui Liu<sup>1,2</sup>, Zheng Lin<sup>1</sup>, Weiping Wang<sup>1</sup>, "Keyphrase Prediction With Pre-trained Language Model", Apr 2020  
[3] Hugo Touvron\*, Thibaut Lavril, etc.. "LLaMA: Open and Efficient Foundation Language Models", Feb 2023



# 도시열섬현상 완화를 위한 그린인프라 공간 최적화

배성민\*, 이서영\*\*, 서지훈\*\*\*

## Optimization of Green Infrastructure Spaces for Mitigating Urban Heat Island Effects

Seongmin Bae\*, and Seoyoung Lee\*\*, and Jihoon Seo\*\*\*

### 요 약

한국의 수도권 도시들은 급속한 성장으로 인한 환경문제 해결에 노력하고 있다. 특히 서울시는 높은 인구밀도에 비해 녹지공간 부족 등으로 도시열섬현상이 발생하고, 일부 지역은 폭염과 관련된 질병 및 사망 사례를 접할 수 있다. 이에 우리는 폭염 위험도를 구성하는 환경적, 효율적, 취약적 변수들을 분석해 도시 녹지 시설을 정량화하고, 이를 바탕으로 서울시 내 폭염 취약 지역을 시각적으로 특정지었다. 이후 입지 추천 알고리즘을 사용해 최적의 그린인프라 입지를 거시적 관점에서 제안하였다.

### Abstract

In the metropolitan area of South Korea, cities are working hard to tackle environmental issues caused by rapid growth. Seoul, in particular, has a shortage of green space compared to its high population density, which leads to urban heat island effects and, in some areas, heat-related illnesses and deaths. To address these challenges, we analyzed various factors contributing to heat risk, measured urban green infrastructure, and visually identified heat-vulnerable areas in Seoul. As a result, we recommend optimal locations for green infrastructure in these areas using a location recommendation algorithm.

### Key words

Urban Heat Island, Green Infrastructure, Apparent temperature, Heat Index

### 1. 서 론

국내 수도권은 급격한 도시화와 서울시의 인구 집중화로 인해 도시열섬현상이 뚜렷하게 나타난다 [1]. 2023년 기준 서울시 인구는 9,400, 인구밀도는 15,533(단위: 천명, 명/km)으로 집계되었다. 이 수치

는 인구밀도가 2번째로 높은 부산보다 대략 4배의 약간 미치지 못하는 수준이었다[2]. 도시화와 인구 밀도가 높아지면 자연스럽게 콘크리트 및 아스팔트를 이용한 인공구조물과 포장된 도로가 많아지게 되고, 건폐율 또한 상승한다. 고층빌딩과 빽빽하게 밀집된 건물들은 열 흡수 면적을 증가시키고, 도심

\* 강남대학교 데이터사이언스전공, 201904207@kangnam.ac.kr,

\*\* 강남대학교 데이터사이언스전공, pinklee831@kangnam.ac.kr,

\*\*\* 강남대학교 데이터사이언스전공, Email: jihoon@kangnam.ac.kr(교신저자)

내 바람길을 가로막아 대기의 순환을 저해시킨다. 특히나 풍속이 낮고 기온이 높은 여름철에는 뜨거운 공기의 순환이 더욱 어려워져 기류 정체 현상이 심화되고, 이는 도시열섬현상으로 이어지게 된다[3]. 반면, 그린인프라는 도시 내 녹지 네트워크를 형성함으로써 바람길의 형성과 순환에 도움을 줄 수 있고, 대기오염 저감으로 도시열섬현상에 냉각효과를 주게 된다. 또한, 도시열섬현상 완화 외 공기 정화, 수량 및 수질 보호, 서식지 제공 및 생태계의 다양성 보전 등의 생태적 효과가 있으며, 사회적 소통 및 건강 증진, 범죄 감소 등의 사회적 효과를 미친다. 그리고 지가상승, 저에너지화, 건설비용 감소 등의 경제적 효과가 있을 수 있다[4]. 이처럼 그린인프라는 인간과 자연이 공존할 수 있는 공간계획의 중요한 수단으로 도시의 지속가능성을 높여준다. 이에 우리는 서울시 도시열섬현상을 완화하기 위한 최적의 그린인프라 입지를 제안하고자 한다.

## II. 본 론

본 연구의 입지 추천 알고리즘의 고려 요인으로는 효율적 변수인 유동 인구나 취약적 변수인 사회적 혹은 신체적 약자가 많으며, 환경적 변수인 열지수와 건폐율이 높은 지역을 기준으로 진행하였다. 입지 추천에 사용된 알고리즘은 서로 다른 목표를 가지는 두 알고리즘의 시나리오를 비교 분석해 최종 공간 최적화를 진행하였다(그림 1 참조).

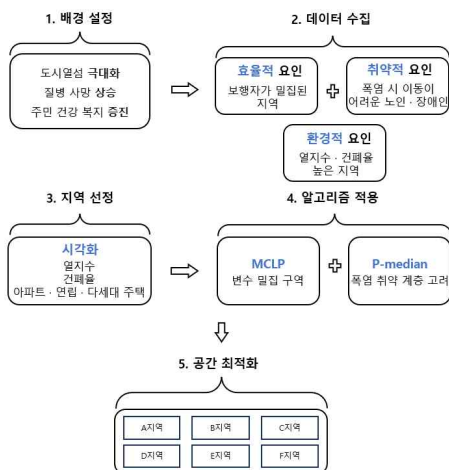


그림 1. 흐름도  
Fig. 1. Flow chart

효율적 변수는 2023년 기준 서울시 지하철역과 버스 정류장 위치 데이터를 사용하였고, 취약적 변수는 2022년 기준 서울시 사회(노인)복지시설 위치 데이터를 사용하였다. 환경적 변수는 서울시의 2024년 5~6월 온도, 습도 데이터와 2022년 기준 서울시 건폐율과 아파트 및 연립주택 데이터를 사용하였다. 마지막으로 알고리즘 비교 분석과 최종 공간 최적화를 기존의 공원 및 녹지대를 기반으로 하고자 2021년 서울시 공원 데이터와 2022년 녹지대 데이터를 추가로 사용하였다. 입지 추천지를 분석할 지역은 환경적 변수를 활용해 서울시의 열지수, 건폐율, 열지수와 건폐율이 모두 높은 행정동 각 5곳을 선정해 도시열섬현상이 뚜렷한 지역 총 15곳을 분석하였다. 열지수는 온도와 습도를 함께 나타내기 위한 것으로 사람이 실제로 느끼는 더위를 수치화한 값을 의미한다.

### 1) MCLP(Maximal covering location problem)

MCLP 알고리즘은 주어진 시설이나 서비스를 정해진 범위 내 최대한 많이 커버하는 위치를 찾는 것이 목적이다. 알고리즘의 구조는 입지 후보지 생성 함수와 메인 입지 추천 함수 2가지로 설계했으며, 그림 1의 효율적 변수와 취약적 변수의 좌표값을 하나로 병합해 일종의 수요지 포인트로 저장시켜주었다. 입지 후보지 생성 함수는 입력값으로 들어오는 모든 수요지 포인트를 포함하는 가장 작은 다각형 경계(Convex Hull)를 생성해준다. 그런 다음 해당 경계선 내의 무작위로 X, Y좌표를 추출해 랜덤 입지 후보지 100개를 생성해준다. 메인 입지 추천 함수에서는 미리 생성된 입지 후보지 100개의 좌표를 입력으로 받아 각 후보지와 수요지 포인트 간의 거리를 계산해 정해진 범위 내에 있으면 1, 아니면 0으로 Matrix 형태로 표시된다. 최종 출력되는 입지 추천지는 지역 크기에 따라 2개 또는 3개로 지정해주었고, 해당 제약조건을 충족하도록 지역마다 커버 수요지가 최대화되는 2, 3곳만 최종 입지 추천지가 되도록 설계하였다. 본 연구에서 입지 추천지의 범위는 선행연구를 바탕으로 냉각범위가 최대 100~150m이며[5], 녹지대와 녹지대 간의 거리를 고려하여 반지름을 대략 260m로 설정해주었다. MCLP 최종 입지 추천지 결과는 아래 표 1과 같다.

표 1. MCLP 시각화

Table 1. MCLP visualization

■ 공원 및 녹지 
 ■ 버스 정류장 
 ■ 지하철역 
 ■ 사회(노인)복지센터

열지수 상위 5				
송파구 A1동	종로구 B1동	종로구 C1동	성동구 D1동	성북구 E1동
건폐율 상위 5				
중구 A2동	중구 B2동	종로구 C2동	중구 D2동	중구 E2동
열지수 + 건폐율 상위 5				
마포구 A3동	강북구 B3동	구로구 C3동	관악구 D3동	관악구 E3동

## 2) P-Median

P-Median 또한 입지 선정과 공간 최적화 알고리즘으로 자주 사용되는데, MCLP와 목적이 살짝 다르다. P-Median은 주어진 수요지 포인트들과 후보지 사이의 총거리가 최소화되는 지점을 찾아 MCLP보다 거리를 더 크게 중요시하여 효율적인 위치를 찾는 것이 목적이다. 알고리즘의 구조는 MCLP와 마찬가지로 입지 후보지를 생성하는 함수와 메인 입지 추천 함수 2개로 설계하였다.

입지 후보지 생성부분은 MCLP와 동일한 방식으로 다각형 경계(Covex Hull)를 만들어준다. 메인 입지 추천 함수의 목적함수는 다르게 설정하였는데,

MCLP에서는 불리언 행렬에서 참(True)이 가장 많이 포함되도록 설계했지만, P-Median의 목적함수에서는 거리행렬 D의 합이 최소가 되도록 설계해주었다. 그리고 P-Median에서는 범위가 아닌 거리만을 이용해 계산하기 때문에 반지름을 입력값으로 넣지 않았고, 시각화 시에만 해당 입지의 냉각범위를 똑같이 대략 260m로 설정해주었다. 또한, 기존의 P-Median 알고리즘에서는 각 수요지 포인트마다 가중치를 다르게 적용해서 거리행렬 계산에 반영시킬 수가 있지만, 우리는 모든 수요지에 같은 중요도를 두었기 때문에 가중치는 생략하였다. P-Median의 최종 입지 추천지는 아래 표2와 같다.

표 2. P-Median 시각화

Table 2. P-Median visualization

■ 공원 및 녹지 
 ■ 버스 정류장 
 ■ 지하철역 
 ■ 사회(노인)복지센터

열지수 상위 5				
송파구 A1동	종로구 B1동	종로구 C1동	성동구 D1동	성북구 E1동
건폐율 상위 5				
종구 A2동	종구 B2동	종로구 C2동	종구 D2동	종구 E2동
열지수 + 건폐율 상위 5				
마포구 A3동	강북구 B3동	구로구 C3동	관악구 D3동	관악구 E3동

### 3) 공간 최적화

마지막으로 그린인프라 입지 공간 최적화는 앞서 분석한 MCLP, P-Median을 비교 분석하여 서로 입지 추천지 범위가 겹치거나 인접한 부분을 찾아서 진행해주었다. 본 연구의 그린인프라 입지에서 고려 요인이었던 효율적, 취약적, 환경적 변수의 정보는 이미 분석을 통한 입지 추천지에 포함되어 있으므로 최종 공간 최적화에서는 기존의 그린인프라와의 간격과 겹치는 입지 추천지 범위만 고려해주면 되었다. 그리고 기준별 도시열섬현상이 가장 뚜렷한 2곳만 추려서 총 6곳을 표3과 같이 공간 최적화를 진행해주었다. 아래 해당 공간의 이미지는 구글맵에서 제공하는 스트리트뷰로 사진마다 촬영 연도가

상이할 수 있다. 공간 최적화된 6곳 모두 그린인프라가 들어가기에 최적의 장소로 인근 그린인프라와 의 생각범위도 고려할 수 있고, 가장 많은 수요지 포인트를 커버할 수 있고, 거리 또한 가까운 장소들이다. C와 E 지역은 주변에 아파트 단지가 있고, 넓은 장소가 있기 때문에 도시공원과 같은 그린인프라를 구축할 수 있을 것 같다. 하지만 나머지 지역은 모두 작은 빌라와 상권이 매우 빽빽이 밀집한 주거지 및 상업지역으로 도시공원이나, 숲길과 같은 그린인프라를 구축하기에는 어려움이 있어 보인다. 이처럼 밀집도가 높은 지역들은 옥상녹화 또는 벽면녹화 등으로 공간을 활용한 그린인프라를 구축해 에너지 절약과 도시열섬현상 완화가 가능할 것으로 예상된다.

표 3. 그린인프라 공간 최적화

Table 3. Optimization of Green Infrastructure Spaces

열지수 상위 2	
A. 	B. 
건폐율 상위 2	
C. 	D. 
열지수 + 건폐율 상위 2	
E. 	F. 

### III. 결 론

우리는 서울시의 도시화로 인한 도시열섬현상을 완화하고자 최적의 그린인프라 공간 최적화를 거시적으로 분석했으며, 도시열섬현상이 가장 높은 지역과 가장 낮은 지역을 파악할 수 있었다. 이후에는 더 다양한 고려 요인과 서울시 도시 정책을 고려해서 미시적인 관점의 분석이 필요할 것 같다.

### 참 고 문 헌

[1] K. I. Lee, J. Ryu, S. W. Jeon, H. C. Jung, and J. Y. Kang, "Analysis of the Effect of Heat Island on the Administrative District Unit in Seoul Using LANDSAT Image," *Korean Journal of Remote Sensing*, vol. 33, no. 5\_3, pp. 821-834, Oct. 2017. <https://doi.org/10.7780/kjrs.2017.33.5.3.6>

[2] Korea Statistics Portal. 지표누리. E-나라지표, "지역별 인구 및 인구밀도", 2024.

[https://www.index.go.kr/unity/potal/main/EachDtIPageDetail.do?idx\\_cd=1007](https://www.index.go.kr/unity/potal/main/EachDtIPageDetail.do?idx_cd=1007)

[3] E. J. Ji, D. B. Kim, Y. G. Kim, and J. A. Lee, "Simulation Analysis of Urban Heat Island Mitigation of Green Area Types in Apartment Complexes". *J Korean Inst Landsc Archit*, vol. 51, no. 3, pp. 153-165, Jun. 2023. <https://doi.org/10.9715/KILA.2023.51.3.153>

[4] Y. S. Seok, K. H. Song, H. J. Han, and J. A. Lee, "Derivation of green infrastructure planning factors for reducing particulate matter - Using text mining," *J. Korean Inst. Landsc. Archit.*, vol. 49, no. 5, pp. 79-96, 2021, <https://doi.org/10.9715/KILA.2021.49.5.079>

[5] J. W. Kang, B. C. Kim, "Analysis of Urban Heat Island(UHI) Alleviating Effect on Lands of Urban Parks in Seoul Using Artificial Neural Network(ANN) Model", University of Seoul. College of Urban Science. Online archive, 2020

# 감성분석에 기반해 대규모 언어 모델을 활용하여 게임 리뷰분석을 통한 개발 전략 수립

정환웅\*, 서지훈\*\*

## Formulating Development Strategies through Game Review Analysis Based on Sentiment Analysis and Large Language Models

Jeong Hwan Woong\*, and Seo Ji Hoon\*\*

### 요약

본 논문은 빠르게 발전하는 게임산업에 따라 게임 이용자들이 보는 게임에 대한 눈높이가 상승하고 있음을 인지하고 이에 따라 게임사들은 이러한 발전에 발맞춰 이용자들의 요구를 확인하고 지속적인 피드백을 통해 게임을 발전시켜야 할 필요가 있음을 확인하였다. 게임 이용자들의 생각과 느낌을 직접 확인할 수 있는 리뷰 데이터를 통해 이용자들이 어떤 관점에서 게임을 평가하는지 감성분석을 통해 확인하고 키워드 추출과 공용 단어 사전을 정립하고 각 키워드에 대한 이용자들의 감성 비율을 파악해 인사이트를 도출하였다.

### Abstract

The paper acknowledges that with the rapid development of the gaming industry, players' expectations for games have risen. Consequently, game companies must respond to these evolving demands and continuously improve their games based on player feedback. By analyzing game reviews, the study identifies the perspectives from which players evaluate games through sentiment analysis, extracting keywords and establishing a common word dictionary. It also determines the sentiment ratio for each keyword to derive insights into player perceptions.

### Key words

Sentiment Analysis, UX, Shallow BERT, Transformer, Keyword Extraction, Game DEV

### 1. 서론

게임산업은 국내외를 막론하고 빠르게 성장하고 있는 산업 중 하나이다. [1] 과거 게임은 단순히 오락이라는 개념으로 특히 연령층이 낮은 학생들에게

는 놀이로써, 어른들에게는 자식들의 성적에 방해물로써 인식되었었다. 하지만 현재 게임은 오락으로의 기능과 더불어 하나의 예술과 문화로써 자리매김하고 있다. 게임을 즐기는 연령대 또한 스마트폰의 보급으로 모바일 게임에 대한 접근성이 낮아지면서

\*강남대학교 데이터사이언스학과 학부생, tree193nn@kangnam.ac.kr(제 1저자)

\*\*강남대학교 데이터사이언스학과 조교수, Email : jihoon@kangnam.ac.kr(교신저자)



다양한 연령대가 게임을 즐길 수 있는 기반을 마련하였다. 이에 따라 게임을 플레이하는 연령대가 다양해지고 게임의 종류와 가지 수 또한 다양해지고 있다. 이제는 단순한 ‘재미’라는 요소로는 사람들은 게임을 플레이하지 않는다. 스토리, 콘텐츠, 최적화, 게임 제작사에 대한 인식 등 다양한 요소들이 게임을 플레이하는 요소로써 작용하고 있다.

본 논문에서는 빠르게 성장하고 변화하고 있는 게임산업에 대비해 이용자들의 게임 리뷰 감성분석을 통해 알아 보고 어떤 부분에서 이용자들이 긍정적인 평가를 남겼고, 어느 부분에서 부정적인 평가를 하였는지 확인하고 어떻게 개선할지를 확인해 볼 예정이다. 이는 게임별 리뷰의 키워드를 추출하여 확인하고 공용 단어 사전을 정립해 체계적인 분석을 하였다. 본 연구에선 중국 H사의 모바일 게임인 H 게임과 한국 N사의 L 게임 분석을 시행하였다. H사의 경우 지속적인 매출과 영업이익이 증가하고 있는 반면에 N사의 경우 주가 하락, 영업이익이 감소하고 있는 것을 확인하였다. 이러한 차이를 보이는 두 회사의 대표적인 게임들의 리뷰 내용을 비교 분석하여 어떤 요인에서 두 게임의 차이가 발생하는지 파악해 보았다.

## II. 관련 연구

### 2.1 분석계획

분석에 앞서 구글 플레이 스토어 내의 리뷰데이터에 접근하여 이용자가 남긴 평점, 받은 추천 수, 작성 날짜를 수집하여 각각 1,000개의 리뷰를 확보하였다. 구글 플레이 스토어의 리뷰는 구글 플레이 스토어에서 자체적으로 가지는 필터링 규칙을 통해 광고, 스팸성 리뷰를 제외하고 데이터를 제공하기에 분석에 있어서 양질의 데이터를 확보할 수 있었다. 현재 두 게임 모두 세계적으로 큰 인기를 보유하고 있는 게임이지만, 전체적인 운영, BM(수익 모델), 사용자 경험(UX)을 미루어 비교하면 두 게임에 대한 장단점을 파악할 수 있을 것이다. 또한 게임 플레이와 장르적인 면에서도 두 게임은 서로 다른 양상을 보인다.

## Sentiment Analysis

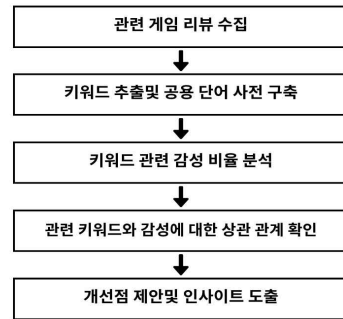


그림 1. 분석 흐름도  
Fig. 1. Analysis flowchart

감성분석에 사용할 모델은 CLOVA-X를 활용하였다. CLOVA-X는 한국어에 특화된 대규모 언어 모델로 불특정 다수가 작성한 리뷰 특성상 다듬어지지 않고 그 당시 느낀 바를 그대로 작성해 두서가 없을 수 있기에 내용의 맥락을 효과적으로 이해할 수 있도록 하여 대규모 언어 모델을 활용해 이러한 애로사항들을 해결할 수 있도록 시도하였다.

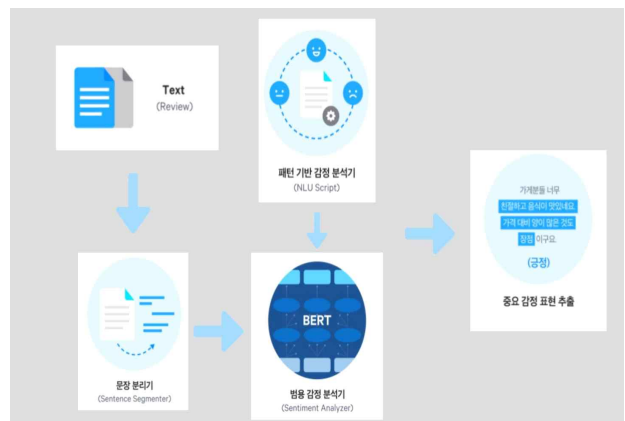


그림 2. Clova-Sentiment 작동 원리  
Fig. 2. Operating principle of Clova-Sentiment

Clova 감성분석[2]의 작동 원리는 수집한 리뷰 내용을 입력하여 텍스트를 정제하고 토큰화 및 형태소 분석을 통해 분석이 가능한 형태로 변환하여 해당 문장에 대한 감성분석을 진행한다. 이는 한국어에 특화된 Pre-Training을 수행하고 Shallow BERT를 적용하여 Attention 정보를 활용해 기존의 언어 모델

과 대비하여 빠른 분석 속도와 정확도를 보인다.  
[3]

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (1)$$

Shallow BERT는 Transformer의 기존 Self-Attention의 구조로 이루어져 있지만 문자들의 양방향 문맥 처리로 보다 정확한 문맥 이해가 가능하다. 이는 각 단어 간의 관계를 효율적으로 분석할 수 있다는 것에서 감성분석에서 큰 장점으로 꼽힌다. 또한 Shallow BERT는 기존의 Base BERT 비해 낮은 계산복잡도로 애플리케이션 같은 사용자와 실시간으로 상호작용을 해야 하는 작업에서 특히 강력한 성능을 보인다.

## 2.2 키워드 추출 및 공용 단어 사전 구성

수집한 리뷰의 명사들을 추출해 TF-IDF 기법으로 상위 키워드 5개를 수집하고 각각의 키워드가 포함된 리뷰를 CLOVA-X로 분석하여 감성 비율을 확인하고 해당 부분에 대한 이용자들의 반응을 살펴볼 것이다. 예를 들어 “과금”이라는 키워드에 이용자들의 감성 비율이 부정에 몰려 있다면 해당 요소에 대해 부정적인 감성을 가지고 있다고 추정한다. 본 연구에선 효과적인 분석을 위해 공용 단어 사전을 정립하였다.

표 1. 공용 단어 사전

Table 1. Common Word Dictionary

분류	단어
BM(수익 모델)	과금, 현질, 결제, 뽑기, 가차
게임 운영	업데이트, 버그, 서버, 이벤트, 패치
콘텐츠	재미, 스토리, 그래픽, 음악, 재미
사용자 경험(UX)	UI, 인터페이스, 조작감, 접근성, 직관, 디자인 화면, 난이도
기술적 성능	최적화, 렉, 끊김, 프레임, 배터리 소모, 튕김

공용 단어 사전은 이용자들이 게임을 플레이할 때 보편적으로 평가하는 요소들을 하나의 그룹으로 만들어 키워드의 집합으로 구성된 사전이다. 이는

게임 내 콘텐츠나 게임사의 운영에 대해 평가하는 기준으로 구성하여 해당 게임에 대한 전반적인 평가를 한 눈에 살펴볼 수 있다. 이러한 평가를 바탕으로 두 게임에 대한 감성 비율의 차이를 확인하고 분석한다.

## 2.3 키워드별 감성 비율 확인

먼저 H 게임에 대한 키워드를 추출하여 각 키워드에 대한 감성 비율을 확인하였다. 감성 비율의 경우 소수점 두 자리를 기준으로 반올림을 진행하였다.

표 2. H 게임 키워드 감성 비율

Table 2. H Game Keyword Sentiment Ratio

키워드	부정	중립	긍정
게임	69.18%	0.48%	30.33%
캐릭터	65.31%	0.76%	33.92%
번역	86.59%	0.77%	12.64%
스토리	67.76%	0.58%	31.65%
과금	59.25%	-	40.74%

H 게임의 경우 ‘게임’, ‘캐릭터’, ‘번역’, ‘스토리’, ‘과금’ 순으로 키워드가 추출되었다. 가장 높은 부정 비율을 보인 키워드는 ‘번역’으로 해외 게임의 특성상 현지어로 현지화 과정 중 번역에 관한 오류로 인해 이용자들이 부정적인 감성을 가지게 된 것으로 보인다. 가장 높은 긍정 비율을 보인 키워드는 ‘과금’이다. H 게임의 경우 기본으로 지급되는 캐릭터를 활용하여 게임 내의 콘텐츠와 스토리를 즐기기에 충분하다는 이용자들의 의견이 반영되었다.

표 3. L 게임 키워드 감성 비율

Table 3. L Game Keyword Sentiment Ratio

키워드	부정	중립	긍정
게임	82.95%	4.60%	12.45%
과금	78.90%	2.22%	18.89%
리니지	77.78%	3.50%	18.71%
유저	87.10%	1.61%	11.29%
접속	89.47%	8.95%	1.58%

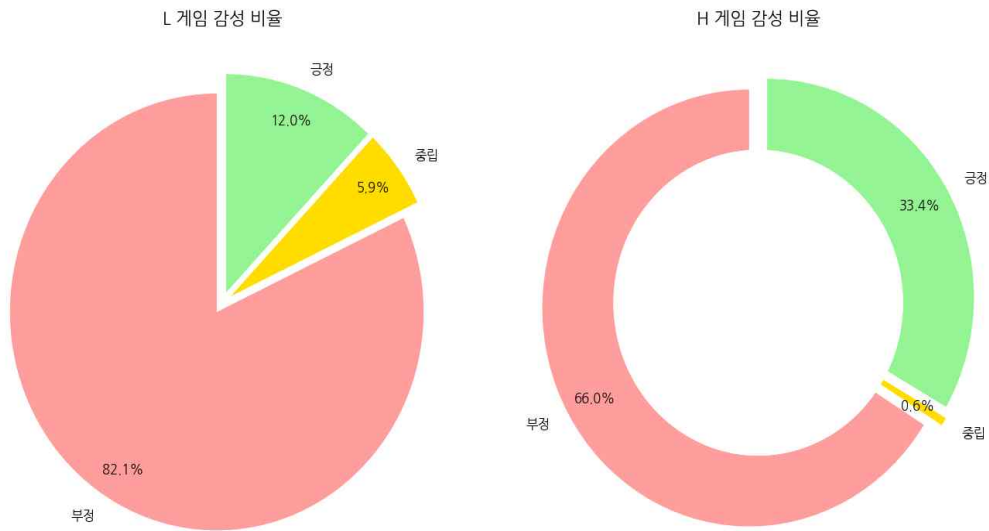


그림 3. L 게임과 H 게임의 공용 단어 사전 감성 비율에 관한 그래프

Fig. 3. Graph on the sentiment ratios of the Common Word Dictionary for L game and H game

L 게임의 경우 추출된 키워드 대부분이 부정 비율이 약 80% 이상을 기록하였고 특히 ‘접속’에 관한 부정 비율이 90%에 근접한 것을 보여주었다. 이는 실시간으로 진행되는 L 게임의 특성상 정체나 튕기는 현상이 빈번히 발생한다는 것을 알 수 있었다. 또한 H 게임과 비교하면 중립 리뷰가 상대적으로 많은 분포를 띄고 있다는 것을 확인할 수 있었다.

#### 2.4 공용 단어 사전을 활용하여 L 게임과 H 게임 비교 분석

키워드 추출을 통해 전반적인 게임에 대한 분석을 마치고 사전에 정립한 공용 단어 사전을 활용해 감성 비율을 확인하여 이를 통해 항목별로 이용자들이 게임에 대해 어떤 감성을 가지고 게임을 이용했는지 한 눈에 이해할 수 있다.

표 4. 공용 단어 사전 기반 분석- L 게임  
Table 4. Common Word Dictionary-Based Analysis - L Game

분류	부정	중립	긍정
BM	81.19%	4.38%	14.42%
게임 운영	87.78%	6.43%	5.79%
콘텐츠	67.38%	5.67%	26.95%
사용자 경험	82.26%	8.06%	9.68%
기술적 성능	84.09%	4.45%	11.36%

표 5. 공용 단어 사전 기반 분석- H 게임  
Table 5. Common Word Dictionary-Based Analysis - H Game

분류	부정	중립	긍정
BM	63.13%	-	36.87%
게임 운영	72.73%	-	27.27%
콘텐츠	65.59%	0.70%	33.70%
사용자 경험	66.40%	-	33.60%
기술적 성능	67.74%	-	32.26%

추가로 두 게임 간의 항목별 부정 비율의 차이를 확인하여 어느 항목이 가장 많은 차이를 보이는지 확인하였다.

표 6. L 게임과 H 게임의 부정 비율 차이  
Table 6. Difference in Negative Ratio Between L Game and H Game

분류	부정 비율 차이
BM	18.06%
게임 운영	15.05%
콘텐츠	1.79%
사용자 경험	15.86%
기술적 성능	16.35%

공용 단어 사전을 통해 분석한 결과를 토대로 두 게임의 부정 비율의 차이를 확인해 보면 가장 큰 비율 차이를 보이는 항목은 ‘BM’으로 18.06%의 차이를 보인다. 그다음으로는 ‘기술적 성능’, ‘사용자 경험’, ‘게임 운영’, ‘콘텐츠 순’으로 차이가 나타났다.

### 2.5 감성분석에 따른 개발 전략 제안

본 연구에서 정립한 공용 단어 사전의 항목별 분석 결과 L 게임은 다섯 개 항목 중 한 개를 제외하고 80% 이상의 부정 감성을 기록하였다. 이 중 가장 높은 수치를 보인 항목은 ‘게임 운영’이다. L 게임은 한국 게임 역사에서 오랜 기간 사랑받아 온 인기 게임 IP(Intellectual Property)였지만 시간이 지남에 따라 기존 BM(수익 모델)을 통한 이윤 추구 주를 이롭과 동시에, 게임의 재미를 증대시키는 업데이트의 부재로 결국 이용자들의 이탈까지 이어지고 있다. 게임 회사가 이윤을 추구하는 것은 자연스러운 현상이지만 이 과정에서 사용자 경험과 충돌해서는 안 된다. 이는 장기적으로 회사의 이미지와 게임에 대한 이용자의 평가에 부정적인 영향을 미칠 수 있기 때문이다. H 게임은 게임 내 비즈니스 모델로 한정 캐릭터 뽑기를 추가하지만, 이 한정 캐릭터를 뽑지 않는다고 해서 H 게임 이용에 제한이 생기지는 않는다. H 게임은 이런 방식으로 이용자들에게 과금하지 않아도 충분히 할 수 있는 게임으로 이용자들 사이에서 평가받고 있다. 반면, L 게임은 지속적인 강화를 통해 몬스터를 사냥하고 이용자들 간의 경쟁을 통해 재미를 보는 것이 메인 콘텐츠이지만 캐릭터가 강화하는 이 과정에서 발생하는 상당한 재화는 대부분이 과금해야만 얻을 수 있

는 재료로 구성되어 있다. 결국 과금 강제성에 의한 악순환이 반복되는 것이다. 이에 대해 L 게임은 기존의 방식에서 벗어나 이용자들이 과금하지 않고도 실력이나 노력으로 충분히 즐길 수 있는 콘텐츠를 추가해야 한다. 또한 지나친 과금 유도에 지친 이용자들의 의견을 반영해 과금 요소를 축소하는 방향으로 개발을 진행하고 부가적으로 게임 내 튕김 현상이나 버그 등 최적화 문제를 해결해 사용자 경험을 높인다면 더 긍정적인 효과를 기대할 수 있다. 이러한 문제들을 개선하고 이용자와 소통한다면 풍부한 사용자 경험을 통해 기존 이용자들에게는 문제점이 완화되어 좋고 신규 이용자 유치에도 유의미한 홍보 효과가 될 수 있을 것으로 생각한다.

### III. 결 론

종합적으로 L 게임은 공용 단어 사전에서 나타난 감성 비율들을 확인한 결과 게임 내외부적으로 발생하고 있는 문제들에 대해 파악하고 BM과 사용자 경험을 개선하는 데에 우선순위를 두고 더 많은 관심을 가져야 한다. 단순히 게임을 개발하는 데에서 그치지 않고 이용자와 소통하여 문제점들을 수용해야 한다. 그리고 이 문제점들을 고치고 발전 시킨다면 이용자와 회사가 상호호혜적 관계로 거듭나 보다 지속 가능한 게임을, 그리고 더 나은 게임산업의 미래를 그려나갈 수 있을 것으로 기대한다.

### 참 고 문 헌

- [1] 한국 콘텐츠 진흥원. “2023 게임 이용자 실태조사”, 13~29p
- [2]CLOVA-SENTIMENT. doi: <https://www.ncloud.com/>
- [3] Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I., “Attention is all you need”, *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998-6008, 2017..

# GPT Reasearcher 기반 대학 BI 보고서 생성

정우제\*, 정유철\*\*

## Generate university BI report based on GPT Researcher

Wooje Jung\*, and Yuchul Jung\*\*

### 요 약

본 연구는 GPT Researcher를 활용하여 프롬프트 조정 기법을 적용한 AI 기반의 BI 보고서 생성 방안을 탐구한다. 기존의 BI 보고서 생성은 데이터 통찰력 정제를 위한 상당한 수작업이 요구되지만, 최근 AI와 프롬프트 엔지니어링의 발전으로 보고서 품질과 관련성을 자동으로 향상시키는 접근이 가능해졌다. 본 연구는 프롬프트를 반복적으로 조정하여 최소한의 개입으로 상세하고 맥락에 맞는 대학 BI 보고서를 생성하는 워크플로우를 제시한다. 연구결과, 체계적인 프롬프트 조정을 통해 높은 수준의 자동화된 BI 보고서를 생성할 수 있음을 확인하였다.

### Abstract

The previous methods for shape from focus (SFF) do not generate accurate depth estimates because they acquire the focus measure over the pixels in only single frame. In the further research, the SFF using artificial neural network method is proposed, and it gives better performance than the traditional methods. For better performance, an artificial neural network for shape from focus is proposed in this paper. The proposed method increase the performance with low local minimization.

### Key words

Prompt Adjustment, Business Intelligence, GPT Researcher, LLM, Data crawling, Report Quality

## 1. 서론

대학의 경쟁력 분석 및 전략 수립에서 비즈니스 인텔리전스(BI) 보고서의 중요성은 갈수록 커지고 있다. 특히, 대학별 언론 동향과 연구 성과를 비교 분석한 보고서는 대학의 중장기 발전 전략을 수립하는 데 중요한 역할을 한다. 그러나 기존의 BI 보

고서 작성 방식은 데이터 수집 및 정제, 분석, 그리고 보고서 생성에 상당한 수작업이 요구되며, 이로 인해 시간과 자원이 많이 소요된다.

최근 GPT와 같은 대규모 언어 모델(LLM)의 발전은 보고서 생성에서 자동화 수준을 높이는 데 기여하고 있으며, 이를 활용한 자동 보고서 생성 방법론이 주목받고 있다. GPT의 답변 품질은 입력된 프

\*금오공과대학교 인공지능공학과. [wjddnwp29@kumoh.ac.kr](mailto:wjddnwp29@kumoh.ac.kr),

\*\*금오공과대학교 인공지능공학과(교신저자), [jyc@kumoh.ac.kr](mailto:jyc@kumoh.ac.kr)

프롬프트에 크게 좌우되므로, 적절한 프롬프트 조정은 보고서의 품질을 높이고, 대학별 특성과 맥락에 맞는 BI 보고서를 자동으로 생성하는 데 필수적이다. 본 연구에서는 대학별 BI 보고서를 자동으로 생성하기 위한 프롬프트 조정 기법을 제안한다. 특히, 대학별 뉴스 헤드라인 데이터를 수집하고 이를 바탕으로 GPT 모델의 프롬프트를 최적화하여, 각 대학의 특성에 맞는 비교 분석 보고서를 생성하는 방법론을 설계하였다.

## II. 방법 설계

본 연구에서는 GPT를 활용한 BI 보고서 자동 생성의 프로세스를 설계하였습니다. 데이터 수집부터 보고서 작성까지의 전체 프로세스는 그림 1에 나타난 바와 같이 크롤링 모듈에서 시작하여 보고서 작성 모듈까지 이어집니다.

### 1. 데이터 수집

연구에 필요한 데이터는 Selenium과 BeautifulSoup을 이용하여 뉴스 헤드라인을 크롤링하는 방식으로 수집하였습니다. 특히, A대학, B대학, C대학을 대상으로 <수집 시기> 동안의 뉴스 데이터를 크롤링하여 데이터베이스를 구축하였으며, 수집된 데이터는 Pandas를 사용하여 Excel 파일로 저장 후 분석에 사용하였습니다.

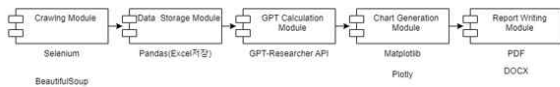


그림 1. 시스템 개요도

### 2. 프롬프트 설계 및 조정

초기 프롬프트는 각 대학의 언론 노출 동향을 간단하게 요약하는 데 초점을 맞춰 설계되었습니다. 이 프롬프트는 특정 카테고리에서 각 대학이 얼마나 자주 언급되었는지에 대한 기본 정보를 제공했으나, 연구 결과의 깊이와 각 대학의 특성을 충분히 반영하지 못한다는 한계가 있었습니다. 이를 보완하기 위해 프롬프트 조정 기법을 적용하여 각 대학의 주요 연구 분야와 사회적 이슈에 대한 기여도, 활동 성과 등을 강조하는 방식으로 프롬프트를 정교하게

다듬었습니다. 이 과정은 단순한 데이터 나열에서 벗어나, 각 대학이 미디어와 사회에 미친 영향력을 더 입체적으로 드러낼 수 있도록 하는 데 중점을 두었습니다.

““

{category} 카테고리의 뉴스 동향 분석:

A대학: {A대학 뉴스 수}건

B대학: {B대학 뉴스 수}건

C대학: {C대학 뉴스 수}건

각 대학의 {category} 관련 활동을 요약하고, 이들이 대학의 발전 전략과 어떻게 연계되는지 논의하십시오.

””

이 프롬프트 설계 및 조정 과정은 단순한 정보 전달을 넘어서, 보다 심층적인 데이터 해석과 각 대학의 고유한 특성을 강조할 수 있는 방향으로 설계되었습니다. 또한, 데이터 분석에 대한 통찰력을 높이고, 각 대학의 전략적 목표와의 연관성을 효과적으로 드러내는 방향으로 최적화되었습니다.

### 3. 데이터 분류 및 시각화

수집된 데이터를 카테고리별로 분류하여 분석한 결과를 바탕으로 각 대학의 언론 노출 빈도와 연구 성과를 시각화하였습니다. 그림 2와 그림 3은 A대학과 B대학의 카테고리별 뉴스 수 비교 및 총 뉴스 수 비교를 보여줍니다.

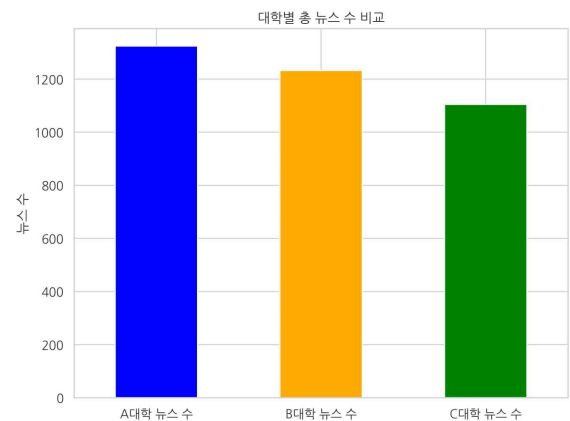


그림 2. 대학별 총 뉴스 비교

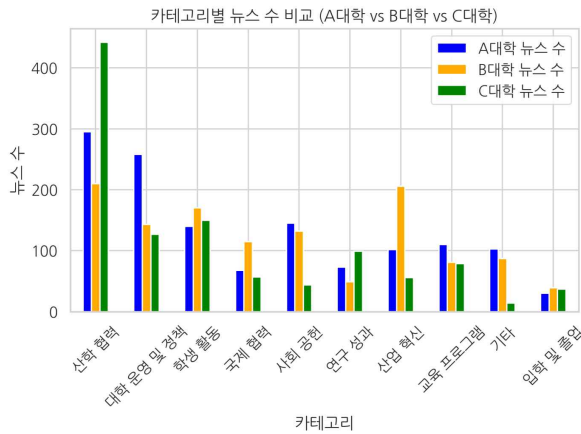


그림 3. 카테고리별 뉴스 수 비교

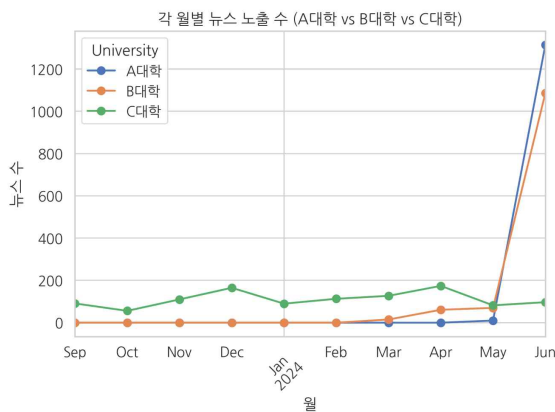


그림 4. 각 월별 뉴스 노출 수

### III. 연구 방법

#### 1. 프롬프트 조정 전후의 보고서 품질 비교

초기 프롬프트로 생성된 보고서와 프롬프트 조정 후 보고서를 비교하여 품질을 분석하였습니다. 초기 보고서는 각 대학의 특성을 충분히 반영하지 못하고 일반적인 내용을 제공하였으나, 프롬프트 조정 후 보고서는 대학별 연구 주제와 사회적 맥락을 구체적으로 반영하여 상세하고 일관된 내용을 제공하였습니다.

#### 2. 비교 보고서 예시 및 결과 분석

그림 2와 그림 3을 통해, A대학은 산학 협력과 연구 성과분야에서 높은 언론 노출 수를 기록하였으며, B대학은 대학 운영 및 정책카테고리에서 다양한 활동을 발표한 것으로 나타났습니다. C대학은 교육 프로그램과 사회 공헌분야에 상대적으로 더 많은 언론 노출 빈도를 보였습니다. 이를 통해 각

대학이 강점을 보이는 분야와 발전 가능성을 확인할 수 있었습니다.

#### 3. 시계열 분석을 통한 동향 파악

그림 4를 통해 A대학, B대학, C대학의 월별 뉴스 노출 수를 시계열로 시각화하여 각 대학의 연구 활동과 협력 주제의 변화 패턴을 분석하였습니다. 이를 통해 각 대학의 전략적 발전 방향을 파악하고, 프롬프트 조정이 이러한 데이터의 맥락을 더 잘 반영하게 함을 확인하였습니다.

### IV. 결론

본 연구에서는 GPT Researcher와 프롬프트 조정 기법을 통해 대학별 BI 보고서를 자동 생성하는 방법론을 제안하고, 이를 통해 프롬프트 조정 기법이 보고서 품질에 미치는 영향을 분석하였습니다. 연구 결과, 체계적인 프롬프트 조정 기법은 대학별 특성과 맥락을 반영한 고품질 BI 보고서 자동 생성을 가능하게 하였으며, 특히 비교 분석 보고서 생성에 있어 매우 효과적임을 확인하였습니다.

본 연구에서 제안한 방법론은 향후 대학뿐 아니라 다양한 산업 분야에서 BI 보고서 생성에 적용될 수 있는 가능성을 가지고 있습니다. AI와 프롬프트 엔지니어링을 활용한 BI 보고서 자동화는 데이터 분석과 인사이트 도출의 효율성을 높이며, BI 보고서 작성에 혁신적인 변화를 가져올 수 있을 것으로 기대됩니다.

### 참 고 문 헌

[1] OpenAIGPT-4 (2023), <https://openai.com/research/gpt-4>

[2] Gpt - researcher, <https://github.com/assafelovic/gpt-researcher>

[3] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A., "Language Models are Few-Shot Learners," Advances in Neural

Information Processing Systems (NeurIPS),  
December, 2020.

- [4] Chen, J., & Zhang, Y., "Automating Business Intelligence Reports: A Case for NLP and AI in Data Analysis," Journal of Business Analytics and AI Applications, April, 2021.



# 다중문서 요약을 위한 Multi - LLM Prompting

김희겸\*, 정유철\*\*

## Multi - LLM Prompting for Multiple Document Summarization

Kim Heekyum\*, and Jung YuCheol\*\*

### 요약

본 연구는 다중 문서 요약(Multi-Document Summarization) 문제에 대하여 복수의 LLM(Large Language Model) 프롬프팅을 활용한 새로운 접근 방식을 제안한다. 대규모 언어 모델은 개별 문서의 요약에 뛰어난 성능을 보였지만, 다양한 출처와 주제를 포함한 다수의 문서를 통합하여 요약하는 데 있어서는 제한적인 성능을 보여왔다. 본 논문에서는 OpenAI의 GPT-4o 모델을 기반으로 하여, 각 문서의 주요 정보를 효과적으로 통합하고 핵심 내용을 추출하는 multi-LLM prompting 기법을 제안하고 이를 평가한다. 연구에서는 LLM을 통해 각 문서의 특성을 반영하는 맞춤형 프롬프트를 생성하고, 이를 종합하여 일관성 있고 유용한 요약을 생성하는 방식으로 다중 문서 요약의 품질을 향상시키고자 했다. 본 접근법의 실험 결과는 기존 단일 모델 기반 요약 기법 대비 더욱 높은 정보 정확성과 가독성을 보이며, 다양한 문서 유형에 대해 강건한 성능을 입증하였다.

### Abstract

This study presents a novel approach to the problem of Multi-Document Summarization by leveraging multiple LLM (Large Language Model) prompting. While large-scale language models have demonstrated impressive capabilities in summarizing individual documents, their performance remains limited when tasked with synthesizing and summarizing multiple documents from diverse sources and topics. This paper proposes and evaluates a multi-LLM prompting technique based on OpenAI's GPT-4o model to effectively integrate key information from each document and extract essential content. In this approach, customized prompts are generated through LLMs, tailored to reflect the characteristics of each document, and then synthesized to produce consistent and meaningful summaries. Experimental results indicate that this approach surpasses traditional single-model summarization techniques in both informational accuracy and readability, demonstrating robust performance across various document types.

### Key words

multiple document, multi agent, LLM

---

\*금오공과대학교 컴퓨터공학과 학부생, hee99521@kumoh.ac.kr,

\*\*금오공과대학교 인공지능공학과 조교수 (교신저자), enthusia77@kumoh.ac.kr

※ "The work is supported by the Korea Basic Science Institute (Project No.A438200)"

## I. Introduction

디지털 정보의 폭발적인 증가로 인해 다수의 문서를 효율적으로 요약하는 기술에 대한 수요가 크게 증가하고 있다. 특히 여러 문서의 핵심 내용을 하나의 요약문으로 통합해야 하는 다중문서 요약(Multi-Document Summarization, MDS)은, 다양한 출처에서 제공되는 방대한 정보를 사용자에게 간결하고 일관되게 전달하는 데 중요한 역할을 한다. 기존의 MDS 연구는 주로 규칙 기반 또는 기계 학습[1] 모델에 의존해왔으나, 대규모 언어 모델(Large Language Model, LLM)의 발전은 이 분야에 혁신적인 변화를 가져왔다.

본 연구에서는 OpenAI의 GPT-4o 모델을 활용하여 다중문서 요약의 성능을 향상시키기 위한 새로운 multi-LLM prompting 기법을 제안한다. 단일 문서 요약에서 LLM은 이미 탁월한 성능을 보였으나, [2]여러 문서에 걸쳐 정보를 통합하고 일관성 있는 요약을 생성하는 데에는 한계가 있었다. 이를 극복하기 위해, 본 논문에서는 각 문서의 특성에 맞춘 프롬프트를 생성하고, 이를 종합하여 다양한 문서로부터 효율적이고 정확한 요약을 생성하는 방법[3], [4]을 탐구한다. 이 방법은 정보 정확성과 가독성을 개선하며, 다중 출처의 문서를 대상으로 강건한 성능을 보이는 것을 목표로 한다.

## II. 기술 구현 방식

이본 연구에서는 다중문서 요약을 위해 여러 개의 LLM 프롬프트를 설계하여 사용하는 multi-LLM prompting 방식을 제안한다. 이 접근 방식은 크게 네 단계로 이루어져 있으며, 각 단계는 OpenAI의 GPT-4o와 GPT-4o-mini 모델을 활용하여 정확하고 일관성 있는 요약을 생성하는 것을 목표로 한다. 각 단계의 세부 사항은 다음과 같다.

### 1. 연구 데이터 추출 (researcher)

researcher 함수는 문서 내의 중요한 정보를 누락 없이 추출하기 위해 PDF를 여러 부분으로 나누어 분석하는 방식을 사용한다. 이 프롬프트는 각 부분

에서 중요한 내용을 한 줄씩 정리하여, 핵심 정보를 철저히 포착하도록 설계되어 있다. 이렇게 나눈 데이터를 통해 각 부분의 핵심 내용이 빠짐없이 수집되며, 이후 통합 요약의 정확성을 높이는 중요한 기초 자료로 활용된다.

### 2. 요약 양식 분석 (mds\_form\_maker)

본 연구에서 사용되는 mds\_form\_maker 함수는 사람이 미리 작성한 동향 보고서 양식을 분석하여 GPT 모델이 해당 양식의 구조를 이해할 수 있도록 한다. 입력된 original 양식을 바탕으로 보고서의 구성 요소를 세밀하게 파악하며, 이를 통해 작성된 보고서가 양식을 최대한 유사하게 따르도록 유도한다. 이러한 양식 분석 과정은 사용자 정의 형식에 맞춘 일관성 있는 요약문을 생성하기 위해 필수적인 단계이다.

### 3. 동향 보고서 생성 (mds\_maker)

mds\_maker 함수는 양식 분석 결과와 추출된 핵심 정보를 바탕으로 한국어 동향 보고서를 작성한다. 프롬프트는 지정된 양식을 철저히 준수하여, 누락되는 정보 없이 모든 주요 내용을 포함하도록 지시한다. 이를 통해 다수의 문서에서 통합된 정보를 기반으로 일관성 있는 요약문을 생성하며, 최종 보고서가 기존 양식에 최대한 유사한 형태로 작성될 수 있도록 한다.

### 4. 결과 평가 (mds\_evaluator)

마지막 단계로, mds\_evaluator 함수는 생성된 요약 보고서가 원본 양식을 얼마나 잘 준수했는지를 평가한다. 이 프롬프트는 보고서와 원본 보고서를 비교하여 일치도를 점수로 산출하며, 보고서의 형식적 일관성과 정보 정확성을 0에서 10점 사이로 평가한다. 평가 결과는 multi-LLM prompting 기법의 성능을 검증하는 중요한 지표로 사용된다.

## III. Conclusion

본 연구에서는 다중문서 요약의 일관성과 정확성을 높이기 위해 multi-LLM prompting 기법을 제안하였으며, 이를 통해 중요한 정보를 누락 없이 통합하여 일관된 형식의 요약문을 생성할 수 있음을 입증

하였다. 현재 사용한 LLM은 오픈 ai사의 GPT 모델로, 외부로의 정보 유출 가능성이 존재한다. 향후 연구에서는 더 가벼운 파라미터를 가진 로컬 환경에서 구동 가능한 LLM 여러 개를 비동기 방식으로 결합하여 구동함으로써 보안성을 강화하고, 효율적인 요약 성능을 실현할 가능성을 탐구하고자 한다.

## Reference

- [1] Congbo Ma, Wei Emma Zhang "Multi-document Summarization via Deep Learning Techniques: A Survey", December 2022.
- [2] Wafaa s. El-Kassas, Cherif R. Salama "Automatic text summarization: A comprehensive survey", March 2021
- [3] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692.
- [4] Gao, L., Biderman, S., Black, S., Golding, L., Hoppe, T., Foster, C., ... & Leahy, C. (2020). The Pile: An 800GB Dataset of Diverse Text for Language Modeling. arXiv preprint arXiv:2101.00027.

# 자동 지식 그래프 구축을 통한 이기종 문서 요약 기법

최승민\*, 정지유\*, 정유철\*\*

## A Method for Summarizing Heterogeneous Documents Through Automated Knowledge Graph Construction

Seung-min Choi\*, Ji-yu Jung\*, and Yu-chul Jung\*\*

### 요약

이기종 문서는 정책/행정문서, 논문, 특허 등의 여러 메타정보로 정의될 수 있는 문서들을 뜻한다. 이러한 이기종 문서들은 수동으로 요약하여 하나의 핵심 정리 문서로 만들어내는 동향 보고서의 형태가 존재한다. 본 논문에서는 최근 많이 연구되는 KG&LLM(Knowledge Graph&&Large Language Model)을 활용한 이기종 문서 MDS(Multi-Document Summarization)를 다룬다. 본 논문은 이기종 문서의 복합지식을 지식 그래프화 하기 위하여, 이기종 문서의 수집 및 메타정보의 관계(Relation) 구조를 정의하고, 정의된 관계 구조를 활용한 자동 복합 지식 그래프 구축 및 이기종 문서의 최종 요약하는 시스템을 제안한다.

### Abstract

Heterogeneous documents refer to various types of documents that can be defined by different metadata, such as policy/administrative documents, academic papers, and patents. These heterogeneous documents are often manually summarized into a single comprehensive summary report. This paper addresses the use of recently studied KG&LLM(Knowledge Graph & Large Language Model) for MDS(multi-document summarization) of heterogeneous documents. To graphize the complex knowledge of these heterogeneous documents, this paper defines the relational structure of the metadata in the documents and proposes a system for automatically building a composite knowledge graph based on the defined relationships, culminating in the final summarization of heterogeneous documents.

### Key words

heterogeneous documents, KG, LLM, MDS

## 1. 서론

이기종 문서는 정책/행정문서, 논문, 특허 등 일관되지 않은 요소나 구성 등을 가진 문서를 일컫는

문서이다. 이러한 이기종 문서를 활용하여 n년차 보고서를 정리하는 업무가 존재하는데, 보통 수동으로 요약한 핵심 보고서 형태로 만들어진다. 하지만 이기종 문서의 정의인 일관되지 않는 구성 및 요소는

\*금오공과대학교 컴퓨터공학과, csm1018csm@gmail.com,

\*금오공과대학교 컴퓨터공학과, snow\_winter7@naver.com,

\*\*금오공과대학교 인공지능공학과, jyc@kumoh.ac.kr

여러 메타정보로 정의될 수 있는데, 이기종 문서를 활용하여 공통 메타정보를 관계(Relation)를 기반으로 정의하고, 정의된 메타정보를 활용하여 지식 그래프(KG, Knowledge Graph)를 자동 구축한다.

구축된 지식 그래프는 거대 언어 모델(LLM, Large Language Model)과 함께 다중 문서 요약(MDS, Multi-Document Summarization)을 진행[1]하여 정책 동향 보고서를 생성한다.

최종적으로 본 연구에서는 다음과 같은 순서로 최종 동향 보고서를 생성하고 평가한다.

1. 이기종 문서 수집 및 메타정보 관계 정의
2. 정의된 정보 기반 자동 지식 그래프 구축
3. 구축된 지식 그래프 활용한 동향 보고서 생성

## II. 실험

### 2.1 자동 복합지식 그래프 구축

본 연구에서는 해당 연구를 위한 정책/행정문서, 논문, 특허에서 이기종 문서 각 300건을 수집하였다. 수집한 이기종 문서의 각 문서 목적, 주제에 맞는 메타정보를 정의를 추출함과 동시에, 문서에서 다르게 표현되는 관계 정규화(Normalization)를 진행한다.

그와 동시에, 수집된 이기종 문서에서 관계 정규화와 더불어 핵심 내용(기술, 기법 등)을 포함하는 객체(Entity) 및 관계 정의와 추출 또한 진행한다. 다음은 메타정보에 따른 관계를 정리한 표이다.

표 1. 메타정보에 따른 관계(Relation) 정리

Table 1 Relations Organization Based on Metadata

메타정보	관계	설명
연구 목표	define	연구의 중심 목표
연구 요약	summarize	연구 전반적인 요약
키워드	relate	연구 관련 핵심 용어
연구 방법	produce	연구 사용된 방법
...	...	...

위와 같이 정의된 이기종 문서의 메타정보를 관계 정의는 Triple Sentence의 관계로 쓰이고, 문맥에 따른 Entities로 Triple을 만들어 복합지식 그래프를

구축한다. 구축된 복합지식 그래프는 다음과 같다.

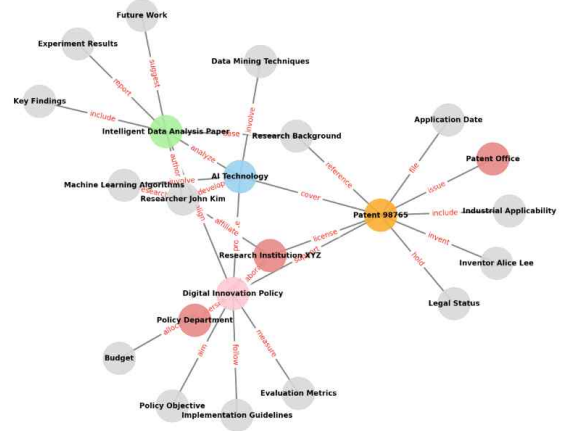


그림 579. 구축된 복합지식 그래프

Fig. 1. Composite Knowledge Graph Constructed

### 2.2 이기종 문서 다중 문서 요약

본 논문에서는 구성된 복합지식 그래프를 토대로 이기종 문서의 다중 문서 요약을 LLM을 사용하여 진행한다[2]. 다중 문서 요약은 해당 업무에 맞는 여러 개의 문서를 입력받고, 입력받은 문서의 키워드나 구성된 복합지식 그래프를 기반으로 진행된다.

진행된 복합지식 그래프 내의 복합지식을 기반으로 문서 내 문맥에서 중요도 높은 정보를 요약한다[3]. GPT4o-mini, Llama-3.2-3b, Llama-3.2-11b 모델을 활용하여 복합지식 그래프가 구성된 모델과 기존의 모델의 성능 비교를 진행한다.

아래의 평가는 이기종 문서를 대상으로 수동으로 요약한 보고서의 기술, 기법 등의 핵심 요소 언급도로 평가한다. 수동 평가의 경우, 총 27개이다.

표 2. 복합지식 그래프 사용/미사용 성능 비교

Table 2. Performance Comparison with and without Composite Knowledge Graph

언급도 평가	KG 사용	KG 미사용
GPT4o-mini	24	18
Llama-3.2-3b	19	15
Llama-3.2-11b	21	17

최종적으로 평가를 진행한 모델은 GPT4o-mini,

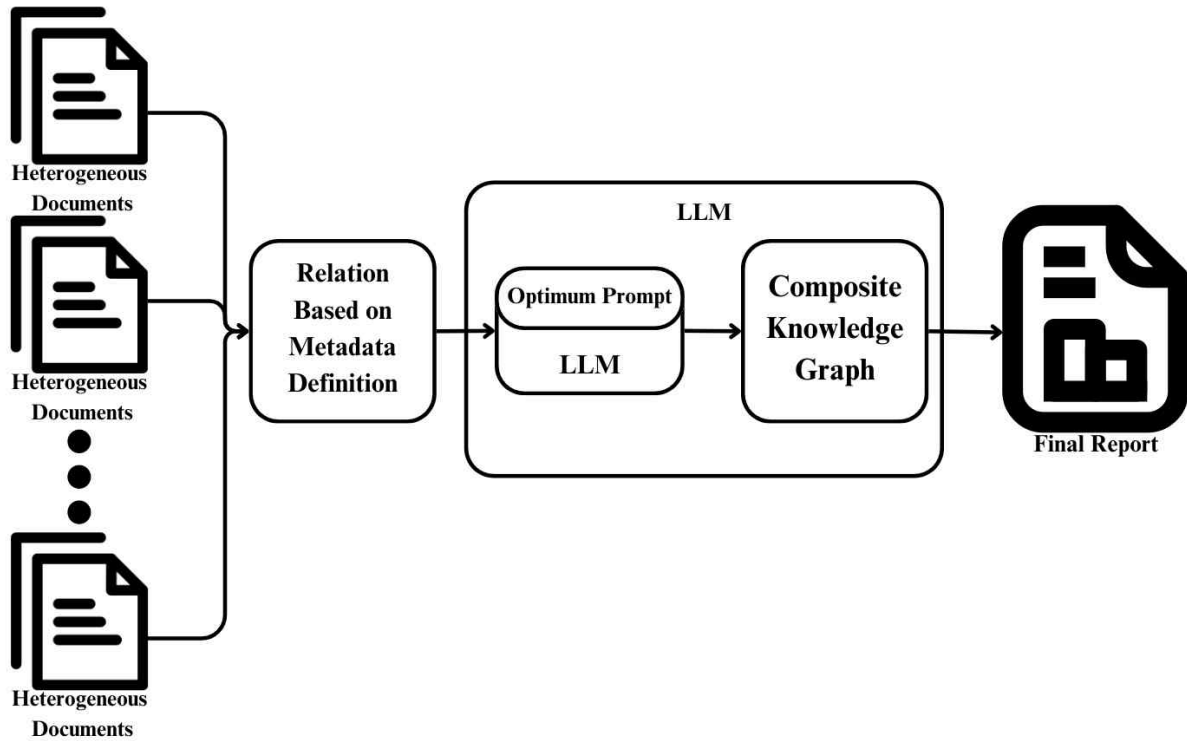


그림 580. 자동 복합지식 그래프 구축 및 다중 이기종 문서 요약 시스템

Fig. 2. Automated Composite Knowledge Graph Construction and Multi-Heterogeneous Document Summarization System

Llama-3.2-3b, Llama-3.2-11b로 복합지식 그래프를 사용한 경우, 전후 18/24, 15/19, 17/21개를 언급도가 높아졌음을 알 수 있다.

### III. 결론

본 연구에서는 이기종 문서의 관계를 기반으로 복합지식 그래프를 구축했고, 이러한 그래프를 활용하여 다중 문서 요약을 진행했다.

결과적으로, 복합지식 그래프를 사용하는 경우, 기존 모델의 다중 문서 요약성능이 개선됐음을 알 수 있었다. 이는 다중 문서 요약에서도 지식 그래프를 활용한 성능 개선을 이룰 수 있음을 알 수 있었다. 이러한 성능 개선은 복합지식 그래프의 형성에서 Triple 정제 등 개선이 이뤄진다면 더 좋은 언급도와 요약성능 향상을 이룰 수 있다.

### 참고 문헌

- [1] Godbole, Aditi, Jabin Geevarghese George, and Smita Shandilya. "Leveraging Long-Context Large Language Models for Multi-Document Understanding and Summarization in Enterprise Applications." arXiv preprint arXiv:2409.18454 (2024).
- [2] Pan, Shirui, et al. "Unifying large language models and knowledge graphs: A roadmap." IEEE Transactions on Knowledge and Data Engineering (2024).
- [3] Wang, Pancheng, et al. "Multi-document scientific summarization from a knowledge graph-centric view." arXiv preprint arXiv:2209.04319 (2022).

# Gmapping 알고리즘 개선을 통한 로봇의 장애물 인식 및 지도 작성 정확도 향상 연구

전종식\*, 강민철\*, 정동원\*\*

## Improving Obstacle Recognition and Mapping Accuracy of Robots through Enhancement of the Gmapping Algorithm

Jongsik Jeon\*, Mincheol Kang\* and Dongwon Jeong\*\*

### 요약

본 연구에서는 기존 Gmapping 알고리즘의 성능을 개선하여 로봇의 장애물 인식 및 지도 작성 정확도를 높이는 새로운 접근 방안을 제안한다. 기존 Gmapping의 문제점인 이동 거리와 회전 각도를 단일 값으로 처리하는 방식을 개선하기 위해 로봇의 이동 거리와 방향 변화를 개별적으로 반영하여 상태 예측의 신뢰도를 향상시켰다. 또한 기존의 맵을 업데이트하는 수식에 가우시안 커널을 적용하여 센서 측정값의 신뢰도를 반영하고 지도 품질을 향상시켰다. 실험 결과, 제안한 방법으로 작성된 맵은 기존 방식보다 장애물을 더 정확하게 인식하여, 지도 작성의 정확도가 기존보다 높은 것을 확인하였다. 이를 통해 SLAM 알고리즘의 실용성과 정확성을 높이는 데에 기여할 수 있음을 보여준다.

### Abstract

This study proposes a new approach to improve the performance of the existing Gmapping algorithm, enhancing obstacle recognition and mapping accuracy for robots. To address the issue in the traditional Gmapping algorithm, which processes travel distance and rotation angle as single values, we individually incorporated the robot's travel distance and direction changes, improving the reliability of state prediction. Additionally, a Gaussian kernel was applied to the map update formula to reflect the reliability of sensor measurements, thereby enhancing map quality. Experimental results show that the maps created using the proposed method recognize obstacles more accurately than previous methods, confirming improved mapping accuracy. This demonstrates that the proposed approach can contribute to the practicality and accuracy of SLAM algorithms.

### Key words

Gmapping Algorithm, Map Creation, SLAM, Robot Navigation, Mapping Accuracy

---

\*국립군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과 {jongsik5077@gmail.com, glsamw960@gmail.com}

\*\*국립군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과 djeong@kunsan.ac.kr(교신저자)

\* 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음

## I. 서론

최근 자율주행 기술의 발달로 로봇이 실내 환경에서도 자율적으로 이동하고 여러 작업을 수행할 수 있는 능력이 중요해지고 있다. 이를 위해서는 동시 위치 추정 및 지도 작성(SLAM; Simultaneous Localization and Mapping) 기술이 필수적이다. SLAM이란 로봇이 자신의 위치를 실시간으로 파악하고 주변 환경의 지도를 작성함으로써 미지의 환경에서도 정확하게 이동할 수 있도록 한다. SLAM은 주로 2D 또는 3D 센서 데이터를 기반으로 이루어지며, 자율주행 자동차, 모바일 로봇, 증강 현실 등 여러 분야에서 필수적인 역할을 수행한다.

SLAM 기술에서 사용되는 여러 알고리즘 중 대표적으로 Gmapping (Grid-based Fast SLAM) 알고리즘이 있다. Gmapping 알고리즘은 2D 환경에서 격자 기반의 지도(Grid map)를 생성하고, 파티클 필터(Particle Filter)를 활용하여 로봇의 위치를 추정한다. 이는 비교적 간단하면서도 효과적인 방식으로 실시간 맵핑을 가능하게 한다.

그러나 기존 Gmapping 알고리즘은 로봇의 이동 거리와 회전 각도를 단일 값으로 처리하는 방식으로 인해 상태 예측의 신뢰도가 낮아지는 문제점이 존재한다. 이는 로봇의 장애물 인식 및 지도 작성 정확도에 부정적인 영향을 미칠 수 있다. 따라서 본 연구에서는 이러한 문제를 해결하고자 기존 Gmapping의 방식에 대한 개선안을 제안한다. 로봇의 이동 거리와 방향 변화를 개별적으로 반영하여 상태 예측의 신뢰도를 높이는 동시에, 맵 업데이트 과정에서 가우시안 커널을 적용하여 센서 측정값의 신뢰도를 반영함으로써 지도 품질을 향상시키고자 하였다.

## II. 관련 연구

### 2.1 Gmapping 알고리즘

Gmapping알고리즘은 SLAM(Simultaneous Localization and Mapping) 기술에서 사용되는 여러 알고리즘 중 대표적인 기법이다. Gmapping은 실내 자율주행 로봇과 같은 다양한 응용 분야에서 널리 사용되고, ROS에서도 기본 SLAM 솔루션으로 제공된다. Gmapping은 2D 환경에서 격자 기반의 지도(Grid map)를 생성한다. 격자 기반 지도란, 환경을 일정한 크기의 셀로 나누어 각 셀에 공간의 정보를 저장하는 방식으로, 로봇이 주변 환경의 구조를 효율적으로 표현하는 것을 의미한다. 또한, Gmapping 알고리즘은 파티클 필터(Particle Filter)를 활용하여 로봇의 위치를 추정한다. 여기서 파티클이란 로봇의 잠재적 위치와 방향을 나타내는 가상의 점들로, 이 점들은 환경에서 로봇이 있을 수 있는 위치를 여러 개의 가능성으로 추정하는 역할을 한다. 파티클 필터는 이러한 다수의 파티클을 통해 센서 데이터를 반영하여 가장 신뢰할 수 있는 위치를 찾고 업데이트함으로써 로봇의 위치를 정밀하게 추정할 수 있다. 이렇게 Gmapping 알고리즘은 격자 기반의 지도를 생성하고, 파티클 필터를 활용하여 로봇의 위치를 추정한다.

이를 위해 다음과 같은 상태 업데이트 수식을 사용한다.

$$x_t^i = f(x_{t-1}^i, u_t) + \epsilon_x \quad (1)$$

$x_t^i$ 는 t시점에서 i번째 파티클의 상태를 의미한다.  $u_t$ 는 로봇에 대한 제어 입력을 나타낸다. 제어 입력에는 로봇이 이동할 방향과 속도가 포함된다. 상태 전이 함수  $f$ 는 이전 상태와 제어 입력을 바탕으로 새로운 상태를 계산하는 함수이다. 여기에 평균이 0인 가우



시안 노이즈  $\epsilon_x$ 를 추가한다.

### III. 제안 방안

본 연구에서는 기존의 Gmapping 알고리즘을 개선하여 로봇의 장애물 인식 및 지도 작성 정확도를 높이는 방법을 제안한다. 기존의 Gmapping은 이동 거리와 회전 각도를 단일 값으로 처리한다. 따라서 로봇의 상태를 업데이트할 때, 두 가지 요소를 구분하지 않기 때문에 위치 추정의 정확성이 떨어진다.

제안 방안에서는 로봇의 이동 거리와 방향 변화를 단일 값이 아닌 두 개의 값으로 처리함으로써 상태 예측의 신뢰도를 높인다. 제안 방안은 파티클 수의 증가, 상태 예측의 신뢰도 향상, 맵 업데이트 수식의 최적화 기법을 포함한다.

#### 3.1 파티클 수 증가 및 상태 예측 보완

Gmapping 알고리즘은 파티클 필터를 기반으로 로봇의 위치를 추정한다. 파티클의 수를 증가시키는 이유는 로봇의 위치 추정 정확도를 높일 수 있기 때문이다. 파티클의 수가 많아질수록 더 많은 위치 추정이 이루어지기 때문에, 로봇이 주어진 센서 데이터를 기반으로 더 정밀한 위치를 찾아낼 수 있다. 또한 파티클의 수를 늘리면 여러 가지 가능한 경로를 동시에 탐색할 수 있으므로, 로봇이 최적의 경로를 찾는 데 유리하다. 그러나 파티클의 수를 증가시킴으로써 계산 비용이 증가하기 때문에, 이를 최적화하기 위해서 이전 상태를 기반으로 현재 상태를 예측하는 상태 예측 수식을 개선한다.

로봇은 특정 시간 간격마다 상태를 업데이트하는데, 이 과정은 다음과 같은 단계를 포함한다. 먼저, 로봇은 이동하는 동안 주변 환경에 대한 센서 데이터를 지속적으로 수집한

다. 이 데이터는 LiDAR 센서, IMU(관성 측정 장치), GPS에서 얻는다. 그 후, 로봇은 이동한 거리와 회전 각도를 계산하여 이전 위치 및 방향에서 새로운 위치와 방향을 예측한다. 이러한 예측은 아래의 상태 업데이트 수식을 통해 이루어진다:

$$x_t = x_{t-1} + \Delta d \cdot \cos(\theta_{t-1}) + \epsilon_x \quad (2)$$

$x_t$ 는 현재 시점  $t$ 에서의  $x$ 좌표를 나타낸다. 이전 시점인  $t-1$ 의  $x$ 좌표에 이동 거리  $\Delta d$ 의 코사인 값을 더한다. 이때 코사인 값은  $x$ 축 방향의 이동량을 나타내며, 여기에 평균이 0인 가우시안 노이즈  $\epsilon_x$ 를 추가한다. 가우시안 노이즈는 센서 오차 및 환경 변화에 따른 예측 오류를 반영하기 위해 포함되며, 이는 로봇의 위치 추정에서 발생할 수 있는 불확실성을 고려하여 보다 현실적인 이동 경로 추정을 가능하게 한다.

$$y_t = y_{t-1} + \Delta d \cdot \sin(\theta_{t-1}) + \epsilon_y \quad (3)$$

$y_t$ 는 현재 시점  $t$ 에서의  $y$ 좌표를 나타낸다. 이전 시점인  $t-1$ 의  $y$ 좌표에 이동 거리  $\Delta d$ 의 사인 값을 더한다. 이때 사인 값은  $y$ 축 방향의 이동량을 나타내며, 여기에 평균이 0인 가우시안 노이즈  $\epsilon_y$ 를 추가한다.

$$\theta_t = \theta_{t-1} + \Delta\theta + \epsilon_\theta \quad (4)$$

$\theta_t$ 는 현재 시점  $t$ 에서의 회전 각도를 나타낸다. 이전 시점인  $t-1$ 의 회전 각도에 이동 중의 회전 각도 변화를 의미하는  $\Delta\theta$ 를 더한다. 그 후, 평균이 0인 가우시안 노이즈  $\epsilon_\theta$ 를 추가한다.

이로써 각  $x, y$  좌표의 이동 거리와 로봇의 회전 각도를 단일 값이 아닌 두 개의 값으로 처리를 하게 된다. 이러한 예측 결과를 기반

으로, 로봇은 각 파티클을 업데이트하고, 각 파티클의 확률을 계산하여 위치 추정을 더욱 정확하게 한다. 이는 로봇이 최적의 경로를 탐색하고 장애물 인식 및 지도 작성의 정확도를 높이는 데 기여한다.

### 3.2 맵 업데이트 수식 최적화

본 연구에서는 Gmapping 알고리즘의 맵 업데이트 과정에서 발생하는 한계를 극복하고자, 기존의 점유 확률 계산 방식을 개선하는 방안을 제안한다.

점유 확률(Occupancy Probability)은 특정 지역이 장애물로 점유되어 있을 확률을 나타내는 지표이다. 이는 로봇이 수집한 센서 데이터를 기반으로 특정 지점이 장애물인지, 비어 있는 공간인지를 판단하는 데 사용된다. 일반적으로  $[0, 1]$  범위의 값을 가지며, 0은 해당 셀이 비어 있음을, 1은 장애물이 있음을 나타낸다. 기존 Gmapping에서는 단순한 조건부 확률을 곱하여 점유 확률을 계산하였으나, 이 방식은 센서 측정값의 불확실성을 충분히 반영하지 못하는 문제점이 있다. 따라서 제안 방안에서는 가우시안 커널을 추가하여 센서 측정값의 불확실성을 반영하여 점유 확률 계산의 정확성을 높였다.

기존 Gmapping 알고리즘에서의 점유 확률 계산은 다음과 같은 수식을 통해 이루어진다:

$$P(m_{i,j}|z_{1:t}, x_{1:t}) = \frac{P(z_t|m_{i,j}) \cdot P(m_{i,j}|z_{1:t-1}, x_{1:t-1})}{P(z_t|z_{1:t-1}, x_{1:t-1})} \quad (5)$$

분모의  $P(z_t|z_{1:t-1}, x_{1:t-1})$ 은 이전 측정값과 현재 위치에서 새로운 측정값이 나올 확률을 의미한다. 이는 새롭게 관측된 데이터에 대한 신뢰도를 고려한 확률이다.  $P(z_t|m_{i,j})$ 은 현재 셀의 점유 여부에 따라 특정 측정값이 나올 확률을 의미한다. 이는 현재 측정값의 척도를 반영한다.  $P(m_{i,j}|z_{1:t-1}, x_{1:t-1})$ 은 이전 측정값과 위치를 바탕으로 셀이 점유될 확률을

의미한다.

이를 바탕으로 제안 방안에서는 가우시안 커널을 적용하여 센서 측정값의 불확실성을 확률적으로 반영하여  $P(z_t|m_{i,j})$ 를 더욱 정밀하게 계산한다. 수식은 아래와 같다.

$$P(z_t|m_{i,j}) = \exp\left(-\frac{(z_t - \hat{z})^2}{2\sigma^2}\right) \quad (5)$$

여기서  $z_t$ 는 현재 측정값,  $\hat{z}$ 는 예측된 거리,  $\sigma$ 는 측정 값의 표준 편차를 나타낸다.

## IV. 실험 및 결과

### 4.1 실험 환경

실험 환경은 실내 환경에서 수행이 되었으며, 실험 장소의 사진을 통해서 환경의 구조와 장애물 분포를 확인할 수 있다. 로봇에는 2d LiDAR, IMU(관성 측정 장치), GPS가 장착되어 있으며, 이를 통해 로봇의 위치 추정 및 지도 작성 과정에서 수집된 데이터를 활용하였다.

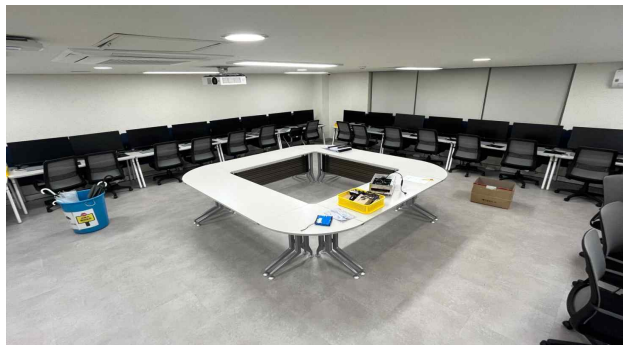


그림 581 실제 로봇 실험 환경

### 4.2 실험 결과

제안한 알고리즘을 사용하여 생성된 지도와 기존 Gmapping 알고리즘으로 생성된 지도를 정성적으로 비교하였다. 그림 2은 기존의 Gmapping으로 작성된 맵이고, 그림 3는 제안한 방식으로 작성된 맵이다.

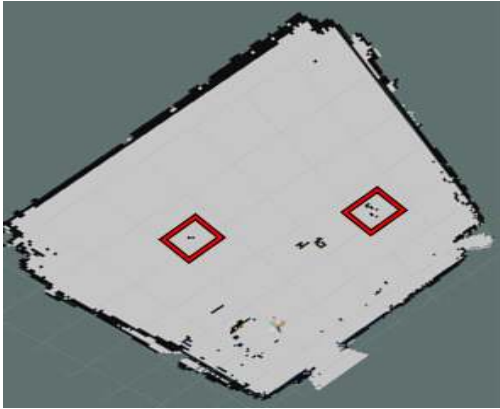


그림 582 Gmapping으로 작성된 맵

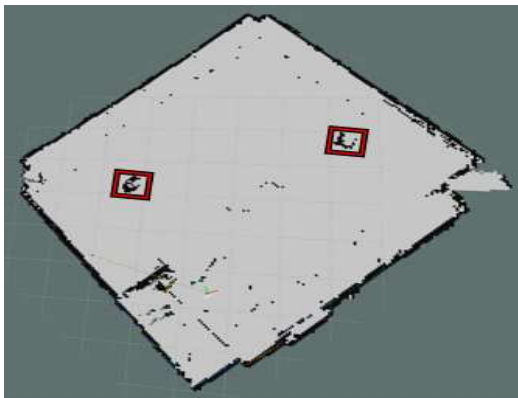


그림 583 제안한 방식으로 작성된 맵

결과 맵을 비교하였을 때, 제안한 방식으로 작성된 맵이 기존의 방식으로 작성된 맵에 비해서 두 가지 장애물과 벽에 대해서 더 정확하게 인식한 것을 확인할 수 있다.

## V. 결론

본 연구에서는 기존 Gmapping 알고리즘의 한계를 보완하여 로봇의 장애물 인식과 맵 작성 정확도를 향상시키기 위한 새로운 접근 방안을 제시하였다. 로봇의 이동 거리와 방향 변화를 개별적으로 반영함으로써 상태 예측의 신뢰도를 높였다. 또한 파티클 수 증가시키고, 가우시안 커널을 적용해 맵 업데이트 수식을 개선하여 맵의 품질을 개선하였다. 실험 결과, 제안 방안이 기존보다 장애물

을 정확하게 인식하여 맵을 작성한 것을 확인하였다. 본 연구에서 제안된 기법은 향후 다양한 SLAM 응용 분야에서 정밀한 지도 작성과 위치 추정이 요구되는 환경에 적용 가능할 것으로 기대된다.

## 참 고 문 헌

- [1] 송지민, 이상준. 실내의 주행환경 및 LiDAR의 화각에 따른 SLAM 성능 분석 연구. 대한전기학회. pp. 2,171-2,172. 2024.
- [2] 이기현, 조동진, 강길환, 신영식, 조영근. 실내 환경에서의 평면 팩터 기반 그래프 SLAM 및 의미론적 물체 표현 방법. 제어로봇시스템학회. pp. 412-413. 2024.
- [3] 이명희, 정영재, 장건, 서희, 이성진. Loam SLAM을 활용한 교내 Mapping. 한국통신학회. pp. 963-964. 2023.
- [4] 조재형, 안우주, 공우식, 지승한. 센서 퓨전을 통한 범용 환경에서의 SLAM 성능 향상. 한국정보기술학회. pp.889-894. 2022.
- [5] 유찬희, 이종원, 이현우, 김태혁, 임현민, 이충호. ROS기반 SLAM을 활용한 장애물 인식 및 모빌리티 자율주행. 한국기계기술학회. pp.451. 2022

# 로봇을 활용한 어르신 대상 키오스크 사용자 교육의 자유도 향상을 위한 구현 방법론 및 실험 설계

김하람\*, 신예준\*, 이의진\*, 이원형\*\*

## Implementation Methodology and Experimental Design to Enhance Flexibility in Robot-Based Kiosk Training for Older Adults

Haram Kim\*, Yejun Sin\*, Uijin Lee\*, and Wonhyung Lee\*\*

### 요약

COVID-19 이후 비대면 서비스의 확산으로 키오스크 사용이 증가하면서, 디지털 소외계층인 고령층이 키오스크 사용에 어려움을 겪고 있다. 이에 따라 고령층을 위한 다양한 교육 프로그램이 제공되고 있으나, 접근성 및 자율성이 부족하다는 한계가 있다. 본 연구는 이러한 기존 교육 방식의 한계를 보완하기 위해 자율 학습 기반의 교육 앱과 AI 로봇 연동 방안을 제안하였다. 로봇이 학습자에게 칭찬과 격려, 상황별 가이드를 제공함으로써 학습 효과를 높이며, 고령층이 키오스크 사용 과정을 더욱 유연하고 실제에 가깝게 경험할 수 있도록 돕는다. 이를 통해 학습자의 자율성과 문제 해결 능력을 향상시키고, 고령층의 디지털 접근성을 개선할 수 있을 것으로 기대한다.

### Abstract

With the expansion of contactless services post-COVID-19, kiosk usage has increased, leading elderly individuals in the digitally marginalized population to face challenges in its use. In response, various educational programs have been provided for older adults, but they still lack accessibility and autonomy. This study proposes a self-directed learning app integrated with AI and robotic assistance to address these limitations in existing educational approaches. By providing encouragement, situation-specific guidance, and positive reinforcement, the robot enhances the learning effectiveness, allowing elderly learners to experience kiosk use in a flexible and realistic manner. This approach is expected to improve learner autonomy, problem-solving skills, and ultimately increase digital accessibility for the elderly.

### Key words

Elderly education, kiosk usage, digital accessibility, autonomous learning, AI-assisted learning, digital divide.

---

\*한동대학교 전산전자공학부 학생, 제1저자

\*\*한동대학교 전산전자공학부 교수, 교신저자, whlee@handong.edu

## I. 서 론

COVID-19 이후 비대면 서비스가 증가하면서 키오스크 사용 비율이 크게 증가하였다. 하지만 키오스크 사용에 익숙하지 않은 디지털 소외계층에 포함되는 고령층 이용자는 키오스크 사용에 많은 어려움을 겪고 있으며 이로 인해 일상의 어려움을 겪고 있다. 이에 따라 어르신들을 위한 키오스크 교육과 관련한 서비스가 증가하고 있으며 그 필요성이 강조되고 있다. 이와 관련하여 정부 교육 프로그램, 교육 서비스, 정책 등이 도입되어 고령층 이용자의 키오스크 사용을 도움으로써 문제를 해결하고자 노력하고 있다. 하지만, 이와 관련한 기존의 교육제도, 서비스들은 접근이 제한적이고, 자율성이 떨어진다는 문제가 있다. 본 연구는 고령층 이용자가 기존의 키오스크 교육이 미치는 교육 효과에 주목하며, 기존의 교육 방식과 한계를 보완하는 방안에 관한 연구를 제시하고자 한다. 또한, 고령층 이용자에게 자율성을 부여하여 교육 효과를 높이고, 로봇과 연동된 교육 앱을 사용함으로써 더 효과적인 학습 환경 및 교육 방식을 제안하고자 한다.

## II. 선행연구 고찰

현재 수행되고 있는 키오스크 교육 사례 연구를 사례에 따라 4가지 유형으로 정리하면, 다음과 같다. 첫째로 프로그램 운영 유형은 이론과 실습을 겸비하여 실제와 유사한 경험을 제공할 수 있지만, 지속적인 교육 진행이 어려우며, 제한된 키오스크 기기 보급으로 다수 인원수용이 불가능하다는 한계가 있다. 둘째로, 온라인 콘텐츠 제공은 장소의 제약이 없으며 다양한 콘텐츠를 제공할 수 있으나, 실제 키오스크 이용과 다르며, 체계가 부족하다는 한계가 있다. 셋째로, 교육용 서비스 개발은 스스로 학습이 가능하다는 장점이 있으나, 범위와 유형을 특정하지 않아 구체적 학습에 한계가 존재한다. 마지막으로, 체험존 설치는 다양한 학습이 가능하다는 장점이 있으나, 설치 공간 및 기기 확보에 한계가 존재한다. [1]

결론적으로, 기존에 수행되고 있는 선행연구는

대부분 포괄적이며 자율성이 부족하다. 또한, 주입식이거나 수동적인 접근으로 인해 그 학습 효과에 한계가 존재하며, 이러한 이유로 실제적인 학습 효과를 일상생활에서 적용하기에는 어려움이 있다.

## III. 연구모형 및 가설

앞서 선행연구 고찰에서 살펴보았듯이 기존에 수행되고 있는 교육 방식에는 접근 제한성이 존재하며 자율성이 떨어진다는 한계가 있음을 보여주고 있다. 본 연구에서는 고령층을 위한 키오스크 교육의 한계를 보완하려는 방안을 제시하고자 한다. 기존의 한계 중 하나인 접근 제한의 어려움을 극복하고 능동적인 접근을 제공함으로써 학습 효과를 높이기 위해 다음과 같은 연구모형을 설계하였다.

기존의 교육용 서비스 개발을 보완함으로써 스스로 학습이 가능하다는 장점을 유지하되, 구체적 학습의 한계를 극복하기 위해 AI와 로봇을 도입함으로써 그 한계를 보완하고자 한다. 대부분의 교육용 서비스는 사용자가 취해야 하는 행동이 고정되어 있어 학습자의 자율성이 떨어진다. 이를 보완하고자 무작위로 학습 목표를 제공함으로써 더욱 다양한 흐름을 경험하고, 능동적인 학습을 돕는다. 또한, 잘못된 버튼을 누르거나, 틀린 답을 선택하더라도 다음 페이지로 넘어가도록 허용함으로써 자율성을 제공하여 실제와 유사한 상황을 경험하도록 하며 답을 유추하지 못하도록 한다. 더불어 사용자에게 친근함을 주는 AI와 연결된 로봇을 서비스와 연동하여 사용자에게 칭찬과 격려, 가이드를 제시함으로써 로젠탈 효과를 기대할 수 있으며[2], 이러한 효과로 더욱 학습 효율을 높일 수 있다. 그뿐만 아니라, 학습자의 학습 과정을 기록하여 분석해 보여주고 이를 학습자가 인지함으로써 더욱 높은 학습 효과를 기대할 수 있다. [3]

## IV. 연구방법

### 1. 연구 대상

- 본 연구의 대상은 키오스크 사용 경험이 적거

나, 디지털 기기에 익숙하지 않은 65세 이상의 고령층으로 설정한다.

- 참여자는 지역사회 복지관 혹은 디지털 교육 센터의 지원을 받아 모집하며, 참여자에게 연구의 목적과 절차를 설명한 후 동의를 받는다.
- 참여자의 교육 효과 평가를 위해 기본적인 디지털 리터러시 수준, 키오스크 사용 경험, 자기 주도 학습에 대한 선호도를 사전에 조사해 연구에 반영한다.

## 2. 연구 도구

### 2.1 교육 앱과 AI 로봇 연동 학습 환경

연구는 고령층의 자율적 학습 효과를 증진하기 위해 교육 앱과 AI 로봇을 연동한 학습 환경을 제공하였다. Flutter로 개발된 앱과 ChatGPT API 기반의 AI 시스템, 그리고 로봇을 결합하여 고령층이 스스로 다양한 시나리오를 학습할 수 있도록 유도하였다. AI 로봇과 사용자는 상호 작용하며 학습 진행 상황을 실시간으로 모니터링하고 피드백을 제공하여 학습 몰입도와 성취감을 높였다.

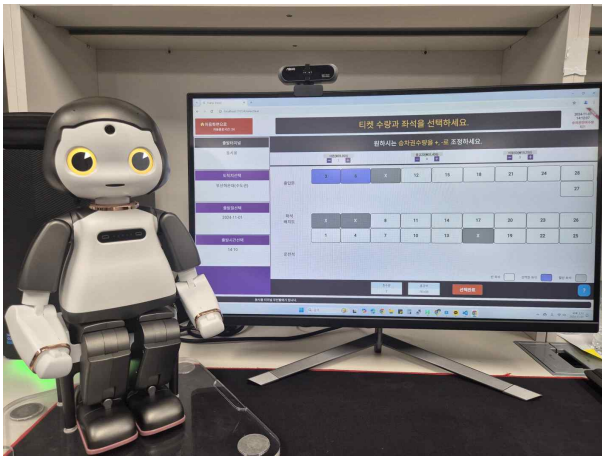


그림 584. AI로봇과 교육 앱 활용 모습  
Fig 1. Utilizing AI Robots and Educational App

#### 2.1.1 키오스크 UI 기반 앱의 목표 설정 및 자율 학습 강화

교육용 앱은 동서울 터미널의 실제 키오스크 UI를 본떠 Flutter로 개발되었으며, 고령층이 실제 키오스크와 유사한 학습 경험을 하도록 설계되었다.

주요 화면은 [처음 화면], [목적지 선택], [버스 시간 선택], [좌석 선택], [결제], [완료]의 6개로 구성되며, 각 화면은 직관적인 흐름을 통해 사용자가 손쉽게 학습할 수 있도록 구성되었다.

처음 화면: 보라색 버튼(당일/예매 발권)을 선택하면 [목적지 선택 화면]으로 이동.

목적지 선택 화면: 사용자는 상단의 지역 선택과 중단 초성 검색 기능을 통해 목적지 목록을 좁혀나가고, 목적지를 선택하면 [버스 시간 선택 화면]으로 이동한다.

버스 시간 선택 화면: 출발 시간에 맞는 버스 리스트가 있고 각 버스의 오른쪽에 있는 선택 버튼을 누르면 [좌석 선택 화면]으로 이동한다.

좌석 선택 화면: 사용자가 승차권 수량과 좌석을 선택한 후 결제 화면으로 이동하며, 실제 결제는 애니메이션으로 대체하여 학습 경험을 제공한다.

결제 화면 및 완료 화면: 출발지, 도착지, 출발 시간, 요금 정보, 버스 시간 등의 정보가 표시되고 (카드결제) 버튼을 누르면 [완료 화면]으로 이동한다.

완료 화면 페이지: 실제 키오스크에서는 카드를 삽입하여 결제가 진행되지만 교육용 앱에서는 결제 애니메이션으로 대체하였다.

무작위 목표 설정 기능: 학습자가 매 세션마다 새로운 목표를 무작위로 설정하여 고정된 학습 흐름에서 벗어나도록 하였다. 잘못된 버튼을 선택해도 학습을 이어갈 수 있는 유연한 흐름을 제공하여 자율성을 강화하였다.

#### 2.1.2 로봇 피드백을 통한 학습 동기 부여

로봇은 학습자가 특정 과제를 완료할 때마다 칭찬과 격려를 제공하고, 필요한 경우 단계별 가이드를 제시하여 로젠탈 효과를 유도하였다. 이를 통해 학습자는 심리적 안정감을 느끼고 학습 지속성을 유지할 수 있도록 지원받는다.

로봇은 학습자가 특정 과제를 완료했을 때 칭찬과 격려를 제공하고, 필요한 경우 단계별 가이드를 제시함으로써 로젠탈 효과를 유도하여 학습자의 동기과 자신감을 높이는 데 이바지할 수 있다.

### 2.2 ChatGPT API 기반 키오스크 가이드 시스템

본 연구에서는 ChatGPT API를 활용하여 학습자

에게 실시간 키오스크 가이드를 제공하였다. 이를 위해 다음과 같은 파일 구성을 통해 학습자가 다양한 시나리오를 자율적으로 학습할 수 있도록 설계하였다.

### 2.2.1 시나리오 파일 구성

시나리오 파일(scenario.txt): 프로그램은 실행 시마다 204개의 목적지, 11종 인원, 8개의 시간 선택을 무작위로 조합하여 약 17,952개의 시나리오를 생성하였다. 학습자는 매 세션마다 다른 상황을 경험할 수 있다.

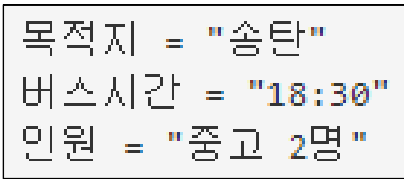


그림 585. 시나리오 파일의 예시  
Fig 2. Example of a Scenario File

### 2.2.2 키오스크 화면 구성 정보

화면 구성 정보(kiosk.json): 각 화면의 구성 요소(버튼 이름, 기능 등)를 JSON 형식으로 정의하여 AI가 각 화면의 구조와 기능을 이해하고, 상황에 맞는 피드백을 제공할 수 있도록 하였다.

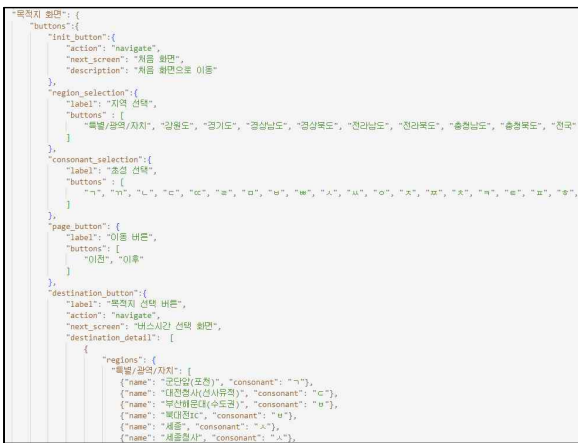


그림 586. 화면 구성 정보 파일의 예시  
Fig 3. Example of a Screen Layout Information File

### 2.2.3 학습 기록 및 사용자 매뉴얼 제공

매뉴얼(manual.txt) 및 학습 기록(history.txt): manual.txt 파일은 학습자가 각 화면의 기능을 쉽게

이해하고 적절하게 사용할 수 있도록 돕는다. history.txt 파일은 학습자가 선택한 버튼과 AI의 응답을 기록하여, AI가 학습자의 선택을 평가하고 적절한 피드백을 제공하는 데 사용된다.



그림 587. 매뉴얼 파일의 예시  
Fig 4. Example of a manual file

## 2.3 실시간 통신 및 AI와 사용자 상호작용 시스템

### 2.3.1 앱과 Azure 서버 간 통신 구조

앱은 학습자의 입력 데이터를 HTTP POST 요청으로 서버의 FastAPI에 전송한다. 이 과정에서 sendMessageToServer 함수가 사용자 입력을 JSON으로 인코딩하여 Azure 서버로 전달하는 역할을 수행한다.

Azure 서버는 FastAPI를 사용하여 앱과 RESTful 방식으로 통신하며, 사용자 입력을 수신한 후 LangChain을 통해 이를 처리하여 AI 응답을 생성하고, 이를 JSON 형식으로 앱에 반환한다. 또한 Azure 서버는 MQTT broker를 실행하여 Liku 서버로 텍스트 데이터를 전달하는 역할도 수행한다.

### 2.3.2 LangChain 기반 맞춤형 AI 응답 생성

FastAPI는 LangChain과 OpenAI API를 활용하여 사용자의 질문에 대해 일관성 있는 AI 응답을 생성한다. 이 과정에서 PromptTemplate을 활용해 시나리오, 매뉴얼, 학습 기록을 바탕으로 맞춤형 응답을 제공한다.

### 2.3.3 Azure - MQTT를 통한 실시간 음성 피드백 제공

생성된 응답은 MQTT 프로토콜을 통해 Liku 서버로 전송되며, 이를 통해 사용자에게 실시간 음성

피드백이 제공된다.

2.3.4 Liku 서버 및 Liku 로봇을 통한 사용자 피드백

Liku 서버는 Mqtt 서버로부터 수신한 메시지를 Liku 로봇에 전달하고, 로봇은 해당 메시지를 바탕으로 사용자에게 음성 피드백을 제공하여 학습자의 몰입과 참여를 높인다.

### 3. 실험 설계

#### 3.1 무작위 목표 설정과 자율 학습

학습자에게 매번 다른 학습 목표를 무작위로 제시하여 다양한 학습 경험을 유도하고, 능동적인 학습 참여를 유도한다. 이로써 기존의 고정된 학습 흐름에서 벗어나 사용자가 상황에 따라 유연하게 학습을 조작할 수 있도록 돕는다.

#### 3.2 AI 연동 로봇의 피드백 제공

학습 중 로봇이 학습자가 버튼을 잘못 누르거나 틀린 답을 선택했을 때도 긍정적 피드백과 간단한 가이드를 제공함으로써 심리적 안정감을 높이고, 학습 도중 발생할 수 있는 스트레스를 완화한다.

#### 3.3 실제 상황과 유사한 시나리오 제공

실제 키오스크 상황과 유사한 시나리오를 포함하여 잘못된 선택을 해도 학습을 이어가도록 설계해 학습자의 자율성과 문제 해결 능력을 증진한다.

### 4. 자료 수집 방법

#### 4.1 설문조사

교육 후 학습자에게 설문조사를 시행하여 자율 학습 방식과 로봇의 피드백에 대한 만족도, 학습 경험의 효과, 학습 후 키오스크에 대한 심리적 불안감의 정도의 차이에 대해 평가한다.

#### 4.2 학습 결과 제공

학습자가 키오스크 사용하는 과정을 관찰하여 학습 중 발생하는 어려움을 기록한다. 이를 통해 학습자가 어떤 부분에서 틀린 답을 선택했고, 어떤 부분에서 옳은 답을 선택했는지를 로봇을 통해, 친근하게 전달함으로써 학습 효과를 높인다.

### 5. 분석 방법

설문조사 결과를 통해 자율 학습 기반 교육 방식

과 로봇 피드백의 효과를 분석한다. 교육 전후의 만족도와 자신감 수준의 변화를 비교하여 학습 효과를 평가한다.

## V. 예상 결과 및 기대 효과

본 논문에서는 실험 설계에 대한 제안을 먼저 하였으며, 이를 기반으로 고령층이 자율적으로 학습할 수 있는 교육 방식을 제안하였다. 로봇을 활용한 자율 학습 기반 교육 방식은 고령층이 자율적으로 학습할 수 있게 하여 실제 상황에서도 적용할 수 있는 교육 효과가 높아질 것으로 기대한다.2) 추후 연구를 통해 사용자 테스트를 확대하여 제안된 교육 방식의 실질적인 효과를 검증할 계획이다. 로봇의 칭찬과 격려는 학습자의 학습 동기를 유발하고, 상황에 따라 제공되는 가이드는 학습자의 학습 효과를 증진하고 지속해서 교육에 참여할 수 있도록 유도할 것으로 예상된다. 기존 교육 방식의 한계를 극복할 뿐 아니라, 능동적인 학습 참여와 로봇을 통한 상호작용은 학습 효과에 큰 긍정적 효과를 가져올 것으로 기대한다.

## VI. 결론

본 연구는 COVID-19 이후 비대면 서비스가 증가함에 따라 키오스크 사용이 일상화되었으나, 디지털 소외계층인 고령층이 이러한 기술에 적응하는 것에 어려움을 겪고 있다는 문제에서 출발하였다. 본 연구의 목적은 고령층 이용자가 기존의 키오스크 교육이 미치는 교육 효과에 주목하며, 기존의 교육 방식과 한계를 보완하는 방안 관련 연구를 제시하고자 하였다. 기존의 주입식 교육 방식은 고령층 이용자에게 충분한 자율성과 유연성을 제공하지 못해 학습 효과가 제한적이라는 한계가 있다. 이에 본 연구에서는 고령층이 자율적으로 학습할 수 있는 교육 방식과 로봇과 연동되어 제공되는 피드백 기능을 결합하여 고령층의 키오스크 교육 학습효과를 향상할 방법을 제시하였다.

2) 시연 영상: <https://youtu.be/iVV-e8ptLw>



본 연구에서 제안하는 자율 학습 기반 교육 앱과 AI 로봇의 연동은 고령층들이 학습 과정에서 스스로 조작하고 실수를 교정하며 더 유연하게 학습할 수 있도록 설계되었다. 로봇의 칭찬과 격려는 학습자의 학습 동기를 유발하고, 상황에 따라 제공되는 가이드는 학습자의 학습 효과를 증진 시키며 지속해서 교육에 참여할 수 있도록 유도한다. 또한, 고령층 이용자가 틀린 답을 선택했을 때에도 다음 행동을 제한하거나 통제하지 않고 학습을 이어가도록 설계해 실제 흐름을 경험할 수 있으며, 학습자의 자율성과 문제 해결 능력을 증진한다. 또한, 학습자가 키오스크 사용하는 과정을 관찰하여 학습 중 발생하는 어려움을 기록해 학습자가 틀린 답을 선택했는지, 옳은 답을 선택했는지를 정리해 보여줌으로써 학습 효과를 높인다.

향후 연구에서는 실제 사용자 테스트(User Testing)를 통해 본 연구에서 제안한 교육 방식의 실질적인 효과를 검증할 필요가 있다. 이를 통해 고령층의 키오스크 접근성을 개선하고, 디지털 소외계층을 위한 맞춤형 교육 프로그램의 발전에 이바지할 수 있을 것이다.

## 참 고 문 헌

- [1] 윤예진 and 하광수, "중 · 고령층의 디지털 역량 향상을 위한 키오스크 교육 사례 연구," 한국콘텐츠학회 종합학술대회 논문집, pp. 99-100, 대전, 2021-08-27.
- [2] 강마리아, "Influence of Rosenthal Effect Using Praise on the Improvement of Concentration of the Elderly in Mobile Serious Game Learning," 2018.
- [3] 김상홍, "AI기반 학습분석 시스템의 교육적 효과성 분석," 한국콘텐츠학회논문지, vol. 24, no. 6, pp. 55-65, 2024. (10.5392/JKCA.2024.24.06.055)

# 미세먼지 보간 모델을 위한 지역적 배치 정규화

윤진용\*, 안석호\*\*, 서영덕\*\*

## Local Batch Normalization for Air Pollution Interpolation Models

JinYong Yun\*, Seokho Ahn\*\*, and Young-Duk Seo\*\*

### 요약

배치 정규화(Batch Normalization, BN) 기법을 도입하여 미세먼지 보간 모델을 개선한다. 기존 보간 연구는 배치 정규화를 활용한 공간적 특징의 안정화를 고려하지 못했다. 본 연구에서는 미세먼지 보간 모델에 글로벌 배치 정규화와 로컬 배치 정규화 계층을 적용하여 이러한 문제를 해결하고자 한다. 실험 결과, 다양한 수준에서 배치 정규화 계층을 적용하면 보간 정확도가 개선됨을 확인했다.

### Abstract

We propose a method to enhance spatial dependencies in air pollution interpolation models by applying Batch Normalization (BN) layers. Global BN and Local BN are applied to all stations and individual stations, respectively, to improve spatial feature stability. Experiments using PM2.5 data from Antwerp demonstrated that models with BN layers achieved higher interpolation accuracy compared to conventional models. This confirms that BN layers can effectively enhance the performance of air pollution interpolation models.

### Key words

Air pollution, Interpolation, ConvLSTM, Batch Normalization, PM2.5

## 1. 서론

도시 내 미세먼지 농도를 정확히 측정하기 위해 곳곳에 미세먼지 측정소가 설치되고 있지만, 모든 지역을 커버할 만큼 충분한 미세먼지 측정소를 설치하는 것은 상당한 비용과 공간적인 제약이 따른다. 따라서 미세먼지 측정소가 없는 지역의 농도 추

정을 위해 다양한 보간 연구가 활발히 진행되고 있다 [1-3]. 이러한 연구에서는 주로 시간과 공간의 상호작용을 고려한 시공간(spatiotemporal) 모델이 사용되며, 특히 시계열적 성능 개선에 초점을 맞춰 보간 성능을 향상시켜왔다.

시계열적 성능 개선 방식으로 배치 정규화를 적용할 경우, 각 시점에서 개별적으로 평균과 분산을

\* 인하대학교 컴퓨터공학과, 12191632Y@inha.edu,

\*\* 인하대학교 전기컴퓨터공학과 sokho0514@inha.edu, mysid88@inha.ackr(교신저자)

※ 본 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2022R1C1C1012408, 우수신진 연구)과 정보통신기획평가원의 지원(No.2022-0-00448/RS-2022-II220448, 인간처럼 회상이 가능한 인공 신경망 지속학습 플랫폼 개발, No.RS-2022-00155915, 인공지능융합혁신인재양성(인하대학교))을 받아 수행한 연구임.

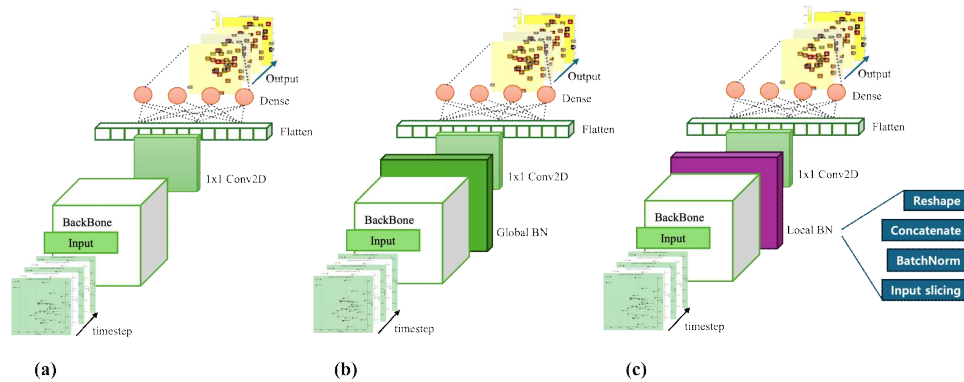


그림 588. 배치 정규화 적용 모델 구조도  
Fig. 1. Model Architecture with Batch normalization 있다.

계산해 정규화하기 때문에 시간 축에 따른 데이터의 일관성이 깨질 위험이 있다 [4]. 반면, 공간적 성능 개선을 위해 배치 정규화를 적용하면 각 측정소의 미세먼지 농도의 범위와 변화 패턴을 일정하게 조정하여 미세먼지의 공간적 분포가 왜곡되지 않도록 방지하며, 공간적 특성을 안정화시킬 수 있다.

본 논문에서는 지역적으로 보간 모델에 배치 정규화를 진행하여 미세먼지 농도 보간 모델의 성능을 향상시킨다. 또한 본 방식은 데이터의 양이나 모델의 크기를 증가시키지 않으면서도 보간 성능을 비용 효율적으로 향상시킬 수 있는 방식이다. ConvLSTM, CNN 보간 모델을 백본 모델로 실험한 결과, 보간 정확도가 지역 배치 정규화, 전역 배치 정규화, 기본 모델 순으로 지역 배치 정규화가 가장 높은 보간 성능을 보였다.

## II. 제안 방법

본 논문에서 제안하는 방식은 그림 1(c)와 같다. 구체적으로, 그림 1(a)와 같은 기존의 미세먼지 보간 모델에 지역적으로 배치 정규화를 수행하여 공간적인 특성을 반영한다. 기존 전역적인 배치 정규화 방식인 그림 1(b)와의 차이점은 각 공간적 위치에 대해서 배치 정규화를 독립적으로 수행한다는 점이다. 이는 측정소별 공간적 특성을 반영하지 못하는 전역적 배치 정규화 방식의 문제를 해결할 수

## III. 실험 및 결과

### 3-1. 데이터셋

본 실험에서 사용한 데이터셋은 Antwerp 지역 32개의 미세먼지 측정소에서 OPCN3PM25 센서를 이용해 1분 단위로 측정된 PM2.5 데이터이다 [5]. 모든 측정소에서 동시에 측정된 시간대의 데이터만을 사용했다. 데이터의 결측치는 IDW (Inverse Distance Weighting) 기법으로 보간했다. 각 측정소의 위도와 경도는 정규화하여 32x32 그리드에 미세먼지 측정값과 함께 저장된다. 본 실험에서 사용된 학습, 검증, 평가 데이터의 비는 8:2:2이다.

### 3-2. 실험 환경

본 논문에서는 실험을 위해 미세먼지 보간을 위한 백본 모델로 ConvLSTM [1]과 CNN [6]을 사용하였다. 각 백본 모델에 대해서 전역 정규화와 지역 정규화를 수행하고 그 결과를 비교하였다. 본 연구에서 보간 정확도를 측정하기 위해 사용한 평가 지표는 spRMSE로, 이는 측정소별로 보간 값과 실제 미세먼지 값의 RMSE값의 평균으로 계산된다. 나머지 실험 환경에 대한 세부 설정은 [1]과 [6]에서 제안한 방법을 따랐다.

표 1. 배치정규화 실험 결과 (spRMSE)  
Table 1. Batch normalization experiment result (spRMSE)

Backbone	ConvLSTM [1]	CNN [6]
Baseline	16.6354	18.6550

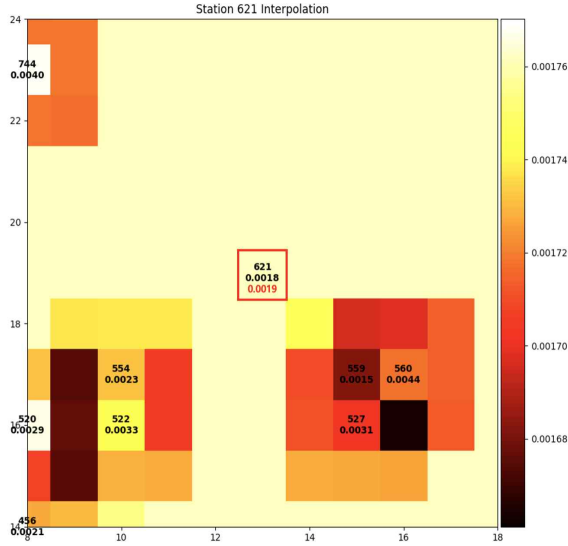


그림 2. 621 측정소 보간 결과 그리드 이미지  
Fig. 2. 621 station's Interpolation result grid image

### 3-3. 실험 결과

표 1은 두 백본 모델에 대해서 배치 정규화를 수행하지 않았을 때 (Baseline), 전역 배치 정규화를 수행했을 때 (Global BatchNorm), 그리고 지역 배치 정규화 (Local BatchNorm)를 수행했을 때 정량적인 비교실험 결과를 나타낸다. 표 1에서 볼 수 있듯이 두 백본 모델에 대해서 보간 정확도는 로컬 배치 정규화 모델, 전역 배치 정규화 모델, 기본 모델 순으로 높았다. 따라서 미세먼지 보간 모델에 전역 배치 정규화를 적용하는 것이 성능에 유의미한 영향을 끼칠 뿐만 아니라, 측정소별 지역 배치 정규화를 적용하는 방식이 더 효과적임을 보였다.

### 3-4. 사례 연구

그림 2는 특정 시점에서 621번 측정소의 주변 측정소를 기반으로 지역 배치 정규화를 적용한 보간 결과를 히트맵으로 나타낸 이미지이다. 그림에서의

색은 주변 측정소의 미세먼지 농도를 나타내며, 색이 짙을수록 미세먼지 농도가 높다. 텍스트는 측정소의 번호와 실제 미세먼지 농도를 나타낸다. 621번에서의 텍스트는 미세먼지 보간값이다. 621번 측정소의 보간값은 0.0018로 해당 시점의 실제 미세먼지 농도 값인 0.0019와 매우 근소한 차이를 보였다.

## IV. 결론

본 논문에서는 배치 정규화를 미세먼지 보간 모델에 활용했을 때 보간 성능이 올라간다는 가설을 입증한다. 특히 각 측정소별로 배치 정규화를 적용하면 각 측정소별 공간적 특성에 맞게 정규화가 일어나서 더욱 정밀한 보간 결과를 얻을 수 있다.

## 참고 문헌

- [1] V. D. Le and T. C. Bui, "Spatiotemporal Deep Learning Model for CityWide Air Pollution Interpolation and Prediction", IEEE BIGCOMP, pp. 55-62, February 2020.
- [2] D. Wong, L. Yuan and S. Perlin, "Comparison of spatial interpolation methods for the estimation of air quality data", nature journal of Exposure Science & Environmental Epidemiology vol. 14, pp. 404-415, September 2004.
- [3] S. H. Kim, et al. "Explainable AI-driven high-fidelity IAQ prediction model for subway stations", Building and Environment., vol. 263, pp. 112220, July 2024
- [4] C. S. Laurent, et al. "Batch Normalized Recurrent Neural Networks", in Proc. IEEE ICASSP, pp. 2657-2661, October 2015.
- [5] M. V. Poppel, et al. "SensEURCity: A multi-city air quality dataset collected for 2020/2021 using open low-cost sensor systems", nature scientific data, vol. 10, pp. 322, May 2023.
- [6] V. D. Le and S. K. Cha, "Real-time Air Pollution prediction model based on Spatiotemporal Big data", The International Conference on Big data, IoT, and Cloud Computing (BIC), April 2018

# Galois LFSR 기반 매칭 퍼즐 회로

한형규, 김현종, 한정윤, 공병용\*

## Galois LFSR-Based Matching Puzzle Circuit

Hyeong Gyu Han, Hyeon Jong Kim, Jeong Yun Han, and Byeong Yong Kong\*

### 요 약

노약자의 약 복용 보조와 치매 예방 및 뇌 훈련을 위한 색상 퍼즐 기능이 포함된 스마트 약통에 퍼즐 매칭 회로를 구현하였다. 퍼즐 기능에 용이한 Galois linear feedback shift register (LFSR)을 이용하여 난수 생성을 구현하였고, 저장과 매칭 회로를 추가한 구조를 제안하였다. 타이머의 클럭 타이밍 조절을 통해 난수 저장 기능을 구현하였고, 조합 논리회로를 병렬로 설계하여 입력 시퀀스와 난수 매칭을 효율적으로 확인하도록 최적화하였다. 매칭 여부에 따라 finite state machine (FSM)을 활용하여 타이머와 LFSR의 동작을 제어하고, 퍼즐 매칭 기능을 동작하는 방법과 구조를 제안하였다.

### Abstract

A puzzle matching circuit is implemented in a smart medicine cabinet with a color puzzle function for assisting the elderly in taking their medicine, preventing dementia, and training the brain. Random number generation is implemented using the Galois linear feedback shift register (LFSR) that is easy to use for the puzzle function, and a structure with additional storage and matching circuits is proposed. A random number storage function is implemented by adjusting the clock timing of the timer, and a combinational logic circuit is designed in parallel to optimize the efficient checking of the input sequence and the random number matching. Depending on the matching, the operation of the timer and LFSR is controlled using a finite state machine (FSM), and a method and a structure for operating the puzzle matching function are proposed.

### Key words

Circuit optimization, finite state machine, Galois linear feedback shift register (LFSR), timer clock control

## I. 서 론

노약자의 약 복용을 보조하고 치매 예방을 위한 뇌 훈련 기능이 포함된 퍼즐 매칭 회로 기반의 스마트 약통을, 난수 생성에 주로 사용되는 linear

feedback shift register (LFSR) 기반으로 구성하고자 하였다. 그러나 LFSR은 난수 생성 기능에만 집중되어 있어 퍼즐 매칭 회로 구현에 한계가 있다. 이를 보완하기 위해 LFSR의 난수 생성 기능을 최적화하고, 저장 기능과 매칭 및 제어 기능을 설계하여 구

\* 국립공주대학교 전기전자제어공학부 전자공학전공, bykong@kongju.ac.kr (교신저자)

※ 이 논문은 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1I1A3064200).

현하였다.

본 논문에서는 구현한 퍼즐 매칭 회로의 개념을 전체 회로 다이어그램을 바탕으로 설명하고, 설계 및 구현 과정을 단계별로 정리하여 기술하고자 한다.

## II. LFSR 기반 매칭 퍼즐 회로의 개념

그림 1을 바탕으로 전체 회로의 개념을 설명하고자 한다. 먼저 난수 생성에만 집중된 LFSR에 저장 기능이 필요하다. 타이머의 클럭제어 회로를 이용하여 Galois LFSR이 난수를 생성과 동시에 클럭 신호를 제어하여 생성된 난수를 저장한다.

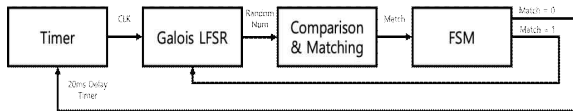


그림 1. 전체 회로 구성도  
Fig. 1. Overall circuit diagram

이후 사용자는 발생된 난수 신호를 랜덤한 퍼즐 형태로 확인하여 입력 시퀀스를 이용해 정답을 풀어나간다. 이 과정에서 매칭 여부를 판단하기 위하여 비교 매칭 회로가 필요하다, 비교 매칭 회로가 생성된 난수 신호를 확인하고 입력 시퀀스가 일치하는지 확인한다. 그 후에 퍼즐 회로의 정답 여부에 따른 이후 제어 과정이 필요하다. FSM을 이용하여 매칭 결과에 따라 타이머와 LFSR의 동작을 조정하며, 매칭 실패 시에는 새로운 난수를 생성하여 사용자가 다시 시도하도록 제안하고, 성공 시에는 시스템이 다음 단계로 넘어가서 다음 대기상태로 전환한다.

앞서 설명한 매칭 퍼즐 회로를 구현하기 위해, 효율적인 난수 생성과 저장, 비교 매칭이 가능한 퍼즐 회로의 핵심적인 기능을 제안한다.

## III. LFSR 기반 매칭 퍼즐 회로의 설계 및 구현

### 3.1 타이머의 클럭 제어를 활용한 저장 회로

LFSR은 난수를 생성 기능에 집중되어 있어 난수 저장 및 출력 기능이 필요하다. 이를 해결하고자 난수를 특정 시간 동안만 저장할 수 있는 회로가 필

요하다. 이를 구현하기 위해 불안정 모드의 555 타이머를 사용하여 High와 Low 신호를 반복적으로 출력한다. 타이머의 신호는 LFSR을 구성하는 D 플립플롭의 Enable 신호와 연결되어 High 신호에서는 난수가 진행되고, Low 신호에서는 LFSR 동작이 정지하여 일시적인 저장 회로를 대신할 수 있다.

그림 2는 타이머와 클럭 제어 회로 구성을 보여준다. 기존의 타이머 회로는 50% 이하의 듀티 사이클을 만들기 어렵기 때문에 [3], Rb에 바이패스 다이오드를 병렬 연결하였다.

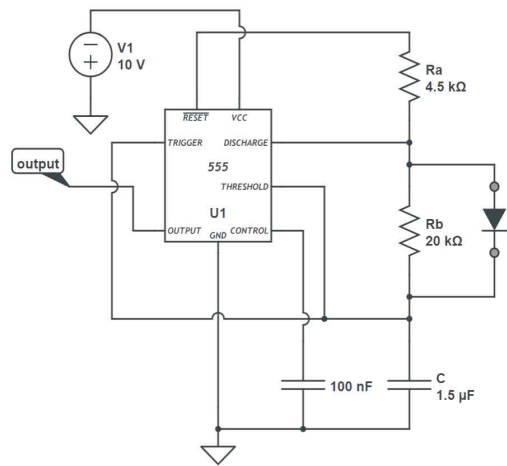


그림 2. 타이머 및 클럭 제어 회로  
Fig. 2. Circuit for timer and clock control

회로를 분석하면, 충전 경로에서 충전 커패시터의 전류는 Ra와 Rb를 통해 흐르고, 커패시터 전압이 0.667Vcc에 도달하면 방전된다. 충전 주기 동안 전류는 Ra와 다이오드를 통해 흐르고, 방전 주기 동안 전류는 Rb와 내부 방전 트랜지스터를 통해 흐르게 된다.

각 파라미터를 식 (1)에 대입하여 듀티 사이클을 계산할 수 있으며, 이 회로는 약 20% duty 사이클을 가진다. 이를 통해 난수의 출력 및 저장 시간과 딜레이를 조정할 수 있다 [4].

$$DutyCycle = \frac{(R_1)}{(R_1 + R_2)} \times 100(\%) \quad (1)$$

이 구조는 전체 주기 동안 LFSR 회로가 정지되어 있다가 충전 주기에만 LFSR이 활성화되고 새로운 난수를 생성하므로, 효율적으로 난수 생성을 제

어할 수 있다.

### 3.2 6비트 Galois LFSR 난수 생성 회로

퍼즐 회로의 랜덤 패턴을 6비트 난수로 생성하고 저장 및 매칭하는 회로를 설계하며, 스마트 약통의 퍼즐 매칭 회로 구현에 적합한 Galois LFSR를 적용한다. 피드백 다항식은 아래 (2)와 같고, n은 LFSR 플립플롭의 피드백 위치를, i와 j는 XOR 게이트가 연결된 난수 생성 피드백 비트의 위치를 나타낸다 [1].

$$P(x) = x^n + x^i + x^j + 1 \quad (2)$$

(2)를 기반으로 6비트 피드백 경로를 구성할 때, 두 번째와 네 번째 비트에 XOR 피드백을 적용하면 (3)과 같은 피드백 다항식이 도출된다.

$$P(x) = x^6 + x^4 + x^2 + 1 \quad (3)$$

(3)은 6비트 LFSR 구성을 나타내며, XOR 게이트가 두 번째 비트(Q2)와 네 번째 비트(Q4)에 연결되어 있음을 의미한다.

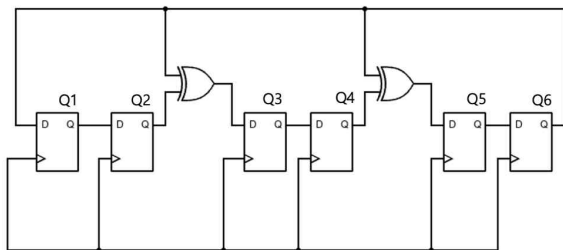


그림 3. 6비트 Galois LFSR 난수 생성 회로  
Fig. 3. 6-bit Galois LFSR random number generation circuit

그림 3은 6비트 Galois LFSR 난수 생성 회로 구성도이다. 6개의 D 플립플롭으로 구성되며, XOR 게이트는 두 번째 비트(Q2)와 네 번째 비트(Q4)의 출력과 연결된다. XOR 게이트의 출력은 첫 번째 플립플롭으로 피드백되며, 각 클럭 주기마다 시프트 연산을 통해 플립플롭 상태가 다음 플립플롭으로 이동한다. 이러한 과정에서 XOR 연산을 통해 새로

운 상태가 생성되고, 결과적으로 난수 비트열이 형성된다 [2]. 회로의 초기 상태에 따라 생성되는 난수 패턴은 달라지며, XOR 게이트의 위치와 개수에 따라 난수의 비트 패턴 역시 달라진다. Q2와 Q4에 연결된 XOR 게이트를 포함하는 이 회로는 최대 63 주기 동안 고유한 난수열을 생성할 수 있으며, 모든 상태를 순회할 때 고유한 비트 패턴을 출력한다.

이와 같은 6비트 Galois LFSR회로는 간단하면서도 빠른 난수 생성을 가능하게 하며, 난수 생성을 기반으로 랜덤한 패턴의 퍼즐 회로 구성에 적합하다.

### 3.3 입력 시퀀스와 매칭을 최적화한 회로

타이머의 클럭 제어를 통하여 저장된 난수 형태의 퍼즐과 사용자의 입력 시퀀스가 일치하는지 여부를 확인하는 효율적인 회로를 설계한다. 이 과정에서 난수와 입력 시퀀스의 매칭 과정 최적화에 중점을 둔다.

그림 4에서는 입력 시퀀스를 High와 Low 값으로 구분하여 설계된 회로 구성을 보여준다. 이 회로는 각 비트마다 난수와 입력 시퀀스의 일치 여부를 확인하고 그 결과를 출력한다. 매칭 검증의 속도를 높이기 위해 상위 비트부터 일치 여부를 검사하며, 상위 비트에서 불일치가 발생하면 즉시 매칭되지 않았음을 출력한다. 반대로 상위 비트가 일치하면 하위 비트까지 매칭 검증을 계속 진행한다.

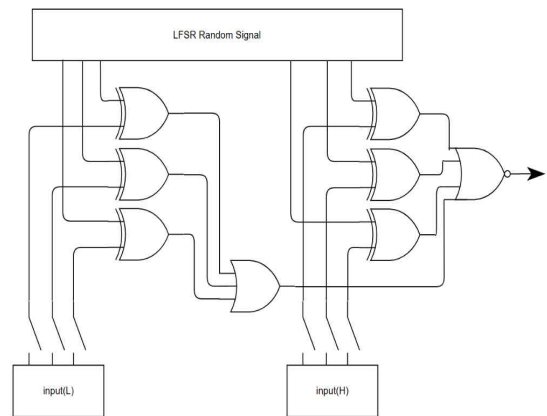


그림 4. 입력 시퀀스와 LFSR 난수 매칭 확인 회로도  
Fig. 4. Circuit diagram for matching the input sequence with LFSR random number sequence

이러한 구조를 통해 비교 매칭 과정에서 효율성을 향상시키고, 빠르고 정확한 결과를 얻을 수 있도록 설계하였다. 퍼즐 매칭 회로에 최적화된 이 설계는 퍼즐 매칭의 정확도를 높이고 회로의 성능을 극대화한다.

### 3.4 매칭 확인 후 제어를 위한 FSM 설계

퍼즐 매칭 회로의 효율성을 높이기 위해 FSM을 사용하여 입력 시퀀스와 생성된 난수 퍼즐 간의 정답 매칭 여부에 따른 상태 변화를 제어한다. FSM은 매칭 여부에 따라 상태를 변화시키며, 각 상태에서는 타이머와 LFSR의 동작이 달라진다.

그림 5는 FSM을 이용한 제어 방식을 나타내고 있다. 우선 입력 시퀀스와 생성된 난수가 일치하지 않을 경우, FSM은 새로운 난수 생성을 위해 타이머를 정지하고 LFSR의 상태를 초기화한다. 반대로 매칭이 성공하면 일정 시간 동안 대기한 후 새로운 난수를 생성하도록 한다. 이 과정에서 EN 신호는 LFSR의 활성화 상태를 제어하며, FSM이 매칭 과정에서 불필요한 연산을 방지하도록 설계되어 시스템의 효율성을 높인다.

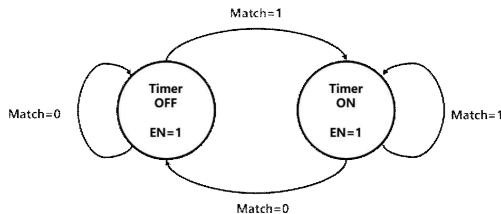


그림 5. 난수 매칭 확인 후 제어 FSM 다이어그램  
Fig. 5. Control FSM after random number matching

## IV. 구현 방식 비교

Fibonacci LFSR은 순차적으로 계산이 이루어져 회로가 복잡해질수록 처리 속도가 감소하는 단점이 있다. 저장하기 위해 많은 클럭 사이클을 기다려야 한다는 단점이 있지만 Galois LFSR은 병렬 계산을 통해 회로 길이와 관계없이 일정한 처리 속도를 유지할 수 있다. JK 플립플롭은 많은 게이트와 복잡한 제어를 필요로 하므로, 비교 매칭 회로에 적합하지 않다. 그러나 딜레이 회로를 사용하면 일정 시간

동안 값을 저장할 수 있으므로 회로 복잡도를 최적화하여 처리 속도를 개선할 수 있다. 난수를 저장하기 위한 별도의 저장 회로 대신 딜레이 회로를 사용하여 시프트 작동을 중지시킴으로써 효율성을 높일 수 있다 [6].

표 1. 구현 방식 비교

Table 1. Comparison of implementation methods

	기존 방식	제안한 방식
LFSR 방식	Fibonacci	Galois
저장 회로	J-K플립플롭 기반	NE555기반 클럭제어
비교, 매칭 방식	직렬형	병렬형 비트제어
비교 매칭 회로	IC 칩	논리 게이트
매칭 후 제어방식	조합논리	FSM 상태전이
속도	지연된 처리	빠른 처리

기존의 비교 매칭 회로는 4비트 비교기 IC 칩인 IV74LS85를 사용했다. 이 칩은 세 가지의 크기 비교 기능을 제공했지만, 제안한 회로에서는 크기 비교가 불필요했다. 구현한 방식에서는 일치 여부만을 처리하는 8개의 논리 게이트를 사용해 회로를 간소화하고 최적화했다. 조합 논리 회로는 설계가 단순하지만 복잡해질 가능성이 있어, 유한 상태 기계 (FSM)를 도입해 사용하는 소자의 수를 줄이고 하드웨어적 변수를 최소화하여 시스템의 안정성을 높였다.

종합하면 Galois LFSR의 병렬 처리가 처리 속도와 효율성을 크게 개선했다. 저장 방식 간소화, 비교 방식 병렬화, 그리고 FSM을 통한 제어 도입이 시스템 성능 향상에 기여했다.

## V. 결론

본 논문에서는 Galois LFSR 기반의 난수 생성 회로를 개선하고 저장 및 매칭 기능을 추가하여 퍼즐 형태의 매칭 회로를 제안하였다. FSM을 통해 입력 시퀀스와 생성된 난수의 매칭 상태에 따라 LFSR과 타이머의 동작을 제어하는 방식은 기존의 복잡한 구조를 간소화하면서도 퍼즐 회로의 랜덤성을 향상시켰다. 설계된 회로는 하드웨어와 소프트웨어에서



모두에서 효율적으로 동작하며, LFSR이 생성한 난수와 사용자의 입력 시퀀스를 비교하고 매칭하는 회로를 구성하였다. FSM을 추가하여 매칭 여부에 따른 시스템 제어를 하도록 구성하였다. 향후 연구에서는 LFSR의 패턴 예측 가능성을 개선하여 초기 상태와 선형 피드백 구조를 개선하여, 시스템 보안에 적용할 수 있도록 퍼즐 회로를 고도화할 예정이다. 또한 다른 상태를 가진 여러 개의 LFSR의 난수를 서로 비교하는 방식으로, LFSR의 문제를 해결할 수 있을 것으로 기대된다. 이를 통해 복잡하고 예측 불가능한 패턴을 구성하여 퍼즐 매칭 회로의 성능 개선과 보안 분야에 적용하여 발전할 수 있을 것이다.

### 참 고 문 헌

- [1] A. Chakraborty, B. Mazumdar, and D. Mukhopadhyay, "Fibonacci LFSR vs. Galois LFSR: Which is more vulnerable to power attacks?," in Proc. Int. Conf. Secur. Priv. Appl. Cryptogr. Eng. (SPACE), Pune, India, pp. 14-27, Oct. 2014.
- [2] P. Mittal and D. Shah, "Linear feedback shift register-based test pattern generators: A comparative study," Int. J. Adv. Res. Electron. Commun. Eng., vol. 9, no. 8, pp. 60-62, Aug. 2020.
- [3] H. H. Ward, "Introduction to the 555 timer," Mast. Digit. Electron., pp. 409-436, Dec. 2023.
- [4] M. Sulthan, S. R. Chowdury, R. Garg, and A. Tripathi, "Design of a power efficient pulse latch circuit as a solution for master slave flip-flop," in Proc. Int. Symp. VLSI Des. Test (VDAT), Indore, India, pp. 532-540, Jul. 2019.
- [5] N. Camaron, "Signal generation with 555 timer IC," Electron. Proj. ESP8266 ESP32, pp. 467-503, Dec. 2020.
- [6] Conversion between Galois and Fibonacci polynomials of linear-feedback shift register, <https://www.01signal.com/other/lfsr-galois-fibonacci/>
- [7] Texas Instruments, "NE555 data sheet, product information and support," Sep. 2023.
- [8] Microcontrollers Lab, "8-bit comparators example with 74LS85," Jul. 2014. <https://microcontrollerslab.com/74ls85>
- [9] P. Rogers, "Design low-duty-cycle timer circuits," Aug. 2002. <https://www.edn.com/design-low-duty-cycle-timer-circuits/>

# AR 기술과 Unity를 활용한 한국 멸종 위기 동물 콘텐츠

안정빈\*, 양인지\*, 변공규\*\*, 유선진\*\*\*

## Korean Endangered Animal Content Using AR Technology and Unity

Jeongbin Ahn\*, Inji Yang\*, Gongkyu Byeon\*\*, and Sunjin Yu\*\*\*

### 요 약

최근 급속한 산업화와 도시화로 인해 많은 동식물이 서식지를 잃고 멸종 위기에 처하였다. 본 논문은 AR(Argumented Reality) 기술을 활용하여 한국의 멸종 위기 동물을 소개하는 콘텐츠를 제작하고, 유니티(Unity) 엔진과 뷰포리아(Vuforia)의 이미지 추적 기술을 활용하여 멸종 위기 동물과 관련한 정보를 전달한다. 이를 통해 관람객에게 멸종 위기에 처한 동물들의 현 상황에 대해 알리고, 그들이 직면한 위험 요소들을 소개하여 인식의 폭을 넓힌다.

### Abstract

Due to the recent rapid industrialization and urbanization, many animals and plants have lost their habitats and are in danger of extinction. This paper produces content that introduces endangered animals in Korea using AR (Argued Reality) technology and delivers information related to endangered animals using Unity engine and Vuforia's image tracking technology. Through this, visitors are informed of the current situation of endangered animals and the risk factors they face are introduced to broaden their awareness.

### Key words

endangered animal, augmented reality, unity, vuforia

## 1. 서 론

생태계의 파괴와 오염으로 해마다 수만 종이 사라지고 있다. 서식지 파괴는 인간이 자연을 개발하

는 과정에서 주로 일어난다[1]. 특히 현대에 와서는 대량 멸종사태에 접어들어 이런 속도라면 머지않아 전체의 25%가 멸종할 것으로 어떤 과학자들은 전망하기도 한다[2].

이에 따라, 본 논문은 관람객들에게 멸종 위기

\* 국립창원대학교 문화테크노학과, gajm7373@gs.cwnu.ac.kr, 20214518@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화융합기술협동과정 박사과정, bgk9517@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr (교신저자)

※ “이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.”

동물이 처한 상황과 그들이 직면한 위험 요소를 효과적으로 전달하고 인식의 폭을 높일 수 있도록 AR(Argumented Reality), 유니티(Unity) 엔진, 뷰포리아(Vuforia) 기술을 활용하여 콘텐츠를 구상한다.

## II. AR 콘텐츠 제작

첫 번째 콘텐츠는 ‘멸종 위기 동물’ 을 주제로 한 AR 포스터 6종이다. 인간의 영향을 많이 받는 멸종 위기 1급 동물들로 각각 포유류, 조류, 어류에서 선정한다.

그림 1은 선정한 멸종위기 동물의 포스터 디자인이다. 관람객이 안드로이드(Android)와 IOS 태블릿 PC를 사용하여 포스터를 인식하면 각 동물의 이름, 서식지, 분포, 특성 등에 관한 정보가 뜬다.

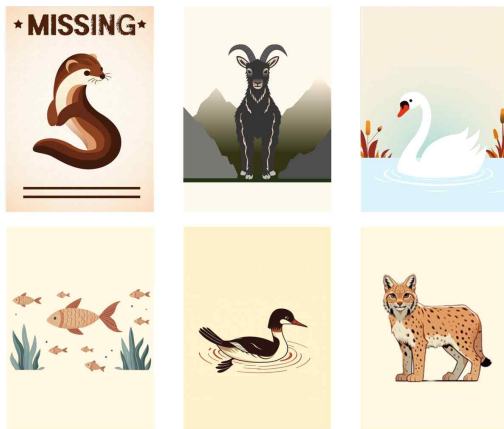


그림 1. AR 포스터  
Fig. 1. AR poster

두 번째 콘텐츠는 ‘멸종 위기 동물 보호의 필요’ 에 대한 애니메이션이다. 이미지 추적용 포토카드 3개를 이용하여 각 애니메이션을 재생한다. 애니메이션은 멸종위기 동물을 보호해야 하는 이유, 보호 방법, 보호 방법을 통해 개체수가 증가한 사례로 구성된다. 유니티의 각 씬에서 특정한 포토카드 이미지가 인식되면, VideoPlayer가 해당 이미지 타겟에 등록된 영상을 재생하도록 구현한다.

## III. 유니티 퀴즈 게임

그림 2는 관람객이 플레이 할 수 있는 멸종위기 동물에 관한 O/X 퀴즈 게임이다. 시작 버튼을 누르면 유니티의 버튼 기능이 활성화되며 씬이 전환되고 퀴즈가 진행된다. 힌트 보기 버튼을 누르면 힌트가 제공되고, 정답을 맞추면 점수가 카운팅된다. 퀴즈가 끝나면 자신의 점수가 화면에 뜨며 다시 하기 버튼이 활성화된다.

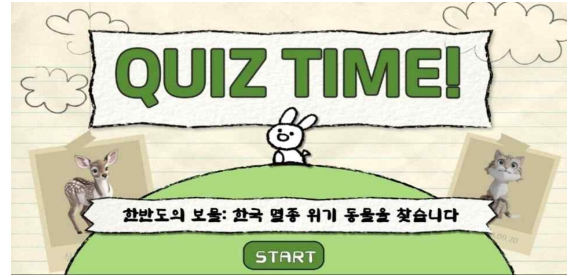


그림 2. 멸종 위기 동물 퀴즈 게임  
Fig 2. Endangered animal quiz game

이 콘텐츠를 통해, 관람객들은 멸종위기 동물에 대한 인식과 지식을 자연스럽게 점검할 수 있고, 학습 효과를 높일 수 있다.

## IV. 결론

본 논문에서는 AR 기술, 유니티 엔진, 뷰포리아를 활용하여 관람객들에게 한국의 멸종위기 동물을 소개하고 관련된 정보를 제공하는 실감 콘텐츠를 제작했다. 이를 통해 한국의 멸종위기 동물 보호에 대한 사회적 인식을 제고하였다. 향후 기존 콘텐츠에서 나아가 관람객이 직접 멸종위기 동물 보호 활동에 참여할 수 있는 체험 프로그램을 마련하여 장기적인 관심을 유도할 수 있다.

## 참 고 문 헌

- [1] H. S. LEE, Biodiversity for Youth 02 ,Korea Institute for Environmental Policy Evaluation, pp.3 , 2020.
- [2] K. H. Ahn, Animals faced with extermination dangers and the restoration methods for them (I) : ASLE-KOREA, vol.8, no.1, pp.29 - 53, 2009.

# 프로야구의 팬 유입을 위한 타자 체험 가상현실 콘텐츠

이유진\*, 우동현\*\*, 유선진\*\*\*

## Enterprise Experience Virtual Reality Content for the Introduction of Professional Baseball

Youjin Lee\*, Donghyun Woo\*\*, and Sunjin Yu\*\*\*

### 요 약

가상현실의 몰입감을 활용하여 사용자가 직접 야구를 체험하고 그 규칙을 이해할 수 있는 VR 콘텐츠를 제안한다. 가상현실 기술을 통해 사용자는 실제 경기장에 있는 것 같은 생동감을 느끼며, 경기의 흐름과 규칙을 자연스럽게 학습할 수 있다. 이를 통해 야구에 대한 관심과 친숙함을 높여 신규 팬의 유입과 기존 팬의 충성도를 강화하고, 나아가 야구 홍보의 효과성을 증진시키고자 한다.

### Abstract

Using the immersion in Virtual reality, it provides VR content to experience and attractive baseball content that allows users to understand the rules. Through virtual reality technology, users can feel the vitality as if they are in a real stadium and learn the flow and rules of the game naturally. Through this, it aims to increase interest and familiarity with baseball, strengthen the influx of new fans and the loyalty of existing fans, and further increase the effectiveness of baseball promotion.

### Key words

virtual reality, interactive media, unity, baseball promotion

### 1. 서 론

최근 1년간 KBO(Korea Baseball Organization) 리그는 팬들의 큰 관심을 받으며 흥행에 성공하였고, 2024년에는 천만 관중을 달성하며 KBO 리그가 국내 최고의 스포츠 콘텐츠 중 하나로 자리 잡았

다.[1] 그러나 현재 팬층뿐 아니라 새로운 팬층 유입을 위해서는 더욱 몰입도 높은 홍보 콘텐츠가 요구된다. 야구는 경기 규칙에 대한 이해가 중요한 스포츠로, 팬층을 확대하기 위해서는 보다 체험적인 콘텐츠의 도입이 필수적이다. 가상현실(Virtual Reality, VR)은 사용자에게 높은 몰입감과 실재감을

\* 국립창원대학교 문화테크노학과 학부생, youjin2224@gs.cwnu.ac.kr

\*\* 국립창원대학교 첨단방위공학과정 박사과정, dhwoo1119@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr(교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.

제공하여 팬들이 야구에 친숙해지고 쉽게 이해하며 즐길 수 있는 환경을 조성하는데 효과적이다. 본 논문은 가상현실 기술을 활용한 타자 체험 콘텐츠를 제안하고자 한다.

## II. 야구 배팅 실감형 콘텐츠

본 가상현실 야구 콘텐츠는 Unity 게임엔진과 Oculus Integration SDK를 기반으로 개발하였다. 사용자는 Meta社의 Meta Quest 2 VR HMD를 착용하고, 대형 화면으로 출력된 가상현실 환경에서 야구 타격을 체험한다. 본 콘텐츠에서 사용하는 야구 배트는 3D 프린터를 통해 제작했으며, 오른쪽 Oculus 컨트롤러와 결합하였다. 컨트롤러 및 설계한 장비는 그림 1과 같다.



그림 597 야구 콘텐츠 사용 장비  
Fig. 1 Baseball content usage equipment

체험자는 가상현실 속에서 투수가 던진 공을 배트로 치는 경험을 하게 된다. 배트를 휘두르면 공이 Unity의 물리 엔진을 기반으로 설정된 방향과 속도로 날아가도록 구현한다. 사용자는 오른손으로 배트를 잡은 상태에서 VR 컨트롤러를 조작하여 체험할 수 있다. 배트의 중심점 위치를 VR 컨트롤러와 정확히 맞춰 타격 동작이 자연스럽게 이루어지도록 한다.

체험 콘텐츠는 타격체험, 몰입감 강화 기능, 결과 판정의 3단계 시나리오로 구성되어 있으며, 각 단계마다 다음과 같은 주요 기능들이 적용된다. 타자가 투수로부터 공을 받아 타격하는 경험을 제공한다. 그림 2와 같이 투수의 공 발사는 BallManager 스크립트를 통해 설정된 속도와 방향으로 이루어지며, 배트와 공이 충돌할 때 공이 날아가는 방향과 거리

를 Unity 물리 엔진으로 계산하여 화면에 표시한다. 야구 타격 시, 타격의 성공 여부에 따라 컨트롤러 햅틱 진동 효과와 화면 밝기 변화를 구현하여 몰입감을 높였다. 타격 후 공이 날아간 거리를 기준으로 홈런, 안타, 헛스윙 결과가 판정되며, 결과가 콘솔에 출력되고, 화면에도 텍스트로 표시된다. Oculus Controller와 연동하여 오른손으로 배트를 휘두를 때 결과가 화면에 표시되도록 설계한다.

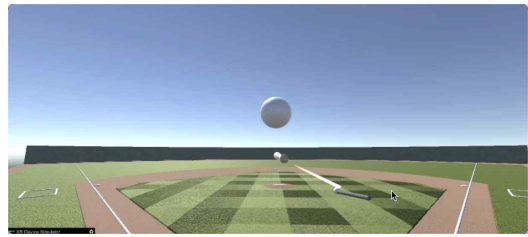


그림 598 공이 날아오는 씬 화면  
Fig. 2 The scene where the ball flies

## III. 결 론

본 가상현실 콘텐츠에서는 사용자가 실제 야구 경기의 상황을 체험함으로써 야구의 규칙과 재미를 직접 경험하고 이해하도록 돕는다. 이를 통해 야구에 대한 친숙함과 흥미를 높이며, 새로운 팬층의 유입과 기존 팬의 충성도 강화를 도모할 수 있다. 또한, 사용자가 실제 경기에 몰입하여 다양한 타격 상황을 경험함으로써 팬 경험의 질을 높이고 스포츠에 대한 집중력과 이해도를 향상시키는 데 기여할 수 있다. 향후에는, 엔터테인먼트 요소를 포함한 교육 및 홍보용 VR 게임 프로그램으로 발전시켜 나갈 예정이다.

## 참 고 문 헌

- [1] Y. S. Kim. Trend Korea 2025 : A new perspective view. The window of the future.
- [2] W. Y. Yoo. "A Study on Reality Enhancement Method of VR Baseball Game". Journal of Korea Game Society, Vol.19, No.2, pp. 23-32, 2019.
- [3] <https://www.sportsworldi.com/newsView/20241029516120> [Accessed: 2024.10.29.]

# VR 기반 나무 심리 테스트와 개인화된 힐링 방법 제안

김가영\*, 강인주\*\*, 유선진\*\*\*

## VR-based Tree Psychology Tests and Personalized Healing Methods Proposed

Gayeong Kim\*, Inju Gang\*\*, and Sunjin Yu\*\*\*

### 요 약

최근 VR을 활용한 심리 치료에 대한 관심이 증가하고 있으며, 본 연구는 VR과 심리 테스트를 결합한 콘텐츠를 개발하여 사용자 맞춤형 힐링 방법을 제공하는 것을 목표로 한다. 테스트는 Unity와 Meta Quest 2를 사용하여 진행되며, 사용자가 나무를 형상화하는 과정을 통해 심리 결과를 도출한다. 이를 통해 사용자들은 자신의 상태를 인지하고 개인화된 힐링 방법을 제안받는다. 향후 연구에서는 인공지능과 데이터 분석 기술을 활용하여 사용자 맞춤형 피드백을 강화하고 심리 상태 변화를 실시간으로 추적하는 시스템을 개발할 예정이다.

### Abstract

Recently, interest in psychotherapy using VR(Virtual Reality) is increasing, and this study aims to provide a customized healing method by developing content that combines VR and psychological tests. The test is conducted using Unity and Meta Quest 2, and psychological results are derived through the process of users shaping trees. Through this, users are aware of their conditions and are suggested personalized healing methods. Future research will develop a system that strengthens user-customized feedback and tracks changes in psychological conditions in real time using artificial intelligence and data analysis technology.

### Key words

virtual reality, psychotherapy, personalized healing methods, unity

### 1. 서 론

최근 VR(Virtual Reality)을 활용한 심리 치료에 대한 관심이 급증하고 있다. VR에서의 경험을 통해

심리적 불안이나, 공포, 트라우마에 대한 치유적 접근이 이루어지고 있으며 기술 발전이 이러한 흐름을 더욱 가속화하고 있다.[1] 시장조사업체 ‘마켓앤마켓’에 따르면, 글로벌 디지털 치료제 시장은

\* 국립창원대학교 문화테크노학과 학부 과정, 20214505@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화융합기술협동과정 석사 과정, 20247306@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@changwon.ac.kr(교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스티원센터의 지원을 받아 수행된 연구입니다.

2021년 34억 달러에서 2026년 131억 달러에 이를 것으로 전망하고 있다.[2] 특히 코로나19 팬데믹 이후, 환자들은 질병 진단 및 치료를 위한 디지털 치료제에 더욱 주목하기 시작했다. 국내외에서 VR을 활용한 다양한 치료 콘텐츠가 개발되고 있는 가운데, 본 연구는 이러한 추세를 반영하여 VR과 심리 테스트를 결합한 콘텐츠를 제작하고, 사용자 개인에 맞는 힐링 방법을 탐구하고 제공하는 것을 목표로 한다.

## II. VR 심리 테스트 및 힐링 방법 제공

본 심리 테스트는 한국아동심리코칭센터에서 제공하는 나무 해설과 CPR 색채 심리 검사를 바탕으로 제작되었으며, Unity 게임 엔진과 Meta Quest 2를 사용한다. 나무는 일상에서 쉽게 접할 수 있는 친숙한 소재이며, 녹색은 안정감과 편안함을 제공하여 테스트의 소재로 선택되었다. 테스트가 실행되면 Meta Quest 2의 컨트롤러를 사용하여 나뭇가지, 나뭇잎, 열매 등을 통해 심리 결과를 도출하는 방법이 설명된 캔버스가 열린다. 이후 그림 1과 같이 이름을 입력할 수 있는 창이 나타나며, 이름 입력을 원하지 않는 사용자에게는 형용사와 동물을 결합한 랜덤 이름 15개 중 하나가 자동으로 생성되어 보여진다. 예를 들어, ‘미소 짓는 고양이’, ‘떠드는 앵무새’와 같은 이름이 제공될 수 있다.

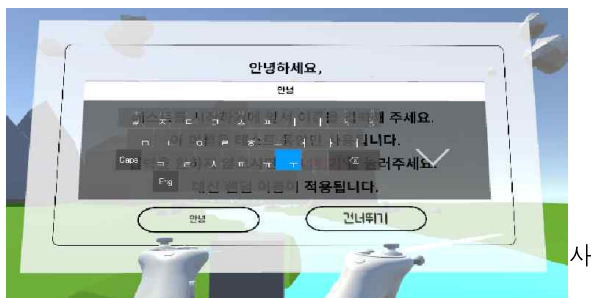


그림 599. 이름을 입력할 수 있는 키보드 창  
Fig. 1. Keyboard window where you can enter a name

용자는 먼저 여러 개의 나뭇가지 후보 중에서 자신이 마음에 드는 하나를 선택한다. 나뭇가지를 선택한 후, 컨트롤러의 Grab 기능을 이용하여 원하는 위치에 나뭇잎을 배치할 수 있으며, 나뭇잎의 색상

과 모양도 자유롭게 선택할 수 있다. 나뭇잎과 나뭇가지의 거리가 0.1f 이하일 경우, 나뭇잎이 나뭇가지에 자석처럼 붙도록 하는 코드를 각 나뭇가지에 적용하였다. 다음 단계에서는 사용자가 열매나 꽃을 선택할 수 있으며, 두 가지 모두 원하지 않는 경우, 테스트를 종료할 수 있다. 심리 테스트는 나뭇가지의 굵기와 개수, 나뭇잎의 색상과 개수, 열매와 꽃 등의 다양한 요소를 통해 결과를 도출한다. 모든 테스트가 완료되면 사용자는 자신이 완성한 나무와 해설을 확인할 수 있으며, 그에 맞는 개인화된 힐링 방법이 제공된다.

표 1. 테스트 결과에 따른 힐링 방법 예시  
Table 1. Examples of healing methods based on test results

테스트 결과	힐링 방법
- 번덕과 소심함	1) 일기 쓰기 & 명상하기
- 도전적이고 자신감 넘치는 성격	2) 단기 목표와 장기 목표 설정
- 무리한 도전이나 과도한 열정 스트레스 유발	3) 순간순간을 즐기는 법을 배우기

## III. 결 론

본 연구에서는 VR을 활용한 심리 테스트를 통해 사용자들이 자신의 상태를 인지하고, 각 개인에 맞는 힐링 방법을 제공하는 것을 목표로 제작하였다. 향후 연구에서는 인공지능 및 데이터 분석 기술을 접목하여 사용자 맞춤형 피드백을 강화하고, 심리 상태 변화를 실시간으로 추적하는 시스템을 개발할 예정이다.

## 참 고 문 헌

- [1] S. R. Rim, "Analysis of Characteristics of Therapeutic Factors of VRAT(Virtual Reality Art Therapy) Contents", The Korea Contents Society Vol.22, pp. 2, April 2022.
- [2] <http://www.atlasreview.co.kr/news/articleView.html?idxno=1591> [Accessed: 2024.10.28]

# 유니티를 활용한 한글 게임 교육 콘텐츠

김솔\*, 배지우\*, 이세진\*\*, 유선진\*\*\*

## Korean game education contents using Unity

Sol Kim\*, Jiwoo Bae\*, Sejin Lee\*\*, and Sunjin Yu\*\*\*

### 요약

SNS와 방송 매체에서 신조어를 확산시키면서 한글 파괴를 가속화하고 있다. 이는 언어의 본질인 '소통'을 저해하는 원인이다. 본 논문에서는 이를 완화 시키고자 한글 게임을 제작하고 올바른 한글을 사용할 수 있는 게임 콘텐츠를 제안한다. 게임 엔진 유니티(Unity)를 사용하여, 안드로이드(Android) 태블릿 PC 환경으로 개발하였다. 이를 통해 사용자는 교육적인 접근으로 바람직한 언어 사용의 중요성을 인식한다.

### Abstract

The spread of neologisms on social media and broadcasting media is accelerating the destruction of the Korean language, which hinders the essence of language: communication. This paper proposes the development of a Korean language game to mitigate this issue and suggests game content that encourages the proper use of the Korean language. Using the Unity game engine, the game was developed for the Android tablet PC environment. Through this, users can recognize the importance of using language appropriately through an educational approach.

### Key words

unity, VR contents, neologisms, korean language educational game

## 1. 서론

오늘날 현대사회는 SNS와 방송 매체의 발달로 신조어를 확산시키며 한글 파괴 현상을 가속화하고 있다[1]. 이러한 비표준적 언어의 남용은 언어의 본질인 '소통'을 저해할 뿐 아니라, 세대 간 의사소통의 단절을 초래한 사회적 문제로 이어지고 있다[2].

따라서 한글 파괴 현상에 대한 인식을 파악하고자 20대 140명을 대상으로 설문조사를 실시하였다. 그 결과, 응답자의 81%가 일상에서 신조어를 자주 사용하며, 65%가 이러한 표현을 SNS와 미디어로 습득하는 것을 확인하였다.

본 연구는 이러한 문제의식에서 출발하여, 비표준적 언어 사용이 한글의 가치를 훼손하지 않도록

\* 국립창원대학교 문화테크노학과, 20200247@gs.cwnu.ac.kr, 20200115@gs.cwnu.ac.kr

\*\* 국립창원대학교 문화융합기술협동과정 석사과정, 20237203@gs.cwnu.ac.kr,

\*\*\* 국립창원대학교 문화테크노학과 교수, siyu@changwon.ac.kr (교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스티원센터의 지원을 받아 수행된 연구입니다.



게임 개발 엔진 유니티(Unity)를 활용한 한글 교육 게임을 제작하였다. 이는 흥미를 기반으로 한 교육적 접근으로 관람자가 올바른 언어 사용의 필요성을 체감하고, 올바른 언어를 사용하는 것을 목표로 한다.

## II. 유니티를 활용한 한글 게임 콘텐츠 제작

이 콘텐츠의 개념트는 사용자가 국어 영역 모의 고사를 치는 것이다. 이 콘텐츠를 시작하기 전, 현재 한글 사용 실태의 부정적인 모습에 대한 영상을 본 이후 콘텐츠를 실행하며, 본인의 언어능력을 실감할 수 있게 된다.

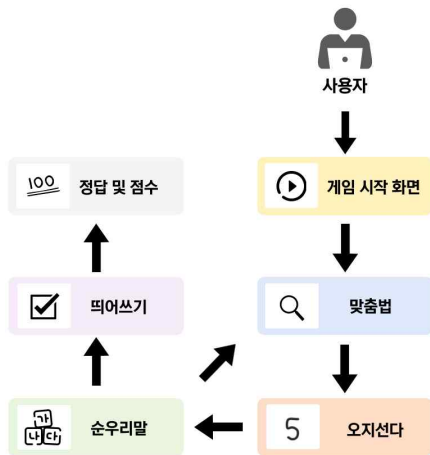


그림 600 콘텐츠 진행 순서도  
Fig. 1. The flow chart of the contents

그림 1은 콘텐츠가 진행되는 순서이다. 사용자가 태블릿 PC를 이용하여 게임을 시작한다. 시작 화면에서 성명을 작성하고, 확인란에 제시된 문구를 동일하게 쓰면 자동으로 첫 번째 문제로 넘어간다. 몰입도를 증가시키기 위해 1초 간격을 두고 대화하는 형식으로 진행되며, 대화 끝에 문제가 제시된다. 또한, 사용자는 올바른 띄어쓰기와 맞춤법, 순우리말을 선택하여 문제를 풀면 정답 여부가 즉시 나타나는 방식이다. 이렇게 문제 4가지가 반복되며 16문제가 진행된다. 정답을 확인 후 ‘다음’ 버튼을 누르면 다음 문제로 넘어간다. 마지막 문제를 풀 이후 모든 문제의 정답과 함께 본인의 점수가 나타나며 게임을 마무리함으로써 콘텐츠가 종료된다.

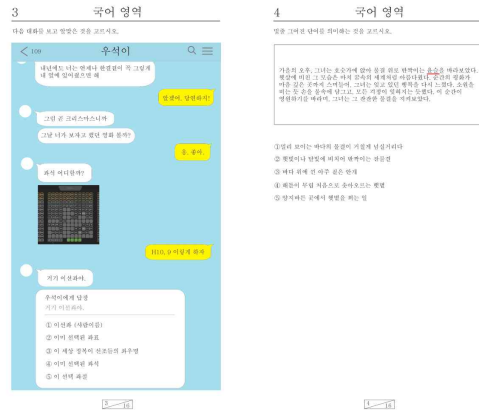


그림 2. 콘텐츠 문제 형식 예시  
Fig. 2. Content problem format

그림 2는 콘텐츠 문제 형식 예시로 대화 형식 문제 6개, 오지선다 문제 8개, O/X 문제 2개로 구성되어 있다. 이러한 형식으로 문제의 구성은 순우리말, 줄임말, 맞춤법, 띄어쓰기 총 16문제로 다양한 문제를 제공한다. 이 모든 과정을 체험한 20대 사용자는 본인의 평소 한글 사용에 대한 경각심을 느꼈다.

## III. 결 론

본 콘텐츠에서는 문제를 통해 본인의 한글 사용의 능력을 진단하며 올바르지 않은 한글 사용에 대한 경각심을 제공한다. 이를 통해 올바른 언어 사용을 높이는 데 도움을 줄 수 있다. 향후 어린이들을 대상으로 난이도를 설정하여 콘텐츠를 체험하게 하고 어린이들의 어휘력 향상에 도움이 되도록 교육용 콘텐츠로 발전할 예정이다.

## 참 고 문 헌

[1] K. T. Koo. "Study on the audience's perception of the use of sanctioned language in broadcasting: Focusing on program genre and demographic factors." Region and world vol.44, no.4 2020.

[2] M. Y. Bang. and K. W. Lee, "Problems of using new words in the media and ways to improve them." Convergence of cultural technologies vol.8, no.1, pp.191-200 2022.

# 펫로스 증후군 극복을 위한 PC 콘솔 게임 콘텐츠

박재환\*, 이준호\*, 소우주\*, 강인주\*\*, 유선진\*\*\*

## PC Console Game Contents to Overcome Pet Loss Syndrome

Jaehwan Park\*, Junho Lee\*, Woojoo So\*, Injoo Kang\*\*, and Seonjin Yoo\*\*\*

### 요약

펫로스 증후군은 반려동물의 죽음 후 겪는 심리적 고통을 의미하며, 현대인에게 정서적 회복이 필요한 영역으로 부각되고 있다. 본 연구는 Unity 게임 엔진을 활용하여 불교적 사후 세계를 배경으로 한 반려견의 여정을 통해 상실과 추억을 되새기며 죽음을 수용하도록 설계된다. 인벤토리 알고리즘, 플랫폼 알고리즘, UI 알고리즘 Scene 이동 알고리즘을 사용한다. 게임의 진행은 환생으로 마무리되어, 죽음이 새로운 시작임을 전달하며 심리적 치유를 유도한다.

### Abstract

Pet loss syndrome refers to the psychological distress experienced after the death of a pet, highlighting a need for emotional recovery in modern society. This study utilizes the Unity game engine to create a journey through a Buddhist-inspired afterlife, where players, embodying a dog, revisit memories of their owner and come to terms with loss and acceptance of death. The game employs inventory algorithms, platform algorithms, UI algorithms, and scene transition algorithms. The journey concludes with reincarnation, conveying death as a new beginning and guiding players toward emotional healing.

### Key words

pet loss syndrome, unity game engine, psychological distress, game contents

## 1. 서론

펫로스 증후군은 반려동물을 가족처럼 여기는 현대인의 정서적 애착이 깊어짐에 따라, 반려동물의 죽음 이후 경험하는 상실감과 슬픔, 그리고 죄책감과 같은 복합적 감정에서 비롯되는 심리적 고통을 의미한다. 이 증후군은 반려동물과의 강한 유대

있을수록 더 심화 되는 경향이 있으며, 최근 연구에서는 약 49.8%의 반려동물 보호자가 반려동물을 잃은 후 펫로스 증후군을 겪는 것으로 나타났다[1]. 이는 일상 생활과 정신 건강에 심각한 영향을 미칠 수 있다. 이러한 증후군은 인간 상실에서 경험하는 고통과 유사하게, 깊은 애도와 그리움을 동반하는데, 반려동물을 단순한 애완동물이 아닌 '가족 구성

\* 국립창원대학교 문화테크노학과, 20180251@gs.cwnu.ac.kr, 20170242@gs.cwnu.ac.kr, 20234618@gs.cwnu.ac.kr,

\*\* 국립창원대학교 문화융합기술협동과정 석사 과정, 20247306@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수, sjyu@ghangwon.ac.kr (교신저자)

※ 이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.

원'으로 여기는 현대적 관점의 증가와 관련이 깊다. 특히 반려견을 상실한 반려인의 85.7% 한 가지 이상 고통을, 6개월 이후 35.1%, 1년이 지난 후 22.4% 반려인이 복합 비애의 고통을 경험한다[2].

## II. 본 론

본 연구에서는 Unity 게임 개발 엔진을 사용하여 불교적 사후 세계를 기반으로 한 반려견의 여정을 구현하였으며, 이를 통해 펫로스 증후군을 겪는 플레이어가 상실을 극복하고 삶의 의미를 재발견할 수 있도록 설계한다. 게임의 주요 흐름은 반려견이 죽음을 맞이한 후 주인을 찾기 위해 사후 세계인 극락에서 여정을 떠나는 내용을 중심으로 한다.

게임은 반려견이 죽음 후 극락에서 주인을 찾기 위한 여정을 통해 상실과 추억을 경험하는 내용으로 구성된다. 각 스테이지에서 사라지는 발판을 지나며 스테이지를 통과한다. 주인과의 추억을 떠올리게 하는 아이템을 수집하며 관계의 깊이를 체험한다. 스테이지를 완료할 때마다 추억의 사진이 나타나며, 마지막 스테이지에서는 반려견과 주인이 재회하고 함께 환생함으로써 불교적 윤회의 의미를 담아 새로운 삶의 시작을 전달한다. 플레이어는 이를 통해 상실을 수용하고 현재의 소중함을 되새기며, 미래에 대한 희망을 얻게 된다.

다음은 게임에 포함된 알고리즘을 설명한다. 1) 인벤토리 시스템 : 반려견이 수집한 기억 아이템을 기록하여 모든 아이템을 수집한 후에 다음 스테이지로 진행할 수 있도록 구성한다. 아이템은 한번에 하나만 획득할 수 있다. 아이템을 획득하면 우상단 UI 이미지가 획득한 아이템 이미지로 바뀐다. 아이템 이미지가 UI에 이미지로 있는 상태에서는 다른 아이템을 획득할 수 없다. UI 아이템 이미지와 같은 아이템을 한번 더 획득하면 기본 UI 이미지로 바뀐다. 기본 UI 이미지에서 다른 아이템을 획득할 수 있다. 2) 플랫폼 관리 알고리즘 : 반려견이 발판을 밟을 때마다 해당 발판을 제거하여 사라지도록 한다. 잘못된 길로 게임을 진행하면 스테이지를 클리어할 수 없게 만든다. 3) 강아지와 주인의 추억 사진 표시 알고리즘 : 각 스테이지가 완료될 때마다 강아지와 주인이 이승에서 함께 했던 추억이 사진

으로 표시되도록 하여 반려견과 주인의 추억을 하나씩 상기시킨다. 4) 스테이지 완료 조건 : 스테이지 모든 발판과 아이템이 제거되면 다음 스테이지로 가도록 설정한다.

본 게임은 플레이어가 반려견과의 소중한 추억을 회상하며 상실과 재생의 과정을 체험할 수 있도록 구성된다. 이를 통해 플레이어는 불교의 윤회 개념을 바탕으로 죽음을 또 다른 시작으로 이해하게 되며, 현재의 삶을 더욱 소중히 여길 수 있도록 유도하는 것이 본 연구의 목표이다.



그림 602. 게임 1 스테이지 플레이 화면  
Fig. 1. Game 1 stage play screen

## III. 결 론

펫로스 증후군은 반려동물을 잃은 후 겪는 깊은 심리적 고통으로, 이에 대한 치유 과정이 필요하다. 본 연구에서 제안한 게임은 불교적 사후 세계를 통해 반려견과의 추억을 되새기고 죽음을 수용하는 과정을 담아, 정서적 회복을 돕고자 한다. 플레이어는 게임 내 아바타와 동일시하며 자신의 감정을 탐구하고, 일상의 스트레스와 불안을 해소할 수 있다. 이를 통해 게임이 펫로스 증후군을 겪는 이들에게 심리적 치유의 방법이 될 가능성을 제시한다.

## 참 고 문 헌

- [1] [https://www.hani.co.kr/arti/animalpeople/companion\\_animal/1027755.html](https://www.hani.co.kr/arti/animalpeople/companion_animal/1027755.html) [Accessed 2024.10.28.]
- [2] J.H. Shin, "A Narrative Inquiry on the Pet Persons' LossExperience and the Nuclear Feelings" Pyeongtaek University, 2023

# 아두이노와 뷰포리아를 활용한 경남 특산물 홍보 방안

임성민\*, 정경택\*, 우동현\*\*, 유선진\*\*\*

## Gyeongnam Local Specialties Rebranding Plan using Arduino and Vuforia

Sungmin Im\*, Gyeongtaek Jeong\*, Donghyun Woo\*\*, and Sunjin Yu\*\*\*

### 요 약

최근 지역 브랜드와 정체성을 강화하는 핵심 요소로 지역 특산물이 부각되고 있다. 본 연구에서는 경상남도 특산물을 계절에 맞게 브랜딩하여 증강현실과 아두이노 RFID 기술을 접목한 새로운 체험형 마케팅 방안을 제안한다. 그 결과 잠재적 소비자는 기존 일방향적인 홍보 방안이 아닌 상호작용이 가능한 콘텐츠를 경험하여 특산물에 대한 정보를 인지하고 소비할 수 있다.

### Abstract

Recently, local specialties have emerged as a key factor in strengthening local brands and identity. In this study, we propose a new experiential marketing plan that combines augmented reality and Arduino RFID technology by branding Gyeongsangnam-do specialties according to the season. As a result, potential consumers can recognize and consume information on special products by experiencing interactive content rather than the existing one-way promotion plan.

### Key words

arduino rfid, vuforia, unity, android, local specialties

### 1. 서 론

'특허청'의 지자체 연도별 농산물 상표 출원현황에 따르면 전체 12,117건에서 2007년부터 2016년까지 10년 동안 4,561건이 출원되어 전체 출원의 30.3%를 차지하고 있다[1]. 이처럼 지역 특산물은 지역

의 문화와 경제에 큰 영향을 미치는 중요한 자원이 다. 또한, 지역 특산물과 경제의 합성어인 로코노미(Loconomy)가 가치소비와 결합하여 주목받고 있다 [2]. 본 연구도 경상남도 지역 특산물을 새롭게 브랜딩하고, 다양한 인터랙티브 기술을 통해 쌍방향적 소통을 할 수 있는 홍보 방안을 제안한다. 아두이노

\* 국립창원대학교 문화테크노학과, 20190260@gs.cwnu.ac.kr, 20190261@gs.cwnu.ac.kr

\*\* 국립창원대학교 첨단방위공학과정 박사과정 dhwoo1119@gs.cwnu.ac.kr

\*\*\* 국립창원대학교 문화테크노학과 교수 sjyu@changwon.ac.kr(교신저자)

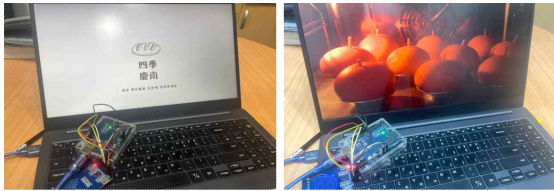
※ “이 연구는 「과학기술정보통신부 방통통신발전기금」의 재원으로 과학기술정보통신부, 정보통신산업진흥원, 경상남도, 경남메타버스지원센터의 지원을 받아 수행된 연구입니다.”

(Arduino)는 센서를 이용해 관람객과 상호작용하며 체험적 요소를 제공한다. 또한 증강현실(Augmented Reality, AR)의 특성을 활용하여 브랜드 메시지를 시각적으로 강화하고 몰입감을 제공한다. 이러한 기술을 통해 경남 특산물을 효과적으로 홍보할 콘텐츠를 제안한다[3].

## II. 아두이노 RFID 활용 관람객 체험 콘텐츠

경상남도 특산물과 펀슈머(Funsumer) 마케팅을 결합한 새로운 홍보 콘텐츠를 제작한다. 각 계절마다 수확하는 특산물을 나눠 새롭게 브랜딩하고, 펀슈머를 결합해 관람객의 흥미를 유발하도록 만든다.

본 콘텐츠는 아두이노 RFID와 윈도우 폼(Window Form)을 사용하여 제작하였다. 아두이노와 윈도우 폼은 시리얼 통신을 통해 연결한다. 관람객이 지정한 오브젝트를 아두이노 RFID 리더기에 놓으면, 오브젝트에 삽입된 태그가 인식되어 경상남도 특산물 브랜드 체험 영상이 재생된다.



RFID 인식 전                      RFID 인식 후

그림 1. 아두이노 RFID를 인식해 영상 재생  
Fig. 1. Recognize Arduino RFID and play the video

그림 1은 아두이노 RFID 태그가 인식되었을 때 영상이 재생되는 그림이다. 태그가 인식되기 전 기본 화면으로 대기 영상이 출력되며 태그를 인식하면 특산물에 대한 브랜드 체험 영상이 재생된다.

## III. 경남 특산물 AR 콘텐츠 제작

경상남도 지역마다 대표 특산물을 소개하는 AR 포스터를 제작한다. 관람객은 증강되는 특산물 정보에 대해 습득하고 생동감 있는 콘텐츠를 경험한다.

본 콘텐츠는 유니티(Unity) 엔진과 뷰포리아(Vuforia)를 사용하여 제작하였다. 안드로이드(Android) 및

태블릿 PC를 사용하여 콘텐츠 경험이 이루어진다. 경남의 주요 특산물 생산지 6곳을 지도에 그림으로 표시하여 뷰포리아를 통해 이미지 타겟으로 등록, 증강될 아이콘을 제작하여 연결한다. 아이콘을 터치하면 경남 특산물에 대한 정보가 있는 팝업창으로 연결된다.

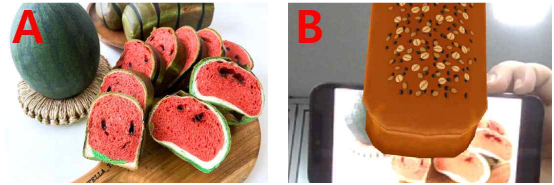


그림 2. 뷰포리아를 통해 증강된 개체  
Fig. 2. Objects Augmented by Vuforia

그림 2는 뷰포리아의 이미지 추적을 통해 설정된 이미지가 인식되었을 때 객체가 증강되는 그림이다. A는 이미지 추적 기능을 사용해 인식될 사진이며, B는 특정 그림이 인식되어 객체가 증강된다.

## IV. 결론

본 연구에서는 경상남도 특산물을 활용한 새로운 마케팅 방안에 아두이노 기술과 AR을 접목하여 인터랙티브 콘텐츠를 제작하였다. 관람객은 자신이 특산물에 대한 상호작용으로, 콘텐츠로의 몰입을 경험하며 잠재적 소비자와의 감성적 연결을 강화할 수 있다. 이를 통해 단순 특산물 홍보를 넘어, 지역 경제 활성화와 특산물의 가치를 재조명할 수 있다.

## 참 고 문 헌

[1] [https://d-research.or.kr/wi\\_files/paper\\_pdf/1243\\_paper.pdf](https://d-research.or.kr/wi_files/paper_pdf/1243_paper.pdf) / [Assessed, 2024-03-10]  
[2] [https://go.seoul.co.kr/news/newsView.php?id=20240809021006&wlog\\_tag3=naver](https://go.seoul.co.kr/news/newsView.php?id=20240809021006&wlog_tag3=naver) / [Assessed, 2024-09-21]  
[3] Yunji Moon, "The Effect of Quality Factors of Augmented Reality Technology on User Experience: Based on Experience Economy Theory", 2023

# 포트홀 검출을 위한 YOLO 모델의 성능 비교 평가

유동주\*, 이석호\*, 최보람\*, 김다정\*, 정동원\*\*

## Comparative Evaluation of Performance of YOLO Models for Pothole Detection

Dongju Yu\*, Seokho Lee\*, Boram Choi\*, Dajeong Kim\*, and Dongwon Jeong\*\*

### 요 약

이 논문에서는 대형 물류 차량으로 인해 발생하는 포트홀 문제를 해결하기 위해 YOLOv5부터 YOLOv11까지의 모델 성능을 비교하였다. AIHUB의 도로 장애물 데이터를 활용해 경량화된 YOLO 모델들을 학습시킨 결과, YOLOv9이 포트홀 검출에서 가장 높은 성능을 보였다. 실시간 처리가 가능한 경량 모델을 중심으로 한 성능 평가를 통해 실제 도로 환경에서의 적용 가능성을 검증하였다. 향후 연구에서는 데이터 전처리와 증강 기법을 통해 포트홀 검출 정확도를 더욱 개선할 계획이다.

### Abstract

This paper compares the performance of YOLO models from YOLOv5 to YOLOv11 to address the pothole issues caused by large logistics vehicles. The lightweight YOLO models were trained using road obstacle data from AIHUB, with results showing that YOLOv9 achieves the highest performance in pothole detection. Through the performance evaluation focusing on lightweight models capable of real-time processing, this paper verifies their applicability in actual road environments. In future research, the accuracy of pothole detection is expected to be further improved through data preprocessing and augmentation techniques.

### Key words

pothole detection, YOLO, deep learning, object detection, road maintenance, computer vision

### 1. 서 론

급속한 경제 성장과 물류 산업의 확대로 인해 대형 화물차량의 도로 이용이 증가하면서, 도로 기반

시설의 체계적인 유지보수와 안전 관리의 중요성이 높아지고 있다. 특히 대형 화물차량은 일반 차량에 비해 도로에 가하는 하중이 현저히 크기 때문에 도로 파손의 주요 원인으로 작용하며, 이는 빈번한 포

\* 군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과, {ydj5762, chincun77, 2201342, 2403123}@kunsan.ac.kr

\*\* 군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과 교수(교신저자), djeong@kunsan.ac.kr

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065)

트홀(Pothole) 발생으로 이어진다.

포트홀은 아스팔트 도로 표면에 발생하는 함몰 현상을 의미하며, 이는 단순한 도로 손상을 넘어 차량 파손, 교통사고, 물류 지연 등 다양한 사회적 문제를 야기할 수 있다[1]. 특히 대형 화물차량과 관련된 포트홀 사고의 경우 대형 사고로 이어질 위험성이 매우 높다.

포트홀은 조기 발견 및 신속한 보수가 이루어지지 않을 경우 규모가 점차 확대되어 더 큰 피해를 초래할 수 있다. 그러나 현재의 포트홀 관리 체계는 주로 육안 검사나 민원 접수에 의존하고 있어, 많은 시간과 비용이 소요될 뿐만 아니라 즉각적인 대응에도 한계가 있다. 이러한 문제점을 해결하기 위해서는 효율적이고 자동화된 포트홀 관리 시스템의 도입이 시급하다.

최근 컴퓨터 비전과 딥러닝 기술의 비약적인 발전으로 인해 포트홀 검출 모델의 실용화 가능성이 높아지고 있다. 본 연구에서는 대형 화물차량으로 인한 포트홀 발생에 효과적으로 대응하기 위해 다양한 YOLO[2] 시리즈 모델을 학습하고 비교 분석하여, 포트홀 관리 시스템에 최적화된 검출 모델을 탐색하고자 한다.

## II. 포트홀 데이터셋을 통한 검출기 학습

본 연구에서 사용한 데이터셋은 AIHUB에서 제공하는 ‘도로장애물/표면 인지 영상(수도권 외)’이다. 이 데이터셋은 도로 이미지에 나타나는 장애물(크랙, 포트홀, 맨홀, 보행자, 공사 표지판)에 대한 라벨링이 포함되어 있어, 포트홀 검출 모델 학습에 적합하다.

본 연구에서는 포트홀이 포함된 이미지만을 추출하여 포트홀 데이터셋을 구성하였으며, 표 1은 학습에 사용된 데이터셋 정보를 보여주며 그림 1은 데이터 예시를 보여준다.

표 1. 훈련 및 테스트 데이터셋  
Table 1. Train and Test Dataset

Data Type	Sample Size
Train	33,162
Test	8,290
Total	41,452



그림 1. 도로장애물/표면 인지 영상(수도권 외)  
Fig. 1. Obstacle/Surface Recognition Images (Outside the Metropolitan Area)

본 연구의 포트홀 검출 모델 학습을 위해 Ultralytics에서 제공하는 YOLOv5, YOLOv8, YOLOv9, YOLOv10, YOLOv11 모델을 활용하였다. Ultralytics 프레임워크는 하이퍼파라미터 자동 저장, 이미지 전처리, 데이터 증강 등의 기능을 제공하며, COCO 데이터셋[3]으로 사전 학습된 YOLO 계열 모델의 전이학습(Transfer Learning)을 통해 효율적으로 모델을 학습할 수 있다.

YOLO 계열 모델은 파라미터 규모에 따라 Nano(Tiny), Small, Medium, Large, Xlarge로 구분되는데, 본 연구에서는 실시간 검출이 가능한 경량화 모델인 Nano(Tiny) 모델을 중심으로 학습을 수행하고 검출 성능을 비교 분석하였다. 실험 환경 및 하이퍼파라미터에 대한 정보는 각각 표 2와 표 3에 기술한다.

표 2. 실험 환경

Table 2. Experiment Environment

Component	Specifications
CPU	Intel i9-13900K 5.5GHz
RAM	32GB
GPU	NVIDIA RTX A6000
OS	Ubuntu 22.04.4 LTS

표 3. 하이퍼파라미터

Table 3. Hyper Parameters

Parameter	Value
Batch Size	64
Epoch	100
Learning Rate	0.01

하이퍼파라미터와 데이터 증강(회전, 크롭, 확대/축소, 좌우 반전 등)은 Ultralytics 에서 제공하는 모델의 기본값을 사용하여 학습을 진행하였다.

표 4. 실험 결과  
Table 4. Experiment Results

Model	mAP50	mAP50-95
YOLOv5n	0.3833	0.17834
YOLOv8n	0.38906	0.18037
YOLOv9t	<b>0.39475</b>	<b>0.18481</b>
YOLOv10n	0.37198	0.17416
YOLOv11n	0.39061	0.18135

객체 검출의 성능 평가지표로는 mAP[4]가 사용된다. mAP(mean Average Precision)는 검출 모델이 객체를 올바르게 검출한 비율을 측정하는 지표로, 여러 클래스의 평균 정밀도(Precision)를 계산한 값이다. 구체적으로, AP(Average Precision)는 Precision-Recall 곡선의 아래 면적을 계산한 값으로, IoU(Intersection over Union) 임계 값에 따른 검출 정확도를 나타낸다. mAP는 이러한 AP 값들을 모든 클래스에 대해 평균한 값으로, 객체 검출 모델의 종합적인 성능을 평가하는 데 사용된다. 이 값은 0과 1 사이의 값을 가지며, 1에 가까울수록 모델의 성능

이 우수함을 의미한다. 본 연구에서는 mAP50의 기준으로 IoU가 0.5 이상인 경우를 정답으로 간주하여 포트홀 검출기의 정량 평가를 진행하였다.

그림 2는 각각의 모델의 학습 속도와 성능을 보여준다.

### III. 결론 및 향후계획

본 연구에서는 YOLOv5부터 YOLOv11까지의 다양한 YOLO 계열 모델들의 포트홀 검출 성능을 비교 분석하였다. 실험 결과, YOLOv9의 mAP50이 0.3945로 가장 우수한 성능을 보여주었으며, 이는 향후 포트홀 관리 시스템 구축에 있어 중요한 참고 자료가 될 것으로 기대된다.

다만, mAP 평가 지표에서 기존 연구[5]들과 비교하여 상대적으로 낮은 수치가 관찰되었다. 이는 데이터셋 분석 결과, 아스팔트의 긴 균열이 포트홀로 잘못 레이블링 된 사례들이 존재하는 것으로 확인되었다. 이러한 데이터의 품질 문제가 전반적인 성능 저하에 영향을 미친 것으로 판단된다.

향후 연구에서는 포트홀 검출의 정확도 향상을 위해 데이터 증강 및 전처리 기법에 대한 연구를 진행할 예정이다. 영상의 명암비 조정, 노이즈 제거

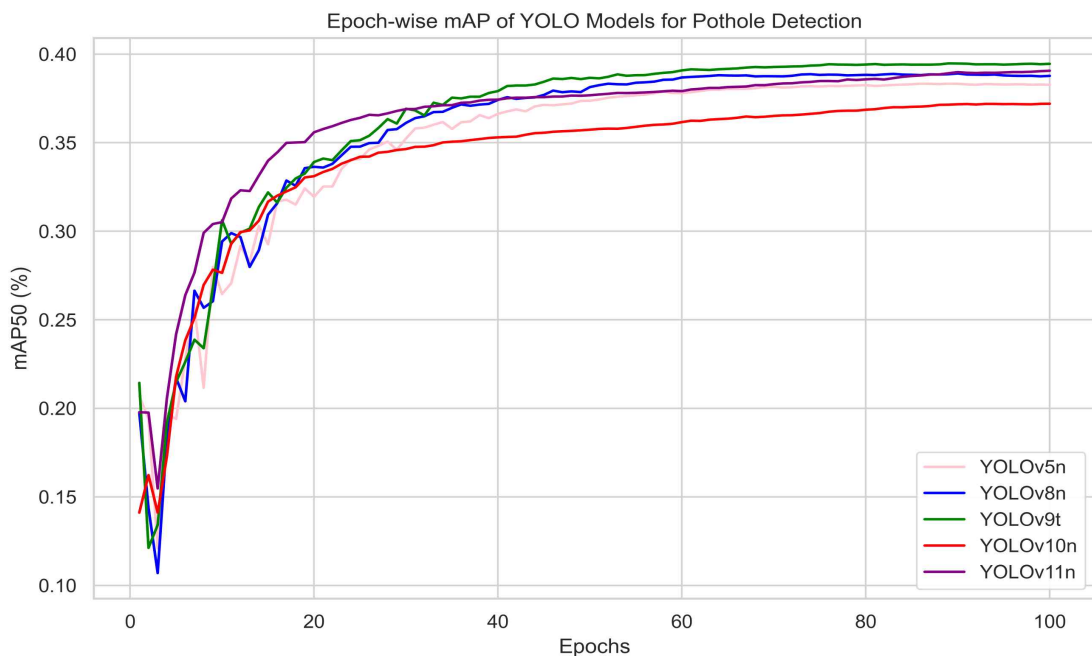


그림 2. mAP50 성능 그래프  
Fig. 2. mAP50 Performance Graph



등 기본적인 영상 처리 기법을 활용한 전처리와 함께, 변이형 오토인코더(VAE)[6], 적대적 생성 신경망(GAN)[7], 확산 모델(Diffusion)[8]과 같은 최신 생성형 AI 기술을 활용하여 다양한 환경의 포트홀 데이터를 증강할 계획이다. 이러한 데이터 품질 개선 접근법은 모델의 검출 정확도를 향상시키고, 실제 도로 환경에서의 포트홀 검출 성능을 개선하는 데 크게 기여할 것으로 기대된다.

본 연구의 결과는 향후 포트홀 관리 시스템 구축 시 최적의 검출 모델을 선정하는 데 중요한 지표로 활용될 수 있을 것이며, 이는 실질적으로 도로 안전성 확보와 유지보수 효율성 향상으로 이어질 것으로 기대된다. 또한 본 연구에서 도출된 각 YOLO 계열 모델들의 성능 비교 결과는 향후 유사한 도로 파손 검출 시스템 개발에서 유용한 참고 자료가 될 것이다.

## 참 고 문 헌

- [1] J. H. Im, S. C. In, T. W. Han, and B. G. Lee, "Suggestion of Future Pavement Maintenance by Pothole Analysis in Gyeonggi-do," International journal of highway engineering, vol. 21, no.1, pp. 51-59, 2019.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," Proc. CVPR, pp. 779-788, 2016.
- [3] T-Y. Lin, M. Maire, and S. Belongie, et al., "Microsoft COCO: Common Objects in Context," European Conference on Computer Vision (ECCV), 2014.
- [4] M. Everingham, L. V. Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes Challenge," International Journal of Computer Vision, vol. 88, no. 2, pp. 303-338, 2010.
- [5] J. H. Shin, "Development of Pothole Detection Model on Road Surface Using YOLOv4", 한국정보과학회 학술발표논문집, 제주, pp. 1820-1822, June 2021.
- [6] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," International Conference on Learning Representations (ICLR), 2014.
- [7] I. J. Goodfellow, J. Pouget-Abadie, and M. Mirza, et al., "Generative Adversarial Nets," Advances in Neural Information Processing Systems (NIPS), vol. 27, 2014.
- [8] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," Advances in Neural Information Processing Systems (NeurIPS), 2020.



# 인공지능을 활용한 날씨 기반 의류 추천 웹사이트 설계 및 구현

권혁찬\*, 최태용\*, 윤성건\*, 조대제\*\*

## A Weather based Clothing Recommendation Website using Artificial Intelligence Design and Implementation

Hyukchan Kwon\*, Taeyong Choi\*, Seong-Geon Yun\*, and DaeJe Jo\*\*

### 요약

본 논문은 사용자의 나이, 성별, 현재 날씨와 같은 다양한 정보를 바탕으로 날씨 기반 맞춤형 의상 추천 서비스를 제공하는 시스템을 제안한다. 특히, 심층 신경망(Deep Neural Network, DNN) 기반의 맞춤형 추천 시스템을 사용하여 의류 이미지 데이터와 날씨 데이터를 학습하며, 소프트 보팅(Soft Voting) 앙상블 기법을 통해 예측 성능을 향상한다. 학습 데이터는 AiHub 의류 통합 데이터를 사용하였고, 웹스크래핑으로 부족한 데이터를 보완한다. 구축된 모델은 상의, 하의, 아우터의 각 카테고리에서 평균 95% 이상의 정확도를 기록하였으며, OpenWeather API와의 연동으로 실시간 날씨 데이터를 반영하여 사용자의 현재 환경에 맞는 의류 조합을 추천한다. 이 시스템은 웹사이트 형태로 제공되어 사용자들이 손쉽게 접근할 수 있으며, 편안함과 스타일을 동시에 만족시키는 실용적 의상 추천 서비스를 제공한다.

### Abstract

This paper proposes a system that provides a weather-based personalised clothing recommendation service based on various information such as user's age, gender, and current weather. In particular, a personalised recommendation system based on Deep Neural Network (DNN) is used to learn clothing image data and weather data, and the soft voting ensemble technique is used to improve the prediction performance. The training data uses AiHub apparel aggregation data, and web scraping is used to supplement the lack of data. The built model has an average accuracy of over 95% in each category of top, bottom, and outerwear, and thanks to the integration with the OpenWeather API, it reflects real-time weather data to recommend clothing combinations for the user's current environment. The system is available as a website, making it easily accessible to users, and provides a practical outfit recommendation service that satisfies both comfort and style.

### Key words

Weather-based recommendation, Deep Neural Network (DNN), Soft voting ensemble, OpenWeather API, Machine learning

---

\* 국립안동대학교 SW융합학과, kwonhc226@gmail.com ,

\*\* 국립안동대학교 SW융합학과, djcho@anu.ac.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

## I. 서론

현대사회에서 패션은 몸을 보호하는 수단, 단순한 의복을 넘어 개인의 개성을 표현하고 사회적 인상을 형성하는 중요한 수단으로 자리 잡았다. 이에 따라 날씨, 시간대, 장소, 분위기 등 다양한 요인이 의상 선택에 영향을 미치며, 특히 날씨는 변화가 잦고 예측하기 어려운 요소로, 이에 적절히 대응하는 의상은 개인의 편안함과 컨디션을 유지하는 데 중요한 역할을 한다[1]. 예를 들어 여름철에는 통기성이 좋은 옷을, 겨울철에는 체온 유지를 위해 보온성이 뛰어난 옷을 착용함으로써 신체적 편안함을 유지할 수 있는 반면 날씨와 맞는 옷을 착용하지 않았을 경우에는 불편함을 초래할 뿐만 아니라 당일 컨디션을 좌지우지할 수 있다. 이러한 문제를 해결하기 위해 심층 신경망(Deep Neural Network, DNN)을 기반으로 의류 이미지 데이터와 나이, 성별, 날씨 등 다양한 요인들 간의 관계를 학습하여 날씨 기반 스타일 추천 서비스를 제안한다.

또한, 현대인들의 일상에서의 편리함을 높이고 효율적인 선택을 돕기 위해 웹사이트 형태로 사용자들에게 쉽게 접근 가능한 형태로 서비스를 제공하고자 한다.

## II. 관련 연구

### 2.1 DNN 기반 개인화 추천 시스템

DNN(Deep Neural Network)은 추천 시스템에서 사용자의 특성을 반영해 맞춤형 추천을 제공하는 고도의 딥러닝 아키텍처이다. 기존 추천 시스템이 사용자와 항목 간 상호작용을 단순히 예측하는 데 반해, DNN은 다층 퍼셉트론(MLP) 구조를 통해 다양한 입력 데이터를 심층적으로 학습할 수 있다. 특히, 사용자 특성과 기상 데이터를 결합하여 더욱 정밀한 맞춤형 추천을 수행하는 데 효과적이다. 예를 들어, 최근 연구에서 제안된 TD-DNN(Time Decay-DNN)은 시간 가중치를 적용하여 사용자 선호도 변화에 민감하게 반응할 수 있는 모델을 개발하였다. 이는 최근 행동에 더 높은 가중치를 부여해

추천 정확도를 향상시키는 방식으로, 시간에 따른 선호 변화까지 반영할 수 있는 장점을 가진다[2]. DNN 기반 추천 시스템은 또한 대규모 희소 데이터를 다루는 데 적합하며, 이러한 특성은 사용자별 맞춤형 추천에서 특히 유리하다. Facebook은 사용자와 항목 간의 희소 데이터를 밀집 데이터로 변환하는 방식으로 대규모 추천 시스템의 성능을 개선한 바 있다. 본 연구의 DNN 모델 또한 사용자 프로필과 날씨 데이터를 임베딩하여 사용자 맞춤형 스타일 추천에 적용하고 있다[3].

### 2.2 소프트 보팅을 활용한 앙상블 기법

소프트 보팅(Soft Voting)은 여러 모델의 예측 결과를 확률적으로 결합하여 최종 결론을 도출하는 앙상블 방법으로, 개별 모델의 단점을 상쇄하며 예측 성능을 향상 시킨다. 최근 다양한 연구에서 소프트 보팅 앙상블을 적용하여 성능을 높인 사례가 있으며, 특히 다중 변수를 포함하는 복잡한 예측 모델에서 효과적인 성능을 발휘한다. 예를 들어, 호흡기 질환의 중증도를 예측하는 연구에서는 소프트 보팅을 통해 개별 모델보다 높은 정확도를 달성하였고, 이는 의료와 같은 민감한 분야에서 보다 정확한 진단을 가능하게 하였다[4]. 또 다른 연구에서는 심장 질환의 예측에 소프트 보팅을 적용하여 기존 방법보다 높은 예측 정확도를 확보하였으며, 이는 예측 안정성과 신뢰성을 크게 개선한 사례로 주목받고 있다[5]. 본 프로젝트에서는 DNN과 기타 AI 모델을 소프트 보팅으로 통합하여, 기상 데이터와 사용자의 선호 데이터를 반영하는 방식으로 앙상블 모델을 구축하여 보다 정교한 추천을 제공하고자 한다.

## III. 설계 및 구현

### 3.1 데이터 수집

본 논문에서는 제안하는 날씨 기반 옷 스타일 추천 인공지능 학습을 위해 AiHub 의류 통합 데이터를 사용하고 부족한 데이터는 웹스크래핑을 통해 보완한다. 의류 데이터셋은 카테고리로 아우터, 상

의, 하의로 분류되어 있으며 10034개의 이미지와 메타데이터(Json)으로 구성된다. 메타데이터는 특징 학습에 필요한 계절, 옷사이즈 등의 정보가 포함된다.

수집한 의류 이미지 데이터는 상의 카테고리 스웨터, 셔츠, 블라우스, 베스트, 티셔츠, 하의 카테고리는 긴바지, 반바지, 숏스커트, 롱스커트, 아우터는 코트, 재킷, 가디건, 점퍼로 구성한다.

### 3.2 데이터 전처리

데이터 전처리 단계에서는 JSON 파일 형식으로 각 이미지에 대한 메타데이터를 수정하여 라벨링 작업을 진행한다. 기존 의류 데이터셋의 하의 데이터셋은 2가지(바지, 스커트)로 구성되어 있어 세분화 작업을 진행한다. 메타데이터의 바지, 스커트 길이를 파싱하여 길이값이 일반적인 기준으로 40 이상이면 긴바지, 롱스커트로 분류한다.

학습 데이터셋을 구축하기 위해 상의, 하의, 아우터 카테고리 별 메타데이터로부터 성별, 나이, 계절, 의류 종류를 파싱하여 Csv파일로 구성한다. 메타데이터로부터 정보를 파싱하는 코드는 다음 그림 1과 같다.

```

Algorithm 1: Extract Unique Clothes Information from Directory of JSON Files
Input: Directory path containing JSON files
Output: Unique clothes types and seasons
Function extract_unique_clothes_info_from_directory(directory):
1. Initialize unique_types and unique_seasons as empty sets;
2. foreach file_name in directory do
   if file_name ends with '.json' then
     3. json_file_path ← os.path.join(directory, file_name);
     4. clothes_type, clothes_season ← extract_clothes_info_from_json(json_file_path);
     5. Add clothes_type to unique_types if not empty;
     6. Add clothes_season to unique_seasons if not empty;
   end
end
7. return unique_types, unique_seasons;
;
Function extract_clothes_info_from_json(json_file):
1. Open the JSON file and parse the data;
2. weather ← data.get('metadata.clothes', {});
3. clothes_type ← weather.get('metadata.clothes.type');
4. clothes_season ← weather.get('metadata.clothes.season');
5. return clothes_type, clothes_season;
    
```

그림 1. JSON 파일로부터 의류 정보 추출 코드  
Fig. 1. Extract Clothes Information from JSON File Code

### 3.3 인공지능 학습

날씨에 기반한 옷 스타일을 추천하기 위해 구축한 학습 데이터로 인공지능 학습을 진행한다. 인공지능 입력 형식을 맞추기 위해 범주형 변수인 성별, 나이, 계절 정보를 One-Hot Encoding 방식으로 변환하고, 의류 종류는 LabelEncoder를 사용하여 인코딩한다.

인공지능 모델은 DNN(Deep Neural Network)를 기반으로 입력층을 포함하여 6개의 은닉층, 활성화 함수, 배치 정규화층, 드롭아웃으로 구성한다. 출력층은 Softmax 함수를 활용하여 다중 분류 작업에 맞게 설정한다. 다음 그림 2는 모델 구조를 시각화한 것이다.

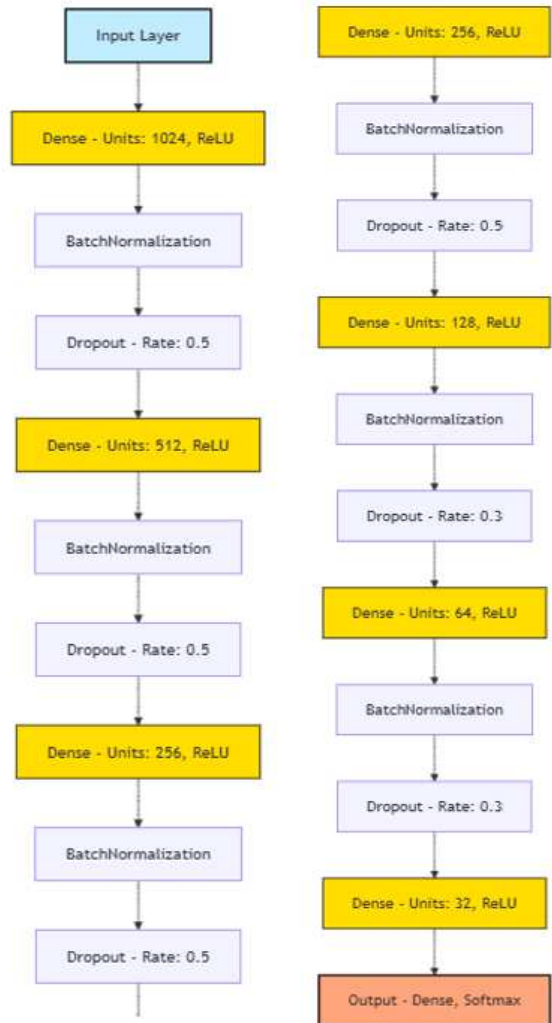


그림 2. 모델 구조 시각화  
Fig. 2. Model structure visualization

학습 설정으로 활성화 함수는 ReLU, Softmax 함수, 옵티마이저는 Adam, Loss는 다중 분류 손실 함수(Sparse Categorical CrossEntropy), Epoch=100, 학습률 0.001, 배치 사이즈는 16으로 설정하고 훈련 시 K-fold 교차 검증 방식으로 5개의 폴드로 설정한다.

최종적으로 교차 검증에서 학습한 5개의 모델을 기반으로 앙상블 예측을 통해 각 모델의 예측 결과를 평균하여 최종 추론을 진행한다. 다음 표 1은 모델 평가 결과이다.

표 1. 모델 평가 세트  
Table 1. Model evaluation sets

모델 종류	Loss	Accuracy
상의	1.31	0.94
하의	1.26	0.97
아우터	1.10	0.95

### 3.4 실행 결과

제안된 날씨 기반 의상 추천 시스템의 실행 결과 화면은 그림 3과 같다. 사용자로부터 입력받은 성별, 나이 등의 기본정보와 OpenWeather API를 통해 수집한 현재 날씨 데이터를 AI 서버로 전송하여 추천해준 의류 이미지를 각 의상 카테고리에 해당하는 웹 화면에 출력한다. 이와 같이 제안된 시스템은 사용자에게 현재 날씨에 맞는 실용적이면서도 개성 있는 의상 조합을 실시간으로 제공함으로써 사용자 경험을 개선할 수 있다.

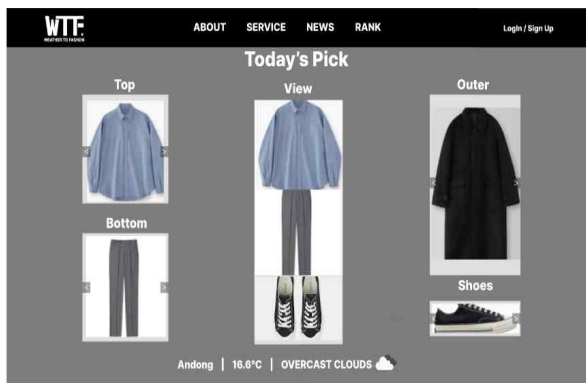


그림 3. 실행 결과  
Fig. 3. Execution Results

## IV. 결 론

본 논문에서는 DNN(Deep Neural Network) 기반 추천 시스템을 소프트 보팅 앙상블 기법을 활용하여 날씨 맞춤형 의류 추천 서비스를 제안하였다. 다양한 요인에 영향을 받는 의류 선택의 중요성을 고려하여, 날씨와 의류의 특징 요소들을 통해 인공지능을 학습하여 추천 알고리즘을 개발하였다. 제안한 시스템은 각 카테고리에서 비교적 높은 성능을 달성하였으며, 특히 날씨 데이터를 결합하여 의류 추천 기능을 통해 사용자 편의성을 향상하였다. 이를 통해 사용자가 날씨에 적합한 옷을 선택하는데 시간 효율성과 맞춤 기능을 제공할 수 있다.

향후에는 Fine-Tuning과 하이퍼-파라미터 조정을 활용한 추천 정확도 향상과 데이터 보강 및 데이터 다양화 방안을 통해 모델의 일반화와 향상된 사용자 경험을 제공할 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1] N. G. Jang, "패션디자인에서 기후변화 적응의 유형과 특징", 한국패션디자인학회 vol. 15, no. 3, pp. 35-5, September 2015.
- [2] Tripti Mahara, Subhash Chander Sharma, Saurabh Agarwal, Hyunsung Kim, "TD-DNN: A Time Decay-Based Deep Neural Network for Recommendation System," Applied Sciences, 2022.
- [3] "The Architectural Implications of Facebook's DNN-based Personalized Recommendation," arXiv preprint arXiv:1906.03109.
- [4] Ali, S., Hussain, A., Aich, S., et al., "A Soft Voting Ensemble-Based Model for the Early Prediction of Idiopathic Pulmonary Fibrosis (IPF) Disease Severity in Lungs Disease Patients," Life, 2021.
- [5] Tabia Tanzin Prama, et al., "An AI-Based Clinical Recommendation System Using Ensemble-Based Soft Voting Classifier," PLOS ONE, 2021.

# YOLO 모델을 이용한 시각 장애인용 식품 부패 탐지 시스템

황순규\*, 안상우\*, 강부진\*, 모아사랑\*, 정기현\*\*

## Food Spoilage Detection System for the Visually Impaired Using YOLOv9 Model

SoonGyu Hwang\*, SangWoo An\*, BuJin Kang\*, A Sa Rang Mo\*, and KiHyun Jung\*\*

### 요 약

본 연구에서는 시각장애인과 저시력자들이 식품의 부패 상태를 인지하지 못하여 겪는 어려움과 위험을 해결하고자 인공지능을 활용한 식품 부패 탐지 시스템을 제안한다. 제안된 시스템은 YOLO 모델을 기반으로 실시간으로 식품의 부패 상태를 분석하는 인공지능이다. 약 1,400장의 라벨링된 이미지 데이터셋을 사용했으며, 현재 모델의 정확도는 약 50%로 제한적이거나 이는 데이터양과 모델 구조의 한계로 분석된다. 향후 연구에서는 데이터셋을 확장하고 모델 성능을 개선하는 방향으로 나아갈 계획이다. 본 연구의 목표는 인공지능 기술을 통하여 시각장애인과 저시력자들이 식품 상태를 안전하게 파악할 수 있도록 정확도를 높이는 것으로, 이를 통해 부패 식품으로 인한 위험을 예방하고자 한다. 앞으로 더 다양한 식품 종류와 부패 상태에 대한 추가 학습을 통해 인공지능의 정확도와 성능을 향상시킬 계획이며, 이를 통해 사회 전반에 걸쳐 보다 안전한 식품 관리가 가능해질 것으로 기대된다.

### Abstract

This study proposes a food spoilage detection system using artificial intelligence to address the challenges and risks faced by visually impaired and low-vision individuals who cannot perceive the spoilage status of food. The proposed system is an AI-based model that analyzes food spoilage in real time, built upon the YOLO framework. Approximately 1,400 labeled images were used in the dataset, and the current model achieves around 50% accuracy, which is limited due to the constraints of data volume and model architecture. Future research will focus on expanding the dataset and enhancing model performance. The goal of this study is to help visually impaired and low-vision individuals safely assess the condition of food through AI technology, thereby preventing risks associated with spoiled food. By further training the AI on a wider range of food types and spoilage conditions, we aim to improve accuracy and performance, enabling safer food management across society.

### Key words

Artificial Intelligence, YOLO, Detection System

---

\* 국립안동대학교 멀티미디어공학과, hskng@naver.com

\*\* 국립안동대학교 소프트웨어융합학과, kingjung@anu.ac.kr (교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

## I. 서 론

오늘날 기술은 매우 빠르게 발전하고 있지만, 시각장애인과 저시력자들이 일상생활에서 겪는 불편함은 여전히 해결되지 않고 있다. 특히 식사 과정에서 시각장애인들이 겪는 어려움은 매우 중요한 문제로 다루어져야 한다[1]. 시각장애인들은 음식에 대한 시각적 정보를 얻기 어려우므로, 식품을 섭취하기 전 그 상태를 정확히 알 수 없다. 외부의 도움이 없을 때, 부패한 식품을 섭취하게 될 위험이 커질 수 있으며, 이는 심각한 건강 문제로 이어질 수 있다. 이런 상황에서 시각장애인들은 식품의 신선도나 부패 상태를 스스로 판단하기 어렵기 때문에, 식사의 안전성은 크게 저하될 수밖에 없다. 이에 따라 시각장애인들의 삶의 질 또한 크게 영향을 받을 수 있다. 식품의 상태를 적시에 정확히 파악하는 것은 누구에게나 중요한 일이지만, 특히 시각적 정보가 제한된 시각장애인들에게는 더욱 절실한 문제이다[2]. 이를 해결하기 위하여 본 연구에서는 인공지능(AI) 기술을 기반으로 식품 부패 상태 탐지 시스템을 제안하고자 한다. 현재 시중에는 시각장애인을 위한 여러 보조 기술들이 존재하지만, 기술의 대부분은 물체의 기본적인 정보를 인식하는 데 그치고, 식품의 부패 여부와 같은 세부적인 정보를 제공하지 못한다[3]. 이는 시각장애인들이 여전히 부패한 식품을 섭취하게 될 위험을 방지하지 못하는 큰 한계점으로 작용한다. 본 연구는 이러한 한계를 극복하기 위해, YOLO(You Only Look Once)라는 인공지능 모델을 활용하여 식품의 상태를 실시간으로 분석하고, 부패 여부를 판단하는 시스템을 제안한다. YOLO 모델은 빠르고 정확하게 객체를 인식할 수 있는 기술로, 입력된 이미지를 분석해 식품이 신선한지 부패했는지를 즉시 판단한다[4]. 이 시스템은 사용자가 식품을 섭취하기 전, 부패 여부를 미리 확인할 수 있게 도와줌으로써 시각장애인의 건강을 보호하는 데 중요한 역할을 할 수 있다. 또한 이 시스템은 시각장애인들뿐만 아니라, 공공기관, 학교, 복지시설 등 다양한 환경에서 식품 안전을 관리하는 데 활용될 가능성이 크다. 특히 대량으로 식품을 관리해야 하는 곳에서는 부패한 식품을 사전에 감지하여 식중독과 같은 사고를 예방하는 데 기여할

수 있을 것으로 기대된다. 이는 단순히 시각장애인 개인의 안전을 보장하는 것뿐만 아니라, 공공의 건강을 지키는 데에도 중요한 역할을 할 수 있다. 향후 연구에서는 다양한 식품의 종류와 부패 상태를 학습시켜, 인공지능의 정확도와 성능을 지속해서 개선할 계획이다. 이를 통해 더욱 정밀한 탐지와 분석이 가능해질 것이며, 시각장애인들이 더욱 안전하고 편리한 식생활을 할 수 있을 것으로 기대된다.

## II. 관련 연구

### 2.1 시각장애인을 위한 AI 기반 보조 기술

현재 시각장애인을 위한 AI 기반 보조 기술에는 다양한 연구들이 이루어지고 있다. 본 연구에서 다루지는 연구와 관련된 보조 기술과 관련되는 기술은 Alris가 있다. Alris는 AI 기반 웨어러블 기기로서 안경에 장착된 정교한 카메라와 자연어 처리 인터페이스를 결합해 실시간 청각적 설명을 받을 수 있도록 하는 기능이 있다. 이 기기는 식품 감지 이외에도 노트 작성, 돈 세기, 텍스트 읽기 및 바코드 스캐닝 등의 다양한 기능을 포함하고 있으나 아직 프로토타입 단계의 기술이다[5].

### 2.2 YOLOv9

본 연구에서는 이미지 내에서 부패한 부분을 탐지하기 위해 활용한 YOLOv9은 인스턴스 분할 분야에서 다른 버전과 비교해 높은 성능을 보여주는 YOLO(You Only Look Once) 시리즈의 최신 버전이다[6]. YOLOv9은 PGI(Programmable Gradient Information)와 GELAN(Generalized Efficient Layer Aggregation Network)을 도입하여 YOLOv8 대비 성능과 효율성을 크게 개선한 객체 탐지 모델이다.

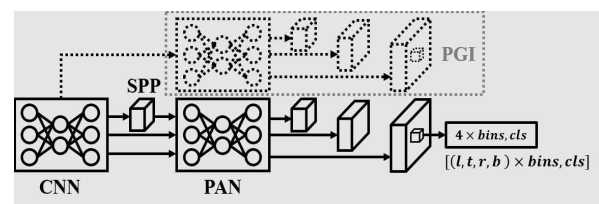


그림 1. YOLOv9의 아키텍처

Fig. 1. YOLOv9 architecture



[그림 1]은 PGI를 담은 YOLOv9의 아키텍처를 담은 그림이다. 이 아키텍처는 모델의 해석성, 견고성 및 다양성을 향상할 수 있다. PGI의 설계는 가역적 아키텍처와 다중 레벨 정보의 개념을 사용하여 모델이 유지할 수 있는 원래 데이터와 대상 작업을 완료하는데 필요한 정보를 극대화하는 것이고, GELAN은 PGI가 매개변수 수가 적은 모델에서 어떻게 뛰어난 정확도, 안정성 및 추론 속도를 달성할 수 있는지 보여주는 것이다[7]. 본 연구에 사용되는 YOLOv9 모델은 더 나은 평균 정밀도(mAP)와 빠른 추론 시간을 제공하며, 실시간 객체 탐지 및 다양한 환경에서 뛰어난 성능을 보여준다. YOLOv9은 이전 YOLO 모델들과 비교했을 때, 파라미터 수와 연산 비용에서 더 나은 성능으로 개선한 버전으로 알려져 있다[8]. 이 모델은 파이썬 패키지에서 쉽게 적용할 수 있으며, YOLOv9은 높은 처리 속도와 효율성을 자랑한다. 이는 객체 분할에 있어서 실시간으로 수행하는 데 효과적이고, 식품의 부패한 부분을 감지하고 부패 상태를 다양한 조건에서 정확하게 식별할 수 있도록 한다.

### 2.3 객체 분할(Instance Segmentation)

본 연구에서 사용되는 객체 분할(IS)은 객체 탐지와 객체 분할의 결합한 형태로, 이미지 내에서 각각의 객체를 픽셀 단위로 구분하는 작업을 말한다. 객체 탐지는 이미지에서 객체의 경계를 바운딩 박스(Bounding Box)로 표시하지만, 객체 분할(IS)은 객체의 윤곽선을 정확하게 추출하여 객체의 형태와 위치를 더욱 정밀하게 표현하는 것이 특징이다. 이 기술은 자율 주행 자동차, 의료 영상 분석, 영상 편집 등 다양한 분야에서 필수적인 역할을 한다[9]. 본 연구에서는 컴퓨터 비전 작업을 위한 플랫폼인 roboflow(<https://app.roboflow.com/>)에서 객체 분할(IS)을 위한 라벨링 작업(Labeling) 및 데이터셋을 생성했다.

[그림 2]의 라벨링 작업과 같이 경계를 정확하게 정의하는 것은 개별 객체를 구분하여 정확한 픽셀 단위로 분할하여 정확한 데이터를 제공하고 다양한 객체를 구분하며 객체의 특성을 반영하는 데에 큰 도움이 된다. 이처럼 라벨링된 데이터를 통해 모델

이 각 객체를 정확히 학습하고 예측할 수 있는 기반이 마련이 될 수 있다[10].

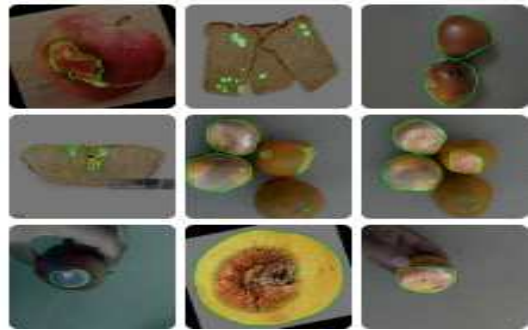


그림 2. Roboflow 라벨링  
Fig. 2. Labeling in roboflow

## III. 시스템 설계 및 구현

### 3.1 아키텍처 설계

본 연구에서는 인공지능 구축 설계 및 구현 방법으로 파이프 필터 패턴을 사용하였다. 파이프 필터 패턴은 소프트웨어 아키텍처 패턴 중 하나로 데이터 처리를 위한 구조를 나타낸다. 처리 단계는 데이터 수집, 데이터 전처리, 인공지능 학습 및 검증, 결과 확인 순서로 설계하였다. 이 패턴에 대한 다이어그램은 [그림 3]과 같다.

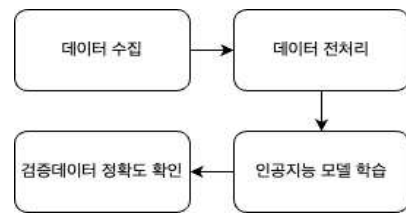


그림 3. 데이터 처리 다이어그램  
Fig. 3. Data processing diagram

[그림 4]는 제안한 시스템의 아키텍처를 입력, 출력 단계로 나뉘어서 사용자의 데이터를 처리하도록 설계한 모습이다.

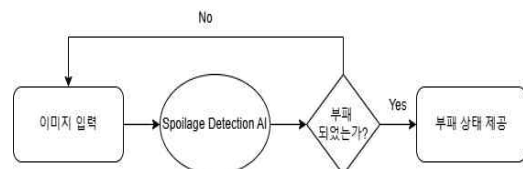


그림 4. 시스템 아키텍처 다이어그램  
Fig. 4. System architecture diagram

인공지능은 식품의 부패 여부를 판단하는 부패 탐지 모델로 설계하여, 사용자가 사진을 전송하면 학습된 인공지능 모델의 판단에 따라 결과 데이터를 사용자에게 실시간 탐지 결과를 전송한다.

### 3.2 데이터 수집

본 연구에서 사용되는 데이터는 컴퓨터 비전 작업을 위한 플랫폼인 로보플로우의 공개데이터셋과 인터넷에서 크롤링 기술을 활용하여 수집하였다.

데이터셋의 구성은 라벨링 된 이미지와 텍스트 데이터가 각각 1,381개로 이루어져 있다. 각 식품 이미지 파일에는 해당 식품 이미지에서 부패에 해당하는 부분에 대한 라벨을 텍스트 파일 형태로 제공한다. 이러한 라벨은 각 부패 유형에 대해 객체 감지를 위한 경계 상자 좌표와 클래스 정보를 포함한다. 이러한 구성은 모델이 학습 과정에서 각 식품 이미지의 상태를 부패 부분 라벨과 연관 짓게 하여, 추후 새로운 식품 이미지 데이터에 대해 부패한 부분을 정확히 예측할 수 있도록 한다. 라벨링 작업은 식품 부패의 유형을 '상함', '곰팡이', '마름' 등 15개의 클래스로 구분하여 구성되어 있으며, 특히 위 세 개의 라벨의 경우 식품별 부패 형태가 다르므로 각각의 이미지에 구분된 라벨링을 진행하였다. 이에 따라 인공지능이 조금 더 정교한 분석이 가능하다고 판단한다. 해당 데이터셋을 활용하여 식품 이미지에서 부패한 부분을 인식하고 분류하는 인공지능을 구현하였다.

image_id	category	lxmin	lymin	xmax	ymax	segmentation
0	9	246	275	392.25	467.5	[375, 409.375, 388.75, 383.125]
0	9	254	244	297.75	279.625	[[296.25, 261.875, 288.75, 255.4]
0	9	269	218	286.5	246.125	[[273.75, 218.125, 268.75, 241.4]
0	9	286	194	294.75	209.625	[[292.5, 209.375, 295, 203.125,]
0	9	304	176	325.25	201.625	[[308.125, 200, 318.75, 201.875]
0	9	368	124	393	157.125	[[376.875, 123.75, 367.5, 133.1]
0	9	280	155	299.375	183.125	[[281.25, 159.375, 280, 164.375]
0	9	245	214	260	234	[[245, 219.375, 245, 229.375, 2]
0	9	301	221	313.5	240.375	[[306.25, 240.625, 313.75, 231.4]
0	9	355	186	377.5	222.875	[[371.875, 186.25, 358.125, 192]
0	9	240	284	258.75	310.875	[[243.75, 286.875, 240, 299.375]
0	9	378	451	419.25	487.25	[[395.625, 451.25, 390.625, 458]
1	6	94	104	510.695	518.375	[[508.262, 343.125, 510.916, 26]
1	4	94	109	344.813	517.125	[[334.418, 424.375, 345.034, 41]
2	1	163	233	246.333	285.5	[[246.667, 245, 226.667, 232.5]

그림 5. 부패 식품 데이터셋 CSV 파일  
Fig. 5. Spoilage food datasets csv file

[그림 5]는 라벨링된 데이터셋을 CSV 파일로 정형화한 사진이다. 이 그림에는 각 이미지와 객체의

클래스에 해당하는 ID와 객체의 경계 좌표를 포함하고 있다. 또한 분할(Segmentation) 부문에서 객체 분할(IS) 작업을 위한 각 객체에 대한 마스크 정보가 담겨져 있다.

### 3.3 데이터 전처리

본 연구에서는 객체 분할(IS) 작업을 위한 전처리 과정으로, 학습 데이터와 검증 데이터를 분리하는 절차를 거친다. 이는 모델이 다양한 데이터에 노출되게 하고 일반화된 성능을 보장하며, 과적합을 방지하기 위한 중요한 단계이다. 본 연구에서의 데이터의 전처리는 객체 분할(IS) 학습을 위해 본 연구에서는 식품의 부패 상태를 정확히 감지하기 위해, 이미지 데이터와 해당 라벨로 구성된 데이터셋을 활용하였다. 우선, 수집된 이미지와 라벨 데이터의 수가 일치하는지 검증하고, Scikit-learn 라이브러리의 내장함수인 train\_test\_split() 함수를 이용해서 데이터셋을 훈련 데이터(80%)와 검증 데이터(20%)로 나누었다. 이 과정은 [그림 6]에 나타나 있는 것과 같이 모델이 학습 중 다양한 데이터 패턴을 학습하게 되어 보다 일반화된 예측 능력을 갖출 수 있게 해준다.

```
# 학습, 검증용 데이터 분리
from sklearn.model_selection import train_test_split

train_img_list, val_img_list = train_test_split(
    sda_img_list,
    test_size=0.2,
    random_state=19)

print(f"#학습 이미지: {len(train_img_list)}, 검증 이미지: {len(val_img_list)}")

# 학습 데이터 경로가 작성된 텍스트 파일 생성
train_txt = r"C:\Users\hskng\Desktop\Cell_project\FORVISION\Spoilage_Detection
val_txt = r"C:\Users\hskng\Desktop\Cell_project\FORVISION\Spoilage_Detection

# 텍스트 파일 생성 코드
with open(train_txt, "w") as f:
    f.write("\n".join(train_img_list))

with open(val_txt, "w") as f:
    f.write("\n".join(val_img_list))

import yaml
```

그림 6. 부패 식품 데이터 처리 코드  
Fig. 6. Spoilage Food data processing code

본 연구에서는 식품 부패 유형을 15개의 클래스로 정의했다. 각 클래스는 '상함', '곰팡이', '마름', '벌레', '변형' 등 다양한 부패 상태를 포함하고 있으며, 라벨링 작업은 YOLO 형식에 맞춰 각 객체의 경계 상자 좌표와 클래스 인덱스를 포함하도록 수

행했다. 이 정보는 모델 학습 시 필요한 YAML 파일로 저장하였으며, YAML은 가독성이 높고, YOLO 모델을 사용하여 인공지능 모델을 만들 때 자주 사용되는 형식으로, 구조화된 데이터를 효율적으로 저장할 수 있는 장점이 있으며, 모델이 학습하는 동안 각 클래스에 맞는 라벨 정보를 참조할 수 있도록 파일을 구성했다[11].

### 3.4 부패 탐지 구현 및 학습

[그림 6]은 부패 탐지 AI의 파인 튜닝(fine tuning) 코드이다. 본 연구에서 사용한 YOLOv9c-seg 모델은 객체 분할 작업에 최적화된 최신 딥러닝 아키텍처로, 본 연구에서는 부패 상태를 인식하고 해당 객체의 위치를 감지하는 데 사용되었다.

```
# 모델 파일 경로
yolov9 = r"C:\Users\hskng\Desktop\Cell_project\FORVISION\yolov9c-seg.pt"
# yaml 파일 데이터 경로
yaml_path = r"C:\Users\hskng\Desktop\Cell_project\FORVISION\Spoilage_Detection\yolov9c-seg.yaml"

# 데이터셋 경로 확인
train_path = r"Spoilage_Detection_AI\SDA_DATA\train"
val_path = r"Spoilage_Detection_AI\SDA_DATA\valid"

command = [
    "yolo",
    "task=segment", # segmentation 작업
    "mode=train", # 학습 모드 지정
    f"model={yolov9}", # 사전 학습된 yolo 모델 사용
    f"data={yaml_path}", # 전처리된 yaml 파일 데이터 사용
    "batch=16", # 배치 크기를 설정
    "epochs=1000", # 에포크 설정
    "imgsz=640", # 이미지 사이즈 조절(640*640이 YOLOv9 수용할 수
    "conf=0.8", # 임계값 설정
    "name=sda_predict", # 학습 결과를 저장할 폴더 이름 설정
    "exist_ok=True", # 동일한 이름 디렉토리가 존재할 경우 덮어쓰기
    "device=0", # GPU장치 n번 사용하여 학습
    "iou=0.8" # 정밀도 설정
]
```

그림 6. 부패 탐지 AI 파인튜닝 코드  
Fig. 6. Spoilage detection AI fine-tuning code

YOLOv9c 모델은 YOLOv7 AF보다 매개변수가 42% 작고 연산 요구량이 21% 적지만 비슷한 정확도를 달성하여 YOLOv9의 상당한 효율성 개선을 보여준다[12]. 이러한 이유로 YOLOv9c-seg 모델을 채택하여 식품의 부패 상태를 감지하는 인공지능 모델을 학습시켰다. 배치 크기는 학습 환경에 맞춰 16으로 지정하였고 에포크는 1,000으로 과적합 상황이 나타나면 자동으로 학습이 종료되도록 설정했다. 그리고 식품 부패도와 관련된 일이므로 정확도가 중요하다고 생각해 임계값과 정밀도를 80%로 책정했으며 이후 학습 데이터를 점차 늘릴 예정이다.

## IV. 성능 평가 및 결과

본 연구의 시스템 모델 학습 결과의 mAP (Mean Average Precision) 성능 지표를 [그림 7]에서 보여주고 있다. B는 바운딩 박스(Bounding Box), M(세분화 작업) 분할을 의미하는데, 세분화 작업에는 다소 낮은 정밀도(50%)와 정확도(60%)로 아쉬운 결과를 보였다.

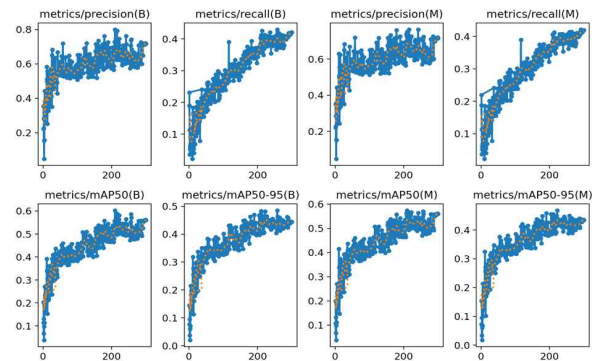


그림 7. mAP 성능 지표 그래프  
Fig. 7. mAP performance metrics graphs

[표 1]은 YOLOv9에서 instance segmentation작업을 하는 기존 모델인 경량형 모델 c모델의 epochs 300 과 1000, 그리고 거대한 모델 e모델의 epochs 300의 수행 능력을 비교한 표다. 각각의 정밀도와 재현을 종합 성능을 비교하면 e모델과 c모델의 수치가 큰 차이가 나지 않는 걸로 보여주고, 이로써 수행속도가 더 빠른 c모델을 채택하였다.

Model	Epoch	Precision	Recall	mAP50@
YOLOv9c-seg	300	0.6212	0.331	0.4759
YOLOv9c-seg	1000	0.6198	0.3365	0.4756
YOLOv9e-seg	300	0.6217	0.3294	0.4749

차트 1. YOLOv9 모델 비교  
Chart. 1. YOLOv9 model comparison

## V. 결 론

본 연구에서는 부패한 식품에 관련해 어려움에 처해있는 시각장애들을 위해 안전하고 더욱더 나은 식생활을 제공하고자, AI 기반 보조 기술을 제시하였고, 식품 부패도 탐지 인공지능 모델을 설계하고 구현하였다. 기존에 존재하는 YOLOv9 모델을 활용

하여 부패 탐지 AI 모델을 구현하였으며, 사용자는 식품의 이미지를 입력하면 부패 상태를 분석하여 준다. 이후 더 나은 구현을 위해 스마트폰에서 사용할 수 있는 어플리케이션이나 웹앱과 접목하여 시각장애인들이 어려움없이 이 시스템을 활용할 수 있도록 만들 것이고, 다소 아쉬운 결과를 나타냈지만, 성능 지표를 보면 그래프들이 계속해서 상승해 나가는 것을 볼 수 있다. 이는 성능이 충분히 향상될 수가 있다는 것을 의미하며, 이를 토대로 더 많은 데이터와 자료를 수집하고 더 좋은 모델을 연구하는 방법을 통해 본 시스템의 성능 및 정확도를 더 높이는 방안을 모색할 예정이다. 그러므로 이를 통해 시각장애인 및 사회적 약자들이 부패한 식품 섭취로 인한 위험을 줄이고, 자율성을 높일 가능성을 제시한다.

## 참 고 문 헌

- [1] Jones, N., & Bartlett, H. (2018). The impact of visual impairment on nutritional status: A systematic review. *British Journal of Visual Impairment*, 36(1), 17-30. <https://doi.org/10.1177/0264619617730860>
- [2] Franklin Mingzhe Li, Jamie Dorst, Peter Cederberg, and Patrick Carrington. 2021. Non-Visual Cooking: Exploring Practices and Challenges of Meal Preparation by People with Visual Impairments. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 30, 1-11. <https://doi.org/10.1145/3441852.3471215>
- [3] Okolo, G.I.; Althobaiti, T.; Ramzan, N. Assistive Systems for Visually Impaired Persons: Challenges and Opportunities for Navigation Assistance. *Sensors* 2024, 24, 3572. <https://doi.org/10.3390/s24113572>
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 779-788, 2016. [<https://doi.org/10.1109/CVPR.2016.91>]
- [5] Danai Brilli, D., Georgaras, E., Tsilivaki, S., Melanitis, N., and Nikita, K., "AIris: An AI-powered Wearable Assistive Device for the Visually Impaired", <i>arXiv e-prints</i>, Art. no. arXiv:2405.07606, 2024. doi:10.48550/arXiv.2405.07606.
- [6] Wang, C.-Y., Yeh, I.-H., and Liao, H.-Y. M. 2024. YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information. *arXiv preprint*. <https://arxiv.org/abs/2402.13616>.
- [7] Wang, C.-Y. and Liao, H.-Y. M. 2024. YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems. *Institute of Information Science, Academia Sinica, Taiwan*. arXiv:2408.09332v1 [cs.CV], 18 Aug. 2024.
- [8] Muhammad Yaseen, "What is YOLOv9: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector", *arXiv preprint arXiv:2409.07813*, 2024. Available at: <https://arxiv.org/abs/2409.07813>
- [9] R. Sharma, M. Saqib, C. T. Lin, and M. Blumenstein, "A Survey on Object Instance Segmentation," *SN Computer Science*, vol. 3, no. 6. Springer Science and Business Media LLC, 29-Sep-2022.
- [10] F. Porcher, C. Couprie, M. Szafraniec, and J. Verbeek, "Better (pseudo-)labels for semi-supervised instance segmentation," *Practical ML for Low Resource Settings Workshop at ICLR 2024*, arXiv:2403.11675, Mar. 2024. [Online]. Available: <https://arxiv.org/abs/2403.11675>
- [11] Wikipedia, *YAML*, <https://en.wikipedia.org/wiki/YAML>, 2024
- [12] Ultralytics, *YOLOv9*, <https://docs.ultralytics.com/models/yolov9/#performance-on-ms-coco-dataset>, Jan. 2024.

# 자율주행 차량의 안전 강화를 위한 멀티모달 상황 인식 및 컨텍스트 추론 연구

정소미\*, 이선우\*\*, 신종현\*\*\*, 김남훈\*\*, 윤현중\*\*\*\*

## Research on Multimodal Situational Awareness and Context Reasoning for Enhancing Safety in Autonomous Vehicles

So-Mi Jeong\*, Sun-Woo Lee\*\*, Jong-Hyeon Sin\*\*\*, Nam-Hun Kim\*\*, and Hyun-Joong Yoon\*\*\*\*

### 요 약

본 연구는 자율주행 차량의 안전성을 강화하기 위해 멀티모달 데이터를 활용한 상황 인식 및 컨텍스트 추론 시스템을 제안하였다. 텍스트 기반 설명을 생성하여 자율주행 시스템의 의사결정을 지원하였다. 이를 통해 다양한 주행 환경에서 높은 정확도의 상황 인식이 가능해져 자율주행 차량의 안전성 향상에 기여하였다. 실험 결과, 자율주행 시스템은 실시간 경고 메시지를 통해 운전자에게 직관적인 정보를 제공할 수 있었다. 본 연구는 탑승자 상호작용, 교통 관제 시스템 연동 등 다양한 응용 분야에서의 잠재적 활용 가능성을 제시한다. 향후 연구에서는 V2X 통신 시스템과의 연동을 통해 차량 간 정보 공유를 강화하여 자율주행의 안전성과 효율성을 더욱 높일 계획이다.

### Abstract

This study introduces a multimodal situation awareness and context inference system designed to enhance the safety of autonomous vehicles. By generating text-based descriptions, the system aids decision-making in autonomous driving and ensures precise situation awareness across various environments. Experimental results demonstrate that the system provides real-time warning messages, offering intuitive information to drivers. The study also explores applications in passenger interaction, traffic control integration, and vehicle-to-everything (V2X) communication. Future research will focus on improving safety and efficiency through enhanced information sharing between vehicles via V2X systems.

### Key words

autonomous driving, multimodal sensor fusion, Gemma language model, situational awareness, V2X communication, situational awareness, natural language processing

---

\* 국립안동대학교 컴퓨터공학과, someee0706@naver.com

\*\* 국립금오공과대학교 기계시스템공학부, gnsskazla@naver.com, amway1213@naver.com

\*\*\* 계명대학교 기계공학과, sd499@naver.com

\*\*\*\* 대구가톨릭대학교 기계공학과, yoon@cu.ac.kr(교신저자)

※ 본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신사업의 결과입니다.(2022RIS-006)

## I. 서 론

자율주행 기술은 현대 교통 체계의 패러다임을 변화시키는 혁신적 기술로 주목받고 있다[1]. 그러나 도로 환경의 복잡성과 불확실성으로 인해 안전성 확보에 있어 여전히 다양한 도전 과제가 존재한다[2]. 특히, 동적으로 변화하는 도로 상황에서의 정확한 상황 인식과 신속한 대응 능력은 자율주행 시스템의 핵심 요소로 부각된다[3].

현재의 자율주행 시스템은 개별 센서 데이터의 처리와 해석에 있어 괄목할 만한 발전을 이루었으나, 다중 센서 정보의 통합적 해석과 직관적 전달 측면에서는 여전히 한계점을 노출하고 있다[4,5]. 자율주행 차량의 안전성을 높이기 위해서는 단순한 센서 데이터 처리를 넘어, 주행 환경에 대한 포괄적 이해와 해석이 필요하다[3]. 이는 개별 센서들이 제공하는 정보를 유기적으로 통합하고, 이를 기반으로 현재 상황을 정확히 인식하며, 잠재적 위험 요소를 사전에 식별할 수 있는 고도화된 능력을 요구한다.

따라서 본 연구는 멀티모달 센서 데이터를 활용한 포괄적 상황 인식과 컨텍스트 추론 시스템을 제안한다. 제안된 시스템은 다양한 센서 데이터를 통합적으로 분석하여 주행 환경을 인식하고, 이를 토대로 직관적인 상황 설명과 안전 경고를 생성한다. 본 연구의 주요 목표는 다음과 같다. 첫째, 다양한 센서 모달리티의 효과적인 통합을 통해 포괄적이고 정확한 상황 인식 능력을 개발한다. 둘째, 고도화된 컨텍스트 추론 알고리즘을 통해 주행 환경에 대한 심층적 이해와 잠재적 위험 요소의 사전 식별 능력을 향상시킨다. 셋째, 실제 도로 환경에서의 적용 가능성과 확장성을 고려한 실용적 시스템을 구현한다. 이를 통해 자율주행 시스템의 안전성 향상에 기여하고, 시스템의 의사결정 과정을 보다 투명하고 해석 가능하게 만드는데 일조할 것으로 기대된다.

## II. 관련연구

### 2.1 센서 기술 및 데이터 융합의 진화

자율주행 기술의 발전은 센서 기술의 혁신과 데이터 융합 기법의 고도화에 기반을 두고 있다. 현대

의 자율주행 차량은 카메라, 라이다, 레이더 등 다양한 센서의 복합적 활용을 통해 주행 환경을 정밀하게 인식한다. 특히, 라이다 기술의 발전으로 초당 1,000회 이상의 스캔 빈도와 300m 이상의 인식 범위가 가능해져 고속 주행 환경에서도 신뢰할 수 있는 데이터를 제공한다[6]. 센서 융합(sensor fusion) 기술은 개별 센서의 한계를 보완하고, 복잡한 도로 환경에서 보다 정확한 상황 인식을 가능케 한다. 예를 들어, 카메라와 레이더의 데이터 융합은 악천후 상황에서도 안정적인 객체 탐지를 가능하게 한다[6]. 이러한 기술적 진보는 자율주행 시스템의 신뢰성과 안전성 향상에 크게 기여하고 있다[5].

### 2.2 멀티모달 학습 기반 상황 인식 연구 동향

멀티모달 학습은 자율주행 분야에서 상황 인식의 정확성과 포괄성을 높이는 핵심 기술로 부상하고 있다. 최근 연구는 다양한 센서 데이터와 텍스트 정보를 통합적으로 처리하는 대규모 언어 모델의 활용에 초점을 맞추고 있다[7]. EMMA(End-to-End Multimodal Model for Autonomous driving)와 같은 혁신적 접근법은 raw 카메라 센서 데이터를 직접 다양한 주행 관련 출력으로 매핑하는 능력을 보여주었다. 이러한 멀티모달 모델들은 시각적 정보와 텍스트 기반 정보를 동시에 처리함으로써, 복잡한 도로 상황에 대한 심층적 이해를 가능케 한다[2][3]. 고수준의 주행 명령과 역사적 컨텍스트를 고려한 주행 계획은 단순 객체 인식을 넘어 도로 환경의 의미론적 해석과 미래 상황 예측까지 가능하게 한다.

### 2.3 컨텍스트 추론 시스템의 최신 발전

최신 연구에서는 end-to-end 학습 방식을 통해 센서 데이터로부터 직접 주행 행동을 생성하는 방법이 제안되었다. 이는 모듈 간 심볼릭 인터페이스의 필요성을 제거하고, raw 센서 입력으로부터 주행 목표의 공동 최적화를 가능하게 한다[8]. 대규모 언어 모델 기반 컨텍스트 추론 시스템은 센서 데이터, 내비게이션 지시, 교통 규칙과 같은 다양한 입력 소스를 종합적으로 분석하여 보다 인간과 유사한 의사결정을 수행할 수 있게 되었다[8]. Chain-of-thought (CoT, 사고 과정의 연쇄) 추론 기법의 도입은 의사결정 과정의 투명성과 설명 가능성을 크게 향상시켰

다[6].

본 연구는 이러한 최신 기술 동향을 바탕으로, NuScenes 데이터셋을 활용한 멀티모달 데이터 통합과 Gemma 모델을 적용한 컨텍스트 추론을 결합하여 보다 포괄적이고 정확한 상황 인식 시스템을 제안한다.

### III. 멀티모달 데이터 처리 및 상황 인식 시스템 구현

#### 3.1 다중 센서 데이터 융합을 통한 상황 인식 기술

본 연구에서 제안하는 자율주행 차량용 멀티모달 상황 인식 시스템은 다양한 센서 데이터의 통합적 처리와 컨텍스트 기반 상황 해석을 통해 차량의 안전성을 향상시키는 것을 목표로 한다. 기존의 자율주행 시스템들이 개별 센서 데이터의 독립적 처리에 중점을 두었던 것과 달리, 본 시스템은 다중 센서 데이터의 통합적 해석을 통해 보다 정확하고 포괄적인 상황 인식을 가능하게 한다.

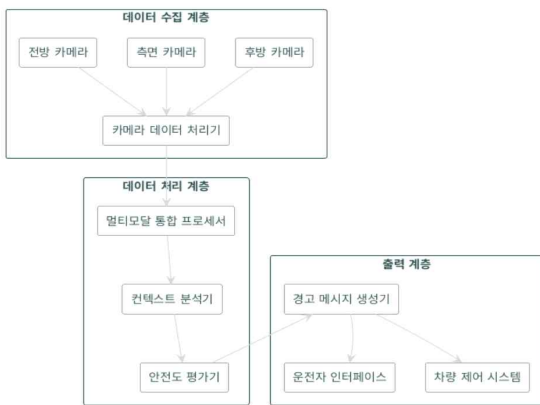


그림 1. 자율주행 시스템 데이터 처리 구조  
Fig. 1. Autonomous System Data Flow

#### 3.2 데이터셋

본 연구에서는 자율주행 차량의 멀티모달 상황 인식 시스템을 검증하기 위해 NuScenes v1.0-mini 데이터셋을 사용하였다. 이 데이터셋은 보스턴과 싱가포르의 도심 환경에서 수집된 다양한 주행 시나리오를 포함하고 있으며, 특히 주간 및 야간, 다양한 기상 조건, 복잡한 교통 상황을 반영하고 있어 자율주행 시스템의 성능을 평가하는 데 매우 적합하다.

NuScenes v1.0-mini는 총 10개의 상세 주행 시나리오로 구성되어 있으며, 각 시나리오는 다양한 교통 밀도와 환경 조건을 포함하고 있다. 카메라 구성은 차량에 장착된 6개의 동기화된 카메라를 통해 이루어지며, 각 카메라는 1600×900 해상도의 이미지를 제공한다. 이를 통해 차량 주변의 모든 방향에서 시각 정보를 수집할 수 있으며, 멀티뷰 기반의 상황 인식이 가능하다. 특히 이 데이터셋은 23개의 객체 클래스와 8개의 속성에 대한 주석 정보를 포함하고 있어 자율주행 시스템이 다양한 객체(예: 차량, 보행자 등)를 인식하고 분석할 수 있게 한다. 또한, 데이터는 2Hz의 시간 해상도로 기록되었으며, 총 약 1,100개의 키프레임이 제공된다.



그림 2. 멀티뷰 카메라 구성 예시

Fig. 2. Multiview camera configuration example

그림 2는 NuScenes 데이터셋에서 제공되는 멀티뷰 카메라 구성을 보여준다. 각 이미지 프레임은 차량의 전방(CAM\_FRONT), 후방(CAM\_BACK), 좌우 측면(CAM\_FRONT\_LEFT, CAM\_FRONT\_RIGHT, CAM\_BACK\_LEFT, CAM\_BACK\_RIGHT)에서 동시에 촬영되며, 이를 통해 차량 주변의 모든 방향에서 시각 정보를 수집할 수 있다. 이러한 멀티뷰 카메라 구성은 자율주행 시스템이 도심 환경에서 발생할 수 있는 다양한 교통 상황을 포괄적으로 인식하는데 중요한 역할을 한다.

#### 3.3 멀티모달 상황 인식 시스템 구현

멀티모달 데이터는 다양한 센서로부터 수집된 정보를 통합하여 처리해야 하므로, 각 센서 데이터에 대한 전처리 과정이 필수적이다. 본 연구에서는 NuScenes 데이터셋으로부터 수집된 영상 데이터를 전처리하여 노이즈 제거 및 해상도 조정 작업을 수행하였다. 이 과정에서 각 카메라 뷰 간의 시간 동기화가 이루어졌으며, 이를 통해 일관된 시점에서의 다중 뷰 데이터를 확보하였다. 또한 학습을 위해

Train/Test 데이터셋을 구성하였으며, 전체 데이터셋 중 80%는 학습용으로 사용되었고 나머지 20%는 테스트용으로 할당되었다. 이 과정에서 다양한 교통 상황과 환경 조건을 반영한 데이터를 학습에 활용함으로써 자율주행 시스템의 성능을 평가하였다.

본 연구는 대규모 언어 모델인 Gemma 2B를 활용하여 컨텍스트 기반 추론을 수행하였다. Gemma 모델은 다중 센서로부터 수집된 데이터를 바탕으로 상황을 해석하며, 이를 자연어로 설명하는 기능을 갖추고 있다. 특히 본 연구에서는 LoRA 최적화를 적용하여 모델의 파라미터를 효율적으로 조정하였고, 이를 통해 추론 속도와 정확도를 향상시켰다.

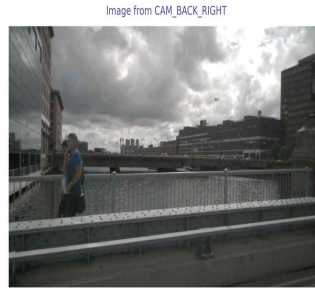
### 3.4 실험 및 분석

본 연구에서는 자율주행 차량이 다양한 도심 환경에서 직면할 수 있는 여러 가지 상황에 대해 멀티모달 데이터를 기반으로 효과적인 상황 인식을 수행하였다. 자율주행 차량은 전방, 후방, 좌우측면에 장착된 6개의 카메라를 통해 주변 환경을 실시간으로 모니터링하며, 각 뷰에서 감지된 차량, 보행자 등의 객체 개수를 바탕으로 안전 상태를 평가한다. 자율주행 시스템이 다양한 교통 상황에서 객체를 인식하고 경로 안전성을 평가하는 과정을 시각적으로 확인할 수 있다.



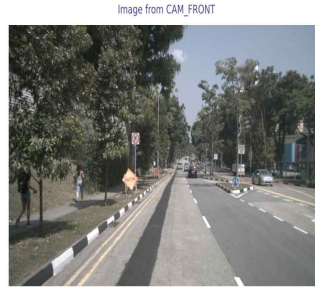
그림 3. 멀티뷰 카메라 기반 객체 인식 및 안전 설명  
Fig. 3. Multi-view camera-based object recognition and safety explained

그림 3은 멀티뷰 카메라 구성을 통해 자율주행 차량이 주변 환경을 인식하는 과정을 시각적으로 보여준다. 전방, 후방, 좌측면, 우측면에서 동시에 촬영된 이미지를 종합하여 차량 주변의 교통 상황을 분석한 결과, "Clear path ahead"라는 결론이 도출되었다. 이는 경로 상에 위험 요소가 없고 정상적인 운행이 가능하다는 것을 의미하며, 이 정보는 운전자와 차량 제어 시스템에 전달된다.



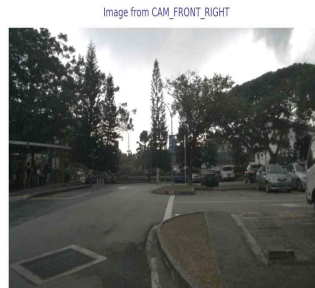
Safety Description

Detected:  
19 pedestrian(s)  
Nearest at 1729.3m  
  
Observed:  
4 vehicle(s)  
Closest at 1735.2m  
  
Status:  
Clear path ahead  
Maintain normal operation



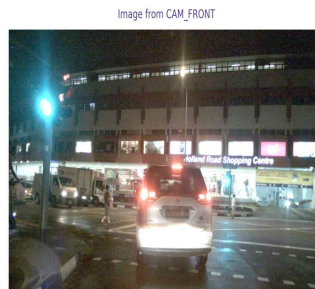
Safety Description

Observed:  
15 vehicle(s)  
Closest at 3100.4m  
  
Detected:  
5 pedestrian(s)  
Nearest at 3102.6m  
  
Note:  
1 movable object(s)  
At 3095.9m  
  
Status:  
Clear path ahead  
Maintain normal operation



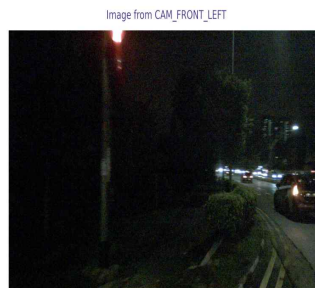
Safety Description

Observed:  
19 vehicle(s)  
Closest at 1908.3m  
  
Detected:  
20 pedestrian(s)  
Nearest at 1911.6m  
  
Status:  
Clear path ahead  
Maintain normal operation



Safety Description

Detected:  
37 pedestrian(s)  
Nearest at 1929.4m  
  
Observed:  
15 vehicle(s)  
Closest at 1917.0m  
  
Status:  
Clear path ahead  
Maintain normal operation



Safety Description

Observed:  
14 vehicle(s)  
Closest at 2060.2m  
  
Status:  
Clear path ahead  
Maintain normal operation

그림 4. 객체 인식 및 안전 설명 결과 (주·야간 환경)  
Fig. 4. Object recognition and safety evaluation results (Day & Night)



그림 4는 주간 및 야간 환경에서 수집된 데이터를 바탕으로 각 카메라 뷰에서 감지된 객체와 경로의 안전 상태를 설명하는 예시이다. 후측면 뷰에서는 여러 보행자와 차량이 감지되었으며, 가장 가까운 객체는 약 1700m 거리 내에 위치해 있다. 전방 뷰에서는 다수의 차량과 보행자가 감지되었고, 가장 가까운 객체는 약 3000m 거리 내에 있다. 우측면 뷰에서도 비슷한 상황으로, 차량과 보행자가 감지되었으며, 가장 가까운 보행자는 약 1900m 거리 내에 위치해 있다. 마지막으로 좌측면 뷰에서는 야간 환경에서 촬영된 데이터로 차량이 감지되었으며, 가장 가까운 차량은 약 2000m 거리 내에 있다.

모든 경우에서 시스템은 "경로가 명확하며 정상적인 운행이 가능하다"는 결론을 도출하였다. 특히 주간 및 야간 환경 모두에서 시스템은 멀티뷰 카메라 데이터를 성공적으로 통합하여 주변 객체의 움직임과 위치를 정확하게 파악하였으며, 이를 기반으로 실시간 경고 메시지를 생성하였다. 실험 결과는 자율주행 시스템이 다양한 도심 환경에서 높은 정확도로 객체를 인식하고 이를 바탕으로 안전 상태를 평가할 수 있음을 보여준다.

#### IV. 결 론

본 연구에서는 자율주행 차량의 안전성을 강화하기 위해 멀티모달 데이터를 활용한 상황 인식 및 컨텍스트 추론 시스템을 제안하였다. 이를 통해 자율주행 차량이 복잡한 도심 환경에서 다양한 교통 상황에 효과적으로 대응할 수 있는 능력을 확보하였다. 특히, Gemma 2B 모델과 LoRA 최적화를 통해 텍스트 기반 설명을 생성하여 자율주행 시스템의 의사결정 과정을 지원하였다.

자율주행 차량은 카메라, 라이다, 레이더 등의 센서 데이터를 실시간으로 통합하여 주행 환경을 모니터링하며, 주간 및 야간 환경 모두에서 높은 정확도를 달성하였다. 실험 결과, 자율주행 시스템은 감지된 객체의 위치와 개수를 바탕으로 경로의 안전 상태를 평가하고, 실시간 경고 메시지를 생성하여 운전자에게 직관적인 정보를 제공할 수 있었다.

또한, 생성된 텍스트 설명을 통해 교차로, 차선 변경, 회전 등의 의사결정 과정에서 추가적인 정보

를 제공함으로써 탑승자의 안전성을 높이고 사용자 경험을 향상시킬 수 있었다. 더 나아가, 교통 관제 시스템과의 연동을 통해 도시 전체 교통 흐름을 모니터링하고 관리할 가능성도 확인되었다.

향후 연구에서는 V2X 통신 시스템과의 연동을 통해 자율주행 차량 간 정보 공유를 강화하고, 도로 인프라 개선 방안을 모색할 예정이다. 또한, 다양한 도로 환경에서의 실시간 성능 향상을 통해 실제로 도로 주행 환경에서의 적용 가능성을 높이는 방향으로 시스템을 발전시킬 계획이다.

#### 참 고 문 헌

- [1] Sapien, "The Role of Multimodal AI in Autonomous Vehicles," Sapien Blog, 2023. [Online]. Available: <https://www.sapien.io/blog/the-role-of-multimodal-ai-in-autonomous-vehicles>
- [2] M. van Rijmenam, "Understanding Self-Driving Vehicles with an Innovation Futurist's Help," The Digital Speaker, 2023. [Online]. Available: <https://www.thedigitalspeaker.com/understanding-self-driving-vehicles-innovation-futurist-help/>
- [3] N. Katiyar et al., "AI in Autonomous Vehicles: Opportunities, Challenges, and Regulatory Implications," Educational Administration: Theory and Practice, vol. 30, no. 4, pp. 6255-6264, 2024. [Online]. Available: <https://kuey.net/index.php/kuey/article/download/2373/1375/6326>
- [4] TNO, "Situational Awareness for Autonomous Driving," TNO, 2023. [Online]. Available: <https://www.tno.nl/en/digital/artificial-intelligence/safe-autonomous-systems/situational-awareness-autonomous-driving/>
- [5] A. Sauer et al., "A Survey of Deep Learning Techniques for Autonomous Driving," arXiv preprint arXiv:2202.02703, 2022. [Online]. Available: <http://arxiv.org/abs/2202.02703>
- [6] M. Aguilar, "Vehicle Sensors for Autonomous Driving Safety," LinkedIn, 2023. [Online]. Available: <https://www.linkedin.com/pulse/advances-vehicle-sensor-technology-autonomous-driving-aguilar>
- [7] S. Bhattacharyya et al., "EMMA: An End-to-End Sequence-to-Sequence Multimodal Model for Autonomous Driving," arXiv preprint arXiv:2410.23262, 2023.
- [8] Z. Liu et al., "Driving with LLMs: Fusing Object-Level Vector Modality for Explainable Autonomous Driving," arXiv preprint arXiv:2312.04316v3, 2023.

# 시각장애인을 위한 AI기반 식품 정보 제공 서비스

안상우\*, 강부진\*, 황순규\*, 모아사랑\*, 정기현\*\*

## Design and Implementation of an AI-based Food Information Service for the Visually Impaired

SangWoo An\*, BuJin Kang\*, SoonGyu Hwang\*, A Sa Rang Mo\*, and KiHyun Jung\*\*

### 요 약

본 논문에서는 식품에 관한 정보를 얻기 힘든 시각장애인 및 사회적 약자들이 식품의 종류를 인지할 수 없어 생기는 불편함과 건강에 위협을 받는 문제를 해결하기 위해, 인공지능을 활용한 식품 정보 제공 시스템을 제안한다. 제안된 시스템은 YOLO 모델을 활용하여 실시간으로 식품을 탐지한 후, 음성으로 결과를 제공한다. 본 연구에서는 데이터 수집 및 전처리를 위해 Roboflow를 사용하였으며, YOLO 모델을 활용한 객체 탐지 및 세그멘테이션 학습을 진행하였다. 또한, React와 ONNX Runtime을 사용하여 PWA를 구현하고, 실시간 비디오 스트림 처리 및 자동 촬영 기능을 통합하여 사용자 편의성을 증대하였다. 본 연구는 식품에 관한 정보를 얻기 힘든 시각장애인을 포함한 사회적 약자들이 안전하고 독립적인 생활을 영위할 수 있도록 돕는 데 기여할 수 있을 것으로 기대하며, 지속적으로 성능 향상을 시킬 계획이다.

### Abstract

This paper proposes a food information system using artificial intelligence to address the challenges faced by visually impaired individuals and other vulnerable groups who may struggle to identify types of food, leading to inconvenience and potential health risks. The proposed system employs the YOLO model to detect food in real time and provides results via audio output. For data collection and preprocessing, Roboflow was used, and object detection and segmentation were performed using the YOLO model. Additionally, a Progressive Web App (PWA) was implemented with React and ONNX Runtime, integrating real-time video stream processing and an automatic capture function to enhance user convenience. This research aims to support a safe and independent lifestyle for visually impaired individuals and other vulnerable groups who struggle to access food information.

### Key words

Artificial Intelligence, Roboflow, YOLO, PWA

---

\* 국립안동대학교 멀티미디어공학과, asw421@naver.com

\*\* 국립안동대학교 소프트웨어융합학과, kingjung@anu.ac.kr (교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

## I. 서 론

기술의 발전에도 불구하고 장애를 가진 사람들의 불편함은 여전히 완전히 해결되지 않고 있다. 특히 시각 장애인들은 일상적인 식사 과정에서 많은 어려움을 겪고 있다. 비장애인은 눈으로 식품을 확인하여 그 종류와 상태를 쉽게 파악할 수 있지만, 시각 장애인은 외부의 도움 없이는 식품이 무엇인지 알 수 없다. 결과적으로 시각 장애인들은 식품을 섭취한 이후에 정보를 인식하게 되는 불편함이 있다 [1].

본 연구에서는 시각 장애인들이 겪는 이러한 불편함을 해소하기 위해, 식품 정보 제공 서비스를 제안한다. 기존에도 스마트폰 앱을 통해 사물 정보를 알려주는 서비스가 존재하지만, 대부분이 단순히 사물의 종류만을 제공하는 데 그친다. 식품 정보나 보관 방법, 부패 상태와 같은 세부적인 정보를 제공하는 플랫폼은 많지 않다. 시각 장애인들은 이러한 정보를 스스로 쉽게 얻기 어려우며, 이로 인해 원하지 않는 식품을 섭취하는 경우가 발생하기도 한다.

본 연구에서 제안하는 시스템은 시각 장애인이 식품의 종류를 손쉽게 파악하도록 도와줌으로써, 식품에 관한 불편함과 위험을 줄이고자 한다. 예를 들어, 시각 장애인 앞에 사과가 있을 경우, 시스템은 해당 식품이 사과임을 인식하고, 음성으로 정보를 제공한다. 이를 통해 사용자는 식품을 직접 섭취하지 않고도 식품 정보를 확인할 수 있고, 자율성을 가지게 할 수 있다. 이는 시각 장애인들이 더 안전하고 편리한 식생활을 유지할 수 있도록 돕는다[2].

본 연구에서는 YOLO(You Only Look Once)모델을 사용하여 실시간으로 식품을 탐지할 수 있는 인공지능 시스템을 설계하였다. 이 시스템은 스마트폰을 통해 사용자가 촬영한 사진에서 식품 정보를 분석하고, TTS(Text-To-Speech), 경고음, 시각적 경고 표시를 통해 사용자가 음성 및 시각 정보를 모두 받을 수 있도록 설계되었다. 또한, 시각 장애인들이 보다 쉽게 카메라를 활용할 수 있도록, 식품이 탐지되면 자동으로 사진을 촬영하는 기능도 제공하여 접근성을 높였다.

이 인공지능 시스템은 YOLO 모델을 파인튜닝하여 설계한 부패 탐지 AI로 구성되며, 각각 식품의

정보를 판단하여 제공한다. 기존의 사물 인식 서비스들은 단순한 정보 제공에 그친 반면, 본 시스템은 식품의 구체적인 정보와 상태까지 함께 수행하는 점에서 차별성이 있다. 이를 통해 시각 장애인들은 건강과 안전에 보다 큰 도움을 받을 수 있다[3].

본 연구에서 개발한 식품 정보 제공 서비스는 시각 장애인뿐만 아니라, 식품 관련 공공기관, 복지시설, 학교 급식소 등에서도 활용될 수 있을 것으로 기대된다. 또한, 이 시스템은 시각 장애인을 넘어 노약자나 다양한 사용자 그룹이 안전하게 식품을 관리하고 소비할 수 있도록 확장 가능성이 크다. 향후 기술 발전에 따라 더 많은 데이터를 기반으로 다양한 식품 정보를 제공할 수 있으며, 이를 통해 식품 폐기로 인한 환경적, 경제적 손실을 줄이는 데도 기여할 수 있다. 본 연구는 시각 장애인들의 식생활에 실질적인 도움을 주고, 그들의 자율성과 안전성을 향상시킬 수 있는 새로운 기술적 기반을 마련함으로써, 사회적 가치와 경제적 효용성 또한 크게 증대시킬 수 있을 것이다.

## II. 관련 연구

### 2.1 ONNX

ONNX(Open Neural Network Exchange) 라이브러리는 딥러닝 모델을 여러 프레임워크와 환경에서 사용할 수 있도록 변환하고 활용할 수 있도록 만들어진 개방형 표준이다. Pytorch, TensorFlow, Keras 등 다양한 딥러닝 프레임워크에서 개발된 모델들을 ONNX 형식으로 변환하여, 호환되는 다른 프레임워크나 환경에서 사용할 수 있다. 또한 딥러닝 모델을 최적화하여 효율적으로 모델을 실행할 수 있도록 지원을 해주어, AI가 실행되는 장치와 상관없이 원활하게 작동할 수 있도록 해준다.

ONNX(Open Neural Network Exchange) 라이브러리는 ONNX 형식을 실행 할 수 있는 런타임도 지원을 한다. 대표적으로 마이크로소프트에서 제공하는 ONNX Runtime이 있다. 이 런타임은 ONNX 모델의 성능을 극대화 시켜주며, 플랫폼 간 호환성을 향상시켜준다.

본 연구에서는 TensorFlow에서 제작된 딥러닝 모

델을 ONNX 라이브러리를 활용하여 JavaScript 환경에 적용시켰다. 이는 기존의 프론트에서 정보를 얻어 AI 서버로 데이터를 전송하고 다시 정보를 받아오는 과정을 생략시켜 프론트 서버 내에서 직접 모델을 실행시키고 정보를 처리함으로써 기존보다 빠른 결론 도출이 가능했다. React 환경에서 개발된 플랫폼에서 모바일 기기의 카메라를 통해 객체 탐지를 하는 분야에 ONNX 라이브러리를 적용시켜 보았을 때, ONNX 라이브러리를 사용하지 않고 AI 서버와의 통신을 통해 객체 탐지를 할 때는 서버와 통신을 하고 결론을 도출해서 다시 서버를 통해 정보를 받아오는 시간이 다소 많이 소모된 반면, ONNX 라이브러리를 사용했을 때 통신 없이 모델을 불러오고 결론을 도출해내는데 소모되는 시간이 매우 감소된 모습을 확인할 수 있었다. 이를 통해, 사용자들의 편의성을 향상시킬 수 있었다.

## 2.2 시각 장애인의 식품 확인 문제

시각 장애인들은 음식의 상태를 육안으로 확인하는 데 어려움을 겪기 때문에 원하지 않고, 위험한 음식을 섭취할 가능성이 상대적으로 높다. 일반적으로 사람들은 시각을 통해 음식의 상태를 쉽게 파악하고 확인할 수 있지만, 시각 장애인들에게는 이러한 과정이 매우 복잡하고 힘들다. 이는 식욕 부진 및 심각한 건강 문제로 이어질 수 있다. 이에 따라 시각 장애인들이 안전한 식생활을 유지할 수 있도록 도와주는 기술적 해결책이 꾸준히 연구되고 있다.

기존의 사물 인식 기술들 중 하나인 Microsoft Seeing AI는 스마트폰 카메라를 통해 주변 사물을 인식하고, 이를 음성으로 안내하여 시각 장애인들에게 도움을 주고 있다. 그러나 이와 같은 시스템은 주로 사물의 기본적인 정보를 제공하는 데 그치며, 음식의 정보와 같은 세부적인 정보를 제공하는 데에는 한계가 있다. 시각장애인들이 원하는 식품을 섭취할 수 있도록 하는 구체적인 식품 확인 시스템은 여전히 필요하다.

## 2.3 YOLO 기반 객체 탐지 모델

최근 몇 년 동안 YOLO(You Only Look Once) 모델을 기반으로 한 객체 탐지 기술은 놀라운 발전을 이루었다. YOLO는 이미지에서 여러 객체를 동시에 탐지할 수 있는 모델로, 특히 실시간 탐지가 필요한 응용 분야에서 큰 주목을 받고 있다. 최신 버전인 YOLOv8은 이전 모델보다 더욱 정밀한 탐지 성능과 빠른 처리 속도를 제공하며, 다양한 산업 분야에서 활용되고 있다. 본 연구에서는 YOLOv8 모델을 활용하여 시각 장애인들이 음식의 종류와 상태를 인식하는 FDA(Food Detection AI)를 개발하였다. 이 시스템은 사용자가 스마트폰 카메라를 통해 음식을 촬영하면 실시간으로 음식의 종류를 탐지하고, 음성으로 안내하는 기능을 제공한다.

## 2.4 객체 탐지 성능 향상용 라벨링 데이터 구성

객체 탐지 모델의 성능을 최적화하기 위해서는 정확한 라벨링 데이터가 필수적으로 본 연구에서는 바운딩 박스(Bounding Box) 방식과 세그멘테이션(Segmentation) 방식을 모두 활용하여 라벨링 데이터를 구성하였다. 각각의 라벨링 기법은 특정 목적에 맞춰 모델의 성능을 향상시키는 역할을 했다.

바운딩 박스는 이미지 내에서 객체의 위치를 사각형으로 감싸는 방식으로, 객체의 경계를 정의하는 가장 기본적인 라벨링 기법이다. 이 방법을 통해 모델은 객체의 크기와 위치를 빠르고 정확하게 파악할 수 있으며, 여러 객체가 혼재된 이미지에서도 각 객체의 위치를 효과적으로 인식할 수 있다. 이러한 바운딩 박스를 기반으로 YOLOv8 모델은 음식의 위치와 크기를 빠르게 탐지하고, 객체 인식의 기본 구조를 형성하였다.

세그멘테이션은 객체의 경계를 더욱 세밀하게 구분하는 기법으로, 객체의 픽셀 단위까지 분석하여 모양과 세부적인 영역을 구분할 수 있다. 세그멘테이션은 특히 음식의 표면 상태나 부패 여부를 정확하게 인식하는 데 중요한 역할을 했다. 예를 들어, 사과와 같이 표면이 부패한 경우, 세그멘테이션을 통해 사과의 부패된 부분을 정확하게 식별할 수 있다. 이러한 기법은 단순히 객체의 위치를 탐지하는 것을 넘어, 객체의 상태 변화를 구체적으로 분석할

수 있는 중요한 요소로 작용했다.



그림 1. 바운딩 박스와 세그멘테이션 기법  
Fig. 1. Bounding box and segmentation techniques

바운딩 박스와 세그멘테이션을 결합한 라벨링 방식은 YOLOv8 모델의 탐지 정확도와 세부 인식 능력을 크게 향상시켰다. 이 두 가지 라벨링 기법을 통해 음식의 크기, 위치, 모양뿐만 아니라 상태 변화를 세밀하게 파악할 수 있었으며, 특히 부패 상태 탐지에서 높은 성능을 발휘할 수 있었다. 이를 통해 시스템은 다양한 환경에서도 일관성 있는 성능을 유지할 수 있었고, 시각 장애인들이 음식을 보다 안전하게 관리할 수 있는 기반을 제공할 수 있었다.

### III. 시스템 설계 및 구현

#### 3.1 아키텍처 설계

본 연구에서는 인공지능 구축 설계 및 구현 방법으로 파이프필터 패턴을 사용하였다. 파이프 필터 패턴은 소프트웨어 아키텍처 패턴 중 하나로 데이터 처리를 위한 구조를 나타낸다. 처리 단계는 데이터 수집, 데이터 전처리, 인공지능 학습 및 검증, 결과 확인 순서로 설계 하였다. 이 패턴에 대한 다이어그램은 [그림 2]와 같다.

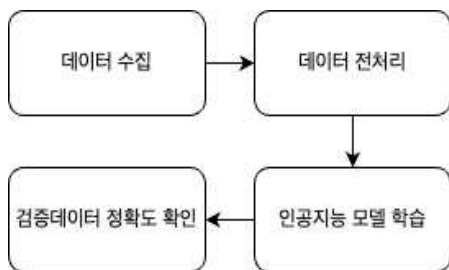


그림 2. 파이프필터 패턴 다이어그램

Fig. 2. Pipe-and-Filter Pattern Diagram

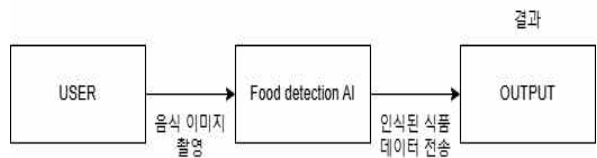


그림 3. 시스템 아키텍처 다이어그램  
Fig. 3. System Architecture Diagram

[그림 3]의 시스템 아키텍처 다이어그램은 소프트웨어 시스템의 구성 요소와 그들 간의 관계를 시각적으로 표현한 것이다. 본 연구는 사용자, 인공지능 모델, 출력 단계로 나누어서 사용자의 데이터를 처리하도록 설계하였다. 인공지능은 식품을 탐지하는 식품 탐지 모델을 설계하여, 사용자가 사진을 전송하면 인공지능 모델의 판단에 따라 결과 데이터를 제작한 웹앱 사이트로 전송하여 실시간 탐지 결과 전송과 저장 등의 단계를 실행한다.

#### 3.2 AI 모델 학습 및 데이터셋 구성

본 연구에서는 Roboflow 플랫폼을 통해 수집된 30,000장 이상의 음식 이미지 데이터셋을 사용하여 부패 탐지 AI를 학습시켰다. 이 데이터셋은 음식의 다양한 종류와 조건을 반영하여 모델이 실제 상황에서도 높은 정확도로 음식을 탐지할 수 있도록 구성되었다.

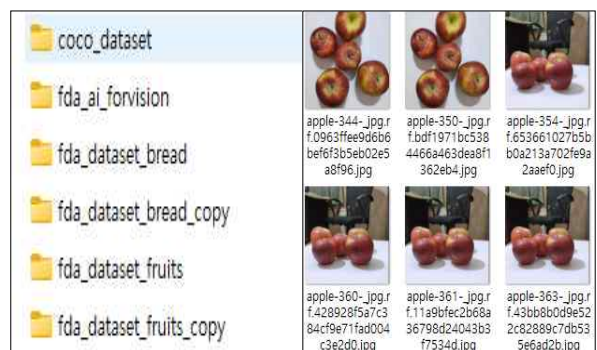


그림 4. YOLOv8 모델 학습에 사용된 데이터셋  
Fig. 4. Dataset used for training the YOLOv8 model

추가적으로, 전이 학습(Transfer Learning) 기법을 적용하여 기존에 학습된 YOLOv8 모델을 바탕으로 추가적인 학습을 진행하였다. 이를 통해 기존의 학

습 경험을 활용하면서 모델의 성능을 더욱 높일 수 있었다. 시스템은 Node.js 서버와 MySQL을 사용하여 분석된 결과를 저장하고, 사용자에게 실시간으로 피드백을 제공한다. 이러한 자동화된 과정 덕분에 시각장애인들은 음식의 상태를 빠르고 쉽게 확인할 수 있으며, 부패된 음식을 섭취하는 위험을 줄일 수 있다.

### 3.3 구현

본 연구에서는 사용자가 카메라를 활용해 식품을 바라보면 음식 탐지 AI인 FDA와 ONNX 라이브러리를 활용한 자동 촬영 기능을 통해 사진을 전송하면 인공지능 모델의 결과 데이터를 웹앱 사이트로 전송하여 실시간 탐지 결과 저장 및 음성으로 안내 등의 단계를 구현하였다.

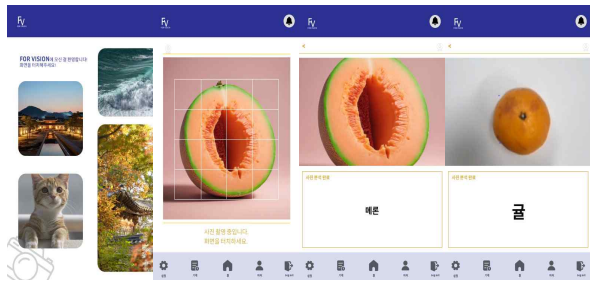


그림 5. PWA 구현  
Fig. 5. PWA implementation

## IV. 성능 평가 및 결과

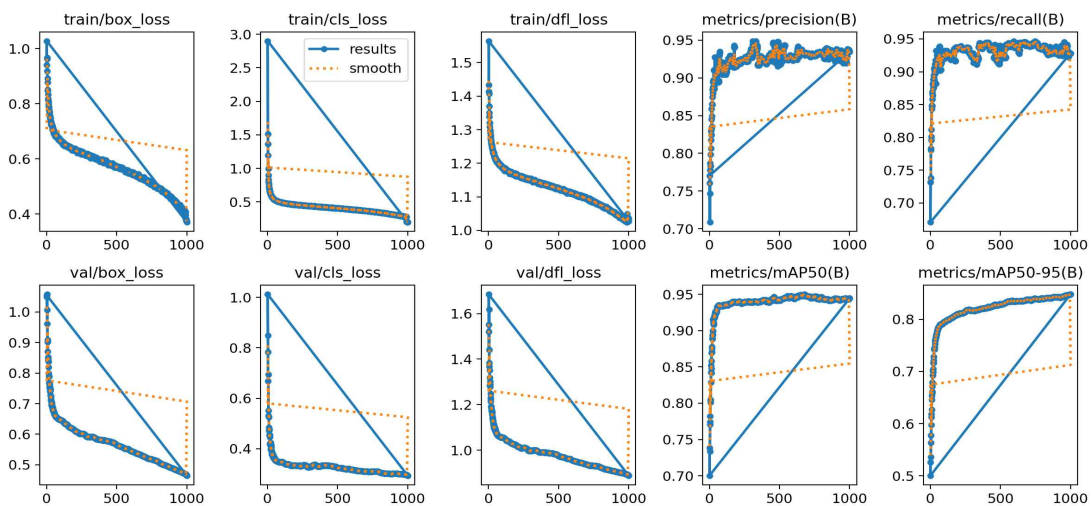


그림 6. 성능 평가 결과

### 4.1 Food Detection AI 성능

시스템의 성능을 평가하기 위해 정확도 (Accuracy), 정밀도(Precision), 재현율(Recall), mAP50, mAP50-95와 같은 주요 지표를 사용하여 YOLOv8 모델을 테스트했다.

FDA는 매우 높은 성능을 보였으며, 정밀도는 0.95, 재현율은 0.9 이상의 결과를 기록하였다. 이는 본 모델이 다양한 음식 종류를 정확하게 탐지할 수 있음을 나타낸다.

### 4.2 모델 학습 과정 분석

모델 학습 중 박스 손실(box loss)과 클래스 손실(cls loss)은 훈련이 진행될수록 꾸준히 감소하였으며, 이는 모델이 점점 더 정확하게 객체를 탐지하고 분류하는 능력을 향상시켰음을 보여준다. mAP50 값은 0.9 이상에 도달하여 매우 높은 탐지 성능을 보여주었으며, 정밀도와 재현율 역시 상승하였다. 이를 통해 시스템이 안정적이고 신뢰할 수 있는 결과를 실시간으로 제공할 수 있음을 입증하였다.

결과적으로, 본 연구에서 개발된 AI 기반 식품 탐지 시스템은 시각 장애인들이 음식을 안전하게 관리할 수 있도록 돕는 데 매우 유용하며, 공공기관이나 복지시설에서도 식품 관리의 효율성을 높이는 데 기여할 수 있을 것으로 기대된다.

Fig. 6. Performance evaluation results

## V. 결론 및 향후 과제

본 연구에서는 식품에 관한 정보를 얻기 힘든 시각장애인 및 사회적 약자들에게 일상생활에서 스스로 식품을 식별하고 확인하는 데 도움을 주어 안전하고 편리한 식생활을 지원하고자하는 목적으로 AI 기반 식품 정보 제공 서비스를 설계하고 구현하였다. YOLOv8 모델을 활용한 FDA 모델을 구현하여 웹앱 서비스를 제공함으로써, 사용자는 카메라를 이용해 식품을 바라보기만 하면, 제안한 FDA가 자동으로 식품을 탐지하고 촬영한 후 분석하여 결과를 음성으로 안내하며, 이를 통해 시각장애인 및 사회적 약자들이 부패된 식품을 섭취하여 발생할 수 있는 위험을 줄이고, 자율성과 생활의 질을 향상시킬 수 있는 가능성을 보여준다.

제안한 인공지능 모델은 데이터를 활용하여 빠르게 탐지할 수 있는 능력을 갖추었다. 이는 식품을 탐지할 뿐만 아니라 향후 추가적인 기능을 통해 다양한 분야에서 활용할 수 있음을 보여준다.

본 연구가 가진 한계점은 첫째, 식품 데이터셋이 제한적이다. 둘째, 식품에 대해 다양한 각도에서의 확인이 어렵다. 이러한 한계점을 극복하기 위하여, 향후 연구에서는 보다 다양하고 많은 식품 데이터셋을 구축하여 연구를 진행할 필요가 있다. 또한, 식품에 대해 다양한 각도에서의 확인을 위해 영상 촬영 방안을 모색하고, 이에 따른 AI 시스템을 구축하여 모델 활용 가능성을 확장하여 다양한 각도 및 촬영 변수에서 성능 및 정확도를 더 높일 수 있는 방안과 다각도 탐지 및 실시간 처리 성능을 더욱 강화할 계획이다.

## 참 고 문 헌

[1] 에이블뉴스, "시각장애인이 식사할 때 어려움 아시나요", 2024.  
[2] 예기훈, 위대광, 이완석, "시각 장애인의 기본적·수단적 일상생활수행능력에 관한 연구", 대한시과학회지 학술대회, 2017.  
[3] 이균호, 문미경, "YOLO알고리즘을 활용한 시각

장애인용 식사보조 시스템 개발", 한국전자통신학회논문지 16(5), 2021.

[4] 배선영, "시각장애인을 위한 인공지능 관련 연구 동향: 1993-2020년 국내·외 연구를 중심으로", 한국콘텐츠학회논문지 20(10), 2020.  
[5] 김은정, 박성제, 김종원, "시각장애인의 스마트폰 수용에 관한 연구", 한국산업정보학회논문지 21(4), 2016.  
[6] 박시웅, 김정은, 강현서, 박형준, "시각장애인의 라이프 사이클을 지원하는 인공지능 웨어러블 플랫폼", Journal of Platform Technology 8(4), 2020.  
[7] 박상헌, 전태재, 김상혁, 이상운, 김주완, "시각장애인을 위한 딥러닝기반 심볼인식", 한국정보통신기술학회논문지 9(3), 2016.

# 또래 음성 TTS를 활용한 한국어 학습 서비스 설계 및 구현

김유진\*, 홍윤기\*, 권태현\*, 이석준\*, 정기현\*\*

## Design and Implementation of Korean Language Learning Service Using Peer Voice TTS

Yu Jin Kim\*, Yun ki Hong\*, Tae Hyun Kwon\*, Seok Jun Lee\*, and Ki Hyun Jung\*\*

### 요 약

본 연구는 다문화 가정과 교육 소외계층 아동들의 한국어 학습을 지원하기 위해 또래 음성 TTS와 생성형 AI를 활용한 한국어 학습 서비스를 제안한다. 제안된 시스템은 아동 음성 TTS와 STT 기술을 통해 또래와의 상호작용을 모방하여 사용자에게 친근한 학습 환경을 제공하고 대화를 통해 자연스러운 한국어 학습을 진행한다. 이 서비스는 단어 학습, 상황별 대화, 자유 대화 기능을 통해 학습자가 어휘력과 의사소통 능력을 향상할 수 있도록 도울 수 있으며, 피드백 제공을 통하여 아동들의 언어적·문화적 이해를 돕는데 활용 가능하다.

### Abstract

This study proposes a Korean language learning service using peer voice TTS and generative AI to support Korean language learning for children from multi-cultural families and educationally marginalized groups. The proposed system mimics peer interaction through child voice TTS and STT technology to provide users with a friendly learning environment and learn Korean naturally through conversation. The service helps learners improve their vocabulary and communication skills through word learning, contextual dialog, and free conversation, and helps children improve their linguistic and cultural understanding by providing feedback.

### Key words

Korean Learning Speech-To-Text(STT), Text-to-speech(TTS), ChatBot, Peer Voice TTS, Generative AI

## 1. 서 론

다문화 가정은 서로 다른 국적 또는 문화의 사람이 만나 구성되는 가족을 말한다. 한국 다문화 가정의 수는 매년 증가하여 2015년 다문화 가구의 수는

약 29만명, 2023년 다문화 가구의 수는 약 41만명으로 확인된다. 다문화 가정의 증가로 다문화 아동의 수도 2010년을 이후로 매년 증가하고 있다. 다문화 학생들이 경험한 학교폭력 피해 경험 비율은 전체 학생과 비교하여 약 8배 이상 높게 나타난다. 이러

\* 국립안동대학교 소프트웨어융합학과, jean71798@gmail.com

\*\* 국립안동대학교 소프트웨어융합학과, kingjung@anu.ac.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)



한 문제는 다문화 학생들의 적응 어려움, 교우관계 등의 문제로 이어져 학업 중단으로 이어지고 있다. 다문화 학생들의 한국어 학습 부족 문제가 학교 폭력으로 이어지고 있다.

청소년기 또래와의 관계가 제대로 되지 않는다고 자각하게 되면 청소년들은 심리적으로 어려움을 느낀다[1]. 청소년기 또래와의 상호작용은 삶의 만족도와 밀접한 관련이 있으며 언어 발달 수준 향상에 효율적이다.

본 연구에서는 또래 아동과의 상호작용을 통해 한국어를 학습할 수 있는 서비스를 제안한다. 생성형 AI와 아동 음성을 모방한 TTS(Text-To-Speech)를 활용하여 한국어를 학습할 수 있는 서비스를 제공한다. STT(Speech-To-Text)를 활용하여 생성형 AI의 프롬프트(Prompt)를 음성으로 전송하고 아동 음성으로 응답을 제공하는 음성 대화 시스템을 구현한다. 청소년기의 아동들에게 또래 대화를 모방한 형식의 학습을 통해 의사소통 문제를 해결하고 학습의 효율을 높이고자 한다.

## II. 관련 연구

### 2.1 TTS (Text-To-Speech)

TTS 기술은 텍스트 입력을 음성 출력으로 변환하여 인간의 음성을 모방하는 시스템으로 인공지능과 음성 상호작용 기술의 핵심 분야이다[2]. 초기 TTS 시스템은 파라메트릭 방식으로 개발되었으며, 특히 포먼트 합성(Formant Synthesis) 방식과 통계적 파라메트릭 합성(Statistical Parametric Speech Synthesis) 시스템이 주요 방법론으로 활용되었다. 이러한 초기 시스템들은 음성의 자연스러움을 어느 정도 제공했지만, 한계점도 가지고 있다[3]. 최근 딥러닝의 발전으로 TTS 기술이 크게 향상되었다. Tacotron 2, Transformer TTS, WaveNet 등과 같은 딥러닝 기반 시스템은 텍스트 분석부터 음향 모델링, 보코딩에 이르는 모든 과정을 통합하여 음성 합성의 품질을 높이고 사람의 목소리와 유사한 수준의 음성을 생성한다. 딥러닝 모델은 수많은 텍스트-음성 페어 데이터를 활용하여 텍스트와 음성 간의 특징을 효과적으로 학습함으로써 기존 기술보다 자연

스러운 음성을 제공한다. TTS는 미디어를 포함해 다양한 상호작용 시스템에서 중요한 역할을 수행하고 있다.

### 2.2 TTS API

API는 정의 및 프로토콜 집합을 사용하여 두 소프트웨어 구성 요소가 서로 통신할 수 있게 하는 메커니즘이다. 기업의 서비스나 소프트웨어, 데이터를 외부 기업이 이용할 수 있도록 해주는 소프트웨어 간 인터페이스이다[4]. 공개된 API를 활용하면 외부의 제3자가 자신의 소프트웨어나 애플리케이션을 작동하도록 명령어와 데이터를 통신할 수 있게 된다[2]. 소프트웨어 개발자들이 활용하면 개발 비용 및 시간을 줄이는데 효과적이다. 대표적인 API들 중 Kakao Map API는 지도 서비스를 제공하고 Google Cloud Api는 인공지능 서비스와 함께 대량의 데이터 분석을 제공하며 OpenAI API는 대화형 채팅 시스템 제작에 사용된다. TTS API인 SK open API는 데이터를 실시간으로 전달하고 다양한 TTS 목소리와 함께 발화 조절, 오디오 저장 등의 기능을 포함하고 있다.

## III. 시스템 설계 및 구현

### 3.1 시스템 설계

시스템은 Flask 서버를 중심으로 STT 변환에는 React를 활용한다. 생성형 AI로 OpenAI API를 사용하여 맞춤형 응답을 생성하고 아동 음성 TTS를 활용하여 대화 시스템을 설계하였다.

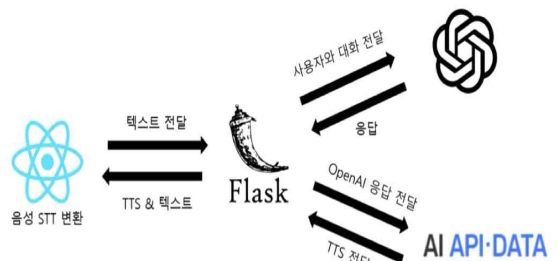


그림 1. 시스템 아키텍처  
Fig. 1. System Architecture

사용자가 녹음한 음성을 텍스트로 변환하여 백엔드 서버로 전달한다. 백엔드는 OpenAI API에 해당 텍스트를 전달해 결과값을 받아온다. 다음으로는 텍스트 결과값을 SK openAPI를 활용한 TTS로 변환해 음성 데이터를 받아와 텍스트 결과값과 음성 데이터를 화면에 출력한다. 사용자는 이러한 시스템을 통해서 음성과 함께 생성형 AI와 대화하며 학습을 진행할 수 있다.

본 연구에서 활용한 외부 데이터는 두 가지이다. 단어 학습을 위한 단어 데이터는 국립국어원의 '표준국어대사전' 데이터를 활용하여 단어 발음, 단어 의미 등을 구성하였다. 아동 음성 TTS 제작을 위해 AI-Hub의 '한국어 아동 음성 데이터'를 활용하였다. 해당 데이터는 WAV파일과 JSON 파일로 구성되어 학습하고자 하는 아동 음성 데이터를 정제하여 사용하였다. 오픈 소스인 Tortoise TTS, SK openAPI, 아동 음성 데이터를 활용하여 음성 생성 시간을 단축하고 실제 또래의 목소리처럼 자연스럽게 친근한 음성을 제공한다.

### 3.2 시스템 구현

서비스는 언어 선택, 단어 학습, 상황별 대화, 자유 대화로 기능을 분류할 수 있다. 사용자가 메인 페이지에서 언어를 선택하면 해당 언어를 React 내 변수로 저장해 다른 페이지에 접속하여도 단어를 제외한 모든 창이 해당 언어로 변환된다.

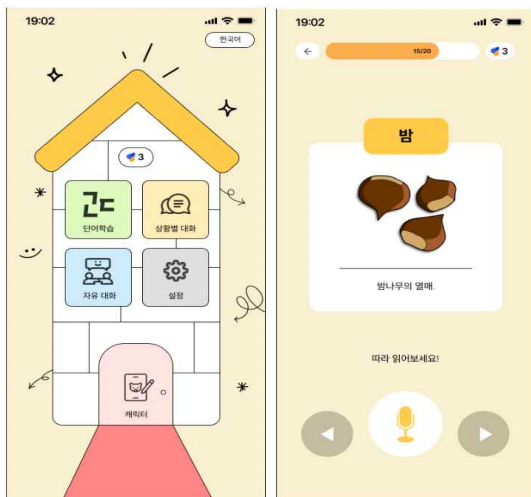


그림 2. 메인과 단어학습 서비스  
Fig. 2. Main and Word Learning Service

단어 학습은 사용자가 단어에 대한 음성 녹음을 완료하면 React 내 react-speech-kit 라이브러리에 있는 AI를 통해 음성 데이터를 텍스트 처리하여 사용자가 학습을 원활하게 진행할 수 있도록 구현했다.

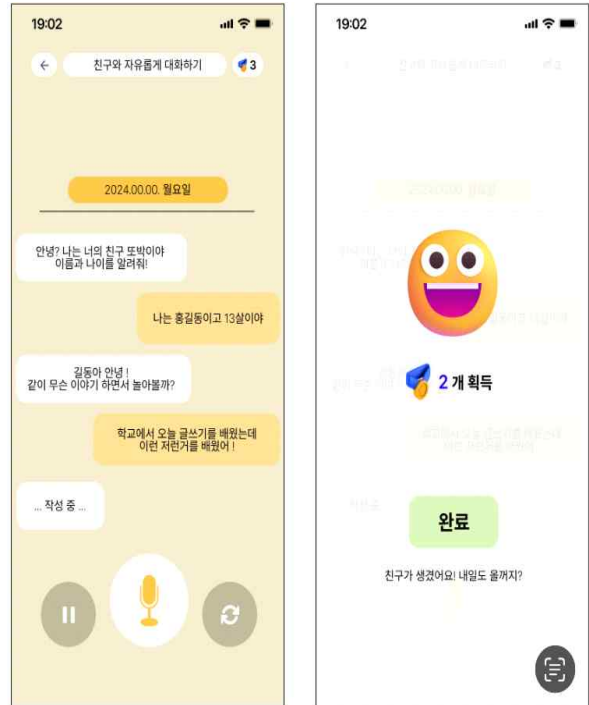


그림 3. 자유 대화 서비스  
Fig. 3. Free Talk Service

자유 대화 기능은 사용자의 첫 대화 부분에서 나이, 이름, 친구 유형 등을 필터링하기 위해 자체 함수를 제작하였다. korea\_to\_num() 함수를 통해 나이를 숫자로 변환하고 extract\_name\_from\_num() 함수를 통해 문장 내에서 이름이 아닌 조사, 명사를 구분하여 필터링한다. 사용자 프로필을 기반으로 생성형 AI는 난이도 조절 및 발화 속도를 조절한다. 이후 사용자의 음성 인식 내용을 텍스트로 변환하여 생성형 AI 전송한다. 생성된 응답을 텍스트로 받아와 또래 음성 TTS를 통해 음성으로 사용자에게 응답을 전달하는 음성 대화 시스템을 사용한다.

자유 대화 기능은 사용자가 다양한 주제에 대해 대화를 나누며 어휘력을 향상시키고, 새로운 단어와 표현을 익힐 수 있도록 돕는다. 대화를 통해 풍부한 어휘를 접하고 반복적으로 사용해 봄으로써, 자연스럽게 어휘 사용 능력을 높일 수 있다.

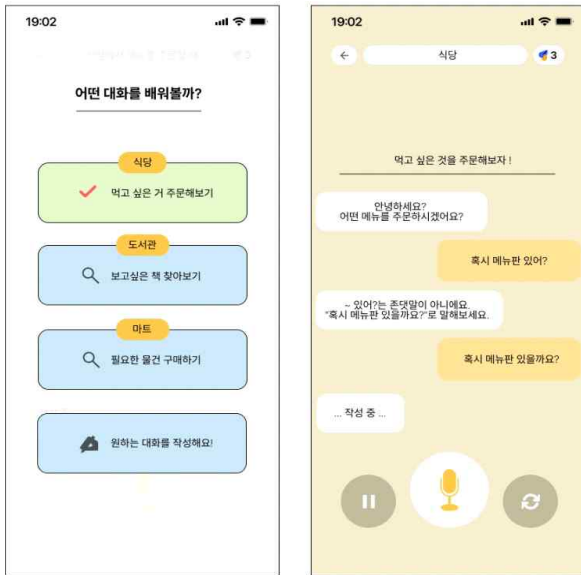


그림 4. 상황별 대화 서비스  
Fig. 4. Contextual Talk Service

상황별 대화 학습 기능은 사용자가 여러 상황 중 하나를 선택할 수 있도록 설계되었다. 사용자가 특정 상황을 선택하면 해당 상황에 맞게 대화를 시작할 수 있다. 사용자의 발화를 인식하여 자동으로 음성 인식을 중단하고 텍스트로 변환한다. 사전에 생성형 AI에게 제공한 프롬프트를 기반으로 생성형 AI는 부적절한 의사 표현, 상황에 맞지 않는 의사소통을 분석하여 피드백을 즉각적으로 제공한다. 상황별 대화도 음성 대화 시스템을 기반으로 구현되어 사용자와 생성형 AI는 음성을 사용하여 대화를 진행하고 피드백을 전달한다.

#### IV. 결 론

본 연구에서는 다문화 가정, 난민, 교육 소외계층의 아동들의 한국어 학습 부족 문제를 해결하고자 또래 음성 TTS와 생성형 AI를 활용한 한국어 학습 챗봇을 구현하였다. 또래 음성을 활용하여 아동들에게 친근하고 편안한 학습 환경을 제공하고 발음 교정 및 언어적, 문화적 피드백을 통해 맞춤형 학습이 가능하게 하였다. 사용자는 자유로운 대화 속에서 다양한 대화 맥락과 상황을 경험하고 사회적 상호작용의 기본 원리를 깨달아 적절한 반응과 표현을 학습할 수 있다. 또래와의 대화처럼 자연스

럽게 주고받는 음성 대화 시스템을 통해 타인의 감정을 이해하고 공감하는 능력을 키워 의사소통 기술을 점진적으로 발전할 수 있도록 한다. 이는 사용자가 실제 생활에서도 자신감을 가지고 효과적으로 소통할 수 있는 능력을 갖추는 것에 도움을 줄 수 있을 것이다.

추후 지원 언어를 확장하여 다양한 국적의 다문화 아동을 대상으로 한 학습 시스템을 개발한다. 양질의 학습 데이터를 확보하고 지속적인 콘텐츠 업데이트를 통해 사용자 연령층을 확대하고 사용대상을 넓혀 다문화 아동뿐만 아니라 외국인 노동자, 또래와의 소통이 필요한 아동들 등과 같은 사용자에게 보급하고자 한다.

#### 참 고 문 헌

- [1] 룡운, 김태연, "남녀 유아의 또래놀이상호작용과 사회적 유능성의 관계에서 언어발달의 매개효과", *Korean Journal of Child Care and Education Policy*, Vol. 16, No. 2, pp. 131-155, 2022.
- [2] M. J. U. Chowdhury, and A. Hussan, "A review-based study on different Text-to-Speech technologies", arXiv:2312.11563, 2023.
- [3] M. R. Hasanabadi, "An overview of text-to-speech systems and media applications", arXiv:2310.14301, 2023.
- [4] EBA Working Group on Electronic Alternative Payments, *Understanding the business relevance of Open APIs and Open Banking for banks*, 2016.

# HoughLine 그룹화 기반 차선 감지 및 주행 제안 HUD 시스템

김현종, 지동준, 공병용\*

## HoughLine Grouping-Based Lane Detection and Driving Suggestion HUD System

Hyeon Jong, Kim, Dong Jun, Ji, and Byeong Yong, Kong\*

### 요약

최근 운전인구가 증가함에 따라 교통법규 미준수로 인한 안전사고 사례가 늘어나고 있다. 특히 도로의 차선 종류에 따른 교통법규를 지키지 않아 발생하는 사고에 주목하여 새로운 차선인식 시스템을 제안한다. 이 시스템은 영상 전처리, HoughLine 및 그룹화를 이용해 교통법규에 해당하는 다양한 차선의 종류를 인식하고 분류한다. 또한 주행이벤트에 따른 적절한 차선 주행 방향성을 head-up display (HUD)를 통해 운전자가 실시간으로 인지할 수 있게 하여 주행 안정성을 높이는 보조시스템을 제안한다.

### Abstract

Recently, as the number of drivers increases, the number of safety accidents due to noncompliance with traffic laws is increasing. In particular, we propose a new lane recognition system by focusing on accidents caused by non-compliance with traffic laws according to the type of road lane. This system recognizes and classifies various types of lanes corresponding to traffic laws using image preprocessing, HoughLine, and grouping. In addition, we propose an auxiliary system that enhances driving stability by allowing drivers to recognize the appropriate lane driving direction according to driving events in real time through the head-up display (HUD).

### Key words

Advanced driver assistance system (ADAS), Hough line transform, lane recognition system, vehicle safety technology

## 1. 서론

현재 사용되는 차선 인식 시스템은 점선과 실선을 인식하여 단순 차선 변경 안내에 그치는 한계를 가지고 있어 실제 주행상황에서 다양한 교통법규를 반영하기 어렵다. 특히, 주정차선 및 차로변경 교통

사고 위험을 줄이기 위한 고도화된 차선 인식 시스템이 필요하다. 제안하는 차선 인식 시스템은 다양한 차선 종류를 인식하고 분류한다. 이를 바탕으로 주행이벤트에 따른 주행 방향성을 인식된 차선에 따라서 제안하며, head-up display (HUD) 형식으로 운전자에게 직관적으로 도로 정보를 파악할 수 있

\* 국립공주대학교 전기전자제어공학부 전자공학전공, bykong@kongju.ac.kr(교신저자)

※ 이 논문은 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1I1A3064200).

도록 보조한다.

본 논문은 영상 처리와 그룹화 기법을 활용해 차선 인식을 구현하고, 주행이벤트에 따른 주행제한 기능을 보완하였으며, 이를 기반으로 HUD 시스템의 구성을 제시한다.

## II. HoughLineP 그룹화 및 차선인식

차선 인식을 위해 직선 검출 기법으로 영상처리 과정이 필요하다. 본 논문에서는 C++ 환경의 OpenCV 기반 HoughLineP를 이용해 직선을 검출한다. 차선을 정확하게 인식하기 위해 직선 그룹화를 구현한다 [1]. 그림 1의 HoughLineP Transform에서 검출된 직선을 그룹화한 Calculate LineGroups 과정을 정리하였다.

```
#include OpenCV, "LineGroups.h"
HoughLineP Transform
Calculate LineGroups

ColorLaneType(LineGroups, Pixel_Color)
Pixel Color Type Calculate

SingleLaneType(LineGroups, Pixel_Size)
Dashed pixel Threshold
SingleLane Process

DoubleLaneType(ROI, LineGroups, SingleLaneType)
Left Right Lane Groups Count

DrawLaneDectection
```

그림 1. 차선 종류 인식 및 분류 코드  
Fig. 1. Lane type recognition and classification code

그룹화 과정에서 선의 각도 조건과 거리 조건을 구성하여 그룹화 여부를 판단하도록 설정하였다.

식 (1), (2), (3)처럼 검출된 선의 정보를 바탕으로 중간점과 각도를 계산하여 평균 각도를 계산한다.

$$mid = \left( \frac{x_1 + x_2}{2}, \frac{y_2 + y_1}{2} \right), Cal Gradi = \frac{y_2 - y_1}{x_2 - x_1} \quad (1)$$

$$Cal Angle = \tan^{-1} \left( \frac{y_2 - y_1}{x_2 - x_1} \right) \times \frac{180}{\pi} \quad (2)$$

$$Avg Angle = \frac{1}{n} \sum_{i=1}^n \theta_i \quad (3)$$

그 다음 식 (4), (5)처럼 선의 각도와 전체 평균

각도의 차이를 계산한 후 설정한 ThresHoldAngle 값을 비교하여 첫 번째 그룹화 여부를 판단한다.

$$\Delta \theta = |\theta_{addAngle} - Avg Angle| \quad (4)$$

$$\Delta \theta \leq Thres Hold Angle \quad (5)$$

새로운 선의 좌표와 그룹화 할 선의 좌표를 식 (6), (7), (8)처럼 시작점과 끝점의 거리 차이를 유클리드 거리를 이용하여 계산하고 d처럼 직선의 거리를 계산한다. 설정한 ThresHoldDis 조건을 비교하여 두 번째 그룹화 여부를 판단한다. 두 가지 그룹화 여부가 일치할 경우 정확한 차선이 인식되도록 구현하였다.

$$dis Start = \sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2} \quad (6)$$

$$dis End = \sqrt{(x_4 - x_2)^2 + (y_4 - y_2)^2} \quad (7)$$

$$d = dis Start + dis End, d \leq Thres Hold Dis \quad (8)$$

인식된 차선의 픽셀 색상을 감지해 차선의 색상을 판단한다. 직선의 픽셀 길이를 점선의 픽셀 임계값과 비교하여 차선의 직선과 점선을 구분하도록 구현하였다. 영상의 중심을 기준으로 좌우 차선 그룹을 구분하여 각 차선 중에 실선이 3개 이상이면 이중 실선으로 인식한다. 차선 종류를 인식 및 분류하는 과정을 간단하게 그림 1의 코드로 작성했다.

그림 2처럼 실제 주행환경에 차선의 종류가 감지되어 좌우 차선 그룹의 개수, 차선의 종류, 색상, 주행 정보까지 화면에 표시하도록 구현하였다 [2].



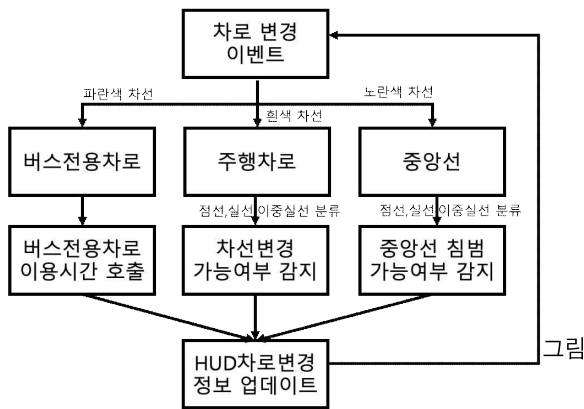
그림 2. 그룹화 기반 차선인식 결과  
Fig. 2. Grouping-based lane detection result

### III. 차선 종류에 따른 주행제한 기능

#### 3.1 차로변경 이벤트

주행이벤트는 차로변경 상황, 주정차선 상황을 감지하는 것이 필요하다.

먼저 차로변경 이벤트는 그룹화된 차선 인식을 이용하며, 방향 지시등의 방향에 따라 차선의 종류를 인식한다. 그림 3과 같이 차선은 색상에 따라 분류된다. 파란색 차선은 버스 전용차로로, 흰색 차선은 주행차로로, 노란색 차선은 중앙선으로 구분한다. 버스 전용차로에서 일반 차량이 접근할 수 있는 시간대를 감지하여 이용 가능 여부를 안내한다. 주행차로와 중앙선은 점선, 실선, 이중 실선의 종류에 따라 접근 가능, 접근 불가, 절대 불가로 분류한다.



3. 차로변경 이벤트 흐름도  
Fig. 3. Flowchart of lane change event

#### 3.2 주정차선 이벤트

주정차선 이벤트는 자동차 변속기가 주차 모드로 작동할 때, 그룹화된 우측 차선의 종류를 감지하여 그림 4와 같이 안내 정보를 업데이트한다. 흰색 실선이 감지되면 주정차로로 인식하여 주정차가 가능함을 안내한다. 노란색 점선이 인식되면 HUD에 5분 타이머가 업데이트되어 작동하며, 이는 임시 주정차 구역임을 알린다. 타이머가 종료되었을 때는 해당 구역을 벗어나라는 안내가 제공된다.

또한, 노란색 실선이 인식되면 탄력적 주정차 구역의 운영 시간을 인식하여 이용 가능 여부를 안내한다.

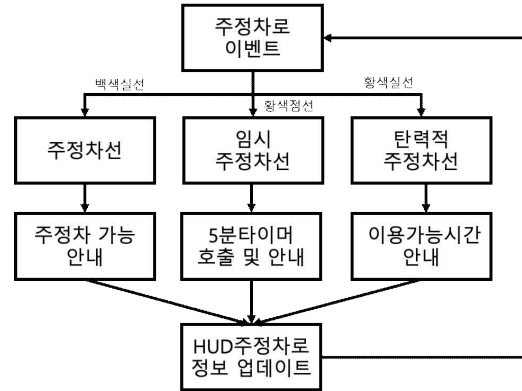


그림 4. 주정차구역 이벤트 흐름도  
Fig. 4. Flowchart of parking zone event

### IV. Head-Up Display (HUD) 시스템

#### 4.1 HUD 인터페이스 구현

주행 정보를 HUD 형태로 실시간으로 제공하여 운전자가 쉽게 인지할 수 있는 인터페이스가 필요하다 [3]. 그림 5는 Qt 환경 기반의 그래픽 인터페이스를 구현한 것이다. 이미지와 텍스트 라벨을 이용해 인터페이스를 구성했으며 차선 주행 정보 안내, 임시 주정차 시간 안내를 실시간으로 안내하도록 구현하였다.

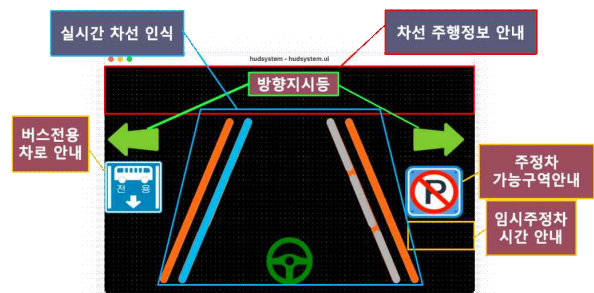


그림 5. Qt 기반 HUD 인터페이스 구성  
Fig. 5. Qt-based HUD interface configuration

#### 4.2 HUD 하드웨어

구현된 인터페이스를 활용하여 모듈 형태의 하드웨어를 구성하는 것이 필요하다. 그림 6처럼 모듈형 하드웨어 구현을 위해 Raspberry Pi 5를 기반으로 제작하였으며, 최적화된 차선 인식 처리를 위해 주행 영상을 720p 해상도와 23fps로 처리한다.

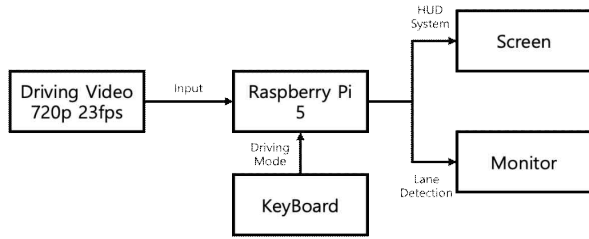


그림 6. HUD 하드웨어 구성도  
Fig. 6. HUD Hardware Configuration Diagram

키보드 입력을 통해 주행 모드를 인식하고, 이를 기반으로 주행이벤트 처리를 수행하도록 설계하였다. 스크린은 HUD 인터페이스를 업데이트하며, 모니터는 실시간 차선 인식 과정을 확인할 수 있도록 구성하였다.

HUD 하드웨어는 그림 7과 같이 설계하였고 그림 8처럼 스크린이 반사될 수 있도록 투명 아크릴 판과 반사 필름을 활용하여 시인성과 가시성을 증가시켰다.

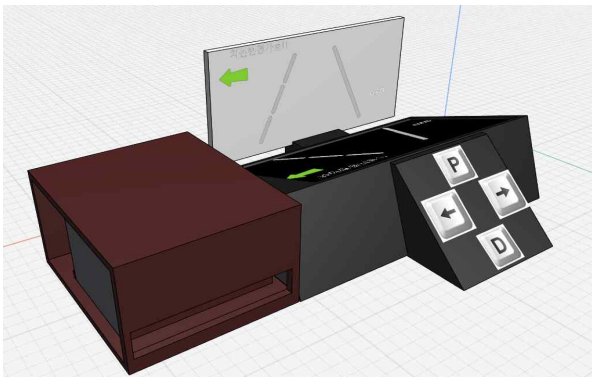


그림 7. 3D 설계 HUD 하드웨어  
Fig. 7. 3D-designed HUD hardware



그림 8. HUD 하드웨어 구현  
Fig. 8. Implementation of HUD hardware

## V. 결론 및 활용 방안

고령 운전자와 초보운전자의 운전 미숙으로 인하여 교통사고를 예방하고자 개발하였다. 이 시스템을 기반으로 차량 제어 분야로 발전시키는 경우, 향상된 주행 보조 시스템으로의 발전 가능성이 크다. 모듈화 시스템 구현으로 구형 차량에도 적용할 수 있어, 불필요한 차량 교체를 줄이고 안전사고를 줄일 수 있다.

아울러 최근 증가하고 있는 개인형 이동장치에도 적용할 수 있을 것으로 기대된다. 이러한 점에서 본 시스템은 향후 운전 보조기술의 발전에 이바지할 수 있을 것으로 판단된다. 본연구에 제안한 시스템 기반으로 차량을 제어하는 방안을 추가적으로 연구하고자 한다.

## 참 고 문 헌

- [1] M. Marzougui, A. Alasiry, Y. Kortli, and J. Baili, "A lane tracking method based on progressive probabilistic Hough transform," *IEEE Access*, vol. 8, pp. 84893-84905, 2020.
- [2] N. J. Zakaria, M. I. Shapiai, R. A. Ghani, M. N. M. Yassin, M. Z. Ibrahim, and N. Wahid, "Lane detection in autonomous vehicles: A systematic review," *IEEE Access*, vol. 11, pp. 1923-1937, 2023.
- [3] J. Korentsides, A. S. Clair, M. Dyapa, and A. Chaparro, "Head-up displays: Analysis of automotive use considerations," *Human Factors Assist. Technol.*, vol. 37, pp. 53-59, 2022.

# 자율주행 환경 인식을 위한 특화 데이터 증강과 Vision Transformer 기반 이미지 분류

정소미\*, 김남훈\*\*, 이선우\*\*, 신종현\*\*\*, 윤현중\*\*\*\*

## Vision Transformer-based Image Classification with Specialized Data Augmentation for Autonomous Driving Scene Recognition

So-Mi Jeong\*, Nam-Hun Kim\*\*, Sun-Woo Lee\*\*, Jong-Hyeon Sin\*\*\*, and Hyun-Joong Yoon\*\*\*\*

### 요약

본 연구는 자율주행 환경에서의 다중 모달 센서 데이터 분류를 위한 Vision Transformer (ViT) 모델의 성능 향상에 초점을 맞추어, 도메인 특화 데이터 증강 기법의 효과를 분석하였다. 극한 기상 조건과 다양한 주행 환경을 고려한 맞춤형 증강 알고리즘을 제안하고, 이를 ViT 모델에 적용하여 기존 모델과의 성능을 비교하였다. 본 연구는 ViT 모델과 특화된 데이터 증강 기법의 통합적 접근을 통해 자율주행 시스템의 환경 인지 능력 향상 및 다양한 주행 조건에서의 안전성과 신뢰성 제고를 위한 기술적 기반을 제시한다.

### Abstract

This study explores the effect of domain-specific data augmentation techniques on Vision Transformer (ViT) models for classifying multi-modal sensor data in autonomous driving. We introduce an augmentation algorithm that addresses extreme weather and diverse driving environments, applying it to the ViT model for performance evaluation against a baseline. Results indicate that the proposed model achieves lower validation loss and more stable early-stage convergence than traditional methods, with improved recognition in challenging weather and nighttime scenarios. This research enhances environmental perception in autonomous driving systems and lays a foundation for improved safety and reliability. Future work will expand the dataset to include more environmental conditions and optimize the model for lightweight use.

### Key words

Autonomous Driving, Vision Transformer, Data Augmentation, Multi-modal Sensor Data, Image Classification, Environmental Perception, Extreme Weather Conditions, Deep Learning

\* 국립안동대학교 컴퓨터공학과, someee0706@naver.com

\*\* 국립금오공과대학교 기계시스템공학부, gnsskazla@naver.com, amway1213@naver.com

\*\*\* 계명대학교 기계공학과, sd499@naver.com

\*\*\*\* 대구가톨릭대학교 기계공학과, yoon@cu.ac.kr(교신저자)

※ 본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신사업의 결과입니다.(2022RIS-006)



## 1. 서론

자율주행 기술의 급속한 발전과 함께, 다양한 도로 환경에서의 정확하고 신뢰성 있는 인지 능력의 중요성이 대두되고 있다[1]. 특히 광학 카메라, 적외선 센서, 깊이 센서 등 다중 모달리티 센서에서 획득되는 데이터의 효과적인 처리 및 해석은 자율주행 시스템의 성능을 결정짓는 핵심 요소이다[3]. 그러나 실제 도로 환경에서 직면하는 극한 기상 조건, 급격한 조명 변화, 센서 오류 등으로 인한 데이터 품질 저하는 인지 시스템의 신뢰성에 중대한 도전과제를 제기하고 있다[4]. 이러한 문제의 해결을 위해서는 자율주행 환경에 특화된 데이터 증강 기법과 같은 체계적인 방법론적 접근이 요구된다[2].

본 연구에서는 자율주행 환경에 최적화된 데이터 증강 기법을 제안하고, 이를 최신 비전 모델인 Vision Transformer (ViT)와 통합하여 다중 모달 센서 데이터의 분류 성능을 향상시키고자 한다[1][2]. 특히, 극한 기상 조건과 센서 오류 상황에서의 모델 강건성 향상에 주안점을 두어, 실제 자율주행 환경에서의 실용성을 제고하고자 한다[4].

## II. 관련연구

### 2.1 Vision Transformer (ViT)

Vision Transformer는 자연어 처리 분야에서 검증된 Transformer 아키텍처를 컴퓨터 비전 도메인에 효과적으로 접목한 아키텍처이다[5]. 기존의 CNN 기반 모델들과 달리, ViT는 이미지를 일정 크기의 패치로 분할하고 이를 시퀀스로 변환하여 처리함으로써, 전역적 특징 추출과 장거리 의존성 포착에서 우수한 성능을 나타낸다[5]. 특히, 자기 주의 메커니즘을 통해 이미지의 각 영역 간 관계를 효과적으로 모델링한다는 특징을 보인다[5].

본 연구에서는 Google이 제안한 'ViT-Base' 사전 학습 모델을 기반으로 자율주행 환경에 최적화된 분류 모델을 구축하였다. 제안된 모델은 입력 영상을 224×224 화소 크기로 정규화한 후, 이를 다시 16×16 화소 단위의 패치로 분할하여 처리한다[5]. 분할된 각 패치는 768차원의 특징 공간으로 매핑되며, 12개의 변환기 인코더 계층(Transformer Encoder Lay

ers)과 12개의 자기 집중 기제(Self-Attention Mechanisms)를 통해 심층 특징을 추출한다[5]. 이러한 계층적 구조는 자율주행 환경의 다양한 시각적 패턴과 특징들을 효과적으로 포착하며, 특히 장기 의존성(Long-range Dependencies)과 전역적 문맥 정보(Global Contextual Information)의 효율적 학습을 가능하게 한다[5][6]. 그림 1은 Vision Transformer (ViT)의 구조도이다.

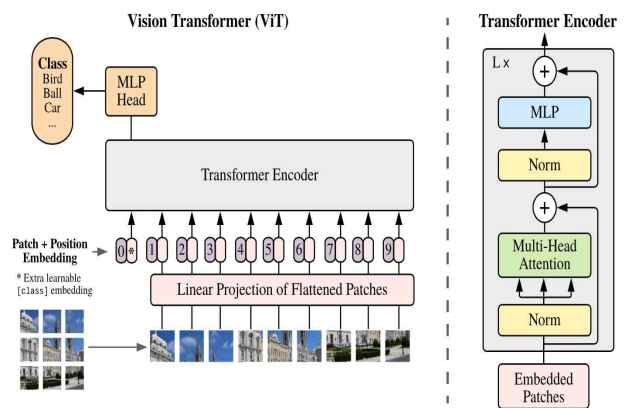


그림 1. Vision Transformer (ViT) 구조도  
Fig. 1. Architecture of Vision Transformer (ViT)

본 연구의 특수성을 고려하여, 모델의 최종 출력층을 자율주행 환경에서 획득되는 다중 센서 데이터의 분류를 위한 3개의 클래스(적외선(IR), 깊이(Depth), 일반(General))에 맞게 재구성하였다[6]. 전이 학습의 효율성을 최대화하기 위해 사전 학습된 모델의 특징 추출 계층의 가중치는 유지하되, 분류기 계층만을 새로 초기화하여 학습을 수행하였다[7]. 이러한 접근은 제한적인 데이터셋에서도 효과적인 학습이 가능하게 하며, 동시에 자율주행 환경에서 획득되는 다양한 센서 데이터의 특성을 잘 반영할 수 있도록 한다.

### 2.2 데이터 증강

데이터 증강은 모델의 일반화 능력을 향상시키고 과적합을 방지하는 학습 전략이다[8]. 특히 자율주행 시스템에서는 다양한 환경 변수에 대한 강건성이 안전성과 직결되므로, 체계적인 데이터 증강 전략의 수립이 필수적이다[8]. 본 연구에서는 고성능 이미지 증강 라이브러리인 Albumentations를 활용하

여 자율주행 특화 데이터 증강 파이프라인을 구축하였다. Albuementations는 최적화된 구현과 다중 처리 파이프라인을 통해 기존 증강 라이브러리 대비 약 5-10배 향상된 처리 효율성을 제공하며[8], 복잡한 증강 기법들의 조합을 확률적으로 적용할 수 있는 유연성을 지니고 있다.

제안하는 데이터 증강 파이프라인은 환경 조건 시뮬레이션과 센서 및 주행 환경 모델링이라는 두 가지 핵심 축을 중심으로 구성되었다. 먼저 환경 조건 시뮬레이션에서는 안개(fog\_coef∈[0.1,0.3]), 강우(drop\_length=20, drop\_width=1), 적설(brightness\_coef=2.5) 등의 다양한 기상 조건을 사실적으로 모사하였으며, 조명 변화 모델링을 위해 밝기 및 대비 조정(brightness\_limit=±0.2, contrast\_limit=±0.2), 그림자 생성(num\_shadows∈[1,3]) 등의 기법을 활용하였다. 각 증강 기법의 파라미터는 실제 기상 데이터와 도로 주행 영상 분석을 통해 실증적으로 최적화되었다.



그림 2. 자율주행 특화 데이터 증강 기법 적용 예시  
Fig. 2. Examples of Specialized Data Augmentation Techniques for Autonomous Driving

그림 2는 제안된 데이터 증강 파이프라인을 통해 생성된 증강 영상의 예시를 보여준다. 안개, 강우, 적설 등 다양한 기상 조건하에서의 도로 주행 장면이 사실적으로 시뮬레이션 되었음을 확인할 수 있다. 이는 자율주행 시스템이 실제 환경에서 접하게 될 다양한 변화 요인들에 대한 강건성을 효과적으로 학습할 수 있게 해준다[9].

그림 3은 실제 자율주행 환경에서 발생 가능한 다양한 센서 및 주행 조건을 모델링한 결과이다. 저조도(Low Light), 높은 주행 속도(High Speed), 센서 간섭(Sensor Interference) 등의 영향으로 인해 영상의 품질이 저하되는 현상을 사실적으로 재현하였으며, 안개와 노이즈가 동시에 발생하는 복합적인 상황(Complex Situation)까지 고려하였다. 뿐만 아니라 주행 속도에 따른 모션 블러 효과와 센서 신호 대 잡음비(SNR)에 기반한 노이즈 특성을 중점적으로 모사하였는데, 이를 위해 저속(30km/h) 및 고속(80km/h) 주행 시의 모션 블러 정도를 차별적으로 적용하고, 센서 노이즈를 가우시안 노이즈(var\_limit∈[10.0,20.0], [40.0,50.0])를 통해 단계적으로 구현하였다.

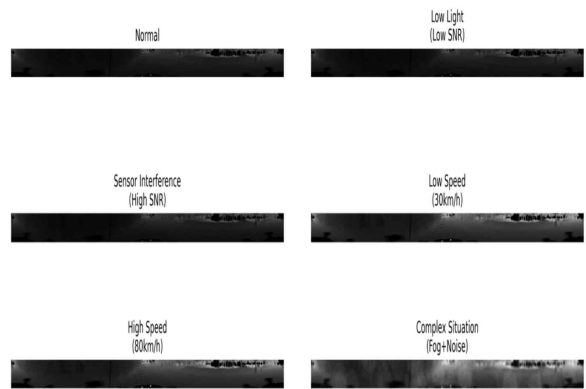


그림 3. 센서 및 주행 환경 시뮬레이션 결과  
Fig. 3. Simulation Results of Sensor Characteristics and Driving Conditions

제안된 데이터 증강 파이프라인은 가혹 조건 하에서도 안정적 동작이 가능한 자율주행 시스템 개발에 기여할 것으로 기대된다. 이는 센서 신호 대 잡음비 분석에 기반하여 가우시안 노이즈(var\_limit ∈ [10.0, 50.0])를 주입하고, 차량 동역학 모델을 고려한 모션 블러(blur\_limit = 7) 효과를 구현함으로써, 센서 장치의 물리적 한계 및 차량 주행 상황이 유발하는 영상 왜곡을 사실적으로 재현하였기 때문이다.

더불어, 베이지안 최적화 알고리즘을 활용하여 증강 기법 간 적용 비율을 최적 조합(각 기법별 확률 0.2)으로 설정하였다. 이를 통해 과도한 변형으로 인한 원본 데이터 손실을 방지하면서도 충분한 다양성 확보할 수 있도록 하였다[10]. 결과적으로

학습 과정에 투입되는 영상이 실제 주행 환경에서 발생하는 다양한 변화 양상을 균형 있게 내포할 수 있게 되었다[10].

### III. 실험

#### 3.1 데이터셋

본 연구의 데이터셋은 자율주행용 다중 모달 영상 데이터셋으로, 표 1과 그림 4는 카테고리별 분포를 보여준다.

표 1. 자율주행 다중 모달 영상 데이터셋 구성  
Table 1. Composition of Multi-modal Image Dataset for Autonomous Driving

	깊이	광학	적외선	합계
학습	5,035	3,572	4,095	12,702
검증	1,339	919	1,037	3,295
테스트	394	308	329	1,031
합계	6,768	4,799	5,461	17,028

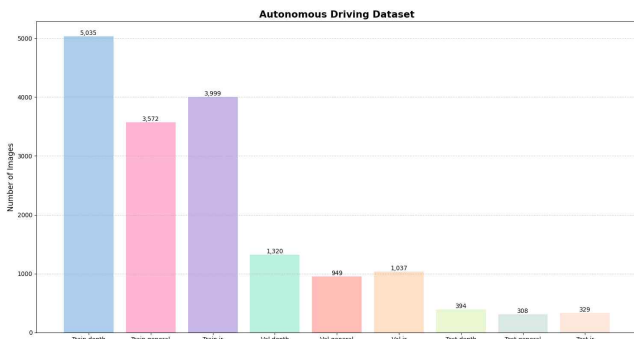


그림 4. 자율주행 데이터셋의 카테고리별 이미지 분포  
Fig. 4. Autonomous Driving Dataset: Image Distribution by Category

#### 3.2 실험환경

본 연구의 실험은 표 2의 시스템 환경과 표 3의 학습 파라미터로 구성하여 수행하였다.

표 2. 실험 시스템 환경  
Table 2. System environment

구분	값
CPU	Intel Xeon 24 Cores / 2.1Ghz
GPU	NVIDIA A100
운영체제	Ubuntu 20.04
프로그래밍 언어	Python

표 3. 학습 파라미터  
Table 3. Training Parameters

항목	설정 값
배치크기/학습횟수	64 / 50epochs
학습률	3e-5
최적화 알고리즘	RAdam
손실함수	CrossEntropyLoss

또한, 학습 효율성을 높이기 위해 혼합 정밀도 학습 (Mixed Precision Training)을 적용하였다. 이 방식은 32비트와 16비트 부동소수점 연산을 혼합하여 메모리 사용량을 줄이고 학습 속도를 향상시킨다.

#### 3.3 실험결과

제안된 데이터 증강 기법의 효과를 검증하기 위해, 증강 적용 모델과 미적용 모델의 성능을 비교하였다. 그림 5는 학습 과정에서 손실 함수(Loss) 변화와 검증 정확도를 보여준다.

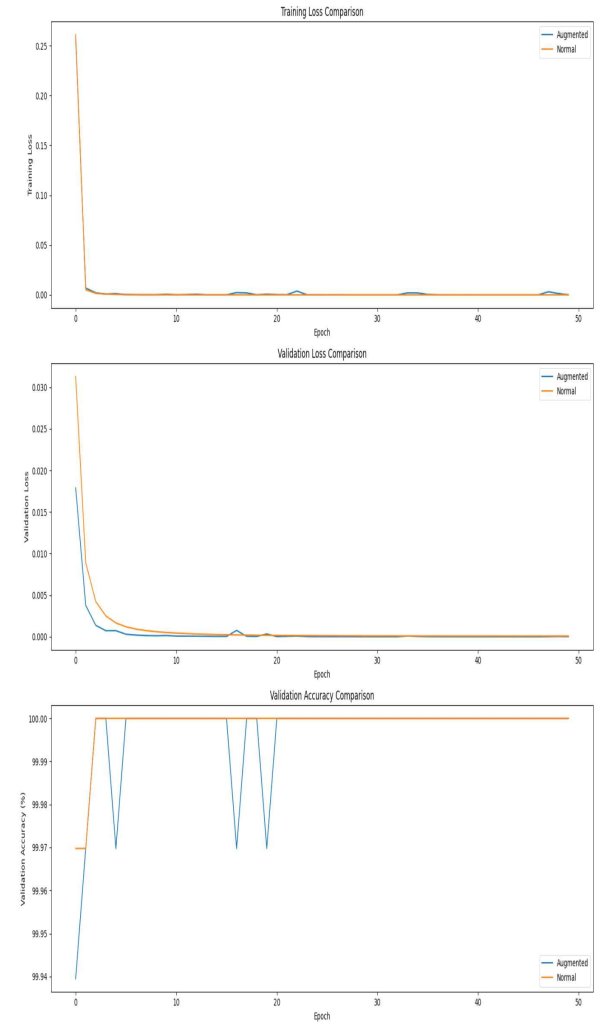


그림 5. 학습 중 손실 함수 변화와 검증 정확도 비교  
Fig. 5. Loss function changes during training vs. validation accuracy

실험 결과, 증강 모델이 기본 모델에 비해 더 안정적인 학습 곡선을 보였으며, 특히 초기 학습 단계에서 더 빠른 수렴 속도를 나타냈다. 검증 정확도 측면에서는 두 모델 모두 99.9% 이상의 높은 성능을 달성했으나, 증강 모델이 학습 과정 전반에 걸쳐 더 일관된 성능을 유지하는 것으로 관찰되었다. 이는 제안된 데이터 증강 기법이 모델의 일반화 능력을 향상시키고, 과적합을 효과적으로 방지하고 있음을 시사한다.

#### IV. 결론

본 연구에서는 Vision Transformer와 자율주행 특화 데이터 증강 기법을 결합하여 다중 센서 자율주행 영상 분류 모델을 제안하였다. 실험을 통한 정량적 평가 결과, 제안된 증강 모델은 기본 모델과 동일한 수준의 높은 정확도를 달성하면서도, 초기 학습 단계에서 더 낮은 검증 손실값(0.017 대비 0.031)과 안정적인 수렴 특성을 보였다. 특히 다중 센서 데이터에 대한 증강 기법의 체계적 적용이 모델의 견고성 향상에 크게 기여했음을 확인하였다. 이러한 결과는 제안된 데이터 증강 기법이 자율주행 시스템의 인지 성능 향상에 실질적으로 기여할 수 있음을 시사한다.

그러나 본 연구는 데이터의 다양성 부족, 극단적 기상 조건 시뮬레이션의 한계, 그리고 모델의 높은 계산 복잡도라는 한계점을 가지고 있다. 향후 연구에서는 이러한 한계를 극복하기 위해 다양한 환경 조건에서의 데이터셋 구축, 센서 간 상호 보완적 특성을 활용한 모델 구조 최적화, 그리고 경량화 모델 개발을 진행할 예정이다.

#### 참 고 문 헌

[1] La, Q. V., et al. "Exploring Vision Transformers in Autonomous Driving: Trends and Applications." arXiv, 2024.  
[2] Waymo. "Using automated data augmentation to advance our Waymo Driver." Waymo Blog, 2020.  
[3] RGBSI. "What is Sensor Fusion for Autonomous

Driving Systems? - Part 1." RGBSI Blog, 2023.  
[4] 박설민. "[르포] 자율주행 최대의 적 '악천후', 해답은 연천에 있다." 시사위크, 2023.  
[5] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.  
[6] La-Dang, Q. V., Nguyen-Mau, T., & Nguyen-Dinh, K. (2024). A Survey of Vision Transformers in Autonomous Driving: Current Trends and Future Directions. arXiv preprint arXiv:2403.07542.  
[7] Nakamura, T., Harada, T., & Kataoka, H. (2023). Pre-training Vision Transformers with Very Limited Synthesized Images. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 19542-19551).  
[8] <https://aws.amazon.com/what-is/data-augmentation/>  
[9] <https://graphite-note.com/the-importance-of-data-augmentation-in-ml/>  
[10] <https://deepdrive.berkeley.edu/project/training-data-augmentation-autonomous-driving>

# GPT 모델 함수 호출 정확도 향상에 대한 연구

이하린\*, 최영림\*\*, 김현석\*\*\*

## Function Calling Accuracy Optimization in GPT Models

Harin Lee\*, Younglim Choi\*\*, and Hyunseok Kim\*\*\*

### 요약

GPT 모델은 외부 도구 및 시스템과 연결할 수 있는 함수 호출 기능을 제공한다. 이를 통해 특정 상황에서 GPT 모델이 데이터를 자동으로 가져오거나 워크플로우를 관리할 수 있다. 이 기능의 핵심은 질문의 컨텍스트를 정확히 이해하고 적절한 함수를 선택하는 것이지만, 질문이 복잡하거나 중의적일 경우 잘못된 함수를 호출할 수 있는 한계가 존재한다. 본 논문에서는 이러한 문제를 해결하기 위해 GPT 모델에서 함수 호출 정확도를 향상시키는 방법을 분석하고, 최적의 방안을 제시한다.

### Abstract

The GPT model offers a function-calling feature that enables integration with external tools and systems, allowing it to automatically retrieve data or manage workflows in specific contexts. The core of this functionality lies in accurately understanding the context of the query and selecting the appropriate function. However, when queries become complex or ambiguous, the model may call incorrect functions due to inherent limitations. This paper analyzes methods to improve the accuracy of function calls GPT models and proposes an optimal solution to address this issue.

### Key words

function calling feature, external tools, context understanding, function all accuraceny, optimal solution

## 1. 서론

LLM(대규모 언어 모델)은 종종 그럴듯해 보이지만 사실이 아닌 정보를 예측하는데, 이를 Hallucination(환각)이라고 한다. [1][2] 이는 주로 단

일 매개 변수 모델과 제한된 컨텍스트에서 통계적 언어 모델링을 수행하도록 훈련된 LLM의 근본적인 한계에서 비롯된다. 이러한 문제는 사용자에게 잘못된 정보 전달 및 의사 결정 오류로 이어질 수 있어 중요한 문제로 인식된다.[3] 이를 해결하기 위해 등

---

\* 동아대학교 컴퓨터공학과 pqud7534@naver.com,  
\*\* 동아대학교 SW혁신센터, cyl88@dau.ac.kr(공동교신저자),  
\*\*\* 동아대학교 컴퓨터공학과, hertzkim@dau.ac.kr(공동교신저자)

※ 본 연구성과물은 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No. 과제 관리 번호 : RS-2023-00242673)과 2023년도부터 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원 (No.2023-0-00076, SW중심대학(동아대학교))된 연구임

장한 모델이 Augmented Language Model(ALM)[4]이다. ALM은 LLM에 외부 도구를 결합하여, 모델의 가중치에 포함되지 않은 추가적인 정보를 활용해 현재의 맥락을 보강한다 [5][6].

GPT 모델은 이러한 ALM 개념의 연장선에서, 외부 도구 및 시스템과의 연결을 통해 모델의 기능을 확장하는 함수 호출 기능을 제공한다. [7] 이를 통해 AI 어시스턴트에 다양한 기능을 부여하거나 애플리케이션과의 통합을 구축할 수 있다. 이 기능의 핵심은 모델이 질문의 컨텍스트를 정확히 이해하고 적절한 함수를 호출하는 것이다. 그러나 질문이 복잡하거나 중의적일 경우, 모델이 함수를 호출하지 않거나 잘못된 함수를 호출하는 문제가 발생할 수 있다.[8][9] 본 연구에서는 이러한 문제를 해결하기 위해 다양한 접근 방식을 시도하고 평가하여, 함수 호출 정확도를 개선하기 위한 최선의 방법을 제안하고자 한다.

성능 분석을 위하여 함수 호출 기능을 지원하는 최신 모델 gpt-4o-2024-08-06을 사용한다.

## II. 실험 설정

표 1. 실험 데이터 세트 예시 1

Table 1. Experimental data sets 1

Item	value
prompt	파워블로거 김슬기씨는 블로그에 포스트를 작성하려고 합니다. 작성하기 전, 포스트를 하나로 묶을 게시판을 생성하기로 했습니다. Campaigns 2024 게시판에 작성한 포스트는 누구나 보고, 댓글을 달고, 좋아요를 남기고, 공유할 수 있도록 하려고 합니다. 김슬기씨의 아이디는 cm789입니다.
solution	create_space(user_id=cm789, name=Campaigns 2023, visibility=public, features=(comments_enabled=True, reactions_enabled=True, shareable=True))

접근 방식을 평가하기 위하여 데이터 세트를 준비하였다. 데이터 세트는 프롬프트와 그에 따른 해결책으로 구성된 60쌍의 문제 세트로 구성된다. 모델은 프롬프트를 받고 하나의 매개변수를 가지는 간단한 함수부터 최대 11개의 매개변수를 필요로

하는 8개의 함수 중 해당하는 함수를 찾아야 한다. 데이터 세트에는 함수를 호출하지 않는 데이터도 존재하며, 이 경우 solution은 pass가 된다.

표 2. 실험 데이터 세트 예시 2

Table 2. Experimental data sets 2

Item	value
prompt	최근에 작성한 Design Studio 2023 관련 글을 다들 괜찮다고 생각하는 것 같아요. 기쁘니다.
solution	pass

## III. 실험 진행

### 1. Measuring Baseline Performance

우선 시스템 프롬프트 없이 독립적으로 GPT-4o의 함수 호출 성능을 평가하였다. 시스템 프롬프트 없이 동일한 데이터셋으로 3회 실험을 실행하여, 데이터셋과 AI의 solution 일치 횟수를 데이터셋 길이로 나눈 값을 사용하여 평균 정확도를 측정된 결과, 각각 0.28, 0.3, 0.28로 평균 0.286의 정확도를 얻었다.

### 2. Flattening Parameters

함수의 중첩된 객체와 서브 속성이 GPT의 함수 호출 정확도에 미치는 영향을 줄이기 위해 매개변수 평탄화를 수행하였다. 예를 들어, 기존에 create\_space의 features 객체 안에 있는 comments\_enabled, shareable과 등의 속성을 각각 독립된 매개변수로 변환하고, 부모 변수 이름에 \_\_를 사용하여 결합하였다. 중첩 구조가 깊을수록 GPT가 매개변수 정보를 제대로 파악하기 힘들어 함수 호출 정확도에 영향을 미치고, 포맷 불일치를 유발할 가능성이 높다고 생각했기 때문이다. 변경된 함수 구조로 3번의 실행 결과는 0.4, 0.383, 0.4로, 평균 정확도 0.394를 기록하였다

### 3. Adding System Prompt

이전 실험까지는 시스템 프롬프트가 없었기 때문에 LLM에게 역할이나 블로그 API와의 상호작용에 대한 설명을 제공하지 않았다. 간단한 시스템 프롬

프트를 추가한 후 함수 호출 기능을 3번 실행한 결과, 0.45, 0.4, 0.45로 평균 정확도는 0.433이며 프롬프트가 없을 때에 비해 정확도가 약간 증가하는 것을 확인할 수 있었다.

#### 4. Improve System Prompt

Adding System Prompt에서 시스템 프롬프트가 정확도 향상에 도움이 된다는 점을 확인하였다. 이에 시스템 프롬프트에 세부 정보를 추가하고 정확도 변화를 관찰하였다. 단계적인 지침과 도구 활용 방법에 대한 설명을 추가하고, 각 도구의 매개변수를 채우기 위해 해당 상황에서 필요한 정보를 추출하는 구체적인 방법을 명시하였다.[10] 예를 들어, 기존의 ‘you must give the solution utilizing your functionalities.’와 같은 간단한 설명을 ‘Second, look into the parameters and extract all the relevant information to fill the parameter with right values.’ 처럼 단계를 나누어 어떤 행동을 해야하는지 구체적으로 명시하였다. 3번의 함수 호출 기능을 실행한 결과, 이전보다 정확도가 15.70% 상승하여 정확도 0.51, 0.533, 0.46을 기록하였다

#### 5. Optimize functions

LLM 역할에 맞게 함수 설명을 개선하고, 명확한 함수 이름을 사용하도록 수정하였다. 예를 들어, 함수의 설명을 ‘Add a new Space to a Workspace.’에서 ‘Creates a new ClickUp space’ 로 변경하고, ‘create\_space’와 같은 함수 명을 ‘create\_a\_new\_posts\_space’ 변경하여 함수의 목적성을 명확히 하였다. 수정 후 3번의 실행 결과, 0.56, 0.533, 0.516으로 평균 정확도는 0.536을 기록하였다.

#### 6. Adding Examples

LLM은 함수 호출에 예시가 제공될 경우 성과가 향상되는 경향이 있어 각 함수 호출과 매개변수에 예시를 추가하였다. 스키마에서 각 함수 호출의 예시와 함수 설명을 추가하고, 각각의 매개변수에도 예시를 추가하였다. 그 결과, 3번의 실험에서 평균 정확도 0.66, 0.63, 0.66을 기록하여, 이전 실험 대비 정확도가 21.27% 상승하였다.

추가로, 데이터셋이 한국어로 되어 있어 프롬프트와 스키마를 한국어로 수정하여 실험을 진행하였으나 오히려 정확도가 떨어지는 경향을 보였다. 이로 인해 데이터셋이 한국어로 작성되어 있더라도, 영어를 사용하여 실험을 진행하는 것이 오히려 더 높은 정확도를 보장함을 확인하였다. 이는 LLM이 특정 언어의 문맥을 이해하는 데 있어 제한이 있을 수 있음을 시사하며, 향후 언어별 데이터셋 최적화가 필요함을 보여준다.

표 3. 실험 결과

Table 3. Experimental Result

Item	Average Accuracy
measuring baseline performance	0.286
Flattening Parameters	0.394
Flattening Parameters + System Prompt	0.433
Flattening Parameters + Improved System Prompt	0.501
Flattening Parameters + Improved System Prompt + Optimize functions	0.536
Flattening Parameters + Improved System Prompt + Optimize functions	0.65

## IV. 결 론

본 연구는 GPT-4 모델을 사용하여 함수 호출의 정확도를 증가시키는 방법에 대해 제시하였다. 본 연구는 GPT 모델의 활용 범위를 넓혀 상호작용성과 실질적인 활용성을 강화하는데 기여 하였으며, 향후 다양한 어플리케이션에서 적용 가능성을 제시한다. 다만, 추가적인 예시 데이터와 복잡한 맥락에서의 함수 호출 정확도 개선 방안에 대한 연구가 필요하다.

## 참 고 문 헌

- [1] 강장묵, 김경민, 조민제, 조동백, 김남선, "자율주행차량 서비스에서 할루시네이션 보안에 대한 소고", 한국정보기술학회 2023년도 하계종합 학술대회 및 대학생논문경진대회 학술대회자료,

pp. 23-26, June 2023.

- [2] 조찬영, 강성준, 정현준, "RAG 기반 랭체인을 이용한 생성형 AI 챗봇 구현", 한국정보기술학회 2023년도 추계종합학술대회 및 대학생논문경진대회 학술대회자료, pp. 460-463, November 2023.
- [3] 김태원, "챗GPT가 촉발한 생성형 AI 시대, 미래 모습과 대응 방안", FUTURE HORIZON, vol.55, pp. 2-9, May, 2023.
- [4] Grégoire Mialon, Roberto Dessi. "Augmented Language Models: a Survey", arXiv:2302.07842., February 2023.
- [5] Timo Schick. (2023). "ToolFormer: Language Models Can Teach Themselves to Use Tools", arXiv:2302.04761., February 2023
- [6] Anish Acharya. "Using Generative AI to Unlock Probabilistic Products", <https://a16z.com/generative-ai-probabilistic-products/> , May 2023.
- [7] OpenAI Official Documentation. "Function calling", <https://platform.openai.com/docs/guides/function-calling>
- [8] OpenAI. "Function calling and other API updates", <https://openai.com/index/function-calling-and-other-api-updates/>, June 2023
- [9] 김민욱, 서용원, 전예찬, & 이원형. "반복적 상호작용을 통한 사용자의 의사전달 모호성 감소 및 언어기반 협동 로봇 팔의 작업 성공률 향상에 관한 연구", 한국정보기술학회 2024년도 하계종합학술대회 및 대학생논문경진대회 학술대회자료, pp. 611-614, June 2023
- [10] 김민근&정유철, "키워드의 도메인 속성을 고려한 고품질 키워드 정제기법", 한국정보기술학회 vol. 22 no. 9, pp. 23-33, September 2024



# 청각장애인을 위한 음성 감정 분석용 채팅 시스템

권건표\*, 홍윤기\*, 이유진\*, 남주영\*, 정기현\*\*

## Chat System with Voice Emotion Recognition for Deaf Users

Geon Pyo Gwon\*, Yun Ki Hong\*, Yu Jin Lee\*, Ju Yeong Nam\*, and Ki Hyun Jung\*\*

### 요약

본 연구는 청각 장애인과 자폐 스펙트럼 장애인을 위한 웹 기반 음성 감정 인식 채팅 시스템을 제안한다. 제안된 시스템은 음성 데이터에서 톤, 높낮이, 강도, 파장 등의 특징을 추출하여 감정을 분석한 후, 실시간으로 텍스트와 함께 감정 상태를 상대방에게 전달한다. 이를 위해 백엔드에서는 Spring Boot JPA와 MySQL을 통해 데이터베이스를 구축하고, STOMP 프로토콜 기반의 pub/sub 구조로 실시간 채팅 기능을 구현하였다. 또한, 프론트엔드는 React로 사용자 인터페이스를 개발하여 사용자가 음성 데이터를 웹 페이지에서 직접 전송하고 실시간으로 분석 결과를 확인할 수 있도록 구성하였다. 감정 분석 서버는 FastAPI와 PyTorch를 기반으로 HuBERT와 Transformer 모델을 결합하여 높은 정확도로 감정 인식을 실현하였다. 본 시스템은 감정 인식 기능을 통해 사용자 간 원활한 소통을 지원하며, 청각 장애인과 자폐 스펙트럼 장애인의 사회적 소외감을 감소시키고 포용적 소통 환경을 제공하는 데 기여할 수 있을 것으로 기대된다.

### Abstract

This study proposes a web-based speech emotion recognition chat system for the hearing impaired and people with autism spectrum disorders. The proposed system extracts features such as tone, pitch, height, intensity and wavelength from speech data, analyses the emotion, and then communicates the emotional state to the other party along with text in real time. At the backend, a database was built using Spring Boot JPA and MySQL, and a real-time chat function was implemented using a pub/sub structure based on the STOMP protocol infrastructure. On the frontend, a user interface was developed in React, allowing users to submit speech data directly from a web page and see the analysis results in real time. The emotion analysis server was based on FastAPI and PyTorch, and combined HuBERT and Transformer models to achieve highly accurate emotion recognition. This system is expected to support smooth communication between users through the emotion recognition function, and contribute to reducing social alienation and providing an inclusive communication environment for people with hearing impairments and autism spectrum disorders.

### Key words

Speech Emotion Recognition, HuBERT-Transformer, STOMP Protocol, Inclusive Communication

---

\* 국립안동대학교 멀티미디어공학과, rjsvy0522@naver.com

\*\* 국립안동대학교 소프트웨어융합학과, kingjung@anu.ac.kr (교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

## I. 서 론

디지털 커뮤니케이션 기술은 대중에게 효율적인 소통 방법을 제공하며, 현대 생활에서 필수적인 요소로 자리 잡고 있다. 특히 스마트폰과 웹 기반의 채팅 시스템은 실시간 소통을 가능하게 하여 시간과 공간의 제약을 뛰어넘는 편리함을 제공한다[1]. 그러나 이러한 기술의 발전에도 불구하고, 청각 장애나 자폐 스펙트럼 장애를 가진 사용자들에게는 여전히 소통의 어려움이 존재한다[2]. 청각 장애인은 비언어적 단서와 감정의 뉘앙스를 놓치는 경우가 많아 상대방의 감정을 충분히 이해하지 못하거나 자신의 감정을 표현하는 데 어려움을 겪는다[3].

이러한 문제를 해결하기 위해 음성 기반 감정 인식 채팅 시스템을 제안한다. 본 연구에서는 음성 메시지에서 감정을 인식하여 텍스트와 함께 전달함으로써 청각 장애인 사용자가 상대방의 감정을 보다 쉽게 이해하고, 자신의 감정 또한 더욱 자연스럽게 표현할 수 있도록 돕고자 한다. 이를 통해 청각 장애인의 사회적 참여를 촉진하고, 의사소통에서 느낄 수 있는 소외감을 줄이는 데 기여하고자 한다.

또한 본 연구에서는 음성 인식과 감정 분석 기술을 결합하여 청각 장애인뿐만 아니라 발달 장애를 가진 사용자 의사소통에 어려움을 겪고 있는 사람들도 대화의 감정적 맥락을 놓치지 않도록 하는 실시간 웹 기반 서비스를 제안한다. 이를 통해 더욱욱 포용적인 소통 환경을 조성하고, 다양한 사용자의 의사소통 경험을 풍부하게 만드는 목표를 가진다.

## II. 관련 연구

### 2.1 Transformer Model

Transformer Model은 자연어 처리(NLP)에서 사용된다. 기존 RNN 방식에서 벗어나, 다층 구조의 인코더-디코더로 구성되어 있다. 이 모델은 Scaled Dot-Product와 Multi-Head라는 두 가지 형태의 Attention 구조를 활용하며, 각 인코더와 디코더 계층에는 Attention 서브레이어 외에도 완전히 연결된 피드 포워드 네트워크가 포함된다. Transformer 모

델의 가장 큰 특징은 세 가지 방식으로 적용되는 Attention 메커니즘에 있다고 할 수 있다. 인코더 내의 Self-Attention, 디코더 내의 Self-Attention 그리고 인코더 - 디코더 Attention으로 위치별로 위치 참조, 토큰 예측, 그리고 토큰 생성을 할 수 있다[4]. 본 연구는 음성 데이터의 복잡한 특성을 효과적으로 포착하기 위해 Transformer 모델을 기반으로 하여 인공지능 학습을 진행하였다.

### 2.2 HuBERT Model

HuBERT Model은 자기 지도 학습 방식을 기반으로 하여 마스킹 된 예측을 통해 효과적인 음성 표현을 학습하는 모델이다. 이 모델의 특징은 음성 데이터에서 정의되지 않고 발견되어야 하는 Hidden Unit이라는 잠재된 음향 단위를 학습 목표로 설정한다는 점이다. 모델은 입력 데이터로부터 자체적으로 특징을 추출하여 Hidden Unit을 식별하며, 이를 통해 음성의 구조와 패턴을 자동으로 파악하고 학습을 수행한다.

HuBERT의 자기 지도 기반 음성 표현 학습 과정은 다음과 같이 진행된다. Libri-Speech와 Libri-Light 데이터셋을 훈련 데이터로 활용하며, Hidden Unit의 추출을 위해 K-Means 클러스터링과 가우시안 혼합 모델(GMMs)을 적용한다. 이 과정에서 K-Means 클러스터링을 여러 차례 수행하여 마스킹 된 구간의 잠재 클러스터를 예측하는 방식으로 학습이 이루어진다. 이러한 반복 학습 방식을 통해 훈련된 모델은 기존의 최신 기술들과 비교했을 때 탁월한 성과를 달성했으며, 과형을 직접 입력으로 사용하고 반복적 개선을 통해 더욱 정교한 유닛을 학습할 수 있다는 장점을 보였다[5]. 본 연구에서는 이러한 음성 인식 성능을 개선하기 위해 HuBERT 모델의 반복 학습 방식과 Fine-Tuning을 활용하였고, 이를 통해 음성 데이터에서 효과적인 감정 추출 학습을 진행하였다.

## III. 음성 감정 판별 모델 구현

본 연구에서는 청각 장애인의 감정 이해를 돕기

위해 음성 데이터의 톤, 강도, 높낮이, 파장의 특징을 추출하여 감정을 분석하는 모델을 구현하고, 이를 상대방에게 감정 판별 결과를 시각적으로 전달하는 웹 서비스를 구현하였다.

이 시스템의 백엔드는 실시간 데이터 처리를 위해 Spring Boot JPA와 MySQL을 사용하여 데이터베이스를 구축하였고, STOMP 프로토콜을 이용한 pub/sub 구조로 실시간 채팅 기능을 지원하였다. 음성 데이터는 사용자가 전송한 후 서버에서 실시간으로 수신되어 감정 분석이 이루어지며, 분석된 감정 상태와 변환된 텍스트 메시지는 클라이언트로 전송된다. 이러한 방식은 WebSocket을 사용함으로써 클라이언트와 서버 간의 빠르고 안정적인 데이터 통신이 가능해져, 즉각적인 피드백과 실시간 소통이 가능해졌다[4].

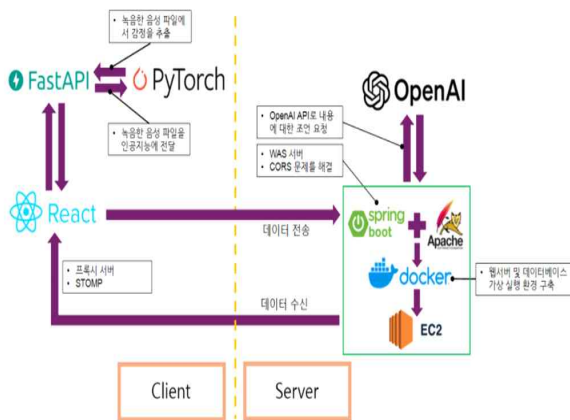


그림 1. 시스템 흐름도  
Fig. 1. System Flow Chart

프론트엔드 개발에는 React를 사용하여 사용자 인터페이스(UI)를 구성하였으며, 사용자가 녹음된 음성 데이터를 웹 페이지 상에서 직접 전송하고 실시간으로 감정 분석 결과를 확인할 수 있다. 음성 데이터를 분석하는 서버는 FastAPI PyTorch를 사용하여 AI 기반 감정 분석 모델을 구현하였고, HuBERT와 Transformer 모델의 결합체를 적용시켜 정확도가 가장 높은 모델을 사용하여 감정 인식을 실현하였다[5].

그림 2에서는 사용자가 웹 서비스를 통해 채팅을 이용하는 과정을 보여준다.

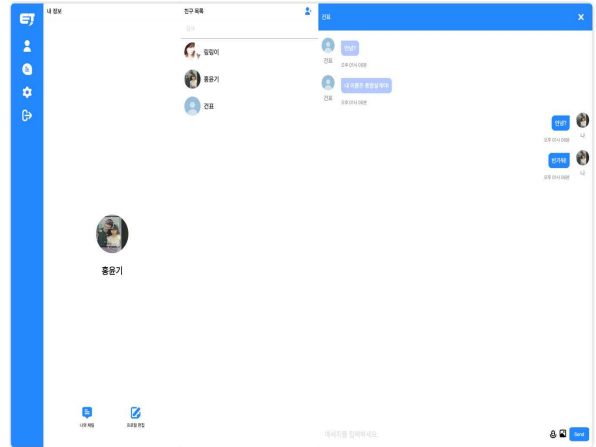


그림 648. 사용자 UI/UX  
Fig. 2. User UI/UX

사용자는 채팅을 통해 음성을 녹음하고, 서버는 이를 수신해 음성 데이터를 추출한 후 감정을 판별한다. 판별된 감정 결과는 상대방에게 채팅 형식으로 전송되며, 전송된 파일은 분석이 완료된 후 최종 감정 결과를 도출한다.

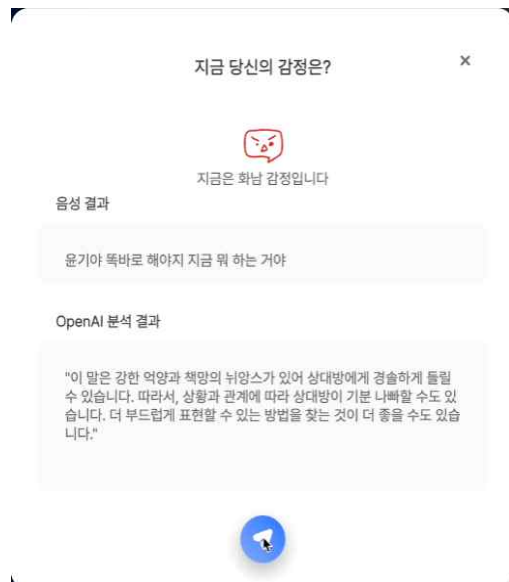


그림 3. 음성 판별 결과화면  
Fig. 3. Voice Results Screen

그림 3에서는 판별된 음성 감정이 총 6가지로 구분된 결과창을 보여준다. 감정은 화남, 두려움, 슬픔, 평범함, 행복함, 싫어함으로 나뉘며, 각 감정에 따라 피드백 문장이 함께 제공된다.

#### IV. 결 론

본 연구는 음성 기반 채팅 시스템을 통해 청각 장애인과 지체 스펙트럼 장애인 사용자에게 원활한 소통을 지원하는 것을 목표로 하였다. 제안된 시스템은 실시간 감정 분석을 통해 사용자가 전달하는 감정 상태를 시각적 피드백으로 변환하여 텍스트 메시지와 함께 상대방에게 전달한다. 이를 통해 의사소통이 원활해지고, 이들의 사회적 자립을 돕는데 기여할 것으로 기대된다.

[7] S. G. Yun, H. C. Kwon, E. Park and Y. B. Cho, "Design and Development of Open-Source-Based Artificial Intelligence for Emotion Extraction from Voice", Journal of The Korea Society of Computer and Information, vol. 29, no. 9, pp. 79-87, September 2024.

#### 참 고 문 헌

- [1] C. Dewes, A. Wichmann and A. Feldmann, "An analysis of Internet chat systems", Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement, pp. 51-64, October 2003.
- [2] L. Pliska, I. Neitzel and U. Ritterfeld, "Toward digital participation in individuals with Autism Spectrum Disorder", Frontiers in Communication, vol. 8, July 2023.
- [3] G. P. Gwon and Y. K. Hong, "Design of a chatting system for the handicapped person", Proceedings of KIIT Conference, pp. 349-352, May 2024.
- [4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, "Attention Is All You Need", Advances in Neural Information Processing Systems, 2017.
- [5] W. N. Hsu, B. Bolte, Y. H. H. Tsai, K. Lakhotia, R. Salakhutdinov and A. Mohamed, "Hubert: Self-supervised speech representation learning by masked prediction of hidden units", IEEE/ACM transactions on audio, speech, and language processing, vol. 29, pp. 3451-3460, October 2021.
- [6] V. Wang, F. Salim, and P. Moskovits, "Using messaging over WebSocket with STOMP," in \*The Definitive Guide to HTML5 WebSocket\*, Berkeley, CA: Apress, pp. 85-108, 2013.

# 스마트 제조환경에서의 작업자 안전을 위한 YOLOv8-CNNLSTM 기반 실시간 위험 감지 및 예측 시스템

정소미\*, 신종현\*\*, 김남훈\*\*\*, 이선우\*\*\*, 윤현중\*\*\*\*

## Real-time Hazard Detection and Prediction System for Worker Safety in Smart Manufacturing Environments Using YOLOv8-CNNLSTM Architecture

So-Mi Jeong\*, Jong-Hyeon Sin\*\*, Nam-Hun Kim\*\*\*, Sun-Woo Lee\*\*\* and Hyun-Joong Yoon\*\*\*\*

### 요 약

본 연구는 제조 현장의 작업자 안전 강화를 위해 YOLOv8과 CNN-LSTM 기반의 하이브리드 딥러닝 아키텍처를 활용한 실시간 위험 감지 및 예측 시스템을 제안한다. 제안된 시스템은 CCTV 영상을 통해 작업자의 위치와 행동 패턴을 실시간으로 분석하고 잠재적 위험을 예측하는 것을 목표로 하며, 실제 제조 현장에 적용 가능한 구조로 설계되었다. 본 연구는 스마트 제조 환경에서의 선제적 안전 관리 체계 구축과 산업재해 예방에 기여할 것이다.

### Abstract

This study presents a real-time risk detection and prediction system designed to enhance worker safety in manufacturing environments through a hybrid deep learning architecture that integrates YOLOv8 and CNN-LSTM models. The proposed system conducts an analysis of worker positions and behavioral patterns in real-time by utilizing CCTV footage, employing edge computing and cloud-based data management to facilitate immediate risk detection and long-term data analysis. The architecture comprises six primary layers: data collection, artificial intelligence processing, risk analysis, response mechanisms, data management, and learning. The YOLOv8 model is responsible for real-time object detection to track workers, while the CNN-LSTM model is utilized to analyze temporal patterns for predicting potential hazards. This system aspires to advance preventive safety management and mitigate industrial accidents within smart manufacturing environments.

### Key words

smart factory, worker safety, artificial intelligence, real-time monitoring, YOLOv8, CNN-LSTM, hybrid deep learning, computer vision, time series analysis, risk prediction, Industry 4.0

\* 국립안동대학교 컴퓨터공학과, someee0706@naver.com

\*\* 계명대학교 기계공학과, sd499@naver.com

\*\*\* 국립금오공과대학교 기계시스템공학부, gnsskazla@naver.com, amway1213@naver.com

\*\*\*\* 대구가톨릭대학교 기계공학과, yoon@cu.ac.kr(교신저자)

※ 본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신사업의 결과입니다.(2022RIS-006)

## I. 서 론

제4차 산업혁명(Industry 4.0)으로 제조 현장의 디지털 전환이 가속화되면서 생산 시스템의 자동화와 지능화가 빠르게 진행되고 있다[1]. 스마트 제조 환경에서는 인간과 기계의 협업이 증가함에 따라 새로운 형태의 안전 위험이 발생하고 있다[2]. 이러한 새로운 위험에는 자동화 시스템과 인간 작업자 간의 상호작용 오류, 복잡한 기계 조작 과정에서 인적 실수, 그리고 예측하기 어려운 기계의 오작동 등이 포함된다.

특히, 제조 현장에서의 작업자 안전 문제는 생산 성과 직결되는 중요한 요소로, 2024년 국제노동기구(ILO) 보고서에 따르면 제조업 분야에서 연간 발생하는 치명적 산업재해가 전체의 63%를 차지하고 있다[3]. 기존의 전통적인 안전 관리 시스템은 사후 대응적 접근 방식을 취하고 있어, 실시간으로 변화하는 스마트 제조 환경에서 발생할 수 있는 잠재적 위험을 효과적으로 감지하고 예방하는 데 한계점을 가진다[4].

본 연구에서는 YOLOv8 객체 탐지 알고리즘과 CNN-LSTM 시계열 분석 모델을 결합한 하이브리드 딥러닝 아키텍처를 기반으로 한 실시간 위험 감지 및 예측 시스템을 제안하고자 한다. 이 시스템은 제조 현장의 CCTV 영상을 실시간으로 분석하여 작업자의 안전을 위협하는 특정 행동이나 상황을 감지하고, 시계열 데이터 분석을 통해 향후 발생 가능한 위험도를 예측함으로써 작업자의 안전을 선제적으로 보장하는 것을 목표로 한다[5].

## II. 관련연구

### 2.1 YOLOv8

YOLO (You Only Look Once) 시리즈는 객체 탐지 분야에서 단일 신경망을 통한 실시간 처리 능력으로 주목받아왔다[6]. 그림 1은 YOLOv8의 전체적인 네트워크 구조를 보여주며, 백본(Backbone), 넥(Neck), 헤드(Head) 부분으로 구성되어 효율적인 특징 추출과 객체 탐지를 수행한다[7].

YOLOv8은 Anchor-free 방식은 객체의 중심점을

직접 예측함으로써 바운딩 박스 예측의 효율성을 높이고 NMS(Non-Maximum Suppression) 연산을 최적화한다[6]. 이는 제조 현장에서 작업자와 장비 등 다양한 크기의 객체를 효과적으로 탐지할 수 있게 한다. GPU 환경에서 초당 100프레임 이상의 처리 속도를 유지하면서도 높은 정확도를 보장한다. 이러한 특징들은 작업자와 장비 등 다양한 크기의 객체를 효과적으로 탐지해야 하는 제조 현장의 실시간 안전 모니터링 시스템에 특히 적합하다[7].

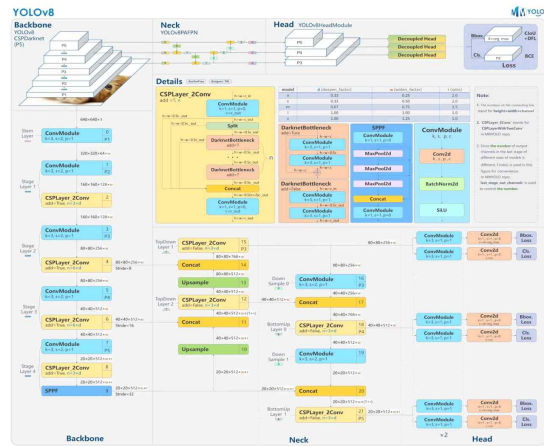


그림 1. YOLOv8 구조  
Fig. 1. YOLOv8 Architecture

### 2.2 CNN-LSTM

CNN-LSTM은 합성곱 신경망(CNN)과 장단기 메모리(LSTM) 네트워크를 결합한 모델로, 시계열 데이터의 시공간적 특징을 효과적으로 분석할 수 있다[8]. 그림 2는 CNN-LSTM의 기본 구조를 보여주며, CNN 층에서 추출된 공간적 특징이 LSTM 층으로 전달되어 시간적 패턴을 학습하는 과정을 나타낸다[9].

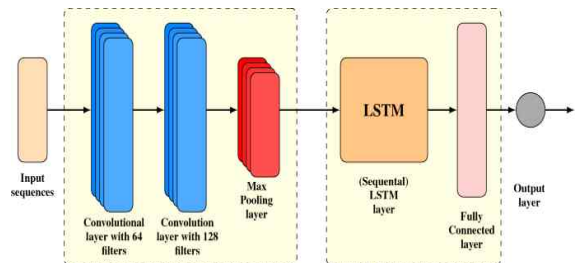


그림 2. CNN-LSTM 구조  
Fig. 2. CNN-LSTM Architecture

행동 인식 분야에서 CNN-LSTM은 우수한 성능을 입증하고 있다. UCI-HAR 데이터셋에서 95.80%, WISDM 데이터셋에서 95.75%의 정확도를 달성했으며[10], 이는 작업자의 행동 패턴을 분석하고 위험 상황을 예측하는 데 효과적임을 보여준다. 제조 현장의 안전 모니터링 측면에서 CNN-LSTM은 작업자의 연속적인 행동 패턴을 분석하고, 시간에 따른 위험 상황의 전개 과정을 학습할 수 있다[11].

본 연구에서는 YOLOv8의 실시간 객체 탐지 능력과 CNN-LSTM의 시계열 데이터 분석 능력을 통합하여 제조 현장의 지능형 안전 관리 시스템을 설계한다. 이를 통해 기존의 수동적 CCTV 모니터링을 자동화된 지능형 감시 시스템으로 발전시켜, 24시간 실시간 모니터링과 즉각적인 위험 상황 대응이 가능한 실용적인 안전 관리 시스템을 구현한다.

### III. 실시간 위험 감지 및 예측 시스템 설계

#### 3.1 YOLOv8-CNNLSTM 기반 실시간 위험 감지 및 예측 시스템 구조

본 연구에서 제안하는 실시간 위험 감지 및 예측 시스템은 스마트 제조 환경에서 발생할 수 있는 다양한 인명사고를 예방하고 효과적으로 대응하는 것을 목표로 한다. YOLOv8 객체 검출과 CNN-LSTM 시계열 분석을 결합한 하이브리드 딥러닝 모델을 기반으로 하여, 작업자의 위치와 행동 패턴을 실시간으로 분석하고 잠재적 위험 상황을 예측할 수 있는 통합 구조를 구현하였다.

제안된 시스템은 데이터 수집, AI 처리, 위험 분석, 대응, 데이터 관리, 학습의 6개 주요 계층으로 구성되어 있으며(그림 3), 각 계층은 독립적으로 작동하면서도 유기적으로 연결되어 실시간 감지와 예측을 수행한다. 특히 엣지 컴퓨팅 기술을 활용한 실시간 처리와 클라우드 기반 데이터 관리를 결합하여 즉각적인 위험 감지와 장기적인 데이터 분석이 가능한 확장 가능한 구조를 채택하였다.

시스템의 각 계층은 명확한 역할과 책임을 가지며, 주요 기능과 세부 구성요소는 표 1에 제시되어 있다. 특히 AI 처리 계층과 위험 분석 계층은 실시간 상황 인식과 예측 분석을 동시에 수행하도록 설계되어 즉각적인 위험 감지와 잠재적 위험 상황에

대한 선제적 대응이 가능하다.

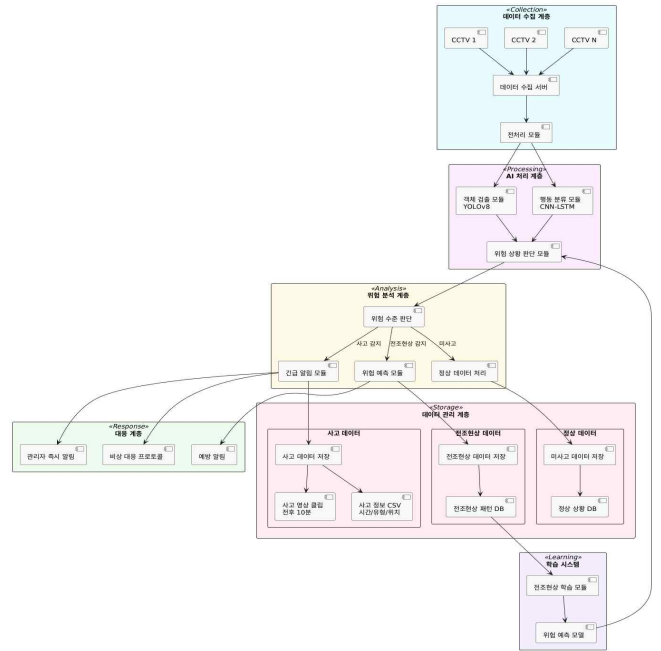


그림 3. 실시간 위험 감지 및 예측 시스템 구조  
Fig. 3. Real-time Risk Detection and Prediction System Architecture

표 1. 시스템 계층별 주요 기능 및 역할  
Table 1. Main Functions and Roles of System Layers

시스템 계층	주요 기능 및 구성 요소
데이터 수집	영상 데이터 획득 및 전처리 • CCTV 1~N, 전처리 모듈
AI 처리	객체 검출 및 행동 분석 • YOLOv8, CNN-LSTM, 판단 모듈
위험 분석	위험수준 평가 • 위험 판단 모듈, 알림 모듈, 예측 모듈
대응	상황별 대응 체계 • 관리자 알림, 대응 프로토콜, 예방 알림
데이터 관리	데이터 저장 및 관리 • 사고 데이터, 전조현상 데이터, 정상 데이터
학습	시스템 성능 개선 • 학습 모듈, 예측 모델

본 시스템은 안전 상태를 사고, 전조현상, 정상의 세 가지로 분류하여 관리하며, 각 상태별로 차별화된 판단 기준, 대응 방식, 데이터 처리 과정을 적용한다. 사고 데이터와 전조현상 데이터를 체계적으로 수집하고 관리함으로써, 지속적인 학습을 통한 시스템의 예측 성능을 향상시키도록 설계되었다.

AI 기반 위험 감지 모듈은 YOLOv8과 CNN-LSTM 모델을 결합한 하이브리드 구조를 채택하여 실시간 객체 검출과 시계열 기반 행동 패턴 분석을 동시에 수행한다. 이를 통해 작업자의 위치 추적, 설비 영역 감지, 위험 행동 예측, 전조현상 감지 등 다양한 안전 관리 기능을 제공한다.

### 3.2 데이터셋 구성 및 시모델

본 연구에서 사용된 스마트 제조 환경 인명사고 데이터셋의 구성은 표 2와 같다. 끼임, 떨어짐, 물체에 맞음, 부딪힘, 쓰러짐 등 다양한 사고 유형별로 데이터를 수집하였으며, 정상 작업 상황도 포함하여 총 180,000건의 데이터를 확보하였다.

표 2. 인명사고 데이터셋

Table 2. Composition of human injury accident dataset

사고 유형	데이터 수량	특징
끼임	40,000	설비 작업 중 발생
떨어짐	40,000	고소 작업 관련
물체에 맞음	34,000	운반/이동 중 발생
부딪힘	28,000	보행/이동 중 발생
쓰러짐	28,000	작업장 내 발생
미사고	10,000	정상 작업 상황

### 3.3 모델 학습 및 현장 적용

수집된 데이터를 활용하여 YOLOv8과 CNN-LSTM 모델을 학습하였다. 그림 4와 5는 학습된 모델을 실제 제조 현장에 적용한 예시로, 작업자의 행동 분석과 위험 상황 감지가 효과적으로 이루어짐을 확인할 수 있다. 표 3은 각 모델의 학습 파라미터를 정리한 것이다. 배치 크기, 학습률, 에폭 수, 최적화 함수 등을 적절히 설정하여 모델의 성능을 극대화하였다.

표 3. 학습 파라미터

Table 3. Training Parameters

항목	YOLOv8	CNN-LSTM
배치크기	32	64
학습률	0.001	0.001
최적화 알고리즘	AdamW	Adam
학습 횟수	100	50

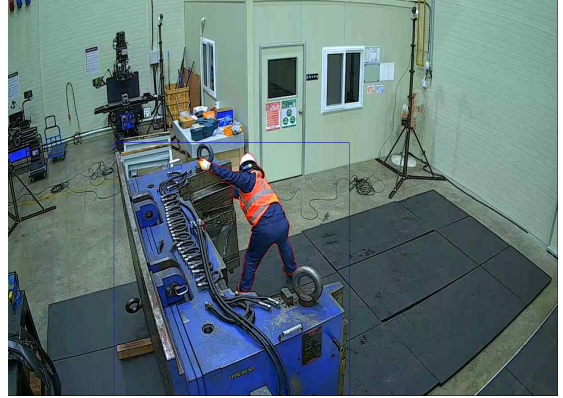


그림 4. 위험 상황 감지

Fig. 4. Hazardous Situation Detection



그림 5. 작업자 행동 분석

Fig. 5. Worker Behavior Analysis

## IV. 결론

본 연구에서는 제조 현장의 인명사고 예방을 위한 AI 기반 실시간 위험 감지 및 예측 시스템을 제안하였다. YOLOv8과 CNN-LSTM을 결합한 하이브리드 구조를 통해 작업자의 위치 및 행동 패턴을 분석하고 잠재적 위험을 예측하는 통합 시스템을 구현하였으며, 실제 제조 현장 적용을 통해 그 실효성을 검증하였다.

현재 연구의 주요 한계점은 다양한 제조 환경에서의 현장 데이터셋 확보가 제한적이라는 점이다. 이러한 한계를 극복하기 위해 향후 연구에서는 다양한 산업 현장에서의 실증적 검증과 함께 설비 이상 감지 등 추가적인 기능 확장을 진행할 계획이다.

본 연구에서 제안된 시스템은 스마트 제조 환경에서 작업자의 안전을 실시간으로 모니터링하고 감



재적 위험을 사전에 예측함으로써, 기존의 사후 대응적 안전 관리 체계를 한 단계 발전시켰다는 점에서 의의가 있다. 본 연구 결과는 제조업의 디지털 전환 과정에서 안전 관리 고도화와 산업재해 예방에 실질적으로 기여할 것으로 기대되며, 이를 통해 작업자 안전 확보와 생산성 향상이라는 측면에서 중요한 시사점을 제공한다.

## 참 고 문 헌

- [1] <http://www.anjunj.com/news/articleView.html?idxno=39238>
- [2] [https://webzine.mynewsletter.co.kr/newsletter/kcplaa/202208-4/issue2\\_%EC%95%88%EC%A0%84%EB%B3%B4%EA%B1%B4%EA%B4%80%EB%A6%AC%EC%B2%B4%EA%B3%84\\_%EA%B5%AC%EC%B6%95\\_%EC%9A%B0%EC%88%98%EC%82%AC%EB%A1%80%EC%A7%91.pdf](https://webzine.mynewsletter.co.kr/newsletter/kcplaa/202208-4/issue2_%EC%95%88%EC%A0%84%EB%B3%B4%EA%B1%B4%EA%B4%80%EB%A6%AC%EC%B2%B4%EA%B3%84_%EA%B5%AC%EC%B6%95_%EC%9A%B0%EC%88%98%EC%82%AC%EB%A1%80%EC%A7%91.pdf)
- [3] <https://www.intel.co.kr/content/www/kr/ko/customer-spotlight/stories/aotu-customer-story.html>
- [4] <https://saige.ai/kr/solutions/safety>
- [5] <https://www.safety1st.news/news/articleView.html?idxno=5137>
- [6] <https://www.kosha.or.kr/oshri/publication/researchReportSearch.do?article.offset=420&articleLimit=10&articleNo=408233&mode=view>
- [7] <https://www.rockwellautomation.com/ko-kr/capabilities/digital-transformation/state-of-smart-manufacturing.html>
- [8] <https://velog.io/@gredora/LSTM-CNN-Architecture-for-Human-Activity-Recognition>
- [9] Livieris, I.E., Pintelas, E. & Pintelas, P. A CNN-LSTM model for gold price time-series forecasting. *Neural Comput & Applic* 32, 17351-17360 (2020). <https://doi.org/10.1007/s00521-020-04867-x>
- [10] <https://www.mdpi.com/1424-8220/24/16/5436>
- [11] <https://pseudo-lab.github.io/Tutorial-Book-en/chapters/en/time-series/Ch5-CNN-LSTM.html>

# 인공지능을 활용한 부동산 왜곡 이미지 탐지 및 보정 서비스 설계 및 구현

권혁찬\*, 김선혁\*, 윤성건\*, 홍윤기\*, 정은미\*\*

## Design and Implementation of Real Estate Distortion Image Detection and Correction Service Using Artificial Intelligence

Hyukchan Kwon\*, Seonhyuk Kim\*, Seong-Geon Yun\*, Yoon-gi Hong\*, and Eunmi Jeong\*\*

### 요 약

본 논문은 부동산 매물 사진의 왜곡 문제를 해결하기 위해 Geocalib 모델을 활용한 AI 기반 부동산 왜곡 이미지 탐지 및 보정 서비스인 Propix를 제안한다. 부동산 거래에서 매물 사진은 소비자에게 중요한 정보로 작용하지만, 많은 경우 사진이 실물과 다르게 보이도록 왜곡된다. 이러한 왜곡은 소비자에게 혼란을 초래하고, 부동산 시장의 신뢰성을 저하시킨다. 기존의 특징점 기반, 직선 검출 기반 방식과 GeoCalib 모델을 비교하여, 왜곡을 효과적으로 감지하고 보정할 수 있는 방법을 탐색하였다. Propix는 사용자가 업로드한 매물 이미지를 GeoCalib 모델을 통해 분석하여 왜곡 여부를 자동 판별하고, 웹 기반 인터페이스에서 결과를 제공한다. 실시간 왜곡 분석 결과와 PDF 보고서를 제공함으로써 거래의 투명성을 높이고 소비자에게 보다 정확한 정보를 전달하는 데 기여하고자 한다.

### Abstract

This paper proposes Propix, an AI-based real estate distortion image detection and correction service using Geocalib model to solve the problem of distortion in real estate listing photos. In real estate transactions, property photos are an important source of information for consumers, but in many cases, photos are distorted to make them look different from the real thing. This distortion is confusing to consumers and undermines the credibility of the property market. By comparing the GeoCalib model to traditional feature point-based and line detection-based methods, we explore how it can effectively detect and correct distortions. Propix analyses user-uploaded property images with the GeoCalib model to automatically determine whether they are distorted and provides results in a web-based interface. By providing real-time distortion analysis results and PDF reports, Propix aims to increase transparency in real estate transactions and provide consumers with more accurate information.

### Key words

Real estate image distortion, Geocalib model, Distortion detection, Image correction, Transparency in real estate

---

\* 국립안동대학교 SW융합학과, kwonhc226@gmail.com ,

\*\* 국립안동대학교 SW융합학과, emjung@anu.ac.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

## I. 서론

부동산 거래는 개인의 생활 방식을 형성하고 경제적 안정성을 결정짓는 중요한 요소로 자리 잡았다. 주택 구매와 임대는 단순한 자산 거래를 넘어 개인의 사회적 지위와 생활 수준에 직결되는 문제로 여겨지며, 이러한 이유로 부동산 시장에서의 투명하고 공정한 거래는 소비자와 공급자 모두에게 필수적이다. 이 과정에서 매물 사진의 신뢰성은 거래의 성패를 좌우하는 중요한 요소로 부각된다. 소비자들은 매물 정보를 판단하는 데 있어 사진에 크게 의존하며, 이는 결국 거래 결정에 중대한 영향을 미친다.

그러나 최근 부동산 광고에서 사용되는 이미지의 신뢰성 문제는 심각한 우려를 낳고 있다. 많은 경우, 매물 사진은 실물을 왜곡하여 보이도록 편집되거나 조작되어 소비자의 기대와 실제 제품 간의 간극을 초래한다. 예를 들어, 공간이 실제보다 더 넓게 보이도록 조작된 이미지나 특정 각도로 촬영된 사진은 매수자에게 잘못된 정보를 제공하여 거래의 혼란과 실망을 초래할 수 있다. 이러한 왜곡된 정보는 소비자에게 경제적 손실을 가져올 뿐만 아니라, 부동산 시장에 대한 신뢰를 저하시켜 궁극적으로 시장의 건강성을 해치는 결과를 초래할 수 있다.

이러한 문제를 해결하기 위해 본 연구는 GeoCalib 모델을 활용하여 매물 이미지의 왜곡을 판별하는 Propix(프로픽스) 서비스를 제안한다. GeoCalib은 기하학적 최적화 기법을 통해 단일 이미지에서 왜곡 요소를 감지하는 혁신적인 인공지능 모델이다. 이 모델은 체커보드 패턴이나 특징점 기반의 교정 작업을 수행하지 않고도 사진 내의 왜곡을 추정할 수 있는 능력을 갖추고 있다. Propix는 사용자가 업로드한 이미지의 왜곡 여부를 자동으로 분석할 수 있는 기능을 제공한다.

## II. 관련 연구

### 2.1 카메라 캘리브레이션(Camera Calibration)

카메라 캘리브레이션은 사진 촬영 시 발생하는 왜곡을 보정하기 위해 카메라의 내·외부 매개변수

를 조정하는 과정이다. 캘리브레이션 기술에는 여러 가지 방법이 존재하며, Single-image Calibration 기술은 단일 이미지로도 보정이 가능하다는 점에서 장점이 있다. 이를 통해 이미지 다수를 촬영하거나 복잡한 설정을 요구하지 않고도 왜곡을 줄일 수 있어 비전 시스템의 효율성을 높여준다[1].

GeoCalib은 이러한 캘리브레이션 기법 중 하나로, 특히 다중 이미지를 사용하여 높은 정확도로 카메라 매개변수를 보정할 수 있다. 이 알고리즘은 반복적 최적화를 통해 이미지의 공간적 왜곡을 정밀히 보정하며, 지리 정보와 같이 높은 신뢰성을 요구하는 이미지 처리에 유용하다. GeoCalib은 다양한 환경에서도 왜곡을 줄이기 위한 안정적인 보정 성능을 제공하는 대표적인 캘리브레이션 방법으로 널리 사용되고 있다[2].

최근에는 심층 학습을 활용한 딥러닝 기반 캘리브레이션 방법이 활발히 연구되고 있다. 딥러닝 캘리브레이션 모델은 대량의 학습 데이터를 통해 비선형 왜곡을 정밀하게 보정할 수 있어, 다양한 환경에서도 높은 정확도를 제공한다. 예를 들어, 심층 신경망(CNN) 기반의 자동 캘리브레이션 기법은 부동산 이미지의 왜곡을 보정하는 데 효과적으로 활용될 수 있다. 딥러닝 기반 캘리브레이션은 실물에 가까운 부동산 이미지 제공에 기여할 가능성이 있다[3].

### 2.2 부동산 사진 왜곡 분석

부동산 사진의 왜곡 문제는 주로 촬영 기법과 장비의 특성에 기인하며, 부동산 정보를 정확히 전달하는 데 큰 영향을 미친다. 광각 렌즈 왜곡은 부동산 사진에서 특히 빈번하게 발생하는 왜곡 문제 중 하나다. 넓은 공간을 한 장에 담기 위해 광각 렌즈를 사용할 경우, 사진 가장자리에 심한 왜곡이 발생하여 실제 공간과는 다른 이미지를 제공할 수 있다. 이러한 왜곡은 소비자에게 부정확한 공간감을 전달하여 신뢰도를 저하시킬 위험이 있다[4].

부동산 사진에서는 또한 특정 촬영 기법을 활용하여 공간감을 극대화하는 경우가 많으나, 이로 인해 사진이 실제보다 왜곡될 수 있다. 예를 들어, 특정 각도나 조명 설정은 사진이 과장되게 보이도록 만

들어 소비자에게 왜곡된 공간 이미지를 전달할 가능성이 있다. 이는 부동산 거래 시 소비자의 기대와 실제 매물 간에 차이를 초래할 수 있다[5].

부동산 앱에서는 사진의 신뢰성이 중요한 요소로 작용하며, 신뢰성 높은 사진을 제공하는 것이 사용자의 경험에 긍정적인 영향을 미친다. 부동산 사진의 왜곡을 감지하고 보정하는 시스템은 부동산 앱의 신뢰성을 높여, 사용자에게 보다 정확한 정보를 제공함으로써 거래의 투명성을 강화하는 데 기여할 수 있다[6].

### III. 인공지능 비교 및 구현

부동산 이미지는 사진의 왜곡이나 렌즈 특성에 의해 정확한 정보를 제공하지 못하는 경우가 많다. 본 연구에서는 특징점 기반 방식, 직선 검출 기반 방식과 AI 모델인 GeoCalib을 비교하여, 부동산 이미지에서 왜곡 탐지와 보정이 가능한 최적의 접근법을 제안한다.

#### 3.1 특징점 기반 왜곡 분석

특징점 기반 왜곡 분석은 ORB (Oriented FAST and Rotated BRIEF) 알고리즘을 활용하여 이미지의 상단과 하단(또는 좌우)에서 특징점을 추출하고, 두 영역의 특징점 매칭 결과를 왜곡 여부의 주요 지표로 활용한다.

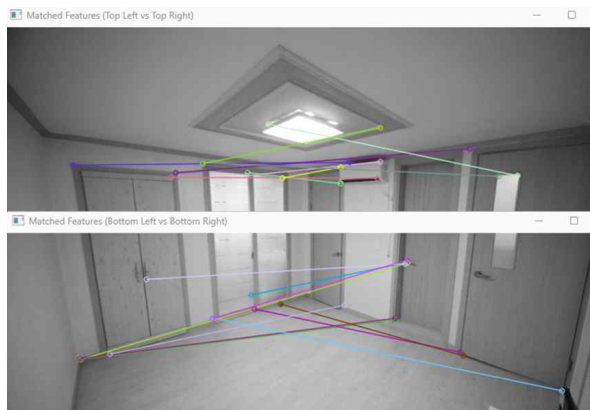


그림 1. 특징점 기반 왜곡 분석 결과 이미지  
Fig. 1. Extract and analyze feature points result image

그림 1과 같이 이미지의 상단과 하단에서 추출한 특징점이 고르게 매칭되지 않는다면 왜곡이 있는 것으로 판단할 수 있다. 그러나 단일 이미지로 왜곡을 탐지하는 본 연구에서는 두 이미지 간의 유사도를 평가하는 특징점 기반 방식은 부적합하다고 판단하여 제외하였다.

#### 3.2 직선 검출 기반 왜곡 분석

직선 검출 기반 왜곡 분석은 이미지 내에서 직선을 검출하고, 기울기와 일관성을 통해 왜곡 여부를 판단하는 방식이다.

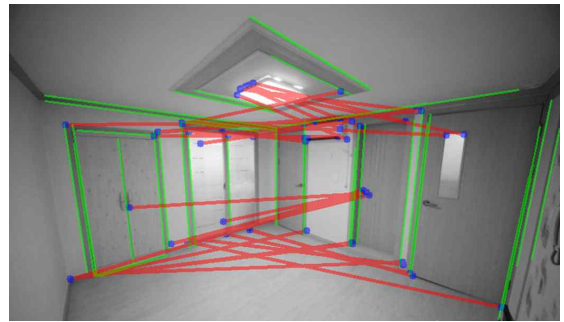


그림 2. 직선 검출 기반 왜곡 분석 결과 이미지  
Fig. 2. Estimating linear-based distortion coefficients result image

그림 2와 같이 벽이나 창문과 같은 요소가 직선으로 표현되는 부동산 이미지의 경우, 이 방법은 이미지 왜곡을 탐지하는 효과적인 방식이 될 수 있다. 간단한 계산으로 왜곡을 탐지할 수 있지만, 이미지 구조에 직선 요소가 적거나 각도가 비정상적인 경우 정확도가 떨어질 수 있다.

#### 3.3 GeoCalib 모델을 활용한 시기반 왜곡 보정

GeoCalib 모델은 단일 이미지에서 카메라 보정을 수행하는 AI 모델로, 다양한 왜곡 상황에서도 높은 정밀도를 제공한다. GeoCalib 모델은 단일 이미지의 초점 거리(focal length), 방사형 왜곡(radii distortion), 중력 방향(gravity direction)을 추정하여 보정 정확도를 높인다.

그림 3과 같이 GeoCalib 모델은 보정의 신뢰도를 평가하기 위해 Calibration Result, Up Confidence, Latitude Confidence와 같은 지표를 제공한다.

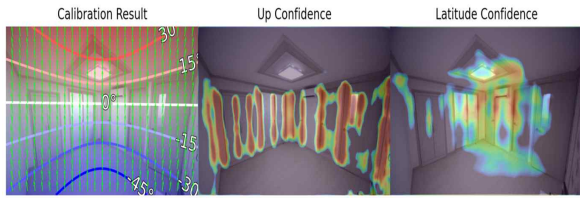


그림 3. GeoCalib 모델 결과 이미지  
Fig. 3. Analyze results for implemented distortions Web

Up Confidence는 상향 방향의 정확도를 나타내고, Latitude Confidence는 위도 방향의 신뢰도를 제공하여 보정 결과의 신뢰성을 높인다. 만약 Up Confidence와 Latitude Confidence가 낮다면 이미지의 왜곡 가능성이 높다. GeoCalib 모델은 직선 검출 방식에 비해 정밀도와 자동화된 보정 성능이 우수하여 최종적으로 이 모델을 선택하였다.

```

알고리즘 1. GeoCalib 모델을 활용한 왜곡 보정 알고리즘
function GeoCalib_analysis(img_path, use_prior=False):
    device = Check if CUDA is available else use CPU
    model = Initialize GeoCalib model with 'pinhole' or
'distorted' weights based on camera model
    img = Load image from img_path
    if use_prior:
        Set prior focal length or gravity for Calibration
        results = Calibrate img with model using prior
information
    else:
        results = Calibrate img with model without
prior information
    undistorted_img = Undistort img using Calibration
results
    if (Confidence scores for up and latitude are high):
        Mark image as undistorted
    else:
        Mark image as distorted
    return undistorted_img
    
```

### 3.4 왜곡 여부 결과 분석 웹 구현

본 논문에서는 부동산 매물 사진의 왜곡도를 분석하고 시각화하는 웹 기반 분석 시스템을 제안한다. 기존 카메라 캘리브레이션 모델들의 비교 분석을 통해 최적의 성능을 보인 GeoCalib 오픈소스 모

델을 선정하여 실제 서비스에 적용 가능한 웹 시스템으로 구현하였다. 제안된 시스템은 웹 인터페이스, 데이터 처리를 위한 백엔드 서버 및 왜곡도 분석을 수행하는 AI 분석 서버로 구성되며, 사용자가 업로드한 매물 이미지는 메타데이터와 함께 데이터베이스에 저장된다. GeoCalib 모델을 통해 분석된 왜곡도 결과는 웹 인터페이스를 통해 즉시 확인이 가능하며, PDF 형식의 상세 보고서로도 제공되어 추후 참고자료로 활용할 수 있다.



그림 4. 구현된 왜곡 여부 결과 분석 웹  
Fig. 4. Analyze results for implemented distortions Web

그림 4는 구현된 부동산 매물 이미지 왜곡도 분석 시스템의 결과 화면을 보여준다. 상단에 배치된 원본 이미지와 AI 보정 이미지의 비교를 통해 전체적인 왜곡 정도를 파악할 수 있으며, 중앙부에서는 이를 정밀하게 분석한 Calibration Result, Up Confidence, Latitude Confidence를 제공한다. 특히 Calibration Result는 격자무늬를 통해 렌즈 왜곡의 정도를, Up Confidence와 Latitude Confidence는 각각 공간의 수직성과 비율 정확도를 색상 기반으로 시각화한다. 하단부에서는 해당 분석 결과를 Roll,

Pitch, vFoV, Focal 등의 정량적 수치로 제시하여 객관적인 평가가 가능하도록 하였다.

현재는 왜곡률 수치의 기준값 도출을 위한 연구가 진행 중이며, 실제 부동산 매물 데이터를 기반으로 한 모델 학습을 통해 더 정교한 왜곡률 측정 기준을 수립할 예정이다. 또한, 분석 결과를 바탕으로 한 경고 메시지 시스템을 구현하여, 부동산 거래 시 발생할 수 있는 시각적 왜곡으로 인한 오해를 사전에 방지하고자 한다. PDF 형식의 분석 보고서는 원본 및 보정 이미지 비교, 세부 분석 결과, 카메라 파라미터 등을 포함하여, 거래 당사자 간 객관적 정보 교환의 기반을 제공한다.

#### IV. 결 론

본 논문에서 제안한 Propix 서비스는 GeoCalib 모델을 활용하여 부동산 매물 사진의 왜곡을 감지하고 보정하는 시스템으로, 거래의 신뢰성과 투명성을 강화하는 데 기여한다. 기존 방식에 비해 GeoCalib 모델은 비교적 단일 이미지에서 왜곡을 정밀하게 분석하고 보정할 수 있는 성능을 보여 최적의 모델로 선정하였다. 구현된 웹 서비스는 사용자에게 왜곡 분석 결과와 보정 이미지를 제공하여 거래 결정을 위한 매물의 투명성과 신뢰성을 향상한다. 향후 실제 부동산 매물 데이터를 통해 모델을 개선하고, 경고 메시지 시스템으로 사용자에게 부정확한 정보를 미리 알림으로써 왜곡으로 인한 부정적 영향을 줄이고자 한다. Propix 서비스는 매물 이미지에 대한 정확한 정보 제공 및 신뢰성 향상 효과로 전반적인 부동산 거래의 질적 향상에 기여할 수 있다.

#### 참 고 문 헌

- [1] Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1330-1334.
- [2] Smith, A., & Lee, J. (2022). GeoCalib: An algorithm for multi-image camera calibration in real estate applications. *Journal of Real Estate*

*Technology*, 4(2), 135-149.

- [3] Brown, T., & Kim, H. (2023). Deep learning-based camera calibration for real estate images. *Computer Vision Journal*, 18(3), 245-259.
- [4] Chen, L., & Patel, S. (2021). The impact of wide-angle lens distortion on real estate photography. *Journal of Visual Communication*, 7(4), 78-91.
- [5] Davis, M., & Wang, R. (2022). Analyzing real estate photography techniques and their effects on image distortion. *International Journal of Real Estate Studies*, 5(1), 22-34.
- [6] Evans, K., & Martinez, D. (2023). Evaluating the reliability of real estate images in online platforms. *Journal of Digital Consumer Trust*, 3(3), 112-124.

# 장애인 자동차 주차표지 위조 탐지를 위한 고찰 연구

황다건\*, 권건표\*\*, 김현기\*\*\*

## A study on detecting counterfeit disabled parking placards

Dagon Hwang\*, Geon-Pyo Gwon\*\*, and Hyen-Ki Kim\*\*\*

### 요 약

장애인 전용 주차구역의 불법 주차는 지속적으로 증가하고 있으며, 특히 장애인 자동차 표지의 위조 및 부정적인 사용이 심각한 문제로 대두되고 있다. 이러한 문제를 해결하기 위한 기술적 접근으로 OCR과 이미지 처리 기술을 분석했다. 본 논문에서는 초기의 단순 패턴 매칭에서 CNN, RNN 등 딥러닝 기반 모델로 발전하여 복잡한 환경에서도 높은 인식률을 보이고 각각 인쇄체 인식, 손글씨 인식, 다국어 지원 등에서 차별화된 강점을 가지고 있다. 이미지 처리 기술은 GAN, YOLO, CLAHE 등 딥러닝 기반 기술은 저품질 이미지 복원, 선택적 문자 영역 탐지, 적응적 명암 대비 조정 등을 통해 더 높은 성능을 제공한다. 본 논문에서 분석된 OCR과 이미지 처리 기술의 조합은 장애인 자동차 표지 탐지의 정확성과 신뢰성을 높일 수 있을 것으로 기대된다.

### Abstract

Illegal parking in disabled parking spaces is continuously increasing, and in particular, forgery and misuse of disabled car signs are emerging as serious problems. As a technical approach to solve this problem, OCR and image processing technologies were analyzed. In this paper, deep learning-based models such as CNN and RNN were developed from the initial simple pattern matching, and they show high recognition rates even in complex environments and have differentiated strengths in printed recognition, handwriting recognition, and multilingual support, respectively. Image processing technologies such as GAN, YOLO, and CLAHE provide higher performance through low-quality image restoration, selective character area detection, and adaptive contrast adjustment. The combination of OCR and image processing technologies analyzed in this paper is expected to increase the accuracy and reliability of disabled car sign detection.

### Key words

Parking signs, OCR, Image processing, Deep learning, Counterfeiting

---

\* 국립안동대학교 대학원 소프트웨어융합학과,

\*\* 국립안동대학교 멀티미디어공학과

\*\*\* 국립안동대학교 소프트웨어융합학과, hkkim@anu.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원에서 지원하는 SW중심대학사업(IITP-2019-0-01113)의 연구 결과로 수행되었음

## I. 서 론

장애인 주차 구역은 이동이 불편한 장애인의 접근성과 편의성을 위해 필요한 시설이다. 장애인 주차구역을 이용하기 위해서는 보행상 장애가 있는 자에게 발급이 되는 장애인 자동차 표지(장애인 차량 스티커)를 차량에 부착해야 하며, 해당 표지를 부착한 차량이라도 보행상 장애가 있는 자가 탑승하지 않은 경우에는 장애인 전용 주차구역 불법 주차로 간주하여 단속의 대상이 된다. 국민권익위원회에 따르면, 장애인 전용구역 불법 주차 민원신고는 2019년 45만9,116건이던 불법 주차 신고가 2021년 125만1,028건, 2022년 84만건(코로나19), 2023년 162만7,195건으로 매년 증가하는 추세이다. 특히, 장애인 자동차 표지를 위조하여 사용하거나, 부정 사용하는 비장애인들이 장애인 주차 공간에 주차하는 사례가 늘어나면서 실제 장애인들이 주차 공간을 이용하는데 큰 불편을 겪고 있다[1].

장애인 자동차 표지의 위조 사례는 다양한 형태로 나타나고 있다. 폐지된 장애인 차량 표지에 기재된 차량번호를 변경하거나, 인터넷에서 장애인 차량 표지 그림 파일을 다운로드 받아 컬러 프린터로 출력하는 등의 방법이 사용되고 있다. 또한 물건으로 반정도 가려 위조방지 홀로그램과 차량번호를 보이지 않게 하는 등의 수법도 발견되고 있다. 이러한 위조 표지 및 부정으로 사용하는 표지는 육안으로 단속할 때 장애인 차량번호와 장애인 자동차 표지에 적혀있는 번호가 동일한 것인지에 확인이 어려울 때가 있다[2]. 일부 지방자치단체는 위조 및 도용 등 유효하지 않은 장애인 자동차 표지를 부착한 차량을 확인하기 위해 IoT(사물인터넷) 기술을 활용한 무인 단속 시스템을 시행하고 있지만 대부분의 장애인 주차구역 단속은 주로 인력에 의존하고 있어 단속 체계의 한계를 드러내고 있다. 이처럼 현재의 육안 단속 방식은 위조 기술의 발전 속도를 따라가지 못하고 있으며, 효과적인 탐지 방법의 도입이 절실하다[1].

이러한 문제를 해결하는 방법으로, 최근 인공지능(AI)과 컴퓨터 비전 기술의 발전이 주목받고 있다. 특히 OCR(광학 문자 인식)과 이미지 처리기술은 장애인 자동차 표지의 진위 여부를 정확하게 판

별하는 데 효과적인 도구로 활용될 수 있다.

이에 본 논문에서는 장애인 주차 스티커 위조 탐지하기 위한 기술로 요구되는 OCR과 이미지 처리 방법에 대한 전반적인 기술을 다루고자 한다.

## II. 연구 동향

### 2.1 OCR(광학문자인식)

OCR기술은 이미지 속에 있는 문자의 특징을 추출하여 기계가 읽을 수 있는 형태로 변환하는 기술로 문서자동화 처리, 의료, 금융, 법률, 공공서비스 등 다양한 분야에서 문서의 디지털화를 위한 기술로 활용되어 있다. 초기 OCR 시스템은 패턴 매칭 알고리즘을 통해 작동했으며 문자 이미지를 사전에 저장된 글리프(glyph)와 비교하여 가장 유사한 문자를 찾아내는 방식이다. 이러한 초기 패턴매칭 알고리즘 방식은 특정 규격의 글꼴이나 형태에만 적합했으며, 변형된 문자, 복잡한 배경이 있는 경우에는 인식하는 데에는 한계가 있었다[3].

컴퓨터 비전 및 머신러닝 기술의 발전에 따라 OCR기술은 CNN과 RNN과 같은 딥러닝 모델을 적용하여 다양한 환경에서의 문자 인식 성능을 크게 향상시켰다. CNN은 이미지의 문자 특징을 추출하고 RNN은 이러한 특징을 시퀀스로 연결해 문맥의 의미를 파악하고 LSTM(Long Short-Term Memory)과 CTC(Connectionist Temporal Classification) 기법은 문장의 흐름을 이해하고, 텍스트의 순서를 고려하여 보다 정확한 문장을 구성할 수 있게 한다[4][5]. 딥러닝 기반의 OCR 기술은 복잡한 배경에서의 문자 인식, 다양한 글꼴 및 손글씨 인식, 휘어진 문자 인식, 다국어 지원, 등 여러 문제를 해결하는 데 중요한 역할을 하고 있다[6]. 휘어진 문자를 정확하게 인식할 수 있는 TextSnake와 Mask TextSpotter와 같은 모델은 배경이 복잡해도 높은 인식 성능을 보여준다. TextSnake는 FCN(Full Convolutional Networks)을 기반으로 글자의 영역과 중심선, 글자의 방향을 예측하고, Mask TextSpotter는 SPN(Segmentation Proposal Network)를 통해 더욱 정교한 문자 인식을 가능하게 하고 있다[3][5].



## 2.2 이미지 처리 기술

이미지 처리기술에는 그레이 스케일, 히스토그램 평준화, 노이즈 제거, 대비 조정, 에지 검출, 이진화, 크기 조정 등이 있다. 다양한 환경에서 문자 인식에 불필요한 요소를 제거하고 이미지 품질을 개선하여 오인식률을 줄여 정확도를 높이기 위한 전처리 과정에 필수적인 역할을 한다[7]. 전통적인 이미지 처리기술은 노이즈 제거와 흑백 변환 등을 통해 이미지 속에 있는 문자 영역을 명확히 구분하여 OCR의 인식률을 높이고, OpenCV와 같은 이미지 처리 라이브러리로 이미지의 밝기, 대비조절, 노이즈를 제거하여 OCR의 인식률을 높인다. 그러나 이 기술들은 복잡한 배경이나 왜곡된 이미지에서는 성능이 저하되는 한계를 보인다[3][6].

이러한 한계를 보완하기 위해 딥러닝 기반 이미지 처리기술은 복잡한 배경의 이미지에서의 문자 검출과 인식에 중요한 역할을 하고, YOLO(You Only Look Once)와 같은 객체 탐지 모델은 이미지 내에서 특정 영역만 선택적으로 탐지하여 OCR이 해당 부분에 집중하여 복잡한 배경이 있는 이미지에서 필요한 문자만 인식할 수 있다[8]. 또한 낮은 품질의 이미지나 노이즈가 많은 스캔 문서에서 인식률을 높일 수 있는 GAN(Generative Adversarial Networks)과 같은 딥러닝 기법도 있다[9].

## III. 비교 분석

### 3.1 OCR 비교 분석

OCR 기술은 다른 엔진과 모델에 따라 성능과 정확도가 다르게 나타나며, 응용 분야에 맞게 OCR 기술을 선택하는 것이 중요하다. 이 논문에서는 OCR의 엔진 및 모델과 이미지 처리기술의 특징에 대해 설명하고자 한다.

표 1은 딥러닝 기반 OCR 엔진 Tesseract OCR, Google Vision API, EasyOCR, ABBYY Fine Reader이다. 딥러닝 기반 OCR 엔진은 다양한 언어와 정형화된 텍스트와 손글씨 인식에서 뛰어난 성능을 보이며, 최근에는 인공지능 기술의 발전으로 더욱 정확해 졌다.

표 1. OCR 엔진 비교 분석

Table 1. OCR Engine Comparison and Analysis

종류	특징
Tesseract	인쇄된 텍스트, 손글씨 정확도 높음. 밝기 값이 다른 이미지에 효과적, 문자와 배경 영역 구분
Google Vision API	인쇄된 텍스트, 손글씨 정확도 높음. 표나 서명, 로고와 같은 특정한 구조에도 명확하게 인식
EasyOCR	문자 영역 인식과 문자 인식을 동시에 활용, 손글씨에 대한 인식률이 낮음
ABBYY Fine Reader	고정밀 인식, 정밀한 다중 언어 및 서체 인식 지원

Tesseract는 오픈소스 기반의 OCR 엔진으로, 인쇄된 텍스트와 손글씨에 인식 정확도가 높다는 장점이 있으며, 이미지에 문자와 배경 영역을 구분하고, 다양한 언어와 글꼴을 지원한다. Google Vision API는 인쇄된 텍스트뿐만 아니라 손글씨에도 높은 정확도를 보여 주며, 다양한 텍스트 레이아웃과 복잡한 문서 구조에서도 높은 정확성을 제공하는 점에서 로고나, 서명 등 금융, 법률 문서 처리에 좋은 성능을 보인다. EasyOCR는 다양한 글꼴과 언어를 지원하며 커스터마이징이 쉽다는 특징을 가지고 있다. 특히 비정형 텍스트와 손글씨 인식에서 성능이 우수하여, 글자 크기나 글꼴의 차이가 큰 이미지에도 안정적인 인식을 제공한다. ABBYY Fine Reader는 고정밀 인식에 강점을 지니며, 다양한 언어와 서체를 정확하게 인식하는 성능이 뛰어나다. 특히 다국어를 포함하는 문서나 복잡한 서체가 포함된 문서에서도 높은 인식률을 제공한다.

### 3.2 이미지 처리 기술 비교 분석

이미지 처리기술은 전통적인 기술과 딥러닝을 활용하여 전처리 과정 성능을 평가할 수 있다.

표 2는 이미지 처리 기술과 모델의 특징에는 OCR 성능 향상을 위한 전처리 기법은 다양한 이미지 품질 개선과 텍스트 인식을 향상을 목표로 한다. 전통적인 이미지 처리방법의 OpenCV와 같은 라이브러리를 활용하여 노이즈 제거, 밝기 조정, 이진화 등을 수행한다. 이 방법은 문자 영역을 명확히 구분하여 OCR 인식률을 높이는 데 기본적으로 필요한 전처리 작업이며, 실시간 처리를 하여 빠른 속도와

효율성을 자랑한다[10].

표 2. 이미지 처리 기술 비교 분석  
Table 2. Analysis of Image Processing Technologies

방법	특징
전통적	OpenCV 간단하고 빠른 전처리 기능 제공
딥러닝	GAN 가짜 이미지를 생성하여 이미지 복원
	YOLO 특정 문자 영역 선택적 탐지 가능
	CLAHE 명암 대비 조정으로 가독성 향상

특히 단순한 배경의 이미지에는 좋은 성능을 발휘하지만 복잡한 배경이나 다양한 색상이 있는 이미지에는 인식률이 낮게 나온다. 낮은 품질의 이미지나 노이즈가 많은 이미지에서 인식 성능이 좋은 GAN(Generative Adversarial Networks)은 고유한 특성을 가지고 있는 이미지에 유사한 가짜 이미지를 만들어 저해상도 이미지의 복원을 통해 문자 인식 정확도를 향상시킬 수 있다.[11]. YOLO는 객체 탐지 모델로, OCR과 결합하여 특정 문자 영역만 선택적으로 탐지할 수 있는 기능을 제공한다. 복잡한 배경에서 특정 텍스트만을 인식하도록 해주며, 이를 통해 불필요한 텍스트 인식을 방지하여 OCR의 인식 효율을 높인다[9]. CLAHE(Contrast Limited Adaptive Histogram Equalization)는 이미지의 명암 대비를 조정하여 텍스트와 배경을 보다 명확하게 구분하고 실시간으로 밝기를 균일하게 조정하여 다양한 문서 환경에 유연하게 적용될 수 있으며, 밝기 차이가 큰 이미지에서 특히 유용하다.

#### IV. 결론 및 향후 연구

본 논문에서는 장애인 주차 스티커 위조 탐지를 위한 기반 기술로 OCR과 이미지 처리 방법에 대해서 알아보았다. 본 논문에서의 다양한 OCR 엔진은 다국어 지원, 복잡한 레이아웃 처리, 손글씨 인식 등에 특화된 성능을 발휘하며, 특정 응용에 적합한 선택이 가능하다. 이미지 처리 기술은 전통적 기법과 딥러닝 기반 모델 등으로 나뉘며, 각각의 장점을 통해 OCR 성능을 향상시킨다. 전통적 처리 기법은 기본적인 전처리를 제공하며, GAN과 YOLO와 같은 딥러닝 모델은 복잡한 배경에서도 문자 인식률을 크게 높여준다. 이러한 분석 결과는 장애인 주차 스티커

위조 탐지에서 정확성과 신뢰성을 높일 수 있는 기술 조합을 제시할 수 있을 것으로 본다.

향후 연구에서는 본 논문에서 고찰한 OCR과 이미지 처리기술을 적용하여 장애인 주차 스티커 위조 탐지 시스템을 설계하고 구현할 계획이다.

#### 참고 문헌

- [1] “장애인 주차증? 그냥 프린트해 써요”, 근절 안되는 암체주차족, <https://news.zum.com/articles/88264097>, 2024.10.25
- [2] “장애인 주차 위반 신고 불법위조 스티커 찾아내는 요령 5가지”, <https://kisstheguitar.com/장애인-주차-위반-신고/>, 2024.10.25
- [3] 김동영 외 11명, “금융 특화 딥러닝 광학문자인식 기반 문서 처리 플랫폼 구축 및 금융권 내 활용”, 한국지능정보시스템학회, vol.29, no.1, pp. 143-174, 2023
- [4] 김정섭 외 2명, “오픈소스 기반 OCR의 한국어 인식성능비교분석에 관한 연구”, 한국정보처리학회, 30권 2호, pp. 471-472, 2023.
- [5] 스노버거 예런, 이충호, “한글 필기체 OCR의 기술과 데이터셋에 대한 체계적 문헌 고찰”, 삶의 질연구회, 제1권 제2호, pp. 27-38, 2023.
- [6] 민기현 외 5명, “딥러닝 기반 광학 문자 인식 기술 동향”, 한국전자통신연구원, 전자통신동향분석 제37권 제5호, 2022.
- [7] 김원준 외 2명, “특허와 논문정보를 활용한 OCR 기술발전 동향예측에 관한연구”, 한국IT서비스학회지, 제21권 제6호, 2022.
- [8] 김성연 외 4명, “Object Detection과 OCR을 활용한 다중 책 제목 검출 모델”, 한국정보기술학회, Vol. 21, No. 2, pp. 1-8, Feb 2023
- [9] 김하운 외 3명, “OCR 시스템에서 YOLOv4를 활용한 텍스트 이미지 전처리 연구”, 한국정보처리학회, 2022.
- [10] 이준, 유민선, “OCR을 통한 비전자기록물의 전자화 연구 - 딥러닝을 중심으로”, 디지털문화아카이브지, Vol. 7, No. 1, pp. 259-276, 2024.
- [11] 백진성 외 3명, “지능형 OCR 시스템을 위한 한글 필기체 생성 및 분류 모델에 관한 연구”, 융합신호처리학회, Vol. 23, No. 4, pp. 222-227, 2022.

# 융합인문학 데이터셋 기반 챗봇 설계 및 구현

정용태\*, 김선혁\*, 홍윤기\*, 김수환\*\*, 김현기\*\*\*

## Designing and implementing a chatbot based on a fusion humanities dataset

Yong Tae Jung\*, Sun Hyeok Kim\*, Yun Ki Hong\*, Su Hwan Kim\*\*, and Hyen Ki Kim\*\*\*

### 요약

본 연구에서는 융합인문학 데이터셋을 기반으로 한 챗봇 설계 및 구현 방법을 제안하였다. 초기 챗봇은 단순 규칙 기반 시스템으로 한정된 대화 기능만을 제공했으나, 본 연구는 자연어 처리와 인문학적 지식을 결합하여 더욱 인간적인 상호작용을 지원하는 지능형 챗봇을 구현하고자 하였다. 이를 위해 고전 문학, 철학적 사상, 역사적 사건, 사회적 관습 등을 포함하는 인문학적 데이터를 수집 및 전처리하여, 감정과 사회문화적 맥락을 반영한 대화가 가능하도록 GPT-3.5 Turbo 모델을 파인튜닝하였다. 이러한 챗봇은 특정 감정이나 시대에 따른 문체를 설정할 수 있는 기능을 통해 사용자와 보다 깊이 있는 대화를 가능하게 하였으며, 교육, 상담, 문화 콘텐츠 제공 등 다양한 분야에서 응용 가능성을 보였다. 본 연구는 융합인문학적 접근을 통한 챗봇 개발의 유의미한 가능성을 확인하였다.

### Abstract

In this study, a chatbot design and implementation method based on the convergence humanities dataset was proposed. Early chatbots provided only limited conversation functions as a simple rule-based system, but this study attempted to implement an intelligent chatbot that supports more humane interaction by combining natural language processing and humanities knowledge. To this end, the GPT-3.5 Turbo model was fine-tuned to enable conversation reflecting emotions and socio-cultural contexts by collecting and pre-processing humanities data including classical literature, philosophical ideas, historical events, and social customs. These chatbots enabled more in-depth conversations with users through the function of setting a specific emotion or style according to the times, and showed application potential in various fields such as education, counseling, and cultural content provision. This study confirmed the meaningful possibility of developing chatbots through a convergent humanities approach.

### Key words

LLM, chatbots, AI, and natural language processing (NLP)

---

\* 국립안동대학교 멀티미디어공학과,  
\*\* 국립안동대학교 유럽문화관광학과,  
\*\*\* 국립안동대학교 소프트웨어융합학과, hkkim@anu.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원에서 지원하는 SW중심대학사업(IITP-2019-0-01113)의 연구 결과로 수행되었음

## I. 서 론

챗봇 기술은 자연어 처리와 기계 학습의 발전을 토대로 인간과 기계 간 상호작용 방식을 새롭게 정의하고 있다. 초기 챗봇은 단순한 규칙 기반 시스템으로 미리 정의된 답변만을 제공했으나, 최근 딥러닝(Deep-Learning)과 빅데이터(Big-Data)의 발전으로 인해 더욱 지능적이고 유연한 대화가 가능해졌다. 현재의 챗봇은 단순 정보 제공을 넘어 고객 서비스, 의료 상담, 교육, 심리 치료 등 다양한 분야에서 인간의 역할을 보조하거나 대체할 잠재력을 보여주고 있으나, 인간의 복잡한 감정과 사회문화적 맥락을 완전히 이해하지 못하는 한계가 있다. 이는 주로 챗봇이 통계적 패턴과 언어 모델에 기반하여 작동하기 때문에 인간의 깊은 사고 과정이나 문화적 배경, 윤리적 판단 등을 반영하기 어렵기 때문이다[1]. 따라서 보다 인간적인 상호작용을 제공하려면 인문학적 지식과 통찰을 통합하는 지속적인 연구가 필요하다.

융합인문학은 인문학과 자연과학, 사회과학 등 다양한 학문 분야의 경계를 허물고 지식을 통합하여 새로운 이해와 해결책을 모색하는 학문적 접근법이다. 해당 접근법은 복잡한 사회 문제나 인간의 행동을 이해하기 위해 다각적인 시각이 필요하다는 인식에서 비롯되었으며, 문학, 철학, 역사, 심리학 등의 인문학적 지식과 기술적 이해를 결합하여 새로운 형태의 기술 개발을 촉진할 수 있다[2].

챗봇 설계에 융합인문학적 접근을 적용하면 단순한 언어적 응답 생성을 넘어 대화 상대의 감정 상태, 문화적 배경, 윤리적 고려 사항 등을 챗봇이 이해하고 반영할 수 있다. 철학적 질문에 대한 답변이나 문학적 인용을 활용해 대화를 풍부하게 만들 수 있고, 심리학적 지식을 통해 사용자의 감정 변화를 감지하고 적절히 대응할 수 있으며, 역사적 사례나 문화적 맥락을 고려한 대화로 사용자에게 더 깊은 공감과 이해를 제공할 수 있다[3]. 그러나 융합인문학적 기능 구현을 위해서는 다양한 인문학적 지식과 데이터를 포함한 융합인문학 데이터셋이 필요하다. 해당 데이터셋은 문학 작품, 철학적 이론, 역사적 사건, 문화적 관습 등 광범위한 인문학적 정보를

포함하며, 챗봇이 다양한 상황에서 적절하고 의미 있는 대화를 생성하는 데 필수적인 자원을 제공한다[4]. 이를 활용하면 챗봇은 단순한 정보 전달자에서 벗어나 사용자의 내면적 요구와 사회적 맥락을 이해하는 지능형 대화 상대로 발전할 수 있다.

본 논문에서는 융합인문학 데이터셋을 기반으로 한 챗봇의 설계 및 구현 과정을 상세히 기술하였다. 챗봇이 인간의 복잡한 감정과 사고를 이해하고, 인문학적 관점을 바탕으로 의미 있는 상호작용을 창출할 수 있는 방법을 탐구한다. 특히 자연어 처리 기술과 인문학적 지식의 통합 방법, 데이터셋의 구성과 활용 방안, 그리고 실제 구현된 챗봇의 성능 평가 등을 다룰 것이다.

## II. 관련 연구

최근 인공지능(AI)과 머신러닝(ML) 기술의 발전은 다양한 산업에 혁신을 불러일으키고 있으며, 특히 자연어 처리(NLP) 분야에서 그 성과가 두드러지고 있다[5]. 해당 기술은 인문학과 융합하여 새로운 응용 분야를 개척하고 있다. 융합인문학 기반 챗봇은 단순한 정보 제공을 넘어서 인간의 감정, 사회적 맥락, 문화적 배경 등을 이해하고 이에 맞춰 대화를 진행하는 고도화된 AI 시스템을 목표로 한다. 최근 언어 모델의 발전은 이와 같은 챗봇 설계에서 핵심적인 역할을 하고 있다[6]. 특히 GPT-3.5 Turbo와 같은 최신 대규모 언어 모델은 다양한 언어 처리 작업에서 매우 높은 성능을 발휘하며, 이러한 융합 챗봇 시스템에서 중요한 역할을 하고 있다. GPT-3.5 Turbo는 인간 수준의 자연어 이해 및 생성 능력을 갖추고 있으며, 다양한 응용 분야에서 그 성능이 입증되었다. 본 모델은 융합인문학의 맥락을 반영한 대화 시스템에서 복잡한 질문과 감성적 소통을 처리할 수 있다. 예를 들어, 역사적 사건에 대한 질문을 단순히 사실 기반으로 답변하는 것이 아니라, 당시 사회적, 문화적 배경을 고려한 답변을 제공할 수 있다. 이와 같은 챗봇은 교육, 상담, 문화 콘텐츠 제공 등 다양한 분야에서 응용 가능성이 높으며, 인문학적 지식을 AI 시스템에 접목함으로써 인간과 더욱 자연스럽게 상호작용할 수 있는 기회를 제공할

다. GPT-3.5 Turbo는 이러한 융합형 챗봇 설계에서 처리 속도와 성능 면에서 주목받고 있으며, 다양한 작업에서 고성능을 발휘한다. 본 모델의 강점은 대규모 데이터 처리 능력과 정교한 언어 처리 능력에 있으며, 이를 통해 복합적인 질문에 대한 적절한 응답을 생성할 수 있다.

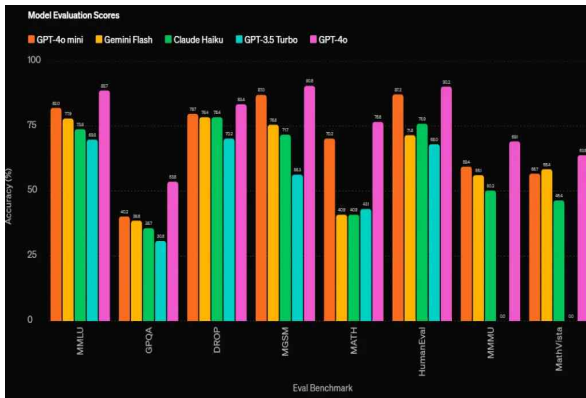


그림 1. GPT-4 및 GPT-3.5 Turbo를 포함한 다양한 모델의 성능 비교

Fig. 1. Compare performance of different models, including GPT-4 and GPT-3.5 Turbo

그림 1은 GPT-4, GPT-3.5 Turbo, Claude Haiku, Gemini Flash, 그리고 GPT-4.0 mini 모델의 성능을 다양한 벤치마크 데이터셋을 통해 비교한 결과를 보여준다[7]. 각 모델은 MMLU, GPOA, DROP, MGSM, MATH, HumanEval 등 여러 평가 지표에서 분석되었으며, GPT-3.5 Turbo는 융합인문학 기반의 언어 처리 작업에서도 높은 성능을 기록하였다. 특히 해당 결과는 대규모 언어 모델이 단순한 텍스트 생성 작업을 넘어서 인문학적 통찰을 반영한 대화 시스템에서 얼마나 중요한 역할을 할 수 있는지를 보여준다. 융합인문학 기반 챗봇 설계에서의 이러한 연구는 인문학적 지식을 기계 학습에 반영하여 사람과의 대화를 더욱 자연스럽게 의미있게 만들 수 있음을 시사한다. 향후 챗봇 시스템은 교육, 문화 콘텐츠 제공, 심리 상담 등 다양한 분야에서 인간과 AI의 상호작용을 혁신적으로 변화시킬 것으로 기대된다. 그림 1에 나타난 성능 지표는 다양한 언어 모델의 성능을 시각적으로 비교하며, 각 모델이 특정 작업에서 얼마나 뛰어난 성능을 발휘하는지를 확인할 수 있는 좋은 사례이다.

### III. 시스템 설계 및 구현

#### 3.1 데이터 수집

본 논문에서는 챗봇의 설계 및 학습을 위해 다양한 인문학적 지식과 사회문화적 정보를 포함하는 데이터셋을 구축하였다. 이를 위해 공공데이터 포털에서 고전 및 현대 문학 작품에서 인간 감정과 관계의 상호작용을 다루는 텍스트, 윤리적 판단과 도덕적 딜레마와 같은 철학적 이론 및 사상 자료, 특정 시대와 지역의 역사적 사건에 대한 설명, 그리고 현대 사회에서 빈번하게 논의되는 주제와 사회적 관습을 포괄하는 데이터를 수집하였다. 또한, AI-hub에서 제공하는 감정이 태깅된 대화 데이터를 추가하여 챗봇이 사용자 감정에 적절히 대응할 수 있도록 구성하였다. 이와 같이 다양한 출처로부터 수집된 데이터셋을 통해 챗봇은 인문학적 배경 지식을 학습하며, 사용자 질문에 인문학적 통찰을 반영한 응답을 제공할 수 있게 된다.

#### 3.2 데이터 전처리

본 논문에서는 수집된 데이터셋을 챗봇 모델 학습에 적합한 형태로 가공하기 위해 데이터 전처리 과정을 수행하였다. 먼저 텍스트 정제를 통해 불필요한 특수 문자와 중복된 표현을 제거하고, 문장의 일관성을 유지하기 위해 잘못된 문장 구조를 수정하였다. 이후 토큰화 작업을 통해 데이터를 어휘 단위로 분할하여 모델이 문장의 문법적 구조를 이해할 수 있도록 하였으며, 각 문장에는 감정 상태, 문화적 맥락, 철학적 배경 등을 반영한 라벨을 추가하여 상황에 맞는 응답을 생성할 수 있도록 구성하였다. 또한, 데이터의 일관성을 확보하기 위해 텍스트 정규화 및 불용어 제거를 수행하여 학습의 정확도를 높이고자 하였다.

#### 3.3 인공지능 학습

본 논문에서는 수집된 데이터셋을 바탕으로 OpenAI의 GPT 모델을 파인튜닝하여 챗봇이 인문학

적 지식과 감정적 반응을 효과적으로 반영할 수 있도록 학습하였다. 먼저 최신 대규모 언어 모델 중 챗봇의 목적에 부합하는 GPT-3.5 Turbo를 선정하고, 융합 인문학 데이터셋으로 파인튜닝을 진행하였다. 감정 상태와 사회적 맥락에 따라 적절한 응답을 생성할 수 있도록 다양한 상황에 맞는 데이터를 학습 시키고, 이를 통해 사용자 질의에 대해 풍부한 인문학 통찰을 반영한 대화를 제공할 수 있도록 구성하였다.

### 3.4 시스템 구현 결과

본 논문에서는 융합인문학 기반 챗봇을 디스코드 봇으로 구현하였다. 이 챗봇은 이미 수집 및 파인튜닝 된 융합인문학 데이터셋을 활용하여 OpenAI의 GPT-3.5-turbo 모델로 응답을 생성하도록 설계되었다. 기본적으로 일반 GPT 모델처럼 동작하나, 사용자가 특정 감정이나 시대에 따른 어투와 어조를 명령어로 지정하면 그에 맞는 방식으로 응답할 수 있도록 구성하였다. 예를 들어 '!기쁨', '!슬픔', '!조선시대', '!근현대시대' 등의 명령어를 입력하면 해당 감정이나 시대적 문체를 반영하여 챗봇이 응답을 생성하도록 하였다.

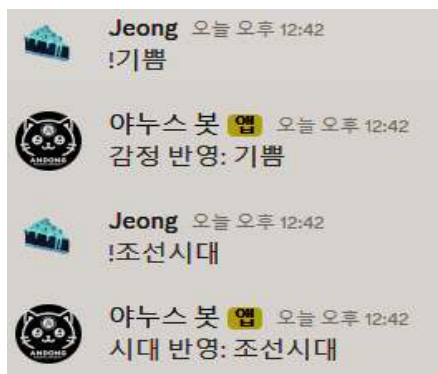


그림 2. 챗봇 감정 및 시대반영 예시  
Fig. 2. Examples of Chatbot Emotions and Times Reflections

디스코드 봇은 OpenAI API를 활용하여 비동기적으로 응답을 생성하며, 다양한 예외 처리를 통해 API 호출 중 발생할 수 있는 오류를 처리하였다. 또한, 'set\_emotional\_tone' 함수와 'adjust\_style\_by\_era' 함수는 사용자가 요청한 감정이나 시대적 스타일에

따라 응답의 어조를 조정하도록 설정하여, 융합 인문학 요소를 반영한 대화가 가능하게 하였다. 이러한 기능은 GPT 모델의 응답이 더욱 유의미하게 다가가도록 하여, 사용자가 감정과 시대에 따른 다양한 문체로 챗봇의 대화를 즐길 수 있게 한다.

"아, 이 얼마나 즐겁고 흐뭇한 광경이옵니까! 김홍도의 이른바 \*\*「서당」\*\*이라 하는 작품 속엔 서당의 아이들이 다채로운 모습으로 앉아 있사옵데, 한 아이는 엷드려 즐고, 또 한 아이는 판생각에 빠져 있으니, 이를 바라보는 소인의 마음이 웃음으로 가득 차옵니다.

저 훈장님께서 진지한 눈빛으로 아이들을 가르치시나, 아이들은 저마다 다른 자세로 앉아 그리도 천진난만하게 시간을 보내고 있사옵니다. 김홍도 어른님께서 여짜 이리도 서민의 일상을 사실적으로 그려낼 수 있었는지, 참으로 경탄할 따름이옵니다. 서당이라는 조선의 배움터가 이렇듯 정겹고 생동감 넘치는 모습이라니, 이를 보는 소인 또한 마음이 한층 밝아지옵니다!"

그림 3. 김홍도의 서당의 작품을 설명한 예시  
Fig. 3. An example of the work of Kim Hong-do's Seodang

이를 통해 융합 인문학 챗봇은 사용자가 감정이나 시대적 문체를 지정하여 다양한 인문학적 질문에 접근할 수 있는 인터페이스를 제공하며, 일반 GPT의 기능을 바탕으로 인문학적 표현의 깊이를 더하였다.

## IV. 결 론

본 논문에서는 융합인문학 데이터셋을 기반으로 한 챗봇 설계 및 구현 과정을 통해 인간적인 대화와 감성적인 소통이 가능한 디스코드 챗봇을 구현하였다. 이를 위해 고전 문학, 철학적 사상, 역사적 사건, 사회적 관습 등 인문학적 자료를 폭넓게 수집하고 이를 전처리하여 OpenAI GPT-3.5-turbo 모델에 파인튜닝하였으며, 감정과 시대적 맥락에 따른 응답을 생성할 수 있는 특화된 기능을 개발하였다.

챗봇은 단순한 정보 전달을 넘어 사용자의 감정 상태와 질문 맥락을 반영하여 더욱 인간적인 상호작용을 제공하며, '기쁨', '슬픔' 등의 감정이나 '조선시대', '근-현대시대'와 같은 시대별 문체를 명령어로 설정할 수 있는 기능을 통해 보다 깊이 있는 대화를 가능하게 하였다. 이러한 구현은 기존 챗봇 시스템의 한계를 보완하고, 인문학적 지식을 기계 학습에 반영함으로써 인문학적 통찰을 담은 자연스러운 대화 경험을 제공하는 데 의미가 있다.

융합인문학 기반 챗봇은 교육, 문화 콘텐츠 제공,

심리 상담 등 다양한 분야에서 폭넓게 응용될 가능성이 높으며, 인간의 복잡한 감정과 사회문화적 맥락을 이해하는 지능형 대화 상대로서의 역할을 기대할 수 있다. 향후 연구에서는 더 다양한 시대적 자료와 감정에 따른 세분화된 응답 생성 기능을 추가함으로써 챗봇의 상호작용 범위를 확장하고, 사용자 만족도를 높이는 방향으로 발전시킬 수 있을 것이다.

### 참 고 문 헌

- [1] Yoon, S. O. (2018). A Study on the Key Issues of AI-based Public Services. *Korean Public Management Review*, 32(2), 83-104.
- [2] Jeon, J. H. (2017). The Unity of Knowledge and Mutual Consilience - For the Coexistence of Humanities and Natural Sciences. *Philosophy • Thought • Culture*, (23), 216-240.
- [3] Park, S. E. (2021). Analysis of the Current Status of Deep Learning-centered Natural Language Processing Technology. *Journal of the Korean Big Data Society*, 6(1), 63-81.
- [4] Kim, H. Y. (2016). On Big Data and Humanities. *Citizen and Humanities*, 30(0), 70-95.
- [5] Kim, J., & Lee, S. (2020). The Impact of Chatbot Service Quality on Customer Experience: Evidence from the Service Industry. *Journal of the Korean Society for Quality Management*, 47(4), 755-770.
- [6] Lin, C.-C., Huang, A. Y. Q., & Yang, S. J. H. (2023). A Review of AI-driven Conversational Chatbots: Implementation Methodologies and Challenges (1999 - 2022). *Sustainability*, 15(5), 4012.
- [7] Koubaa, A. (2024). GPT-4 vs. GPT-3.5: A Concise Showdown. Prince Sultan University.

# Polingo : 다중 입력 기반 통합 번역 및 학습 플랫폼 개발

김민석\*, 이태헌\*, 정유철\*\*

## Polingo: Development of an Integrated Translation and Learning Platform Based on Multiple Inputs

Minseok Kim\*, Taeheon Lee\*, and Yuchul Jung\*\*

### 요약

본 논문은 해외여행 시 언어 장벽을 해결하기 위해 다중 입력 기반의 통합 번역 및 학습 플랫폼 Polingo를 소개한다. Polingo는 이미지와 음성 번역을 포함한 다양한 입력 방식을 지원하며, 학습 콘텐츠를 인공지능을 통해 자동 생성해 사용자 외국어 학습을 돕는다. Polingo는 번역과 학습 기능을 통합하여 편리한 언어 학습 환경을 제공하며, 글로벌 사용자들이 언어 장벽을 효과적으로 극복할 수 있는 솔루션을 제시한다.

### Abstract

This paper introduces Polingo, a multi-input integrated translation and learning platform designed to overcome language barriers during international travel. Polingo supports diverse input methods, including image and voice translation, and leverages artificial intelligence to automatically generate learning content that assists users in foreign language acquisition. By combining translation and learning functionalities, Polingo provides a convenient language learning environment and offers a solution for global users to effectively overcome language barriers.

### Key words

Multi-Input Translation, Language Learning, Large Language Model(LLM), Optical Character Recognition (OCR)

## I. 서론

해외여행 중에 여행객이 현지 언어로 된 메뉴판을 이해 못하거나, 외국인과의 소통의 어려움을 겪는 경우 등 사람들은 번역기 사용과 외국어 학습을 해야 하는 환경에 놓여지게 되었다. 이러한 환경에서 이미지 번역, 음성 번역 등 다양한 입력에 대한 여러 국가의 언어 번역을 제공하고, 상황에 맞는 예문과 학습을 위한 콘텐츠를 제공함으로써 이러한 언

어 장벽 문제를 해결하고자 한다.

## II. 기능 소개

본 프로젝트는 크게 로그인 및 사용자 설정, 텍스트 번역, Google STT API를 이용한 음성 번역, OCR 기법을 사용한 이미지 번역, Daily News를 기반으로한 뉴스 학습 콘텐츠 제공, 나만의 단어장, 나만의 단어장을 기반으로 동적으로 문제를 생성하

\* 국립금오공과대학교 컴퓨터공학과, 20190938@kumoh.ac.kr, kms6290@kumoh.ac.kr

\*\* 국립금오공과대학교 인공지능공학과, jyc@kumoh.ac.kr, 교신저자



는 나만의 문제집, 그리고 날씨, 시간, 요일, 공휴일, 방문 장소에 따른 적절한 예문을 제공하는 상황별 예문 기능을 포함하고 있다. 현재 본 프로젝트가 지원하는 언어는 2가지(영어, 일본어)로, 추가적인 개발을 통해 언어를 편리하게 추가할 수 있도록 확장성을 고려하여 개발되었다. 마지막으로 뉴스 학습 콘텐츠나 상황별 예문, 단어와 같이 사용자에게 제공되는 콘텐츠를 추가 및 수정할 수 있는 관리자 페이지(Back Office) 또한 개발되었다.

### III. 시스템 설계

polingo 프로젝트는 크게 사용자 인터페이스(UI), 서버, 학습 콘텐츠 생성기, 이미지 처리 서버, LLM(Large Language Model) 서버, 관리자 웹 페이지로 구성되어 있으며, 이미지 처리 서버와 LLM 모델 서버는 서버의 특성상 TPS(Transaction Per Second)가 떨어지기 때문에 정적 Scale Out을 통해 TPS를 올리도록 하였다. 이 외에 직접 개발하기 어렵거나, 외부에서 제공하는 기능은 외부 API를 사용해 개발 비용을 줄이는 데 노력하였다. 본 프로젝트의 전체 시스템 구조를 도식화 하면 그림 1과 같다.

### IV. 서버 개발 내용

서버는 모바일 어플리케이션과 관리자 웹 어플리

케이션과 통신을 수행하여 도메인 로직 수행, DB 저장 등 본 프로젝트에서 가장 핵심적인 모듈로, HTTP 1.1 프로토콜을 사용해 통신을 수행한다. 서버는 JAVA 언어로 개발되었으며, JAVA의 여러 웹 프로그램 개발 프레임 워크 중 보편적으로 많이 사용되는 Spring Web MVC를 기반으로 개발되었다. Database는 기본적으로 관계형 데이터베이스인 MySQL을 사용하며, 검색 로직이 복잡한 기능에 대해서는 부분적으로 검색엔진(ElasticSearch)를 사용하여 성능을 개선하였다.

사용자로부터 단어 검색이 가장 많이 발생할 것이라 판단하여 해당 기능을 비동기 처리를 통해 성능 개선을 진행하였다. ElasticSearch와 DeepL 단어 번역을 동시 호출하여 ElasticSearch에 데이터가 없는 경우 DeepL 번역 결과를 사용하도록 하였다. 이러한 처리를 통해 ElasticSearch 조회 후 DeepL을 사용하던 기존의 동기적인 처리방식에서 응답 시간을 30% 줄일 수 있었다.

### V. LLM 서버

LLM 서버는 콘텐츠 내 영/일 문장에 대한 문법 분석을 수행하는 서버로, LLAMA3 및 mistral 모델을 사용하였다. 사용자에게 더욱 적절한 답변을 생성하기 위해 프롬프트에 역할을 부여하거나 프롬프트 엔지니어링 기법을 적용하였다. 프롬프트 엔지니어링 기법은 Few-Shot 프롬프팅을 적용하였다.

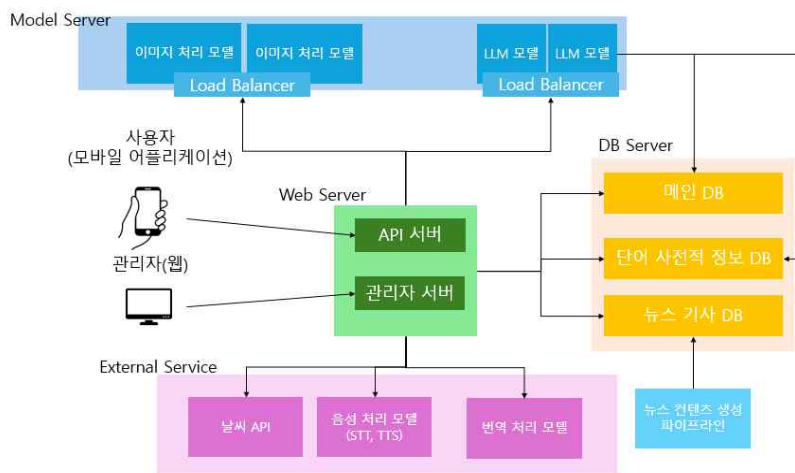


그림 1. 전체 시스템 구조  
Fig 1. Overall System Architecture

이를 통해 기존의 존재한 모델을 사용자에게 문장에 대한 문법적 요소를 분석해주는 현 시스템의 취지에 적합하게 수정 및 보완하여 성능을 향상시켰다. 표 1은 LLM 서버의 문장 분석 응답 예시를 표로 나타낸 것이다

표 1. LLM 서버의 문장 분석 응답 예시  
Table 1. Example of LLM Server Sentence Analysis Response

Input	“The early bird catches the worm”.
Results of LLM Sentence Analysis	“주어 + 동사 + 능동 구성: 문장은 능동태 구조를 사용합니다. 'The early Bird'가 주어 역할을 하고 'catches'가 능동 동사구 역할을 합니다. 이 구성은 동작 자체보다는 동작을 수행하는 사람을 강조합니다. '전치사 문구: 'the worm'은 동작의 대상에 대한 추가 정보를 제공합니다. 일찍 일어나는 새가 무엇을 잡는지 설명해줍니다.”

히, 이미지와 음성을 포함한 다양한 번역 방식을 지원함으로써 여행자들이 현지에서 겪는 언어 문제를 실질적으로 해결할 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1] Spring Documents, <https://spring.io/docs>
- [2] llama-3 Source code  
<https://github.com/meta-llama/llama3>
- [3] React Native Documents,  
<https://reactnative.dev/docs/getting-started>
- [4] Easy-OCR Documents,  
<https://www.jaided.ai/easyocr/documentation/>

## VI. 결 론

본 논문에서는 해외여행 중 언어 장벽을 효과적으로 극복할 수 있는 통합 번역 및 학습 플랫폼인 Polingo의 설계와 구현에 대해 논의하였다. Polingo는 이미지 및 음성 번역을 포함한 다양한 입력 방식을 지원하며, 사용자에게 상황에 맞는 학습 콘텐츠를 제공함으로써 단순한 번역을 넘어선 언어 학습 환경을 제공한다. React Native와 Spring Web MVC 기반으로 구축된 이 플랫폼은 비동기 처리와 Elasticsearch를 활용하여 성능을 최적화하였으며, LLM 서버를 통해 사용자에게 맞춤형 문법 분석 기능을 제공하였다.

Polingo는 기존 번역 서비스와 달리 학습 콘텐츠와 번역 기능을 통합하여, 사용자가 번역뿐만 아니라 외국어 학습까지 하나의 플랫폼에서 경험할 수 있게 한다. 이러한 기능적 차별화는 언어 장벽 해소뿐 아니라, 글로벌 사용자들이 새로운 언어에 대한 접근성을 높이는 데 기여할 수 있음을 기대한다. 특

# 감정인식 학습지원 플랫폼의 설계

차현우\*, 정용성\*\*, 안수윤\*\*, 정윤주\*\*\*, 김현기\*\*

## Design of an Emotion Recognition Learning Support Platform

H. W. Cha\*, Y. S. Jeong\*\*, S. Y. Ahn\*\*, Y. J. Jeong\*\*\*, and H. K. Kim\*\*

### 요 약

본 논문은 학습자의 감정 상태를 실시간으로 인식하고, 이를 기반으로 맞춤형 학습 지원을 제공하는 감정인식 학습지원 플랫폼을 설계하였다. 기존 교육 플랫폼이 학습 콘텐츠와 활동 데이터에 초점을 맞춘 것에 비해, 본 연구의 플랫폼은 학습자가 표현하는 감정을 분석하여 개인화된 피드백과 학습 계획을 제공함으로써 학습자의 몰입도와 학습 효과를 증진시키고자 한다. 감정인식 모듈에서는 학습자의 표정과 음성 신호를 종합적으로 분석하며, 학습지원 모듈은 이 데이터를 바탕으로 적절한 피드백을 제공하고 학습 전략을 조언할 수 있도록 한다.

### Abstract

This paper designs an emotion recognition-based learning support platform that detects learners' emotional states in real-time to provide personalized learning assistance. Unlike conventional education platforms focused mainly on content and activity data, this platform analyzes learners' expressed emotions to offer tailored feedback and learning plans, thereby enhancing engagement and learning outcomes. The emotion recognition module comprehensively analyzes learners' facial expressions and vocal signals, while the learning support module uses this data to deliver appropriate feedback and recommend effective learning strategies.

### Key words

Emotion Recognition, Personalized Learning, Real-time Feedback, Learning Engagement, Facial and Vocal Analysis

## 1. 서 론

디지털 기술의 발전으로 맞춤형 학습 환경을 지원하는 다양한 교육 플랫폼이 등장하고 있다. 온라인 강의, 모바일 앱, AI 기반 학습 시스템 등은 학습자의 요구와 특성에 맞춘 교육을 가능하게 하여 교육의 큰 변화를 이끌었다. 그러나 대부분의 기존 플랫폼은 학습 콘텐츠와 활동 데이터에 집중할 뿐,

학습자의 감정 상태나 몰입도와 같은 실시간 데이터를 활용하는 데에는 여전히 한계가 있다.

감정인식 기술의 발전으로 학습자의 감정 상태를 실시간으로 파악하고, 이를 통해 개별화된 학습 지원을 제공하는 시도가 증가하고 있다. 학습 과정에서 학습자의 감정을 분석하여 맞춤형 피드백과 학습 지원에 활용하면, 학습자의 몰입도와 학습 효과를 크게 높일 수 있을 것으로 기대된다.

\*국립안동대학교 멀티미디어공학과, \*\*국립안동대학교 소프트웨어융합학과, \*\*\*국립안동대학교 SW전문인재양성사업단, hkkim@anu.ac.kr(교신저자)

※본 연구는 과학기술정보통신부 및 정보통신기획평가원에서 지원하는 SW중심대학사업(IITP-2019-0-01113)의 연구 결과로 수행되었음

이에 본 연구는 학습자의 감정을 실시간으로 인식하고 분석하여, 개별 학습자의 학습 상태에 최적화된 피드백과 지원을 제공하는 감정인식 학습지원 플랫폼을 제안한다. 이를 통해 학습자의 몰입도와 학습 효과를 향상시킬 수 있는 새로운 맞춤형 교육 플랫폼의 가능성을 탐색하고자 한다.

## II. 관련 연구

감정 인식 또는 감정 분석 기술은 개인의 심박수, 눈 움직임, 표정, 피부 수분, 보행 추적, 목소리 톤 또는 기타 신체적 특징을 스캔한 생체 데이터를 해석해 정신 상태를 추론하거나 누군가의 느낌을 예측하는 AI 기반 기술을 말한다[1].

### 2.1 감정 인식 기술

IT기술의 발달로 개인화 서비스 기술이 인기를 끌게 되면서, 인간의 감정 상태를 판단하기 위한 감정인식 기술[2]이 다양한 분야에서 활용되고 있다.

얼굴 이미지를 사용한 감정인식 연구는 주로 얼굴 표정의 미세한 변화를 분석하여 감정 상태를 파악하는 데 중점을 둔다. 최근에는 심층 신경망과 컴퓨터 비전 기술을 활용하여 얼굴 이미지를 실시간으로 분석하는 연구가 활발히 진행되고 있으며, CNN(Convolutional Neural Network)과 같은 모델이 많이 사용된다. CNN은 얼굴 표정의 특징을 추출하고 감정을 분류하는 데 유용하며, VGGFace, ResNet, Inception 등의 모델이 특히 높은 인식률을 보인다[3].

또한, 감정이 얼굴 표정의 시간적 변화에 따라 달라지는 특성을 고려하여, 연속된 얼굴 이미지를 분석하는 트랜스포머 모델도 주목받고 있다. 오문흠 등은 트랜스포머 기반의 판소리 소리꾼 검출과 얼굴 표정 인식 연구에서 판소리 비디오에서 얼굴 표정을 자동으로 분류하였다[4].

음성을 통한 감정 인식 기술은 음성에서 추출된 다양한 특징(예: 주파수, 음높이, 속도, 강도)을 분석하여 감정을 파악하는 방식이다. 일반적으로 감정 인식을 위해 음성 데이터를 스펙트로그램으로 변환한 후, 딥러닝 모델(예: CNN, RNN)을 활용하여 음성의 감정적 특징을 추출한다. 이러한 기술은 슬픔, 분노, 기쁨, 중립 등 다양한 감정 상태를 분류하는데 효과적이며, 실시간 피드백 시스템 및 상담 지

원 시스템 등에 폭넓게 활용된다.

### 2.2 인공지능을 활용한 교육

AI를 활용한 교육은 학습자 맞춤형 학습 경로 제공, 학습 성과 예측, 교사 보조 등의 기능을 통해 교육을 혁신하는 방식이다. 예를 들어, AI는 학습자의 학습 속도와 성향을 분석하여 개별 맞춤형 학습 콘텐츠를 추천하며, 학습자의 이해도를 실시간으로 추적해 적절한 피드백을 제공한다. 또한, 학습 데이터 분석을 통해 학습자의 성과와 학습 경로를 예측하여, 보다 효과적으로 학생의 학습능률을 향상시킬 수 있도록 한다.

한컴로보틱스에서 개발한 토키(Toki)는 음성 인식과 얼굴 인식 기능을 통해 학습자와 상호작용할 수 있는 유아교육용 인공지능 로봇이다. 토키는 학습자가 질문을 반복적으로 할 때 이전 대화를 고려해 다르게 답변하며, 학습자가 휴식을 취할 때나 예상치 못한 상황에 맞춰 맞춤형 대화를 제공할 수 있다.

## III. 시스템 설계

본 연구에서의 감정인식 학습지원 플랫폼은 학습자의 감정을 실시간으로 파악하여 맞춤형 피드백과 학습 계획을 제공해주는 시스템이다. 주요 구성 요소와 전체 구조는 다음과 같다.

### 3.1 주요 기능

본 연구의 주요 기능은 다음과 같이 세 가지이다. 첫째 감정인식 학습 도우미는 얼굴 인식을 위해 JavaScript 기반의 face-api.js 라이브러리를 사용한다. 이 라이브러리는 TensorFlow.js 기반으로 얼굴 인식 및 감정 분석을 수행하며, 웹 애플리케이션에서 웹캠을 통해 사용자 얼굴 이미지를 실시간으로 캡처하여 기쁨, 슬픔, 분노 등 주요 감정을 파악하고 이를 통해 학습 환경을 동적으로 조정한다.

둘째, 학습 진행 예측 모델은 Python의 Pandas와 NumPy를 사용해 데이터를 전처리하고, scikit-learn 및 TensorFlow 등 머신러닝과 딥러닝 라이브러리를 활용해 구축된다. 이를 통해 감정 데이터와 학습 데이터를 분석하여, 다음 학습 단계에서 예상되는 성과와 학습자 반응을 예측한다.

셋째, 대화형 AI 학습 멘토는 Google의 T5 모델

을 활용하여 Transformer 기반 자연어 처리 모델을 구현한 것으로, 학습자와 자연스러운 대화를 통해 학습을 지원한다. 학습자의 질문에 응답하고, 학습 과정 중 발생할 수 있는 의문점을 해결하여 학습 효과를 극대화한다.

### 3.2 시스템 구성 및 흐름

본 연구의 시스템 구성 및 흐름은 Fig. 1과 같이 사용자 인증, 감정인식, 데이터베이스 관리, 감정 데이터 분석, 사용자 피드백, 학습 계획 수립, 학습 종료 단계로 설계되었다.

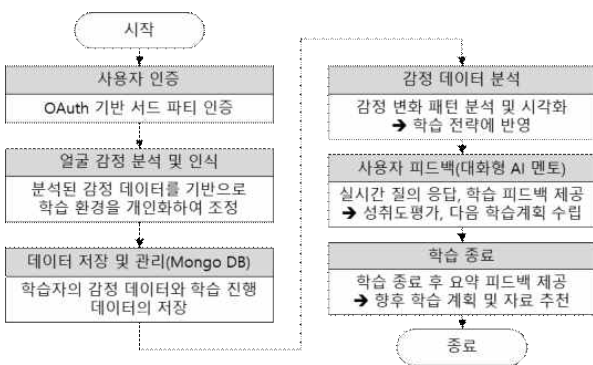


그림 1. 감정인식 학습지원 시스템의 개요도

Fig. 1. Schematic diagram of the emotion recognition learning support system

사용자 인증은 OAuth 기반의 인증 시스템을 통해 Google, Facebook 등 서드 파티 인증을 지원하여 사용자 편의성을 높인다.

감정인식 단계는 face-api.js와 TensorFlow.js를 활용하여 웹캠으로 사용자 얼굴 이미지를 실시간으로 분석하여 표정에서 감정을 식별한다. 여기서 도출된 감정 정보는 학습 환경을 개인화하여 학습자가 효과적으로 몰입할 수 있도록 학습 진행을 조절하는 데 사용된다.

데이터베이스는 MongoDB를 사용하여 감정 데이터와 학습 진행 데이터를 저장 및 관리하여, 사용자별 맞춤형 피드백과 통계적 분석이 가능하게 한다. 효율적인 데이터 관리 구조를 통해 감정 인식 결과와 학습 데이터의 일관성을 유지할 수 있다.

감정 데이터 분석 단계에서는 Python 기반의 데이터 분석 도구(Pandas, Matplotlib 등)를 사용하여 저장된 감정 데이터의 패턴을 분석하고, 감정 변화 추세를 시각화하여 학습자가 학습 중 보이는 감정

변화를 추적한다. 이러한 분석 결과는 학습자에 맞춘 개선된 학습 전략 수립에 활용된다.

사용자 피드백은 대화형 AI 학습 멘토를 통해 실시간으로 제공되며, 학습자는 학습 중 발생하는 질문에 대한 답변을 받을 수 있다. 이를 통해 학습 과정에서의 문제 해결을 지원하고, 학습 성취도를 기반으로 다음 학습 계획을 제안하며, 필요한 학습 자료를 추천하여 학습 동기부여를 강화한다.

이와 같은 시스템 설계는 학습자의 감정 상태에 맞춘 맞춤형 학습 피드백과 동적 학습 계획 수립을 통해 학습 효과를 극대화하는 데 기여할 수 있다.

## IV. 결 론

본 연구에서는 학습자의 감정 상태를 실시간으로 분석하고, 이를 기반으로 맞춤형 학습 지원을 제공하는 감정인식 학습지원 플랫폼을 설계하였다. 이 플랫폼은 학습자의 감정 데이터를 분석하여 학습 진행 상황에 따라 개인화된 피드백과 학습 계획을 수립함으로써 학습자의 몰입도와 학습 효과를 높일 수 있도록 돕는다. 기존의 온라인 교육 시스템이 제공하는 일반적 학습 콘텐츠를 넘어, 학습자 개별의 감정 변화를 반영한 맞춤형 학습 환경을 지원하여 학습 만족도 향상을 기대할 수 있다.

## 참 고 문 헌

- [1] AI 타임스, <https://www.aitimes.com>, (Accessed 2024, October)
- [2] M. J. Seo and M. H. Kim. "Ensemble Method of Emotion Classifier for Speech Emotion Recognition," 한국IT정책경영학회 논문지, Vol. 11, No. 2, pp. 1187-1193, 2019.
- [3] Choksi, Kushan, et al. "SensEmo: Enabling Affective Learning through Real-time Emotion Recognition with Smartwatches." arXiv preprint arXiv:2407.09911, 2024.
- [4] M. H. Oh, H. J. Lee and J. H. Lee, "Transformer-Based Detection and Facial Expression Recognition of a Pansori Singer," The Journal of Korean Institute of Communications and Information Sciences, Vol.47 No.12, 2022.

# 딥러닝 기반의 이미지 분석을 통한 사용자 맞춤형 요리 추천 시스템 설계

윤수환\*, 문현진\*\*, 정윤주\*\*\*, 김현기\*\*

## Design of a Personalized Cooking Recommendation System using Deep Learning-based Image Analysis

S. H. Yoon\*, H. J. Moon\*\*, Y. J. Jeong\*\*\* and H. K. Kim\*\*

### 요 약

본 논문은 YOLO, GAN, BERT와 같은 딥러닝 기법을 활용하여 사용자 맞춤형 레시피 추천 시스템을 개발하는 방법을 제안한다. 사용자가 업로드한 음식 이미지를 YOLO로 분석해 재료를 인식하고, BERT를 통해 텍스트 임베딩을 생성하여 유사도가 높은 레시피를 추천한다. 또한, GAN을 사용해 다양한 음식 이미지를 생성하여 데이터 다양성을 확보하고 추천 정확도를 향상시킨다. 사용자가 보유한 재료를 효율적으로 활용하게 하여 재료 관리와 비용 절감에 기여하고, 다양한 요리 방법을 제안함으로써 요리 지식과 스킬 향상을 돕는다. 또한, AI와 요리의 결합 사례로 교육적 활용 가능성도 제공한다.

### Abstract

This paper proposes a method for developing a personalized recipe recommendation system using deep learning techniques such as YOLO, GAN, and BERT. The system analyzes food images uploaded by users with YOLO to recognize ingredients, then generates text embeddings using BERT to recommend recipes with high similarity. Additionally, GAN is employed to create diverse food images, enhancing data diversity and improving recommendation accuracy. By enabling users to utilize their available ingredients more efficiently, the system contributes to ingredient management and cost savings. Furthermore, it aids in the enhancement of cooking knowledge and skills by suggesting various cooking methods and offers potential for educational applications as an example of AI integration in culinary contexts.

### Key words

YOLO, GAN, BERT, recipe recommendation, personalized cooking

## 1. 서 론

요리에 대한 관심이 증가하면서 많은 사람들이 다양한 요리를 시도하지만, 재료 활용의 어려움과

\*국립안동대학교 AI융합학과, \*\*국립안동대학교 소프트웨어융합학과, \*\*\*국립안동대학교 SW전문인재양성사업단, hkkim@anu.ac.kr(교신저자)

※본 연구는 과학기술정보통신부 및 정보통신기획평가원에서 지원하는 SW중심대학사업(IITP-2019-0-01113)의 연구 결과로 수행되었음

요리 실패로 인해 음식물 쓰레기가 늘고 있다. 이러한 문제를 해결하기 위해 사용자가 보유한 재료를 자동 인식하고 맞춤형 레시피를 추천하는 시스템의 필요성이 부각되고 있다.

본 연구는 YOLO, BERT, GAN과 같은 딥러닝 기법을 활용한 맞춤형 레시피 추천 시스템을 제안한다. 사용자가 업로드한 음식 이미지를 YOLO 모델로 분석하여 재료를 인식하고, BERT 모델로 유사한 레시피를 추천하며, GAN으로 생성된 이미지를 추가해 데이터 다양성을 확보한다. 이를 통해 사용자는 보유한 재료를 기반으로 최적의 요리 방법을 손쉽게 찾을 수 있다.

이 시스템은 요리 실패를 줄이고 음식물 낭비를 감소시키는 데 기여할 수 있다. 또한, 본 연구는 맞춤형 요리 추천을 통해 사용자 경험을 향상시키고 정보 제공의 가능성을 확장하고자 한다.

## II. 관련 연구

### 2-1 YOLOv8

YOLO는 2015년에 처음 소개가 된 후 자율주행 차량 시스템, 농업 분야, 의료 분야, 보안 시스템 등 다양한 분야의 컴퓨터 비전 연구자들의 관심을 받아왔다[1]. 2023년에는 현재 SOTA 성능을 달성한 YOLOv8가 소개되었다. YOLOv8은 YOLOv5를 기본으로 모델 구조를 최적화하는 방식으로 개발되었고, 1개의 백본에 두 개의 중복 헤더를 가지는 구조로 개량되었다[2]. YOLOv8은 YOLO 시리즈의 최신 버전으로, 모델 구조가 개선되어 정확도와 속도에서 더욱 우수하다[3]. 특히 추론 과정에서 전반적인 성능 최적화가 이루어져 있어 음식 재료 인식과 같은 세밀한 작업에서 높은 성능을 기대할 수 있다. 또한 PyTorch와의 호환성이 뛰어나며, 다양한 배포 환경에서 쉽게 적용할 수 있다.

### 2-2 BERT

BERT(Bidirectional Encoder Representations from Transformers) 모델[4]은 Transformer의 인코더를 여러 개 쌓아 구성된 모델로 주어진 맥락에 맞게 특정 단어의 임베딩 벡터를 출력하도록 학습된 모델이다[5]. BERT는 입력된 문장을 양방향에서 동시에 읽어 단어와 문장의 맥락을 깊이 있

게 파악할 수 있다. 트랜스포머 아키텍처를 기반으로 하며, 대량의 텍스트 데이터로 사전 학습 후 특정 작업에 맞게 미세 조정된다. 또한, BERT는 마스크 언어 모델(MLM)과 다음 문장 예측(NSP) 작업을 통해 문맥과 문장 간의 관계를 이해하는 능력을 키워, 다양한 NLP 애플리케이션에서 뛰어난 성능을 발휘한다.

### 2-3 GAN

GAN(Generative Adversarial Networks)은 2014년 이안 굿펠로우(Ian Goodfellow)가 저술한 논문이며, 생성적 적대 신경망이라는 이름처럼 두 개의 신경망의 경쟁을 통해 학습하고 결과물을 만들어낸다[6]. 생성자는 가짜 데이터를 만들어내고, 판별자는 실제 데이터와 가짜 데이터를 구별하려고 하며, 이 두 네트워크가 서로 경쟁하며 학습한다. 생성자는 판별자를 속이려 하고, 판별자는 더 정확히 구별하려 노력함으로써, 생성자는 점점 더 실제와 유사한 데이터를 만들어낸다. GAN은 이미지 생성, 데이터 증강, 스타일 변환 등 다양한 분야에 활용되며, 새로운 데이터를 생성할 수 있는 강력한 도구로 컴퓨터 비전과 인공지능 분야에서 널리 사용되고 있다.

## III. 레시피 추천 시스템 설계

### 3-1 시스템 구성도



그림 1. 레시피 추천 시스템의 개요도  
Fig. 1. Schematic diagram of the recipe recommendation system

본 논문의 시스템은 이미지를 업로드하고 추천 레시피를 확인할 수 있는 사용자 인터페이스와, 이미지 처리, 재료 인식, 레시피 추천, 데이터 저장 및 검색을 담당하는 서버로 구성된다. Fig. 1과 같이 사용자가 음식 이미지를 입력하면 YOLO 모델이 음식 재료를 객체로 탐지하고, 탐지된 객체는 원-핫 인코딩을 통해 1차원 행렬로 변환된다. 변환된 행렬에 가중치를 부여한 후, 내용 기반 필터링을 통해 유사도가 높은 순으로 추천한다. 유사도 계산에는 코사인 유사도를 사용하며, 일정 기준 이상의 유사도를 가진 데이터를 DB에서 조회하여 사용자에게 추천한다.

### 3-2 데이터 수집 및 정제

본 연구에서는 텍스트 학습을 위해 공공데이터포털의 데이터를, 이미지 학습을 위해 Recipe 1M+ API에서 수집한 데이터를 사용하였다. 텍스트 데이터는 요리 제목, 설명, 재료 목록 등으로 구성되며, BERT 모델 학습에 적합하도록 일관성 있게 정제하였다. 구체적으로, 텍스트 데이터는 불필요한 공백 및 특수 문자를 제거하고, 요리 레시피와 관련된 주요 정보를 추출하여 필드별로 정리하였다.

이미지 데이터는 레시피와의 매칭을 위해 Recipe 1M+ API에서 요리 이미지와 함께 수집하였으며, 각 이미지에는 요리 유형, 재료 등의 레이블을 부착하여 학습에 필요한 구조를 갖추도록 하였다. 이를 통해 BERT 모델의 텍스트 임베딩 및 YOLO 모델의 이미지 인식 정확도를 높이고, 데이터의 일관성과 신뢰성을 확보하였다.

### 3-3 BERT 모델 훈련

본 연구에서는 BERT 모델을 사용하여 레시피 텍스트 데이터를 임베딩하고, 이를 추천 시스템에 활용하였다. BERT 모델은 수집된 텍스트 데이터(요리 제목, 설명, 재료 목록)를 임베딩 벡터로 변환하여 단어와 문장 간의 의미를 보존한다.

훈련 과정에서 텍스트 데이터를 전처리하여 불필요한 정보를 제거하고 문장 단위로 토큰화하였다. 생성된 텍스트 임베딩은 코사인 유사도를 계산하는데 사용되며, 이를 통해 사용자 입력과 유사도가 높은 레시피를 추천하도록 설계하였다.

### 3-4 GAN 모델 훈련

음식 이미지 데이터의 다양성을 확보하기 위해 GAN 모델을 활용하여 새로운 음식 이미지를 생성하였다. 생성된 이미지는 기존 데이터와 함께 사용되며, 다양한 음식 종류에 대한 학습을 통해 레시피 추천 시스템의 정확도를 높이는 데 기여한다. 특히, 생성된 이미지들은 레시피 추천 과정에서 유사도 계산에 필요한 데이터의 양과 질을 보완하는 역할을 한다.

이후 GAN 모델이 생성한 이미지와 API를 통해 수집된 이미지를 함께 사용하여 데이터를 통합한다. 통합된 데이터는 다양한 종류의 음식 이미지를 포함하며 결과적으로 내용 기반 필터링 시스템의 성능을 향상시킨다.

### 3-5 생성한 텍스트 정보와 이미지 데이터의 통합

BERT 모델이 생성한 텍스트 임베딩과 이미지 데이터를 함께 사용하여 데이터를 통합한다. 통합된 데이터는 음식 이미지와 텍스트 정보가 모두 포함되어 추천 시스템의 정확도를 향상시킨다.

### 3-6 웹과 연동

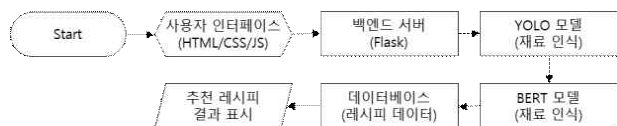


그림 2. 웹과 백엔드 데이터의 연동  
Fig. 2. Integration of web and backend data

Fig. 2는 생성된 레시피 정보와 음식 이미지 데이터를 웹과 연동하여 사용자에게 제공하는 과정을 보여준다. HTML/CSS/JavaScript로 구성된 사용자 인터페이스에서 사용자는 이미지를 업로드하고 추천된 레시피를 확인할 수 있다. 업로드된 이미지는 백엔드 서버(Flask)로 전송되며, 서버는 YOLO 모델로 이미지를 분석해 재료를 인식한다.

인식된 재료 정보는 BERT 모델을 통해 텍스트 임베딩으로 변환되며, 이를 레시피 데이터베이스에 저장된 임베딩과 비교해 유사도를 계산한다. 유사도가 높은 레시피가 선택되면, 해당 레시피를 사용자 인터페이스로 전송하여 결과를 표시한다.



#### IV. 결론

본 논문은 YOLOv8, BERT 및 GAN 모델을 활용하여 사용자에게 맞춤형 레시피를 추천하는 방법을 제안한다. 이를 통해 사용자가 보유한 식재료를 자동으로 인식하고, 다양한 레시피를 제공하여 요리 실패를 줄이고 흥미를 유발할 수 있다. 이 접근은 사용자 경험을 향상시키는 동시에 음식물 쓰레기 감소에도 기여한다. 향후 연구에서는 딥러닝 모델의 정확도를 높이고 사용자 피드백을 반영하여 개인화된 추천 시스템을 구축하고자 한다.

#### 참 고 문 헌

- [1] J. Terven and D. Cordova-Esparza. "A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond." arXiv preprint arXiv:2304.00501, 2023.
- [2] T. H. Lee and C. S. Park, "Real-Time Fire Detection Method Using YOLOv8", Journal of the Semiconductor & Display Technology, Vol. 22, No. 2. June 2023.
- [3] G. Ang, et al. "A novel application for real-time arrhythmia detection using YOLOv8." arXiv preprint arXiv:2305.16727, 2023.
- [4] J. Devlin, M. W. Chang, and K. T. Lee, "BERT: Pre-training of deep bidirectional transformers for language understanding." arXiv:1810.04805, 2018.
- [5] S. H. Hwang and D. H. Kim, "BERT-based Classification Model for Korean Document," The Journal of Society for e-Business Studies Vol.25, No.1, pp.203-214, 2020.
- [6] S. J. Bae, M. G. Kim and H. K. Jung. "GAN system using noise for image generation." Journal of the Korea Institute of Information and Communication Engineering Vol. 24, No. 6, pp.700-705, 2020.

# 빛과 소리를 감지한 ON/OFF 무드등 구현

김은선\*, 석찬비\*, 차하얀\*\*, 고주영\*\*\*, 김현기\*\*\*\*

## Implementation of ON/OFF Mood Light for Detects Light and Sound

Eun Seon Kim\*, Seok Chan Bi Seok\*, Ha Yan Cha\*\*, Jooyoung Ko\*\*\*, and Hyenki Kim\*

### 요 약

본 연구는 아두이노 키트를 활용하여 빛과 소리를 감지해 자동으로 ON/OFF 되는 무드등을 설계하고 구현하는 것을 목표로 한다. 빛 감지 센서와 소리 감지 센서를 사용하여 주변 환경의 조도와 소음을 실시간으로 모니터링하고, 마이크로컨트롤러가 이를 처리하여 무드등을 제어한다. 연구 결과물은 아두이노의 저비용, 고효율 특성을 활용하여 스마트 홈 기술의 접근성을 높이는 데 기여하며, 향후 센서 성능 개선 및 인공지능 도입을 통해 더욱 고도화된 시스템 개발 가능성을 제시한다.

### Abstract

This study aims to design and implement a mood light that automatically turns ON/OFF by detecting light and sound using an Arduino kit. Light and sound detection sensors are used to monitor the illumination and noise of the surrounding environment in real time, and a microcontroller processes them to control the mood light. The research results contribute to increasing the accessibility of smart home technology by utilizing Arduino's low-cost and high-efficiency characteristics, and suggest the possibility of developing a more advanced system through improved sensor performance and introduction of artificial intelligence in the future

### Key words

Smart Home Technology, Light Detection, Sound Detection, Arduino

## 1. 서 론

조명은 인테리어 디자인에서 중요한 역할을 하며, 실내 공간의 분위기와 사용자 경험에 큰 영향을 미친다. 조명의 종류와 배치는 공간의 느낌과 전달

하는 메시지를 결정짓는 주요 요소로, 조명을 조절하거나 변경하는 것은 실내 환경을 변화시키는 중요한 방법중 하나이다. 최근 스마트홈, 스마트 인테리어, 홈 IoT 산업이 급격히 발전하면서 사용자들은 더 높은 편의성과 맞춤형 환경을 제공하는 조명 솔루션

\*국립안동대학교 AI융합학과, \*\*국립안동대학교 응용화학과, \*\*\*국립안동대학교 SW융합교육원, \*\*\*\*국립안동대학교 소프트웨어융합학과, hkkim@anu.ac.kr(교신저자)

※본 연구는 과학기술정보통신부 및 정보통신기획평가원에서 지원하는 SW중심대학사업(IITP-2019-0-01113)의 연구 결과로 수행되었음

루션을 요구하고 있다. 이에 다양한 기능을 가진 무드등은 사용자의 정서적 안정과 라이프스타일에 맞춘 조명 효과를 제공하며 큰 인기를 끌고 있다.

스마트 조명 시스템은 단순한 스위치 제어를 넘어, 사용자의 필요와 환경 변화에 맞춰 자동으로 조명을 제어할 수 있는 기술이 필요하다. 빛과 소리를 감지하여 자동으로 ON/OFF 되는 무드등은 이러한 요구를 충족시켜준다. 본 연구에서는 스마트 조명 시스템을 구현하기 위해 아두이노 키트를 활용했다. 아두이노는 오픈 소스 하드웨어 플랫폼으로, 다양한 센서와 쉽게 통합할 수 있어 저비용으로 고효율의 시스템을 구축할 수 있다.

본 연구에서는 빛 감지 센서와 소리 감지 센서를 활용하여 실내조명을 자동으로 제어하는 무드등을 설계하고 구현한다. 이를 통해 사용자는 별도의 조작 없이도 주변 환경에 맞춰 조명이 자동으로 켜지고 꺼지도록 할 수 있으며, 이는 편의성 증대와 에너지 절약에도 기여할 수 있다.

본 연구에서는 빛과 소리 감지 기술의 원리와 구현 방법을 설명하고, 아두이노를 활용한 시스템 설계와 구현 과정을 다룬다. 실험을 통해 시스템의 성능을 검증하고, 결과를 분석하여 스마트 조명 시스템의 실현 가능성과 향후 개선 방안을 제시한다. 이는 스마트홈 기술의 발전과 더불어 사용자 맞춤형 조명 솔루션 개발에 중요한 기초 자료가 될 것으로 기대한다.

## II. 관련 연구

### 2.1. 조명 최신 트렌드

과거의 조명은 주로 수동으로 제어되었지만, 스마트홈, 스마트 인테리어, 홈 IoT 산업이 발전하면서 스피커, 가습기 등 다양한 기능과 함께 센서를 활용한 자동화 시스템이 적용되어 사용자의 편의성을 높이고 있다. 계절에 따라 조도를 조절하거나 날씨에 따라 색온도를 자동으로 조절하는 기능, 자주 사용하지 않는 조명을 자동으로 제어하는 기능, 스마트폰 어플리케이션을 통해 음성으로 조명을 제어할 수 있는 편리함은 사용자들에게 높은 만족도를 제공하고 있다[1].

### 2.2. 무드등

무드등은 크기가 작고 배선 공사가 필요 없어 어디에서나 쉽게 배치할 수 있으며 사용자의 정서적 안정을 돕는 역할을 한다. 최근에는 자신의 라이프스타일에 맞게 집을 꾸미고자 하는 소비자들이 많아지면서 부담 없는 비용으로 만족스러운 인테리어 효과를 얻을 수 있는 무드등이 큰 인기를 끌고 있다[2].

### 2.3. 조도 센서

조도 센서는 주변의 빛의 세기(조도)를 감지하여 해당 값을 아날로그 신호로 변환한다. 이 신호를 이용하여 무드등을 제어할 수 있다. 대표적인 조도 센서로는 포토레지스터(LDR)가 있으며, 이는 저렴하고 사용이 간편하여 자동차 헤드라이트나 비상조명과 같이 다양한 응용 분야에서 사용된다. 포토레지스터(LDR) 조도 센서는 저항형 센서로 빛의 양이 많아지면 조도 센서의 저항값이 작아지고, 빛의 양이 적어지면 저항값이 커진다. 이러한 변화를 아날로그 신호로 읽어 조명 제어에 활용할 수 있다. 그림 1은 광센서 모듈의 작동원리를 나타낸다.

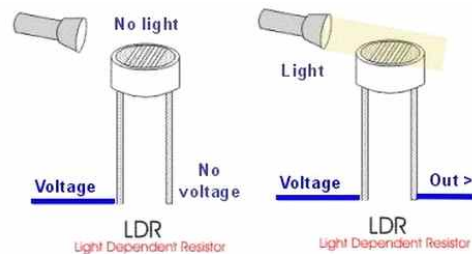


그림 1. 광센서 모듈 작동원리  
Fig 1. LDR(Light Dependent Resistor)

### 2.4. 소리 증폭 감지 센서

소리 증폭 감지 센서는 주변의 소리를 감지하여 이를 전기 신호로 변환한다. 마이크로폰과 신호 증폭기로 구성되어 있어, 소리를 증폭하여 더 민감하게 감지할 수 있으며, 노래방이나 방송에서 사용하는 마이크, 스마트폰, 전화기, 영상 및 소리 녹화 기기 등 소리를 녹음하거나 전달하는 장치와 같이 다양한 분야에서 사용된다.

### III. 시스템 구성

본 연구에서는 스마트홈 on/off 무드등 시스템을 위해 아두이노 보드, 조도 센서, 소리 증폭 감지 센서, 그리고 무드등으로 구성된다. 이 장에서는 각 구성 요소의 역할과 동작 원리를 상세히 설명한다.

#### 2.1. 아두이노 보드

아두이노 보드는 시스템의 중심 제어 장치로, 센서들로부터 데이터를 수집하고 무드등을 제어하는 역할을 한다. 아두이노는 다양한 센서와 액추에이터와 쉽게 연동될 수 있으며, 프로토타이핑에 적합한 플랫폼이다. 본 연구에서는 아두이노 우노(Arduino Uno) 보드를 사용하였다.

#### 2.2. 조도 센서

조도 센서는 주변 환경의 밝기를 측정하여 조명의 상태를 결정하는 데 중요한 역할을 한다.

본 시스템에서는 LDR(Light Dependent Resistor)을 사용하였다. LDR은 빛의 강도에 따라 저항 값이 변하는 특성을 가지고 있으며, 이를 통해 주변 조도를 측정할 수 있다. 조도 센서의 출력은 아두이노의 아날로그 입력 핀에 연결된다.

#### 2.3. 소리 증폭 감지 센서

소리 증폭 감지 센서는 환경의 소리 변화를 감지하여 조명을 제어하는 데 사용된다. 본 연구에서는 소리 증폭 모듈(MIC)을 사용하였다. 이 센서는 마이크로폰을 통해 소리를 감지하고, 증폭된 신호를 아두이노로 전송하여 소리의 유무를 판단한다. 소리 감지 센서의 출력은 아두이노의 디지털 입력 핀에 연결된다.

#### 2.4. 무드등

무드등은 사용자가 원하는 분위기를 조성하는 조명 장치로, 본 시스템의 출력 장치이다. 본 연구에서는 LED 조명을 사용하였으며, 이는 아두이노의 디지털 출력 핀을 통해 제어된다. LED 조명은 저전력으로 다양한 색상과 밝기를 구현할 수 있어 무드등으로 적합하다.

#### 2.5. 전원 공급 장치

시스템의 모든 구성 요소는 안정적인 전원 공급이 필요하다. 아두이노 보드는 USB 포트를 통해 전원을 공급받을 수 있으며, 필요에 따라 외부 전원 어댑터를 사용할 수 있다. 센서와 LED 조명도 아두이노 보드를 통해 전원을 공급받는다.

이와 같은 구성 요소들은 서로 유기적으로 결합되어 스마트홈 환경에서 자동으로 켜지고 꺼지는 무드등 시스템을 구현한다. 다음 절에서는 이러한 구성 요소들이 어떻게 상호 작용하는지 설명한다.

### IV. 구현 및 테스트

스마트 무드등 시스템의 구현은 하드웨어 연결과 소프트웨어 개발 두 가지 측면으로 나뉜다

#### 4.1. 하드웨어 연결

본 시스템의 하드웨어 구성은 아두이노 보드, 조도 센서, 소리 증폭 감지 센서, 그리고 LED 무드등을 포함한다.

조도 센서(LDR)는 아두이노의 아날로그 입력 핀에 연결된다. LDR의 한쪽 단자는 5V 전원에, 다른 한쪽 단자는 아날로그 입력 핀(A0)과 10kΩ 저항을 통해 GND에 연결된다. 이 회로를 통해 LDR의 저항값 변화에 따라 아날로그 입력값이 변하게 된다.

소리 감지 센서(MIC)는 주변 소리를 감지하여 아두이노에 신호를 전달한다. 센서의 출력 핀은 아두이노의 디지털 입력 핀(2번 핀)에 연결된다. 센서의 VCC 핀은 5V 전원에, GND 핀은 아두이노의 GND에 연결된다.

LED 무드등은 아두이노의 디지털 출력 핀(13번 핀)에 연결된다. LED의 양극(+)은 저항을 통해 디지털 출력 핀에 연결되고, 음극(-)은 GND에 연결된다. 저항값은 LED의 전류 제한을 위해 220Ω을 사용하였다.

#### 4.2. 소프트웨어 구현

하드웨어 연결이 완료된 후, 아두이노 보드에서 센서 데이터를 처리하고 LED 무드등을 제어하기 위한 소프트웨어를 개발한다. 소프트웨어는 아두이노 IDE를 사용하여 작성되며, 주요 구현 내용은 다

음과 같다. 그림 2는 아두이노 스케치이다.

는 데 기여할 수 있다고 판단된다.

```

sketch_jun12a.g
void setup() {
  pinMode(7, OUTPUT); // 7번 핀을 출력 모드로 설정 (LED 제어용)
  Serial.begin(9600); // 시리얼 통신을 9600 보드레이트로 초기화
}

void loop() {
  int cds = analogRead(A0); // A0 아날로그 핀에서 값을 읽어 변수 cds에 저장
  Serial.println(cds); // A에서 읽은 값을 시리얼 모니터에 출력

  if (cds > 600) { // 만약 읽은 값이 600보다 크면 (빛이 밝으면)
    digitalWrite(7, HIGH); // 7번 핀에 HIGH 신호를 출력 (LED 켜기)
  } else { // 그렇지 않으면 (빛이 어두우면)
    digitalWrite(7, LOW); // 7번 핀에 LOW 신호를 출력 (LED 끄기)
  }
}

sketch_jun12a.g
int R = 9; // 빨간색 LED 연결 핀
int B = 10; // 초록색 LED 연결 핀
int G = 11; // 파란색 LED 연결 핀

int re;
int sound = A0; // 소리 센서로부터 읽은 값을 저장할 변수
int h = 1000; // 소리 감지 임계값
int count = 3; // 1000 LED 색을 변경하기 위한 카운트 변수

void setup() {
  Serial.begin(9600); // 시리얼 통신을 9600 보드레이트로 초기화
  pinMode(R, OUTPUT); // 빨간색 LED 핀을 출력 모드로 설정
  pinMode(B, OUTPUT); // 초록색 LED 핀을 출력 모드로 설정
  pinMode(G, OUTPUT); // 파란색 LED 핀을 출력 모드로 설정
  pinMode(sound, INPUT); // 소리 센서 핀을 입력 모드로 설정
}

void loop() {
  re = analogRead(sound); // 소리 센서로부터 값을 읽어 변수 re에 저장
  Serial.println(re); // 읽은 값을 시리얼 모니터에 출력

  if (re > h) { // 소리 감지 임계값(h)보다 큰 소리가 감지되면
    digitalWrite(R, HIGH); // count 변수 증가
    count++;
  }

  if (re < h) { // count 2배 (소리 LED 색 변경)
    count = count * 2; // count를 3으로 나눈 나머지를 저장하여 0, 1, 2 /
  }

  // count 2배 (소리 LED 색 변경)
  if (count == 1) { // count가 1이면 빨간색 LED 켜기
    set(255, 0, 0);
  } else if (count == 2) { // count가 2이면 파란색 LED 켜기
    set(0, 0, 255);
  } else { // count가 0이면 LED 끄기
    set(0, 0, 0);
  }
}

// 1000 LED 색을 변경하는 함수
void set(int red, int green, int blue) {
  digitalWrite(R, red); // 빨간색 LED에 주어진 밝기 값 출력
  digitalWrite(B, green); // 초록색 LED에 주어진 밝기 값 출력
  digitalWrite(G, blue); // 파란색 LED에 주어진 밝기 값 출력
}
    
```

그림 668. 아두이노(Arduino) 스케치 (sketch)  
Fig 2. Arduino sketch

이 코드는 조도 센서로부터 주변 밝기를 측정하고, 소리 감지 센서로부터 소리를 감지하여 무드등을 제어한다. 조도 센서값이 일정 임계값 이하이면 무드등을 켜고, 이상이면 끈다. 또한, 소리 감지 센서값이 일정 임계값 이상이면 무드등의 상태를 토글한다.

### 4.3. 시스템 테스트

시스템의 하드웨어와 소프트웨어가 모두 구현된 후, 다양한 환경에서 시스템의 동작을 테스트하였다. 테스트는 조도 센서와 소리 감지 센서의 반응 속도와 정확성을 평가하는 데 중점을 두었다. 테스트 결과, 시스템은 실내 명 조건과 사용자 소리 입력에 따라 안정적으로 동작함을 확인하였다.

이와 같은 하드웨어 연결과 소프트웨어 구현을 통해 스마트홈 on/off 무드등 시스템이 완성되었으며, 다양한 환경에서의 테스트를 통해 시스템의 유효성을 검증하였다.

## III. 결 론

본 연구는 아두이노를 활용하여 빛과 소리를 감지하여 제어되는 무드등을 구현하여 스마트 조명 시스템을 개발하였다. 아두이노는 오픈 소스 하드웨어로서 저비용으로 다양한 센서와 쉽게 통합할 수 있다는 장점이 있다. 이를 통해 일반 사용자나 개발자들이 손쉽게 스마트 조명 시스템을 구축할 수 있게 하였으며, 이는 스마트 홈 기술의 접근성을 높이

특히, 빛과 소리를 동시에 감지하여 상황에 맞는 조명을 제공하는 기능은 기존의 단순한 자동 조명 시스템과 차별화되는 중요한 특징이다. 이러한 기능은 사용자에게 더욱 쾌적한 환경과, 에너지 절약 측면에서도 큰 이점을 제공한다.

향후 연구에서는 다음과 같은 방향으로 연구를 확장할 수 있다.

첫째, 센서의 정확도를 높이기 위해 보다 고성능의 빛 감지 센서와 소리 감지 센서를 도입할 수 있다. 이를 통해 시스템의 신뢰성을 더욱 향상시킬 수 있을 것이다.

둘째, 무드등의 디자인 및 기능을 다양화하여 사용자에게 더욱 다양한 선택지를 제공할 수 있다. 예를 들어, 색상 변화, 밝기 조절 등의 기능을 추가할 수 있다.

또한, 인공지능(AI) 기술을 도입하여 사용자 패턴을 학습하고 예측하는 기능을 추가할 수 있다. 이를 통해 시스템은 사용자에게 더욱 개인화된 조명 환경을 제공할 수 있을 것이다. 마지막으로, 본 연구에서 제안한 시스템을 다른 스마트 홈 기기와 연동하여 보다 종합적인 스마트 홈 솔루션을 개발하는 것도 의미미한 연구 방향이 될 것으로 기대한다.

## 참 고 문 헌

[1 J. Yoon, “The Latest Trend in Architectural Lighting-IOT Solutions“, Review of Architecture and Building Science, vol. 68, no. 2, pp. 30-31,

2024.

- [2] “Mood Light, which makes the space cozy, is popular”,  
<http://www.theliving.co.kr/news/articleView.html?idxno=21279>(accessed OCT. 22, 2024).

# 임베디드 시스템을 활용한 자동 식물 관리 시스템 설계 및 구현

최민동\*, 서승태\*, 배정호\*, 고주영\*\*, 김현기\*\*\*

## Design and Implementation of Automatic Plant Management System Using Embedded System

Mindong Choi\*, Seung Tae Seo\*, Jeong Ho Bae\*, Jooyoung Ko\*\* and Hyenki Kim\*\*\*

### 요 약

현대 사회에서 많은 사람이 일과 바쁜 생활로 식물 관리에 충분한 시간을 할애할 수 없는 경우가 많다. 이에 따라 자동 식물 관리 시스템을 설계 및 구현되었다. 이 시스템은 온도 및 습도 센서를 통해 식물이 자라기에 적합한 환경을 조성하고, 토양 수분 센서를 사용하여 자동으로 물을 공급하는 기술을 적용하였다. 이러한 기술은 탄소 배출 감소와 물 소비 절약을 통해 가뭄에 대한 대응력도 높이는 등 식물 관리에 편의성을 크게 향상시킬 것으로 기대된다.

### Abstract

In the modern society, many people are often unable to devote enough time to plant management due to their busy work. Accordingly, an automatic plant management system was designed and implemented. This system applied a technology that creates an environment suitable for plants to grow through temperature and humidity sensors, and automatically supplies water using soil moisture sensors. These technologies are expected to greatly improve the convenience of plant management, such as increasing the responsiveness to drought by reducing carbon emissions and saving water consumption.

### Key words

Automated plant management system, Water conservation, Automatic watering

### 1. 서 론

현대 사회에서 많은 사람이 일과 생활의 균형을 맞추기 위해 바쁜 일정을 소화하며 살아가고 있다.

이에 따라 식물 관리에 충분한 시간을 할애하지 못하는 경우가 많아, 식물의 성장 상태가 저하되거나 관리가 소홀해지는 문제가 발생한다. 예를 들어, 적절한 물 주기와 일조량의 조절, 토양의 영양 상태

\*국립안동대학교 멀티미디어공학과, \*\*\*국립안동대학교 SW융합교육원, \*\*\*\*국립안동대학교 소프트웨어융합학과, hkkim@anu.ac.kr(교신저자)

※본 연구는 과학기술정보통신부 및 정보통신기획평가원에서 지원하는 SW중심대학사업(IITP-2019-0-01113)의 연구 결과로 수행되었음

관리 등은 식물의 건강에 필수적이지만, 이를 지속해서 모니터링하고 조절하는 것은 많은 시간과 노력이 요구된다. 이러한 배경에서, 자동화된 식물 관리 시스템의 필요성이 대두되었다. 최근 사물 인터넷(IoT)과 센서 네트워크 기술의 발전으로 농장의 상태를 지속적으로 모니터링 할 수 있게 되었다[1]. 그리고 데이터 수집 및 분석으로 농업 생산성을 향상 시킬수 있다[2]. 그러므로 본 연구에서는 임베디드 시스템 기반으로 센서를 이용하여 토양 수분을 측정하고 데이터를 클라우드에 전송함으로써 사용자가 식물관리를 쉽고 편리하게 지속할 수 있는 시스템을 설계하고 구현하였다

## II. 관련 연구

임베디드 시스템은 전자 장치의 특정 기능을 수행하기 위해 하드웨어와 소프트웨어가 결합된 시스템으로, 다양한 산업 분야에서 활용되고 있다. 특히, 스마트팜 기술은 이러한 임베디드 시스템을 활용하여 농업 생산성을 향상시키고 있다. 스마트팜은 센서 네트워크와 IoT(사물인터넷) 기술을 통합하여 실시간으로 환경 데이터를 모니터링하고, 이를 기반으로 자동화된 관리 시스템을 구현한다. 예를 들어, 정밀 농업에서는 토양 수분, 온도, 습도 등의 데이터를 실시간으로 수집하여 작물의 성장 상태를 최적화한다.

스마트팜 타운 식물재배 시스템에서는 토양 수분 센서, 온도 및 습도 센서, 그리고 자동 제어 장치들이 결합하여 식물의 성장 조건을 최적화한다[3]. 이러한 시스템은 기후 변화에 따른 작물 관리의 어려움을 해결하고, 물과 에너지를 효율적으로 사용하여 지속 가능한 농업을 실현하는 데 기여한다. 다양한 연구에서 이러한 스마트팜 기술이 농업 생산성을 높이고, 자원의 낭비를 줄이며, 작물의 품질을 개선하는 데 효과적임을 보여주고 있다. 이러한 사례들은 스마트팜 기술이 농업의 효율성과 지속 가능성을 높이는 데 큰 잠재력을 가지고 있음을 보여준다.

## III. 설 계

본 연구에서 설계한 자동 식물 관리 시스템은 아두이노를 기반으로 한 임베디드 시스템으로, 다양

한 센서를 통해 식물의 성장 환경을 자동으로 조절한다. 시스템은 크게 센서 모듈, 제어 모듈, 및 출력 모듈로 구성되어 있다.

이 시스템은 연구자가 설정한 임계값을 기준으로 자동으로 작동하며, 실시간으로 환경 데이터를 모니터링하고 출력한다. 이러한 자동화 시스템은 혼자서도 간단히 식물의 최적 성장 환경을 유지할 수 있도록 돕는다. 표 1은 각 모듈의 구성 요소와 기능이다.

표 1. 각 모듈의 구성 요소와 기능

Table 1. Components and functions of each module

시스템 명	구성 요소	기능
센서 모듈	토양 수분 센서	토양의 전기 전도도를 측정하여 수분 상태를 모니터링
	온도 및 습도 센서	주변 환경의 온도와 습도를 측정
제어 모듈	아두이노 보드	센서 데이터를 실시간으로 수집
	control logic	토양의 전기 전도도에 따라 펌프모터 작동 온도가 18도 미만이면 LED 실행, 온도가 25도 이상이면 LED 종료
출력 모듈	펌프 모터	적절한 물을 공급
	식물 성장용 LED	적절한 빛 제공
	LDC Display	실시간으로 온도, 습도를 사용자에게 보여줌

## IV. 구 현

본 연구에서는 아두이노를 기반으로 한 자동 식물 관리 시스템을 구현하였다. 시스템은 토양 수분 센서, 온도 및 습도 센서(DHT22), 물 공급 및 배출 펌프, 식물 성장용 LED, 그리고 LCD 디스플레이로 구성되어 있다. 각 구성 요소들은 아두이노 보드와 연결되어 있으며, 센서로부터 수집된 데이터를 바탕으로 자동으로 제어된다. 하드웨어에서는 토양 수분 센서를 이용하여 토양의 수분 상태를 측정하여 아두이노 보드로 데이터를 전송한다. 두 개의 센서가 각각 A0와 A1 핀에 연결된다. 온도 및 습도 센서(DHT22): 주변 환경의 온도와 습도를 측정하여 아두이노 보드로 데이터를 전송한다. DHT22 센서는 디지털 핀 2에 연결된다. 물 공급 및 배출 펌프: 토



양 수분 센서의 데이터를 바탕으로 물을 공급하거나 배출하는 역할을 한다. 두 개의 펌프는 각각 디지털 핀 3과 4에 연결된다. 식물 성장용 LED: 온도에 따라 LED를 켜거나 끈다. LED는 디지털 핀 13에 연결된다. LCD 디스플레이: 실시간으로 온도와 습도 데이터를 사용자에게 보여준다. I2C 통신을 통해 아두이노 보드에 연결된다. 자동 식물 관리 시스템의 소프트웨어는 아두이노 IDE를 사용하여 작성되었다. 데이터는 토양 수분 센서와 DHT22 센서로부터 데이터를 수집한다. 펌프 제어는 토양 수분 값이 임계값보다 낮을 때 펌프를 작동시켜 물을 공급하거나 배출한다. LED 제어는 7시간 간격으로 LED를 켜고 끄도록 하였다. 그리고 수집된 온도와 습도 데이터를 LCD 디스플레이에 출력한다. 표 2는 아두이노 보드와 센서를 이용한 데이터 수집 및 제어 코드의 일부이다.

표 2. 토양 수분값에 따른 물 주기 코드

Table 2. Code of water supply according to soil moisture value

```

// 토양 수분 값 읽기
int soilMoistureValue1 = analogRead(soilMoisturePin1);
int soilMoistureValue2 = analogRead(soilMoisturePin2);
// 토양 수분이 낮으면 물 공급 및 배출
if (soilMoistureValue1 < moistureThreshold ||
soilMoistureValue2 < moistureThreshold) {
    digitalWrite(pumpPin1, HIGH); // 물 공급 펌프 켜기
    delay(waterOperationDuration);
        // waterOperationDuration 동안 물 공급
    digitalWrite(pumpPin1, LOW); // 물 공급 펌프 끄기
    digitalWrite(pumpPin2, HIGH); // 물 배출 펌프 켜기
    delay(waterOperationDuration);
        // waterOperationDuration 동안 물 배출
    digitalWrite(pumpPin2, LOW); // 물 배출 펌프 끄기
}
    
```

## V. 결 론

자동식물 관리 시스템은 바쁜 현대인들에게 필요한 솔루션으로, 적절한 물 주기, 햇빛, 비료 공급 등의 식물 관리를 자동화함으로써 시간과 노력을 절약할 수 있다. 사물 인터넷(IoT)과 아두이노 센서 기술의 발전 덕분에 이러한 시스템이 가능해졌으며, 이는 환경 보호에도 기여한다. 자원 낭비를

줄이고 지속 가능한 식물 관리를 가능하게 하는 자동식물 관리 시스템은 미래의 스마트 홈의 중요한 요소가 될 것이다. 이러한 시스템의 도입은 친환경적인 생활을 촉진할 것이다. 예를 들어, 물 소비를 최적화함으로써 가뭄에 대한 대응력을 높이고, 에너지 사용을 효율화하여 탄소 배출을 줄일 수 있다.

따라서 자동식물 관리 시스템은 현대 사회에서 필수적인 기술로 자리 잡을 것이다. 향후 연구에서는 인공지능을 활용하여 식물의 상태를 분석하고, 예측 모델을 통해, 더 정밀한 관리가 가능하도록 할 수 있을 것이다. 예를 들어, 기계 학습 알고리즘을 사용하여 각 식물의 개별 요구 사항을 학습하고, 이를 바탕으로 맞춤형 관리 전략을 개발할 수 있다. 또한, 다양한 식물 종류에 대한 데이터베이스를 구축하여, 각 식물에 맞춤형 관리가 가능하도록 시스템을 발전시킬 수 있을 것으로 기대된다.

## 참 고 문 헌

[1] Walter, A., Finger, R., Huber, R., & Buchmann, N. "Smart farming is key to developing sustainable agriculture". Proceedings of the National Academy of Sciences, vol. 114, No. 24, pp. 6148-6150, Jun., 2017.

[2] Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. "Big data in smart farming—A review. Agricultural Systems", vol. 153, pp. 69-80. May., 2017.

[3] E. B. Ko, I. W. Jeong, C. H. Lee, "Design and Implementation of Platform for Small-Scale Smart Farm using Deep Learning", Journal of the Korean Society of Manufacturing Process Engineers, vol. 22, no. 11, pp. 59-67, Nov. 2023.

# 데이터 증강 및 CycleGAN 기반 저조도 객체 인식 향상

김지호\*, 문창현\*, 차승훈\*, 황민\*, 신연순\*\*

## Enhancing Low-Light Object Recognition through Data Augmentation and CycleGAN

Ji-ho Kim\*, Chang-hyun Moon\*, Seung-hoon Cha\*, Min Hwang\*, Youn-Soon Shin\*\*

### 요약

본 연구는 저조도 환경에서 객체 인식 성능을 향상시키기 위해 이미지 증강 기법과 CycleGAN을 결합한 시스템을 제안하였다. 밝기 조절 (0%, -5%, -10%, -15%)을 적용한 데이터셋을 통해 모델을 학습하고, CycleGAN을 이용해 고조도로 변환한 테스트셋을 통해 평가한 결과, -10%와 -15% 밝기 조절 모델이 저조도에서 일관된 성능 향상을 보였다. 본 연구 결과는 CycleGAN과 데이터 증강기법이 결합된 시스템이 저조도 환경에서도 높은 객체 인식 성능을 제공함을 시사한다.

### Abstract

This study proposes a system that combines image augmentation techniques and CycleGAN to enhance object recognition performance in low-light environments. Models were trained with brightness adjustments (0%, -5%, -10%, -15%) and evaluated using a test set transformed to higher brightness with CycleGAN. The evaluation results showed that models adjusted to -10% and -15% brightness consistently improved performance in low-light conditions. These findings suggest that a system combining CycleGAN with data augmentation techniques provides high object recognition accuracy even in low-light environments.

### Key words

CycleGAN, YOLO, Low-Light, Situation Recognition, Data Augmentation

### 1. 서론

저조도 환경에서의 상황인식은 자율주행, 보안 감시, 의료 영상 처리 등 다양한 분야에서 중요한 문제이다. 저조도 이미지의 낮은 선명도와 세부 정

보 손실로 인해 정확한 인식이 어렵기 때문에, 이를 해결하기 위해 데이터 증강기법과 적대 신경망(CycleGAN)을 활용한 연구가 활발히 이루어지고 있다 [1].

\*동국대학교 공과대학 컴퓨터공학전공kimjiho1125@gmail.com, ckdgus1713@naver.com, papa021326@naver.com, minmin3194@naver.com, \*\*동국대학교 AI융합대학 AI소프트웨어융합학부, ysshin@dongguk.edu(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음 (2023-0-00049)

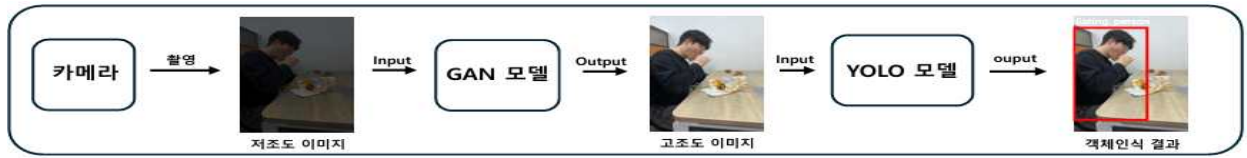


그림 1. 제안하는 시스템 구조  
Fig. 1. Proposal System Architecture

데이터 증강기법은 제한된 데이터 셋을 보완하여 모델의 일반화 성능을 향상시키는 중요한 도구로, 이미지 회전, 밝기 조정 등을 통해 다양한 데이터를 제공한다. 또한 CycleGAN [2]은 저조도

이미지의 밝기와 색상 왜곡을 복원하여 이미지 품질을 개선하는데 효과적이다. YOLO [3]와 같은 객체 탐지 알고리즘도 저조도 환경에서 실시간 성능을 제공하며, 저조도 상황인식 시스템의 핵심 기술로 주목받고 있다. 상황인식에 따른 조도 제어 시스템에서 저조도 이미지로는 정확한 객체 인식이 어려워 적절한 조도 조절이 불가능하다는 문제를 발견하였다. 이에 따라, 저조도 이미지를 고조도로 전환하여 객체 인식의 정확성을 향상시키는 방안을 모색하였다. 데이터 증강기법으로 밝기 조절과 노이즈 추가 등을 활용하여 다양한 학습 데이터를 생성하고, CycleGAN을 통해 입력된 저조도 이미지를 고조도로 변환한 후 YOLO 모델에 입력함으로써 저조도 상황에서의 객체 인식 성능을 높이고자 한다.

## II. 본 론

### 2.1 관련 연구

본 연구는 저조도에서의 객체 인식 정확도를 높이는 것을 목표로 하고 있다. 이와 관련하여 기존 연구들에서는 저조도의 이미지를 고조도의 이미지로 변환하는 방법으로 히스토그램 평활화와 같이 이미지의 밝기와 대비를 조정하는 이미지 향상 기법과 CycleGAN과 같은 딥러닝 기반의 이미지 변환 기법의 평균 정확도(mAP)를 비교하는 연구[4]를 진행하였다. [4]의 히스토그램 평활화의 방법은 전체 이미지의 밝기와 대비를 균등하게 조정하는 과정에서 노이즈가 증폭되어 객체 인식 정확도에 부정적인 영향을 끼치는 문제가 발생하였다. 반면, CycleGAN의 경우 이미지의 특징에 따라 조정되어

객체 인식에 더 나은 성능을 보이는 것으로 확인되었다. 따라서, 본 연구에서는 저조도 상황에서의 객체 인식 정확도 향상을 위한 방법의 하나로 CycleGAN을 채택하였다.

### 2.2 제안하는 시스템

본 연구에서는 이미지 증강기법과 CycleGAN을 결합하여 저조도 상황에서 효과적인 객체 인식 시스템을 제안한다. 그림 1은 저조도의 이미지를 CycleGAN을 활용하여 고조도 이미지로 변환하고, 이미지 증강기법을 적용한 객체 인식 모델을 통해 객체 인식을 진행하는 저조도 객체 인식 시스템을 시각화한 그림이다. 본 연구는 위에서 제안하는 전체 시스템을 구축하고 실험을 통해 그 효율성을 확인하였다.

### 2.3 실험 및 실험 환경

본 연구에서 테스트를 위해 표 1과 같이 실험 환경을 설정하였다. 보통 YOLO를 기반으로 알고리즘을 훈련하려면 많은 데이터가 필요하다. 본 연구에서는 총 7,426개의 이미지로 구성된 데이터셋을 활용하여 학습을 진행하였다. 해당 데이터셋은 표 2와 같이 총 7,042개의 Training Set, 279개의 Validation Set, 105개의 Testing Set으로 구성된다.

표 1. 실험 환경  
Table 1. Experimental Environment

실험환경	버전
운영체제	Ubuntu 22.04.3 LTS
CUDA	12.2
cuDNN	8.9.6
Python	3.10.12
Pytorch	2.5.0 + cu121
GPU	NVIDIA A100

표 2. 데이터셋  
Table 2. Dataset

Training Set	Validation Set	Testing Set
7042	279	105

본 연구에서 테스트를 위해 표 1과 같이 실험 환경을 설정하였다. 보통 YOLO를 기반으로 알고리즘을 훈련하려면 많은 데이터가 필요하다. 본 연구에서는 총 7,426개의 이미지로 구성된 데이터셋을 활용하여 학습을 진행하였다. 해당 데이터셋은 표 2와 같이 총 7,042개의 Training Set, 279개의 Validation Set, 105개의 Testing Set으로 구성된다.

본 실험에서는 다양한 조도 상황에서 객체 인식 모델의 성능을 최적화하기 위해 이미지 증강기법을 통해 조도 변화를 적용한 학습을 진행했다. 특히, 저조도 상황에서의 인식 성능 향상을 목표로, 원본 이미지의 밝기를 0%, -5%, -10%, -15%로 차등 적용했다. 이렇게 밝기를 단계별로 감소시키며 실험한 이유는 저조도 상황에서 더 효율적인 조도 변화를 찾기 위해서였다. 이를 통해 특정 밝기 수준이 저조도에서의 객체 인식 성능에 미치는 영향을 보다 체계적으로 분석하고자 한다.

다양한 조도 환경에서 학습된 객체 인식 모델과 CycleGAN을 활용한 실험을 진행하였다. CycleGAN을 이용해 저조도 이미지로부터 고조도 이미지를 생성하여, 기존의 저조도 상황에서의 상황인식 성능을 더욱 향상시켰다. 구체적으로, 저조도 환경에서 촬영된 이미지를 CycleGAN에 입력으로 넣어 고조도 이미지를 생성하고, 이렇게 생성된 이미지를 활용하여 상황인식을 진행한다. 결과적으로 저조도 상황에서의 상황인식 능력이 개선되는 효과를 확인한다.

표 3. 밝기 조절을 적용한 테스트 데이터셋  
Table 3. A test dataset with brightness adjustment applied

V1	V2	V3
-100%	-200%	-260%

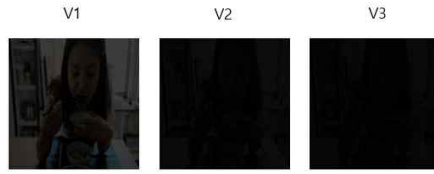


그림 2. 테스트 데이터셋 예시  
Fig. 2. Example of Test Dataset

실험에 사용할 테스트 데이터셋은 총 120장의 저조도 이미지로 구성되어 있으며, 그림 2와 같이 밝기를 원본 이미지 대비 -100%, -200%, -260%로 낮춘 데이터셋을 사용한다. 이를 통해 다양한 저조도 조건에서의 실험을 진행했다.

## 2.4 실험 결과

표 4는 저조도 환경에서의 객체 인식 성능을 개선하기 위해 다양한 데이터 증강기법과 CycleGAN을 적용한 모델의 성능을 V1, V2, V3와 같은 저조도 테스트 데이터를 통해 비교 분석한 결과이다. 모델의 성능평가 지표로는 mAP50-95를 사용하였다.

표 4. 이미지 증강 및 CycleGAN 적용 결과 (mAP50-95)  
Table 4. Image Augmentation Techniques and CycleGAN Application Results (mAP50-95)

	V1		V2		V3	
	전	후	전	후	전	후
0	0.92	0.81	0.54	0.81	0.43	0.66
-5%	0.94	0.87	0.72	0.80	0.48	0.70
-10%	0.94	0.97	0.75	0.93	0.55	0.80
-15%	0.97	0.98	0.68	0.91	0.57	0.77

V1(-100%)에서는 -15% 밝기 조정 모델이 0.97로 최고 정확도를 보였고, V2(-200%)에서는 -10% 모델이 0.75로 가장 높은 성능을 보였다. V3(-260%)에서는 -15% 모델이 0.57로 가장 우수한 성능을 나타냈다. 이 결과는 -10%와 -15% 밝기 조정이 중간에서 극한의 저조도 환경에서 모델의 객체 인식 성능을 향상시키는 데 효과적임을 시사한다.

CycleGAN을 적용한 후, 대부분 모델에서 정확도가 향상되었으며, 특히 -10%와 -15% 이미지로 학습시킨 모델에서 성능 개선이 두드러졌다. 예를 들어,

-10% 모델은 V2 테스트 데이터에서 CycleGAN 적용 전 평균정확도가 0.75였으나, CycleGAN 적용 후 0.93으로 증가하였다. 이는 저조도 조건에서 CycleGAN이 모델의 인식 성능 향상에 긍정적인 영향을 미친다는 점을 보여준다.

한편, 이미지 증강 기법을 적용하지 않은 0% 모델에서는 일부 테스트 데이터(V1)에서 CycleGAN 적용 후 평균정확도가 다소 감소하는 경향이 나타났다. 예를 들어, V1 테스트 데이터에서 0% 모델의 평균정확도는 CycleGAN 적용 전 0.92였으나, 적용 후 0.81로 약간 감소하였다.

가장 높은 평균정확도를 보인 모델은 -10% 모델로, 다양한 저조도 테스트 데이터(V1, V2, V3)에서 일관되게 높은 정확도를 유지하였다. 특히 V2 테스트 데이터에서 -10% 모델의 평균정확도는 CycleGAN 적용 후 0.93로, 다른 밝기 조절 수준의 모델들에 비해 높은 성능을 보였다. -5% 모델과 -15% 모델 또한 CycleGAN 적용 후 평균정확도가 향상되었으나, -10% 모델만큼의 성능은 보여주지 않았다.

결론적으로, 본 연구는 저조도 이미지의 객체 인식 성능을 높이기 위해 CycleGAN과 데이터 증강 기법이 효과적으로 활용될 수 있음을 보여주었다. 특히 -10% 모델은 저조도 테스트 데이터(V2, V3)에서 가장 높은 평균정확도를 기록하며, 저조도 환경에서의 실시간 인식 시스템에 적합한 모델로 평가된다.

### III. 결 론

본 연구에서는 다양한 밝기 조절을 적용한 이미지 증강기법과 CycleGAN을 활용하여 저조도 환경에서의 객체 인식 성능을 향상시키는 시스템을 제안하고 실험을 통해 평가하였다. 실험 결과, CycleGAN을 통한 고조도 변환이 다양한 저조도 조건에서 객체 인식 성능을 크게 향상시킴을 확인하였다. 특히, -10%와 -15% 밝기 조절 모델이 중간에서 극한의 저조도 환경에서 일관된 성능 향상을 보였다. 이 결과를 바탕으로 최종 시스템을 구성할 때, 저조도 이미지의 경우 CycleGAN을 사용해 고조도로 변환하고, -10% 또는 -15% 밝기 조절을 적용

한 학습 모델을 통해 객체 인식을 수행하도록 설계하는 것이 효과적일 것으로 판단된다. 이를 통해 다양한 저조도 환경에서 안정적이고 높은 정확도의 객체 인식 시스템을 구축할 수 있을 것이다. 나아가, 본 연구에서 제시한 방법은 자율주행, 보안 감시, 의료 영상 처리 등 실제 응용 분야에서 저조도 조건에서도 높은 인식 성능을 보장하여 활용성을 높일 수 있으며, 저조도 인식 오류를 최소화할 수 있는 기반을 제공할 것으로 기대된다.

### 참 고 문 헌

- [1] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [2] Wang, L., Zhao, L., Zhong, T., & Wu, C. (2024). Low-light image enhancement using generative adversarial networks. *Scientific Reports*, 14(1), 18489.
- [3] Štancel, M., & Hulič, M. (2019). An introduction to image classification and object detection using YOLO detector. In *CEUR Workshop Proceedings* (Vol. 2403, pp. 1-8).
- [4] Chen, W., & Shah, T. (2021). Exploring low-light object detection techniques. *arXiv preprint arXiv:2107.14382*.
- [5] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

# 생성형 AI를 활용한 인터랙티브 동화책 생성 서비스

정용태\*, 김선혁\*, 김세훈\*, 정기현\*\*

## Interactive Storybook Generation Service Using Generative AI

Jung YongTae\*, Kim SunHyuk, Kim Sehun\*, and Jung KiHyun\*\*

### 요 약

본 연구에서는 생성형 AI를 활용한 서비스를 개발하였다. 이 서비스는 아이들이 동화의 전개와 삽화를 실시간으로 선택할 수 있는 인터랙티브 동화책을 제공한다. 기존 동화책과 달리, 사용자의 선택에 따라 이야기와 이야기에 따른 생성 이미지를 보여준다. 이러한 상호작용 방식은 매번 새로운 이야기를 만들어내며, 아이들이 이야기의 주인공이 된 듯한 몰입감을 느끼게 한다. 본 서비스는 단순한 콘텐츠를 넘어 맞춤형 콘텐츠로 학습 효과를 극대화하고자 한다. 아이들은 다양한 이야기를 생성하며 창의력과 상상력을 키운다. 또한, 개인의 선호도와 반응에 따라 콘텐츠가 조정되어 더욱 개인화된 학습 환경을 제공한다. 본 연구는 시스템의 구현 과정을 상세히 설명하며 인터페이스 설계, 콘텐츠 생성 알고리즘에 대해 논의한다. 이를 통해 이 서비스가 지닌 교육적 가치와 잠재력을 탐구하고자 한다.

### Abstract

In this study, we developed an innovative service using Generative AI. This service provides an interactive fairy tale book that allows children to choose the development of fairy tales and illustrations in real time. Unlike conventional fairy tale books, it shows the story and the generated image according to the user's choice. This interaction method creates a new story every time and makes children feel immersed as if they are the main characters of the story. This service aims to maximize the learning effect with customized content beyond simple content. Children create various stories and develop creativity and imagination. In addition, content is adjusted according to individual preferences and reactions to provide a more personalized learning environment. This study explains the implementation process of the system in detail and discusses interface design and content generation algorithms. Through this, we want to explore the educational value and potential of this service.

### Key words

Interactive children's books, Generative AI, customized contents

---

\* 국립안동대학교 멀티미디어공학과, jyt6640@gmail.com

\*\* 국립안동대학교 소프트웨어융합학과, kingjung@anu.ac.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

## I. 서 론

디지털 기술의 발전과 코로나19 이후 비대면 교육 수요 증가는 아동 교육 분야의 혁신을 요구한다. 특히 스마트 기기 사용이 보편화된 현대 아동들의 학습 환경은 수동적인 콘텐츠 소비에서 벗어나 상호작용이 가능한 형태로 변화가 필요하다[1].

최근 유튜브와 같은 영상 플랫폼의 확산으로 아동의 수동적 콘텐츠 소비가 증가하는 추세이다. 단순 시청 위주의 콘텐츠는 아동의 창의력과 상상력 발달을 저해하며 능동적 사고를 제한한다. 현재 아동 교육용 콘텐츠는 단순 지식 전달에 치중된 전자책이나 일회성 게임 형태가 대부분이다[2]. 일방향적 콘텐츠는 아동의 창의적 사고와 문제 해결 능력 향상에 한계를 보인다.

ChatGPT와 같은 생성형 AI 기술의 발전은 사용자 맞춤형 콘텐츠 제작의 새로운 전환점이 되었다. 생성형 AI는 텍스트와 이미지를 실시간으로 생성할 수 있어 아동 교육 분야에서 새로운 가능성을 제시한다. 특히 아동의 선택과 반응에 따라 스토리와 삽화가 변화하는 인터랙티브 동화책은 수동적 콘텐츠 소비 문제를 해결할 수 있는 대안이 될 수 있다.

본 연구는 생성형 AI 기술을 활용한 인터랙티브 동화책 서비스를 제안한다. 제안된 서비스는 아동이 직접 동화의 전개 방향을 선택하고 이에 따른 스토리와 삽화를 실시간으로 제공받는다. 대규모 언어 모델과 이미지 생성 모델 기반의 맞춤형 콘텐츠는 아동의 창의력과 상상력 발달에 긍정적 영향을 미친다. 본 서비스는 디지털 기기를 창의적으로 활용하는 새로운 방식을 제시하며 아동의 능동적 학습 참여를 유도한다.

## II. 관련 연구

AI 기술의 발전이 교육 콘텐츠의 형태를 변화시키고 있으며, 특히 어린이를 대상으로 하는 창의적 분야에서 새로운 가능성을 제시하고 있다. 인터랙티브 동화책 서비스는 단순한 읽기 자료를 넘어서, 어린이들이 직접 스토리와 삽화를 선택하며 상호작용할 수 있는 참여형 경험을 제공함으로써 창의적 사

고와 흥미를 자극한다[3,4]. 본 논문에서는 생성형 AI 기술을 활용하여 아이들이 직접 이야기를 탐색하고 각자의 상상력에 맞는 동화책을 생성할 수 있는 서비스를 제안한다[5].

OpenAI의 GPT-4 아키텍처를 기반으로 하는 GPT-4o는 최신 기술을 적용하여 언어 이해와 생성 능력을 향상시킨 대규모 언어 모델이다. 본 모델은 사용자 설정에 맞춘 동화 스토리 생성에 강점을 보이며, 동화의 주제나 등장인물 설정에 따라 자연스러운 문맥을 유지하며 이야기를 생성할 수 있다. 또한, Fine-Tuning을 통해 특정 장르나 사용자 취향에 맞는 세밀한 스토리 커스터마이징이 가능하다[6,7].

동화책 삽화 생성에 있어 SDXL 모델은 텍스트 설명을 기반으로 다양한 화풍을 적용할 수 있는 이미지 생성 모델로, 동화 삽화와 같은 예술적 콘텐츠에 적합하다. 그러나 기본 모델은 사용자가 원하는 특정 화풍을 완벽히 반영하는 데 한계가 있다[8]. 본 논문에서는 SDXL 모델에 대한 화풍 Fine-Tuning을 수행하였다. 다양한 화풍 데이터셋을 수집하여 모델을 Fine-Tuning하였으며, 크레용 화풍, 수채화 스타일, 동화적인 블러시 효과 등의 데이터를 학습시켜 사용자가 원하는 삽화 스타일을 실시간으로 제공할 수 있도록 개선하였다[9,10].

## III. 서비스 설계 및 구현

### 3.1 데이터 수집

본 연구에서는 AI-Hub에서 제공하는 데이터셋을 수집하여 Stable Diffusion XL 모델 즉, SDXL 모델과 LLM 모델의 학습에 활용하였다. 이를 통해 사용자 맞춤형 동화 내용 생성 및 삽화 생성을 구현할 수 있도록 설계하였다.

SDXL 모델의 학습을 위해 동화 이미지 데이터셋을 수집하였으며, 이 데이터셋은 다양한 화풍과 장면 구성을 포함한 동화 삽화들로 구성되었다. 일반적으로 각 화풍을 학습하기 위해 약 100~300장의 이미지가 필요하며, 고전 유럽 동화, 일본 신화, 전래동화, 동양 판타지, 모던 판타지 등 테마별로 이미지를 분류하여 학습 데이터셋을 구축하였다. 이를 통해 SDXL 모델은 다양한 스타일의 이미지를 학습하고, 동화의 시각적 요소를 풍부하게 지원한다.

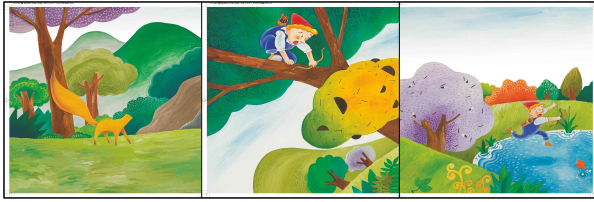


그림 1. 전래동화 이미지 데이터셋  
Fig. 1. Traditional Fairy Tale Image Dataset

또한, LLM 모델은 AI-Hub의 동화 줄거리 데이터셋을 학습하였다. 이 데이터셋은 여러 주제와 스토리 구조를 포함한 동화 내용으로 구성되어 있으며, 이를 바탕으로 LLM 모델은 사용자 설정에 맞는 이야기 생성이 가능하다. 주제와 배경에 맞는 스토리라인을 형성함으로써 아이들이 몰입할 수 있는 서사를 제공한다.

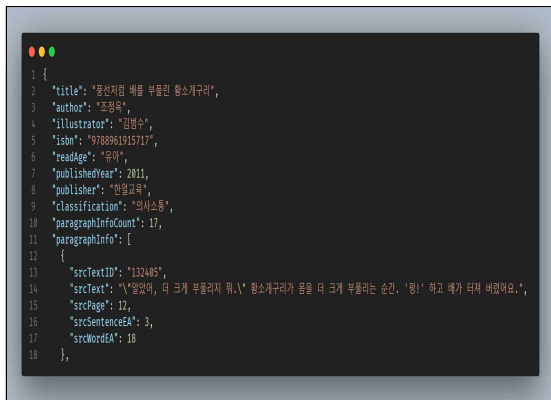


그림 2. 전래동화 내용 데이터셋  
Fig. 2. Traditional Fairy Tale Content Dataset

### 3.2 데이터 전처리

본 연구에서 수집한 동화 이미지 데이터셋과 동화 줄거리 데이터셋은 SDXL 모델과 GPT-4o 모델 학습을 위해 각각 전처리 과정을 거쳤다.

이미지 데이터셋의 경우, 동화 이미지의 화풍에 따라 캡션을 고전 유럽, 동양 판타지, 모던 판타지 등으로 지정하고, 트리거 코드를 각 화풍에 맞춰 할당하였다. 전처리 과정에서는 데이터셋을 화풍별로 정리하여, 모델이 트리거 코드에 따라 특정 화풍의 이미지를 생성할 수 있도록 하였다. 트리거 코드는 영어 단어가 아닌 한글을 그대로 영어로 타이핑하여 활용하였다.

표 1. 각 동화별 전처리 데이터셋

Table 1. Preprocessing Dataset for each fairy tale

화풍	개수	캡션	이미지
고전 유럽	200	rhwijsdbfjq	
동양 판타지	200	ehddidvksxkw	
모던 판타지	200	ahejsvksxkw	
수채화	200	tncoghk	
전래동화	200	wjsfoehdghk	

본 서비스의 동화 LLM 모델을 위해 데이터셋은 대화형 형식으로 구성하였다. 데이터셋은 system, user, assistant 역할로 구분되며, 각 역할에 맞는 내용을 설정하여 사용자 요청에 적합한 동화 이야기를 생성할 수 있도록 설계되었다. 주인공의 특징과 시대적 배경 등의 정보를 포함하여 모델이 맞춤형 동화를 생성할 수 있도록 하였으며, 각 대화가 자연스럽게 이어지도록 구성하였다.

또한, 동화 줄거리 데이터셋은 GPT-4o 모델 학습을 위해 텍스트 토큰화 전처리 과정을 거쳤다. 텍스트 데이터를 토큰화하여 모델이 동화의 흐름과 문맥을 더 효과적으로 파악할 수 있도록 하였으며, 이를 통해 다양한 사용자 설정에 맞춘 동화 내용을 생성한다. 이와 같은 전처리 과정을 통해 SDXL과 GPT-4o 모델은 본 서비스의 사용자 설정에 맞는 맞춤형 동화 내용과 일관성 있는 삽화를 생성할 수 있게 되었다.

### 3.3 인공지능 학습

본 연구에서는 사용자 요청에 맞춘 맞춤형 동화 내용과 삽화를 생성하기 위해 SDXL 모델과 GPT-4o 모델을 학습하였다. SDXL 모델은 전래동화, 동양 판타지, 모던 판타지 등 다양한 화풍을 반영한 이미지 데이터셋을 통해 각 화풍별로 적합한 이미지를 생성하도록 훈련되었으며, 각 화풍에는 전용 트리거 코드와 캡션이 포함되어 특정 스타일을 표현할 수 있도록 구성되었다.



표 2. 모델 파라미터 설정값  
Table 2. Parameters

파라미터	설정값
batch_size	8
learning_rate	5e-5
epochs	10
image_resolution	512x512
style_prompt	["rhwisehdghk"]

GPT-4o 모델은 동화 줄거리를 생성하기 위해 대화형 형식으로 구성된 데이터셋을 기반으로 파인튜닝을 진행하였다. 데이터셋은 system, user, assistant 역할로 구성되며, 사용자가 제공하는 주인공의 특징과 시대적 배경이 포함된 대화 내용을 모델이 자연스럽게 일관성 있게 이어가도록 설계되었다. 각 대화형 데이터셋은 텍스트를 토큰화하는 전처리 과정을 거쳐, 모델이 문맥과 이야기 흐름을 학습할 수 있도록 하였다. 이를 통해 모델은 사용자 설정에 따라 다양한 동화 줄거리를 자연스럽게 생성할 수 있는 역량을 갖추게 되었다.

### 3.4 웹 기반 서비스 구현

본 논문에서는 웹 기반 인터랙티브 동화 제작 시스템을 설계 및 구현하였다. 시스템의 인터페이스는 스토리 전개부와 시각화부로 구성된다. 스토리 전개부는 사용자의 선택지와 이야기 흐름을 제어하며, 시각화부는 생성형 AI가 실시간으로 제작한 삽화를 표시한다.



그림 3. 구현된 동화책 서비스  
Fig. 3. Implemented Storybook Service

그림 3은 구현된 시스템의 사용자 인터페이스를 보여준다. 우측의 스토리 전개부는 동화 내용과 두

가지 선택지를 포함하며, GPT-4o 모델이 사용자의 사전 설정값, 이전 줄거리, 선택한 분기점을 입력으로 받아 다음 스토리와 선택지를 생성한다. 좌측의 시각화부는 GPT-4o가 생성한 스토리를 기반으로 SDXL 모델이 연속적으로 삽화를 생성하여 화면에 출력한다.

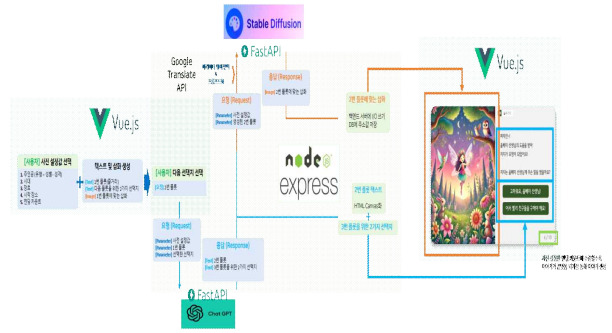


그림 4. 시스템 구조도  
Fig. 3. System Structure Chart

먼저 사용자는 주인공 유형, 시작 장소, 시대 및 장르, 엔딩 페이지 수와 같은 동화의 기본 설정을 선택하며, 선택된 설정값을 바탕으로 GPT-4o 모델은 동화 첫 장의 스토리와 선택지를 생성한다. 생성된 스토리는 SDXL 모델의 입력 프롬프트로 변환되어 삽화 생성에 연이어 사용되며, 텍스트와 삽화가 준비되면 시스템은 다음 페이지로 전환한다. 실시간 상호작용을 통해 사용자는 본인의 선택에 따라 줄거리와 결말이 실시간으로 변화하는 동화를 경험할 수 있다.

## IV. 결론

본 연구에서 개발한 서비스는 SDXL 모델과 GPT-4o 모델을 활용하여 사용자 요청에 맞춘 동화 삽화와 줄거리를 생성하는 시스템을 구축하였다. SDXL 모델은 다양한 화풍에 맞는 이미지 생성이 가능하도록 트리거 코드와 캡션 기반으로 학습하였으며, GPT-4o 모델은 주인공의 특징과 시대적 배경에 따라 자연스러운 동화 이야기를 생성할 수 있도록 대화형 형식으로 파인튜닝되었다. 이러한 모델 통합을 통해 사용자 맞춤형 콘텐츠 제공이 가능해

졌으며, 학습된 AI 모델은 높은 수준의 몰입감과 개인화된 동화 경험을 제공할 수 있는 가능성을 확인하였다. 앞으로 본 시스템의 성능을 향상시키고 사용자 피드백을 반영한 발전을 통해 교육과 창의성 향상에 기여할 수 있는 차세대 동화 콘텐츠 서비스를 구현하였다.

### 참 고 문 헌

- [1] M. K. Moon and H. Y. Jung, "A Study on the Use of Digital Technology in Early Childhood Education through Case Analysis of OECD Countries", Korea Institute of Child Care and Education Research Report 2021-08, pp. 19-20, December 2021.
- [2] J. W. Lee, W. S. Park, and J. W. Eom, "A Study on Policy Measures to Optimize Media Use of Infants and Young Children", Korea Institute of Child Care and Education Research Report 2021-15, pp. 17-20, December 2021.
- [3] D. Y. Yang, "The Effect of AI-based Art Fusion Education on the Creativity of Fifth-Grade Elementary Students." Master's Thesis, Gyeongin National University of Education Graduate School, Incheon, 2022.
- [4] J. H. Song, S. R. Shin, "Proposal for Making Fairy Tale Book Contents to Develop Creativity - Art Textbook Content Development through Constructivist Instructional Methodology." KSDS Conference Proceedings, Seoul, October 8, 2011.
- [5] G. H. Kim, T. Y. Na, "Case Study on Applying Image and Video Generative AI Technology for Automatic Fairy Tale Illustration Generation and Service Planning." Korean Broadcasting and Media Engineering Conference Proceedings, Jeju, June 25, 2024.
- [6] J. Y. Park, J. M. Shin, G. Y. Kim, J. H. Nam, and B. C. Bae, "A Prototype Design of an Interactive Story Authoring Tool Based on a Generative Language Model." Korea Information Science Society Conference Proceedings, Jeju, December 20, 2022.
- [7] J. Y. Park, H. M. Choi, H. S. Kim, K. H. Lim, and Zhongweijia. "TRPG Game Development by ChatGPT Prompting." Korea Information Science Society Conference Proceedings, Jeju, June 26, 2024.
- [8] G. M. Choi, J. H. Heo, and T. Y. Na, "Case Analysis of Constraints in AI-based Fairy Tale Illustration Generation Using Korean Prompts and Quantitative Expression." Korean Broadcasting and Media Engineering Conference Proceedings, Jeju, June 25, 2024.
- [9] M. H. Park, S. H. Lee, and Y. G. Lee, "Techniques for Creating Custom Article Illustrations with Stable Diffusion." Korean Broadcasting and Media Engineering Conference Proceedings, Jeju, June 25, 2024.
- [10] S. M. Myeong, S. M. Kang, D. B. Song, C. Y. Hong, H. Jeong, and S. H. Park, "A Study on Coherent Image Generation Reflecting Story Genres Using Text-based Image Generation Models." Korean Broadcasting and Media Engineering Conference Proceedings, Jeju, June 25, 2024.

# 양식장 배출수를 활용한 수력 터빈 발전 시스템의 전기요금 절감 효과 분석

김응표\*, 김영은\*, 윤신웅\*, 김경민\*, 이희진\*\*

## Hydro-Turbine Generation System Using Fish Farm's Water Analysis of Electricity Cost Reduction Effect

Eung-Pyo Kim\*, and Yeong-Eun Kim\*, Shin-Woong Yun\*, Kyoung-Min Kim\*, and Hee-Jin Lee\*\*

### 요 약

최근 개방형 양식장 내 수온 변화로 인한 어류 폐사와 전기요금 급등으로 인해 수산업이 심각한 위기를 겪고 있다. 이에 따라 개방형 양식장보다 폐쇄형 및 육상형 양식장의 비율이 증가하고 있다. 그러나 전기요금 상승 문제는 여전히 해결되지 않아 지속해서 폐업하는 양식장이 늘어나는 추세다. 따라서 이에 대한 실질적이고 혁신적인 해결책이 필요한 실정이다. 본 논문은 양식장의 전기요금 부담을 해소하고자 물을 순환하는 과정에 수력 터빈을 탑재하는 방안에 대해서 경제성을 분석한 내용을 담고 있다.

### Abstract

The aquaculture industry is currently facing significant challenges due to fluctuations in water temperature causing fish mortality and escalating electricity costs. Consequently, there is a shift from open to closed and land-based aquaculture systems. Despite this, the issue of soaring electricity prices remains unresolved, leading to an increasing number of business closures. Thus, effective solutions are urgently needed. This paper analyze the economic viability of integrating hydro turbine into the water circulation processes of aquaculture facilities to mitigate electricity costs.

### Key words

fisheries industry, hydro turbine, economic

## 1. 서 론

최근 지구온난화로 인해 바다의 수온이 급격하게 올라감에 따라 개방형 양식장보다 폐쇄형 양식장이 많아지고 있다. 통계청에서 발표한 2023년 어류양식 동향 조사 결과에 따르면 전국 양식장 총면적은

30,000m<sup>2</sup>가 감소했지만, 폐쇄형 양식장은 오히려 40,000m<sup>2</sup>의 면적이 증가한 것을 알 수 있다[1].

하지만 폐쇄형 양식장은 개방형 양식장과는 다르게 24시간 동안 펌프를 가동한다는 점 때문에 전기요금이 가중된다는 단점이 있다. 또한 전기요금의 지속적인 인상으로 인해 양식장은 또 다른 위기를

\* 국립금오공과대학교

\*\* 국립금오공과대학교 전자공학부 교수, 교신저자 email: jinlee@kumoh.ac.kr

맞고 있다. 제22대 국회 농림축산식품해양수산위원회 회의록에 따르면 2022년과 2023년에는 양식장 37곳이 전기요금의 부담으로 폐업한 사실을 발표했다[2]. 이처럼 급격한 기후 변화 및 전기요금 인상에 따라 양식장은 경영 악화, 폐업 등 심각한 위기를 맞고 있다. 하지만 양식업은 1차 산업으로써 국민에게 식량을 공급하고 야생 어류의 남획을 막아주는 국가의 주요 산업 중 하나다.

이런 문제를 해결하기 위해 본 논문은 양식장 배출수를 활용하여 수력 터빈을 사용해 전기를 생산하는 방안을 제안하며 발전 시스템의 이익을 경제적으로 분석했다.

본 논문의 시스템은 내륙지역에 분포해 있는 육상 양식장에 도입하기에 어려움이 있다.

하지만 그림 1에서 보이는 바와 같이 제주도 내 양식장은 해안가에 전부 분포해 있어 본 논문의 시스템을 도입하는 데 문제가 없다.

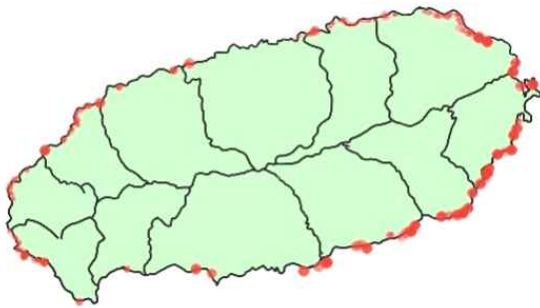


그림 1. 제주도 내 양식장 분포 현황  
Fig. 1. Distribution of fish farms in Jeju Island

## II. 수력 터빈 에너지 시스템

본 논문에서는 그림 2와 같이 양식장 배출수를 통해 수력 터빈을 가동하여 전력을 생산하는 것을 제안한다. 수력 터빈을 사용할 수 있는 환경은 유량이 많고 낙차가 굉장히 높은 댐과 같은 곳이다. 하지만 낮은 낙차로도 많은 유량이 공급된다면 육상형 양식장에서도 수력 터빈을 사용할 수 있는 환경이 된다. 수력 터빈 모델을 선정하기 위해서 수차를 15m로 가정[3], 유량 값은 제주 도내에 있는 331개의 양식장에 대한 평균값으로 계산했다. 제주 시에서 발표한 육상양식시설 배출수 수질기준 보고

서에 따르면 양식장 한 개의 평균 방류량은  $0.1206\text{m}^3/\text{s}$ 임을 알 수 있다[3].

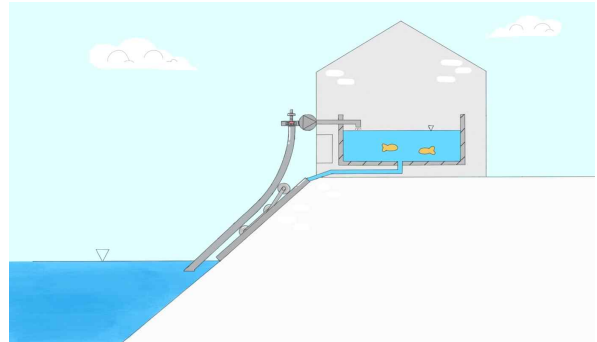


그림 2. 양식장 배출수를 활용한 발전 시스템  
Fig. 2. Power generation systems using fish farm's water

그림 3에서 보이는 M사의 프란시스 수력 터빈은 유량이 최소  $0.11\text{m}^3/\text{s}$ , 수차는 최소 10m의 기준을 요구한다.



그림 3. M사의 프란시스 터빈 3D 모델  
Fig. 3. Power generation systems using fish farm's water

표 1. M사 수력 터빈의 성능표  
Table 1. Performance table for the Company M turbine

터빈 매개 변수			발전기 매개 변수	
수차 [m]	유량 [ $\text{m}^3/\text{s}$ ]	출력 [kW]	생산 전력 [kW]	정격 속도 [r/min]
10	0.10	8	10	750
12	0.11	10	11	750
14	0.12	12	15	750
16	0.13	17	18	750
18	0.13	20	20	1000
20	0.14	22	25	1000

제주에 있는 양식장의 유량과 수차는 위의 가동 조건을 충족하기 때문에 본 논문에서는 M사의 프란시스 수력 터빈을 본 논문의 에너지 발전 시스템에 사용되는 모델로 선정했다.

### III. 수력 터빈 에너지 시스템의 경제성 분석

프란시스 터빈을 가동했을 때의 생산 전력은 다음과 같이 계산한다.

$$P = 9.8 \times Q \times H \times \eta \quad (1)$$

Q: 최대 사용 수량, H:낙차, η:발전기 효율

제주시에서 발표한 육상양식시설 배출수 수질기준 보고서에 따르면 양식장 한 개의 유량은 평균 0.12m<sup>3</sup>/s였다. 양식장 배출수의 낙차는 15m로 가정하고 M사의 프란시스 발전기의 효율은 90%로 표기되어있어 식에 대입하여 계산하면 양식장 내에 설치된 한 개의 발전기가 연간 약 139,215kWh의 전력을 생산한다는 결과가 도출됐다.

양식장은 농사용(을) 고압용 전기를 사용하는 것이 일반적이기 때문에 해당 전기요금으로 계산한다.

표 2. 2025년 기준 농사용 전기요금 표[5]  
Table 2. Table of agricultural electricity rates of 2025[5]

구분	전력량 요금 (원/kWh)	
	1월	4월
여름철 (6 ~ 8월)	66	68.6
봄·가을철 (3~5, 9 ~ 10월)	64	66.6
겨울철 (11 ~ 2월)	66	68.6

위의 요금표를 바탕으로 발전기의 전력 생산에 대한 소비 대체 요금을 계산하면 연간 약 8,542,240원의 전기요금 절감 효과를 한 개의 수력 터빈 설치로 얻을 수 있다. M사의 수력 터빈은 설계 수명이 자체적으로 40년이 넘기 때문에 한 개에 약 60,150,000원인 수력 터빈은 약 7년 4개월 이후부터는 순수익이 발생한다.

### IV 결 론

1차 산업은 한 나라의 식량 주권과 같다. 따라서 국가는 이를 보호하고 지원해야 한다. 하지만 1차 산업 중 하나인 양식업이 급격한 기후 변화와 지속적인 전기요금 인상으로 어려움을 겪고 있는 것이 현실이다. 이에 따른 해당 산업을 지원하고 지킬 방안이 필요하다.

본 논문은 전기요금에 대한 양식업의 부담을 해소하고자 양식장 배출수를 활용하여 수력 터빈을 탑재하는 아이디어를 제안하며 분석을 통해 경제적으로 연간 약 8,542,240원의 전기요금 절감 효과를 얻을 수 있어 약 7년 4개월 이후부터는 순수익이 발생한다는 것을 도출했다. 또한 제주도 내 있는 모든 양식장은 내륙지역이 아닌 해안지역에 자리 잡고 있으므로 본 논문의 시스템을 도입할 수 있는 환경이다. 도출된 결론을 바탕으로 설치할 수 있는 제주 관내의 모든 양식장에 수력 터빈을 설치하는 것을 제안한다. 물론 수력 터빈의 설치 비용이 많이 드는 것이 사실이고 제주 도내에 있는 모든 양식장에 설치하기에는 어려운 현실이다. 하지만 정부에서 이를 지원하고 도와준다면 국가 주요 산업 중 하나인 양식업을 지키는데 큰 도움이 될 것이다.

### 참 고 문 헌

- [1] 통계청 “2023년 어류양식 동향 조사 결과”, 2024. 03.
- [2] 제22대 국회 제416회 국회(임시회) 제1차, “농림축산식품해양수산위원회 회의록”, 2024. 07. 10
- [3] 김형호, 서상호, 김경엽 “양식장용 수력발전 수차의 설계에 관한 연구”, 한국유체기계학회 2020. 11.
- [4] 제주특별자치도 “육상양식시설 배출수 수질기준 설정 연구용역 보고서”, 2023.
- [5] 한국전력공사 “2025년도 전기요금 표 (종합)”, 2024.

# 고분자 물질이 함침된 탈리그닌화 목재를 이용한 광케이블

박제언\*, 김승준\*, 이희진\*\*

## Optical Cable Using Delignified Wood Impregnated with Polymer Material

Je-Eon Park\*, Seung-Jun Kim\*, and Hee-Jin Lee\*\*

### 요 약

고분자 물질이 함침된 탈리그닌화 목재를 사용하여 광케이블을 개발하는 방안을 제안한다. 목재에서 리그닌을 제거함으로써 투명성과 유연성을 크게 향상되었으며, 고분자 물질인 PVP와 PDMS를 함침시켜 코어와 클래딩 사이의 굴절률 차이를 증대시켰다. 이를 통해 빛의 손실을 줄이고 효율성을 높여 통신 산업에서 혁신적인 소재로 활용될 수 있음을 확인하였다. 이러한 결과는 제조 공정에서의 효율성 개선과 함께 통신 기술의 미래 발전에 기여할 것으로 기대된다.

### Abstract

A method for developing optical cables using delignified wood impregnated with polymer materials is proposed. By removing lignin from the wood, transparency and flexibility were significantly improved. Polymer materials such as PVP and PDMS were impregnated to enhance the refractive index difference between the core and the cladding. This approach reduces light loss and improves efficiency, confirming its potential as an innovative material in the telecommunications industry. These results are expected to contribute to the advancement of future communication technologies along with improvements in manufacturing efficiency.

### Key words

critical angle, delignified wood, polymer material, refractive index

## I. 서 론

전 세계적으로 통신 및 데이터 전송 수단으로서 광케이블에 대한 수요는 꾸준히 증가하고 있다. 기존의 광케이블은 주로 실리카를 기반으로 한 코어와 클래딩 구조로 구성되며, 이 두 층의 굴절률 차이를 이용한 전반사 원리로 빛을 전송한다. 실리카

는 우수한 광학적 성질과 내구성 덕분에 오랜 기간 광케이블의 주요 재료로 사용되어 왔으나, 제조 비용이 높고 물리적 특성에 한계가 있어 성능 개선을 위한 새로운 재료에 대한 연구가 요구되고 있다.

현재 탈리그닌화 목재를 이용한 연구가 다양한 분야에서 이루어지고 있다. 탈리그닌화된 목재에 에폭시 수지와 퍼플루오로데실트리에톡시실란을 적용

\* 국립금오공과대학교 전자공학부

\*\* 국립금오공과대학교 전자공학부 교수, 교신저자 email: jinlee@kumoh.ac.kr

하여 친수성과 투명성을 갖춘 목재를 제작하는 연구가 이루어졌으며, 이를 통해 스마트 창문, 윈드실드, 투명 장식 분야에서의 응용 가능성이 제안되었다.[1].

또 다른 연구에서는 탈리그닌화 목재에 폴리비닐알코올(PVA)과 메틸 메타크릴레이트(MMA)를 함침시켜 높은 투과율과 기계적 강도를 가진 투명 목재 복합재를 개발하였으며, 이를 디스플레이 패널 및 의료기기용 보호 커버와 같은 분야에 응용 가능성을 제시하였다[2]. 탈리그닌화 과정은 목재에서 리그닌을 제거하여 높은 투과성과 투명성을 가진 구조를 형성하는 공정으로, 이를 통해 목재는 빛의 투과율을 향상시키고 유연성과 내구성을 갖춘 신소재로 변모한다.

본 논문에서는 광케이블의 새로운 재료로써 탈리그닌화 목재를 활용하고 탈리그닌화 목재를 기반으로 한 광케이블이 가진 가능성을 높이기 위해 특정 고분자 물질을 목재에 함침시키는 공정을 통해 굴절률을 조정하고, 빛의 산란을 최소화하여, 광학적 전송 효율을 극대화할 수 있는 방법을 제시한다.

## II. 광케이블의 원리

광케이블은 빛을 이용한 데이터 전송을 가능하게 하는 기술로, 전반사 원리를 활용하여 정보를 전달한다. 광케이블의 주요 구성 요소인 코어와 클래딩 간의 굴절률 차이가 전반사를 발생시켜, 빛이 손실 없이 장거리로 전송될 수 있게 한다.

그림 1을 보면 알 수 있듯이, 빛은 한 매질에서 다른 매질로 이동할 때 굴절된다. 이때 스넬의 법칙이 적용되며, 코어의 굴절률 ( $n_1$ ), 클래딩의 굴절률 ( $n_2$ ), 입사각( $\theta_1$ ), 굴절각( $\theta_2$ )에 의해 표현된다.

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (1)$$

임계각은 빛이 코어에서 클래딩으로 나가려고 할 때, 빛이 굴절하지 않고 전반사되는 최소 입사각이고, 스넬의 법칙에서 굴절각이  $90^\circ$  가 될 때 입사각은 임계각( $\theta_c$ )이 된다.

$$\theta_c = \sin^{-1} \frac{n_2}{n_1} \quad (2)$$

임계각보다 큰 각도로 입사한 빛은 클래딩으로 굴절되지 않고, 전반사 현상이 발생하여 빛이 코어 내부에서 계속 반사하면서 이동하게 된다. 이 과정에서 빛은 클래딩을 통과하지 않고 코어 내에서만 전송되므로, 외부로의 손실이 거의 발생하지 않게 된다. 즉, 전반사 원리에 의해 빛은 코어를 따라 효율적으로 이동하며, 이를 통해 손실 없는 장거리 전송이 가능해진다. 이러한 전반사의 효과는 광케이블의 성능에 매우 중요한데, 임계각이 작을수록 더 큰 범위의 입사각에서 전반사가 발생할 수 있어, 빛이 코어 내부에서 더 잘 유지되고 손실이 더욱 줄어들게 된다. 결과적으로, 임계각이 작아지면 빛의 전송 효율이 높아지고, 광통신 시스템에서 더 긴 거리의 전송이 가능해진다.

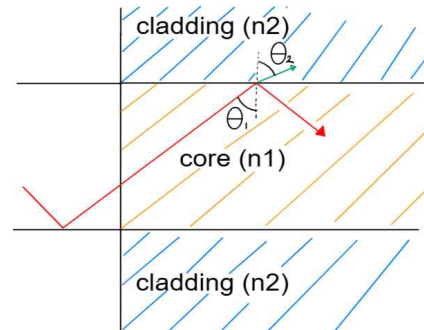


그림 1. 스넬의 법칙을 적용한 광케이블의 원리  
Fig. 1. Principle of optical cable applying Snell's law

## III. 고분자 함침 물질에 따른 굴절률

리그닌과 저분자 물질들이 제거된 목재는 셀룰로오스( $n=1.56$ )의 원래 투명한 성질을 나타내야 하지만, 셀룰로오스 세포벽과 공기층 사이의 굴절률 차이로 인해 빛이 산란되어 하얗게 보인다. 그러므로 고분자 물질을 함침시켜 탈리그닌화 목재의 투명성을 부여하고 고분자 물질에 차이를 두어 코어와 클래딩의 굴절률 차이를 최대화시킨다[3].

그림 2에서 볼 수 있듯이, 본 논문에서는 PVP가 첨가된 탈리그닌화 목재를 코어에 사용하고 PDMS가 첨가된 탈리그닌화 목재를 클래딩에 사용한다. PVP는 수용성 고분자이고 PDMS는 실리콘 기반 고분자로 PVP와 PDMS를 함침시키면 탈리그닌화 목재 내부의 미세한 공극을 채우면서 빛의 산란이 줄어들고 투명성이 증가하게 된다[4].

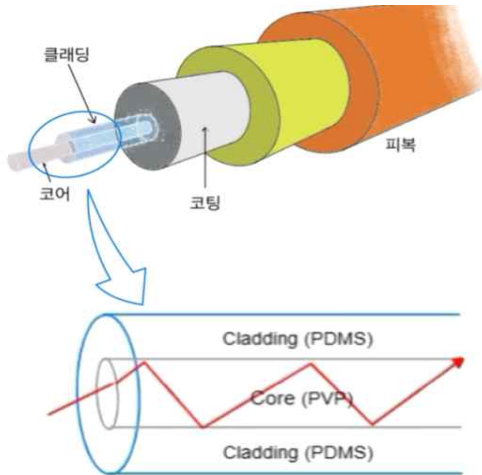


그림 2. 고분자 함침 물질을 적용한 탈리그닌화 목재 광케이블 구조

Fig. 2. Delignified wood optical cable structure using polymer impregnated material

표 1을 통해서 게르마늄과 불소가 소량 도핑된 기존 실리카 기반의 광케이블에서 임계각은 82.95도이다[5]. 본 논문에서 제시한 물질을 함침시킨 탈리그닌화 목재 광케이블의 임계각은 66.2도임을 알 수 있다.

기존 실리카 기반의 임계각 보다 약 16.75도 차이가 난다. 따라서 빛의 손실을 줄여 광학적 성능을 개선할 수 있다.

표 1. 고분자를 함침시킨 탈리그닌화 목재와 실리콘의 굴절률과 임계각 차이

Table 1. Differences in refractive index and critical angle of delignified wood and silicon impregnated with polymer

	탈리그닌화 목재		실리콘	
	PVP( $n_1$ )	PDMS( $n_2$ )	Germanium doped ( $n_1$ )	Fluorine doped ( $n_2$ )
굴절률( $n$ )	1.53	1.40	1.454	1.443
임계각( $\theta_c$ )	66.2°		82.95°	

#### IV. 결 론

본 논문에서는 기존 실리카 기반 광케이블의 한계를 극복하기 위한 대안으로, 탈리그닌화된 목재를

활용한 새로운 광케이블을 제안한다. 탈리그닌화 목재에 대한 연구는 다양한 분야에서 이미 수행되어 왔으나, 이를 광케이블 구조에 적용하고, 고분자 물질을 함침시키는 방법을 통해 광학적 성능을 극대화한 연구는 이루어지지 않았다. 특히, PVP와 PDMS를 함침하여 코어와 클래딩 간 굴절률 차이를 증대시키고, 그 결과 기존 실리카 기반 광케이블보다 빛의 손실을 효과적으로 감소시킬 수 있음을 확인하였다. 따라서 탈리그닌화 목재를 기반으로 한 광케이블은 향후 통신 기술의 효율성 증대뿐만 아니라, 고성능 광케이블 개발에 있어 중요한 기초 자료로 활용될 수 있을 것으로 기대된다.

#### 참 고 문 헌

- [1] Linhu Ding, Xiaoshuai Han, Lian Chen, Shanohug Jiang, "Preparation and properties of hydrophobic and transparent wood", Journal of Bioresources and Bioproducts, Vol. 7, No. 4, pp. 295-305, November 2022
- [2] Sivabalan Kr, Ramaraj E., "Surface Segmentation and Environment Change Analysis using Band Ratio Phenology Index Method-Supervised Aspect", IET Image Process, Vol. 14, No. 9, pp. 1813-1821, 2020
- [3] 윤제문, 김영훈, 정찬영, "다공구조 내에 고분자 충전된 유연한 투명나무 복합재 패널 제조 연구", 한국고분자학회 폴리머, 2023
- [4] 김영훈, 오원태, 윤제문, "투명나무 제조를 위해 리그닌 제거된 나무 패널에 다양한 고분자 함침에 관한 연구", 한국산학기술학회 논문지, 2024
- [5] Sanjeev Sharma, Arun Kumar "Analysis of silica based single-mode fiber doped with germanium at different transmission window", Silicon, Vol. 14, pp. 1023-1028, 2022



# 데이터 불균형을 처리하기 위한 증강 기법 비교

이민지\*, 김동현\*, 김재엽\*, 이상민\*, 이석훈\*†

## Comparison of Data Imbalance Handling Techniques

Minji Lee\*, Donghyun Kim\*, Jaeyeob Kim\*, Sangmin Lee\*, and Sukhoon Lee\*†

### 요약

의료 데이터의 특성상 정상인과 환자 데이터 간의 불균형이 발생하며, 이는 모델의 성능을 저하시키는 주요 요인이 된다. 본 연구에서는 치매 진단을 위한 웨어러블 라이프로그 데이터의 불균형 문제를 해결하기 위해 SMOTE와 VAE 두 가지 데이터 증강 기법을 적용하고 비교 분석하였다. 실험 결과, VAE는 Precision 측면에서, SMOTE는 Recall과 F1-score 측면에서 우수한 성능을 보였으며, 두 기법 모두 원본 데이터셋 대비 유의미한 성능 향상을 달성하였다.

### Abstract

Due to the nature of medical data, there is an imbalance between normal people and patient data, which is the main factor that degrades model performance. In this study, two data augmentation techniques, SMOTE and VAE, were applied and comparatively analyzed to solve the imbalance problem of wearable life log data for dementia diagnosis. Experimental results show that VAE performs better on the Precision side and SMOTE performs better on both Recall and F1-score, and both techniques achieve significant performance improvements compared to the original dataset.

### Key words

Data imbalance, SMOTE, VAE, Data augmentation

## 1. 서론

최근 딥러닝 기술의 발전과 함께 라이프로그 데이터를 활용한 연구가 활발히 진행되고 있다. 그러나 실제 데이터 수집 과정에서 발생하는 클래스 간 데이터 불균형(Class Imbalance) 문제가 발생하고 있으며 이는 모델의 성능을 저하시키는 주요 요인으로 지적되고 있다.

데이터 불균형이란 특정 클래스에 속하는 샘플 수가 다른 클래스에 비해 현저히 적은 상황을 의미하며, 이는 모델이 다수 클래스에 편향되어 학습되는 결과를 초래할 수 있다[1].

특히 의료 분야에서는 이러한 데이터 불균형 문제가 더욱 두드러진다. 예를 들어, 치매 연구에서는 정상인 데이터에 비해 환자 데이터의 수집이 제한적일 수밖에 없어 필연적으로 데이터 불균형이 발

\* 국립군산대학교 소프트웨어학과

† 교신저자

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음 (2023-0-00065).

생한다[2]. 이러한 불균형은 희귀 질환이나 특정 증상의 발생 빈도가 낮은 경우, 데이터 수집의 물리적 제약, 프라이버시 보호 문제 등 다양한 요인에 기인한다.

데이터 불균형 문제는 단순히 데이터의 양적 차이를 넘어, 모델의 예측 성능에 직접적인 영향을 미쳐 의료 진단과 같은 중요한 의사결정 과정에서 심각한 결과를 초래할 수 있다.

따라서 이 논문에서는 데이터 불균형 문제를 해결하기 위해 라이프로그 데이터를 활용하여 다양한 기법들을 비교 분석한다.

## II. 데이터 불균형 처리 기법

### 2.1 데이터셋

이 연구는 한국지능정보사회진흥원의 AI-Hub에서 제공하는 '치매 고위험군 웨어러블 라이프로그' 데이터셋을 활용하였다[3]. 데이터는 55세 이상 성인 300명을 대상으로 수집되었으며, 전문의 진단을 통해 정상인지군(CN), 경도인지장애(MCI), 치매군(Dem)으로 분류되었다. 최종 데이터셋은 174명의 12,183건의 일일 라이프로그로 구성되어 있다. 데이터의 품질 향상을 위해 Isolation Forest 알고리즘을 적용하여 상위 10%의 이상치를 제거하였으며, 그 결과 CN 클래스 7,075개, MCI 클래스 3,374개, Dem 클래스 515개로 구성된 불균형 데이터셋이 도출되었다.

### 2.2 SMOTE

SMOTE(Synthetic Minority Over-sampling Technique)[4]는 소수 클래스의 데이터 분포를 확장하는 증강 기법이며 소수 클래스의 각 샘플에 대해 k-최근접 이웃을 식별하고, 해당 샘플들 간의 특성 공간에서 선형 보간을 통해 새로운 합성 샘플을 생성한다. 이 논문에서는 SMOTE를 활용하여 소수 클래스인 MCI와 Dem 클래스를 각각 5,000개로 증강하였으며, 다수 클래스인 CN은 무작위 추출을 통해 5,000개로 조정하여 클래스 간 균형을 맞추었다.

### 2.3 VAE

VAE(Variational Autoencoder)[5]는 확률적 생성 모델로, 기존 오토인코더의 확장된 형태이다. 이 모델

은 인코더가 입력 데이터를 잠재 공간(latent space)의 확률 분포로 매핑하고, 디코더가 이 잠재 변수로부터 원본과 유사한 새로운 데이터를 생성하는 구조를 가진다. SMOTE와 마찬가지로 VAE를 통해 MCI와 Dem 클래스를 각각 5,000개로 증강하였으며, CN 클래스는 무작위 추출을 통해 동일한 수로 조정하였다.

### 2.4 실험 및 평가

이 논문에서는 XGBoost 알고리즘을 기반으로 분류 모델을 구축하였다. 실험은 이상치가 제거된 원본 데이터셋, SMOTE 적용 데이터셋, VAE 적용 데이터셋을 대상으로 수행되었다. 각 데이터셋은 8:1:1의 비율로 훈련, 검증, 테스트 세트로 분할하였다.

모델의 하이퍼파라미터는 최대 트리 깊이 6, 학습률 0.05로 설정하였으며, 과적합 방지를 위해 subsample 0.7, colsample\_bytree 0.8을 적용하였다. 또한, 검증 세트의 다중 클래스 로그 손실을 기반으로 한 조기 종료 전략을 채택하여, 10회 연속으로 성능 개선이 없을 경우 학습을 중단하였다.

모델의 성능은 Precision, Recall, F1-score를 통해 평가하였다. Precision은 모델의 양성 예측 중 실제 양성 비율을, Recall은 실제 양성 중 모델이 정확히 예측한 비율을 의미한다. F1-score는 이 두 지표의 조화평균으로, 모델의 전반적인 성능을 나타낸다.

표 1은 이상치를 제거한 원본 데이터로 XGBoost 모델을 평가한 결과를 보여준다. CN 클래스는 높은 성능을 보였으나, MCI와 Dem 클래스는 낮은 재현율과 F1-score를 기록하였다. 특히, Dem 클래스의 낮은 재현율은 불균형 클래스에 대한 성능 저하를 나타낸다.

표 2는 SMOTE로 증강한 데이터를 XGBoost로 평가한 결과를 보여준다. CN 클래스는 가장 높은 성능을 보였고, MCI는 상대적으로 낮은 정밀도와 F1-score를, Dem은 MCI보다 나은 성능이지만 여전히 CN에 미치지 못했다.

표 3은 VAE로 증강한 데이터의 성능을 보여준다. CN 클래스의 F1-score는 가장 높았으나, MCI와 Dem 클래스는 SMOTE(표 2)에 비해 정확도는 높지

표 1. 원본 데이터 성능 비교

Table 1. Original data performance comparison

Class	Precision	Recall	F1-score
CN	0.8483	0.9433	0.8933
MCI	0.8333	0.6769	0.7470
Dem	1.0000	0.5918	0.7436

표 2. SMOTE 증강 데이터 성능 비교

Table 2. SMOTE Augmented Data Performance Comparison

Class	Precision	Recall	F1-score
CN	0.8623	0.9025	0.8820
MCI	0.7846	0.7219	0.7519
Dem	0.8696	0.7692	0.8163

표 3. VAE 증강 데이터 성능 비교

Table 3. VAE Augmented Data Performance Comparison

Class	Precision	Recall	F1-score
CN	0.8592	0.9321	0.8942
MCI	0.8265	0.7211	0.7702
Dem	0.9706	0.6471	0.7765

표 4. 데이터 증강 기법별 성능 비교

Table 1. Performance Comparison of Different Data Augmentation Methods

Method	Precision	Recall	F1-score
Original Data	0.8939	0.7374	0.7646
SMOTE	0.8388	0.7979	0.8167
VAE	0.8854	0.7667	0.8136

만 낮은 재현율로 성능 개선이 제한적이었다.

최종 실험 결과는 표 4과 같다. VAE 기반의 데이터 증강 방법이 0.8854의 Precision으로 가장 높은 예측 정확도를 보였으며, SMOTE 기법은 0.7979의 Recall과 0.8167의 F1-score로 가장 우수한 종합 성능을 나타냈다.

### III. 결 론

이 논문에서는 치매 진단을 위한 웨어러블 데이터의 불균형 문제를 해결하기 위해 SMOTE와 VAE 두 가지 데이터 증강 기법을 적용하고 그 효과를 비교 분석하였다. 실험 결과, 두 증강 기법 모두 원본 데이터셋 대비 유의미한 성능 향상을 보였다.

SMOTE는 0.8167의 F1-score로 가장 높은 종합 성능을 보였는데, 이는 기존 데이터의 특성을 잘 보존하면서도 효과적으로 소수 클래스의 샘플을 생성했기 때문으로 분석된다. VAE의 경우 0.8854의 높은 Precision을 기록하여, 생성된 데이터의 정확성 측면에서 강점을 보였다.

향후 연구에서는 두 증강 기법의 장점을 결합한 하이브리드 접근법이나, 데이터의 특성을 더 잘 반영할 수 있는 증강 기법의 개발이 필요할 것으로 보인다. 또한, 생성된 데이터의 품질을 더욱 개선하고 평가하기 위한 추가적인 방법론 연구도 필요할 것으로 판단된다.

### 참 고 문 헌

- [1] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *Journal of Big Data*, Vol. 6, No. 1, pp. 27, 2019. <https://doi.org/10.1186/s40537-019-0192-5>
- [2] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 21, No. 9, pp. 1263-1284, 2009. <https://doi.org/10.1109/TKDE.2008.239>
- [3] AI-Hub: <https://aihub.or.kr/>. [accessed: Nov. 04, 2024]
- [4] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, Vol. 16, pp. 321-357, 2002.
- [5] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *Proc. International Conference on Learning Representations (ICLR)*, arXiv:1312.6114, 2014.

# 한국어 검색 증강 생성을 위한 오픈소스 벡터 스토어의 성능 비교

우민식\*, 시종욱\*\*, 김성영\*\*\*

## Performance Comparisons of Open Source Vector Stores for Retrieval-Augmented Generation in Korean

Minsik Woo\*, Jongwook Si\*\*, and Sungyoung Kim\*\*\*

### 요약

sLLM의 한계를 보완하기 위해 제안된 RAG는 실시간으로 외부 데이터를 검색하여 생성 과정에 반영함으로써 맥락에 맞는 응답을 생성할 수 있다. 이에 따라 본 논문에서는 한국어 RAG를 구현하는 데 적합한 오픈소스 벡터스토어의 성능을 평가한다. 이를 위해 4가지 오픈소스 벡터스토어와 4가지 임베딩 모델을 사용하여 데이터 업로드 시간, 검색 시간, 정확도를 기준으로 정량적 및 정성적 분석을 수행하였다. 분석 결과, FAISS는 삽입 및 검색 속도에서 가장 뛰어난 성능을 보였으며, 정확도에서는 ChromaDB가 가장 우수한 결과를 기록했다.

### Abstract

RAG proposed to address the limitations of sLLMs, enhances the generation process by retrieving external data in real-time to produce contextually relevant responses. This paper evaluates the performance of open-source vector stores for implementing Korean RAG. Four open-source vector stores and four embedding models were employed to conduct quantitative and qualitative analyses, focusing on data upload time, search time, and accuracy. The results demonstrated that FAISS exhibited superior performance in both data insertion and search speed, while ChromaDB outperformed the others in terms of accuracy.

### Key words

LLM, sLLM, RAG, Vector Store, NLP

## 1. 서론

가벼운 모델 구조로 대규모 데이터를 처리할 수 있도록 설계된 sLLM은 전문적인 작업이나 특정 언어에서, LLM에 비해 성능이 떨어지는 경우가 많다.

따라서, sLLM의 전문적인 작업에 대한 파인튜닝이 필수적이다. 그러나 파인튜닝 과정은 많은 자원과 시간이 소모되는 등의 단점이 있으며, 이러한 한계점들을 극복하기 위해 RAG[1] 기술이 제안되었다. RAG[1]는 외부 데이터를 실시간으로 검색해 생성

\* 국립금오공과대학교 컴퓨터공학과

\*\* 국립금오공과대학교 컴퓨터·AI융합공학과

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

※ 본 연구는 2023년도 중소벤처기업부의 기술개발사업 지원에 의한 연구임 IS33448821

과정에 반영하는 방식으로 학습 데이터에 포함되지 않은 정보를 실시간으로 가져와, 보다 정교하고 맥락에 맞는 답변을 생성할 수 있는 장점이 있다.

## II. 벡터 스토어 성능 평가

본 논문에서는 한국어를 대상으로 RAG 모델을 활용하기 위한 다양한 오픈소스 벡터 스토어의 성능을 평가하고, 각 벡터 스토어의 적합성을 분석한다. 실험에 사용된 벡터 스토어는 FAISS[2], ChromaDB[3], Milvus[4], Weaviate[5]로 선정하였으며, 성능 평가를 위한 임베딩 모델은 bge-m3[6], kr-sbert[7], bert[8], kobert[9]을 사용하여 수행하였다. 평가에는 각 벡터 스토어에 데이터를 업로드하는 시간과 검색 시간을 기준으로 하는 정량적 평가, 검색된 문서의 관련성을 판단하여 정확도를 측정하는 정성적 평가로 진행되었다. 벡터 스토어는 코사인 유사도를 기준으로 상위 5개의 문서를 검색하도록 구성하였으며, 모든 조합에 대해 실험하였으며 결과는 표 1과 같다.

실험에 사용된 데이터는 AIHUB[10]의 문서 요약 데이터셋 중 신문 요약 데이터셋에서 1,000개의 데이터를 추출하여 벡터 스토어에 삽입하였으며, 쿼리는 "교통 사고 관련 내용을 알려주세요"로 설정하였다. 실험 결과, 업로드 시간과 검색 시간 측면에서는 FAISS[2]가 가장 우수한 성능을 보였으며, 검색된 문서의 정확도 면에서는 ChromaDB[3]가 가장 성능이 높았다.

표 1. 벡터 스토어와 임베딩 모델의 성능 비교

Table 1. Performance Comparisons with vector store and embedding models

	Metric	FAISS [2]	ChromaDB [3]	Milvus [4]	Weaviate [5]
bge-m3 [6]	Upload Time	<b>972.56</b>	1030.84	1006.71	997.85
	Search Time	<b>0.0006</b>	0.0304	0.0421	0.0174
	Accuracy	4/5	<b>5/5</b>	4/5	4/5
kr-sbert [7]	Upload Time	<b>295.83</b>	314.87	312.24	309.68
	Search Time	<b>0.0005</b>	0.0120	0.0092	0.0089
	Accuracy	1/5	<b>3/5</b>	2/5	2/5
bert [8]	Upload Time	316.67	314.42	314.62	<b>312.51</b>
	Search Time	<b>0.0011</b>	0.0130	0.0051	0.0094
	Accuracy	1/5	0/5	0/5	0/5
kobert [9]	Upload Time	<b>306.64</b>	322.16	310.16	311.41
	Search Time	<b>0.0005</b>	0.0091	0.0043	0.0090
	Accuracy	0/5	<b>1/5</b>	0/5	0/5

## III. 결 론

본 논문에서는 한국어 RAG를 효율적으로 수행하기 위해 적합한 오픈소스 벡터 스토어를 평가하였다. 성능 평가 결과로 FAISS는 업로드 및 검색 시간 측면에서, ChromaDB는 정확도 측면에서 가장 우수한 성능을 나타냈다.

## 참 고 문 헌

- [1] P. Lewis, et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems*, Vol. 33, pp. 9459-9471, 2020.
- [2] <https://ai.meta.com/tools/faiss> [accessed: Oct. 14, 2024]
- [3] <https://docs.trychroma.com> [accessed: Oct. 14, 2024]
- [4] <https://milvus.io>, [accessed: Oct. 14, 2024]
- [5] <https://weaviate.io> [accessed: Oct. 14, 2024]
- [6] <https://huggingface.co/BAAI/bge-m3>, [accessed : Oct. 14, 2024]
- [7] <https://huggingface.co/snunlp/KR-SBERT-V40K-klue-NLI-augSTS> [accessed ed: Oct. 14, 2024]
- [8] <https://huggingface.co/google-bert/bert-base-uncased> [accessed: Oct. 14, 2024]
- [9] <https://huggingface.co/skt/kobert-base-v1> [accessed: Oct. 14, 2024]
- [10] <https://aihub.or.kr/aihubdata/data/view.do?currMenu=&topMenu=&aihubDataSe=data&dataSetSn=97>, [accessed: Oct. 14, 2024]

# RF Heating 기반 초고온 플라즈마 에너지 가열 효율 예측을 위한 기계 학습 모델링

김규민\*, 이정표\*\*, 김유섭\*\*\*

## Machine Learning Modeling for Predicting Heating Efficiency of High-Temperature Plasma Based on RF Heating

Kyu-Min Kim\*, Jung-Pyo Lee\*\*, and Yu-Seop Kim\*\*\*

### 요 약

본 논문은 KSTAR 토카막(Tokamak, 핵융합로)에서의 핵융합 플라즈마(Plasma) 가열 효율을 기계학습(Machine Learning, ML)으로 예측하고 최적화하는 것을 목표로 한다. KSTAR의 토카막 내에서 플라즈마는 최대 1억°C까지 가열되며, 주로 Ohmic Heating, Neutral Beam Injection, Radio Frequency(RF) Heating 세 가지 방법으로 가열된다. 그중 RF Heating은 전자기파를 이용하여 플라즈마 내 전자와 이온에 에너지를 전달하는 효율적인 가열 방식이지만, 비선형적인 환경 요인으로 인해 가열 효율이 변할 수 있다. 본 논문에서는 기계학습을 통해 RF Heating의 효율과 위치에 따른 에너지 전달 프로파일을 예측하고자 한다. 구체적으로, RF Heating 환경 변수에 따른 0-D 효율 예측 모델, 위치에 따른 1-D 파워 프로파일 모델, 역문제(Inverse problem) 해결을 위한 기계 학습 기법을 성능 평가하였다.

### Abstract

This paper aims to predict and optimize the efficiency of fusion plasma heating in the KSTAR Tokamak through machine learning methodologies. In the KSTAR Tokamak, plasma is heated to temperatures up to 100 million degrees Celsius, primarily employing three methods: ohmic heating, neutral beam injection, and radio frequency (RF) heating. Among these, RF heating is particularly efficient as it transfers energy to electrons and ions in the plasma via electromagnetic waves. However, the efficiency of RF heating can be affected by nonlinear environmental factors. This study leverages machine learning models to predict RF heating efficiency and spatial energy distribution profiles. Specifically, models for 0-dimensional (0-D) efficiency prediction based on RF heating environmental variables, 1-dimensional (1-D) power profile estimation, and inverse problem approaches were developed and evaluated in terms of performance.

### Key words

KSTAR Tokamak, Radio Frequency Heating, Machine Learning Prediction

\* 한림대학교 정보과학대학 소프트웨어학부, kyumin6379@gmail.com

\*\* 한양대학교 공과대학 원자력공학과 교수, jungpyo@hanyang.ac.kr

\*\*\* 한림대학교 정보과학대학 소프트웨어학부 교수, yskim01@hallym.ac.kr (교신저자)

※ 본 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원[RS-2021-II212068, 인공 지능 혁신 허브 연구 개발 (서울대학교, AI연구원)]과 2023년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2023-00254695).

## 1. 서 론

핵융합 에너지는 차세대 에너지원으로서 환경친화적이고 안전한 특성을 보인다. 현재 한국핵융합연구원(KFE)에서 가용 중인 핵융합연구장치 KSTAR는 실증로를 건설하기 위한 기반 기술을 확보하고 있으며, 이를 통해 국내 핵융합 연구 역량을 강화하는 것을 목표로 하고 있다. 특히 초고온 플라즈마(Plasma) 실험은 핵융합 에너지 개발에 있어 중요한 단계이다. 이러한 KSTAR의 성공적인 운용을 위해서는 Tokamak 내부에 존재하는 플라즈마를 1억℃ 이상으로 높이는 것이 필수적이다. 따라서 플라즈마를 효율적으로 가열하기 위한 다양한 방법에 대한 연구가 필요하다. 핵융합을 위한 플라즈마의 가열 방식은 전류에 의한 가열 (ohmic heating), 중성입자빔 주입(Neutral beam injection), Radio frequency(RF)의 세 가지 방식이 존재한다 [1]. 이 중 RF heating은 외부로부터 전자기파를 조사하여 내부의 전자 및 이온을 공명시키는 것으로 플라즈마를 가열하는 방법이다. RF heating은 매우 효과적인 가열 방식이나, 전자기파의 편향이나 내부 플라즈마의 환경에 의하여 가열 효율이 100%를 밑돌며, Tokamak 내부의 위치에 따라서, 전자가 RF Beam으로부터 받을 수 있는 에너지의 양 또한 달라질 수 있다 [2].

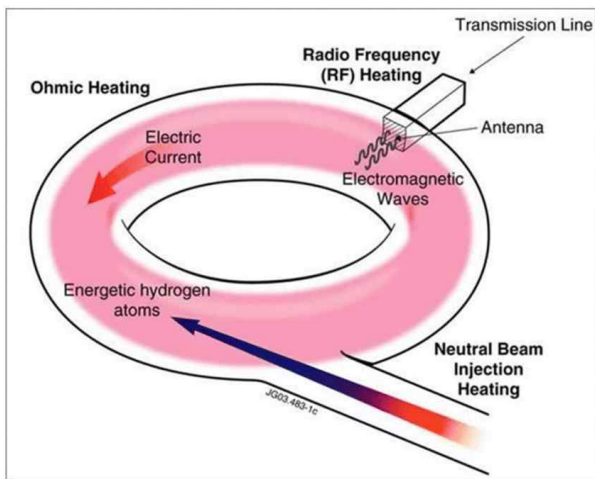


그림 1. 핵융합로에서의 플라즈마 가열 방법  
Fig. 1. Plasma heating methods in a tokamak

이러한 토카막 내부의 RF heating의 효율 및 에너지 전달에 대한 계산은 Ray tracking 코드를 이용하여 계산을 진행할 수 있으나, 여러 가지 환경변수

에 대한 RF heating의 관계식은 매우 비선형적이기 때문에, 이를 계산하기 위해서는 많은 자원을 필요로 한다. 기존 연구들은 신경망 모델을 주로 사용했고[3], 따라서 본 논문에서는 기계학습(Machine Learning, ML)을 이용하여 비선형적인 관계를 예측할 수 있는 모델을 개발하고, 이를 비교하고자 한다.

## II. 방법론

### 2.1 Dataset

#### 2.1.1 0-D, Inverse problem

KSTAR Tokamak의 전자공명가열 장치는 RF 주파수와 Gyromotion 주파수가 맞을 때, 공명이 일어나서 RF 에너지가 플라즈마에 전달된다. 플라즈마에 전달되는 파워 효율 예측을 위해 0-D와 Inverse problem에 대해서는 웨이브 편극과 전자공명가열 장치의 쏘는 각도, 플라즈마의 환경에 따른 에너지 가열 데이터들을 이산화하여 기계학습을 진행한다. (총 1250(2\*5\*5\*5\*5)개의 데이터)

#### - RF Beam

- RF heating mode: O or X mode (2개의 이산화 데이터)
- Toroidal injection angle: 1.60 ~ 1.80 (5개의 이산화 데이터)
- Poroidal injection angle: 22 ~ 30 (5개의 이산화 데이터)

#### - Plasma 환경

- Central density:  $1 \cdot 10^{-19} m^{-3} \sim 5 \cdot 10^{-19} m^{-3}$  (5개의 이산화 데이터)
- Central temperature: 1 ~ 5 keV (5개의 이산화 데이터)

#### 2.1.2 1-D problem

1-D 문제는 전자가 가열되는 위치를 고려하여, 반경 방향 공간에 대한 가열 효율의 프로파일을 데

이터로 사용한다. 0-D Problem의 Dataset에서 Tokamak의 부반경(rho)을 고려한다.

(총 250,000(2\*5\*5\*5\*5\*200(rho))개의 데이터)

- Tokamak 구조

- rho (r/a): 0.0025 ~ 0.9975 (200개의 이산화 데이터, Step by 0.0005)

## 2.2 Model

### 2.2.1 0-D problem

비선형적인 RF 에너지 흡수 효율을 예측하기 위해 다양한 기계학습 모델을 사용하였다. 앙상블(Ensemble) 기반의 회귀(Regressor) 모델[4]들을 사용하였고, 사용한 모델은 표 1과 같다.

표 1. 0-D, Inverse problem 예측 사용 모델

Table. 1. 0-D, Inverse problem Prediction model

ML Model(Ensemble based Regressor)
RandomForestRegressor
AdaBoostRegressor
BaggingRegressor
ExtraTreesRegressor
GradientBoostingRegressor
HistGradientBoostingRegressor
StakingRegressor
VotingRegressor

모델의 평가는 평균제곱오차(Mean Square Error, MSE)로 진행했다. MSE의 수식은 (1)과 같다.

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y_i^{real} - Y_i^{predicted})^2 \quad (1)$$

### 2.2.2 1-D problem

1-D 문제에서는 추가로 늘어난 feature와 데이터를 고려하여 기존 기계학습 모델이 아닌 LSTM[5]과 Transformer Decoder 기반의 GPT-2 모델[6, 7]을 사용했다.

1-D 문제 해결에는 두 가지 방법으로 진행하였다. 첫 번째로 0-D와 똑같은 방식으로 새로 추가된 rho를 포함한 6개의 feature로 label인 파워 프로파일

을 예측하는 방법이고, 두 번째로는 rho를 시계열 데이터의 순서로 사용하여 예측하는 방법이다. 5-fold 교차 검증을 수행하면 그림 3과 같이 나온다.

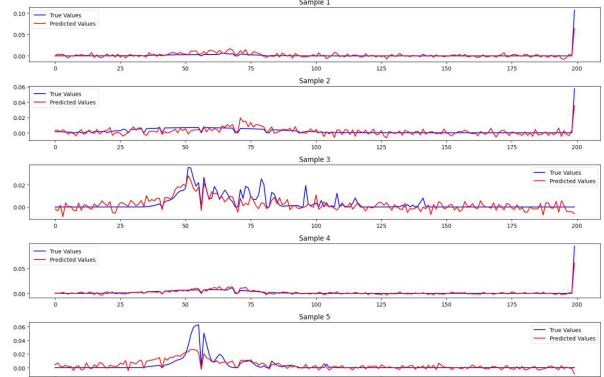


그림 3. 5-Fold 교차 검증(LSTM Model)

Fig. 3. 5-Fold Cross Validation using the LSTM Model

그림 3을 보면 190~200의 rho index에서 실제 물리적 현상으로 보이지 않는 Peak noise가 등장하여, 제거하는 전처리과정을 진행하고 rho를 시계열 데이터의 순서로 사용하여 예측하였다.

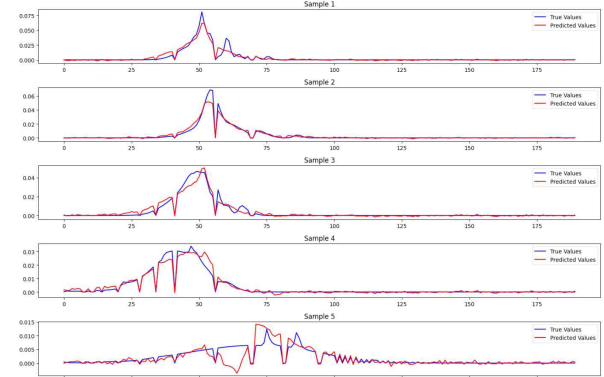


그림 4. 전처리 후 5-Fold 교차 검증(LSTMt Model)

Fig. 4. 5-Fold Cross Validation using the LSTMt Model after preprocessing

1-D problem은 평균절대오차(MAE)를 평가 지표로 활용했다. 첫 번째 방법에 대해서는 (2)의 수식을 사용해 평가했고, 두 번째 방법에 대해서는 (3)의 수식에 사용해 평가했다.

$$MAE = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_i|}{n} \quad (2)$$

$$MAE = \frac{1}{N} \frac{1}{H} \sum_{i=1}^N \sum_{t=T+1}^{T+H} |y_{i,t} - f_{i,t}| \quad (3)$$



### 2.2.3 Inverse problem

Inverse problem은 기존 문제의 입력 값과 출력 값을 바꿔주는 것으로, power absorption, 즉 기존 label에 대한 feature 값을 추측하는 함수 f를 최적화하는 문제이다. 하지만, 기존 1개의 출력값으로 나머지 5개 feature 값을 예측하는 것은 사실상 불가능하다.

본 논문은 그런 점을 고려하여, 1개의 출력값과 4개의 파라미터를 feature로 두고 나머지 1개의 파라미터를 예측하는 것과 1개의 출력값과 3개의 파라미터를 가지고 나머지 2개의 파라미터를 예측하는 시도를 하였다. 사용한 모델은 0-D problem과 동일한 표 1의 모델들을 사용했다. 0-D, 1-D problem과 다르게 RF 에너지 흡수 효율을 예측하는 문제가 아니고, 5개의 이산화 데이터 중 예측하는 문제이기 때문에, 평가 지표는 정확도, 수식 (4)를 따른다.

$$ACC = \frac{\sum_{i=1}^N I(y_i = \hat{y}_i)}{N} \quad (4)$$

## III. 결 과

### 3.1 0-D problem

0-D 에너지 흡수 효율 예측 평가 결과는 표 2와 같다.

표 2. 0-Dimension Problem 모델 예측 평가 결과  
Table 2. Evaluation Results of Model Predictions for the 0-Dimension Problem

ML Model	MSE
<b>RandomForest</b>	<b><math>3.18 \times 10^{-4}</math></b>
AdaBoost	$2.33 \times 10^{-3}$
Bagging	$3.46 \times 10^{-4}$
<b>ExtraTrees</b>	<b><math>3.06 \times 10^{-4}</math></b>
GradientBoosting	$8.25 \times 10^{-4}$
HistGradientBoosting	$3.42 \times 10^{-4}$
Staking	$3.06 \times 10^{-3}$
Voting	$5.13 \times 10^{-4}$

그림 2는 MSE가 가장 높은 두 모델에 대해 산점도(Scatter Plot)를 나타낸 그림이다.

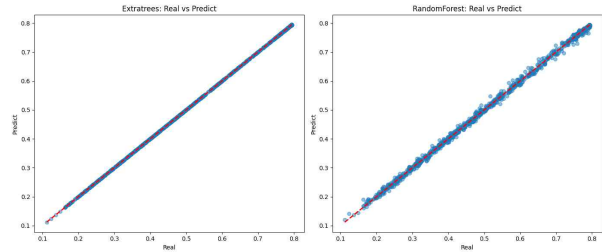


그림 2. 0-D 산점도(좌:Extratrees, 우:RandomForest)  
Fig. 2. 0-D Scatter Plot (Left: ExtraTrees, Right: RandomForest)

### 3.2 1-D problem

표 3. 1-Dimension Problem 모델 예측 평가 결과  
Table 3. 1. Evaluation Results of Model Predictions for the 0-Dimension Problem

Model	MAE
LSTM	$2.95 \times 10^{-3}$
GPT-2	$2.49 \times 10^{-3}$
<b>LSTMt(time series)</b>	<b><math>1.27 \times 10^{-3}</math></b>
<b>GPT-2t(time series)</b>	<b><math>2.25 \times 10^{-3}</math></b>

표 3에서 볼 수 있듯이, 기존 조건에서 추가로 peak noise를 제거하고 rho를 time series로 적용한 뒤 MAE가 절반 정도로 줄어든 것을 확인할 수 있다.

### 3.3 Inverse problem

표 4. Inverse problem 1개의 특성 예측  
Table 4. 1. Inverse problem 1 feature prediction

Feature	Best Model	Acc
ox_mode	RandomForest	1.000
ne center	RandomForest	0.964
te center	RandomForest	0.808
tor_angle	RandomForest	0.984
pol_angle	RandomForest	0.864

표 4를 보면 1개의 특성 예측은 비교적 간단한 작업이라 대부분 좋은 성능을 보인다. 그 중 RandomForest 모델[8]이 모든 Feature에 대해서 가장

좋은 성능을 보였다. feature 중 Tor\_angle과 Pol\_angle 두 RF beam의 입사각을 예측하는 것은 중요하다. 그래서 2개의 feature를 예측하는 실험은 두 RF beam 입사각에 대해 진행했다.

표 5. Inverse problem 2개의 특성 예측

Feature	Best model	Train Acc	Test Acc
Tor_angle	RandomForest	0.842	0.464
Pol_angle	Extratrees	0.891	0.188

Table 5. 1. Inverse problem 2 feature prediction

2개의 특성 예측에서는 표 5를 보면 Train set과 Test set의 정확도 차이가 큰 것을 확인할 수 있다. 이는 과대 적합(Overfitting)의 대표적인 현상인데, 다양한 이유로 인해 생긴다. 하지만 이번 문제에서는 신경망 모델을 사용하지 않은 간단한 알고리즘의 모델에서도 과대 적합이 일어나는 것으로 보아, 데이터의 부족이 주된 사유라고 생각한다.

#### IV. 결론

본 논문에서는 KSTAR Tokamak에서 RF 가열 방식을 이용한 초고온 플라즈마 가열 효율을 예측하고 최적화하기 위해 다양한 기계 학습 모델을 적용하여 성능을 평가하였다. 0-D problem에서는 앙상블 기반 회귀 모델을 통해 RF 가열 효율을 성공적으로 예측하였고, 1-D problem에서는 LSTM과 Transformer 모델을 사용하여 플라즈마 위치에 따른 가열 프로파일 예측을 두 가지 방법으로 모델링하였다. 또한, Inverse problem을 통해 RF 가열 과정의 주요 변수들을 추정함으로써 플라즈마 가열 효율을 개선하기 위한 가능성을 확인하였다. 향후 연구에서는 더 다양한 모델과 데이터 증강을 통해 모델의 예측 정확도를 높이고, 핵융합 에너지 상용화를 위한 실질적인 기여를 목표로 한다.

#### 참고 문헌

[1] Chen, F. F. "Introduction to plasma physics and controlled fusion." Springer Science & Business Media, 2012

[2] Lamalle, P. U., et al. "Ion cyclotron resonant

frequency heating on the JET tokamak: a review." Fusion Science and Technology 53.2 (2008): 355-387

[3] 조성현, 박지호, 함성원. "원자력전산해석응용 보고서", 한양대학교 2024.

[4] Zhou, Z. H. "Ensemble methods: foundations and algorithms." CRC press, 2012

[5] Hochreiter, S., and Schmidhuber, J. "Long short-term memory." Neural Computation, 1997.

[6] Vaswani, A., et al. "Attention is all you need." Advances in neural information processing systems. 2017.

[7] Radford, A., et al. "Language models are unsupervised multitask learners." OpenAI, 2019.

[8] Breiman, L. "Random forests." Machine learning 45.1 (2001): 5-32.

# 재활의학적 통증 환자를 위한 AI기반 기록관리 솔루션

허찬호\*, 문경서\*, 박희은\*, 현우정\*, 사공운\*\*, 김민희\*\*\*

## AI-based record management solution for rehabilitation pain patients

Chanho Hur\*, Kyeongseo Moon\*, Hee-Eun Park\*, Woojeong Hyun\*, Woon Sagong\*\*, and Minhee Kim\*\*\*

### 요약

본 연구는 환자의 통증 기록을 체계적으로 관리하고, AI를 통해 맞춤형 통증 보고서를 생성하는 웹 솔루션을 개발하는 것을 목표로 한다. 현대 사회의 고령화와 만성 통증 환자의 증가, 그리고 국내 의료 환경에서 짧은 진료 시간으로 인해 발생하는 의사와 환자 간 소통 문제를 해결하기 위해 제안된 이 솔루션은 환자가 자신의 통증을 효과적으로 기록하고, AI 분석을 통해 의료진에게 보다 정확한 정보를 제공한다. 이를 통해 진료의 정확성과 효율성을 높이고, 의료진과 환자 간의 소통을 개선하는 데 기여할 수 있다.

### Abstract

This study aims to develop a web solution that systematically manages patients' pain records and generates customized pain reports using AI. In response to the challenges of an aging society, the increase in chronic pain patients, and communication issues between doctors and patients caused by short consultation times in the Korean medical environment, this solution allows patients to effectively record their pain. Through AI analysis, more accurate information is provided to medical professionals, enhancing the accuracy and efficiency of diagnoses and improving communication between healthcare providers and patients.

### Key words

AI-based solution, Pain management, Rehabilitation, Personalized reports, Medical consultation efficiency, Patient documentation

## I. 서론

현대 사회에서 고령화와 더불어 통신 기술의 급격한 발전은 의료 분야에 새로운 도전 과제를 제시하고 있다. 특히, 근골격계 통증을 겪는 만성 질환

환자가 급증하고 있으며, 이로 인해 의료진의 진료 부담은 가중되고 있다. 국내 병원에서 의사 1인당 진료 건수가 매우 많고, 2023년 7월 기준 서울대병원 진료과별 진료 시간[5]에 따르면 외래환자 평균 진료 시간이 5분 내외로 짧은 상황에서 환자의 증

\*영남대학교 컴퓨터공학과, chhur3376@gmail.com

\*\*영남대학교 컴퓨터공학과 교수, wsagong@yu.ac.kr

\*\*\*(주)신라시스템 대표이사, mhkim@sillasystem.com

상을 충분히 파악하고 진단을 내리는 것은 어려운 현실이다. 환자 역시 짧은 시간 동안 자신의 증상을 명확하게 전달하는 데 어려움을 겪고 있으며, 기억에 의존한 부정확한 표현으로 인해 의료진과의 소통이 원활하지 않은 경우가 빈번하다.

이러한 문제를 해결하기 위해 본 프로젝트는 AI 기반 재활의학적 통증 환자를 위한 기록 관리 솔루션 개발을 목표로 한다. 특히, 이 솔루션은 환자가 통증을 보다 정확하고 체계적으로 기록하고, 이를 바탕으로 AI가 맞춤형 분석과 보고서를 제공함으로써 환자와 의료진 간의 소통을 원활하게 하는 데 중점을 둔다. 이를 통해 환자는 자신의 통증 상태를 보다 구체적이고 명확하게 의료진에게 전달할 수 있으며, 의료진은 더 나은 진단과 치료 계획을 세울 수 있다.

근거기반 임상간호실무지침[6]의 근거 기반 임상간호 실무 지침에 따라, 통증의 강도, 유형, 치료 효과 및 부작용을 정기적으로 평가하고 기록하는 것이 중요하다. 본 시스템에서는 이러한 지침을 반영하여, 사용자들이 통증의 세부 정보를 쉽게 기록하고 AI 분석을 통해 자동으로 보고서를 생성할 수 있는 기능을 제공한다.

더불어, 모바일 환경에서 접근 가능한 이 시스템은 환자가 언제 어디서나 통증을 기록하고, 의료진과 실시간으로 데이터를 공유하여 진료의 정확성과 효율성을 높이는 것을 목표로 한다.

## II. 재활의학적 통증 환자를 위한 AI기반 기록관리 솔루션

### 2.1 시스템 구조 설계

본 연구에서는 재활의학적 통증 환자를 위한 체계적인 관리와 효율적인 의료진과의 소통을 위해 AI 기반 기록관리 솔루션을 설계하였다. 시스템은 사용자 인터페이스, 데이터 관리, 서비스 계층, 백엔드 서버, 외부 서비스와의 연동으로 구성되며, 각 구성 요소는 상호 유기적으로 통증 기록과 분석, 보고서 생성을 담당한다.

[그림 1]은 본 솔루션의 시스템 아키텍처를 시각적으로 표현한 것이다.

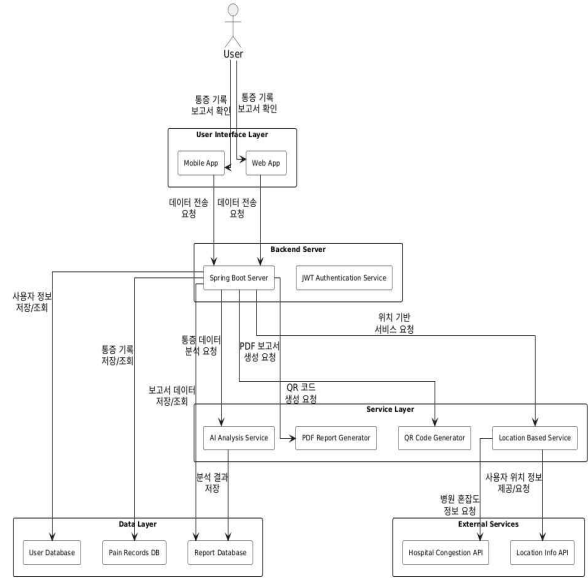


그림 1. 시스템 아키텍처 다이어그램  
Fig. 1. System Architecture Diagram

### 2.2 통증 기록 관리의 핵심 기능 흐름

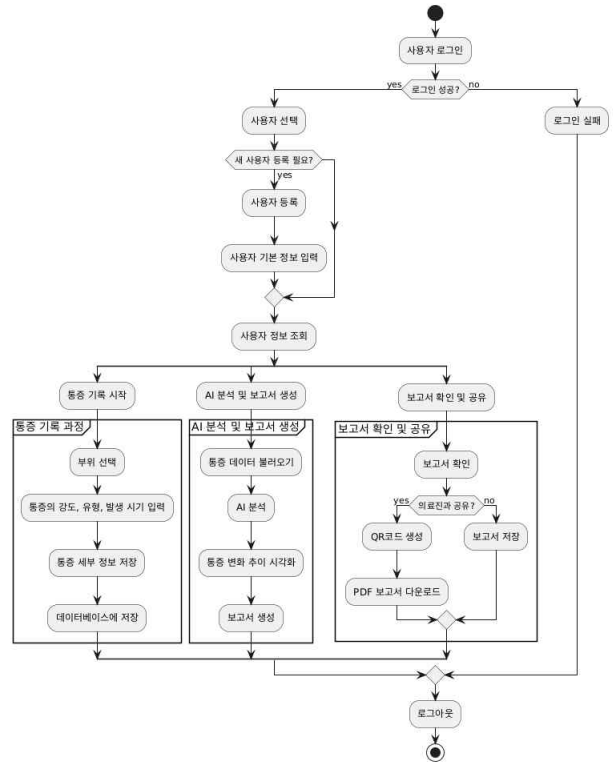


그림 2. 서비스 흐름도  
Fig. 2. Flow Chart

[그림 2]는 통증 기록 관리 서비스의 플로우 차트로, 사용자가 통증을 기록하고 AI 분석을 통해 결과를 확인하는 전체 과정을 나타낸다. 본 시스템은

사용자 인증 후 여러 기능을 선택할 수 있는 구조로 설계되었으며, 통증 기록, AI 분석, 보고서 확인 및 공유가 주요 기능이다.

### 2.3 통증 경과 시각화

사용자가 기록한 통증 데이터는 AI 분석을 통해 통증 강도 및 변화 추이를 시각화할 수 있으며, 이를 통해 사용자는 통증의 경과를 쉽게 이해할 수 있다. [그림 3]은 사용자의 통증 데이터를 기반으로 생성된 시각화 그래프의 예시이다. 전자기록관리 업무 및 기록정보서비스에서의 생성형 기술 활용 AI [4]의 연구에 따르면, 전자기록 관리 업무에서 생성형 AI 기술의 활용이 증가하고 있으며, 이를 통해 기록정보서비스의 효율성을 향상시킬 수 있다. 본 연구에서는 이러한 기록 관리의 자동화 기술을 통증 기록 시스템에 적용하여, 사용자가 손쉽게 통증 정보를 기록하고 관리할 수 있는 인터페이스를 제공하였다.

### III. 시스템 설계 화면

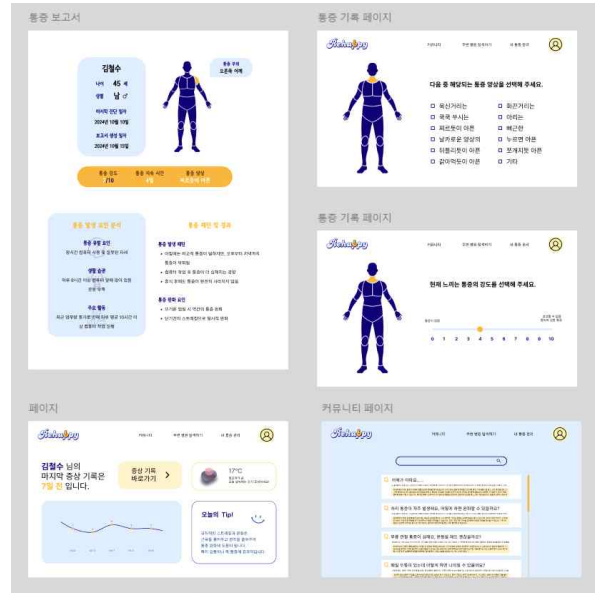


그림 4. 웹 서비스 화면  
Fig. 4. Web service view

### IV. 결론

본 연구에서 제안한 AI 기반 재활의학적 통증 기록 관리 솔루션은 환자와 의료진 간의 소통을 혁신적으로 개선하고, 진료의 효율성을 크게 향상시킬 것으로 기대된다.

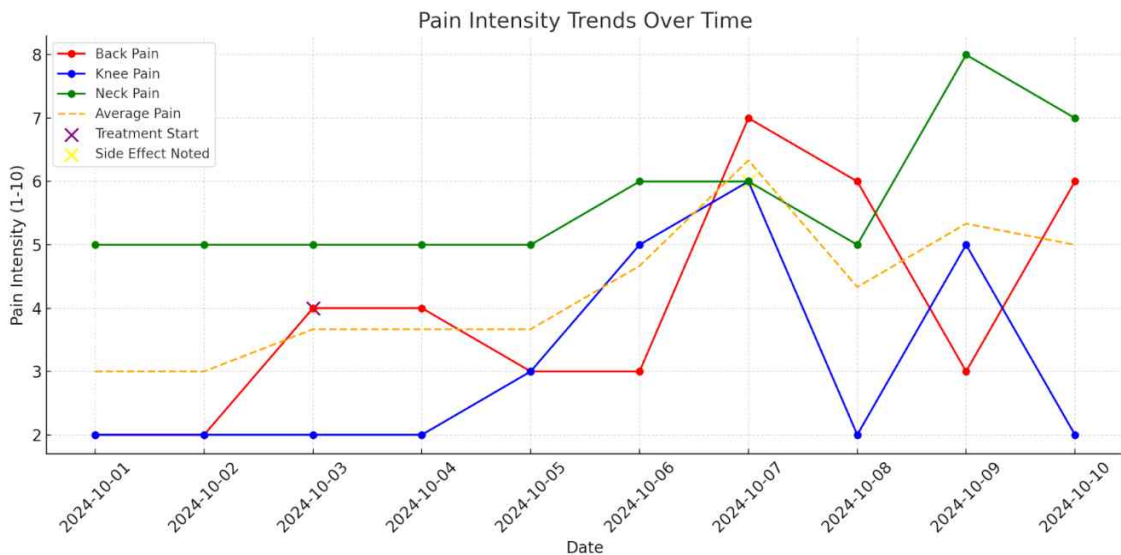


그림 3. 통증경과 기록 그래프  
Fig. 3. Pain history graph

환자는 어플리케이션을 통해 자신의 통증을 체계적으로 기록하고, AI 분석을 기반으로 한 맞춤형 보고서를 제공받아 진료 시간 동안 의료진과 효과적으로 소통할 수 있다.

이는 특히 통증 관리가 중요한 재활의학 분야에서 환자의 상태를 보다 정확하게 파악하고, 이에 맞는 치료 계획을 세우는 데 큰 도움을 줄 것이다.

또한, 이 솔루션은 병원 혼잡도를 실시간으로 조회할 수 있는 기능을 제공하여 응급 상황에서 신속한 치료를 받을 수 있도록 지원하며, 이를 통해 환자의 삶의 질이 향상될 것으로 예상된다.

본 연구에서 제시한 통증 기록 관리 어플리케이션은 환자와 의료진 간의 정보 교환을 원활하게 하고, 보다 정확한 진단과 효과적인 치료를 가능하게 하여 의료 서비스 품질을 높이는 데 중요한 역할을 할 것으로 기대된다.

## V. 감사의 글

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ‘2024년도 SW중심대학사업’ 지원으로 수행되었음.

## 참 고 문 헌

1. Bell, L., & Duffy, A. (2009). Pain assessment and management in surgical nursing: A literature review. *British Journal of Nursing*, 18(3), 153-156.
2. Cornellius, R., Herr, K. A., Gordon, D. B., & Kretzer, K. (2016). Evidence Based Practice Guideline: Acute Pain Management in Older Adults. 1-116. Iowa city, IA:Csomay Center for Gerontological Excellence, The University of Iowa, College of Nursing.
3. 한국기록관리학회지 23(4), 179-200, 2023.11  
<https://doi.org/10.14404/JKSARM.2023.23.4.179>  
pISSN 1598-1487 eISSN 2671-7247, 강윤아 (Yoona Kang)1, (Hyo-Jung Oh) 오효정 2
4. 전자기록관리 업무 및 기록정보서비스에서의 생성형 기술 활용 AI

<https://accesson.kr/jksarm/assets/pdf/33856/journal-23-4-179.pdf>

5. 서울대병원 비뇨기과 '3분' 정형외과·안과 '4분'

[https://www.dailymedi.com/news/news\\_view.php?wr\\_id=903431](https://www.dailymedi.com/news/news_view.php?wr_id=903431)

6. 근거기반 임상간호실무지침 - 병원간호사회

# 머신러닝 기반 멀티 IoT 센서를 사용한 대학 캠퍼스의 재실 감지 및 스마트 에너지 절감 솔루션

김윤성\*<sup>1</sup>, 문채영\*<sup>2</sup>, 이상돈\*<sup>3</sup>, 조민준\*<sup>4</sup>, 신연순\*\*

## Machine Learning Based Multi-IoT Sensor Solution for Occupancy Detection and Smart Energy Saving in University Campuses

Yoonseong Kim\*<sup>1</sup>, Chaeyoung Moon\*<sup>2</sup>, Sangdon Lee\*<sup>3</sup>, Minjoon Cho\*<sup>4</sup> and Younsoon Shin\*\*

### 요 약

본 연구는 Mobius 플랫폼에 연결된 IoT 센서 데이터를 활용하여 머신러닝 모델로 강의실의 재실 여부를 판단하고 이를 통해 전력 사용 장치를 제어함으로써 에너지 절감을 극대화하는 스마트 전력 관리 시스템을 제안한다. 센서 값만 의존했던 기존 연구와 달리, 본 연구는 각 강의실의 강의 시간과 대관 현황 등을 종합적으로 고려한다. 또한, 수집된 센서 데이터를 가우시안 필터(Gaussian Filter)로 전처리하여 기존 연구 방법보다 예측 정확도를 평균 9.4% 향상시켰다. 본 연구는 대학 캠퍼스와 같은 고전력 소비 환경에서 멀티 IoT 센서와 AI 기반 에너지 관리 시스템의 가능성과 지속 가능성을 효과적으로 제시한다.

### Abstract

This study proposes a smart power management system that maximizes energy savings by using Machine Learning to determine occupancy in lecture rooms through IoT sensors connected to the Mobius platform. Unlike previous studies that relied solely on sensor data, this research considers various factors, such as lecture schedules and room reservations, for comprehensive analysis. Additionally, data preprocessing using a Gaussian Filter resulted in an improved average prediction accuracy of 9.4%. This study effectively demonstrates the potential and sustainability of multi-IoT sensor and AI-based energy management systems in high-energy-consumption environments, such as university campuses.

### Key words

internet of things, machine learning, occupancy detection, sustainable management, green campus

---

\*1 동국대학교 AI융합대학 AI소프트웨어융합학부 컴퓨터공학전공, kimys\_21261@dgu.ac.kr

\*2 동국대학교 AI융합대학 AI소프트웨어융합학부 컴퓨터공학전공, 2021112037@dgu.ac.kr

\*3 동국대학교 AI융합대학 AI소프트웨어융합학부 컴퓨터공학 · 산업시스템공학전공, 2019112491@dgu.ac.kr

\*4 동국대학교 AI융합대학 AI소프트웨어융합학부 컴퓨터공학전공, wayou0822@dgu.ac.kr

\*\* 동국대학교 AI융합대학 AI소프트웨어융합학부(교신처), ysshin@dongguk.edu

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음 (2023-0-00049)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2024-2020-0-01789)

## 1. 서 론

### 1.1 배경

전 세계적으로 전력 수요가 급증하면서 에너지 절약의 중요성이 커지고 있다. 대학교는 상시 운영되는 연구 시설과 저렴한 전기료로 인해 전력 소비가 가장 높은 건물로 확인되었다 [1]. 한국전력공사의 조사에 따르면, 대학 내에서 사람이 없는 강의실에 불필요하게 점등되는 형광등과 전기기구는 시간당 약 320kWh를 소비하며, 매달 약 1,370만원의 비용 낭비를 초래한다. 여러 대학이 이 문제를 해결하기 위해 정부의 ‘그린캠퍼스 사업’ 등으로 전력 절감 프로젝트를 진행했으나, 대부분 일회성에 그쳐 절감 효과가 지속되지 않는 한계가 있다.

이러한 배경에서, OneM2M 표준을 따르는 Mobius 플랫폼 [2]을 도입함으로써 대학 내 전력 절감 시스템을 구축할 수 있다. Mobius 플랫폼을 통해 강의실의 재실 여부를 실시간으로 파악하고, 사람이 없는 상황에는 전기 시설들을 자동 제어함으로써 불필요한 전력 낭비를 줄일 수 있다. 본 연구는 이를 기반으로 대학 내 전력 절감 시스템 구축에 필요한 고려 사항과 지속 가능한 관리 방안을 제시하고자 한다.

따라서 본 연구에서는 강의실 사용 시간과 멀티 IoT 센서 데이터를 활용하여 AI 기반의 재실 판단 기법을 제안함으로써, 사람이 없는 강의실에서 불필요하게 낭비되는 전력을 줄이고, 에너지 절감 효과를 극대화하고자 한다.

### 1.2 관련 연구

전력 절감 시스템 구축에 있어 가장 중요한 것은 재실 여부를 판단하는 것이며, 재실 여부 판단과 관련된 다양한 연구가 있다. [3]은 카메라와 온습도 센서를 기반으로 머신러닝 알고리즘에 적용해 재실 여부를 판단하는 알고리즘을 제시했지만, 카메라의 제한된 FoV(Field of View), 온습도 센서의 불확실성으로 인해 80%대의 낮은 정확도를 보인다. [4]는 열화상 카메라를 이용해 인원 수를 확인하는 방식을

제시했다. 열화상 카메라를 이용하기 때문에 정확하지만, 초기 비용이 많이 들고 사각지대가 존재한다. 특히 카메라는 개인정보 및 초상권 문제로 인해 캠퍼스의 전력 절감을 위한 장비로 적절하지 않다.

본 연구는 기존 연구들과는 다르게, 저비용의 멀티 IoT 센서를 활용하고, 데이터 전처리 과정을 통해 재실 판단 정확도를 향상시키고자 한다.

## II. 재실 여부 판단을 위한 데이터 확보 및 인공지능 학습

### 2.1 재실 판단 로직 및 방법론

본 연구는 그림 1에서와 같이, IoT 디바이스에서 수집된 온습도, 전력 소모량 등의 IoT 센서 데이터를 게이트웨이를 통해 Mobius 플랫폼에 적재한다.

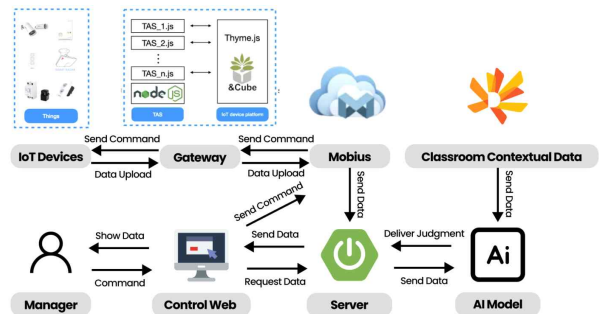


그림 1. 에너지 절감 솔루션 아키텍처  
Fig 1. Energy Saving Solution Architecture

Mobius는 데이터를 분석해 전력 제어 명령을 IoT 디바이스로 전송함으로써 실시간 에너지 관리가 가능하게 한다. 관리자는 웹 인터페이스(Control Web)를 통해 데이터를 모니터링하고 제어 명령을 인가함으로써 에너지 절감 시스템을 운영한다.

기존 연구와 달리, 본 연구에서 AI 모델은 IoT 센서 데이터뿐 아니라 강의 시간표와 대관 정보와 같은 상황적 데이터를 종합하여, 재실 여부를 판단한다. 예를 들어, 강의 시간표를 통해 수업 중에는 자동으로 재실로 판단하고, 대관 정보와 캠퍼스 일정(방학, 시험 기간, 공휴일 등)을 반영해 사용 가능성이 낮은 시간대는 비재실로 간주한다. 또한, 예약된 회의나 세미나 시간에는 재실로 판단하여 불필



요한 에너지 낭비를 줄인다. 이와 같이 강의실의 상황적 정보와 IoT 센서 데이터를 결합하여, 사람이 없는 상황에서의 불필요한 에너지 낭비를 효과적으로 차단함으로써 에너지 관리의 효율을 한층 더 높이고자 한다. 본 연구에서 제안하는 재실 여부 판단 로직은 그림 2에 나타나 있다.

```

Algorithm 1: Occupancy Detection based on IoT and Classroom Contextual Data


---


Input: IoT Sensor Data, Classroom Contextual Data
Output: Occupancy Status
Receive IoT sensor data and classroom contextual data;
if Classroom Contextual Data indicates a scheduled class then
  if Current time is within class hours or within 30 minutes before/after then
    return Occupancy status as "Occupied";
  end
else
  if Classroom Contextual Data indicates a reserved room then
    return Occupancy status as "Occupied";
  end
else
  if Radar sensor value exceeds threshold for 10 continuous minutes then
    Use all IoT Sensor Data in AI model to predict occupancy status;
  end
else
    return Occupancy status as "Unoccupied";
  end
end
end

```

그림 2. 재실 여부 판단 로직 알고리즘  
Fig 2. Algorithm for Determining Occupancy

수집된 데이터를 기반으로 최적의 재실 예측 모델을 도출하기 위해 scikit-learn [5]의 Decision Tree, Regression, Ensemble Model 등 총 9가지 알고리즘을 적용하여 각 모델의 성능을 비교하였다. 9가지 모델의 특성과 예측 정확도를 평가하여 가장 우수한 성능을 가진 모델을 선정하고, 이를 재실 판단 모델로 도입함으로써 최적화된 에너지 절감 효과를 달성하고자 한다.

## 2.2 데이터 수집 및 전처리

먼저 데이터 수집을 위해 동국대학교 서울캠퍼스 강의실에서 실험을 진행했다. 해당 강의실에서 2일에 걸쳐 상주하고 재실 유무를 기록하여 label 데이터를 확보했다. 여러 종류의 IoT 센서 데이터를 기반으로 재실 여부를 판단할 수 있도록 각 모델의 학습을 진행했다. 사용된 IoT 센서는 스마트 전등 스위치, 스마트 콘센트, 레이더, 온습도 센서, 스마

트 에너지미터가 있다. 각 센서에 대한 역할과 설정은 [6-8]을 참고했다. 실험은 다양한 시간대 뿐만 아니라 수업 중, 쉬는 시간, 야간 등 강의실 사용 패턴을 반영한 여러 상황을 포함하여 수집하였다. 수집된 실험 데이터 세트의 구성과 특성값은 각각 표 1, 표 2와 같다.

표 1. 실험 데이터 세트  
Table 1. Experimental data sets

데이터개수	train	test
119,587	83,711	35,876

표 2. IoT 디바이스와 특성값  
Table 2. IoT device and feature values

디바이스	데이터	내용
스마트 전등 스위치	ls1_onoff	1번 스위치 on/off
	ls2_onoff	2번 스위치 on/off
스마트 콘센트	co_power	전력(kW)
	co_current	전류(A)
	co_voltage	전압(V)
레이더	radar	움직임 정도
온습도 센서	th_temperature	온도
	th_humidity	습도
스마트 에너지미터	em1_voltage	에너지미터1 전압
	em2_voltage	에너지미터2 전압

본 연구에서는 멀티 IoT 센서 데이터를 이용해 재실 여부를 판단한다. IoT 디바이스 데이터를 활용하여 결측치를 선형보간법으로 보완하고, 이상치 제거 및 정규화를 통해 데이터를 전처리하였다. 이를 통해 여러 센서 데이터 간 일관성을 확보하고, 재실 판단 알고리즘의 정확성을 높이고자 하였다. 특히, 레이더는 재실 여부 판단의 핵심 센서이지만, 그림 4에서 보이듯이 움직임이 포착된 순간에는 큰 값을 보이다가 이후에는 사람이 있어도 낮은 값을 나타낸다. 'is\_person'은 사람의 재실 여부로, 재실 인원이 있다면 1000, 없다면 0으로 표시된다.

따라서 가우시안 필터(Gaussian Filter)를 사용하여 재실 여부에 따른 레이더 센서 값의 패턴을 강화하고, 이를 모델의 학습 데이터로 사용하여 인공지능 모델의 성능을 높이고자 했다. 그림 5는 가우시안 필터로 전처리한 결과이다.

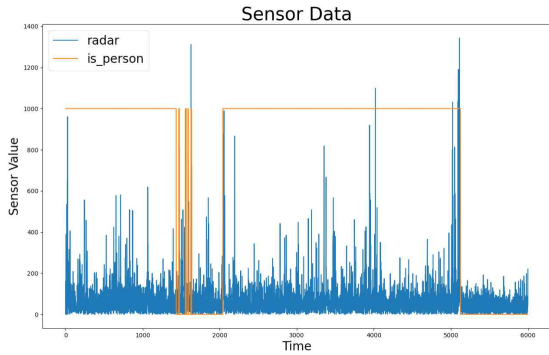


그림 4. 레이더 데이터의 전처리 필요성  
Fig. 4. Need for preprocessing radar data

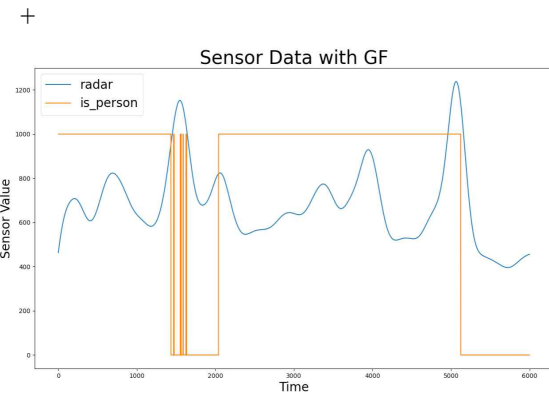


그림 5. 가우시안 필터를 사용한 데이터 전처리  
Fig. 5. Data preprocessing using Gaussian Filter

전처리 결과 노이즈가 제거되고 세부 정보가 부드럽게 표시되는 것을 확인할 수 있었다. 이를 통해 재실 여부에 따라 레이더 센서 값의 패턴을 강화하고, 이는 학습 모델의 성능 향상에 도움을 줄 수 있다.

### 2.3 인공지능 학습 및 평가

본 논문에서는 데이터를 기반으로 강의실에 사람이 재실 중인지 예측하기 위해 다양한 인공지능 모델을 학습 및 검증하였다. 사람이 없는 경우 불필요한 전력을 차단하는 피드백 시스템을 구현하는 것을 목표로 하며, 각 모델은 재실 여부 판단의 정확도(accuracy)로 성능을 평가하였다. 여기서 정확도는 Confusion Matrix에서 도출된 지표로, 실제 재실 여부와 모델의 예측 결과가 일치한 비율을 의미한다.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

특히, 본 연구에서는 멀티 IoT 센서에서 수집된 데이터를 결합하여 재실 상태를 판단하고자 하였으며, 이를 통해 각 모델의 특성과 성능을 비교하여 최적의 알고리즘을 도출하였다. AdaBoost, Random Forest 등 다양한 AI 모델을 적용하여 예측 정확도를 평가하였다.

표 3. 데이터 전처리 전 후 머신러닝 모델 학습 결과 비교

Table 3. Compare machine learning model learning results before and after data preprocessing

Model	Before	After
Decision Tree	0.7261	0.8136
Random Forest	0.7037	0.8232
SVM	0.7177	0.8752
Logistic Regression (LR)	0.5769	0.6120
LR with L2	0.5769	0.6210
Bagging Classifier	0.6896	0.7621
AdaBoost Classifier	0.7976	0.9051
XGBoost Classifier	0.8041	0.9122
GradientBoosting Classifier	0.7793	0.8935

결과를 보면, 데이터 전처리 후 전체적으로 정확도가 높아진 것을 확인할 수 있었으며, 전처리 후 XGBoost Classifier 모델에서 91.22%로 가장 좋은 예측 정확도를 보이는 것을 확인했다. XGBoost Classifier는 어셈블 모델로서, 틀린 예측에 대한 가중치를 조절하며 학습을 진행한다. 이에 따라 더 정교한 판단과 가장 높은 정확도를 보였으며, 결과적으로 성능이 가장 우수한 XGBoost Classifier 모델을 최종 재실 판단에 활용하였다.

## III. 결 론

본 논문에서는 강의실 재실 여부 판단을 위한 머신러닝 기반 멀티 IoT 센서 메커니즘을 제안했다. 특정 강의실의 강의 시간표, 대관 예약 정보 등 정보를 종합적으로 활용하고, Mobius와 연동된 센서 데이터로 학습한 인공지능을 통해 재실 여부를 판단함으로써 그린 캠퍼스 구축을 위한 전력 세이빙

솔루션을 제안했다. 카메라를 이용한 기존 연구와는 달리 레이더, 콘센트 등 센서 측정값을 사용함으로써 개인정보 문제도 고려하였다. 또한 가우시안 필터를 사용해 민감한 레이더 값을 전처리함으로써 기존 논문에서 제시한 방법에 비해 모델의 성능이 평균 9.4% 높아진 모습을 보였으며, 그 중 XGBoost Classifier가 91.22%의 정확도를 보였다. 이는 다양한 종류의 센서의 특성값을 통해 재실 여부를 정확하게 파악하고, Mobius 플랫폼을 이용한 IoT 제어를 통해 전력 낭비 문제를 해결할 수 있다는 가능성을 제시한다. 추후 본 논문의 결과를 바탕으로 실제 환경에 적용하여 에너지 절감 효과에 대해 검증할 계획이다.

## 참 고 문 헌

- [1] 서울특별시 에너지정보, 에너지 다소비 건물 현황 (서울특별시, 2022).
- [2] J. Kim, S.-C. Choi, J. Yun, and J.-W. Lee, "Towards the oneM2M standards for building IoT ecosystem: Analysis, implementation and lessons," *Peer-to-Peer Networking and Applications*, vol. 11, pp. 139-151, Sep. 2016.
- [3] Zemouri, Sofiane, Yinannis Gkoufas, and John Murphy. "A machine learning approach to indoor occupancy detection using non-intrusive environmental sensor data." *Proceedings of the 3<sup>rd</sup> international conference on big data and internet of things*. 2019.
- [4] Acquaaah, Yaa, et al. "Occupancy detection for smart HVAC efficiency in building energy: a deep learning neural network framework using thermal imagery." *2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. IEEE, 2020.
- [5] Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." *the Journal of machine Learning research* 12 (2011): 2825-2830.
- [6] Koichi Kuzume, Tomonori Tabusa, and Haruka Sawa, "Electric Power Saving Awareness System at School Using ICT," *MATEC Web of Conferences*, vol. 55, pp. 03002, April 2016.
- [7] Mohd Wafi Nasrudin, Nur Asyikin Nordin, Izaidly Ismail, Mohd Ilman Jais, Amir Nazren Abdul Rahim, and Wan Azani Mustafa, "Smart Classroom for Electricity-Saving with Integrated IoT System," *Journal of Physics: Conference Series*, vol. 2107, pp. 012019, October 2021.
- [8] Imran Ahmad, Dr. Waseem Abassi, and M. Bilal Iqbal, "Energy Conservation Smart Classroom System using IOT," *International Journal of Advancements in Mathematics*, vol. 1, no. 1, pp. 32-46, 2021.

# 딥러닝 기반 객체 인식을 위한 증강기법 비교 분석

오성우\*, 김경서\*, 송우현\*, 이창근\*, 김한솔\*\*

## Comparison Analysis of Data Augmentation Techniques for Deep Learning-based Object Detection

Sungwoo Oh\*, Gyeong Seo Kim\*, Woo Hyun Song\*, Changgeun Lee\*, Han Sol Kim\*\*

### 요 약

본 논문에서는 4가지 증강기법에 따른 성능향상과 각 증강기법의 성능을 분석한다. YOLO11m(You Only Look Once) 모델과 현실 유해조수 데이터셋에 photometric과 geometric transformation 기반의 증강기법을 적용하여 다양한 조합을 실험하고 측정된 성능향상을 비교한다. 실험을 통해 증강기법 중 가장 큰 향상도를 가지는 증강기법과 모든 증강기법을 적용하는 것이 항상 좋은 결과를 보여주지 않는다는 것을 알아본다. 최종적으로는 Flip과 brightness 증강기법을 조합하여 YOLO11m 모델의 정확도를 최대 10% 향상시킬 수 있다.

### Abstract

In this paper, we present individual performance based impacts of four distinct augmentation techniques. We applied photometric and geometric transformation-based augmentation methods to the YOLO11m(You Only Look Once) model and real-world varmint dataset, experimenting with various combinations and comparing performance improvements. Our experiments found that applying all augmentation yields slight improvement not as a common expectation, and combining some methods can lead to better results than applying all. Ultimately, we improved the accuracy of the YOLO11m model by up to 10% through a combination of flip and brightness augmentations.

### Key words

Data augmentation, photometric transform, geometric transform, object detection, yolol1

### 1. 서 론

Object detection은 이미지나 비디오에서 모델이 학습한 객체의 종류와 위치를 추정하는 기술이다. 이러한 점으로 농업, 산업 그리고 군사적인 분야까

지 여러 분야에서 object detection을 활용하고 있다. 그중 최근에 발표된 Ultralytics사의 YOLO11는 기존 모델에 비해 가벼운 크기와 실시간성에 장점이 있는 모델이다[1].

여러 object detection의 SOTA(State Of The Art)

+ 교신저자

\* 단국대학교, (ohsungwoo, 32190322, 32192257, 32193601, hansol)@dankook.ac.kr

※ 이 연구는 2023년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (과제번호 : RS-2023-00251621).

모델들이 경쟁적으로 출시되고 있지만 모델의 학습을 위한 데이터셋을 구성하는 방법은 여전히 사람이 크게 관여하는 등 달라지지 않았다. 데이터셋은 인식모델의 성능에 큰 영향을 끼치기에 과적합과 과소 적합이 되지 않도록 모델에 크기에 알맞게 크고 다양한 데이터셋을 구성하는 것이 이상적이다. 하지만 현실적인 제약으로 인해 이런 데이터셋을 구현하기는 어렵다.

이러한 문제를 완화하기 위해 이미지에 photometric과 geometric transformation을 가해 새로운 이미지를 만들되 추가적인 라벨링은 필요로 하지 않는 증강기법들이 사용되고 있다. Photometric과 geometric transformation은 이미지를 자르는 crop 처럼 이미지에 변화를 가해 object detection 모델이 더욱 다양한 형태의 객체를 학습할 수 있도록 한다 [2]. 앞에서의 논문처럼 증강기법이 개별적으로는 소개되고 있지만 photometric과 geometric transformation을 나눠서 실질적인 성능을 비교하는 연구는 미비한 편이다[3].

앞선 분석에서 착안하여, 본 논문에서는 데이터셋에 photometric과 geometric transformation[4]을 적용하여 데이터셋을 증강하고 증강기법끼리 조합하여 가장 mAP50-95(mean Average Precision50-95)를 향상시키는 기법을 제안한다. mAP50-95는 인식모델이 validation 값을 predict 할 때 label과 겹치는 정도인 IOU(Intersection Over Union)를 50%부터 95%까지 평균을 내어 구한 지표이다. 50% 이상이면 정답이라고 보는 mAP50보다 엄격한 기준이다.

## II. 주요 기법 소개

이러한 4가지 종류의 증강 방식을 선택한 이유는 시연데이터와 학습데이터 간의 차이를 고려하였을 때 4가지 증강 방식을 통해 밝기, 그림자, 동물의 다양한 위치를 추가로 증강하는 것이 시연데이터에서 더 좋은 결과를 얻기에 적합하다고 예상했기 때문이다.

이러한 점들은 지금의 유해조수 데이터셋뿐만 아니라 실제 환경에서 데이터를 수집할때 공통적으로 적용될 수 있다. 예를 들어 특정 물건을 인식하

려 한다면 물건이 다양한 밝기, 구도, 그리고 위치를 가질 것이고, 그 물건을 촬영하는 카메라 또한 다른 초점, roll, yaw, 그리고 pitch를 가질 것이다. 데이터셋 수집할 때 이러한 점을 모두 고려하면 좋겠지만 현실적인 제약으로 불가하다. 본 논문이 제작하고 활용한 유해조수 데이터셋은 사람이 제어할 수 없는 동물의 다양한 움직임과 위치를 인식하기 위해 여러 방향으로 augmentation을 테스트해 볼 수 있었다.



그림 1. Photometric transformation에 대한 예시 사진  
왼쪽은 원본, 가운데는 밝기를 -15%, 그리고 오른쪽은 흑백으로 조정했다.

Fig. 1. Examples of photometric transformation:  
The left is the original image, the middle is adjusted to -15% brightness, and the right is converted to grayscale.

Photometric transformation은 객체의 색상과 관련된 변형들을 의미한다. 이러한 대표적인 transformation에는 이미지를 흑백으로 바꾸는 grayscale, 밝기를 조정하는 brightness, 그리고 선명도를 조정하는 blur 등이 있다. 그림 1처럼 원본 이미지에 대해 조정이 된 이미지들이 기본 데이터셋에 추가된다. 해당 데이터셋에서는 brightness는 밝기를  $\pm 15\%$ 로 조정했고, grayscale은 전체 데이터셋을 흑백으로 조정하고 기존데이터 세트에 추가하였다.

Geometric transformation은 객체의 위치와 모습에 관련된 변형들을 의미한다. 이러한 대표적인 transformation에는 이미지를 대칭으로 뒤집는 flip, 이미지를 x축을 기준으로 회전(roll)시키는 rotation, 그리고 이미지를 y축(yaw)과 z축(pitch)을 기준으로 동시에 회전시키는 shear 등이 있다. 그림 2처럼 원본 이미지에 대해 조정이 된 이미지들이 기본 데이터셋에 추가된다. Flip는 horizontal, vertical 모두 뒤집을 수 있게 조정했고, rotation은  $\pm 15\%$ 로 이미지를 회전시켰다.



그림 2. Geometric transformation에 대한 예시 사진  
왼쪽은 원본, 가운데는 vertical flip, 그리고 오른쪽은  
+15도 회전으로 조정했다.

Fig. 2. Examples of geometric transformation:  
The left is the original image, the middle is vertically  
flipped, and the right is rotated 15 degrees clockwise.

### III. 실험 결과



그림 3. 유해조수 데이터세트의 일부 사진.  
왼쪽은 늑대 사진, 오른쪽은 고라니 사진

Fig. 3. Pictures from varmint dataset.  
Left side is picture of wolf, right side is waterdeer

유해조수 데이터세트는 현실에서의 멧돼지, 늑대, 고라니 각각 150장으로 8(train) : 2(validation)로 나뉘어 진행하였다. 실험은 데이터세트를 모두 동일한 A100 GPU 환경에 놓고 2가지로 진행한다. 1) 각각의 transformation을 적용하여 기본 데이터세트와 비교하여 성능향상 폭을 확인하는 것과 2) 그중 성능이 좋았던 기법을 동시에 적용해 보고, 4가지 기법을 모두 적용해 보는 것으로 진행하였다. 각각을 mAP50-95 지표로 평가하여 object detection 성능향상 정도를 기록하였다.

Photometric에서 널리 사용되는 brightness, greyscale transformation과 geometric에서 rotation, flip transformation을 그림 3처럼 실제 유해조수를 촬영하여 구성된 데이터세트에 적용하여 실험을 진행한다. 증강의 양은 기본 데이터세트를 제외하고는 2배 수로 고정하였다.

모델은 YOLO11m pretrain된 모델을 사용하였고 환경 파라미터로 VRAM 사용량을 90%, epoch 40으로 설정한 후 데이터세트만 변인으로 진행하였다. 각 데이터세트의 학습은 약 30분씩 소모되었다.

표 1. 증강기법에 따른 mAP50-95 지표

Table 1. Table of mAP50-95 and augmentation techniques

증강기법	mAP50-95
없음(기본)	0.717
Flip + Brightness	0.817
Flip	0.813
Brightness	0.810
Rotation	0.796
All(4가지)	0.789
Greyscale	0.774

실험을 통해 3가지를 확인할 수 있었다. 가장 성능이 좋을 것으로 예상한 1) 4가지 증강기법을 모두 적용한 데이터세트가 항상 폭이 7%로 뒤에서 2등인 것을 확인했다. 2) Flip과 brightness 2가지를 적용한 조합 증강기법이 10% 성능향상으로 가장 좋은 성능을 가지는 것을 확인할 수 있었다. 3) 마지막으로 어떤 증강기법을 사용하더라도 기본에 비해 성능이 향상되는 것을 확인하였다. 각 증강기법에 따른 mAP50-95는 표 1에 정리되었다.

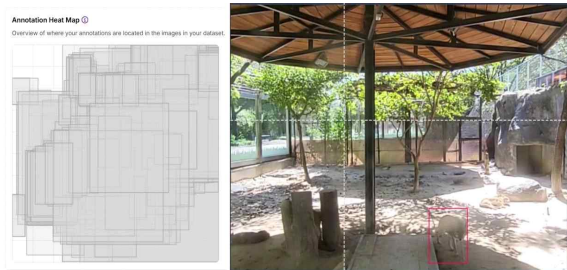


그림 4. 왼쪽은 Flip 증강으로 다양해진 바운딩박스의 모습, 오른쪽은 validation에서 그림자에 가려진 유해조수의 그림

Fig. 4. Proof of increased mAP metrics.  
Left side is boundingbox heatmap, right side is shadowed varmint from validation

이러한 결과가 나오는 원인은 그림 4처럼 1) flip 증강으로 인해 바운딩 박스가 이미지의 여러 위치에 분포하고 이러한 점이 NMS(Non Maximum Suppression)에 영향을 주어 다양한 위치의 유해조수를 잘 파악하게 하는 것과 2) validation에서 유해조수가 나무 그림자에 가려져 잘 안 보이는 부분들이 brightness 증강으로 train 데이터세트에서도 미리 학습할 수 있었기에 모델이 이를 학습하여 더 높은 mAP50-95를 달성한 것으로 예상된다.

#### iv. 결 론

논문에서는 4가지 증강기법에 따른 성능향상과 각 증강기법의 성능을 분석한다. YOLO11m 모델과 현실 유해조수 데이터세트에 photometric과 geometric transformation 기반의 증강기법을 적용하여 다양한 조합을 실험하고 측정된 성능향상을 비교했었다. 실험을 통해 증강기법 중 가장 큰 향상도를 가지는 증강기법과 모든 증강기법을 적용하는 것이 항상 좋은 결과를 보여주지 않는다는 것을 알아보았다. 최종적으로는 Flip과 brightness 증강기법이 가장 좋은 향상도로 YOLO11m 모델의 정확도를 최대 10% 향상시킬 수 있었다. 추후 연구에선 스스로 정확도 향상에 필요한 photometric, geometric transformation 을 결정하여 적용하는 알고리즘의 연구를 진행할 예정이다.

#### 참 고 문 헌

- [1] R. Khanam and M. Hussain, "YOLOv11: an overview of the key architectural enhancements," arXiv preprint arXiv:2410.17725, 2024.
- [2] L. Taylor and G. Nitscke, "improving deep learning with generic data augmentation" Institute of Electrical and Electronics Engineers, vol. 10, pp. 1542-1547, 2018
- [3] S. Sivamani, S. I. Chon, and J. H. Park, "selective augmentation to create a balance dataset based on photometric variation," Institute of Electrical and Electronics Engineers, access.2024.3437678, 2024.
- [4] P. Kaur, B. S. Khehra, and B. S. Mavi, "data augmentation for object detection: A review," Institute of Electrical and Electronics Engineers, vol. 65, pp. 537-543, 2021.





# MMPose를 활용한 신체 균형 및 자세 판단 알고리즘 설계

김대민\*, 시종욱\*\*, 김성영\*\*\*

## Design of Algorithm for Body Balance and Pose Assessment Using MMPose

Daemin Kim\*, Jongwook Si\*\*, and Sungyoung Kim\*\*\*

### 요 약

현대 사회에서 잘못된 자세와 신체 불균형의 문제는 점점 심화되고 있으며, 이는 걸음걸이에 영향을 미쳐 무릎과 허리에 부담을 주고, 장기적으로는 만성 통증과 질환으로 발전할 수 있다. 본 논문은 이러한 문제를 해결하기 위해 MMPose 골격 추출 모델을 활용하여 자세와 균형을 정확하게 판별하는 알고리즘을 제안한다. 제안 방법은 골격 좌표 사이의 각도를 계산함으로써 걸음걸이와 자세의 균형을 평가할 수 있음을 보인다.

### Abstract

In modern society, the issues of poor posture and body imbalance are becoming increasingly severe. These problems can affect gait, placing strain on the knees and lower back, and potentially leading to chronic pain and conditions over time. This paper proposes an algorithm to accurately assess pose and balance using the MMPose skeletal extraction model as a solution to these issues. The proposed method suggests that by calculating the angles between skeletal coordinates, it is possible to evaluate the balance of gait and pose.

### Key words

Algorithm, MMPose, Body Balance, Pose Assessment, Skeletal Extraction

## 1. 서 론

현대 사회에서는 오랜 시간 앉아서 생활하거나 모바일 기기를 장시간 사용하는 경우가 많아, 잘못된 자세로 인해 신체 불균형이 발생하는 경우가 흔하다. 이러한 신체 불균형은 걸음걸이에도 영향을

미쳐 발목, 무릎, 허리 등 여러 부위에 무리를 줄 수 있으며, 장기적으로는 만성 통증이나 근골격계 질환으로 발전할 위험이 있다 [1]. 이러한 문제를 해결하기 위해 다양한 골격 추출 모델을 활용하여 균형 상태를 평가하거나 낙상을 감지하는 연구들이 활발히 진행되고 있다 [2-3].

\* 국립금오공과대학교 컴퓨터공학과

\*\* 국립금오공과대학교 컴퓨터·AI융합공학과

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학연협력 선도대학 육성사업 (LINC 3.0)의 연구결과입니다.

## II. 신체 균형 및 자세 판단 알고리즘 설계

본 논문에서는 골격 추출 모델 중 하나인 MMPose[4]를 활용하여 신체 균형을 감지하고 올바른 자세의 걸음걸이를 판별하는 알고리즘의 설계를 소개한다. 이 알고리즘은 카메라를 통해 사람의 움직임을 촬영한 후, 관절 포인트를 추출하고 각도를 계산하여 걸음걸이와 관련된 자세의 정확성을 평가한다. 이를 통해 사용자의 걸음걸이에서 나타나는 잘못된 자세를 조기에 식별하고 교정할 수 있다.

그림 1은 MMPose를 통해 추출한 골격 좌표들에 대하여 키포인트를 도식화한 내용이다. 제안하는 알고리즘에서는 MMPose를 활용하여 왼쪽 얼굴, 오른쪽 얼굴, 왼쪽 어깨, 오른쪽 어깨, 그리고 골반 선상의 두 점을 중점으로 설계한다. 이러한 키포인트를 기반으로 여섯 개의 각도를 계산하며, 이를 각각 짝을 이루는 각도 쌍으로 비교하여 신체의 균형 상태를 평가할 수 있다.

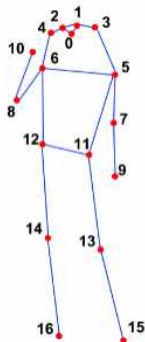


그림 1. MMPose의 골격 좌표 도식화

Fig. 1. Visualization of Skeletal Coordinates Using MMPose

첫째, 얼굴과 어깨 사이의 좌우 각도를 비교하여 신체 균형을 판단한다. 얼굴의 왼쪽과 왼쪽 어깨를 잇는 선과, 왼쪽 어깨와 오른쪽 어깨를 잇는 선 사이의 각도를 계산하고, 얼굴의 오른쪽과 오른쪽 어깨를 잇는 선, 오른쪽 어깨와 왼쪽 어깨를 잇는 선 사이의 각도를 동일하게 계산하여 비교한다. 둘째, 어깨와 골반 사이의 좌우 각도를 비교한다. 이는 왼쪽 어깨와 골반 사이의 각도와 오른쪽 어깨와 골반 사이의 각도를 계산하여 비교한다. 이 과정에서, 각도 쌍 중 하나라도 일정 각도 이상 차이가 발생하

면 해당 자세는 신체 균형이 맞지 않는 것으로 판별된다. 즉, 얼굴과 어깨 사이의 각도 또는 어깨와 골반 사이의 각도에서 일정 각도 이상의 차이가 나타나면 신체의 좌우 균형이 불균형한 것으로 간주한다.

걸음걸이의 경우, 골반 선상의 점, 무릎, 발의 위치를 분석하여 각도가 일정 각도 미만이면 걸음을 걸은 것으로 인식한다. 그러나 앞서 계산한 균형 상태에서 문제가 발견될 경우, 해당 걸음은 카운트되지 않으며 사용자가 균형 잡힌 자세로 걸을 수 있도록 해야한다.

## III. 결론

본 논문에서는 MMPose 골격 추출 기술을 활용하여 사용자가 신체 균형을 유지하면서 올바른 걸음걸이를 하고 있는지를 확인할 수 있는 알고리즘을 설계하였다. 이 알고리즘은 걷기 기능에만 국한되지 않고, 다양한 운동에 적용될 수 있도록 추가 개발이 필요하며, 이를 통해 더 많은 사용자가 활용할 수 있는 확장 가능성을 제시한다. 추후에는 이 기술을 가상 환경의 운동 콘텐츠와 연계하여 실제로 적용할 수 있는 방안을 고려해볼 예정이다.

## 참고 문헌

- [1] <https://www.k-health.com/news/articleView.html?idxno=66112>
- [2] S. Son, J. Si, and S. Kim, "Development of Metaverse Rehabilitation Content for Body Balance and Gait Analysis", Proceedings of KIIT Conference, pp. 414-415, Nov. 2023.
- [3] J. Si, S. Kim *et al.*, "Fall Detection using Skeletal Coordinate Vector and LSTM Model", The Journal of Korean Institute of Information Technology, Vol. 18, No. 12, pp. 19-29, Dec. 2020.
- [4] A. Sengupta, F. Jin, R. Zhang, and S. Cao, "mm-Pose: Real-time human skeletal posture estimation using mmWave radars and CNNs", IEEE Sensors Journal, Vol. 20, No. 17, pp. 10032-10044, May. 2020.

# 이중토큰을 이용한 효율적인 Wi-Fi 보안 프로토콜에 대한 정형검증 연구

강경아\*, 김지윤\*\*

## Study on Formal Verification for Efficient Wi-Fi Security Protocol Using Dual Tokens

KyeongA Kang\* and Jiyeon Kim\*\*

### 요 약

기존의 WPA2 방식은 4-way 핸드셰이크 방식을 활용하여 보안세션을 설정하였다. 이중토큰 기반 Wi-Fi 보안 프로토콜은 WPA2와 WPA3 등의 기존 방식들의 효율적 문제와 보안성 문제를 개선하기 위해 설계되었다. 이 방식은 이중토큰 기술을 활용하여 4-way 핸드셰이크 방식을 사용하지 않고 보안세션을 설정한다. 토큰을 이용하여 지속적인 인증 요청을 완화하고 보안세션 설정에 걸리는 시간을 감축하였다. 본 논문은 기존 Wi-Fi 보안세션 설정 프로토콜의 문제점을 보완하기 위해 제안된 이중토큰 기반 Wi-Fi 보안 프로토콜을 정형 검증 도구인 BAN-Logic을 통해 검증하고 그 결과를 분석한다.

### Abstract

The existing WPA2 method utilized a 4-way handshake to establish a secure session. The dual-token based Wi-Fi security protocol was designed to improve the efficiency and security issues of existing methods such as WPA2 and WPA3. This method uses dual-token technology to establish a secure session without using the 4-way handshake. It alleviates continuous authentication requests using tokens and reduces the time taken to set up a secure session. This paper will verify the proposed dual-token based Wi-Fi security protocol using a formal verification tool called BAN-Logic and analyze the results to address the issues of existing Wi-Fi secure session establishment protocols.

### Key words

Wi-Fi security protocol, BAN-Logic, Formal cerification

### 1. 서 론

Wi-Fi는 전파를 이용한 무선랜 기술로 전파를 통해 정보를 전송하기 때문에 도청공격, 변조공격을 방지하는 것이 중요하다. 초기에는 WEP가 사용되

었는데 이는 짧은 키 길이와 취약한 설계로 인해 통신패킷을 수집하면 공유키를 찾아낼 수 있었다. 이를 개선하기 위해 설계된 WPA2는 4-way 핸드셰이크 방식을 사용한다[1]. 이는 4번의 통신을 요구하는 복잡한 프로토콜이기 때문에 효율성 측면에서

\* 경상국립대학교 컴퓨터공학과, kka2418@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, jykim92@gnu.ac.kr(교신저자)

개선이 필요했으며 공유 비밀키가 노출될 경우 여전히 취약점이 존재한다. WPA2[2]의 취약점을 개선한 WPA3은 모듈러 승산 연산을 이용하여 보안적 측면을 개선하였으나 여전히 4-way 핸드셰이크 방식을 사용한다.

Wi-Fi 보안 프로토콜에서는 클라이언트 입장에서 접속하고자 하는 AP에 대한 신뢰성을 확인하기 어렵다는 문제점이 존재하며, 4-way 핸드셰이크 방식으로 인해 복잡하고 시간이 소요된다. 또한, 공개 Wi-Fi 등에서 비밀키가 노출되는 경우 통신 도청이 가능해지는 취약점이 존재한다.

AP의 인증과 빠른 보안 세션 설정 등을 위해 토큰 인증 방식과, 이중토큰을 사용한 것이 이중토큰 기반 Wi-Fi 보안 프로토콜이다[3]. 토큰인증 방식은 초기 인증을 성공한 사용자에게 토큰을 발급하여 인증상태를 유지하는 방식이며, 이중토큰을 이용한 난수화 토큰인증 기술은 공개토큰과 비밀토큰을 활용하여 인증유지 시에 비밀토큰을 이용하여 일회용 인증값을 계산하고 공개토큰과 현재시간을 함께 전달하여 인증된 상태를 증명한다. 이중토큰 기반 Wi-Fi 보안 프로토콜은 Wi-Fi 보안의 효율성을 높이고, 기존의 보안 취약점을 보완하며, AP 인증 문제를 해결하기 위해 설계되었다.

본 논문은 본론에서 이중토큰 기반 Wi-Fi 보안 프로토콜에 대한 정형 검증을 BAN-Logic [4]을 통해 진행하고, 결론에서 검증 결과를 분석하고 결론을 맺는다.

## II. 본론

### 2.1 이중토큰 기반 Wi-Fi 보안 프로토콜 절차

클라이언트와 AP가 세션을 설정하는 과정은 초기인증 및 토큰 발급 과정과 보안세션 설정 과정으로 구분된다. 그림 1은 초기인증 및 토큰 발급 과정으로 처음 연결을 시도하거나 인증 유효기간이 지난 경우 실행된다.

- 1-1) 클라이언트가 AP에게 무선랜 서비스를 요청함.
- 1-2) AP는 클라이언트에게 인증서  $Cert_{AP}$ 를 전달함.
- 1-3) 클라이언트는 AP의 인증서를 검증하고 난수  $r$

를 생성하고  $\{r, puk\}_{puk}$ 를 전달함.

- 1-4) AP는  $\{r, puk\}_{puk}$ 를 복호화하여  $r$ 을 획득하고 난수값  $c$ 를 클라이언트에게 전달함.
- 1-5) 클라이언트는  $\{PSK, c, r\}_r$ 을 전달함.
- 1-6) AP는  $\{PSK, c, r\}_r$ 을 복호화하고  $T_p$ 와  $T_s$ 를 발급하고 암호화하여 클라이언트에게 전달함.
- 1-7) 클라이언트는 복호화하여  $T_p$ ,  $T_s$ 를 획득하고 시스템에  $[AP, T_p, T_s]$ 를 저장함.

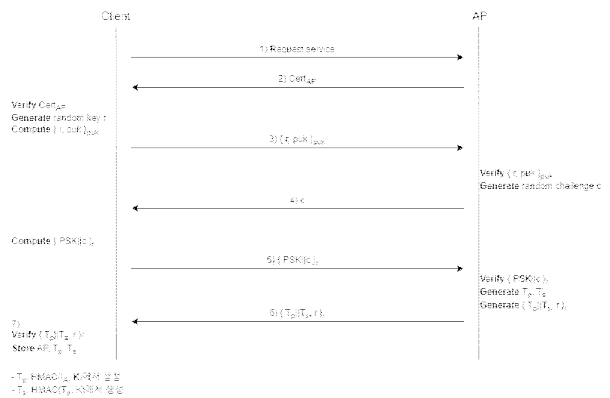


그림 1. 초기인증 및 토큰 발급 과정

[그림 2]는 인증과 토큰 발급 후에 보안세션을 설정하는 과정이다.

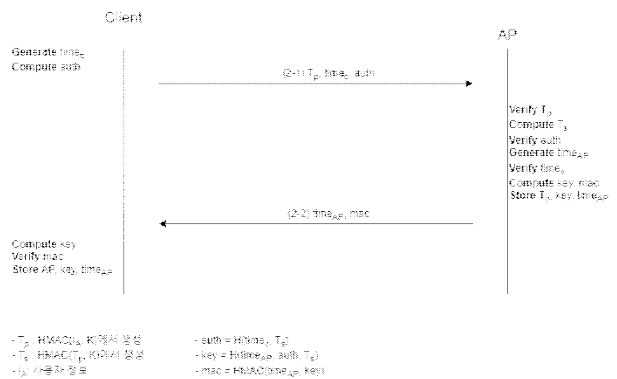


그림 2. 토큰을 이용한 보안세션 설정 과정

- 2-1) 클라이언트가 일회용 인증정보  $auth$ 를 계산하고 AP에게  $T_p$ ,  $time_c$ ,  $auth$ 를 전달함.
- 2-2) AP는  $key$ 를 계산하고  $time_{AP}$   $mac$ 를 클라이언트에게 전송한 후,  $[T_p, key, time_{AP}]$ 를 저장함.
- 2-3) 클라이언트는  $key$ 를 계산하고  $mac$ 을 검증한 후에  $[AP, key, time_{AP}]$ 를 저장하고 보안세션 설정이 완료됨.

## 2.2 BAN-Logic 기반 정형검증

그림 3은 초기 인증 및 토큰 발급 과정을 BAN-Logic을 통해 검증한 결과이다. (D8)에서 PSK에 대한 믿음을 도출하고, (D12)에서  $r$ 에 대한 믿음을 도출할 수 있다. 또한, (D13), (D14)에서  $T_p$ ,  $T_s$ 에 대한 믿음을 도출할 수 있다. AP의  $r$ 에 믿음은 Hypothesis인 (H1)이 존재해야 도출이 가능하다.

```

< Idelization >
(11) Client → AP : {r, puk}puk
(12) Client → AP : {PSK, c, r}r
(13) AP → Client : {Tp, Ts, r}r

< Assumption >
(A1) AP |≡ puk
(A2) AP |≡ Client |⇒ r
(A3) AP |≡ # (c)
(A4) Client |≡ r
(A5) Client |≡ # (r)
(A6) Client |≡ AP |⇒ Tp, Ts

< Goals >
(G1) AP |≡ Client |≡ r
(G2) AP |≡ r
(G3) AP |≡ Client |≡ PSK
(G4) Client |≡ AP |≡ r
(G5) Client |≡ AP |≡ Tp, Ts
(G6) Client |≡ Tp, Ts

< Hypothesis >
(H1) AP |≡ # (r)

< Delivation >
(D1) AP <| {r, puk}puk
(D2) AP |≡ Client |~ r
(D3) AP |≡ Client |≡ r
(D4) AP |≡ r
(D5) AP <| {PSK, c, r}r
(D6) AP |≡ Client |~ (PSK, c, r)
(D7) AP |≡ Client |≡ (PSK, c, r)
(D8) AP |≡ Client |≡ PSK
(D9) Client <| {Tp, Ts, r}r
(D10) Client |≡ AP |~ (Tp, Ts, r)
(D11) Client |≡ AP |≡ (Tp, Ts, r)
(D12) Client |≡ AP |≡ r
(D13) Client |≡ AP |≡ Tp, Ts
(D14) Client |≡ Tp, Ts
    
```

그림 3. 초기 인증 및 토큰 발급 과정 BAN-Logic 검증

```

< Idelization >
(14) Client → AP : {Is, K}K, {timec, Ts}Ts
(15) AP → Client : {timeap, timec, Ts, key}key

< Assumption >
(A7) AP |≡ K
(A8) AP |≡ Ts
(A9) AP |≡ # (Ts)
(A10) Client |≡ key
(A11) Client |≡ # (timec)

< Goals >
(G7) AP |≡ Client |≡ K
(G8) AP |≡ Client |≡ Ts
(G9) Client |≡ AP |≡ key
(G10) AP |≡ Client |≡ key

< Hypothesis >
(H2) AP |≡ # (K)

< Delivation >
(D15) AP <| {Is, K}K
(D16) AP |≡ Client |~ {Is, K}
(D17) AP |≡ Client |≡ {Is, K}
(D18) AP |≡ Client |≡ K
(D19) AP <| {timec, Ts}Ts
(D20) AP |≡ Client |~ (timec, Ts)
(D21) AP |≡ Client |≡ (timec, Ts)
(D22) AP |≡ Client |≡ Ts
(D23) Client <| {timeap, timec, Ts, key}key
(D24) Client |≡ AP |~ (timeap, timec, Ts, key)
(D25) Client |≡ AP |≡ (timeap, timec, Ts, key)
(D26) Client |≡ AP |≡ key
    
```

그림 4. 보안세션 설정 과정 BAN-Logic 검증

그림 4는 토큰 교환 후 보안세션을 설정하는 과정을 BAN-Logic을 통해 검증한 결과이다. (D22)에서  $T_s$ 에 대한 믿음을 도출하고, (D26)에서 key에 대한 믿음이 도출된다. AP의 K에 대한 간접믿음은 Hypothesis인 (H2)가 존재해야 도출이 가능하다. 또한, AP의 key에 대한 간접적인 믿음을 도출할 수 없다. 따라서, 키 신뢰성에 대한 취약점이 발생할 수 있다. 이를 보완하기 위해 key에 관한 확인 절차가 필요하다.

## III. 결 론

본 논문에서 살펴본 이중토큰을 이용한 Wi-Fi 보안 프로토콜은 4-way 핸드셰이크 프로토콜에서의 보안 문제점들과 반복적인 상호작용 요청을 개선하기 위하여 이중토큰 기술을 이용하여 설계되었다. BAN-Logic 검증을 통해 발견한 보안성 문제에는 AP가 랜덤키  $r$ 과 비밀키  $K$ 에 대한 fresh함을 믿는다고 생각하기 어렵고 AP의  $K$ 에 대한 간접믿음이 도출되지 않아 비밀키  $K$ 에 대한 키 신뢰성이 부족하다. 랜덤키  $r$ 과 비밀키  $K$ 에 대한 fresh함은 Hypothesis로 넣어 검증하고 비밀키  $K$ 에 대한 키 신뢰성 문제는 키 확인 절차를 추가하여 보완할 수 있으며, 이 절차를 검증하여 간접적인 믿음을 도출할 수 있다. 향후 연구에서는  $r$ 과  $K$ 의 fresh함을 보장할 수 있으면서 key에 대한 키 확인 절차를 추가한 프로토콜을 설계하고 구현하고자 한다.

## 참 고 문 헌

- [1] K. Benton, "The Evolution of 802.11 Wireless Security," UNLV Informatics, Spring 2010. [https://benton.pub/research/benton\\_wireless.pdf](https://benton.pub/research/benton_wireless.pdf)
- [2] IEEE 802.11i-2004, "Wi-Fi Protected Access II (WPA2)," 2004.
- [3] M. Jones and D. Hardt, "The OAuth 2.0 Authorization Framework: Bearer Token Usage," RFC 6750, Oct. 2012.
- [4] M. Burrows, M. Abadi, and R. Needham, "A logic of authentication," ACM Trans. Comput. Syst., vol. 8, no. 1, pp. 18-36, Feb. 1990.

# 객체 인식 기반 지능형 신호등 시스템 설계 및 제작

김도현\*, 박하은\*\*, 조현준\*\*\*, 강현수\*\*\*\*

## The Design and Implementation of an Intelligent Traffic Light System Based on Object Recognition

Kim Do-hyeon\*, Park Ha-eun\*\*, Cho Hyun-jun\*\*\*, and Kang Hyun-soo\*\*\*\*

### 요 약

본 연구에서는 객체 인식 기술을 활용한 지능형 신호등 시스템을 설계하였다. 라즈베리파이 5와 YOLOv5 모델을 사용해 교차로의 차량 수를 실시간으로 감지하고, 아두이노 메가를 통해 신호등 시간을 동적으로 조절하였다. 1,945장의 BMW i4 1/65 스케일 모형 차량 데이터를 사용해 모델을 학습한 결과, precision(정밀도) 0.99, recall(재현율) 0.99 등의 높은 성능을 기록하였다. 실시간 차량 수에 따른 신호 조정은 기존 고정 시간 신호 체계보다 교통 흐름을 개선하는 데 효과적임을 확인하였다.

### Abstract

This study designed an intelligent traffic light system based on object recognition technology. Using Raspberry Pi 5 and the YOLOv5 model, the system detects the number of vehicles at an intersection in real time and dynamically adjusts the traffic light timing via Arduino Mega. The model was trained on 1,945 images of a 1/65 scale BMW i4 model dataset, achieving high performance with a precision of 0.99 and a recall of 0.99. The system's dynamic signal adjustment based on real-time vehicle detection was found to improve traffic flow efficiency compared to conventional fixed-time traffic systems.

### Key words

Object recognition, YOLOv5, Intelligent traffic light, Real-time traffic control, Signal control

## 1. 서 론

최근 교통량 증가와 혼잡 문제는 점점 심각해지고 있다. 하지만, 기존 신호등 시스템은 주로 고정된 타이머를 기반으로 동작하여 실시간 교통 상황

을 반영하지 못하는 한계를 가지고 있다[1]. 이러한 한계는 불필요한 대기 시간 증가와 교통 혼잡을 초래한다.

따라서, 본 연구는 객체 인식 기술을 활용하여 교통 상황에 맞게 신호를 동적으로 제어할 수 있는 지능형 시스템을 설계하고 구현하는 것을 목표로

\*,\*\*,\*\*\* 충북대학교 정보통신공학부, {doneunim, daniel1530, tonycho98}@chungbuk.ac.kr,

\*\*\*\* 충북대학교 정보통신공학부, hskang@chungbuk.ac.kr (교신저자)

한다[2]. 이를 통해, 기존의 신호 체계와 비교하여 새로운 시스템의 성능과 효율성을 제시한다.

## II. 연구 내용

본 연구의 핵심 아이디어는 교차로에서 실시간으로 교통량을 분석하고, 가장 혼잡한 구간에 더 많은 신호 시간을 할당함으로써 차량 흐름을 최적화하는 것이다.

### 2.1 소프트웨어 설계 내용

이 시스템의 소프트웨어 설계는 차량 감지 모듈과 신호 제어 모듈로 나눌 수 있다.

차량 감지 모듈의 카메라는 신호등이 바뀌기 5초 전(노란불)에 활성화되며, 각 도로별 차량 수를 분석한다. 이때 사용되는 객체 인식 알고리즘으로는 YOLOv5 모델을 사용한다.

신호 제어 모듈에서는 각 도로의 초록불이 꺼지기 직전에 해당 도로의 차량 수가 다른 도로보다 많을 경우, 해당 도로의 초록불 시간을 20초 연장한다. 반면, 다른 도로의 차량 수가 더 많다면 다음 신호로 넘어간다.

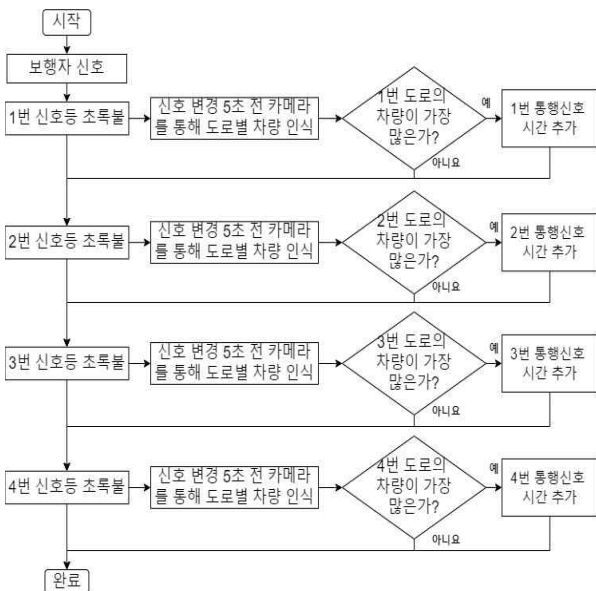


그림 1. 객체 인식 기반 신호등 시스템 플로우 차트  
Fig. 1. Flowchart of Object Recognition-Based Traffic Light System

### 2.2 하드웨어 설계 내용

라즈베리파이5는 객체 인식 및 데이터 처리를 담당하는 핵심 장치이다. 라즈베리파이5에 AI Kit를 결합하여, 차량 인식에 필요한 YOLOv5를 실행시킨다. 이 장치는 교차로에 설치된 카메라 모듈을 통해 실시간으로 각 도로의 차량 수를 감지하고 분석한다.

아두이노 메가는 USB 시리얼 통신을 통해 라즈베리파이5와 연결되며, 라즈베리파이5에서 송신된 차량 정보에 따라 각 도로의 신호등을 제어한다. 아두이노 메가는 각 신호등에 연결된 DM45 LED 모듈을 통해 교차로의 신호등을 조작하며, TM1637을 이용하여 신호등의 남은 시간을 표시한다.

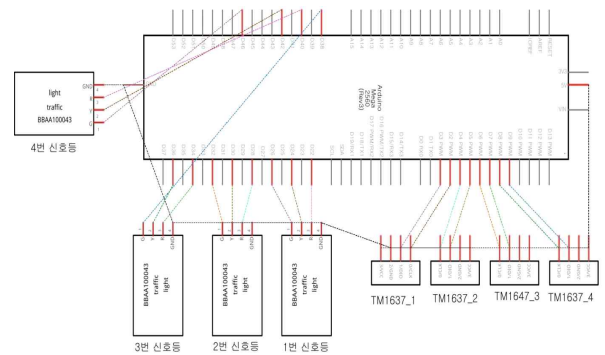


그림 2. 하드웨어 회로 설계도  
Fig. 2. Hardware Schematic

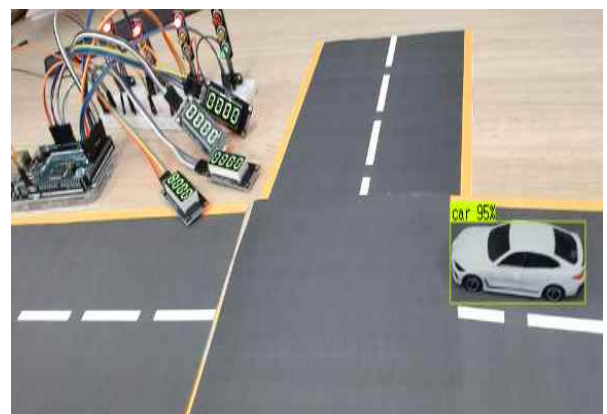


그림 3. 하드웨어 설계 및 구현  
Fig. 3. Hardware Implementation

### III. 실험(동작 테스트) 내용

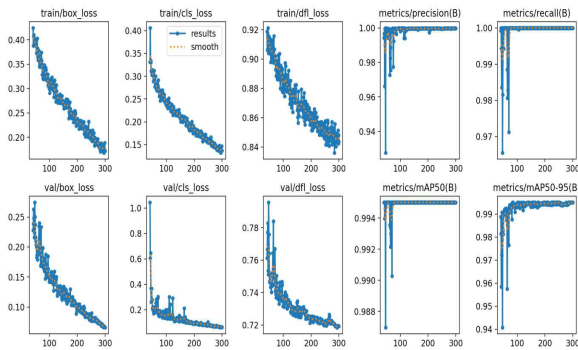


그림 4. YOLOv5 모델 성능 지표

Fig. 4. YOLOv5 Model Performance Metrics

본 연구에서는 BMW i4 1/65 스케일 모형 차량의 1,945장 이미지 데이터를 사용하여 YOLOv5 모델을 학습시켰다. 데이터셋은 train set, valid set, test set, 7:2:1의 비율로 학습 및 검증, 테스트를 수행하였다. 이를 통해 모델의 일반화 성능을 확인하고, 학습 중 과적합 여부를 검토하였다.

먼저, box\_loss는 학습 초반 약 0.40에서 시작하여 학습이 진행됨에 따라 약 0.20까지 감소하였다. 이는 모델이 점차 정확한 경계 상자를 예측하도록 학습되었음을 의미한다. cls\_loss는 초기 약 0.40에서 시작하여 점차 감소하여 학습 후반에는 약 0.15로 수렴하였다.

또한, DFL(Distribution Focal Loss)은 0.92에서 시작하여 0.84로 감소하였다. 검증 데이터에서도 비슷한 경향을 보였으며, validation set의 손실은 학습과 유사하게 감소하여 모델이 과적합 없이 안정적으로 학습되었음을 보여준다.

성능 지표를 살펴보면, Precision(정밀도)는 학습 초반 소폭 변동이 있었으나 학습이 진행됨에 따라 약 0.99 수준에서 안정적으로 유지되었으며, Recall(재현율)은 0.99로 유지되어 모델이 대부분의 객체를 정확히 검출하고 있음을 확인할 수 있다.

mAP50은 학습 초기부터 약 0.98 이상을 기록하여 매우 높은 성능을 보였고, 다양한 IoU 임계값에서 측정된 mAP50-95는 학습 중 소폭의 변동이 있었으나 결국 0.99 수준에서 수렴하였다.

### IV. 기대효과 및 활용

본 연구에서 제안한 객체 인식 기반 스마트 신호등 시스템은 기존의 고정 타이머 방식 신호등 시스템과 비교해 실시간으로 교통 상황을 반영하고, 차량 흐름을 효율적으로 관리할 수 있는 혁신적인 시스템이다. 이를 통해 교통 혼잡을 줄이고, 대기 시간을 단축할 수 있다.

### V. 결론

실험은 총 1,945장의 이미지를 사용하여 데이터셋을 train set (70%), valid set (20%), test set (10%)으로 분리해 진행되었으며, YOLOv5 모델은 높은 성능을 발휘했다. Precision(정밀도)과 Recall(재현율)은 약 0.99로, mAP50은 0.98, mAP50-95는 0.99로 수렴하여 모델이 차량을 정확하게 인식하는 데 탁월한 성능을 보였다. 특히, 실시간으로 교차로의 각 도로별 차량 수를 감지하여, 차량이 가장 많은 도로의 초록 불 시간을 20초 연장하는 방식은 차량 흐름을 크게 개선하는데 도움이 될 것으로 기대한다.

### 참고 문헌

- [1] Baek Kwang-Moo, Shin Ji-Hwan, & Park Mu-Hun, "Development of Auto Traffic Light Control System for Prevention of Traffic Jam". The Journal of Korea Institute of Convergence Signal Processing, vol. 15, no. 4, pp. 148-154, 2014.
- [2] Yeong-Pil An, Dong-choon Kim and Seung-kwon Na. "A Algorithm on Optimizing Traffic Network by the Control of Traffic Signal Timing." The Journal of Korea Navigation Institute, vol. 21, no. 5, pp. 472-478, 2017.



# 유니티 엔진을 활용한 국립금오공과대학교의 AR 길 안내 어플리케이션 개발

정혜리\*, 시종욱\*\*, 김성영\*\*\*

## Development of AR Navigation System Application for Kumoh National Institute of Technology Using Unity Engine

Hyeri Jeong\*, Jongwook Si\*\*, and Sungyoung Kim\*\*\*

### 요 약

본 논문에서는 국립금오공과대학교의 신입생과 방문객이 캠퍼스 내 건물을 쉽게 찾을 수 있도록 돕기 위한 AR 기반 길 안내 시스템을 제안한다. GPS 데이터, 네이버 지도 API, AR 기술을 통합하여 실시간으로 3D 화살표와 미니맵을 통해 직관적인 방향 안내를 제공한다. 제안방법은 캠퍼스 내에서 사용자의 만족도를 높이며, AR 기술을 활용한 위치 기반 서비스의 가능성을 확인하였다.

### Abstract

This paper proposes an AR-based navigation system designed to assist freshmen and visitors at Kumoh National Institute of Technology in easily locating campus buildings. By integrating GPS data, Naver Map API, and AR technology, the system provides intuitive directional guidance through real-time 3D arrows and a mini-map. The proposed method enhances user satisfaction within the campus and demonstrates the potential of AR technology in location-based services.

### Key words

AR(Augmented Reality), Navigation, Unity, LBS(Location-Based Service), GPS(Global Positioning System)

## 1. 서 론

최근 대학 캠퍼스 내에서 위치를 찾기 어려워하는 학생들이 증가하고 있으며, 특히 신입생과 방문

객들은 처음 방문하는 건물의 위치를 쉽게 찾지 못하는 경우가 많다. 이러한 문제를 해결하기 위해 다양한 방법이 시도되고 있지만, 기존의 정보 제공 방식은 사용자가 직관적으로 위치를 파악하기 어려운

\* 국립금오공과대학교 컴퓨터공학과

\*\* 국립금오공과대학교 컴퓨터·AI융합공학과

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학연협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

경우가 많다[1]. 따라서, 보다 상세하고 정확한 위치 정보를 제공하는 방식이 필요하다.

## II. 제안 방법

본 논문에서는 국립금오공과대학교의 길 안내를 제공하기 위해 유니티를 활용한 AR 기반의 시스템을 제안한다.

첫째, 사용자가 원하는 건물을 선택하고 길 찾기 버튼을 누르면, AR 카메라 화면에 3D 화살표가 나타나 건물의 방향을 안내한다. 이 화살표는 사용자의 현재 위치를 기반으로 실시간으로 조정되며, 목적지에 도착하면 도착 메시지를 출력한다. 이를 구현하기 위해 AR Plane Manager 컴포넌트를 통해 카메라로 바닥을 인식하고, AR Raycast Manager 컴포넌트를 통해 화살표 오브젝트를 사용자가 가야 할 방향에 배치한다. 사용자의 현재 위치와 목적지 간의 방위를 계산하기 위해서는 두 지점 간의 방위 계산이 필요하다. 이를 위해 해버사인 공식과 함께 CalculateBearing 메서드를 사용하여 방위 각도를 계산하고, PlaceOnPlane을 통해 화살표가 올바른 방향을 가리키도록 한다.

둘째, GPS와 위도, 경도 데이터를 이용하여 사용자의 현재 위치를 파악하고, 목적지까지의 남은 거리를 계산하여 실시간으로 표시한다. GPS로 받은 좌표 정보를 바탕으로 출발지와 도착지 간의 거리를 계산하였고, 이를 통해 사용자는 남은 거리와 이동 방향을 확인할 수 있다. 이 기능은 SetDestination 메서드를 사용해 사용자가 목적지를 설정하고, GetLocation 코루틴을 통해 GPS 위치를 주기적으로 갱신하는 방식이다. 또한, 목적지에 도착하면 화면에 도착 메시지를 출력하여 사용자가 목적지에 도달했음을 알릴 수 있도록 CheckArrival 메서드를 활용하였다. 이때, 거리가 일정 거리 이내일 경우 목적지에 도착한 것으로 간주하여 메시지를 출력한다.

셋째, 본 연구에서는 네이버 지도 API[2]를 활용하여 학교 지도를 받아와 미니맵을 구현한다. 이 미니맵은 화면 하단에 표시되며, 사용자가 자신의 위치와 경로를 쉽게 파악할 수 있도록 돕는다. 이를 이용해 지도 이미지와 상세 건물 데이터를 받아올

수 있었으며, 출발지와 목적지 간의 경로를 실시간으로 표시한다. 미니맵은 Static Map 기능을 사용해 유니티 내에서 지도 이미지를 RawImage 컴포넌트로 화면에 표시하고, 터치 입력을 통해 지도를 이동시키거나 확대/축소할 수 있도록 한다. 터치 입력을 감지하고 이를 처리하기 위해 HandleTouch Input 메서드를 사용하며, 터치의 각 단계에 따라 지도를 조작할 수 있는 메서드들을 호출하여 미니맵을 직관적으로 사용할 수 있도록 한다. 그림 1은 제안하는 방법에서의 길 찾기 UI와 도착 시 결과의 예시를 나타낸다.

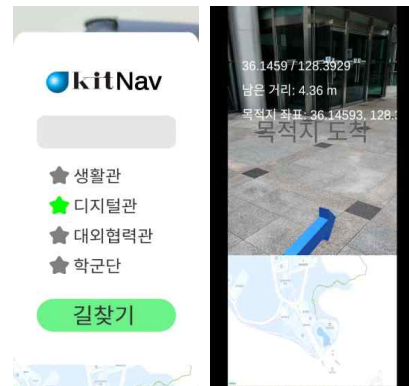


그림 1. 제안 방법의 예시

Fig. 1. Examples of proposed method

## III. 결 론

본 논문에서는 AR 기반 길 안내 시스템은 국립금오공과대학교의 신입생과 방문객들이 학교 내에서 목적지 건물을 빠르고 쉽게 찾을 수 있도록 하는 시스템을 제안하였다.

## 참 고 문 헌

- [1] J. Si, G. Kim, J. Kim, and S. Kim, "Enhanced Location-based Facility Management in Mobile Environments using Object Recognition and Augmented Reality", Journal of Korean Institute of Information Technology, Vol. 21, No. 11, pp. 183-192, Nov. 2023.
- [2] 네이버 지도 API, <https://www.ncloud.com/product/applicationService/maps>

# ControlNet을 활용한 Stable-Diffusion 기반의 면접관 전신 영상 생성

김대민\*, 시종욱\*\*, 김성영\*\*\*

## Full-Body Image Generation of Interviewer Based on Stable Diffusion Using ControlNet

Demin Kim\*, Jongwook Si\*\*, and Sungyoung Kim\*\*\*

### 요 약

가상 면접 플랫폼에서는 현실감 있는 다양한 가상 면접관을 생성하는 것이 중요한 과제로 부각된다. 그러나, Stable-Diffusion 기반의 영상 생성은 전신 면접관 영상을 일관성 있게 생성하는 데 어려움이 있다. 이를 해결하기 위해, 본 연구에서는 골격 정보를 반영한 사전 학습된 ControlNet 모델을 Stable-Diffusion에 적용하여, 정면을 보는 전신의 가상 면접관을 생성할 수 있는 방안을 제시한다.

### Abstract

In virtual interview platforms, generating realistic and diverse virtual interviewers has emerged as a critical challenge. However, image generation based on Stable-Diffusion faces difficulties in consistently producing full-body interviewer images. To address this issue, this study proposes the application of a pre-trained ControlNet model, which incorporates skeletal information into Stable-Diffusion, enabling the generation of front-facing, full-body virtual interviewers.

### Key words

Image Generation, ControlNet, Stable-Diffusion, Virtual Interview, Prompt Engineering

## 1. 서 론

최근 인공지능과 메타버스 기술의 급격한 발전은 다양한 분야에서 변화를 보이고 있으며, 그 중에서도 가상 면접 시스템의 발전은 이러한 기술 융합의

대표적인 사례로 주목받고 있다. 인공지능을 활용한 가상 면접 시스템은 보다 생동감 있고 몰입할 수 있는 새로운 면접 준비 방식을 제시하고 있다[1]. 이러한 시스템에서 면접관 영상을 어떻게 생성하느냐는 사용자가 느끼는 전체 경험의 질을 결정짓는

\* 국립금오공과대학교 컴퓨터공학과

\*\* 국립금오공과대학교 컴퓨터·AI융합공학과

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

※ 본 연구는 2024년도 과학기술정보통신부와 한국전파진흥협회의 메타버스랩지원사업에 의한 연구임

핵심 요소이다[2]. 그러나 영상 생성 기술은 전신 면접관을 일관성 있게 생성하는 데 한계가 있다. 이에 본 논문에서는 가상 면접 시스템의 사용자 경험을 극대화하기 위해 현실감 있고 일관된 전신 면접관 영상을 생성할 수 있는 방법을 제안하고자 한다.

## II. 가상 면접관의 전신 영상 생성

본 논문에서는 가상 면접관 영상을 생성하기 위해 Stable-Diffusion[3]을 기본 모델로 사용하고, 면접자가 현실감을 느낄 수 있는 체크포인트를 적용한다. 가상 면접관은 정장을 착용하고 서 있는 자세를 취하는 것으로 가정하며, 이러한 설정을 반영한 프롬프트로 "Interviewer, Wearing a suit, Clean background, Standing"을 사용한다. Openpose-editor를 통해 자세의 주요 키포인트를 설정하고, 이를 기반으로 ControlNet[4]을 활용하여 면접관 영상을 생성하는 방법을 제안한다. 이때 영상의 크기는 (512, 1024)로 고정한다. 이러한 방식은 원하는 자세를 취한 가상 면접관을 안정적으로 생성할 수 있다. 그림 1에서는 가상 면접관 생성을 위한 포즈더미와 이를 기반으로 생성된 면접관 이미지를 보여준다. 생성된 영상은 프롬프트와 포즈더미를 잘 반영하였으며, 다양하고 현실감이 높은 결과임을 확인할 수 있다.

## III. 결 론



그림 1. 포즈더미 예시와 생성한 가상 면접관

Fig. 1. Example of Pose Dummy and Generated Virtual Interviewer

본 논문에서는 정면의 전신 자세를 설정한 후, Stable Diffusion을 사용하여 가상 면접관을 생성하는 방법을 제안하였다. 이를 통해 일관성 있는 면접관 이미지를 생성할 수 있음을 확인하였다.

## 참 고 문 헌

- [1] C. Yoon, *et al.*, "Metaverse Virtual Interview Platform Leveraging Generative AI and Speech Recognition", *The Journal of Korean Institute of Information Technology*, Vol. 22, No. 6, pp. 163-173, 2024.
- [2] J. Si, S. Yang, D. Kim, and S. Kim, "Metaverse Interview Room Creation with Virtual Interviewer Generation Using Diffusion Model", *IEEE Asia-Pacific Conference on Computer Science and Data Engineering*, Dec. 2023.
- [3] R. Rombach, B. Ommer, *et al.*, "High-Resolution Image Synthesis with Latent Diffusion Models", In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 10684-1065, Jun. 2022.
- [4] L. Zhang, A. Rao, A., and M. Agrawala, "Adding conditional control to text-to-image diffusion models", In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3836-3847, 2023.

# 언리얼 엔진을 이용한 가상 면접 콘텐츠의 프로토타입 개발

조재혁\*, 시종욱\*\*, 김성영\*\*\*

## Development of Virtual Interview Contents Prototype Using Unreal Engine

Jaehyuck Jo\*, Jongwook Si\*\*, and Sungyoung Kim\*\*\*

### 요약

본 논문에서는 사용자가 환경의 제약을 받지 않고 몰입할 수 있는 가상 면접 시스템을 언리얼 엔진을 통해 구현한 콘텐츠를 제안한다. 언리얼 엔진의 고품질 렌더링과 상호작용 기능을 활용하여, 현실감 넘치는 면접 콘텐츠를 제공함으로써 실제와 같은 면접을 경험하고, 반복적인 훈련을 통해 면접에 대한 실력을 향상시킬 수 있을 것으로 기대한다.

### Abstract

This paper proposes a virtual interview contents implemented using Unreal Engine, which allows users to immerse themselves in a realistic interview environment without being limited by physical constraints. Utilizing Unreal Engine's high-fidelity rendering and interactive capabilities, the system delivers highly realistic interview content, enabling users to experience interview scenarios that closely resemble real-world situations. This immersive approach is expected to enhance interview skills through repeated, structured practice sessions.

### Key words

Virtual Reality, Metaverse, Unreal Engine, Interview, Metahuman

## 1. 서론

청년 취업의 어려움이 증가함에 따라 면접의 중요성 또한 크게 부각되고 있다. 이러한 상황에 맞춰, 본 논문에서는 면접 준비에 몰입할 수 있는 언리얼 엔진 기반의 가상 면접 콘텐츠를 제안한다. 언

리얼 엔진은 고품질의 실시간 렌더링 기능을 제공하여 더욱 현실감 있는 시각적 경험을 가능하다는 장점이 있다. 또한, 캐릭터 및 환경의 동작을 더욱 정교하고 자연스럽게 구현할 수 있다. 이를 통해 사용자는 실제와 유사한 면접 환경에서 더욱 높은 몰입감을 느낄 수 있는 콘텐츠를 제안한다.

\* 국립금오공과대학교 컴퓨터공학과

\*\* 국립금오공과대학교 컴퓨터·AI융합공학과

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학연협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

## II. 가상 면접 콘텐츠의 프로토타입

본 논문에서는 현실감 있는 몰입감을 제공하기 위해 언리얼 엔진을 기반으로 구현한 가상 면접 시스템을 제안한다. 기존의 가상 면접 시스템[1]과 유사한 흐름을 따르지만, 사용자가 더욱 실제에 가까운 면접 경험을 느낄 수 있도록 콘텐츠를 설계하였다. 특히, 메타휴먼[2]을 면접관으로 설정하고, 자연스러운 립싱크 및 표정 변화를 구현하기 위해 Nvidia의 Audio2Face[3] 기술을 활용하였다. 그림 1은 개발한 언리얼 엔진 내의 면접장을 나타낸다.

면접이 시작되면, 사용자는 UI 창을 통해 이름, 지원하는 회사, 전공을 입력하고 이 정보는 클라우드 서버로 전송된다. 면접이 진행되면서 사용자가 녹음한 답변은 실시간으로 서버에 전송되며, 서버는 답변을 분석해 질문을 생성하여 면접을 이어간다. 생성한 질문은 TTS(음성 합성) 기술을 통해 음성 파일로 변환하여 클라이언트로 전달한다.

콘텐츠로 전달된 음성 파일은 Audio2Face 파이프라인을 통해 처리되며, 메타휴먼 면접관의 얼굴 애니메이션과 자연스러운 립싱크를 구현할 수 있다 [4]. 이 기술을 통해 사용자는 실제 면접에 가까운 몰입감을 경험할 수 있으며, 면접관의 립싱크와 발음이 자연스러워 면접 환경의 품질이 크게 향상됨을 확인할 수 있다. 또한, 면접관의 자연스러운 행동을 위해 다양한 Idle Animation을 적용하여 면접자가 답변을 하는 동안 면접관은 시선을 돌리거나 필기하는 등의 동작을 수행한다. 이를 통해 생동감 있는 면접 환경을 제공할 수 있다.

## III. 결 론

본 논문에서 제안한 가상 면접 시스템은 언리얼 엔진과 메타휴먼, Audio2Face를 활용하여 현실감 있는 면접 환경을 구현하였다, 이를 통해 기존 가상 면접 시스템의 한계를 극복하였으며, 실제 면접에 가까운 몰입감을 제공할 수 있음을 보였다. 또한, 다양한 Idle Animation 기능을 통해 면접관의 행동을 더욱 자연스럽게 표현할 수 있음을 확인하였다.

## 참 고 문 헌

- [1] C. Yoon, S. Yang, J. Park, J. Si, Y. Jung, and S. Kim, "Metaverse Virtual Interview Platform Leveraging Generative AI and Speech Recognition", The Journal of Korean Institute of Information Technology, Vol. 22, No. 6, pp. 163-173, 2024.
- [2] EpicGames Metahuman, <https://www.unrealengine.com/en-US/metahuman> [Accessed: Oct. 6, 2024]
- [3] Nvidia Audio2Face, <https://build.nvidia.com/nvidia/audio2face> [Accessed: Oct. 6, 2024]
- [4] J. Jo, J. Si, and S. Kim, "Real-Time Facial Expression Changes for Interviewers in Virtual Interviews: Application of Audio2Face in Unreal Engine", Proceedings of KIIT Conference, pp. 404-405, May. 2024.

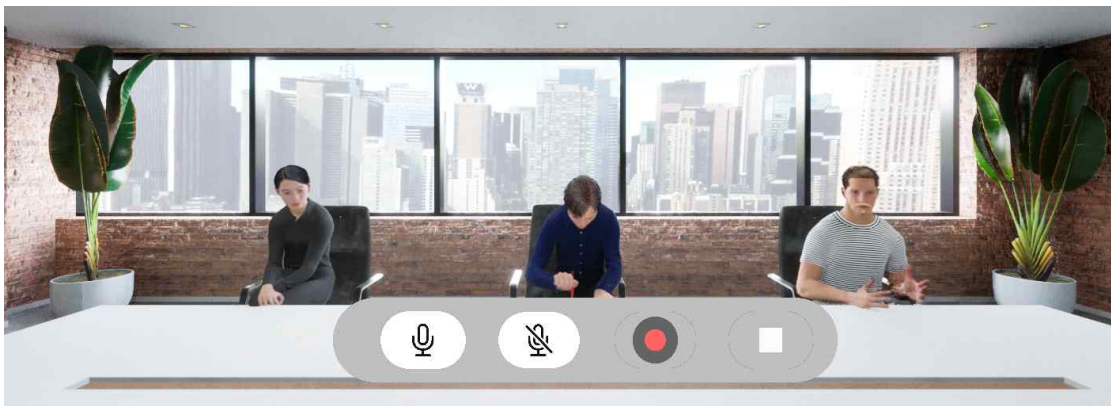


그림 1. 언리얼 엔진 내의 면접장 예시  
Fig. 1. Example of interview room in unreal engine

# 객체 인식을 활용한 메타버스 상점의 자동 변환 인터페이스 개발

정혜리\*, 김대민\*, 시종욱\*\*, 김성영\*\*\*

## Development of an Interface for Automatic Transformation of Metaverse Stores Using Object Detection

Hyeri Jeong\*, Daemin Kim\*, Jongwook Si\*\*, and Sungyoung Kim\*\*\*

### 요 약

디지털 기술의 발전과 함께 오프라인 소매업계는 온라인으로의 확장을 가속화하고 있으며, O2O 서비스와 메타버스 환경에서의 쇼핑 경험에 대한 관심이 증가하고 있다. 본 논문은 객체 검출 기술을 사용하여 상품을 자동으로 인식하고, 이를 바탕으로 Unity 환경에서 가상의 상점을 자동으로 구축하는 인터페이스를 제안한다. 객체 검출 모델을 활용하여 추론 결과를 JSON 파일로 저장하고, 이를 바탕으로 Unity에서 3D 오브젝트를 생성하여 실제 상점과 유사한 형태로 가상 공간을 구성하는 방법을 제시한다.

### Abstract

With the advancement of digital technology, the offline retail industry is accelerating its expansion into the online domain, leading to increased interest in O2O services and shopping experiences within the metaverse. This paper proposes an interface that leverages object detection technology to automatically recognize products and, based on this data, constructs virtual stores within the Unity environment. The method involves storing the inference results of the object detection model in JSON format and utilizing this data to generate 3D objects in Unity, thereby creating a virtual space that closely resembles a physical store.

### Key words

O2O Stores, Metaverse, Object Detection, Interface, Unity

## 1. 서 론

디지털 기술의 발전은 오프라인 소매업계의 온라

인 확장을 촉진하며, O2O(Offline to Online) 서비스로의 전환을 가속화하고 있다[1-2]. 이와 함께 메타버스 환경에서의 쇼핑 경험이 중요해지고 있으며,

\* 국립금오공과대학교 컴퓨터공학과

\*\* 국립금오공과대학교 컴퓨터·AI융합공학과

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

이러한 환경에서 가상 상점을 현실감 있게 구성하기 위한 기술 개발이 필요해지고 있다.

환경에서 구현할 수 있음을 확인하였다.

## II. 자동 변환 인터페이스 기술

본 논문에서는 실제 상점의 상품 정보를 바탕으로 메타버스 환경에서 가상의 상점을 자동으로 구축할 수 있는 방법론을 제안한다. 객체 검출 기술인 YOLO-v8[3]을 활용하여 상점 내 진열대의 상품 위치를 검출하고, 검출된 객체의 좌표 및 이미지를 JSON 파일로 저장하는 방식이다. 이러한 JSON 데이터를 기반으로 Unity 환경에서 3D 오브젝트를 생성하고, 실제 상점의 진열대와 유사한 형태로 가상의 상점을 구성한다.

YOLO-v8 모델은 사전 학습된 모델을 파인튜닝하여 다양한 상품을 인식할 수 있도록 하였으며, 인식된 상품의 좌표 정보는 JSON 파일로 저장된다. 이 좌표 정보는 Unity에서 3D 오브젝트를 생성하고 배치하는 데 활용된다. JSON 파일의 데이터는 Unity의 `Resources.Load<TextAsset>` 메소드와 `Newtonsoft.Json` 라이브러리를 사용하여 로드되며, 각 상품의 위치와 크기에 맞춰 3D 오브젝트가 자동으로 배치된다. 이를 통해 그림 1과 같이 실제 상점의 진열대와 유사한 가상 상점을 자동으로 구축할 수 있다.

객체 검출 결과를 바탕으로 Unity에서 생성된 3D 오브젝트는 JSON 파일에 기록된 좌표를 기반으로 배치되었으며, 실제 상점과 유사한 진열대를 가상

## III. 결 론

본 논문은 YOLO-v8 객체 인식 기술을 활용하여 메타버스 환경에서 가상의 상점을 자동으로 구축할 수 있는 인터페이스를 제안하였다. 이를 통해 오프라인 상점의 상품 정보를 가상 환경으로 효과적으로 변환하여, 사용자가 메타버스에서 현실감 있는 쇼핑 경험을 할 수 있도록 지원할 수 있다.

이러한 가상 상점의 사용자 경험을 더욱 향상시키기 위해 인터페이스의 최적화 및 추가 기능 개발이 필요할 것으로 보인다.

## 참 고 문 헌

- [1] D. Kim, J. Si, S. Lee, and S. Kim, "Calculation of Product Location Based on Object Detection and Product name recognition through Image Similarity Measurement", Proc. Of Korean Institute of Information Technology Conference, pp. 500-501, Nov. 2023.
- [2] M. Woo, J. Si, J. Jo, and S. Kim, "Development of a Chatbot Prototype for Customer Intent Classification in O2O Stores", Proc. Of Korean Institute of Information Technology Conference, pp. 446-447, Nov. 2023.
- [3] YOLOv8, <https://github.com/ultralytics/ultralytics>



그림 1. 자동 변환된 상점의 진열대  
Fig. 1. Automatically Transformed Store Displays



# 위치 요소 기반 사용자 인증 프로토콜에 대한 정형검증 연구

이수왕\*, 이주현\*, 김지윤\*\*

## Study on Formal Verification for Location Factor Based User Authentication Protocol

Soowang Lee\*, JooHyeon Lee\*, and Jiyeon Kim\*\*

### 요 약

본 논문은 모바일 환경에서의 위치 요소 기반 사용자 인증 프로토콜에 대해 정형 검증 도구인 BAN-Logic을 이용해 보안성을 검증한다. 제안된 프로토콜은 총 2단계로, 1단계에서 위치를 확인하고 2단계에서 위치를 검증하는 방식으로 진행된다. 프로토콜은 인증 과정에서 NFC와 HTTPS를 활용하지만, 주고받는 정보들의 내부 요소에 대한 보안성 검증이 부족하다. 따라서 해당 프로토콜을 BAN-Logic을 활용해 검증하였고 개인키 신뢰성과 난수의 freshness 문제 등의 보안 취약점이 발견되었다.

### Abstract

This paper validates the security of location factor based user authentication protocols on mobile environment using BAN-Logic, a formal verification tool. The proposed protocol comprises two steps, where step 1 checks the location and step 2 verifies it. The authentication process involves NFC and HTTPS, but the security verification of internal elements in the exchanged information is inadequate. Consequently, the protocol underwent verification using BAN-Logic, revealing security vulnerabilities related to private key reliability and random number freshness.

### Key words

Authentication protocol, mobile environment, multifactor authentication, Security vulnerability, Verification, BAN-logic

## I. 서 론

스마트폰과 스마트 워치 등의 모바일 기기들이 보편화됨에 따라 모바일 환경에서의 사용자 인증 수단의 중요성은 갈수록 부각되고 있다[1]. 근래의 스마트폰의 경우 지문 인식과 얼굴 인식 등의 사용자의 고유한 특징을 활용하여 인증 과정에서의 보안성을 확보하고자 하고 있다[2]. 사용자의 고유 특징을 활용함으로써 네트워크 연결이 필요 없게 되어 강력한 사용자 인증이 가능하게 된다. 하지만 이

러한 인증 방식은 사용자의 생체 특징을 인식하는 기능이 필요하기 때문에 여전히 많은 기기에서 지원하지 못하는 문제가 존재한다.

본 논문에서 검증하는 프로토콜은 모바일 기기의 생체 인증 수단에 의존하지 않고 대부분의 모바일 기기에서 지원하는 GPS를 활용하여 사용자 인증을 진행한다. 인증 프로세스는 총 2단계로 이루어져 있는데, 인증 프로세스 1단계에서 사용자의 디바이스에 설치된 MA(Mobile Application)와 사용자 근처에 존재하는 WA(Witness Application)가 NFC를 통

\* 경상국립대학교 컴퓨터공학과, id0311sw@gnu.ac.kr, wngysl12@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, ivkim92@gnu.ac.kr(교신저자)

해 각자의 정보를 주고받아 암호화하여 AS (Authentication Service)에 정보를 전달한다. 2단계에서는 MA와 POS(Point Of Sale)는 NFC를 통해 서로의 정보를 암호화하여 AS 서버로 전송한다. AS는 전달받은 정보들을 복호화, 검증하여 무결성을 확인한 후 결과를 FCM 서버를 거쳐 MA에 인증 결과를 알린다[3].

본 논문은 본론에서는 프로토콜의 절차를 알아보고, BAN-Logic을 이용하여 정형 검증을 진행하여 결론에서는 BAN-Logic의 결과에 대한 분석으로 결론을 맺는다.

## II. 본론

### 2.1 위치 요소 기반 사용자 인증 프로토콜 절차

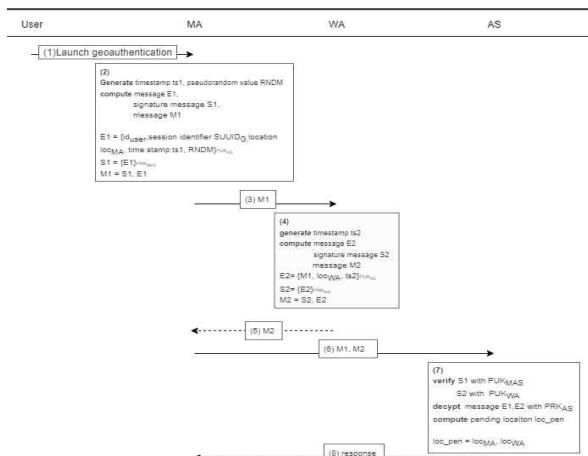


그림 1. 신뢰할수 있는 사용자 선언 프로토콜

위치 요소 기반 사용자 인증 프로토콜은 두 단계로 이루어져 있다. [그림 1]은 위치 요소 기반 프로토콜의 1단계인 신뢰할 수 있는 위치를 선언하는 과정이다. 절차의 세부 내용은 다음과 같다.

- (1) 사용자는 모바일 애플리케이션에서 버튼을 눌러 지리적 위치 정보를 활용한 인증 방법을 시작한다.
- (2) 시작 신호를 받은 MA는 E1과 S1을 만든다.
- (3) E1과 S1을 합쳐 M1을 WA에 보낸다.
- (4) M1을 받은 WA는 E2와 S2를 만든다.
- (5) E2와 S2를 합쳐서 M2를 MA에 보낸다.
- (6) MA는 M1, M2를 AS에게 보낸다.

- (7) AS는 S1과 S2의 서명을 검증하고, E1과 E2를 해독하여, loc\_pen을 만든다.
- (8) 처리된 요청의 상태를 MA에 반환한다.

[그림 2]는 1단계에서 선언한 위치를 확인하는 단계이다. 2단계의 세부 설명은 다음과 같다.

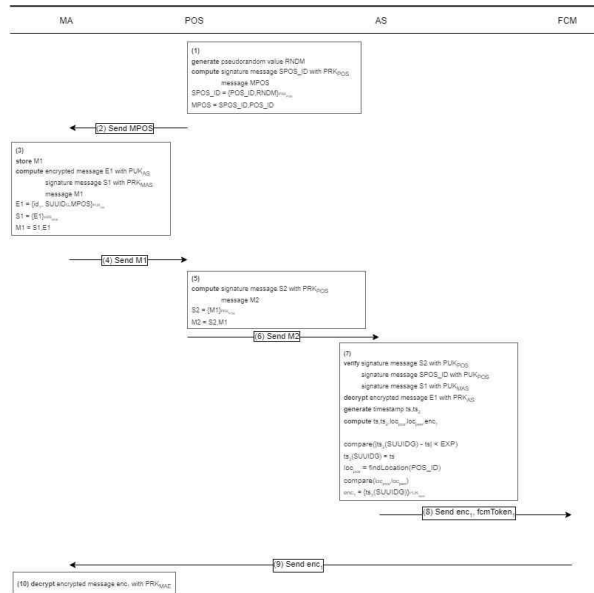


그림 2. 위치 확인 프로토콜

- (1) POS는 SPOS\_ID 과 MPOS를 만든다.
- (2) POS는 MPOS를 MA로 전송한다.
- (3) E1과 S1을 만든다.
- (4) MA는 M1을 만들어서 POS에게 전달한다.
- (5) POS는 S2와 M1으로 M2 만든다.
- (6) M2를 AS로 전송한다.
- (7) AS는 전송받은 S2와 S1을 검증한다. E1을 복호화한다. ts2와 ts1의 차이가 EXP보다 크지 않은 경우 검증이 성공적이니 ts3을 설정한다. SPOS\_ID의 무결성을 검증하고 POS의 위치를 AS의 레지스트리에서 찾아 1단계에서 저장한 임시 위치와 비교한다. 만약 값이 같다면, AS는 POS의 위치를 올바른 사용자 위치로 설정한다. 위치 확인 이후 enc1을 만든다.
- (8) enc1과 fcmToken1을 FCM 서버로 보낸다.
- (9) FCM은 전달받은 enc1을 MA로 보낸다.
- (10) MA는 전달받은 enc1을 복호화한다. 복호화가 끝나면 사용자는 인증 과정의 결과를 볼 수 있다.

## 2.2 BAN-Logic 기반 정형 검증

BAN-Logic은 보안 프로토콜의 보안성을 검증하기 위해 고유한 기호와 규칙을 사용하여 정규화, 가정, 목표, 유도 과정을 수행하는 강력한 도구이다 [4]. 본 논문에서는 제안된 프로토콜에 대해 BAN-Logic을 활용한 정형 검증을 진행한다.

<Idealization>

(I1) MA  $\rightarrow$  AS :  $\{(ID_{user}, SUUID_G, loc_{MA}, loc_{WA}, ts_1, RNDM_{MA})\}_{P_{UK_{AS}}}\}_{P_{UK_{MA}}}$

<Assumptions>

(A1) AS believes  $\frac{P_{UK_{AS}}}{P_{UK_{AS}}}$  AS

(A2) AS believes fresh( $ts_1$ )

<Hypothesis>

(H1) AS believes  $\frac{P_{UK_{MA}}}{P_{UK_{MA}}}$  MA

<Derivation>

(D1) AS sees  $\{(ID_{user}, SUUID_G, loc_{MA}, loc_{WA}, ts_1, RNDM_{MA})\}_{P_{UK_{AS}}}\}_{P_{UK_{MA}}}$

(D2) AS believes MA said  $(ID_{user}, SUUID_G, loc_{MA}, loc_{WA}, ts_1, RNDM_{MA})$  by D1, A1, H1, MM

(D3) AS believes MA believes  $(ID_{user}, SUUID_G, loc_{MA}, loc_{WA}, ts_1, RNDM_{MA})$  by D2, A2, FR, NV

(D4) AS believes MA believes  $loc_{MA}$  by D3, BC

(D5) AS believes MA believes  $loc_{WA}$  by D3, BC

(D6) AS believes MA believes  $SUUID_G$  by D3, BC

<Goals>

(G1) AS believes MA believes  $loc_{MA}$

(G2) AS believes MA believes  $loc_{WA}$

(G3) AS believes MA believes  $SUUID_G$

<Idealization>

(I2) MA  $\rightarrow$  AS :  $\{(ID_U, SUUID_G, \{POS_{ID}, RNDM_{MA}\}_{P_{UK_{AS}}}\}_{P_{UK_{MA}}}\}$

(I3) AS  $\rightarrow$  MA :  $\{ts_u, SUUID_G\}_{P_{UK_{MA}}}$

<Hypothesis>

(H1) AS Believes fresh( $RNDM$ )

<Assumption>

(A1) AS believes  $\frac{P_{UK_{MA}}}{P_{UK_{MA}}}$  MA

(A2) AS believes  $\frac{P_{UK_{AS}}}{P_{UK_{AS}}}$  AS

(A3) AS believes  $\frac{P_{UK_{POS}}}{P_{UK_{POS}}}$  POS

(A4) MA believes  $\frac{P_{UK_{MA}}}{P_{UK_{MA}}}$  MA

(A5) MA believes fresh( $ts_u$ )

<Derivation>

(D1) AS sees  $\{(ID_U, SUUID_G, \{POS_{ID}, RNDM_{MA}\}_{P_{UK_{AS}}}\}_{P_{UK_{MA}}}\}$  by I2

(D2) AS believes MA said  $(ID_U, SUUID_G, POS_{ID}, RNDM_{MA})$  by D1, A1, A2, A3, MM

(D3) AS believes MA believes  $(ID_U, SUUID_G, POS_{ID}, RNDM_{MA})$  by D2, H1, FR, NV

(D4) AS believes MA believes  $SUUID_G$  by D3, BC

(D5) AS believes MA believes  $POS_{ID}$  by D3, BC

(D6) MA sees  $\{ts_u, SUUID_G\}_{P_{UK_{MA}}}$  by I3

(D7) MA believes AS said  $(ts_u, SUUID_G)$  by D5, A4, MM

(D8) MA believes AS believes  $(ts_u, SUUID_G)$  by D6, A5, FR, NV

(D9) MA believes AS believes  $SUUID_G$  by D7, BC

<Goal>

(G1) AS believes MA believes  $SUUID_G$

(G2) AS believes MA believes  $POS_{ID}$

(G3) MA believes AS believes  $SUUID_G$

1단계 프로토콜은 Goals를 도출하는 과정에서 MA에 개인키에 대한 신뢰성이 증명되지 않는다. 2단계 프로토콜은 Goals를 도출하는 과정에서 가설이 발생하였는데, 이는 재사용 가능성에 취약점을 가지고 있다. 따라서, 위치기반 프로토콜은 취약점을 가지고 있다.

## III. 결 론

본 논문에서 살펴본 프로토콜은 생체 기반 인증 방식 대신 위치 기반 인증 방식을 채택하였다. 생체 인증을 지원하지 못하는 디바이스가 많기 때문에 대부분의 디바이스에서 지원하는 GPS 등을 이용하여 위치를 통해 인증을 진행하는 방식이다. 프로토콜의 세부 단계는 디바이스 간의 위치 정보를 확보하는 1단계와 모바일 디바이스와 POS 기를 통한 위치 검증 및 확인 과정인 2단계로 진행된다.

본 논문에서 BAN-Logic으로 해당 프로토콜을 확인한 결과 프로토콜의 1단계에서는 MA의 개인키에 대한 신뢰성 문제와 2단계에서는 난수의 freshness 문제가 발견되었다. 이를 통해 프로토콜의 보안성에 취약점이 있음을 알 수 있으며 추후에 취약점을 해결할 방안과 프로토콜 최적화에 관한 연구를 진행하고자 한다.

## 참 고 문 헌

- [1] D. Hintze, P. Hintze, R. D. Findling, and R. Mayrhofer, "A large-scale, long-term analysis of mobile device usage characteristics", Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 1, no. 2, Article 13, June 2017, pp. 21.
- [2] C. Wang, Y. Wang, Y. Chen, H. Liu, and J. Liu, "User authentication on mobile devices: Approaches, threats, and trends", Computer Networks, vol. 170, 2020, pp. 107118.

- [3] M. Bartłomiejczyk, I. El Fray, M. Kurkowski, S. Szymoniak, and O. Siedlecka-Lamch, "User authentication protocol based on the location factor for a mobile environment", *IEEE Access*, vol. 10, 2022, pp. 16439-16453.
- [4] M. Burrows, M. Abadi, and R. Needham, "A logic of authentication", *ACM Transactions on Computer Systems*, vol. 8, no. 1, pp. 18-36, Feb. 1990.

# 3D 손 추적 기반의 저비용 혼합 현실(MR) 시스템

연수민\*, 이민재\*\*, 이수원\*\*\*

## Low-cost Mixed Reality System based on 3D Hand Tracking

Sumin Yeon\*, Minjae Lee\*\*, and Suwon Lee\*\*\*

### 요 약

본 논문은 저비용 스마트폰 기반 HMD에서의 상호작용 수단 부재 문제를 해결하기 위해, 실시간 3D 손 추적 시스템을 제안한다. 손을 추적하여 얻은 위치를 카메라, 손, 월드 좌표계 간 변환을 통해 가상 환경에 반영함으로써, 사용자가 손을 이용해 가상의 객체와 상호작용할 수 있도록 구현하였다. 이 시스템은 추가적인 장비 없이도 저비용 MR 환경에서 자연스럽게 직관적인 3D 상호작용을 제공하여 사용자 경험을 향상시키며, 다양한 MR 시나리오 적용 가능성을 제시한다.

### Abstract

This paper proposes a real-time 3D hand tracking system to address the lack of interaction methods in low-cost smartphone-based HMDs. By tracking the hand and mapping its position through transformations between camera, hand, and world coordinate systems, the system allows users to interact with virtual objects in the environment. Without the need for additional hardware, this system enables natural and intuitive 3D interactions in low-cost MR environments, enhancing the user experience and offering potential for application in various MR scenarios.

### Key words

mixed reality, low-cost hand tracking, 3D interaction, coordinate transformation

## I. 서 론

최근 혼합 현실 (Mixed Reality, MR) 기술의 발전은 사용자 경험을 크게 향상시키며 그 수요를 증가시켰다. 그러나 MR 기기의 높은 비용은 접근성을 제한하여, 저비용 MR 환경의 필요성이 대두되고 있다[1]. Cardboard 등의 스마트폰 기반 헤드 마운티드 디스플레이 (HMD)는 더 많은 사용자가 저비용으

로 MR을 경험할 수 있게 하지만, 상호작용 수단의 부족으로 MR 시나리오의 활용 범위를 제한한다[2].

이를 해결하기 위해, 본 논문에서는 저비용 환경에서 손을 3D로 추적하는 시스템을 개발하고자 한다. 이 시스템은 추가적인 기기 없이 사용자에게 가상 환경과 상호작용을 할 수 있는 수단을 제공하여 사용자 경험을 향상시키는 것을 목표로 한다. 손 추적 기술을 통합함으로써 사용자는 가상 객체

\* 경상국립대학교 컴퓨터공학과, sumin.yeon@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, wjdchs0129@gnu.ac.kr

\*\*\* 경상국립대학교 컴퓨터공학과, leesuwon@gnu.ac.kr (교신저자)

를 보다 직관적이고 자연스럽게 조작할 수 있다. 구현 방법에 있어서 특정 수단에 제한되지 않는 범용성을 지니고 있어 다양한 MR 시나리오에 활용될 수 있다.

본 논문에서는 이러한 시스템의 구현과 저비용 MR 환경의 혁신 가능성을 논의한다.

## II. 좌표계 변환을 통한 손 추적 시스템

그림 1은 제안된 방법의 개요를 보여준다. 사용자가 보게 될 시야를 담당하는 스마트폰이 실시간으로 영상을 촬영하고, 이를 통해 MR 콘텐츠와 상호작용 할 수 있게끔 마커와 손의 위치 정보를 추적한다.

본 논문에서는 월드라는 객체가 렌더링 될 가상 공간을 정의하기 위해 마커를 도입하였으며, 이는 MR 시나리오에 따라 임의의 평면, 특정 객체, 또는 건물과 같은 다양한 형태를 취할 수 있다. 또, 2D 영상에서 손을 추적하기 위한 모듈로 MediaPipe의 hand landmark detection module[3]을 사용하였다. 이 외에도 다양한 손 추적 모듈을 사용할 수 있다.

MediaPipe의 모듈을 사용하면, 손의 각 관절에 대한 21개의 2D 랜드마크 좌표를 얻게 된다. 3D 좌표의 경우, 손의 기하학적 중심이 미터 단위의 현실 세계 좌표로 반환되는데, 이는 앞서 정의된 월드에서는 직접 렌더링할 수 없다. 따라서 월드 좌표계로 변환하기 위해, 중지의 기저를 나타내는 9번째 랜드마크

마크의 좌표를 각 관절의 3D 좌표에서 빼서 손 좌표계로 변환한다.

그림 2에서는 카메라, 손, 월드 좌표계 간의 변환 과정을 보여준다. 마커의 3D-2D 점 대응을 사용하여, PnP(perspective-n-point) 자세 계산 알고리즘을 통해 월드에서 카메라 좌표계로의 변환 행렬(W2C)을 계산한다[4]. 이는 3D-2D 점 대응에서 발생하는 재투영 오차를 최소화하는 회전 및 평행 이동을 해결하는 방식이다. W2C의 역행렬을 취하면 카메라에서 월드 좌표계로의 변환 행렬(C2W)이 되며, 이를 카메라에 적용한다. 또한, 손 좌표계를 카메라 좌표계로 변환하는 변환 행렬(H2C)은 손의 3D 및 2D 좌표를 사용하여 계산된다. 손에서 월드 좌표계의 직접 변환은 불가능하므로, 카메라 좌표계를 중간 매개체로 변환 행렬(H2W)을 계산하고, 이를 모든 손 관절 좌표에 적용한다.

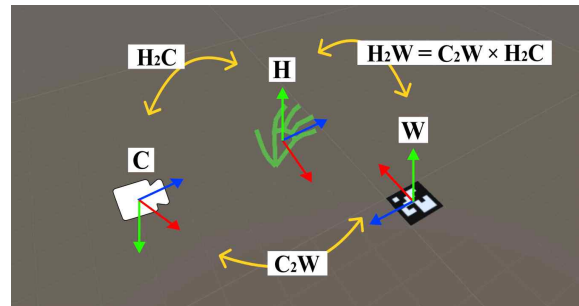


그림 2. 카메라, 손, 월드 좌표계 간의 변환 과정  
Fig. 2. Transformation process among camera, hand, and world coordinate systems

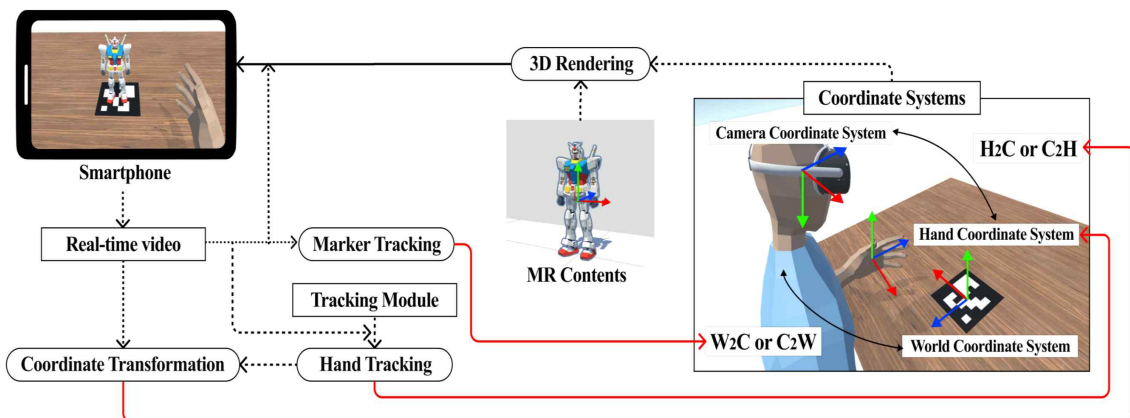


그림 1. 개발 과정 개요  
Fig. 1. Development process overview

### III. 구현

월드를 정의하기 위해 ArUco 마커 방법[5]을 사용하였으며, 이는 다른 방법에 비해 사용자 친화적이고 접근성이 뛰어나다. 현실 세계에 임의의 크기의 마커를 배치하여 현실과 월드 상의 위치를 연결한다. 마커 추적을 위해, 마커의 네 모서리의 2D 좌표를 반환하는 마커 인식 과정을 거친다. 마커의 3D 좌표는 첫 번째 모서리를 월드 상의 좌표 (0, 0, 0)로 직접 지정하고, 나머지 모서리의 좌표는 첫 번째 모서리부터 시계 반대 방향으로 마커 변의 길이를 더하여 결정된다. 손 추적을 위해 특별한 장비는 필요하지 않으며, 월드 좌표계로 변환된 좌표를 렌더링 하기위한 3D환경으로는 Unity3D를 선택하였다.

그림 3은 파란색으로 월드 상에 렌더링된 손 뼈대가 객체 뒤로 이동할 때 가려지고, 객체 앞에 있을 때는 객체를 가리는 모습을 보여주며, 월드와의 3D 추적이 이루어지고 있음을 증명한다. 덧붙여 손의 각 관절과 객체와의 충돌 작용 등 실시간 상호작용을 위한 추가적인 기능이 사용될 수 있다.

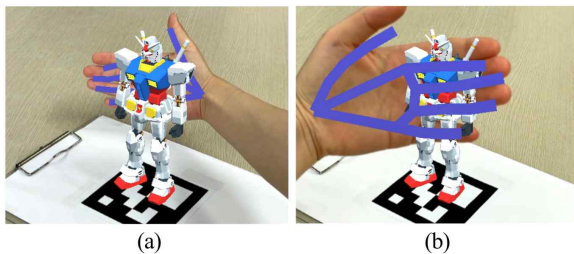


그림 3. 월드에 렌더링된 손과 객체:

(a) 객체에 가려진 손과 (b) 객체를 가리고 있는 손

Fig. 3. Rendered hand and objects in the world:

(a) hand occluded by the object and (b) hand covering the object

### III. 결론

본 연구에서는 상호작용 수단이 부재한 저비용 스마트폰 기반 MR 환경에서 손 추적을 통해 가상 객체와 상호작용 할 수 있는 MR 시스템을 개발하였다. 스마트폰을 활용한 저비용 MR 환경으로, 제안된 시스템은 교육 및 게임 등 다양한 분야에서

더 많은 사용자가 기존의 상호작용 수단의 부재로 제한되었던 다양한 MR 시나리오를 체험함으로써 더욱 향상된 경험을 갖도록 활용될 수 있다. 향후 연구를 통해 더욱 정교한 상호작용과 다중 사용자 경험을 지원하여 사용자의 만족도와 몰입감을 높이는 방향으로 개선될 것으로 기대된다.

### 참고 문헌

- [1] Anna Syberfeldt, Oscar Danielsson, and Patrik Gustavsson. 2017. Augmented reality smart glasses in the smart factory: Product evaluation guidelines and review of available products. *Ieee Access* 5 (2017), 9118-9130.
- [2] Kent T Espejon, Jaymer M Jayoma, and Elbert S Moyon. 2023. Low-Cost Mobile-Based Augmented Reality Viewer: Bridging the Gap Between Affordable AR Glasses and Smartphones. In *Proceedings of the 2023 12th International Conference on Networks, Communication and Computing*. 254-259.
- [3] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. 2020. Mediapipe hands: On-device real-time hand tracking. *arXiv preprint arXiv:2006.10214* (2020).
- [4] Eric Marchand, Hideaki Uchiyama, and Fabien Spindler. 2015. Pose estimation for augmented reality: a hands-on survey. *IEEE transactions on visualization and computer graphics* 22, 12 (2015), 2633-2651.
- [5] Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco José Madrid-Cuevas, and Manuel Jesús Marín-Jiménez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47, 6 (2014), 2280-2292.

# 인공지능을 활용한 분리수거 장치

김민호\*, 박원규\*, 신민혁\*, 백경호\*\*, 강병권\*\*\*

## Waste Separation Device using Artificial Intelligence

Min Ho Kim\*, Won-gyu Park\*, Min Hyuck Shin\*, Kyung-Ho Paik\*\* and Byeong-Gwon Kang\*\*\*

### 요 약

본 논문에서는 플라스틱 재활용률 향상을 목표로, YOLO v8 모델과 ESP32 기반의 이중 모터 제어 시스템을 결합한 분리수거 보조 장치를 개발하였다. 이중 모터 시스템을 통해 정밀한 분리 작업을 구현하며, YOLO v8 모델을 사용해 약 97.1%의 mAP50과 96.1%의 Precision을 기록하는 높은 객체 인식 성능을 달성하였다. 또한, Google Apps Script와 AppSheet를 이용해 실시간으로 데이터를 모니터링하고 시각화하여 사용자 패턴 분석을 가능하게 한다. 본 연구는 분리수거 보조 장치의 정확도와 효율성을 높여, 재활용 시스템 개선에 기여하고 환경 보호 및 자원 관리의 지속 가능성을 제고 할 수 있는 가능성을 제시한다.

### Abstract

This paper aims to improve the plastic recycling rate by developing a recycling assistance device that combines the YOLO v8 model with an ESP32-based dual-motor control system. The dual-motor system enables precise sorting operations, while the YOLO v8 model achieves high object recognition performance with approximately 97.1% mAP50 and 96.1% precision. Additionally, Google Apps Script and AppSheet are utilized to monitor and visualize data in real-time, enabling user pattern analysis. This study enhances the accuracy and efficiency of the recycling assistance device, contributing to the improvement of recycling systems and promoting the sustainability of environmental protection and resource management.

### Key words

Plastic Recycling Rate, YOLOv8, ESP 32, Real-Time Data Monitoring

## 1. 서 론

우리나라는 환경 보호와 지속 가능한 자원 관리

를 목표로 다양한 재활용 정책을 시행해왔지만, 여전히 플라스틱 재활용률은 낮은 수준에 머물러 있다. 2021년 기준, 가정과 사무실에서 배출된 1,192만

\* 순천향대학교 정보통신공학과 학부과정 kiminho9946@gmail.com

\*\* 순천향대학교 전자공학과 겸임교수, pkhtow@naver.com

\*\*\* 순천향대학교 정보통신공학과 교수 bgkang@sch.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 연구 결과로 수행되었음”(2021-0-01399)



톤의 플라스틱 쓰레기 중 약 69.3%는 재활용되지 못하고 폐기되었다[1]. 이는 주로 재활용 과정의 효율성 부족과 분리 작업의 어려움에서 기인한다. 본 연구에서는 이러한 문제를 해결하기 위해 기존 재활용 시스템을 개선하고자 하며, 이를 위해 하드웨어와 소프트웨어 시스템을 확장한다.

본 연구의 주요 기여는 다음과 같다. 첫째, 이중 모터 제어 시스템을 도입하여 정밀한 분리 작업을 수행함으로써 재활용 효율성을 향상시킨다. 둘째, 기존의 연구[를 확장하여 보다 다양한 재활용품을 인식할 수 있도록 개선하고, 데이터 수집 및 분석 기능을 강화하여 사용자 패턴을 파악할 수 있게 한다. 셋째, AppSheet를 통한 시각화를 통해 시스템 성과를 직관적으로 평가할 수 있는 환경을 제공한다. 또한, YOLO v8 모델과 다른 객체 탐지 모델 간의 성능을 비교하여 최적의 모델을 선택하고, 사용자 피드백을 반영하여 실시간 알림 기능을 포함한 인터페이스를 개선한다. 이러한 연구를 통해 지속 가능한 자원 관리와 환경 보호를 위한 재활용 시스템의 실질적 개선을 도모하고자 한다.

## II-1 YOLO v8 객체 인식

본 장치의 알고리즘은 [그림 1]에 제시된 형태로 구성되었으며, 훈련 데이터셋은 기존 연구에서 확보한 사진 데이터에 ‘라벨’ 칼럼을 추가하여 준비하였다. YOLO v8 모델을 이용해 학습을 진행하였고, 최적의 성능을 얻기 위해 다양한 하이퍼파라미터를 조정하면서 실험을 수행했다[2]. 총 7회에 걸쳐 훈련을 진행했으며[그림 2], 한 가지 파라미터를 고정한 상태에서 나머지 파라미터를 순차적으로 조정하여 가장 성능이 우수한 조합을 찾아냈다.

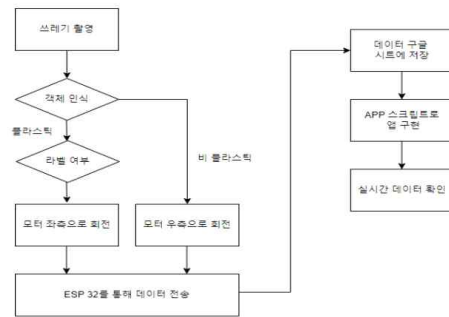


그림 1. 분리수거 보조 장치의 알고리즘  
Fig. 1. Recycling aid algorithm

## II. 본문

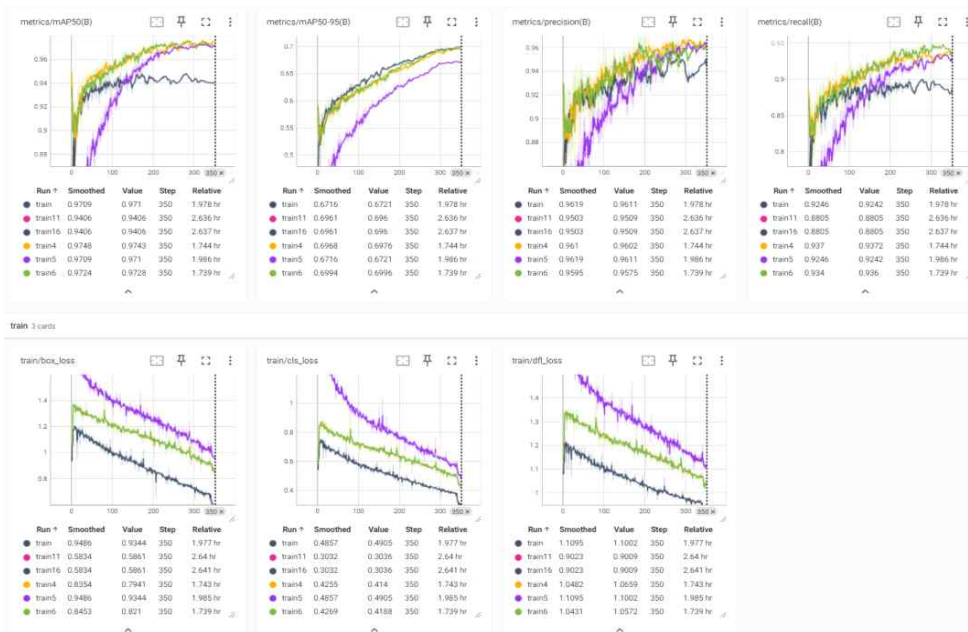


그림 2. 훈련 결과 플로팅  
Fig. 2. Train result floating

훈련 결과는 TensorBoard를 통해 분석했으며, 주요 성능 지표는 다음과 같다. 첫째, mAP50 및 mAP50-95 지표는 각각 97.1%와 67.2%로 나타나, 모델이 다양한 IoU 임계값에서도 높은 객체 감지 정확도를 유지함을 확인할 수 있었다. 둘째, Precision(정밀도)은 약 96.1%로, 모델이 예측한 객체 중 실제로 올바른 객체를 식별한 비율이 매우 높음을 보여준다. 셋째, Recall(재현율)은 92.4%로, 실제 객체 중 모델이 성공적으로 감지한 비율이 높아 놓친 객체가 적음을 의미한다. 마지막으로, box\_loss, cls\_loss, dfl\_loss 등의 Loss 지표는 낮은 수치를 보이며, 이는 모델의 예측 오류가 적다는 점을 시사한다[3].

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN}$$

		실제 정답	
		True	False
분류 결과	True	True Positive	False Positive
	False	False Negative	True Negative

그림 3. 훈련 결과 플로팅  
Fig. 3. Train result floating

정밀도와 재현율의 공식은 식(1)과 [그림 3]의 혼동 행렬을 통해 나타낼 수 있다.

### II-2 모터 제어

본 시스템의 [그림 1]의 알고리즘을 따르며, 객체 인식은 YOLOv8 모델이 장착된 카메라를 통해 수행 된다. 객체가 감지되면, ESP 32 마이크로컨트롤러가 해당 정보를 수신하여 분류 작업을 시작한다. 입력 값이 0일 경우, SG90 서보 모터가 0도에서 140도로 회전하여 객체의 라벨을 확인하는 위치로 조정된다. 이 초기 단계는 객체의 라벨 여부를 확인하여 플라스틱 재활용률을 높이기 위한 작업을 수행한다. 플라스틱 객체가 감지되면 입력 값이 1로

설정되며, SG90 모터가 위치를 조정된 뒤 360도 모터가 작동 하여 플라스틱 분류 작업을 진행한다. 비 플라스틱 객체가 감지되면 입력 값이 2로 설정되고, 같은 방식으로 SG90 모터가 조정된 후 360도 모터가 비 플라스틱 분류 작업을 수행한다. 이러한 모터 제어 로직은 객체 유형에 따라 적절한 동작을 실행함으로써 분리수거 장치의 정확도와 효율성을 높인다. 또한, ESP 32는 각 분류 작업 완료 시 Google Apps Script를 통해 클라우드에 데이터를 전송하여[4], 현재까지 분류된 플라스틱 및 비 플라스틱 수량을 실시간으로 모니터링 할 수 있도록 한다[그림 4]. 모터 제어 시스템과 ESP 32의 결합으로 장치는 정확하고 효율적으로 작동한다.

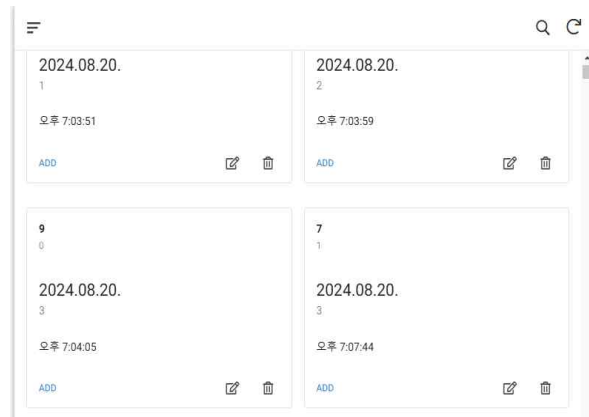


그림 4. 실시간 모니터링  
Fig. 4. Monitoring in real time

### III. 결론

본 연구는 YOLO v8 객체 인식 모델과 ESP32 기반 이중 모터 제어 시스템을 적용하여 보다 정밀하고 효율적인 분리수거 보조 장치를 구현하는 데 중점을 두었다. 이중 모터 시스템을 통해 다양한 각도에서 정밀하게 분류 작업을 수행함으로써 분류 효율성을 높였으며, YOLO v8 모델의 우수한 객체 인식 성능을 활용하여 재활용 품목의 인식 정확도를 향상시켰다. 또한, Google Apps Script를 이용한 클라우드 데이터 전송과 AppSheet 기반의 데이터 시각화 기능을 통해 실시간 모니터링과 성과 분석이 가능해져 재활용 과정에서 발생할 수 있는 문제를

사전에 해결할 수 있는 기반을 마련했다.

향후 연구에서는 객체 탐지 모델 성능 비교와 사용자 피드백을 반영한 인터페이스 확장 등의 개선 작업을 통해 분리수거 보조 장치의 성능을 지속적으로 강화할 예정이다.

## 참 고 문 헌

- [1] 김민호, et al. "YOLO v8 을 이용한 분리수거 보조 장치." 한국통신학회 학술대회논문집 (2024): 1205-1206..
- [2] W. Lan, J. Dang, Y. Wang and S. Wang, "Pedestrian Detection Based on YOLO Network Model," 2018 IEEE International Conference on Mechatronics and Automation (ICMA), Changchun, China, 2018, pp. 1547-1551, doi: 10.1109/ICMA.2018.8484698.
- [3] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934 (2020).
- [4] Pravalika, V., and Ch Rajendra Prasad. "Internet of things based home monitoring and device control using Esp32." International Journal of Recent Technology and Engineering 8.1S4 (2019): 58-62.

# O2O 상점에서의 상품 인식을 위한 구글 클라우드 플랫폼 기반 시스템 구현

안지섭\*, 김대민\*, 시종욱\*\*, 김성영\*\*\*

## Implementation of Google Cloud Platform-Based System for Product Recognition in O2O Stores

Jiseop Ahn\*, Daemin Kim\*, Jongwook Si\*\*, and Sungyoung Kim\*\*\*

### 요 약

본 논문에서는 O2O(Online to Offline) 상점에서의 상품 인식을 위한 YOLO-v8 모델을 GCP(Google Cloud Platform) 기반의 환경에서 시스템을 소개한다. 제안하는 시스템은 GCP의 VPC(Virtual Private Cloud)와 VM(Virtual Machine) 인스턴스를 사용해 모델을 배포하고, Instance group과 Autoscaling 기능을 통해 자원을 효율적으로 관리한다. 또한, 트래픽 분배를 위해 Load Balancer와 Instance group을 연동하고, Health check로 인스턴스 상태를 모니터링이 가능하다. 이를 통해 CPU 과부하를 방지하고 비용 절감과 최적의 시스템 성능을 확보할 수 있음을 보인다.

### Abstract

This paper introduces a system for product recognition in O2O (Online to Offline) store environment utilizing the YOLO-v8 model deployed on Google Cloud Platform (GCP). The proposed system employs GCP's Virtual Private Cloud (VPC) and Virtual Machine (VM) instances for model deployment, and efficiently manages resources using Instance Groups and Autoscaling capabilities. Additionally, the system integrates Load Balancers with Instance Groups for traffic distribution and uses Health Checks to monitor instance status. This approach prevents CPU overload, reduces costs, and ensures optimal system performance.

### Key words

O2O(Offline to Online), Object Recognition, Cloud System, GCP(Google Cloud Platform), VPC(Virtual Private Cloud)

### I. 서 론

O2O(Online to Offline) 서비스의 부상과 함께,

O2O 상점 내에서 인공지능 기술의 활용에 대한 관심이 증가하고 있다[1]. 특히, 객체 인식 기술은 O2O 상점에서 상품을 인식하고 관리하는 데 중요

\* 국립금오공과대학교 컴퓨터공학과

\*\* 국립금오공과대학교 컴퓨터·AI융합공학과

\*\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자)

한 역할을 한다. 이를 위해 YOLO-v8 모델을 학습시켜 이미지에서 상품의 영역을 검출하는 기술[2]을 클라우드 시스템에 구축하여 확장성과 접근성을 활용함으로써 시스템 운영의 유연성과 자원 관리의 효율성을 높일 수 있다. 이러한 방식을 통해 O2O 상점에서의 객체 인식 서비스를 클라우드 기반으로 구현한 내용에 대하여 소개한다.

## II. 클라우드 시스템 설계 및 구현

본 논문에서 제안하는 O2O 상점에서의 객체 인식을 위한 클라우드 기반 시스템은 Google Cloud Platform(GCP)을 기반으로 설계된다. 제안된 시스템의 설계에서는 보안성과 관리 편의성을 고려하여 Virtual Private Cloud(VPC)를 생성하고, 이 VPC를 네트워크로 사용하여 Virtual Machine(VM) 인스턴스를 구축한 후, YOLO-v8 모델을 배포하여 객체 인식을 수행한다.

특히 다수의 사용자가 동시에 시스템을 사용할 수 있도록 VM 인스턴스를 기반으로 한 Instance

Group을 구성하며, Autoscaling 기능을 적용하여 자원을 효율적으로 관리할 수 있도록 설계한다. CPU 사용률이 60%를 초과하면 VM 인스턴스를 자동으로 추가하고, 사용량이 줄어들면 인스턴스를 제거하여 필요한 자원만을 사용하도록 함으로써 비용 절감과 시스템의 효율적인 운영을 도모한다. 이를 통해 CPU 과부하로 인한 서비스 중단을 방지하며, 저비용으로도 최적의 시스템 성능을 유지할 수 있다.

또한, Instance Group 내 VM 인스턴스가 두 개 이상인 경우, 트래픽을 효율적으로 분배하기 위해 Instance Group과 Global External Application Load Balancer를 연동한다. 이와 함께 Health Check 기능을 설정하여 Instance Group 내의 모든 인스턴스가 정상적으로 작동할 때만 트래픽이 해당 인스턴스로 분배되도록 한다. 이러한 구조를 통해 시스템의 가용성과 안정성을 보장하며, 최적의 성능을 제공할 수 있도록 한다. 그림 1은 GCP의 Global External Application Load Balancer와 Instance Group의 연동 구조를 시각적으로 나타낸다.

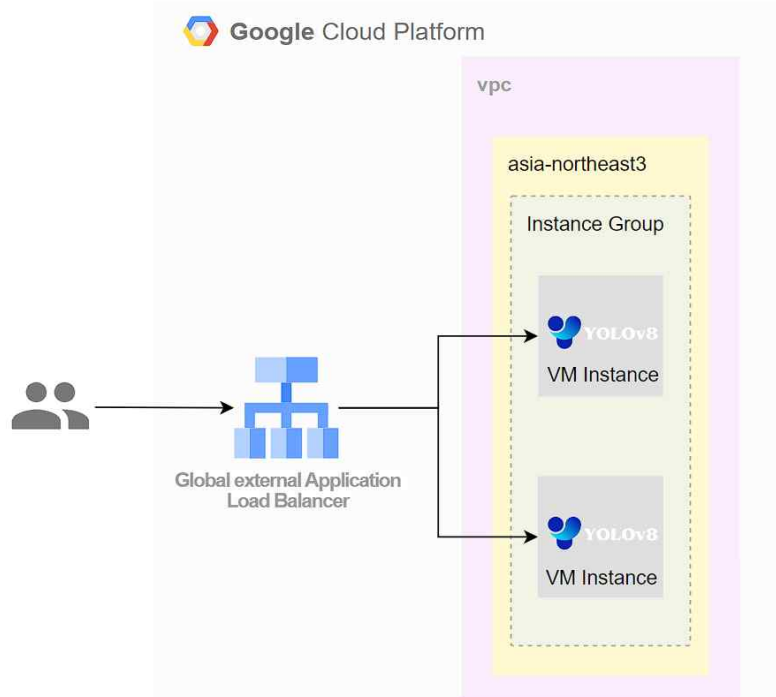


그림 1. Global external Application Load Balancer와 Instance group의 구조  
Fig. 1. Structure of Global external Application Load Balancer and Instance group

### III. 결 론

본 논문에서는 O2O 상점에서 객체 인식을 위해 GCP 기반 시스템을 설계하고 구축하였다. 클라우드 환경에서 서비스를 제공함으로써 높은 접근성과 효율적인 트래픽 관리가 가능해졌으며, 이를 통해 O2O 상점의 운영 효율성을 크게 개선할 수 있었다. 이러한 클라우드 기반 시스템 설계를 통해 객체 인식 서비스의 유연성과 확장성을 확보하고, 클라우드 자원의 효율적인 사용을 통해 비용 절감과 최적의 시스템 성능을 달성하였다.

### 참 고 문 헌

- [1] J. Si, S. Lee, and S. Kim, "Recognition of Korean Menu for Online to Offline Stores: VGG-ResNet Fusion Model with Attention Mechanism". Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol. 17, No. 4, pp. 190-197, Aug. 2024.
- [2] J. Si, D. Kim and S. Kim, "Automation of Online to Offline Stores: Extremely Small Yolov8 and Feature-Based Product Recognition", Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol 17, No. 3, pp.121-129, Jun. 2024.

# 임계값 필터링을 활용한 sA2C-T 강화학습 방법

장승균\*, 소훈\*, 정찬우\*, 서정훈\*, 홍연우\*, 김성빈\*, 신은성\*, 정현준\*\*†

## sA2C-T Reinforcement Learning Method Using Threshold Filtering

Seungkyun Jang\*, Hoon So\*, Chanwoo Jeong\*, Jeonghoon Seo\*, Yeonwoo Hong\*, Seongbin Kim\*, Eunseong Shin\*, Hyunjun Jung\*\*

### 요약

컴퓨터와 인공지능 기술이 발전함에 따라 복잡한 학습 문제를 해결하기 위한 방안으로 강화학습이 개발되었다. 강화학습은 에이전트가 다양한 행동을 수행하며 목표를 달성하기 위한 행동 패턴을 학습하도록 하는 방법이다. 최근에는 다중 에이전트를 학습시켜 각 에이전트의 학습 성능에 따른 중요성 가중치를 산출하여 가중치만큼 공유 모델에 반영하는 연구가 진행되고 있다. 그러나 이 과정에서 지나치게 안 좋은 학습 환경이어도 공유 모델에 반영된다는 문제점이 있다. 이 연구에서는 이 중요성 가중치에 임계값 필터링 방법을 사용하여 성능이 지나치게 떨어지는 학습 환경을 제외하는 방법을 제안하였으며, Unity 가상환경을 통해 다중 협지 환경에서 특정 목표 지점까지 이동하는 자율주행 실험을 통해 이전 sA2C 알고리즘과 제안한 sA2C-T 알고리즘의 성능 비교를 진행하였다.

### Abstract

Reinforcement learning has been developed as a way to solve complex learning problems as computer and artificial intelligence technologies develop. Reinforcement learning is a method that allows agents to perform various actions and learn behavioral patterns to achieve goals. Research has recently been conducted to train multiple agents to calculate the importance weight according to each agent's learning performance and reflect it in the shared model as much as the weight. However, in this process, there is a problem that even if the learning environment is too bad, it is reflected in the shared model. This study proposed a method of excluding learning environments with excessively poor performance by using a threshold filtering method for this importance weight, and the performance comparison between the previous sA2C algorithm and the proposed sA2C-T algorithm was conducted through an autonomous driving experiment that moves from a multi-tidal environment to a specific target point through the Unity virtual environment.

### Key words

reinforcement learning, unity, multi-agent, importance weighted, threshold filtering

---

\* 국립군산대학교 소프트웨어학부 학사과정, 1901604@kunsan.ac.kr, hoonso20@kunsan.ac.kr, didam00@naver.com, jcw0690@gmail.com, mondmoonlight@naver.com, napo987600@gmail.com, plmjuiio123@naver.com

\*\*국립군산대학교 소프트웨어학부 교수, junghj85@kunsan.ac.kr

†교신저자

※이 연구는 정부(과학기술정보통신부)의 재원으로 한국 연구재단의 지원을 받아 수행되고 있습니다.(No. NRF-2022R1G1A1008493).

※본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음(2023-0-00065)

## I. 서 론

컴퓨터의 성능이 높아짐에 따라 인공지능 기술이 다양한 분야에 적용되고 있다[1]. 하지만 현실에서 인공지능을 기술을 적용하기 위해서는 복잡한 환경을 컴퓨터가 이해해야 하며, 사용자가 원하는 답을 얻기 위해서는 다양한 분류나 정확한 정답을 찾는 연산이 진행되어야 한다. 이 외에도 자율주행이나 로봇 제어, 게임같이 특정 상황에서 컴퓨터가 상황에 맞는 행동을 선택하는 경우는 다양한 변수와 문제점이 발생하기 때문에 지도학습이나 비지도 학습으로 컴퓨터를 학습시키기 어렵다. 이를 보완하기 위해서 강화학습(Reinforcement Learning)이라는 인공지능 기술이 도입되었다. 강화학습은 머신러닝 기술 중 하나로 컴퓨터가 직접 환경에서 상호작용하면서 특정 목표에 도달하기 위한 행동 패턴을 찾고 그에 따른 보상을 얻어서 이 중 가장 최대의 보상 값을 얻는 최적의 행동 패턴을 찾는 학습 방법이다. 구글의 딥마인드에서는 강화학습 연구를 진행하여 2015년에 Deep Q-Network(DQN) 알고리즘을 개발하였으며, 이는 2016년에 '알파고'라는 바둑을 두는 게임 인공지능이 개발되어 강화학습의 기술을 전세계적으로 알리게 되었다[2]. DQN 알고리즘이 나온 뒤, 강화학습 기술은 하나의 인공지능만 학습시키는 것이 아닌 다중 환경에서 여러 개의 에이전트를 학습하는 멀티 에이전트 강화학습 방법이 등장하였다. 이 기술은 같은 목표를 가지는 서로 다른 에이전트들이 독립적인 환경과 행동을 통해 최적의 행동 패턴을 찾는 학습이 진행된다. 이후 학습한 데이터를 하나의 공유 모델에 전달하여 서로 다른 학습을 하나의 모델에 통합시킬 수 있다. 이 방법을 사용한 대표적인 알고리즘으로 A3C 알고리즘이나 PPO 알고리즘 등이 있다[3][4]. 하지만 Local Network에서 학습한 데이터 중 목표가 떨어지는 학습이 진행되고 있으면 공유 모델에 데이터를 전송하게 될 때 부적절한 데이터의 값으로 인해 공유 모델의 학습 성능이 떨어지게 된다. 이를 보완한 방법으로 sA2C 알고리즘이나 IMPALA 알고리즘이 연구되었다. 이 알고리즘은 공유 모델에 데이터를 전달하기 전에 각 Local Network의 학습 성능을 확인하고 이를 중

요성 가중치로 산출하여 전달되는 파라미터의 값을 조절하는 역할을 해준다[5]. 이 논문에서는 중요성 가중치의 값을 임계값 필터링 방법을 사용하여 불필요한 중요성 가중치를 배제하는 sA2C-T 강화학습 방법을 제안한다. 이전 알고리즘에서는 중요성 가중치를 적용하여도 공유 모델에 적용되기 때문에, 학습의 안정성을 떨어트리는 경우가 발생한다. 하지만 sA2C-T 알고리즘에서는 공유 모델에 적용할 수 있는 특정 임계값을 설정하여 임계값 이하의 가중치를 필터링하여 학습 성능이 떨어지는 환경은 공유 모델에 적용하지 않게 하여 학습의 안정성을 높여주는 역할을 해줄 수 있다. 이 방법을 바탕으로 Unity 가상환경을 통해 자율주행 로봇이 호출한 사용자에게 이동하는 실험을 진행하였으며, 이전 알고리즘이 sA2C 알고리즘과 sA2C-T의 성능 비교 평가를 진행하였다.

이 논문에서는 2장에서는 sA2C 알고리즘에 대한 방법과 임계값을 적용한 강화학습 알고리즘 방법을 소개한다. 3장에서는 sA2C-T 알고리즘을 설명하고 4장에서는 실험한 환경에 대한 설명과 sA2C 알고리즘과 sA2C-T 알고리즘에 대한 성능 비교한 내용을 설명하고 5장에서 결론 및 향후 과제로 마무리한다.

## II. 관련 연구

### 2.1 IMPALA 알고리즘

IMPALA 알고리즘은 중요성 가중치를 사용하여 공유 모델에 학습시키는 멀티에이전트 강화학습 방법이다[6]. 이 알고리즘은 각 에이전트가 환경에서 얻은 데이터를 공유 모델에 전달하고 학습할 때 에이전트에서 얻은 값과 이전 공유 모델에서 학습한 값을 비교하여 중요성 가중치 산출해 공유 모델의 학습 성능을 높여주는 연구 결과를 보여주었다. 하지만 학습 환경이 공유 모델의 네트워크보다 매우 다른 형태로 학습이 진행되어도 중요성 가중치는 0이 되지 않는다면 공유 모델에 악영향을 미칠 수 있으므로 공유 모델의 학습 성능이 떨어지게 되며, 이로 인한 학습의 안정성이 떨어지는 한계점이 발



생하게 된다.

## 2.2 임계값을 활용한 멀티에이전트 강화학습

임계값을 활용하여 멀티에이전트를 학습하는 연구가 진행되었다[7]. 이 연구에서는 각 에이전트에서 학습한 데이터를 공유 모델에 전달할 때 중요성 가중치 중 지나치게 낮은 값을 제외하고 높은 가중치 값을 추가로 데이터를 반영하여 학습 성능을 높이는 실험을 진행하였다. 이 실험을 통해 모델의 학습 안정성을 높이고 분산의 폭을 감소하여 공유 모델의 학습 성능이 높아졌다는 실험 결과가 있었다.

### III. 임계값 필터링을 활용한 sA2C-T 알고리즘 학습 방법

sA2C-T 알고리즘은 임계값 기반 중요성 가중치 필터링을 핵심으로 하여, Global Network가 성능이 높은 환경의 데이터에 집중할 수 있도록 설계되었다. 각 에이전트는 자신이 속한 환경과 상호작용하며 상태, 관찰값, 보상 등을 수집하고, 이를 Local Network에 입력하여 정책 파라미터와 가치 함수  $V(s)$  값을 산출한다. 이때 각 환경에서 구한  $V(s)$

값은 환경의 학습 성능을 반영하는 지표로 사용된다.

모든 환경에서 얻은  $V(s)$  값들은 리스트로 수집된 후 정규화 과정을 통해 중요성 가중치로 변환된다. 학습 성능이 높은 환경일수록 큰 가중치를 가지게 되며, 성능이 낮은 환경은 작은 가중치를 갖는다. 하지만 성능이 지나치게 낮은 환경의 경우, 낮은 가중치만으로는 학습에 미치는 부정적 영향을 완전히 방지하기 어렵다. 이때 사용자가 설정한 임계값을 적용한 필터링 방법을 사용하여 중요성 가중치가 임계값 이상인 환경의 데이터만 Global Network에 반영되게 하여 성능이 지나치게 낮은 환경이 전체 학습에 미치는 영향을 최소화한다. 이는 중요성 가중치가 임계값보다 낮은 환경의 경우 중요성 가중치의 값을 0으로 변환하여 학습 성능이 떨어지는 환경의 파라미터 값을 배제하게 한다. 이를 통해 Global Network에서는 학습 성능이 높은 환경의 데이터에 집중할 수 있으며, 이는 Global Network의 학습 불안정성을 방지할 수 있다.

이 방식을 통해, 임계값 필터링을 활용한 sA2C 알고리즘은 에이전트들이 다양한 환경에서 최적의 행동 정책을 방해받지 않고 빠른 시간 내에 학습 목표를 달성하게 도와주고 전체 학습의 효율성과 안정성을 동시에 보장할 수 있다.

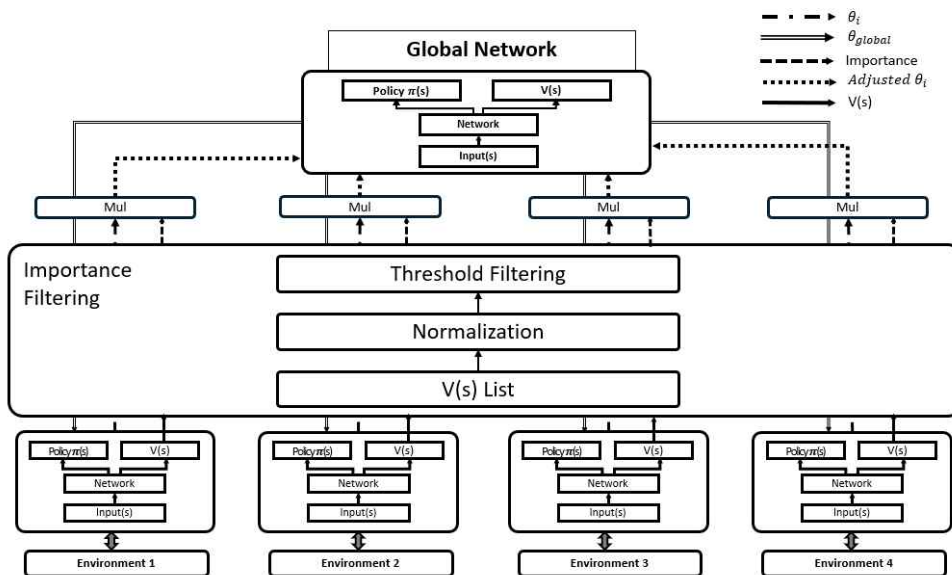


그림 1. 임계값 필터링을 활용한 sA2C-T 알고리즘 아키텍처  
Fig. 1. sA2C-T algorithm architecture using threshold filtering

## IV. 실험 및 결과

### 4.1 실험 환경 구축

제안한 알고리즘을 테스트 하기 위해서 험지 환경에서의 자율주행 학습으로 제안한 알고리즘 실험을 진행하였다. 험지 환경은 기존 평지 환경에 비해 굴곡의 형태를 가지고 있어 로봇의 기울기나 중력으로 인한 물리적인 영향으로 로봇에게는 복잡한 환경을 가지게 된다. 이는 로봇이 이동하는데 다양한 상황들이 나타나며, 사용자까지 이동하는데 수많은 행동 패턴이 나타나게 된다. 이를 학습하기 위해 로봇은 자신의 위치와 사용자의 위치, 그리고 자신과 사용자의 거리를 상태값으로 받으며 목표 지점까지 가까워지도록 앞뒤나 양옆으로 이동하는 행동을 선택한다. 그리고 올바른 학습을 진행하기 위해, 로봇이 사용자에게 가까워질 때마다 가까워진 거리 \* 0.01의 값을 주고 목표 지점에 도달하였을 경우, +10점이라는 양의 보상 함수를 설계하였다. 이 외에도 거리가 이전보다 멀어졌을 경우, 멀어진 거리 \* -0.01의 보상 값을 주거나 영역 밖으로 나가게 되는 경우 -10의 보상을 줘서 로봇이 빨리 사용자에게 도달할 수 있게 음의 보상 함수도 설계하였다.

설계한 환경을 다중 환경으로 구성하기 위해 이 논문에서는 총 4개의 똑같은 환경을 구성하였으며, 구성된 환경은 그림 2과 같다.

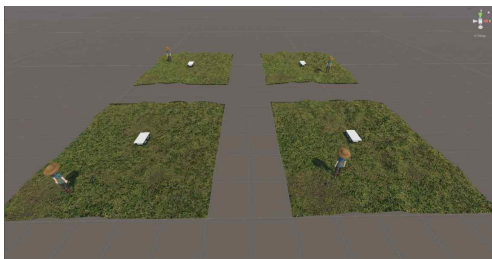


그림 729 다중 험지 환경 구축  
Fig. 2. Building multi-hazard environments

### 4.2 sA2C 알고리즘과 제안 알고리즘 비교 실험

이 논문에서 설계한 험지 환경을 sA2C 알고리즘

과 제안한 알고리즘을 동일한 하이퍼파라미터를 사용하였으며 각 알고리즘에 대한 Loss 값과 평균 보상 값을 비교하였다.

표 1. 학습에 사용된 파라미터

Table 1. Parameters used for training

Parameter	Value
Discount Factor	0.99
Episode	1500
Learning Rate	0.0001
Maximum steps per Episode	50000
Actor size	4
Importance threshold	0.2

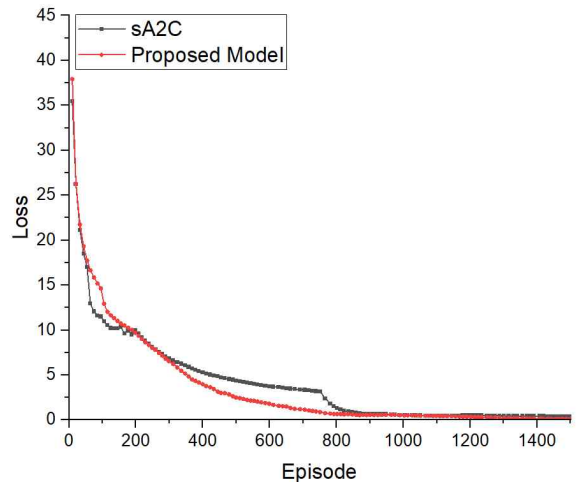


그림 730. sA2C 모델과 sA2C-T 모델의 Loss 그래프  
Fig. 3. Loss graph of sA2C model and sA2C-T model

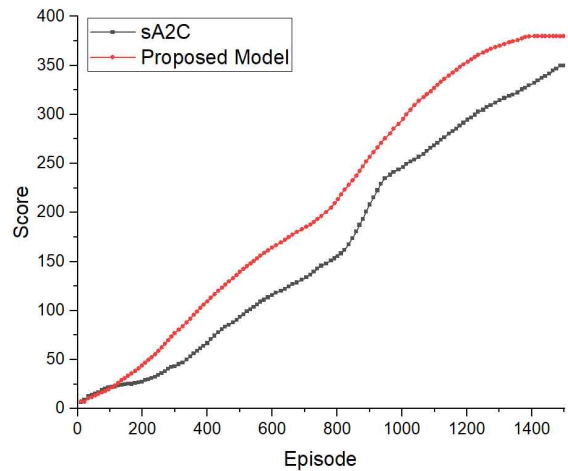


그림 731 A2C 모델과 sA2C-T 모델의 보상 그래프  
Fig. 4. Reward graph for sA2C model and sA2C-T model

먼저 Loss 값은 그림 3같이 결과를 보였다. 기존 sA2C 알고리즘 같은 경우 100번째 Episode나 800번째 Episode에서 학습의 불안정성으로 인해 Loss 값에 대한 변곡점이 존재하여 모델 학습의 안정성이 떨어지게 되었다. 반면, sA2C-T 모델에서는 학습 성능이 떨어지는 환경에서의 파라미터를 공유 모델에 배제함으로써 학습의 안정성이 증가하였으며 이는 이전 모델보다 손실 그래프의 변곡점이 자주 나타나지 않는 형태로 그려지게 되었다. 또한 그림 4의 보상 그래프를 확인하면 이전 sA2C 알고리즘에서 학습의 불안정성으로 생긴 변곡점으로 인해 공유 모델의 학습 성능이 떨어지게 되어 최종 Episode 수치인 1500번째 Episode에서 350의 보상 값으로 수렴하였지만, 임곗값 필터링을 활용한 sA2C-T 알고리즘에서는 학습 안정성을 통해 공유 모델의 학습이 떨어지는 것이 이전 모델보다 적었으며 이는 380에 가까운 보상값이 수렴하였고 이는 이전 모델보다 8.57% 높은 값으로 수렴하는 것을 확인할 수 있다.

## V 결론

이 논문에서는 특정 임곗값을 포함한 동기식 액터-크리틱 알고리즘으로 다중 환경에서 각 에이전트의 학습 성능을 나타내는  $V(s)$  값을 다른 환경과 비교하여 중요성 가중치를 구하고 만약 가중치가 특정 임곗값보다 낮은 값을 형성할 경우, 공유 모델에 적용하지 않으며, 이 외에는 중요성 가중치만큼 공유 모델에 데이터를 전달하는 멀티에이전트 학습을 진행하였다. 이러한 학습 방법은 이전 sA2C 알고리즘과 비교하였을 때, 낮은 학습 성능을 보이는 학습 환경이여도 공유 모델에 데이터를 전달하는 이전 알고리즘과 달리 제안한 알고리즘은 성능이 낮은 환경을 배제함으로써, 공유 모델 학습의 안정성을 높이는 것을 가정으로 이 연구가 진행되었다. 이를 증명하기 위해 복잡한 환경을 가진 험지 환경을 다중 환경으로 구성하여 기존 sA2C 알고리즘과 제안한 알고리즘을 성능 비교를 진행하였으며 제안한 알고리즘이 이전 알고리즘보다 학습 성능과 안정성을 높이는 것을 확인하였다.

추후에, 임곗값만 설정하는 것뿐만 아니라 임곗값을 더욱 세부적인 공식을 통해 동적인 임곗값을 설정하는 방식을 진행하거나 이전 실험인 IR 알고리즘 방식처럼 임곗값으로 구별한 환경을 다시 한번 샘플링하는 방법을 통해 공유 모델의 학습 성능을 높이는 연구를 진행할 것이다.

## 참 고 문 헌

- [1] GARG and Pradeep Kumar. Overview of artificial intelligence. Artificial intelligence. Chapman and Hall/CRC, 2021. pp. 3-18.
- [2] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. Playing atari with deep reinforcement learning, arXiv preprint arXiv:1312.5602, Dec, 2013
- [3] Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., ... and Kavukcuoglu, K. Asynchronous methods for deep reinforcement learning. In International conference on machine learning. Vol. 48, pp. 1928-1937, Jun, 2016.
- [4] Schulman J, Wolski F, Dhariwal P, Radford A, and Klimov O. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347. Jul, 2017.
- [5] Kapturovski, S., Ostrovski, G., Quan, J., Munos, R., and Dabney, W. Recurrent experience replay in distributed reinforcement learning. In International conference on learning representations. ICLR. Dec. 2018.
- [6] Espeholt, L., Soyer, H., Munos, R., Simonyan, K., Mnih, V., Ward, T., ... and Kavukcuoglu, K. . Impala: Scalable distributed deep-rl with importance weighted actor-learner architectures. In International conference on machine learning. Vol 80 pp. 1407-1416, Jun, 2018
- [7] Schlegel and Matthew. Importance resampling for off-policy prediction., Advances in Neural Information Processing Systems, 2019, Nov.

# 도플러 레이더와 손동작 인식을 이용한 YOLO 기반 디스플레이 제어 시스템

남현수\*, 최인식\*

## YOLO-based Display Control System using Doppler Radar and Hand Gesture Recognition

Hyeon-su Nam\*, and In-Sik Choi\*

### 요 약

본 논문에서는 24GHz 도플러 레이더와 YOLO(You Only Look Once) 네트워크를 활용하여 HGR(Hand Gesture Recognition) 기반의 무선 디스플레이 제어 시스템을 설계하였다. 24GHz 도플러 레이더를 통해 서로 다른 6개의 손동작을 2초 주기로 측정하고, 측정된 데이터에 대해 STFT(Short-Time Fourier Transform)를 적용하여 시간-도플러 스펙트로그램을 영상을 획득하였다. 획득한 스펙트로그램 영상을 YOLOv5 기반 네트워크로 학습을 시키고, 학습된 YOLO 분류기를 STM32 블루투스 모듈 시스템을 결합하여 사생활 침해 없이 가정에서 사용할 수 있는 디스플레이 제어 시스템을 설계하였다.

### Abstract

In this paper, a wireless display control system based on Hand Gesture Recognition (HGR) was designed using a 24GHz Doppler radar and the You Only Look Once (YOLO) network. Each of the six hand gestures was measured over a two-second interval using the 24GHz Doppler radar, and the time-Doppler spectrogram was obtained by applying the Short-Time Fourier Transform (STFT) to the measured data. The obtained spectrogram images were used to train a YOLOv5-based network, and the trained YOLO classifier was integrated with an STM32 Bluetooth module system to design a display control system that can be used at home to protect privacy

### Key words

YOLO, Micro-Doppler, 24GHz Doppler radar, Spectrogram, Hand gesture recognition

## 1. 서 론

AI 기술의 발전으로 가정에서 비접촉 사용자 인터페이스의 필요성이 증가하고 있다. 카메라를 사용하는 방법은 사생활 침해 우려가 있으며, 음성 인식

은 환경에 따라 제한적일 수 있다. 이로 인해 레이더 기반 HGR 기술이 주목받고 있다. 도플러 레이더는 저렴하고 구현이 간단하며, 속도 측정에 있어 장점을 가진다[1]. 따라서 본 논문은 24GHz 도플러 레이더를 활용해 실시간으로 시간-도플러 맵을 추

---

\* 한남대학교 전기전자공학과(Dept of Electrical & Electronic Engineering Hunnam University)  
Corresponding author : In-Sik Choi(necog@hnu.kr)

출하고, 빠른 분류 성능을 가지는 YOLOv5 네트워크를 통해 6가지 손동작을 인식하여 디스플레이 제어를 수행하는 시스템을 제안한다.

## II. 본론 및 결과

실험에 사용된 레이더 모듈은 24GHz의 중간 주파수를 가지는 IPS-937 모델을 사용했으며, 손이 레이더로 접근할 때 도플러 주파수는 양의 값을 가져 스펙트로그램에서 positive의 주파수 성분을 가진다. 반대로, 손이 레이더에서 멀어질 때 도플러 주파수는 음의 값을 가지며, 이는 스펙트로그램에서 negative의 주파수 성분을 가진다. 이러한 도플러 주파수의 특성[2]을 이용하여 gesture 1부터 6까지 각각 손목 접기, 손목 펴기, 손바닥 밀고 당기기, 손가락 순차적으로 접기, 손목 돌리기, 손가락 2개 순차적으로 펴기 등을 사용하였다.

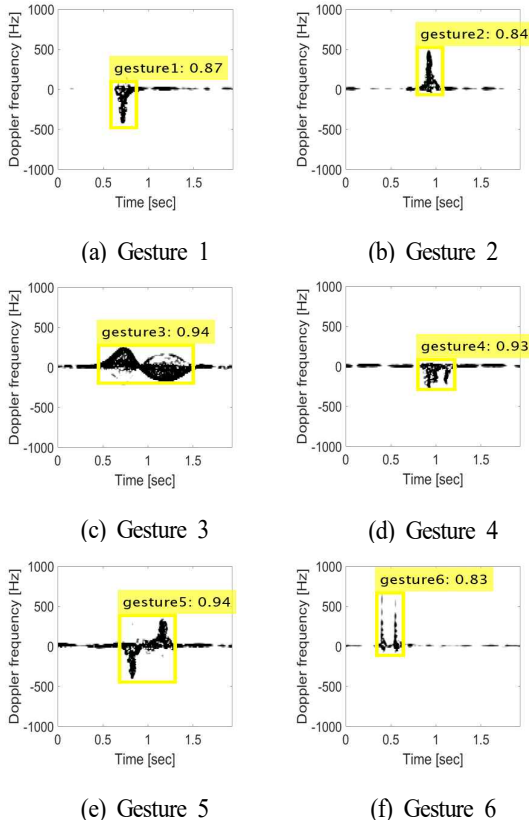


그림 1. Bounding box가 적용된 손동작 스펙트로그램  
Fig. 1. Hand gestures spectrogram with bounding box

수집된 데이터는 1.5m에서 2m 범위에서 측정되었으며, I/Q 데이터에 대해 STFT를 적용하여 스펙트로그램 영상을 생성하였다. 생성된 스펙트로그램 영상을 annotation를 적용하고, YOLOv5 네트워크를 학습하였다. 그림 1은 각 손동작에 bounding box가 적용된 결과이다.

분류기를 학습시키기 위해 총 5명의 피실험자를 대상으로 6개의 손동작 클래스를 100개씩 수집하여 총 3000개의 데이터를 수집하였다. 측정은 2초의 주기로 진행되었으며 7:2:1 비율로 training data, validation data, test data로 나누어 학습을 진행했으며, 그 결과 표 1과 같이 구분 성능을 확인하였다.

표 1. 손동작 구분에 사용된 시뮬레이션 환경  
Table. 1. Simulation spec used for the hand gesture recognition

# of Class	6
# of Images	300
Precision	0.996

학습된 분류기를 이용하여 각 손동작을 인식하는 시스템과 STM32의 블루투스 모듈을 결합하여 HGR 디스플레이 제어 시스템을 설계하였다.

## 참 고 문 헌

- [1] Seung-Tak Ra, Tae-Yong Doh, Hee-Yeol Lee, Young-Ji Lee, Jeong-Yoon Kim,\*Hong-Jik Kim, Seung-Jin Oh, Gi-Beom Kim, Hye-Ji Park, \*Dong-Geun Lee, "Deep learning based left and right Gesture recognition using reflector and CW radar", Summer Annual Conference of IEIE, pp. 1080-1083, Jun. 2022.
- [2] Chen, V. C., Fayin Li, Shen-Shyang Ho, & Wechsler, H "Micro-Doppler Effect in Radar: Phenomenon, Model, and Simulation Study," IEEE Transactions on Aerospace and Electronic Systems Vol. 42, No. 1, pp 2-21, Jan. 2006.

# 단일 이미지 감정 분석을 위한 이미지 캡셔닝 기반 양방향 크로스 어텐션 멀티모달 모델

김주영\*, 김건우\*\*

## Image Captioning-based Bidirectional Cross-attention Multimodal Model for Single Image Sentiment Analysis

Ju-Young Kim\* and Gun-Woo Kim\*\*

### 요약

감정은 인간의 의사 표현의 일부로, 추상적이고 복잡적으로 나타나며 주관적이고 모호하게 표현되어 의도와 다르게 해석되기도 한다. 특히 단일 이미지만으로는 감정을 이해하기 위한 상황 맥락 정보가 부족하여 정확한 감정 분석에 한계가 있다. 본 논문에서는 단일 이미지 감정 분석을 위한 이미지 캡셔닝 기반 양방향 크로스 어텐션 멀티모달(Captioned-Multimodal Bi-Cross-attention, CMBC) 모델을 제안한다. CMBC 모델은 이미지 캡셔닝을 통해 이미지 내 객체와 배경의 상호 관계를 설명하는 텍스트를 생성하고, 이를 이미지와 결합하는 양방향 크로스 어텐션 메커니즘 기반의 감정 분류를 수행한다. 실험 결과, 제안된 CMBC 모델은 F1 Score 0.787을 달성하여 기존 단일모달 대비 최대 8.4%의 성능 향상을 보여주었다. 이를 통해 이미지 캡셔닝과 양방향 크로스 어텐션의 결합이 효과적인 감정 분류 방법임을 입증하였다.

### Abstract

Emotion, as a part of human expression, manifests in complex and abstract ways, often leading to varied interpretations due to its subjective and ambiguous nature. Single images particularly lack contextual information necessary for accurate emotion analysis. This paper proposes the Captioned-Multimodal Bi-Cross-attention (CMBC) model for single-image emotion analysis. The CMBC model generates descriptive text explaining object-background relationships through image captioning and performs emotion classification using a bidirectional cross-attention mechanism that combines this text with image features. Experimental results show that the CMBC model achieved an F1 Score of 0.787, demonstrating up to 8.4% improvement over existing unimodal approaches. This proves the effectiveness of combining image captioning with bidirectional cross-attention for emotion classification.

### Key words

Sentiment Analysis, Multi Modal, Bidirectional Cross-Attention, Image Captioning

---

\* 경상국립대학교 컴퓨터공학부, wndudwkd003@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, gunwoo.kim@gnu.ac.kr

※ 본 논문은 2023년도 정부(교육부)의 지원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2021R1G1A1006381)

## I. 서 론

감정(Sentiment)이란, 인간이 의사 표현을 하기 위한 중요한 심리적 반응으로 표정과 행동을 통해 명확하게 표출되기도 하지만, 때로는 추상적이고 복합적으로 나타나기도 한다. 이러한 감정은 현대 사회에서 소셜 미디어, 마케팅, 예술 등 다양한 분야에서 각기 다른 방법으로 전달되고 있다[1].

인간의 감정을 이해하는 것은 사회적 흐름을 이해하고, 시장의 요구에 신속히 대응하는 데 중요한 요소이다. 이를 위해 감정을 분류 및 예측하는 감정 분석(Sentiment Analysis) 분야가 발전해 왔으며, 주로 자연어처리(NLP)의 일환으로 텍스트 데이터를 통해 사람의 감정을 분석하는 기술이 발달해왔다. 그러나, 이미지 기반의 SNS가 급속도로 성장하면서 콘텐츠 형태가 텍스트에서 이미지로 변화하는 추세이며, 이미지를 통한 감정 전달이 더욱 중요해지고 있다[2]. 이미지 감정 분석은 추가적인 상황 정보 없는 경우 모델의 한계가 존재한다. 기존의 합성곱 신경망(CNN)과 같은 모델은 이미지의 표면적 특징을 추출하는 성능은 뛰어나지만, 감정의 심리적 요소나 맥락적 의미를 파악하는 데 제한이 있다.

본 논문에서는 단일 이미지의 감정 분석의 한계를 극복하기 위한 멀티모달 감정 분석 모델을 제안한다. 본 모델은 이미지 캡셔닝(Image Captioning) 모델을 통해 이미지 내의 객체와 배경을 분석하여 이미지에서 나타나는 상황을 설명하는 텍스트 정보를 생성하고, 이미지와 정보와 어텐션 메커니즘을 수행하는 Bidirectional Cross-attention 형태로 서로 다른 형태의 정보를 결합하여 시각적 정보와 상황 맥락적 요소를 고려한 감정 분석을 가능하게 한다. 이러한 구조를 통해 모델이 단일 이미지에서 효과적인 감정 분류를 수행할 수 있다.

## II. 관련 연구

기존의 감정 분석 분야에서는 텍스트, 이미지, 음성 등 단일 유형의 정보만 사용하는 단일모달 학습과 다른 유형의 데이터를 결합하는 멀티모달 학습의 형태로 연구가 진행되고 있다[3]. 그러나, 감정의 복잡한 특징으로부터 정확한 감정 예측을 하기 위

해서는 멀티모달 형태의 학습 방법이 유리할 수 있다. 멀티모달 학습을 통한 감정 분석 모델에는 Multi-Cross-Attentive Model(MCAM)과 Image Captioning Joint Attention Mechanism(ICAM) 등이 제안되었으며 텍스트 정보를 기반으로 이미지 정보를 결합하는 형태로 연구가 진행되었다.

### (1) Multi-Cross-Attentive Model(MCAM)

Hongchan Li 등(2024)은 SNS에서 감정 분석을 위해 이미지와 텍스트를 결합한 MCAM을 제안하였다[4]. 해당 연구에서는 DenseNet121[5]을 사용하여 이미지에서 특징을 추출하고, CBAM(Convolutional Block Attention Module)[6]을 통해 감정과 관련된 중요한 영역에 집중하도록 하는 어텐션 메커니즘으로 결합하였다. 그리고 ALBERT[7]와 BiLSTM[8]을 통해 텍스트에서 감정의 맥락적 요소를 추출하였다. 마지막으로, 이미지와 텍스트에서 추출된 특징 벡터는 크로스 어텐션(Cross-Attention) 메커니즘[9]을 통해 결합하여 효과적으로 감정을 분류할 수 있었다. MCAM은 TumEmo 데이터셋에서 F1 Score 0.767을 달성하였으며, BERT 단일모달 학습 대비 7.72% 개선하였다.

### (2) Image Captioning Joint Attention Mechanism (ICAM)

Ye Sun 등(2023)은 Captioning with Transformers (CATR) 모델을 사용하여 텍스트 정보를 보충하고, BERT 인코더로 텍스트 특징을 추출하였다. 그리고 ResNet[10]과 CBAM으로 이미지 특징을 추출하고 크로스 어텐션으로 결합하며, 다시 BERT 인코더로 추출된 텍스트 정보를 더하는 형태인 ICAM을 제안하였다. ICAM은 Twitter2017 데이터셋에서 BERT 단일모달 학습 대비 3.3% 향상된 0.64 Mac-F1 Score를 달성하였다[11].

## III. 제안하는 내용

본 논문에서는 단일 이미지의 부족한 상황 정보를 보완하기 위해 캡셔닝 모델을 사용하여 이미지에서 나타나는 객체와 배경 관계를 설명하는 텍스트 정보를 생성하고, 추출된 텍스트 특징과 이미지

특징을 양방향 크로스 어텐션 메커니즘으로 결합하는 Captioned-Multimodal Bi-Cross-attention(CMBC) 모델을 제안한다.

(1) Captioned-Multi-modal Bi-Cross-attention 모델

CMBC 모델은 그림 1과 같이 이미지로부터 캡션 정보를 생성하고, ALBERT와 BiLSTM으로 결합된 텍스트 특징 추출기로부터 감정과 맥락 특징을 추출한다. 또한, EfficientNetV2-L과 CBAM으로부터 감정과 관련된 특징을 추출하며, 각 특징 벡터는 수식 1 및 2와 같이 비대칭 구조의 Bidirectional Cross-attention 메커니즘과 수식 3과 같이 Weighted Sum으로 결합된다. 수식 3에서 W는 학습으로 정해지는 가중치를 의미하며 F는 Weighted Sum의 결과인 특징 벡터를 의미한다.

$$Z_{text-img} = softmax\left(\frac{Q_{text} K_{img}^T}{\sqrt{d_k}}\right) \times V_{img} \quad (1)$$

$$Z_{img-text} = softmax\left(\frac{Q_{img} K_{text}^T}{\sqrt{d_k}}\right) \times V_{tea} \quad (2)$$

Weighted Sum으로 결합된 정보는 이전에 추출된 이미지 특징 정보를 다시 사용하여 Self-attention 메커니즘을 적용함으로써 특징 정보를 정교하게 강화한다. Self-attention은 이미지 특징 내에서 중요한 부분들 간의 상호작용을 더 세밀하게 모델링하는 역할을 하며, 마지막으로 추출된 특징 벡터는 Fully Connected Layer를 통해 감정 예측을 수행한다. 본 모

델에서 Bidirectional Cross-attention 메커니즘은 텍스트와 이미지의 상호 정보 전달을 효과적으로 수행하여, 서로 다른 모달 데이터에서 중요한 특징이 풍부하게 통합되도록 하는 것을 목적으로 사용된다. 또한, Weighted Sum은 두 모달 데이터 간의 중요도를 동적으로 조절하여 최적의 결합을 이루게 하며, 이를 통해 감정 분류를 효과적으로 수행할 수 있다.

$$F_w = W_{img} \times Z_{img-text} + W_{text} \times Z_{text-img} \quad (3)$$

(2) Image Captioning 모델

본 연구에서는 이미지의 부족한 상황 정보를 보완하기 위해 이미지 캡션 정보를 생성한다. 캡션 정보는 Pre-Trained MiniCPM-2.6[12] 모델을 사용하였으며, 기존의 CPM을 소형화하여 속도와 효율성을 높인 자연어처리 모델로 우수한 생성 능력이 입증된 모델이다. 모델이 생성하는 캡션은 상황 및 맥락 정보가 중요하며, 맥락 정보 없는 단순한 이미지 설명 텍스트는 멀티모달 학습에 노이즈를 증가시킨다. 본 연구에서는 이미지의 맥락 정보와 문장 수의 관계를 분석하기 위해 프롬프트를 작성하여 텍스트를 생성하였다. 감정 분류를 수행하였을 때 표 1과 같이 문맥 정보를 포함하고 5문장 이내의 길이로 감정 분류를 수행하였을 때 F1 Score 0.733으로 성능이 가장 높았다. 그러나, 문맥 정보를 포함하지 않고, 문장 길이를 3문장 이내로 설정하였을 때 모델이 제대로 학습하지 못하는 것을 확인할 수 있다.

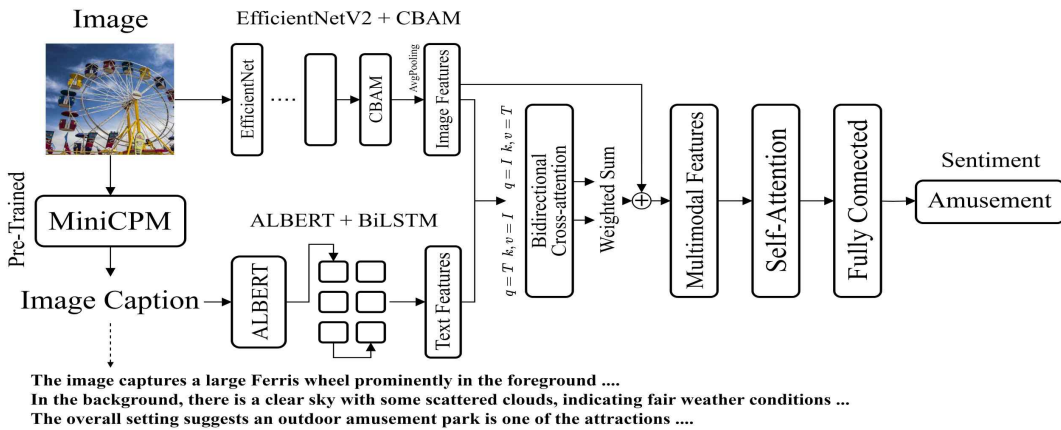


그림 1. Captioned-Multimodal Bi-Cross-attention(CMBC) 모델  
Fig. 1. Captioned-Multimodal Bi-Cross-attention(CMBC) Model



표 1. 텍스트 생성 적합 실험

Table 1. Experimenting with text generation fit

모델	문맥 정보	문장 제한	Accuracy	F1 Score
ALBERT + BiLSTM	미포함	3문장	0.048	0.167
	포함	3문장	0.713	0.713
	미포함	5문장	0.681	0.683
	포함	5문장	0.737	0.736

본 연구에서는 MiniCPM에 문맥 정보를 포함하고 5문장으로 이미지를 설명할 수 있도록 프롬프트를 작성하였으며, 본 모델의 정확도를 검증하기 위해 단일 이미지 감정 분류 실험을 진행하였다.

#### IV. 실험 및 결과

##### (1) 데이터 세트 소개

본 연구에서는 표 2와 같이 즐거움, 분노, 경외, 만족, 혐오, 흥분, 공포, 슬픔 등 8가지 감정이 라벨링된 이미지 데이터 세트인 EmoSet[13]으로 학습을 진행하였으며, 과적합 방지와 모델 성능을 비교하기 위해 데이터 세트를 7.5:2.5 비율로 나누었다.

##### (2) 감정 분류 정확도 평가

본 논문에서 제안하는 CMBC 모델이 정확하게 감정 분류를 수행하는지 검증하기 위해 단일모달 및 멀티모달 방법과 비교 실험을 진행하였다. 이미지 단일모달 학습 모델은 CBAM을 추가로 적용해 보았으며, 기존 모델보다 성능개선의 효과를 확인할 수 있었다. 그리고 생성된 캡션 데이터만을 이용하였을 때는 기존 이미지 분류 모델보다 성능이 하락한 것을 확인할 수 있었다.

멀티모달의 경우, 데이터 세트의 텍스트 데이터가 없으므로 생성된 텍스트와 이미지 결합을 적용하는 ICAM을 구현하였으며, 본 논문에서 제안된 CMBC 모델이 F1 Score 0.787로 3.69% 성능개선 효과를 보였고, EfficientNetV2-L2 이미지 단일모달 학습 대비

표 2. 모델 학습 데이터 세트

Table 2. Model Learning data sets

종류	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness
개수	19,445	10,660	15,037	16,337	10,666	19,828	13,453	12,676
비율	16%	9%	13%	14%	9%	17%	11%	11%
Train	15,556	8,528	12,030	13,069	8,533	15,862	10,762	10,141
Test	3,889	2,132	3,007	3,268	2,133	3,966	2,691	2,535

2.74% 성능개선 효과를 확인할 수 있었다. 이는 멀티모달 데이터를 Bidirectional Cross-attention 및 Weighted Sum으로 결합하고, 기존의 이미지 정보를 다시 사용하는 방식이 효과가 있다는 것을 알 수 있었다.

표 3. 성능 비교 실험

Table 3. Performance comparison experiments

종류	형태	모델	Accuracy	F1 Score
단일 모달	이미지	ResNet50	0.742	0.743
		EfficientNetV2-L	0.765	0.766
		ResNet50+CBAM	0.753	0.752
		EfficientNetV2-L +CBAM	0.773	0.772
	텍스트	RoBERT+BiLSTM	0.725	0.726
		ALBERT+BiLSTM	0.737	0.736
멀티모달		ICAM	0.759	0.759
		<b>CMBC</b>	<b>0.786</b>	<b>0.787</b>




#### V. 결론

본 논문에서는 단일 이미지의 감정 분류를 위해 Captioned-Multi-modal Bi-Cross-attention(CMBC) 모델을 제안하였다. 본 모델은 이미지 내의 객체와 배경 관계를 설명하는 캡션 텍스트를 생성하여 이미지의 부족한 상황 정보를 보완하였으며, Bidirectional Cross-attention 메커니즘과 Weighted Sum으로 결합하였다. 논문에서 제안된 CMBC 모델은 Amusement, Contentment 등 긍정적인 감정에 대해서는 정확한 분류를 수행하였다. 그러나, Fear, Sadness 등 부정적인 감정에 대해서는 표 4의 (b)와 같이 일부 올바르게 않은 결과가 나타났으나, 모델의 전반적인 성능을 측정하였을 때는 Accuracy 0.786, F1 Score 0.787로 최대 8.4% 성능개선 효과를 가져왔다.

이를 통해 논문에서 제안하는 모델이 기존 모델보다 우수한 성능으로 감정 분류를 수행하는 것을 입증하였다.

표 4. CMBC 모델의 감정 예측 수행 결과

Table 4. Results of CMBC model's sentiment prediction

	이미지	이미지 캡션	Predict	Label
(a)		The image depicts an outdoor setting with two women and two dogs. The overall atmosphere suggests a moment of affectionate interaction between humans and animals ...	Contentment	Contentment
(b)		The image captures a sculpture of a human figure engulfed in flames. The background consists of a clear sky with scattered clouds, providing a stark contrast between the natural elements and the dramatic spectacle of the burning sculpture ...	Anger	Anger
(c)		The image displays an interior corridor of a dilapidated building. The walls are made of rough, showing signs of age and neglect with visible cracks and stains. The floor is appear damaged, contributing to the overall atmosphere of abandonment ...	Sadness	Fear

### 참 고 문 헌

- [1] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4), 1093-1113.
- [2] 이미경, 박진완. (2016). 이미지 기반 SNS의 사용성 평가 연구. *한국콘텐츠학회논문지*, 16(8), 508-516.
- [3] Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, 13(7), 4550.
- [4] Li, H., Lu, Y., & Zhu, H. (2024). Multi-Modal Sentiment Analysis Based on Image and Text Fusion Based on Cross-Attention Mechanism. *Electronics*, 13(11), 2069.
- [5] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [6] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).
- [7] Lan, Z. (2019). Albert: A lite bert for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942*.
- [8] Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11), 2673-2681.
- [9] Vaswani, A. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.
- [10] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [11] Sun, Y., Jin, G., Zhao, Y., & Cui, R. (2023, October). Multimodal Sentiment Analysis Based on Image Captioning and Attention Mechanism. In *2023 IEEE 5th International Conference on Civil Aviation Safety and Information Technology (ICCASIT)* (pp. 296-301). IEEE.
- [12] Yao, Y., Yu, T., Zhang, A., Wang, C., Cui, J., Zhu, H., ... & Sun, M. (2024). Minicpm-v: A gpt-4v level mllm on your phone. *arXiv preprint arXiv:2408.01800*.
- [13] Yang, J., Huang, Q., Ding, T., Lischinski, D., Cohen-Or, D., & Huang, H. (2023). Emoset: A large-scale visual emotion dataset with rich attributes. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 20383-20394).

# XGBoost 기반 데이터 증강과 하이퍼파라미터 최적화를 통한 부모-자녀 신장 예측 기법

고건민\*, 하유진\*\*, 김건우#

## Parent-Child Height Prediction Technique Using Data Augmentation and Hyperparameter Optimization Based on XGBoost

Keonmin Ko\*, Yujin Ha and Gun-Woo Kim\*\*

### 요 약

본 연구는 부모의 신장을 기반으로 자녀의 신장을 예측하는 XGBoost 회귀 모델을 개발하고, 이의 성능 향상을 위한 방법론을 제시한다. 기본 모델의 한계를 극복하기 위해 부모 신장의 평균값과 차이를 새로운 변수로 추가하였으며, 노이즈 기반 데이터 증강을 통해 데이터의 다양성을 확보하였다. 또한 Optuna를 활용한 하이퍼파라미터 최적화를 수행하여 모델의 예측 정확도를 개선하였다. 실험 결과, 제안된 방법은 MSE 5.6, MAE 1.6, 결정 계수 0.92를 기록하며 기존 모델 대비 MSE 68%, MAE 50%, 결정 계수 18%의 성능 향상을 달성하였다.

### Abstract

We develop an XGBoost regression model to predict children's height based on their parents' height and propose methodologies for improving its performance. To overcome the limitations of the basic model, we added new variables including the average and difference of parental heights, and secured data diversity through noise-based data augmentation. Additionally, we improved the model's prediction accuracy through hyperparameter optimization using Optuna. The experimental results show that the proposed method achieved MSE of 5.6, MAE of 1.6, and  $R^2$  of 0.92, demonstrating performance improvements of 68% in MSE, 50% in MAE, and 18% in  $R^2$  compared to the existing model.

### Key words

XGBoost regression, height prediction, data augmentation, hyperparameter optimization, parental height

---

\* Dept. of Computer Science & Engineering, Gyeongsang National University, 2024080064@gnu.ac.kr

\*\*Dept. of AI Convergence Engineering Gyeongsang National University, nrt5077@gnu.ac.kr

# Dept. of Computer Science & Engineering, Gyeongsang National University, gunwoo.kim@gnu.ac.kr

※ 본 논문은 2023년도 정부(교육부)의 지원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2021R1G1A1006381)

## 1. 서론

부모의 키를 기반으로 자녀의 키를 예측하는 문제는 유전적 요인을 이해하고, 건강과 성장을 예측하는 문제는 자녀를 둔 부모들의 중요한 관심사 중 하나이다[1]. 그렇기 때문에 자녀의 키를 예측하는 것은 부모들에게 도움을 줄 수 있는 중요한 연구 분야이다. 자녀의 키는 부모의 유전적 특성을 반영하기 때문에[2] 예측 모델로 자녀의 키를 추측하여 부모가 자녀의 성장과 발달에 대한 불안감을 줄이고 적절한 운동이나 영양 섭취와 같은 건강 관리 결정을 내려줄 수 있다[3][4].

본 연구에서는 아버지의 키, 어머니의 키, 자녀의 키, 자녀의 성별을 이용하여 자녀의 키를 예측하는 회귀 모델을 개발하고 성능 개선을 위한 다양한 방법을 적용하였다. 특히, 회귀 모델에서는 그래디언트 부스팅(Gradient Boosting)을 통해 이전 모델이 만든 예측의 오류를 보완하여 강화 학습이 이루어진다. L1(Lasso) 및 L2(Ridge) 정규화를 통해 모델의 복잡성을 제어하여 과대 적합을 방지하며 학습률(learning\_rate), 트리 개수(n\_estimators), 최대 깊이(max\_depth)를 포함한 총 9개의 하이퍼파라미터를 사용하여 모델의 성능을 최적화하는 XGBoost 알고리즘을 사용하였다. 이를 통해 모델의 성능을 향상하고자 하였으며[5] 변수 추가[6], 데이터 증강[7], 하이퍼 파라미터 튜닝[8] 등의 기법을 통해 예측 정확도를 높이는 방법을 탐구하였다.

## 2. 데이터 수집 및 전처리

본 연구에서 사용된 데이터는 부모의 키와 자녀의 성별, 자녀의 키로 구성된 데이터 셋을 사용하였다. 데이터 셋의 구성은 여자 자녀의 수가 2,162명, 남자 자녀의 수가 2,005명으로 총 4,167명이었고, 부모의 수는 4,167명으로 NULL 값이 없는 데이터임을 확인하였다. 표 1은 전체 데이터 셋에 대한 정보이다.

데이터의 변수가 부족하다고 판단하여 모델이 과대 적합 되는 문제를 방지하고, 예측 성능을 향상하기 위해 부모의 키 평균값과 부모의 키 차이를 변수로

표 1. 실험 데이터 세트

table 1. Experimental data sets

		Train	Test	Total
부모	부	3,334	833	4,167
	모	3,334	833	4,167
자녀	남	1,604	401	2,005
	여	1,730	432	2,162

추가하였다[2]. 부모의 평균 키를 토대로 자녀의 예상 키를 추정할 수 있기 때문에, 부모 키의 평균값은 자녀의 유전적인 신장을 반영하는 중요한 지표이다. 하지만 부모의 평균 키가 동일하더라도 부모의 키가 동일하지 않기 때문에 부모의 키 차이를 변수로 추가하였다. 이를 통해 모델이 부모의 평균 키에 대한 학습을 강화하고 평균 키와 키 차이를 통한 상관관계를 이해하여 성능을 높이는 데 집중하였다. 추가한 변수들은 자녀의 키 예측을 위한 데이터 셋의 정보량을 증가시키고 모델이 부모와 자녀의 키 상관관계를 학습할 수 있도록 돕는다. 표 2와 같이 기존의 데이터만을 사용한 XGBoost 모델보다 변수를 추가한 XGBoost (Add Feature) 모델의 성능이 향상된 것을 확인하였다 이를 통해 변수 추가가 데이터의 질과 유용성을 향상시켰다.

## 3. 실험

### 3.1 데이터 증강

적은 양의 데이터는 모델 학습에 있어 과대 적합 문제를 야기할 수 있어 충분한 데이터를 확보하는 것이 중요하다. 표 1의 데이터 수에서 데이터 증강 방법을 활용하여 데이터 수를 늘려 과대 적합은 방지하고 성능을 개선하고자 하였다. 이러한 문제점을 해결하기 위해 노이즈 추가를 통한 데이터 증강 기법을 적용하였다[2]. 노이즈 기반의 데이터 증강 방법은 기존 데이터에 미세한 노이즈를 추가해 새로운 샘플 데이터를 생성하고 데이터 셋을 확장하였다. 확장된 데이터는 모델이 훈련할 때 일반화된 패턴을 학습하게 되어 과대 적합은 방지하면서 성능을

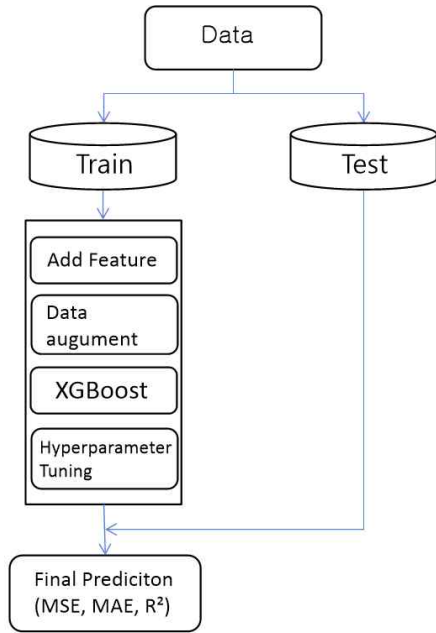


그림 743. 실험 증강도  
Fig 1. experimental flowchart

향상하는 데 기여한다. 노이즈 수준이 0.1일 때 부모의 키에  $\pm 0.1$ 을 추가하여 원본 데이터와 유사한 데이터를 생성하게 된다. 본 연구에서는 노이즈를 0.1부터 0.00001까지 증가시키며 성능을 비교하였다. 이 과정에서 노이즈 레벨이 0.00001일 때 최적의 성능을 나타내는 것을 확인하였다. 적절한 노이즈 레벨을 통한 증강은 모델이 일반화된 학습을 수행하고, 과대 적합을 방지하고 모델의 성능을 향상시켰다.

데이터를 한 번 증강할 때마다 데이터의 양이 2배로 늘어나므로, 증강을 여러 번 반복하면 데이터 수가 기하급수적으로 증가하여 과적합의 위험이 있다. 이에 본 연구에서는 그리드 서치를 사용하여 데이터 증가 배수가 성능에 미치는 영향을 비교 분석하였다.

### 3.2 모델 선정 및 학습 방법

회귀모델에서 성능을 비교하기 위해 7가지의 머신러닝 기반 모델 LR(Linear Regression), SVR, DT (Decision Tree), XGBoost, EN(Elastic Net), RF(Random Forest), LG(LightGBM) 모델을 사용하여 데이터

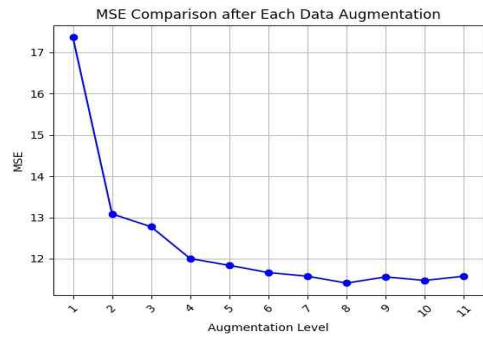


그림 744. 증강에 따른 성능 변화  
Fig 2. Performance variation according to the number of augmentations

분석을 진행하였다. 성능지표로 MSE, MAE,  $R^2$  를 사용하여 각 모델 간의 성능을 비교·분석하였다. 표 2에서 볼 수 있듯이 XGBoost가 다른 모델에 비해 제일 좋은 성능을 보여 XGBoost 모델을 사용하였다.

표 2. 회귀모델 간의 성능 비교

Table 2. Comparison of Regression Models' Performance

Model	MSE	MAE	$R^2$
LR	24.3	3.89	0.67
SVR	23.9	3.77	0.67
DT	22.8	3.48	0.61
EN	24.4	3.88	0.67
RF	18.1	3.17	0.75
LGBM	19.2	3.31	0.74
XGBoost	17.7	3.21	0.78
XGBoost (Add Feature)	13.5	2.9	0.84
XGBoost (Performance Improvement)	5.6	1.6	0.92

XGBoost 모델의 성능을 최적화하기 위해 Optuna 라는 하이퍼파라미터 튜닝 라이브러리를 활용하였다. Optuna는 최적의 하이퍼파라미터값을 자동으로 탐색하여 파라미터값을 이용하여 연구하였다. 이러한 방법으로 XGBoost 모델의 주요 파라미터인 결정 트리 개수( $n\_estimators$ )를 257, 학습률( $learning\_rate$ )을 0.016, 트리 깊이( $max\_depth$ )를 8, 샘플링 비율( $subsample$ )을 0.873으로 튜닝하였다. Optuna의 자동화된 하이퍼파라미터 튜닝을 통해 최적화된 결

과는 모델의 과대 적합을 방지하였다.

#### 4. 실험 결과

본 연구에서는 유전적인 요소인 부모의 키, 성별, 을 바탕으로 자녀의 키를 예측하였다. 적은 양의 데이터와 변수로 인해 모델 성능 향상에 한계가 있어, 이 문제를 해결하기 위해 표 2에서 성능이 가장 좋았던 XGBoost를 사용하였으며, 추가 변수를 생성하고 데이터 증강과 하이퍼파라미터 튜닝을 통해 모델의 성능을 개선할 수 있었다. 또한 표 2에서 XGBoost 모델 성능에 비해 XGBoost(Performance Improvement) 모델의 성능이 MSE가 약 68% 향상된 5.6, MAE가 약 50% 향상된 1.6, 결정계수가 약 18% 향상된 0.92라는 성능을 보였다.

#### 5. 결론

본 연구에서는 부모의 키를 활용하여 자녀의 키를 예측 모델의 성능 향상을 위한 방법을 제안하였다. 모델의 성능을 향상시키기 위한 방법으로 추가 변수를 생성하고 데이터 증강을 통해 데이터를 늘려 과대 적합을 방지하였으며 최적의 하이퍼파라미터 튜닝을 진행하였다. 이러한 방법들을 통해 다른 모델들보다 우수한 성능을 보였다. 이를 통해 부모의 키를 활용한 자녀의 키 예측에 대해 보다 신뢰할 수 있는 접근법을 마련할 수 있을 것으로 기대한다. 하지만 본 연구에서는 식습관, 질병과 같은 외부적 요소를 포함하지 않았기 때문에 다음 연구에서는 외부적인 요인까지 추가하여 성능 개선을 시도할 계획이다.

#### 참 고 문 헌

- [1] 조혜선, “키 크는 주사’ 처방, 5년간 4배 급증… 부작용도 5배↑”, 동아일보, 2024.10.10., <https://www.donga.com/news/Health/article/all/20241010/130187952/2>
- [2] 최민형, 김덕곤, 이진용. (2010). 성장에 대한 유전적·환경적 요인의 영향. 대한한방소아과학회

지, 24(3), 138-149.

- [3] 신민우, “또래보다 작은 우리 아이, 혹시 ‘저신장’은 아닐까?”, 헬스경향, 2016.03.02., <https://www.k-health.com/news/articleView.html?idxno=23996>
- [4] 이보람, “고민되는 아이 저신장 해결법”, 경향신문, 2014.02.05., <https://www.khan.co.kr/life/health/article/201402051448572>
- [5] Kim, J., Won, J., Kim, H., & Heo, J. (2021). Machine-learning-based prediction of land prices in Seoul, South Korea. Sustainability, 13(23), 13088.
- [6] JOVIĆ, Alan; BRKIĆ, Karla; BOGUNOVIĆ, Nikola. A review of feature selection methods with applications. In: 2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO). Ieee, 2015. p. 1200-1205.
- [7] 이태영, 전승배, 이희영, 정명훈. (2023-10-19). Generative Adversarial Network를 이용한 데이터 증강 기반 머신러닝 예측 모델 성능 향상. 대한토목학회 학술대회, 전남.
- [8] Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019, July). Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (p. 2623-2631).

# Bi-LSTM-Transformer 모델을 활용한 시계열 기반 농산물 가격 예측

이규혁\*, 박수진\*, 박지홍\*, 김주영\*, 김건우\*\*

## Time Series-Based Agricultural Price Prediction Exploiting Bi-LSTM-Transformer Model

Gyu-Hyeok Lee\*, Su-Jin Park\*, Ji-Hong Park\*, Ju-Young Kim\*, and Gun-Woo Kim\*\*

### 요약

농산물은 실생활의 필수품으로, 농산물 가격 상승은 소비자의 생활에 직접적인 영향을 미친다. 시계열 데이터를 활용한 농산물 가격 예측에서는 주로 LSTM이 사용되어 왔으나, 긴 시계열 데이터에서의 장기 의존성 학습에 한계가 있어 예측 성능이 저하되는 문제가 있다. 본 연구에서는 이러한 한계를 극복하기 위해 양방향 시퀀스 학습이 가능한 BiLSTM과 Self-Attention 메커니즘 기반의 Transformer를 결합한 BiLSTM-Transformer 모델을 제안한다. 제안된 모델의 성능 평가를 위해 2018년부터 2021년까지의 10개 농산물 품목 가격 데이터를 사용하였으며, Normalized MAE를 기준으로 LSTM 대비 27.40%, Transformer 대비 22.56%의 성능 향상을 달성하였다. 이는 제안된 모델이 장기 의존성 학습을 통해 농산물 가격 예측의 정확도를 효과적으로 개선할 수 있음을 보여준다.

### Abstract

Agricultural products are essential commodities in daily life, and their price fluctuations directly impact consumer living standards. While LSTM has been widely used in agricultural price prediction using time series data, it shows limitations in learning long-term dependencies from extended time series data, resulting in decreased prediction accuracy. To overcome these limitations, this study proposes a BiLSTM-Transformer model that combines BiLSTM, capable of bidirectional sequence learning, with Transformer based on the Self-Attention mechanism. The model's performance was evaluated using price data from 10 agricultural products between 2018 and 2021, achieving performance improvements of 27.40% and 22.56% compared to LSTM and Transformer models, respectively, based on Normalized MAE. These results demonstrate that the proposed model can effectively enhance agricultural price prediction accuracy through improved learning of long-term dependencies.

### Key words

Agricultural Commodity, Predictive models, Time Series Data, Deep Learning, Price Forecasting

---

\* 경상국립대학교 컴퓨터과학부, gyuhyeok\_lee0527@gnu.ac.kr, park\_11412@gnu.ac.kr, hong\_0002@gnu.ac.kr, wndudwkd003@gnu.ac.kr

\*\* 경상국립대학교 교신저자 컴퓨터과학부, gunwoo.kim@gnu.ac.kr

※ 본 논문은 2023년도 정부(교육부)의 지원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2021R1G1A1006381)

## I. 서론

오늘날 농산물은 사람들의 실생활에서 필수적인 품목으로 자리 잡아 사람들의 소비 생활에 큰 영향을 미치고 있다. 2023년 가구당 월평균 가계수지[1]에 따르면 식료품 및 비주류 음료 품목에서 '곡물', '곡물 가공품', '과일 및 과일 가공품', '채소 및 채소가공품' 등 농산물이 차지하는 비중은 대략 31%이다. 그 외 육류는 20%, 수산물은 10% 안팎의 비중을 가진다. 이를 통해 소비자가 식료품에서 농산물을 소비하는 비중이 매우 높다는 것을 알 수 있다. 또한, 농가판매가격지수[2]에 따르면 2023년 농산물 판매가격은 2019년에 비해 약 15.5% 상승하였다. 특히 청과물은 가격이 32.6%로 급격하게 상승하였다. 이러한 통계는 소비자의 농산물 소비 비중이 높아지고 있으며 가격도 가파르게 증가하고 있음을 보여준다. 이에 따라 소비자의 경제적 부담도 점차 커지고 있다.

이러한 상황에서 문제의 심각성을 인지한 정부는 여러 대책을 마련하고 있다. 그 예로 농림축산식품부는 주요 원예농산물 중 수급 불안 우려가 있는 품목에 대해서는 안정 대책을 추진할 계획을 세웠다[3]. 이는 농산물 가격 상승이 일시적인 문제가 아니라 장기적인 문제로 발전할 수 있음을 시사하며 보다 근본적인 해결책이 필요하다는 점을 강조한다.

본 논문에서는 위에서 언급한 내용을 바탕으로 농산물 가격 예측을 위해 시계열 분석을 진행하였다. Bi-LSTM과 Transformer를 결합한 모델을 활용하여 시계열 분석을 진행하여 농산물 가격을 예측함으로써 농산물 가격 상승에 대한 대응 방안을 마련하는데 기여하고자 한다.

## II. 관련 연구

기존의 가격 예측에 관한 연구에서는 시계열 분석 기법이 많이 사용되고 있다. 특히 LSTM(Long Short-Term Memory) 모델은 과거의 데이터를 바탕으로 미래를 예측하는 데 강점을 가진 딥러닝 모델로서 가격 변동의 추세 예측에 활용되고 있다.

### (1) 가격 예측 연구

Sungho Shin 등(2018)은 LSTM 네트워크를 활용하여 다섯 가지 주요 농산물에 대한 가격 예측을 수행하였으며, 그 결과 일반적인 회귀 분석 모델보다 더 높은 예측 성능을 보였다[4]. 그러나 가격 변동성이 큰 데이터를 예측할 때는 변동성을 제대로 학습하지 못하는 한계가 나타났다.

### (2) 이상치 제거 방법

Chan Park 등(2023)은 LSTM 모델의 성능을 높이기 위해 이상치를 제거하는 연구를 진행하였다[5]. 이 연구에서는 월별 사분위 범위를 적용하여 이상치를 탐지하고 이동평균법을 통해 시계열 데이터를 정상화한 후 LSTM 모델에 적용하였다. 그리고 ARIMA 및 Prophet 모델과의 성능을 비교한 결과 이상치를 제거한 LSTM 모델이 가장 높은 예측 성능을 보였다. 이상치 처리는 시계열 데이터의 불규칙성을 완화하여 모델 성능을 향상하는 중요한 역할을 했지만, 비정상 시계열 데이터에서는 이상치를 올바르게 탐지하지 못했다는 한계가 있었다.

## III. 제안하는 방법

본 연구에서는 Bi-LSTM과 Transformer를 결합한 BiLSTM-Transformer 모델을 제안한다. 제안하는 방법은 기존 LSTM 모델이 가진 긴 시계열 데이터에서의 장기 의존성 문제를 해결할 수 있다. 따라서 가격 예측에 있어서 정확성을 높일 수 있다.

### (1) BiLSTM-Transformer

본 논문에서 제안하는 모델의 구조는 다음과 같다. 그림 1은 본 논문에서 제안하는 BiLSTM-Transformer 모델의 전체 구조와 흐름을 나타낸다. 입력 데이터(좌측)를 받아 다음 시점에 대한 정보인 증감률과 가격을 예측한 후 예측한 데이터를 이용해 새로운 입력 데이터(우측)를 만들어 초기 입력 데이터의 처음 시점 증감률을 예측하여 추가 학습할 수 있는 구조이다.



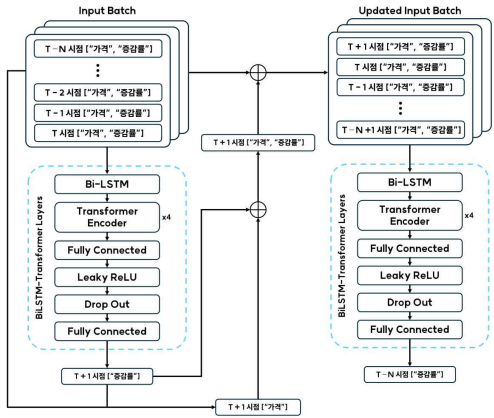


그림 1. BiLSTM-Transformer 모델 구조  
Fig. 1. BiLSTM-Transformer Model Architecture

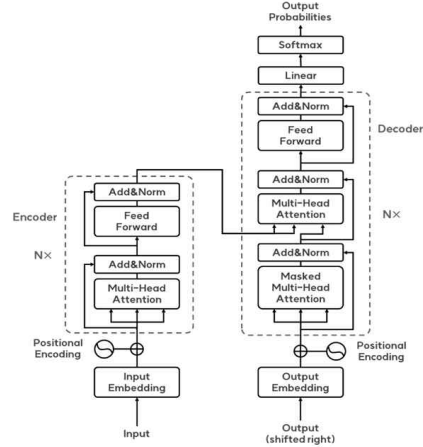


그림 3. Transformer 모델 구조  
Fig. 3. Transformer Model Architecture

(2) Bi-LSTM

Bi-LSTM[6]은 LSTM을 확장한 양방향 LSTM으로 순방향과 역방향으로 시퀀스 데이터를 처리한다. 이 모델을 사용하면 길이가 긴 시퀀스 데이터에서도 모델이 패턴을 잘 파악할 수 있다. 본 논문에서는 제안하는 모델이 예측한 미래의 가격 변동률을 이용해 과거 데이터와 연결하여 재차 학습하는 구조에 사용한다. 그림 2는 Bi-LSTM의 전체 구조를 나타낸다.

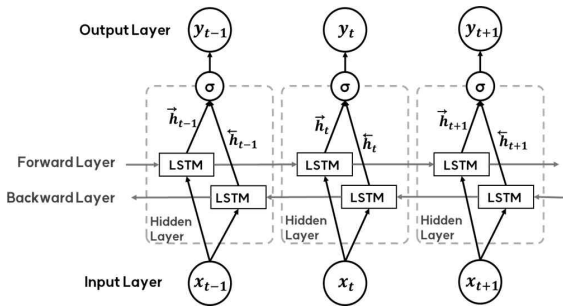


그림 2. Bi-LSTM 모델 구조  
Fig. 2. Bi-LSTM Model Architecture

(3) Transformer

Transformer[7]는 Self-Attention 기법을 핵심으로 이용한 모델이다. Self-Attention은 단일 시퀀스 내의 서로 다른 위치를 연관 지어 시퀀스의 표현을 계산하는 방식이다. 이를 통해 시퀀스 내에서 장기 의존성을 효과적으로 처리할 수 있다. 그림 3은 Transformer의 전체 구조를 나타낸다. 본 모델에서는 Transformer의 Encoder를 사용한다.

IV. 모델 학습

(1) 학습 데이터

본 논문에서 사용한 학습 데이터는 표 1과 같이 예측하고자 하는 10개의 품목에 대해 2018년 1월 중순부터 2021년 12월 하순까지 품목별 143개 시점에서 평균 가격과 증감률로 구성된 데이터 세트를 사용하였다.

표 1. 학습 데이터 세트  
Table 1. Training data sets

속성명	속성값
시점	2018.01.중순 ~ 2021.12.하순
품목명	감자, 건고추 등 10개
품종명	감자 수미, 화건 등 예측할 품종명
거래단위	10개, 30 kg 등 예측할 거래단위
등급	상품, 상 등 예측할 등급
평균가격(원)	각 시점의 품목명 평균가격
증감률	이전 시점 대비 가격 증감률

각 학습 배치는 시점, 품목명 정보 등 모델 학습에 필요하지 않은 데이터를 제거한 후 연속된 11개 시점의 평균 가격과 증감률의 쌍으로 구성했다.

(2) 손실 함수 및 최적화 알고리즘

학습에 사용한 손실 함수는 Huber Loss를 사용하였다. Huber Loss는 Mean Squared Error(MSE)의 장점과 Mean Absolute Error(MAE)의 장점을 결합한

손실 함수이다. 식 (1)은 Huber Loss의 수식이다.

최적화 알고리즘은 AdamW[8]에 학습률 = 2e-5, 가중치 감쇠 = 3e-4,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ 를 사용하여 적용하였다. 식 (2)은 Pytorch 라이브러리에서 AdamW의 가중치 감쇠를 적용한 파라미터 업데이트 수식이다.

$$Huber\ Loss = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i) & \text{for } |y_i - \hat{y}_i| \leq \delta, \\ \delta|y_i - \hat{y}_i| - \frac{1}{2}\delta^2, & \text{otherwise.} \end{cases} \quad (1)$$

$$\theta_t = \theta_{t-1} - \alpha\lambda\theta_{t-1} \quad (2)$$

### (3) 학습률 스케줄러

모델 학습 진행 과정에서 학습의 안정적인 수렴 달성을 위해 CosineAnnealingWarmRestarts[9]에 첫 번째 주기 값은 10, 이후 주기는 2배씩 늘어나게 설정하여 사용하였다. 이를 통해 모델의 과적합을 방지하여 일반화 성능을 높일 수 있다.

### (4) 정규화

본 논문에서 제안한 모델의 Bi-LSTM 레이어를 통과한 출력 결과와 출력 레이어의 출력 결과에 드롭아웃[10] 0.1 비율을 적용하였다. 이를 통해 노드 일부를 비활성화시켜 모델의 일반화 성능을 높일 수 있다.

### (5) 학습 방법

본 논문에서는 10개의 농산물 품목에 대해 품목별로 데이터를 나누어 각각 학습을 진행하였다. 실험 데이터는 과적합 방지 및 비교 분석을 위해 8:2 비율로 학습 및 테스트 데이터로 나누었고, Epoch는 100으로 설정하였으며, 조기 종료율 도입하여 Validation Loss 값이 연속 3번 개선되지 않으면 조기 종료하도록 하였다.

모델 학습의 전체 흐름은 다음과 같다. 1) 11개의 시점 데이터에서 과거 8개의 시점을 분리하여 입력 데이터로 사용하여 다음 시점을 예측한다. 2) 모델을 통해 얻은 예측값은 입력 데이터와 결합하여 다음 반복 단계의 입력으로 사용한다. 3) 모델을 통해

얻은 미래 예측값과 실제값 간의 손실 구하고, 과거 예측값과 실제값 간의 손실을 구하고 이를 통해 전체 손실 값을 구한다. 위 과정을 3번 반복하여 3개의 시점에 대해 예측할 수 있도록 하였다.

## V. 실험 결과

본 논문에서 제안한 BiLSTM-Transformer 모델이 예측한 값의 정확성 비교를 위해 다른 기준 모델인 LSTM, Transformer를 비교 대상으로 하였다. 모델 평가를 위해서 품목별 모델에 Normalized MAE를 평가 지표로 적용한 뒤 평균값을 계산하였다. 식 (3)은 Normalized MAE의 수식을 나타낸다. 본 논문에서는 각 품목의 Normalized MAE 값을 평균 내어 평가 지표로 사용하였다. 표 2는 각 모델이 예측한 값에 대한 평가 결과를 나타내며, 논문에서 제안한 모델이 Normalized MAE 0.151로 가장 낮은 오차를 보였으며, 이는 기존 모델보다 효과적으로 가격 예측을 수행할 수 있다는 것을 보여준다.

$$Normalized\ MAE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (3)$$

표 2. 제안한 모델 결과와 기준 모델들의 결과  
Table 2. Results of the proposed model and baseline models

모델	Normalized MAE
LSTM	0.208
Transformer	0.195
BiLSTM-Transformer	0.151

## VI. 결 론

본 논문에서 농산물 가격 예측을 위해 Bi-LSTM과 Transformer를 결합한 BiLSTM-Transformer 모델을 제안하였다. 실험 결과 제안한 모델은 LSTM 대비 약 27.40%, Transformer 대비 약 22.56% 개선율을 보여주었다. 이러한 결과는 제안한 모델이 농산물 가격 예측에서 더 정확한 예측을 제공할 수 있음을 시사한다. 따라서 농산물 가격 안정화 정책 수립에 기여할 수 있을 것으로 기대된다.

본 논문의 한계와 향후 연구 방향은 두 가지로 요약된다. 첫 번째, 본 논문에서는 평균 가격과 가격 변동률만을 이용하여 예측하였다. 향후 수요와 공급과 같은 다른 데이터도 활용하면 더 높은 정확도를 얻을 것이다. 두 번째, 본 연구에서는 예측하려는 농산물 품목 가격을 구분하여 따로 예측하였다. 그러나 한 품목의 가격 변동률이 다른 품목에 영향을 미칠 수 있다. 따라서 향후 연구에서는 다양한 변인들을 추가로 고려하여 연구할 필요가 있다.

### 참 고 문 헌

- [1] 통계청, 「가계동향조사」, 2024. 2/4., 2024. 10. 20., 가구당 월평균 가계수지 (전국,1인이상), [https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT\\_1L9U001&vw\\_cd=MT\\_ZTITLE&list\\_id=G\\_A\\_10\\_003\\_001&seqNo=&lang\\_mode=ko&language=kor&obj\\_var\\_id=&itm\\_id=&conn\\_path=MT\\_ZTITLE](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1L9U001&vw_cd=MT_ZTITLE&list_id=G_A_10_003_001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE), Accessed: October 19, 2024.
- [2] 통계청, 「농가판매및구입가격조사」, 2024. 2/4, 2024. 10. 20., 농가판매가격지수(2020=100,분기), [https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT\\_1J60&vw\\_cd=MT\\_ZTITLE&list\\_id=P2\\_2&seqNo=&lang\\_mode=ko&language=kor&obj\\_var\\_id=&itm\\_id=&conn\\_path=MT\\_ZTITLE](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1J60&vw_cd=MT_ZTITLE&list_id=P2_2&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE), Accessed: October 19, 2024.
- [3] 농림축산식품부, 「원예농산물 수급 상황 및 대책」, 『KDI 경제정보센터』, 2024. 09. 24., <https://eiec.kdi.re.kr/policy/materialView.do?num=257542>, Accessed: October 19, 2024.
- [4] Sungho Shin, Mikyoung Lee, and Sa-kwang Song, "A Prediction Model for Agricultural Products Price with LSTM Network", The Journal of the Korea Contents Association, Vol. 18, No. 11, pp. 416-420, May. 2018.
- [5] Chan Park and Kyung-Soon Lee, "Deep-Learning-Based Price Prediction by Outlier Detection and Processing for Agricultural Commodity Prices", Journal of Digital Contents Society Vol. 24, No. 8, p. 1899-1906, Aug. 2023
- [6] Mike Schuster and Kuldip K. Paliwal, "Bidirectional Recurrent Neural Networks", IEEE Transactions on Signal Processing, Vol. 45, No. 11, November. 1997
- [7] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser and Illia Polosukhin, "Attention Is All You Need", Advances in neural information processing systems, Jun. 2017
- [8] Ilya Loshchilov and Frank Hutter, "Decoupled Weight Decay Regularization", International Conference on Learning Representations (ICLR), November. 2019
- [9] Ilya Loshchilov and Frank Hutter, "SGDR: Stochastic Gradient Descent with Warm Restarts", conference paper at (ICLR), Feb. 2017
- [10] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever and Ruslan Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", Journal of Machine Learning Research, Vol. 15, pp. 1929-1958, Jan. 2014



# 리그 오브 레전드에서 포지션별 승리 기여도분석 및 승리 예측 모델

정수용\*, 윤웅창\*\*

## Win Contribution Analysis by Position and Win Prediction Model in League of Legends

Suyeong Jeong\* and Woongchang Yoon\*\*

### 요 약

Esports는 전 세계적으로 빠르게 성장하고 있으며 2022년 항저우 아시안 게임에서 정식 종목으로 채택되면서 승리 예측에 대한 중요성이 증가하고 있다. 본 연구는 리그 오브 레전드라는 온라인 게임에서 각 포지션별로 공격, 수비, 라인전, 골드 수급 등의 역할에 따른 기여도를 정의하고, 이를 기반으로 승리 기여도를 계산하여 분석하였다. 피어슨 상관계수를 통해 승리에 영향을 미치는 주요 특성을 선정하고, 로지스틱 회귀 모델과 랜덤 포레스트 모델을 사용하여 승리 예측 모델을 구축하였다. 승리 기여도 분석에서는 Bot 포지션이 공격과 골드 수급에서 가장 높은 기여도를 보여 전체 승리 기여도의 약 30%를 차지하였다. 글로벌 게임 리그별 포지션 기여도의 차이를 분석한 결과, 유의하지 않음을 확인하여 각 리그가 유사한 플레이 스타일을 가진다는 사실을 확인하였다.

### Abstract

Esports is rapidly growing worldwide, and its inclusion as an official event in the 2022 Hangzhou Asian Games has heightened the importance of win prediction. This study defines the contributions in terms of offense, defense, laning, and gold acquisition for each position in the online game League of Legends, and analyzes these contributions by calculating their impact on winning. Key features influencing victory were selected using Pearson correlation coefficients, and win prediction models were developed using logistic regression and random forest models. In the win contribution analysis, the Bot position showed the highest contribution in offense and gold acquisition, accounting for about 30% of the total win contribution. Analysis of the differences in positional contributions across global game leagues revealed no significant differences, confirming that each league has a similar play style.

### Key words

League of Legends, Win Contribution, positional analysis, win prediction, Esports

---

\* 경상국립대학교 컴퓨터공학부, tndyd1emd@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과/바이오의료빅데이터학과, neobits@gnu.ac.kr

## I. 서론

Esports는 전 세계적으로 빠르게 성장하고 있으며, 리그 오브 레전드 (League of Legends, LOL)를 비롯한 여러 인기 게임들이 해당 산업의 성장을 이끌고 있다 [1]. 2021년 월드컵 (LoL World Championship)에서는 동시 시청자 수가 7,386만 명에 달했으며, 2022년 항저우 아시안 게임에서 esports가 정식 종목으로 채택되면서 그 인기는 꾸준히 상승하고 있다 [2, 3].

LOL은 총 10명의 플레이어가 두 팀으로 나뉘어, 각기 다른 포지션에서 공격, 수비, 골드 수급, 라인전 등 다양한 역할을 수행하며 팀 전략을 바탕으로 대결하는 팀 기반 전략 게임이다. 각 포지션에서의 기여도는 팀 승리에 중요한 요소로 작용한다 [2].

LOL의 승리 예측과 관련된 연구들은 게임 시작 전 예측과 시작 후 예측으로 나눌 수 있다. 게임 시작 전 예측 연구에서는 플래티넘 이상의 티어에서 등장하는 챔피언들의 능력치 (공격력, 방어력, 주문력)를 기반으로 4,050개의 데이터를 활용해 장단기 메모리 (Bidirectional LSTM) 모델을 적용하여 58.07%의 예측 정확도를 보여주었다 [4]. 반면, 게임 시작 후 예측 연구에서는 챌린저 티어에서 26,000개의 데이터를 수집하고, 10분과 15분 시점의 주요 지표를 활용해 승리를 예측하였다. 단층 신경망 (Single-Layered Neural Network)과 다층 신경망 (Multi-Layered Neural Network) 모델을 사용하여 특성의 가중치를 평가하였으며, 드래곤 격차, 레벨 격차, 타워 파괴 격차, 전령 처치 여부라는 4개의 주요 특성을 통해 딥 뉴럴 네트워크 (DNN) 모델을 적용해 93.75%의 높은 예측 정확도를 보였다 [5].

다른 시작 후 예측 연구에서는 2020년 여름 국내 프로경기를 대상으로 1,714개의 데이터를 수집하여, 게임이 시작된 후 10분 시점의 지표만을 사용해 승리를 예측하였다 [6]. 이 연구에서는 단계 선택법 (stepwise selection)을 통해 특성을 선정하고, 로지스틱 회귀 (Logistic Regression, LR)와 의사결정 나무 (Decision Tree)를 사용해 각각 95%와 93%의 예측 정확도를 보였다.

LOL은 세계적으로 3개의 상위 게임 리그가 있

며, 각 리그는 고유의 특성을 가지고 있어 우위를 가리기 힘들다. 각 리그별 데이터는 다양한 선수들과 포지션별 전략 구상에 중요한 역할을 한다 [7].

본 연구는 각 포지션별 공격, 수비, 라인전, 골드 수급에 대한 기여도를 정의하고, 이를 바탕으로 승리 기여도를 분석하였다. 또한, 각 리그별 포지션의 승리 기여도의 차이를 확인하였다. 이를 통해 기존 경기 데이터를 활용하여 승리 기여도와 포지션과의 관계를 분석하고 승리 예측 모델을 제안하는 것이 목적이다.

## II. 실험 방법

### 2.1 데이터 수집 및 전처리

본 연구에서는 Oracle's Elixir 사이트를 통해 2023, 2024년도 국내 프로경기 (League of Legends Champions Korea, LCK), 유럽 프로경기 (League of Legends European Championship, LEC), 북미 프로경기 (League of Legends Championship Series, LCS)에 대한 데이터를 수집하였다 [8] (표 1).

표 1. 실험 데이터 세트  
Table 1. Experimental data set

리그	플레이어 레코드		팀	
	2023년	2024년	2023년	2024년
LCK	4,880	4,820	10	10
LEC	2,870	2,940	10	10
LCS	2,640	1,910	11	8
합계	10,390	9,670	31	28

다양한 시간대의 경험치, 크립스코어, 골드 차이 지표 중 라인전 종료 시점에 해당하는 15분대 지표만을 선택하여 사용하였으며, 모든 특성들을 0-1 사이의 값으로 정규화하였다.

여러 특성 중에서 분석과 모델 구축에 중요한 특성을 선정하기 위해 여러 특성과 승리와의 상관관계를 피어슨 상관계수 (Pearson correlation coefficient)를 사용하여 계산하였다 (그림 1).

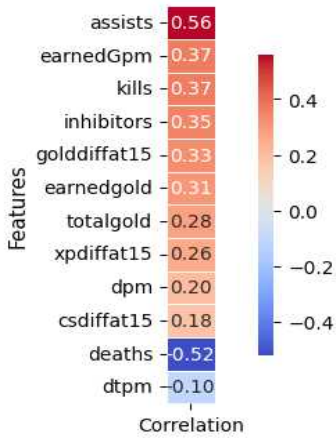


그림 1. 승리와 주요 특성과 상관관계  
Fig. 1. Correlation between victory and key features

## 2.2 중요 특성 선정

이를 통해서 양의 상관관계 중 상위 10개의 특성과 음의 상관관계 중 상위 2개의 특성을 선택하였다. 특히, deaths (죽은 횟수) 특성은 가장 높은 음의 상관관계를 나타내었으나, deaths 특성 만으로는 수비 기여도를 평가하는데 한계가 있어 dtpm (분당 받은 피해량)과 deaths 특성을 이용하여 dpmd (테스 대비 분당 받은 피해량)를 계산하여 사용하였다.

$$dpmd = \frac{dtpm}{deaths + 1} \quad (1)$$

이를 통해 수비 기여도는 단순히 사망 횟수뿐만 아니라 플레이어가 어느 정도의 피해를 흡수하였는지까지도 고려하였다.

## 2.3 승리 기여도와 리그별 포지션 차이 분석

양의 상관관계를 나타내는 상위 10개의 특성을 이용해 공격, 골드, 라인전 기여도를 정의하고, 음의 상관관계를 나타내는 특성을 이용해 수비 기여도를 정의하였다.

공격과 관련된 특성에 해당하는 assists (도움)의 3개 특성들을 공격 기여도 특성으로 선정하였다. 같은 방법으로 골드 획득 기여도는 earnedgpm (분당 직접 획득한 골드)의 2개의 특성을 선정하였으며, 라인전 기여도는 golddiffat15 (15분 시점 상대 라이너와 골드)의 2개 특성들로 정의하였다.

각 포지션은 서로 다른 역할을 수행하며, 공격,

수비, 라인전, 골드 수급 등에서 팀의 승리에 기여한다. 하지만, 각 포지션의 역할 차이를 무시한 킬/데스의 총합으로 일괄적인 평가 방식은 포지션별 기여도를 정확하게 반영하지 못한다 [9]. 때문에, 승리 기여도는 각 포지션별 역할에 따른 순위를 기반으로 가중치를 부여하여 계산하였다. 이때, 가중치 a/b/c/d는 각각 공격/수비/라인전/골드 기여도 순위를 나타낸다. 1위는 0.6의 가중치 값을 가지며, 차 순위부터 0.1씩 낮아져서 5위는 0.1의 가중치 값을 사용한다.

$$VC = a \times AC + b \times DC + c \times LC + d \times GC \quad (2)$$

승리 기여도 (VC)는 특정 선수가 공격 기여도 (AC), 수비 기여도 (DC), 라인전 기여도 (LC), 골드 기여도 (GC) 측면에서 기여한 정도를 각각 측정하여 이를 가중치 a, b, c, d를 곱한 값을 합산한다.

$$FVC = \left( \frac{\sum VC}{TC} \right) \times TG \quad (3)$$

최종 승리 기여도 (FVC)는 모든 포지션별 선수의 기여도 (VC) 합을 팀 전체 기여도 (TC)로 나누고 팀의 총 게임 수 (TG)를 곱하여 정규화 하였다.

마지막으로 LCK, LCS, LEC 등 각 지역별 리그의 차이가 리그에서의 포지션별 승리 기여도에 영향을 미치는지를 확인하고 하였다. 이를 위해서 각 리그의 승리 기여도가 차이가 있는지를 ANOVA 테스트를 통해 유의성을 확인하였다.

모든 통계적 유의수준은 >0.05를 선택 하였다.

## 2.4 기계학습 모델을 통한 승리 예측

본 연구에서는 피어슨 상관관계를 통해서 도출한 중요 12개의 특성을 활용하여 랜덤 포레스트 (Random Forest Model, RF)와 로지스틱 회귀(Logistic Regression Model, LR) 모델을 통해서 승리를 예측하였다 (그림 1).

2023년도 데이터를 이용하여 기본 모델을 랜덤하게 5분할 샘플링을 통해 학습하였으며, 테스트 데이터 세트로는 2024년도 데이터를 사용하였다. 각 모델의 성능은 Accuracy, Precision, Recall, F-1 Score를 통해서 판단하였다. RF모델과 LR 모델 간 성능 차이는 t-test를 통해서 유의성을 검증하였다.

### III. 결 과

#### 3.1 승리 기여도 분석

포지션별 역할에 따른 기여도 순위를 분석한 결과, 공격에서 가장 중요한 역할을 담당한 포지션은 Bot으로 나타났으며, 수비에서는 Jungle이 가장 중요한 포지션임을 확인하였다. 또한, 라인전 역할에서는 Top이 가장 중요하며 골드에서는 Bot이 가장 큰 역할을 하는 것을 확인하였다. Sup 경우에는 다른 포지션을 지원하기 때문에 모든 역할에서 낮은 순위를 보였다 [2].

표 2. 포지션별 역할 기여도 순위

Table 2. Ranking of role contributions by position

포지션	공격	수비	라인전	골드
Top	3	2	1	3
Jungle	4	1	4	4
Mid	2	3	3	2
Bot	1	5	2	1
Sup	5	4	5	5

포지션별 역할 기여도 순위를 통해서 계산된 최종 승리 기여도와 승리와의 관계는 0.46의 상관관계를 가지며, 이는 기존 양의 상관관계에서 두 번째로 높은 값을 확인하였다 (그림 1).

#### 3.2 각 리그에서의 포지션별 승리 기여도 비율 분석

우리는 LCK, LCS, LEC 리그 간 포지션별 승리 기여도 비율에 차이가 있는지 확인하였다.

모든 리그에서 Bot 포지션이 약 30%로 가장 높은 승리 기여도를 기록했으며, 다음으로는 Top과 Mid 포지션이 각각 약 25%로 비슷한 기여도를 보였다. Jungle 포지션은 약 15%, Sup 포지션은 약 7.5%로 가장 기여도를 보였다 (그림 2).

또한, 각 리그의 포지션별 승리 기여도 비율의 차이를 확인한 결과, 통계적으로 유의미한 차이를 보이지 않았다 (그림 2와 표 3). 이를 통해 리그 간 경기에서 각 포지션의 역할이 유사함을 확인할 수 있었다.

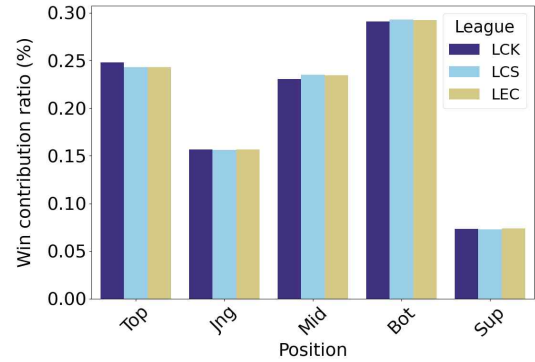


그림 2. LOL 리그의 포지션별 승리 기여도 비율

Fig. 2. Contribution rate to victory by position in the LOL league

표 3. 각 포지션별 리그 간의 차이에 대한 통계분석 결과

Table 3. Statistical analysis results on differences between leagues by each position

포지션	F-통계량	p-값
Top	1.33	0.265
Jngle	0.033	0.968
Mid	1.53	0.216
Bot	0.015	0.985
Sup	0.502	0.605

#### 3.3 LR과 RF 모델을 통한 승리 예측

우리는 피어슨 상관계수를 통해서 선정한 중요 특성 12개를 이용하여 RF와 LR모델을 통한 승리 예측을 시도하였다. 각 모델의 승리 예측의 성능은 Accuracy, Precision, Recall, F1-score의 값이 모두 0.9에 가까운 값을 보여서 예측 성능이 우수함을 확인하였다 (표 4). LR모델과 RF모델간의 성능 차이를 통계적으로 확인한 결과, 유의하지 않음을 확인하였다 (표 4, 5).

표 4. 기계학습 모델들을 통한 승리 예측 결과

Table 4. Win prediction results through machine learning models

평가지표	RF모델		LR모델	
	training	test	training	test
Accuracy	0.9102	0.8784	0.9092	0.8842
Precision	0.9110	0.8824	0.9094	0.8892
Recall	0.9102	0.8784	0.9092	0.8842
F1-score	0.9102	0.8781	0.9092	0.8838



표 5. 통계분석을 통한 각 모델 간의 성능 차이 결과  
Table 5. Results of performance differences between each model through statistical analysis

평가지표	Training		Test	
	F-통계	p-값	F-통계	p-값
Accuracy	0.24	0.81	-1.32	0.22
Precision	0.39	0.70	-0.883	0.40
Recall	0.26	0.80	-1.32	0.22
F1-score	0.24	0.81	-0.78	0.45

#### IV. 결론

본 연구는 포지션별 승리 기여도 분석을 통해 특정 포지션이 승리에 얼마나 중요한 역할을 하는지 확인하였으며, 각 리그의 포지션별 승리 기여도 차이 또한 확인하였다. 이를 통해서 각 포지션이 여러 역할에서 어느 정도의 기여도를 가지는지를 확인하였으며, 각 리그의 포지션별 기여도가 유사함을 확인하였다. 또한, 상관관계 분석을 통해 선정된 중요 특성을 기반으로 구축한 승리 예측 모델은 독립된 데이터에서도 높은 예측 정확도를 보였다.

향후 연구에서는 더 다양한 특성들을 고려하고, 딥러닝 모델을 사용하여 예측 정확도를 향상시킬 수 있을 것이다. 또한, 향후 LOL의 업그레이드에 따른 포지션별 승리 기여도의 변화가 예상되므로, 모델의 지속적인 개선이 필요할 것이다.

결론적으로 본 연구에서 제안한 승리 기여도와 승리 예측 모델을 통해 팀의 승리 전략 수립에 중요하게 사용될 수 있을 것이며 esports 산업의 발전에 도움을 줄 수 있을 것으로 기대한다.

#### 참고 문헌

[1] B. Bickman, R. Farrell, and S. Petrullo, "The Rise of eSports: Insights Into the Perceived Benefits and Risks for College Students," *International Journal of eSports Research*, vol. 1, no. 1, pp. 1-15, Jan.-June 2021.

[2] Riot Games, "Celebrating Worlds 2021," *LoL Esports*, 2021. [Online]. Available: <https://lolesports.com/ko-KR/news/celebrating-worlds-2021>.

[Accessed: Oct. 21, 2024].

[3] Shen, Q. (2022). A machine learning approach to predict the result of League of Legends. *Proceedings of the 2022 International Conference on Machine Learning and Knowledge Engineering (MLKE)*, IEEE, 38-45.

[4] Cheolgi Kim, and Soowon Lee, "Predicting Win-Loss of League of Legends Using Bidirectional LSTM Embedding," *KIPS Transactions on Software and Data Engineering*, Vol.9, No.2, pp. 61-68, 2020

[5] Si-Jae No, Yoo-Jin Moon, and Young-Ho Hwang, "Quantitative Analysis for Win/Loss Prediction of 'League of Legends' Utilizing the Deep Neural Network System through Big Data," *Journal of The Korea Society of Computer and Information*, Vol. 26, No. 4, pp. 213-221, 2021.

[6] Jung-Min Lee, "Design of a Prediction Model for League of Legends Match Outcomes using Machine Learning," *Master's Thesis, Ajou University*, 2021.

[7] Jappe Bulckaen, "Uncovering Differences Between League of Legends Competitions: A Cluster and Discriminant Analysis", *Master's Thesis, Ghent University*, 2021.

[8] Oracle's Elixir, "League of Legends Esports Stats and Data," *Oracle's Elixir*, 2024. [Online]. Available: <https://oracleselixir.com>. [Accessed: Oct. 21, 2024].

[9] Farnod Bahrololloomi, Sebastian Sauer, Fabio Klonowski, Robin Horst, and Ralf Dörner, "A Machine Learning based Analysis of esports Player Performances in League of Legends for Winning Prediction based on Player Roles and Performances," in *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022)*, pp. 68-76, 2022.

# 경두개 집속초음파 시뮬레이션에서 Pseudo-CT의 유효성 분석

권수영\*, 유지훈\*, 임철기\*\*, 서현\*

## Analysis of Pseudo-CT in Transcranial Focused Ultrasound Stimulation

Sue-Young Kwon\*, Ji-Hun Yu\*, Cheolki Im\*\*, and Hyeon Seo\*

### 요약

경두개 집속 초음파 시뮬레이션은 CT(Computed Tomography) 또는 MRI(Magnetic Resonance Imaging)로부터 얻는 의료 영상이 필요하지만, 방사선 노출의 위험성으로 인해 CT 데이터 확보가 어렵다. 본 연구는 Pseudo-CT(PCT)를 활용하여 초음파 시뮬레이션의 정확도를 분석하고, 3D residual U-Net을 이용한 PCT 생성 모델 오픈 소스의 실용성과 일반화 가능성을 검증하고자 한다. 이를 위해 새로운 공개 데이터셋인 synthRAD2023 Grand Challenge를 사용하여 CT와 MRI 데이터를 획득 및 PCT를 제작했고, 0.5 MHz 단일 트랜스듀서를 사용한 집속 초음파 시뮬레이션을 수행했다. 연구 결과, PCT는 CT와 유사한 초점 영역과 음압 분포를 나타냈으며, MRI에 비해 더 정확한 결과를 제공했다. PCT의 최대 음압은 CT에 비해 10.1% 작았고, 집속 영역의 FWHM 넓이는 CT에 비해 2.67% 작았다. 이에 본 연구팀은 집속 초음파 시뮬레이션 시 오픈 소스를 활용한 PCT 사용을 추천한다. 그러나 본 연구는 2D 시뮬레이션과 단일 CT 데이터에 기반하여 일반화에 한계가 있으므로, 향후 연구에서는 다양한 CT 데이터를 이용한 3D 시뮬레이션을 통해 결과의 일반화를 보고하고자 한다.

### Abstract

Transcranial focused ultrasound simulations require both CT and MRI, but the acquisition of CT data is challenging due to the radiation exposure. This study aims to analyze the accuracy of ultrasound simulations using Pseudo-CT (PCT) and to verify the practicality and generalization potential of an open-source PCT generation model based on a 3D residual U-Net. To achieve this, we utilized the new publicly available dataset, synthRAD2023 Grand Challenge, to acquire CT and MRI data, and subsequently generated PCT. We conducted focused ultrasound simulations using a single 0.5 MHz transducer. The results demonstrated that PCT exhibited focal regions and acoustic pressure distributions similar to those of CT, providing more accurate results compared to MRI. The maximum acoustic pressure of PCT was 10.1% lower than that of CT, and the full width at half maximum (FWHM) of the focus area was 2.67% smaller than that of CT. Therefore, we recommend the use of PCT generated via open-source tools for focused ultrasound simulations. However, as this study was based on 2D simulations and a single CT dataset, its generalizability is limited. Future research will report on the generalization of results through 3D simulations using diverse CT datasets.

### Key words

MRI, CT, Pseudo-CT, Transcranial focused ultrasound

\* 경상국립대학교 IT공과대학 컴퓨터공학과 {victoria619, dbwlgns0923}@naver.com, hseo0612@gnu.ac.kr (교신저자)

\*\* 경상국립대학교 기초과학연구소 chelgi47@gmail.com

※ 본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (RS-2023-00280241).

## I. 연구 배경

뇌 신경조절(neuromodulation) 기법은 외부 자극을 통해 뇌의 특정 부위에 자극을 가해 신경세포의 활성화 정도를 조정하는 기술이다. 뇌 신경조절 방법은 크게 자기장을 통해 자극을 주는 TMS (Transcranial Magnetic Stimulation)와 전기장을 통해 자극을 주는 tDCS(transcranial Direct Current Stimulation)로 나눌 수 있다[1]. 하지만 이런 방법들은 해상도가 낮아 넓은 영역에 유도 전류의 영향을 미치기 때문에 국소적인 영역의 자극이 어렵다[2]. 반면 경두개 집속초음파(Transcranial Focused ultrasound, tFUS)는 TMS, tDCS 기술에 비해 높은 공간 해상도를 제공하여 정밀한 자극이 가능하다[3].

tFUS 시뮬레이션은 CT 또는 MRI로부터 얻는 의료 영상을 활용하여 뇌 모델을 만들고 외부 자극으로부터 집속되는 음압 분포를 제공한다. 이때 CT는 뼈와 같은 단단한 조직 구조를 시각화하는데 유용하고, MRI는 연조직, 특히 뇌 조직과 그 형태를 상세히 확인하는데 주로 사용된다. 집속초음파 시뮬레이션 시 두개골이 초음파 전달의 주요 방해 요소이기 때문에 정확한 계산을 위해서는 CT 활용이 필요하다. 하지만 CT를 찍을 때는 방사선 노출 위험이 있기 때문에 MRI에 비해 데이터를 확보하기 어렵다.

PCT란 MR 이미지를 기반으로 CT 이미지를 합성하는 기술로, 주로 컨볼루션 신경망(CNN)을 활용하여 MR 이미지를 CT 이미지로 변환한다. 하지만 이러한 기술은 구현이 용이하지 않아 연구자들 사이에서 널리 사용되지 못하고 있다. 이를 해결하기 위해 최근 Yaakub의 연구에서 3D residual U-Net을 활용해 T1-weighted MRI를 입력으로 PCT를 생성하는 오픈 소스를 공개했다[4]. 이 방법은 특히 초음파 시뮬레이션 연구에서 두개골을 모델링하는데 유용하다. 본 연구팀은 Yaakub의 오픈소스를 활용하여, 해당 모델이 학습 및 테스트에 사용된 데이터셋 외의 다른 공개 데이터셋(MRI)에서도 안정적이고 효과적으로 PCT를 생성할 수 있는지 평가하고자 한다. 또한 이를 통해 집속초음파 시뮬레이션에서 PCT가 CT와 유사한 계산 결과를 도출하는 지 분석하여 모델의 일반화 가능성과 실용성을 검증하는 것을 목표로 한다.

## II. 연구 방법

본 논문에서 Yaakub의 오픈 소스의 실용성과 일반화 가능성을 탐색하기 위해, Yaakub의 모델 학습에서 사용된 CERMEP 데이터셋[5]이 아닌 MRI와 CT를 함께 제공하는 synthRAD2023 Grand Challenge[6]를 사용하였다. CERMEP 데이터셋은 Siemens Sonato 1.5 T scanner를 사용하여 획득되었고 이미지는  $161 \times 192 \times 192$  크기로  $1.2 \times 1.2 \times 1.2 \text{ mm}^3$  voxel 크기를 가지고 있는 반면, synthRAD2023 Grand Challenge 데이터셋은 세 개의 네덜란드 대학 의료센터에서 획득하였기 때문에 Philips/Siemens를 사용하였고 픽셀의 크기 또한 세 개의 센터에 따라 다양하다. 따라서 synthRAD2023 Grand Challenge 데이터셋은 MRI와 CT를  $1 \times 1 \times 1 \text{ mm}^3$  voxel 크기로 resampling 한 후 두 이미지 간의 co-registration을 수행한 데이터를 제공하고 있다.

먼저, 본 연구팀은 Yaakub의 오픈 소스를 활용하여 PCT를 제작하였다. 의료 영상 처리를 위한 오픈 소스 도구로 MONAI, ANTsPy를 사용하였고 synthRAD2023 Grand Challenge 데이터셋의 pretrained\_net\_final\_20220825.pth를 사용하여 PCT 이미지를 생성했다.

CT와 PCT 이미지는 Seg3D software (NIH/NIGMS CIBS, University of Utah, UT, USA)를 이용하여 두개골 영역을 분할하고 MRI 영상은 simNIBS-4을 이용하여 segmentation을 진행하여 두개골 영역을 분할했다[7][8]. 시뮬레이션 도구로 MATLAB (MathWorks, Natick, UNITED STATES)의 초음파 시뮬레이션 오픈 소스인 k-Wave MATLAB toolbox[9]를 이용했다.

트랜스듀서의 주파수는 500kHz, 반지름과 직경은 30mm, 진폭(음압)을 1MPa로 설정하였다. 시뮬레이션 공간은  $0.42 \text{ mm} \times 0.42 \text{ mm} \times 0.42 \text{ mm}$  voxel로  $552 \times 645 \times 514$  공간에서 구현했다. 두개골의 경우 두개골 내에 존재하는 미세한 구멍에 의한 영향을 고려하기 위해 다공성의 개념을 이용하여 밀도, 속도, 감쇠 계수를 정의하였다. 두개골 내에 존재하는 연조직 사이의 초음파 반사는 무시할 정도로 작고 두개골에 의한 영향에 집중하기 위해 두개골을 제외한

영역은 균일한 물성으로 물과 같다고 가정하였다. HU(Hounsfield Unit)을 0~1 사이의 값으로 정규화 후 다음의 식을 따라 불균질 물성 모델을 구축하였다.

$$\psi = 1 - \frac{HU}{1000} \quad (1)$$

$$c_{skull} = c_{water}\psi + c_{bone}(1 - \psi) \quad (2)$$

$$\rho_{skull} = \rho_{water}\psi + \rho_{bone}(1 - \psi) \quad (3)$$

$$\alpha_{skull} = \alpha_{minskull} + (\alpha_{maxskull} - \alpha_{minskull})\psi^{0.5} \quad (4)$$

$c$ 는 음속,  $\rho$ 는 밀도,  $\alpha$ 는 감쇠 계수를 의미한다. CT, PCT의 경우 다공성을 고려한 위의 식을 반영한 물성을 부여하였고 MRI는 하나의 레이어로 구성되어 있어 단일의 물성값을 사용했다. 다공성을 고려한 두개골의 음속, 밀도 그리고 감쇠 계수와 두개골 단일 물성, 그리고 물의 물성은 표 1에 정리하였다. 시뮬레이션은 CT, PCT, MRI를 이용해서 만든 3가지 두개골 모델의 2D 단면에서 수행되었고 트랜스듀서의 경우 머리의 중앙에 위치시켰다.

표 1. 물과 두개골의 음속, 밀도  
Table 1. Speed of sound and density of water and bone

	speed of sound [ $m \cdot s^{-1}$ ]	Density [ $kg \cdot m^{-3}$ ]
CT, PCT	$c_{water} = 1482, c_{bone} = 3100$	$\rho_{bone} = 2200$
MRI	$c_{water} = 1482, c_{bone} = 2850$	$\rho_{bone} = 1735$

### III. 연구 결과

그림 1은 새로운 공개 데이터셋에서 획득한 MRI를 사용하여 제작한 PCT 결과이다. MRI에서 보기 어려운 뼈의 구조가 PCT에서 제작된 것을 확인할 수 있으며 낮은 해상도지만 두개골의 다공성 구조를 보이는 해면골(spongy bone)이 관측된다.

그림 2는 왼쪽부터 CT, PCT, MRI를 이용한 두개골 모델에서 계산된 음압 분포이다. CT와 PCT 모델의 초점 영역과 음압이 집중된 영역이 유사한 반면, MRI의 초점 영역은 CT, PCT에 비해 현저히 작

은 것을 관찰할 수 있다.

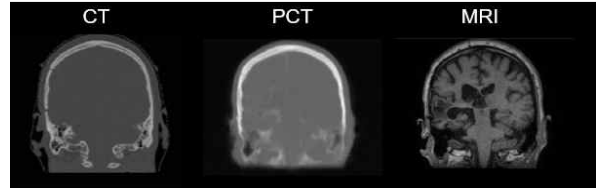


그림 1. CT, MRI, PCT의 머리 단면 이미지  
Fig. 1. Head slice views from CT, MRI, and PCT

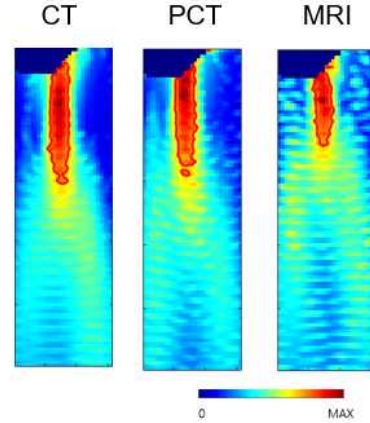


그림 2. CT, PCT, MRI의 초음파 시뮬레이션 결과  
Fig. 2. Ultrasound Simulation Results of CT, PCT, and MRI

모델 간 정량 분석을 위해 두개골을 통과한 이후 뇌 영역에서 음향 최대값을 계산하였고, CT 모델은 1.4037 MPa, PCT가 1.2624MPa, MRI가 1.6053MPa이다. CT에 비해 PCT의 최대 압력이 10.1% 작고 CT에 비해 MRI의 최대 압력이 14.3%가 크다는 것을 확인할 수 있다. 집중 영역은 뇌 영역에서 계산된 음향 최대값의 절반이 되는 영역으로 FWHM 넓이를 정의하였고 CT 모델은  $67.032mm^2$ , PCT 모델은  $65.444mm^2$ , MRI 모델은  $40.572mm^2$ 가 나왔다. FWHM 넓이의 경우 CT에 비해 PCT가 2.7% 작고 CT에 비해 MRI가 39.5% 작다는 것을 확인할 수 있다.

### IV. 결론 및 향후 계획

최근 높은 공간 해상도와 뇌의 심부 영역에 표적으로 자극할 수 있는 집중 초음파 시뮬레이션 기반 신경조절 기술이 주목받고 있으며, 다양한 자극 조건에 따른 시뮬레이션 연구가 활발히 발표되고 있

다. 그러나 주요 방해 요소인 두개골의 구조를 반영하기 위해 CT가 필요하나 데이터 획득의 어려움이 존재하며, 이에 여러 연구팀은 MRI를 활용한 시뮬레이션을 대안으로 수행하고 있다[1]. 해결 방안으로 인공지능을 활용한 PCT 기법이 제안되고 있으나 기술 구현이 용이하지 않았기에 최근 Yaakub의 연구[4]에서 이를 공개하였다. 이에 본 연구팀은 해당 오픈 소스의 활용성을 검증하기 위해 새로운 공개 데이터셋을 활용한 집속 초음파 시뮬레이션을 수행하였다.

Yaakub의 연구에 따르면[4], 최대 음압 기준으로 PCT는 CT와 4.2%, MRI는 41.6%의 차이를 보였으며, 집속 영역에서는 PCT가 17.3%, MRI가 25.2%의 차이를 나타냈다. 공개 데이터셋을 활용한 경우, 최대 음압에서 PCT는 CT와 10.1%, MRI는 14.3%의 차이를 보였고, 집속 영역에서는 PCT가 2.67%, MRI가 39.47%의 감소를 나타냈다. 따라서, Yaakub의 연구보다 본 연구에서 PCT와 MRI 모두 CT 대비 더 큰 오차가 관측되었지만, MRI의 정확도가 PCT보다 떨어지기 때문에 초음파 시뮬레이션에서는 MRI보다는 PCT 사용을 추천한다.

본 연구에서는 시뮬레이션이 2D로 진행해서 실제 결과에 비해 정확도가 떨어진다. 또한 한 사람의 CT 사진을 통해 실험했기 때문에 결과의 일반화가 부족하다. 따라서 향후 다양한 CT 사진을 3D로 시뮬레이션을 진행하고 결과를 일반화할 예정이다.

## 참 고 문 헌

[1] H. Seo and E. H. Lee, "Comparison of Analysis Results According to Heterogeneous or Homogeneous Model for CT-based Focused Ultrasound Simulation", *Journal of Biomedical Engineering Research*, Vol. 43, No. 6, pp. 369-374, Dec, 2022

[2] Huang, Y., et al., "Numerical investigation of the energy distribution of low-intensity transcranial focused ultrasound neuromodulation for hippocampus" *Ultrasonics*, Vol. 124, Aug, 2022.

[3] Legon, W., et al., "Neuromodulation with single-

element transcranial focused ultrasound in human thalamus", *Human brain mapping*, Vol. 39, No. 5, pp. 1995-2006, May, 2018

[4] siti N. Yaakub. " Pseudo-CTs from T1-weighted MRI for planning of low-intensity transcranial focused ultrasound neuromodulation: An open-source tool" *Journal of BRAIN STIMULATION*, vol. 16, Jan-Feb, 2023

[5] Mérida, Inés, et al. "CERMED-IDB-MRXFDG: a database of 37 normal adult human brain [18 F] FDG PET, T1 and FLAIR MRI, and CT images available for research." *EJNMMI research* 11 (2021): 1-10.

[6] Huijben, et al, " Generating Synthetic Computed Tomography for Radiotherapy: SynthRAD2023 Challenge Report ", *Medical Image Analysis*, Vol.97, Oct, 2024

[7] Seg3D C. volumetric image segmentation and visualization. Salt Lake City: Scientific Computing and Imaging Institute (SCI); . www.seg3d.org, 2020.

[8] Oula Puonti, et al, "Accurate and robust whole-head segmentation from magnetic resonance images for individualized head modeling", *NeuroImage*, Vol.219, Oct, 2020

[9] Ji-Hun Yu, et al "A simulation study of low-intensity focused ultrasound simulation according to diameter of transducer" *Journal of KIIT*. Vol.18, No.0, May, 2022

# 청각 자극 기반 실제 발화와 상상 발화의 ERP 분석: ICA를 통한 artifact 제거 평가

백종화\*, 임철기\*\*, 전성찬\*\*\*, 이성환\*\*\*\*, 서현\*

## ERP Analysis of Auditory Stimulus-Based Auditory Speech and Inner Speech: Evaluating Artifact Removal through ICA

Jonghwa Paik\*, Cheolki Im\*\*, Sung Chan Jun\*\*\*, Sunghan Lee\*\*\*\*, and Hyeon seo\*

### 요 약

최근 EEG(Electroencephalogram) 신호에서 일반적인 전처리 과정이 데이터에 부정적인 영향을 미칠 수 있다고 발표되었지만, 눈 깜빡임이나 근육 활동 같은 artifact를 제거하는 ICA(Independent Component Analysis)가 청각 자극에 대한 상상 발화 데이터 질에 미치는 영향은 아직 평가되지 않았다. 이에 본 연구팀은 EEG/ERP를 통해 청각 자극에 대한 실제/상상 발화 데이터에서 ICA 적용 전후의 artifact 제거 효과를 분석하고자 한다. 5명의 피실험자로부터 각각 180회의 소리에 대한 실제/상상 발화 과제를 통해 얻은 총 900개의 데이터를 사용했다. 결과적으로, ICA는 amplitude 감소를 통해 artifact 제거 효과를 확인할 수 있었다. 그러나 ERP 신호의 시간적 일관성 확보에는 한계가 있었다. 피실험자 간의 신경 신호 변동성과 ICA 과정에서의 artifact 제거 한계로 인해, 일부 피실험자에서는 시간 간격의 편차가 증가하는 양상이 나타났다. 결론적으로, ICA는 artifact 제거에는 성공적이었으나, ERP 분석에서 시간적 특징을 유지하는 데는 제한적임을 시사한다.

### Abstract

Recently, it has been reported that conventional preprocessing procedures on EEG(Electroencephalogram) signals can negatively affect the data. However, the impact of ICA(Independent Component Analysis) in removing noise such as eye blinks or muscle activity on the quality of inner speech data in response to auditory stimuli has not yet been evaluated. Therefore, our research team aimed to analyze the effects of artifact removal before and after applying ICA on actual and inner speech data in response to auditory stimuli using EEG/ERP. We used a total of 900 data points obtained from 5 participants, each completing 180 tasks involving actual and inner speech related to auditory stimuli. As a result, ICA was found to be effective in artifact removal through amplitude reduction. However, there were limitations in maintaining temporal consistency of the ERP signals. Due to inter-subject variability in neural signals and the limitations of artifact removal in the ICA process, some participants showed increased variability in timing intervals. In conclusion, while ICA was successful in artifact removal, it was limited in maintaining temporal consistency in ERP analysis

### Key words

EEG (Electroencephalography), ERP (Event-Related Potential), Independent Component Analysis (ICA), Inner Speech

---

\* 경상국립대학교 IT공과대학 컴퓨터공학과 qorwhdghk12@naver.com, hseo0612@gnu.ac.kr (교신저자)

\*\* 경상국립대학교 기초과학연구소 chelgi47@gmail.com

\*\*\* 광주과학기술원 AI대학원 scjun@gist.ac.kr

\*\*\*\* 한림대학교 뇌혈관질환 선도연구센터 sh.lee@halym.ac.kr

※ 본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (RS-2023-00280241).

## 1. 서 론

뇌-컴퓨터 인터페이스는 척수 손상 및 기타 신경 근육 질환으로 고통받는 환자들에게 의사소통 및 제어를 가능하게 하는 기술로 최근 몇십 년 동안 큰 진전을 이루었다. 전통적인 접근 방식은 motor imagery 방식으로 실제로 신체를 움직이지 않고 특정 운동을 상상하여 뇌의 운동 피질에서 발생하는 신경 신호를 활용하는 방식이다[1]. 그 밖에 시각 자극을 활용한 방법도 널리 사용되고 있다[2,3]. 반면, 청각 자극에 대한 상상 발화의 경우 상대적으로 넓은 영역의 피질이 활성화되어 특정 신호를 해석하기 어렵지만, 음성 언어가 가장 일반적인 의사소통 방식이기 때문에 가장 자연스럽게 직관적인 과정으로 작용할 수 있다[4-6].

EEG의 경우 두피에서 비침습적으로 뇌의 전기 활동을 기록하며 높은 시간적 해상도로 다양한 자극과 인지 과제에 대한 실시간 반응을 측정할 수 있다[7]. 청각 자극에 대한 상상 발화 시 생성되는 EEG 신호 패턴을 분석하고 해석하여 BCI 시스템에 활용하려는 연구가 진행되고 있으나 여전히 상대적으로 낮은 분류 정확도가 관찰되고 있다[4-6]. 최근 이러한 분류는 기계 학습 및 딥러닝 기법으로 시행되고 있으며 신호에 대한 전처리 과정을 포함하고 있다.

일반적으로 EEG 측정 중 눈 깜빡임, 근육 움직임 등의 artifact가 포함되기 때문에 신호 전처리 과정을 수행한다. 이때 자주 사용되는 독립 성분 분석(ICA)은 artifact와 신경 신호를 분리하여 신경 신호에 집중할 수 있도록 한다[8]. 최근 Delorme's et al. 연구에서 일반적으로 사용되는 ICA를 포함한 high-pass filtering, referencing 등이 데이터에 부정적인 영향을 미칠 수 있으며 "EEG는 그냥 그대로 두는 것

이 낫다"라고 결론지었다. 그러나 눈 깜빡임이나 근육 활동 같은 artifact를 제거하는 ICA가 청각 자극에 대한 상상 발화 데이터 질에 미치는 영향은 아직 평가되지 않았다. 이에 본 연구팀은 EEG/ERP[9]를 통해 청각 자극에 대한 실제 발화(AS, Auditory Speech)와 상상 발화(IS, Inner Speech) 데이터의 ICA 적용 전후의 artifact 제거 효과를 분석하고자 한다.

## II. 연구 방법

전체 실험은 그림 1 같이 진행되었으며, 총 5명의 피실험자 데이터를 사용했다. 소리 자극에 대한 실제 발화와 상상 발화로 인지 과제를 나눠, 피실험자별로 실제 발화 90개, 상상 발화 90개를 얻어 총 900개의 데이터를 사용하여 분석을 진행하였다. 비교 분석을 위해, 모든 채널 중에서 중앙에 위치하며 EEG 분석에서 많이 사용하는 Cz 채널을 기준으로 실험을 진행하였다[10].

ERP는 각 인지 과제당 하나만 계산하고, ERP 분석에 사용되는 수식을 다음과 같이 정의했다.

$$V_{erp} = \max\left(\frac{1}{N} \sum_{i=1}^N V_i(t)\right)$$

$$V_{peak} = \max(\max(V_1), \max(V_2), \dots, \max(V_N))$$

$$T_{erp} = \operatorname{argmax}_t \left(\frac{1}{N} \sum_{i=1}^N V_i(t)\right)$$

$$T_{peak} = \operatorname{argmax}_t (\max(V_1), \max(V_2), \dots, \max(V_N)) \quad (1)$$

$V$ 는 데이터의 amplitude,  $t$ 는 데이터의 time,  $N$ 은 trial의 개수이다. 각 식을 활용해 ERP amplitude의 최댓값( $V_{erp}$ )과 시점( $T_{erp}$ ), 그리고 EEG 신호에서의 최댓값( $V_{peak}$ )과 최댓값이 관측된 시점( $T_{peak}$ )을 분석하였다.

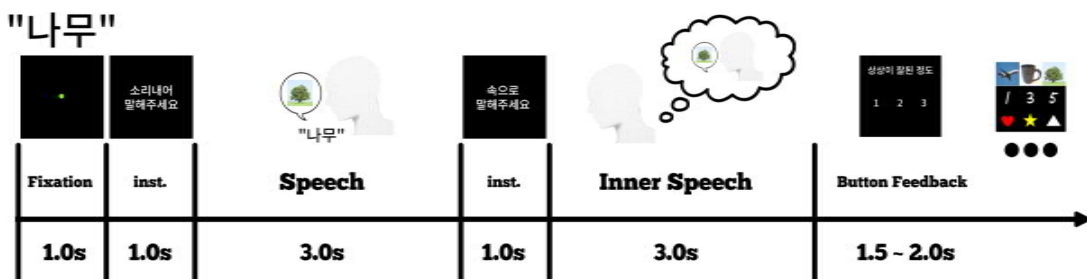


그림 1. 실험 타임라인  
Fig. 1. Experimental timeline

### III. 연구 결과

ICA 적용 전(Before)과 후(After) ERP 그래프를 함께 표시한 결과 (그림 2), ICA 적용 후 실제 발화와 상상 발화 데이터 모두에서 1초까지의 amplitude가 많이 감소하는 경향이 관찰되었다.

다음으로 ICA 적용 전/후에 따른  $V_{erp}$ 와  $V_{peak}$  간의 차이와  $T_{erp}$ 와  $T_{peak}$ 의 차이를 분석하였다 (그림 3). ICA가 효과적으로 artifact를 제거하고 일관된 신호 특징을 유지하는지 지표로서 ERP의 최댓값

( $V_{erp}$ )과 EEG의 최댓값( $V_{peak}$ )의 차이를 분석하였다. 모든 피실험자에서 일관되게 ICA 이후  $V_{erp}$ 와  $V_{peak}$  간의 차이가 감소된 것을 확인하였다. 이는 신호의 특징이 더욱 강조되었을 가능성을 시사한다. 또한, 4번 피실험자는 다른 피실험자들에 비해 더 많은 artifact가 관찰되었고, 이에 따라  $V_{erp}$ 와  $V_{peak}$  간의 차이의 큰 편차를 보였다. 그러나 ICA 적용 후 평균과 표준 편차가 감소하였으며, 이는 해당 피실험자가 실험 중 특정 신체 움직임이나 artifact에 해당하는 활동이 컸을 가능성을 나타낸다.

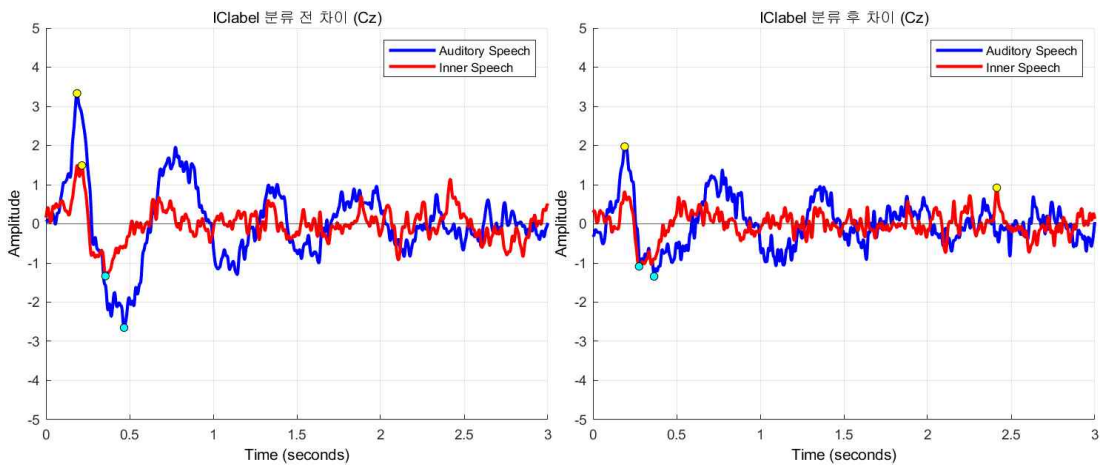


그림 2. ICA 적용 전후 ERP 비교  
Fig. 2. Comparison of ERP Before and After ICA Application

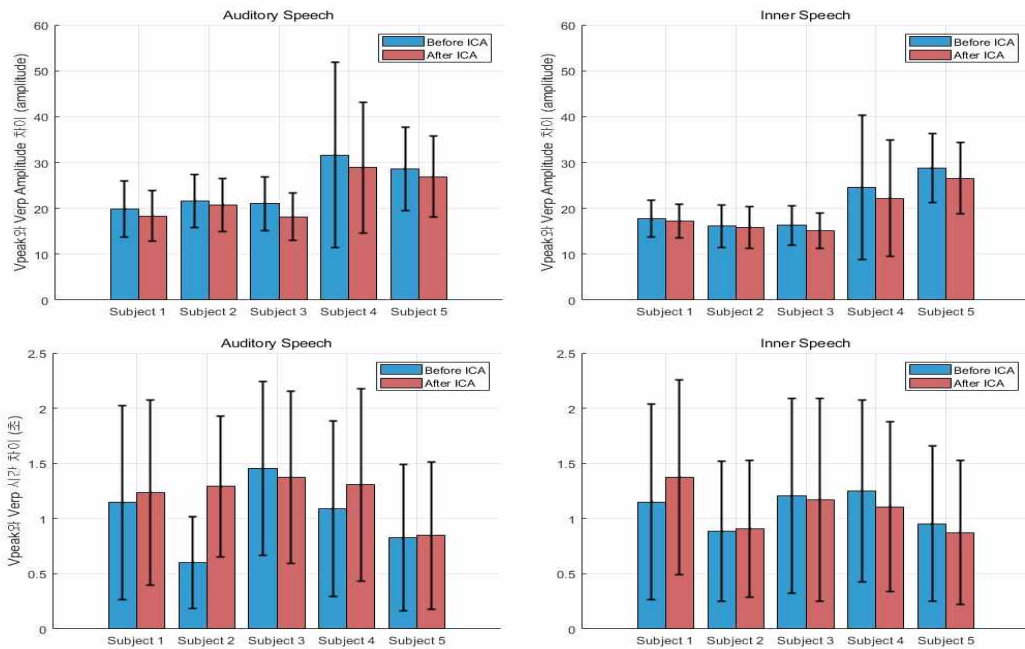


그림 3. ICA 적용 전후 피실험자별 ERP 진폭 및 피크 시간 비교  
Fig. 3. Subject-wise Comparison of ERP Amplitude and Peak Timing Before and After ICA Application



만약, artifact의 감소로 인해 데이터의  $V_{peak}$ 가 ERP 특성을 잘 나타낼 수 있도록 이동했다면, 기존의  $T_{erp}$ 와  $T_{peak}$ 의 시간 차이도 줄어들 것이다. 그러나 시간적 간격에서 일관된 경향을 확인하기 어려웠으며, 일부 피실험자의 경우 시간 간격의 편차가 증가하였다. 이는 피실험자 간 신경 신호의 변동성이나 ICA 적용 과정에서 artifact 제거의 한계 때문일 수 있다.

#### IV. 결론 및 향후 계획

본 연구에서는 EEG 신호를 ICA 적용 전후의 ERP를 비교하여 artifact 제거 효과와 신경 신호 변화를 분석하였다. 실험 결과, ICA는 amplitude 감소를 통해 artifact 제거에 일정 부분 효과적임을 보여주었으며, 특히 상상 발화와 같은 민감한 인지 과제에서도 전반적인 artifact 감소가 관찰되었다. 그러나  $T_{erp}$ 와  $T_{peak}$  간의 시간적 간격에서 일관된 변화를 확인하기 어려웠고, 일부 피실험자에서는 신경 신호의 변동성으로 인해 시간 간격의 편차가 증가하는 양상이 나타났다. 이는 ICA가 artifact 제거에는 성공적이었으나, ERP 분석에서 시간적 일관성 확보에는 한계가 있음을 시사한다. 특히 상상 발화와 같은 복잡한 과제에서는 피실험자 간 신경 신호의 개인차가 크게 작용할 수 있음을 고려해야 한다. 따라서 향후 연구에서는 ICA의 신호 손실 가능성 및 artifact 제거 성능을 더 면밀히 평가하고, 다양한 과제와 채널에서 신경 신호 변동성을 심층적으로 연구할 필요가 있다.

#### 참 고 문 헌

[1] Cho, Hohyun, et al. "EEG datasets for motor imagery brain-computer interface." *GigaScience* 6.7 (2017): gix034.  
[2] Chen, Yiyu, et al. "A high-security EEG-based login system with RSVP stimuli and dry electrodes." *IEEE Transactions on Information Forensics and Security* 11.12 (2016): 2635-2647.  
[3] Lee, S., Jang, S., & Jun, S. C. (2022). Exploring

the ability to classify visual perception and visual imagery EEG data: toward an Intuitive BCI System. *Electronics*, 11(17), 2706.  
[4] Lee, S. H., Lee, M., & Lee, S. W. (2020). Neural decoding of imagined speech and visual imagery as intuitive paradigms for BCI communication. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(12), 2647-2659.  
[5] Yu, S. H., Park, H. Y., Lee, S. J., & Jeong, J. H. (2023). Sentence-Level Neural Language Decoding based on Speech Imagery from EEG Signals. *한국정보기술학회논문지*, 21(8), 153-164.  
[6] Sereshkeh, A. R., Trott, R., Bricout, A., & Chau, T. (2017). EEG classification of covert speech using regularized neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(12), 2292-2300.  
[7] Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement science review*, 2(2), 1-11.  
[8] Sanei, S., & Chambers, J. A. (2013). *EEG signal processing*. John Wiley & Sons.  
[9] Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial psychiatry journal*, 18(1), 70-73.  
[10] Bachmann, M., Lass, J., & Hinrikus, H. (2017). Single channel EEG analysis for detection of depression. *Biomedical Signal Processing and Control*, 31, 391-397.

# Weighted-Voting 앙상블 기반 BERT 및 LLM을 활용한 식품 개체명 인식

임소희\*, 김건우\*\*

## Food Named Entity Recognition Using BERT and LLM with Weighted-Voting Ensemble

So-Hee Lim\*, and Gun-Woo Kim\*\*

### 요 약

최근 자연어 처리(NLP) 분야에서 개체명 인식(NER)은 중요한 연구 과제로 주목받고 있으며, 특히 식품 관련 데이터를 정확하게 식별하고 분류하는 작업은 여러 산업에서 필수적이다. 본 연구에서는 문맥 이해에 강점을 가진 BERT와 범용적 언어 이해 능력이 뛰어난 대규모 언어 모델(LLM)을 결합하여 식품 개체명 인식을 수행하였다. TASTEset과 FoodBase 데이터셋을 통합하고 BIO 태깅 방식으로 전처리한 후, weighted-voting 앙상블 기법을 적용하여 두 모델의 장점을 결합하였다. 실험 결과, 단일 모델 대비 B-FOOD와 I-FOOD 태그에서 각각 0.07, 0.06 정도의 성능 향상을 보였으며, 전체적으로 0.82의 weighted average F1 score를 달성하였다.

### Abstract

Named Entity Recognition (NER) has emerged as a crucial research topic in Natural Language Processing (NLP), with accurate food entity recognition becoming essential across various industries. In this paper, we combines BERT's contextual understanding capabilities with the comprehensive language comprehension of Large Language Models (LLM) for food entity recognition. The integrated dataset from TASTEset and FoodBase was preprocessed using BIO tagging, and a weighted-voting ensemble technique was applied to leverage the strengths of both models. Experimental results showed performance improvements of 0.07 and 0.06 for B-FOOD and I-FOOD tags respectively compared to single models, achieving an overall weighted average F1 score of 0.82.

### Key words

Food, NER, BIO tagging, BERT, GPT-3.5-turbo, Ensemble, weighted-voting

---

\*경상국립대학교 컴퓨터공학과, 220161103\_@gnu.ac.kr

\*\*경상국립대학교 교신전자 컴퓨터공학과, gunwoo.kim@gnu.ac.kr

※ 본 논문은 2023년도 정부(교육부)의 제원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2021R1G1A1006381)

## 1. 서론

최근 자연어 처리(Natural Language Processing, NLP) 분야에서 특정 개체를 인식하고 분류하는 개체명 인식(Named Entity Recognition, NER)은 중요한 연구 주제로 부상하고 있다. 특히 식품 관련 개체를 텍스트에서 정확하게 인식하는 것은 다양한 산업 분야에서 중요한 과제가 되고 있다. 식품 데이터는 요리법, 영양 정보, 식품 성분 등 다양한 형태로 존재하며, 이를 정확히 인식하는 것은 식품 관련 시스템을 구축하거나 다른 분야와의 관계를 분석하는 데 필수적이다[1].

텍스트 마이닝에서는 의미 있는 정보를 추출하고 분석하는 정보 추출(Information Extraction, IE)이 중요하다. 이 과정에서 가장 첫 단계는 개체명 인식(Named Entity Recognition, NER)으로, 특정 개체를 텍스트 내에서 정확하게 식별하고 분류하는 작업이다[3]. 본 연구에서는 이러한 개체명 인식 기술을 식품 관련 텍스트에 적용하여, 식품과 관련된 개체를 효과적으로 분류하는 것을 목표로 한다. 이를 위해 트랜스포머 기반의 BERT 모델과 대규모 언어 모델(Large Language Model, LLM)을 결합한 앙상블 기법을 적용하였다.

## II. 관련 연구

개체명 인식(NER) 기술은 텍스트 마이닝에서 정보 추출의 필수적인 단계로 자리 잡고 있다[2]. 초기에는 규칙 기반 또는 사전 기반 접근이 주로 사용되었으나, 최근에는 심층 학습 기반 모델들이 등장하며 그 성능이 크게 향상되었다[3]. 특히 BERT와 같은 트랜스포머 기반 모델은 텍스트 내 문맥을 깊이 이해하는 데 강점을 가지고 있으며, 식품 데이터와 같은 복잡한 도메인에서도 뛰어난 성능을 보여주고 있다[4].

한편, 대규모 언어 모델(Large Language Model, LLM)은 적은 양의 데이터로도 높은 성능을 발휘할 수 있는 Few-shot Learning과 Zero-shot Learning 기법으로 주목받고 있다. 이러한 모델들은 제한된 데이터에서도 언어적 패턴과 상관관계를 포괄적으

로 이해할 수 있는 장점을 가진다[3]. BERT는 세밀한 문맥 정보를 파악하는 데 뛰어나며, LLM은 보다 범용적인 패턴을 학습하는 데 강점을 보여주기 때문에, 두 모델을 결합하면 BERT의 정밀성과 LLM의 범용성을 모두 활용할 수 있다.

본 연구에서는 이러한 두 모델의 앙상블을 통해 각 모델의 예측값에 가중치를 부여하는 weighted voting 기법을 적용하였다. 이 방식은 더 높은 정확도를 기대할 수 있을 뿐만 아니라, 각각의 모델이 가진 장점을 극대화하여 신뢰성 있는 결과를 얻을 수 있다[5]. 이러한 연구는 향후 식품과 질병 간의 관계를 탐구하는 생물의학 텍스트 마이닝에 중요한 기초 자료로 활용될 것이다.

## III. 실험

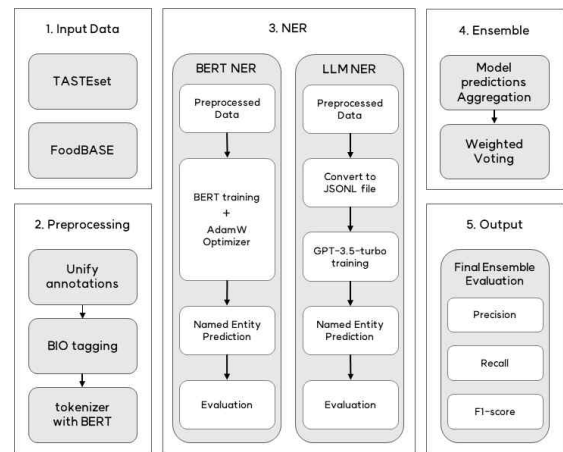


그림 1. 실험 아키텍처  
Fig 1. Experiment Architecture

### 3.1 식품 텍스트 데이터

본 논문에서는 TASTEset[6]과 FoodBase[2] 두 가지 데이터를 사용하였다. TASTEset은 700개의 데이터로 구성되어 있으며, 13,000개 이상의 엔티티를 포함하고 있다. 이 데이터는 Food, Quantity를 포함한 9종의 개체명으로 주석 처리되었다. FoodBase는 Hugging Face에서 제공된 데이터셋을 활용하였으며, BIO 태그를 사용해 B-Food, I-Food, O 3종의 개체명으로 주석 처리되었다. 두 데이터셋을 통합하기 위해 TASTEset의 개체명 중 Food를 제외한 나머지

8종의 개체명을 O로 변경하고, BIO 형식을 적용하여 FoodBase와 동일한 형식으로 주석을 통일하였다. 이후, BERT 모델의 토큰라이저를 사용하여 모델이 학습할 수 있도록 전처리 과정을 수행하였다. 데이터는 학습용과 검증용으로 8:2로 분할하였다.

### 3.2 BERT 모델 학습

트랜스포머 기반의 BERT(Bidirectional Encoder Representations from Transformers) 모델을 활용하여 식품 관련 개체명 인식(NER) 작업을 수행하였다. 사전 학습된 'bert-base-uncased' 모델을 사용하여 BIO 태그(B-Food, I-Food, O)를 기반으로 개체명을 분류하였다. 학습 과정에서는 AdamW 옵티마이저를 사용하여 학습률 5e-5로 20 에폭(epoch) 동안 모델을 학습시켰다. 이후, 정확도, 정밀도, 재현율, F1 점수를 통해 성능을 평가하였다. 결과는 <표 1>에 제시되어 있다. B-FOOD 태그와 I-FOOD 태그의 성능은 O 태그보다 다소 낮지만, 전체적으로 평균 0.81의 성능을 나타냈다.

표 1. BERT 모델 개체명 인식 성능 평가  
Table 1. BERT Model Named Entity Recognition(NER) Performance Evaluation

BERT	precision	recall	f1-score
O	0.87	0.89	0.88
B-FOOD	0.63	0.65	0.64
I-FOOD	0.58	0.56	0.57
macro avg	0.69	0.70	0.70
weighted avg	0.80	0.81	0.81

### 3.3 LLM 모델 학습

LLM 모델은 OpenAI의 GPT-3.5-turbo 모델을 선정하였다. BERT와 동일하게 사전에 토큰화된 데이터를 사용하여 주석 처리된 데이터를 기반으로 식품 관련 개체명 인식(NER) 작업을 수행하였다. OpenAPI를 통해 모델에 적합한 대화형 형식인 JSONL 파일로 변환하였다. 사용자(User) 역할로 입력 문장을 제공하고 모델은 어시스턴트(Assistant) 역할로 각 문장의 BIO 태그를 반환하는 구조로 학습을 진행하였다. 학습용 JSONL 파일을 사용하여 GPT-3.5-turbo 모델을 Fine-tuning하여 식품 개체명

인식 작업에 맞게 모델을 최적화하였다.

Fine-tuning이 완료된 후, BERT와 동일하게 정확도, 정밀도, 재현율, F1 점수를 사용하여 테스트용 JSONL 파일에 대해 평가를 진행하였다. 평가 결과는 <표 2>에 제시되어 있다. BERT 모델에 비해 전반적으로 성능이 다소 낮지만, 평균 0.70의 성능을 나타냈다.

표 2. GPT-3.5-turbo 모델 개체명 인식 성능 평가  
Table 2. GPT-3.5-turbo Model Named Entity Recognition(NER) Performance Evaluation

GPT-3.5-turbo	precision	recall	f1-score
O	0.80	0.76	0.88
B-FOOD	0.47	0.46	0.47
I-FOOD	0.39	0.35	0.37
macro avg	0.55	0.52	0.54
weighted avg	0.69	0.66	0.70

### 3.4 BERT와 LLM 모델의 앙상블

본 연구에서는 BERT와 GPT-3.5-turbo의 성능을 결합하여 더욱 향상된 개체명 인식(NER) 성능을 도출하기 위해 weighted-voting 앙상블 기법을 적용하였다. BERT 모델과 GPT-3.5-turbo 모델은 각 개체명(B-Food, I-Food, O)을 독립적으로 예측한 후, 각 모델의 예측값에 가중치를 부여하여 최종 결과를 도출하였다. 가중치는 예측 결과가 비교적 높았던 BERT 모델에 0.6, GPT-3.5-turbo 모델에 0.4를 부여하였다. 가중치를 부여한 후, 최종 결과는 <표 3>에 제시되어 있다. O 태그의 성능은 기존 BERT 모델과 큰 차이가 없지만, B-FOOD 태그와 I-FOOD 태그에서 각각 0.07과 0.06 정도 성능이 향상되었다.

표 3. 앙상블된 최종 모델 개체명 인식 성능 평가  
Table 3. Ensembled Final Model Named Entity Recognition(NER) Performance Evaluation

Ensemble	precision	recall	f1-score
O	0.88	0.90	0.89
B-FOOD	0.70	0.71	0.71
I-FOOD	0.62	0.65	0.63
macro avg	0.73	0.75	0.74
weighted avg	0.81	0.83	0.82

#### IV. 결론

본 연구는 BERT와 GPT-3.5-turbo 모델을 활용하여 식품 관련 텍스트에서의 개체명 인식(NER)을 수행하고, 이를 통해 식품 관련 데이터를 효과적으로 식별하고 분류할 수 있는 시스템을 제안하였다. 연구에 사용된 TASTEset과 FoodBase 데이터셋을 통합하여 BIO 태그 형식으로 전처리한 후, BERT와 GPT-3.5-turbo 모델을 각각 학습시켰다. 학습된 두 모델에 weighted-voting 앙상블 기법을 적용하여 성능을 개선하였다. 그 결과, 단일 모델을 사용했을 때보다 BERT와 GPT-3.5-turbo의 강점을 결합한 앙상블 모델이 더 높은 정확도를 달성하였다.

본 연구는 식품 관련 텍스트 마이닝 및 NLP 응용 분야에서 NER 모델의 가능성을 확인하는 데 의미가 있다. 특히, 개체명 인식 성능을 높이기 위해 서로 다른 특성을 가진 두 모델을 결합한 접근 방식은 향후 의료, 헬스케어 및 기타 도메인에서의 데이터 분석에도 유용하게 활용될 것으로 기대된다.

향후 연구에서는 데이터셋에 다양한 태그를 적용하여 세부적인 NER을 실행하거나 다양한 종류의 트랜스포머 모델 및 대규모 언어 모델을 사용하여 성능을 개선할 예정이다.

#### 참 고 문 헌

- [1] S. Downer, S. A. Berkowitz, T. S. Harlan, D. L. Olstad, and D. Mozaffarian, "Food is medicine: actions to integrate food and nutrition into healthcare", *BMJ*, vol. 369, p. m2482, 2020, doi: 10.1136/bmj.m2482.
- [2] G. Popovski, B. K. Seljak, and T. Eftimov, "FoodBase corpus: a new resource of annotated food entities", *Database*, vol. 2019, 2019, doi: 10.1093/database/baz121.
- [3] A. Roy, "Recent Trends in Named Entity Recognition (NER)", *arXiv preprint arXiv:2101.11420*, 2021, doi: 10.48550/arXiv.2101.11420.
- [4] S. Yun, S. Kang, and H. Kim, "BERT-based logits ensemble model for gender bias and hate speech detection", *J. Inf. Process. Syst.*, vol. 19, no. 5, pp.

641-651, 2023, doi: 10.3745/JIPS.04.0287.

- [5] S. Jaradat, R. Nayak, A. Paz, and M. Elhenawy, "Ensemble learning with pre-trained transformers for crash severity classification: A deep NLP approach", *Algorithms*, vol. 17, p. 284, 2024, doi: 10.3390/a17070284.
- [6] A. Wróblewska, A. Kaliska, M. Pawłowski, D. Wiśniewski, W. Sosnowski, and A. Ławrynowicz, "TASTEset -- Recipe Dataset and Food Entities Recognition Benchmark", *arXiv preprint arXiv:2204.07775*, 2022, doi: 10.48550/arXiv.2204.07775.

# 공공데이터 API를 활용한 복약 도움 애플리케이션

정미리\*, 정민경\*, 심우진\*, 김황래\*\*

## Application For Medication Assistance Using Public Data API

Miri Jeong\*, Mingyeong Jeong\*, Woojin Sim\*, and Hwangrae Kim\*\*

### 요 약

현대인에게 약이란 필수 불가분의 관계이지만, 올바르게 사용해야만 최고의 효과를 기대할 수 있다. 그러나 노년층처럼 다양한 약물을 오랜 기간 복용해야 하는 사람들은 이러한 다약제 복용으로 인한 문제가 나타날 수 있고, 의약품의 정보나 그 정보를 찾기 위한 방법을 알기 어렵다. 본 논문에서는 이러한 문제점 해결을 위한 복약 도움 애플리케이션 서비스를 구현한다.

### Abstract

For modern people, medications are inseparable, but the best effects can only be expected when used correctly. However, those who need to take various medications over a long period, such as the elderly, may encounter issues due to poly-pharmacy, and it can be challenging to access drug information or know how to find it. In this paper, we implement a medication assistance application service to address these issues.

### Key words

medication assistance, medication management, drug information, digital alienation, application

## 1. 서 론

현대인에게 약이란 필수 불가분의 관계이다. 그러나 이런 약은 이해(利害)라는 양면성을 가지고 있어 올바르게 사용한다면 최고의 효과를 기대할 수 있지만, 잘못 사용한다면 독이 될 수 있다.[1]

한 종류의 약만을 복용하거나 약 복용 기간이 짧은 사람들은 이러한 오남용으로 인한 약의 부작용을 경험할 가능성이 적다. 하지만 다양한 약을 오랜 기간 복용하는 노년층의 경우는 다르다. 한국보건연

구원의 연구에 따르면 다약제 복용자의 비율이 2015년 31.6%에서 2021년 35.4%로 증가하였으며, 10종류 이상의 약을 복용하는 사람도 같은 추세를 보였다.[2] 또한 이러한 다약제의 문제점은 노년층 뿐만 아니라 만성 질환 등의 다양한 약물을 장기적으로 복용해야 하는 이들에게서도 나타날 수 있다.

또한 지난 2월부터 시작된 의사 파업이 장기화되며 진료를 받거나 전문적인 의약품 정보를 제공받기 어려워졌다. 이는 다양한 웹사이트나 애플리케이션 등을 활용하여 스스로 정보를 찾을 수 있는 사

\* 국립공주대학교 소프트웨어학과, d.2mension78@gmail.com, jmingyung1024@smail.kongju.ac.kr, ssqwe2@gmail.com

\*\* 국립공주대학교 소프트웨어학과, plusone@kongju.ac.kr

람들에게는 큰 문제가 아닐지도 모르나, 앞서 말한 노년층을 포함한 디지털 취약계층에는 중대한 문제가 될 수 있다.

따라서 본 연구에서는 의약품 정보 검색 및 복용 알람을 통해 사용자의 건강관리를 도울 수 있는 애플리케이션 개발을 목표로 한다. 더불어 해당 서비스의 주요 대상을 디지털 취약계층으로 하고, 다양한 애플리케이션을 사용하기 어려운 점을 고려하여 사용자의 위치를 기반으로 한 약국 검색, 응급처치 기능 검색 등을 지원한다.

## II. 시스템 설계

기존에도 복약 시간을 안내하는 애플리케이션 자체는 존재해 왔지만, 사용자가 알람을 설정하거나 원하는 의약품의 정보를 찾는 데까지 필요한 버튼 클릭 횟수(depth)가 높아 사용자의 피로도를 높이고 직관적이지 못하다는 단점이 있었다.

또한 의약품에 대한 전문적인 정보를 얻을 수 있는 약국, 응급처치 등 복약 이외의 정보를 얻기 위해서는 타 애플리케이션을 추가로 사용해야 한다.

이러한 문제점들은 본 연구에서 주요 대상으로 하는 디지털 취약계층에는 치명적이라고 생각하였고, 이 모든 서비스(복약 알람, 의약품·주변 약국·응급처치 정보 제공)를 하나의 애플리케이션에서 제공하는 서비스를 구상하였다. 더불어 UI 디자인을 개선하여 사용자가 직관적으로 이해할 수 있도록 그림 1.과 같이 설계하였다.

## III. 시스템 구현

본 시스템은 다양한 편의 서비스에 언제 어디서나 빠르게 접근할 수 있도록 TypeScript, React Native를 이용한 모바일 애플리케이션으로 구현되었다. React Native는 하기 기술할 외부 API와의 통신을 통해 필요한 정보를 받아와 출력하고, Expo를 이용해 이러한 코드들을 Android, iOS 등의 운영체제에서 실행한다.

### (1) API (Application Programming Interface)

API란 서로 다른 소프트웨어 간의 통신을 가능하게 하는 인터페이스로, 필요한 정보를 가진 서버에 요청을 보내고 응답을 통해 정보를 받아온다. API는 내부 구현을 추상화하여 개발의 복잡성을 줄이고 재사용성과 유지보수성을 높일 수 있다. 다양한 시스템 및 기기종 간의 상호작용을 가능하게 한다는 점에서 분산 시스템 개발 환경에서는 핵심적인 기술이다.

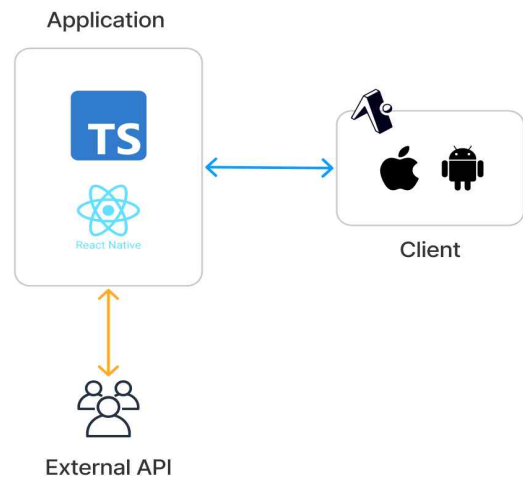


그림 1 시스템 구조도

Fig. 1. System Structure Diagram

본 연구에서 API는 의약품 공공데이터[3]를 가져오는 데에 사용된다. 사용자로부터 의약품 이름을 입력받아 서버 측에 보내고 결과를 반환받는다.

### (2) GPS (Global Positioning System)

GPS란 인공위성을 이용한 위치 파악 시스템으로, 여러 위성으로부터 수신한 신호를 바탕으로 사용자의 위도·경도·고도의 위치 정보를 계산한다. GPS는 군사·교통·지도 제작 등 다양한 분야에 활용되며, 최근 스마트폰의 보급으로 사용자의 실시간 위치 추적 및 목적지까지의 길 안내에도 중요하게 사용된다.

본 연구에서는 사용자의 위도·경도 데이터를 기존 약국 데이터(참고문헌)와 비교하여 일정 거리( $\pm 0.005$ ) 이내에 있는 데이터를 필터링하여 출력하는 데에 사용되었다.

### (3) UI (User Interface)

UI란 사용자가 애플리케이션을 사용하려고 할 때 사용자가 직접적으로 상호작용하게 되는 버튼, 아이콘 등의 인터페이스들의 구성을 의미하며, 본 시스템에서는 그림 2와 같이 디자인되었다.

본 연구의 주요 대상인 디지털 취약계층을 위해 서울디지털재단에서 제공하는 고령층 친화 디지털 접근성 표준을 고려하여 크면서도 정보 전달이 될 수 있는 글씨 및 굵기를 설정하였고, 직관적으로 이해가 가능한 아이콘을 사용하였다. 또한 기능 실행까지 사용자가 수행해야 하는 클릭의 수(depth)를 줄이고 다른 기능으로의 이동 방법을 하나의 버튼으로 고정해 가시적이고 신속하게 조작할 수 있도록 하였다.

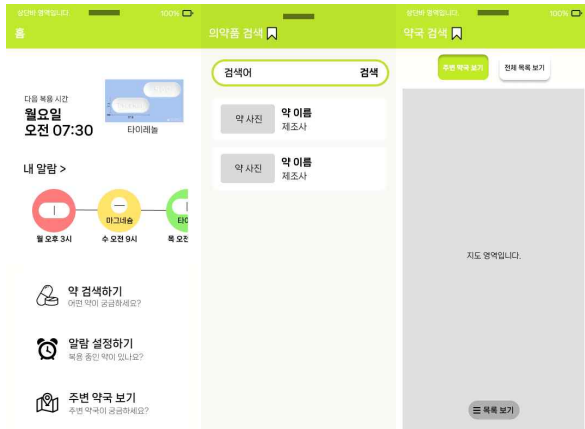


그림 2 UI 디자인  
Fig. 2. UI Design

## IV. 결 론

본 연구는 사용자의 약 복용을 돕는 기능과 함께 다양한 애플리케이션 사용이 어려운 디지털 취약계층을 위해 주변 약국 등의 서비스가 포함된 복약지도 애플리케이션을 설계 및 구현하였다. 이는 사용자 스스로 하여금 자신의 건강 상태를 체계적으로 관리하고 약물 오남용으로 인한 부작용을 예방할 수 있을 것이라 기대한다.

또한 해당 서비스는 상기된 기술적인 내용 이외에도 크고 직관적인 UI 구성 등으로 주요 대상층에 더 나은 서비스를 제공할 수 있을 것으로 보인다. 더 나아가 사용자에게 있어 디지털 애플리케이션

사용에 대한 긍정적인 경험을 가능하게 하여 타 애플리케이션 사용에 대한 두려움을 줄이고, 디지털 정보 격차 해소에 기여하여 공평한 의료 복지 기회를 제공할 수 있을 것으로 예상된다.

## 참 고 문 헌

- [1] Ministry of Food and Drug Safety. (n.d.). What is a drug? National Drug Information System. Retrieved October 15, 2024, from <https://nedrug.mfds.go.kr/cntnts/34>
- [2] Kim, S., & Yoon, J. (2023). Analysis of causes of polypharmacy prescriptions and consumption among the elderly and a review of behavioral economic alternatives. Korea Institute for Health and Social Affairs.
- [3] Ministry of Food and Drug Safety. (2024). Overview of drug information (e-Drug Info). Public Data Portal. Retrieved October 15, 2024, from <https://www.data.go.kr/data/15075057/openapi.do>
- [4] Korea Data Agency. (n.d.). API service for public data. Public Data Portal. Retrieved October 15, 2024, from <https://www.data.go.kr/iim/api/selectAPIAccountView.do>
- [5] Seoul Digital Foundation. (2020). Digital accessibility standards for the elderly.



# 프롬프트 엔지니어링을 통한 LLM 기반 QA 시스템의 안정성 및 응답 일관성 향상

장민규\*, 최상민\*\*

## Enhancing Reliability and Response Consistency of LLM-Based QA Systems through Prompt Engineering

Min-Gyu Jang\* and Sang-Min Choi\*\*

### 요 약

본 논문에서는 LLM(Large Language Models)에 존재하는 신뢰성 문제와 답변 일관성의 문제를 해결하기 위한 방안을 제시한다. LLM은 다양한 영역의 QA(Question Answering) 과제에서 뛰어난 성능을 보이지만, 사전 학습된 정보에 의존하기 때문에 최신 정보를 반영하거나 그 근거를 검증하는 데 한계가 있다. 이러한 한계는 LLM 기반 QA 시스템의 신뢰성을 저해하여 사용자의 모델 활용과 검증에 어려움을 겪게 만든다. 본 연구는 Prompt Engineering 기법의 적용을 통해 답변의 근거를 제공하고 일관성을 높여 답변의 검증 및 활용 가능성을 향상시킨다. 이 접근 방식은 QA과제(BoolQ)에서 모델의 응답 일관성을 높이며, 단일 질문 Prompt보다 4.67%(F1-Score)의 성능을 향상시킨다.

### Abstract

This paper presents a solution to address the issues of reliability and response consistency in large language models (LLMs). While LLMs have exhibited remarkable proficiency in numerous question-answering (QA) tasks, they are constrained in their capacity to reflect current information and verify evidence, given their reliance on pre-trained knowledge. These limitations render LLM-based QA systems unreliable, thereby impeding users' ability to validate and utilize the model effectively. This study employs prompt engineering techniques to provide supporting evidence for answers and enhance consistency, thereby improving the reliability and usability of responses. The proposed approach enhances the consistency of responses in question-answering tasks (such as BoolQ) and achieves a 4.67% improvement in performance (F1-Score) compared to single-question prompts.

### Key words

LLM(Large Language Model), QA(Question Answering), BoolQ(Boolean Question), Prompt Engineering, Response Consistency and Reliability

---

\*경상국립대학교 학사과정, alsrb9936@gnu.ac.kr

\*\*경상국립대학교 조교수, jerassi@gnu.ac.kr (교신저자)

※ 본 논문은 교육부와 한국연구재단의 재원으로 지원 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0), 정부(과학기술정보통신부) 재원의 한국연구재단 지원(RS-2022-00165785)과 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다(2021RIS-003).

## I. 서론

LLM(Large Language Model)의 발전은 QA(Question Answering)를 포함한 다양한 NLP 과제에서 상당한 진전을 보여주었다. 그리고 여러 도메인 영역의 QA 과제에서 뛰어난 성과를 보여준다 [1]. QA는 입력 쿼리에 대해 정해진 올바른 답변을 해야 하는 작업으로 입력 쿼리에 대해 관련 정보를 찾아주는 검색 작업과 유사하다.

하지만 LLM의 발전에도 불구하고 답변의 일관성과 신뢰성 문제는 여전히 남아있다. LLM은 사전 학습된 정보에 의존하기 때문에 최신 정보 반영에 어려움이 존재하며, 잘못된 정보를 정답으로 말하는 Hallucination 문제[2]가 발생할 수 있다. 또한 동일한 문제에 대해 답변의 일관성이 떨어지는 문제[3]가 있다. 이러한 문제들은 LLM 기반 QA 시스템의 신뢰성을 저해하고 답변에 대한 검증 과정을 복잡하게 한다[3].

본 논문에서는 답변의 일관성과 신뢰성 문제 해결을 위해 Prompt Engineering 기법을 적용하였다. Prompt Engineering은 LLM을 효과적으로 활용하기 위해 Prompt를 개발하고 최적화하는 방법을 의미한다[4]. 본 논문에 적용된 Prompt Engineering은 답변 근거를 요구하는 Prompt를 추가하여 일관성과 신뢰성을 향상시킨다. 또한 출력된 답변 근거는 QA 과제의 검증과 성능 향상을 위한 추가적인 과정에 이용할 수 있어 답변의 검증 및 활용 가능성을 향상시킨다[5]. 본 논문의 Prompt Engineering은 BoolQ(Boolean Questions)데이터 셋에서 성능 향상을 보여준다.

## II. 제안 방식

### 2.1 실험 데이터 설명

Table 1. 데이터 셋 구조 및 개수

Question	Answer	Passage
Do iran and afghanistan speak the same language?	True	Persian (/pɜːˈrʃən, -ˈʃɛn/, also known by its endonym Farsi (فارسی fārsī (fɑːrˈsiː) (listen) .....)
Total:15,942 Train:9,430 Validation: 3,270		

실험에 사용된 BoolQ[6] 데이터 구조는 Table 1과 같다. Question, Answer, Passage로 구성되며 Question은 질문, Answer는 True, False 답변, Passage는 Question과 관련된 정보다. 사용한 데이터 질문 수는 전체 데이터 15,942개 중, 3,270개의 Validation 데이터를 사용하였다.

### 2.2 Prompt 방식 설명

총 3개의 Prompt를 설계하였다. 모두 동일한 질문과 답변을 명시적인 예시로 하여 One-shot prompting을 적용하였다. Prompt-1은 본 논문의 Prompt Engineering이 적용되지 않은 Prompt이다. Prompt-2는 답변과 답변 근거를 요구하는 Prompt Engineering이 적용되었으며 Prompt-3은 Prompt-2와 Prompt Engineering이 적용되었지만, 요구하는 답변 형식이 다르다. 답변의 정확성을 향상 시키는 답변 형식을 찾기 위해 순서를 다르게 설계하였다. 자세한 Prompt는 Table 2에서 확인 할 수 있다.

Table 2. Prompt 방식

종류	System	User	Example
Prompt-1	You are a system that answers a given question.	Answers should be 'True', 'False' only. If correct, write 'True', if incorrect, write 'False'. Please use the following example to answer the question	<b>Q:</b> will there be a season 5 of shadowhunters <b>A:</b> False
Prompt-2	You are a system that answers a given question. And explain why you answered that way.	Answers should be 'True', 'False' only. If correct, write 'True', if incorrect, write 'False'. Please use the following example to answer the question	<b>Q:</b> will there be a season 5 of shadowhunters <b>A:</b> False <b>E:</b> The television series "Shadowhunters" was officially canceled after its third season. There are no plans for a season 5, as the show concluded with a two-part finale that aired in 2019.
Prompt-3	You are a system that answers a given question. And explain why you answered that way.	Answers should be 'True', 'False' only. If correct, write 'True', if incorrect, write 'False'. Please use the following example to answer the question	<b>Q:</b> will there be a season 5 of shadowhunters <b>E:</b> The television series "Shadowhunters" was officially canceled after its third season. There are no plans for a season 5, as the show concluded with a two-part finale that aired in 2019. <b>A:</b> False

### III. 실험 및 결과

#### 3.1 실험 계획

실험은 ChatGPT-4o-mini를 이용하여 prompt case 3가지에 따라 각 15회씩, 총 45회 실험을 진행한다. 추가로 타 모델에서의 효과를 확인하기 위해 ChatGPT-3.5-turbo를 사용하여 실험하였고, 평가 방법은 모델이 정답('True', 'False')을 정확히 예측했는지를 accuracy, Precision, Recall, F1-Score를 사용하여 평가한다[7].

#### 3.2 실험 결과

Table 3. 실험1 결과 (gpt-4o-mini, 단위 %)

종류	Acc	Precision	Recall	F1
Prompt-1-avg	70.0543	79.0541	71.8328	74.7393
<b>Prompt-2-avg</b>	73.3192	79.7568	77.3823	<b>78.2330</b>
Prompt-3-avg	72.4305	81.9436	72.0347	76.3640
Prompt-1-var	0.0416	0.1964	0.3232	0.0492
<b>Prompt-2-var</b>	0.0225	0.2090	0.3061	<b>0.0063</b>
Prompt-3-var	0.0829	0.2709	0.5983	0.0731

Table 2는 Prompt-1, 2, 3을 각 15회씩 실험한 결과의 평균과 분산을 보여준다. 실험 결과 Prompt-2가 F1-Score 기준으로 가장 높았으며, Prompt Engineering을 적용하지 않은 Prompt-1보다 4.67% 향상되었다. 또한 가장 낮은 분산을 달성하였다. 이를 통해 Prompt-2와 같이 답변의 근거를 추가한 Prompt가 답변의 정확성을 향상시킬 뿐만 아니라 답변의 일관성 또한 높이는 것을 확인할 수 있다. 그리고 Prompt-3의 결과로 답변의 형식은 정답, 근거의 순서 형식이 더 좋은 결과를 도출하는 것을 확인할 수 있다.

Table 4. 실험2 결과 (gpt-3.5-turbo, 단위 %)

종류	Acc	Precision	Recall	F1
Prompt-1	49.7961	68.2803	35.3096	46.5480
<b>Prompt-2</b>	74.3203	82.1284	74.9340	<b>78.3664</b>
Prompt-3	71.4709	80.4425	71.4285	75.6680

gpt-4o-mini뿐만 아니라 타 모델에서의 효과도 확인하기 위해 gpt-3.5-turbo에서도 실험을 진행해 결

과를 확인한다. 결과는 Table 3에서 확인할 수 있으며 실험1과 동일하게 Prompt-2가 성능을 가장 높은 쪽으로 향상시키는 것을 확인할 수 있다.

### IV. 결론

본 논문에서는 LLM을 이용한 QA 과제에서 질문에 대한 답변뿐만 아니라 근거를 함께 요청하는 것이 성능 향상과 일관성 개선에 효과적임을 입증하였다. 이 접근법은 모델 출력에 대한 일관성을 높이며, 근거를 통해 답변의 신뢰성을 제공한다.

향후 연구에서는 다양한 QA 데이터셋을 활용하여 본 연구에서 제시한 방법론의 일반화 가능성과 성능을 종합적으로 평가할 것이다. 또한 모델이 생성한 근거 정보를 추가적인 학습 및 검증 과정에 통합함으로써, 성능 향상에 기여할 수 있는지에 대한 분석을 수행할 계획이다.

### 참고 문헌

- [1] Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Naveed Akhtar, Nick Barnes, and Ajmal Mian. "A comprehensive overview of large language models." arXiv preprint arXiv:2307.06435, 2023.
- [2] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. "Survey of hallucination in natural language generation." ACM Computing Surveys, 55(12):1-38, 2023
- [3] Jingyuan Yang, Dapeng Chen, Yajing Sun, Rongjun Li, Zhiyong Feng, and Wei Peng. "Enhancing Semantic Consistency of Large Language Models through Model Editing: An Interpretability-Oriented Approach", Findings-ACL , pp. 3343-3353, 2024.
- [4] Louie Giray. Prompt engineering with chatgpt: a guide for academic writers. Annals of biomedical engineering, 51(12):2629-2633, 2023.

- [5] Shehzaad Dhuliawala, Mojtaba Komeili, Jing Xu, Roberta Raileanu, Xian Li, Asli Celikyilmaz, and Jason Weston. “Chain-of-verification reduces hallucination in large language models.” arXiv preprint arXiv:2309.11495, 2023.
- [6] Christopher Clark, Kenton Lee, Ming-Wei Chang, Tom Kwiatkowski, Michael Collins, and Kristina Toutanova. “BoolQ: Exploring the surprising difficulty of natural yes/no questions. ” Association for Computational Linguistics, pages 2924-2936, 2019.
- [7] David MW Powers. “Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation.” arXiv preprint arXiv:2010.16061, 2020.

# 소프트웨어 프로젝트 성과 분석을 위한 LLM 기반 챗봇 개발

박승민\*, 조수희\*, 권령구\*, 권기현\*, 이창을\*\*

## LLM-Based Chatbot for Software Project Performance Analysis

Seungmin Park\*, Suhee Jo\*, Ryeonggu Kwon\*, Gihwon Kwon\*, Changyul Lee\*\*

### 요약

본 연구는 프로젝트 성과를 분석하고 관리 효율성을 높이기 위한 LLM 기반 챗봇을 개발한다. 챗봇은 데이터를 기반으로 팀 성과와 개인 기여도를 체계적으로 평가하고, 개선방안을 제시하여 프로젝트 성공 가능성을 높인다. 프롬프트를 통해 실험한 결과, 챗봇을 통해 효율적인 프로젝트 데이터 분석이 가능했고, 실제 데이터 및 평가표와의 비교를 통해 챗봇을 통한 성과 평가의 신뢰성을 검증하였다.

### Abstract

This study develops an LLM-based chatbot to improve project performance analysis and management efficiency. The chatbot uses data to systematically evaluate team and individual contributions, providing suggestions for improvement to increase project success. Testing showed that the chatbot can analyze project data effectively, and comparisons with real data confirmed the reliability of its performance evaluations.

### Key words

LLM, Chatbot, Scrum, Project Management

## I. 서론

최근 소프트웨어 개발 환경은 복잡한 요구사항과 빠른 기술 변화로 인해 예측이 어려워졌으며, 이에 따라 프로젝트 관리의 중요성이 커지고 있다[1][2]. 효과적인 프로젝트 성과 관리는 팀 협업을 촉진하고 개인 기여도를 파악해 프로젝트 효율성을 높이

는 데 필수적이다[3]. 본 연구는 경기대학교 학부생 프로젝트를 애자일 기반으로 경량화된 K-Scrum 프로세스를 통해서 얻어진 K-Scrum 데이터를 활용한다. K-Scrum 데이터를 분석하여 팀 성과 및 개인 기여도를 평가하고, 자원 배분과 성과 개선을 위한 구체적인 해결책을 제시하도록 챗봇을 설계하는 것을 목표로 한다.

\* 경기대학교 AI컴퓨터공학부,  
psm0811@kyonggi.ac.kr(교신 저자), sujo@kyonggi.ac.kr, rkkwon@kyonggi.ac.kr, khkwon@kyonggi.ac.kr  
\*\* (주)에스피아이디, yulee@espid.com

※ 본 연구는 (주)에스피아이디의 지원을 받아 수행된 과제번호 2024-0316, "LLM 기반 컨설팅 지식 관리 및 의사결정 지원 시스템 구축"의 결과입니다.

기존 프로젝트 관리 도구(Jira, Trello 등)는 널리 사용되고 있으나 한계가 있다[4][5]. 첫째, 데이터 분석이 복잡해 수작업이 필요하며 시간과 자원 낭비를 초래한다. 둘째, 문제 감지 및 의사결정 지원 기능이 부족하다. 마지막으로, 팀원 교체 시 발생하는 지식 손실 문제를 해결하지 못해 프로젝트 연속성에 부정적 영향을 미친다[6].

AI와 챗봇을 활용한 프로젝트 관리 연구는 여러 산업에서 AI 기반 챗봇이 업무 효율성과 프로젝트 관리에 긍정적 영향을 줄 수 있음을 보여주었다 [7][8]. 하지만, 실제 프로젝트 데이터를 통해 AI 챗봇을 활용해 성과 평가와 구체적 개선 방안을 제시한 연구는 부족하다. 챗봇 기반 프로젝트 관리 도구는 기존 도구의 한계를 보완해 프로젝트 관리의 효율성을 높이는 유망한 솔루션으로 평가된다[9]. 챗봇은 실시간 데이터 분석과 자동 보고로 관리자가 수동 처리할 필요를 줄이고, 문제를 조기에 감지해 해결 방안을 제시함으로써 프로젝트 연속성과 효율성을 높인다. 또한 팀원의 업무 지식과 경험을 기록해 지식 손실을 방지하고, 인수인계를 원활하게 하며, 주요 일정과 작업 현황을 실시간으로 업데이트해 협업과 신속한 의사결정을 지원한다.

본 연구는 K-Scrum 데이터를 활용하여 팀과 개인 성과를 체계적으로 분석하고 실질적 개선안을 제공하는 소프트웨어 프로젝트 관리 모델을 제안해 기존 연구와 차별화된다. 이를 통해 데이터 분석, 문제 해결, 지식 관리, 협업 지원이 통합된 효율적 관리와 성과 개선을 기대할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 K-Scrum 데이터와 챗봇의 설계 및 구현 과정을 다루고, 챗봇 성능 평가에 사용된 프롬프트를 설명한다. 3장에서는 챗봇 성능 평가 결과를 성과 평가표와 비교하여 논의하고, 4장에서는 연구의 요약 및 결론과 함께 향후 연구 방향을 제시한다.

## II. 연구 방법론

본 연구에서는 OpenAI의 기능을 활용하여 K-Scrum 데이터를 분석하고, 팀 프로젝트 성과 평가를 지원하는 맞춤형 챗봇을 그림 1처럼 설계하였

다. 챗봇의 역할과 데이터 분석 방법을 구체적인 프롬프트로 설정하여, 데이터를 정확하게 해석하고 사용자 요구에 맞는 결과를 도출할 수 있도록 하였다. 또한, 추가 데이터셋을 제공함으로써 챗봇이 더욱 정교한 분석을 수행할 수 있게 설정하였다. 프롬프트는 챗봇이 데이터 처리와 분석 과정을 지시하는 핵심 요소로, 프롬프트가 구체적일수록 챗봇의 응답 정확성이 높아진다. 이렇게 생성된 챗봇은 설정된 프롬프트와 데이터셋을 바탕으로 팀 성과와 관련된 질문에 대해 적절한 응답을 제공한다.

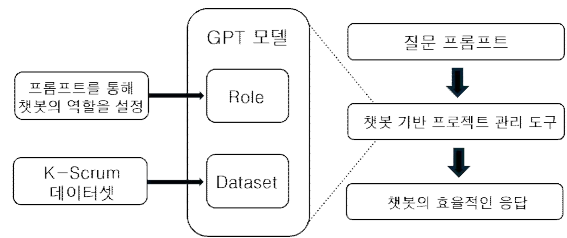


그림 1. 챗봇 흐름도

Fig. 1. Flow Diagram of Chatbot

본 연구는 경기대학교 학부생 프로젝트를 애자일 방식으로 경량화한 K-Scrum 프로세스를 통해 생성된 K-Scrum 데이터셋을 활용하여 소프트웨어 개발 프로젝트의 성과를 분석한다[10]. 이 데이터셋은 프로젝트 백로그 DB, 일감 DB로 구성되며, 그림 2에서 각 DB별 주요 데이터 항목을 확인할 수 있다. 프로젝트 백로그에는 사용자의 요구사항인 사용자 스토리를 통해 정의된 기능과 요구사항이, 일감 DB에는 해당 기능을 구체화한 작업 단위가 포함된다. 이를 통해 각 팀의 스프린트 진행률과 작업량을 모니터링하고, 팀원별 기여도를 평가할 수 있다.

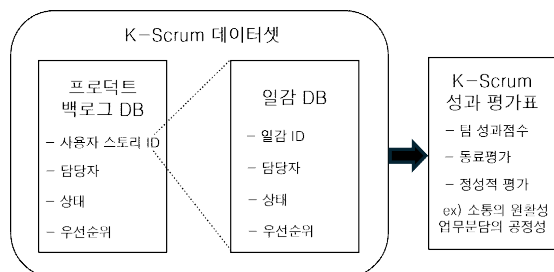


그림 2. K-Scrum 데이터

Fig. 2. K-Scrum Data

K-Scrum 성과 평가표는 학부생 프로젝트의 성과를 정량적 및 정성적 항목으로 평가한 데이터이다. 팀 점수 같은 정량적 평가 항목과 동료 평가, 팀 내 소통, 업무 분배 등의 정성적 항목으로 구성된다. 정량적 평가는 100점 만점으로 평가자가 직접 점수를 부여한 결과이다. 동료 평가는 팀원 간 상호 평가 방식으로, 우수한 팀원에게는 +1점을, 개선이 필요한 팀원에게는 -1점을 부여해 -4에서 4점의 범위를 가진다. 정성적 평가는 팀원이 항목에 대해 자신의 의견을 1-5점으로 평가한 것으로, 높은 점수일수록 긍정적인 평가이다. 이 평가표는 사람이 직접 평가한 데이터를 통해 프로젝트 성과의 전반적 경향을 파악하는 데 유용하다. 표 1과 표 2는 프로젝트 백로그 DB, 일감 DB, 그리고 K-Scrum 성과 평가표의 일부 데이터의 예시이다.

[표 1] 프로젝트 백로그 DB, 일감 DB 예시  
[Table 1] Example of Product Backlog DB, Task DB

프로젝트 DB			
사용자 스토리	담당자	상태	우선순위
영화 추천 기능 구현	A학생	완료	높음
영화 리뷰 기능 구현	B학생	완료	낮음
일감 DB			
일감	담당자	상태	우선순위
웹페이지 CSS 작성	A학생	완료	중간
Login.php 구현	C학생	완료	높음

[표 2] K-Scrum 성과 평가표 예시  
[Table 2] Example of K-Scrum Performance Evaluation

이름	팀점수	동료평가	소통의 원활성	역할분담의 적절성
A	65.98	4	3	2
B	65.98	1	2	2
D	74.34	-3	4	4

본 연구에서는 이를 기준으로 챗봇의 성과 분석 결과와 평가 데이터 간의 유사성을 비교하였다. 다음은 데이터를 기반으로 팀의 성과 및 역할 분담, 개인의 기여도를 평가하고 향후 개선방안을 도출하

기 위해 프롬프트를 설계하여 정리한 것이다.

- 전체 팀 성과 비교 : 모든 팀의 데이터를 종합적으로 분석하여 각 팀별 성과를 계산하여 순위를 나열해보자
- 팀별 역할 분담 평가 : 모든 팀 내에서 업무가 공평한 분배와 협업과 소통이 잘 이루어졌는지를 팀끼리의 순위를 매겨보자
- 개인 기여도 평가 : 데이터에 있는 개인별 작업량, 우선순위 등을 고려하여 최하위 성과 팀에서의 개인의 기여도 순위를 매겨보자
- 팀 성과 비교 : 최상위 팀과 최하위 팀 간의 주요 성과 지표의 차이를 데이터로 분석하여 그 차이를 구체적으로 나타내보자
- 해결책 및 개선 방안 : 다른 팀과 비교했을 때, 최하위 성과 팀이 더욱 우수한 프로젝트가 되려면 어떻게 개선해야 하는지 서술하시오

### III. 결과 분석

챗봇의 성능을 평가하기 위해 앞에서 정의한 프롬프트를 기반으로 챗봇에게 질의하여 팀 프로젝트에 관련된 정보를 분석하도록 하였다.

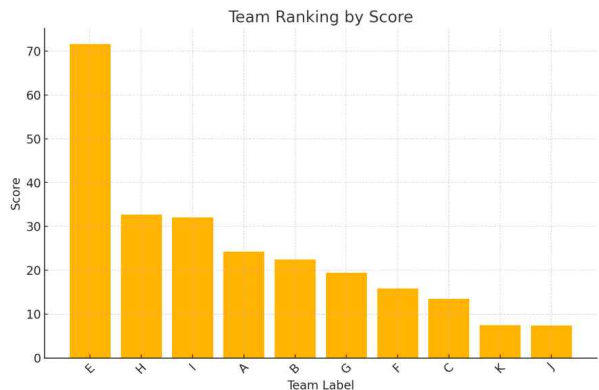


그림 3. 전체 팀 성과 순위표  
Fig. 3. Team Ranking by Overall Score

그림 3은 챗봇이 데이터를 분석하여 응답한 그래프로, 전체 팀의 성과를 점수화하고 이를 순위로 시각화한 것이다. 챗봇의 분석 결과, 팀 E가 가장 높은 성과를 기록한 반면, 팀 K와 J가 가장 낮은 성과를 기록했다. K-Scrum 성과표 데이터에서도 팀 E의 팀 점수가 85.1점으로, 다른 팀에 비해 압도적

으로 높았다. 반대로 팀 K가 56.7점, 팀 J가 61.88점으로 하위권에 위치하였다. 이러한 결과는 챗봇의 정량적 데이터 분석이 성과 평가에 신뢰할 만한 정보를 제공할 수 있음을 보여준다.

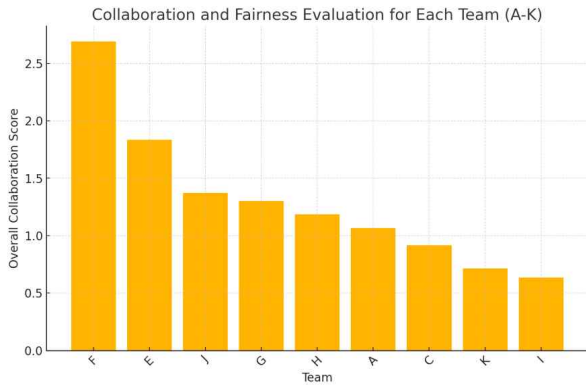


그림 4. 전체 팀 역할 분담 평가 순위  
Fig. 4. Team Ranking by Fairness Evaluation

그림 4는 챗봇이 전체 팀의 데이터를 분석하여, 역할 분담의 적절성과 팀 내 협업 및 소통 수준을 순위로 나타낸 그래프이다. 챗봇의 분석 결과, 팀 F의 협업과 소통 순위가 가장 높았고, 팀 I의 순위가 가장 낮았다. 그러나 실제 K-Scrum 평가표에서는 팀 F가 중위권에 머물렀고, 가장 높은 순위를 기록한 팀은 챗봇이 두 번째로 예상한 팀 E였다. 이는 챗봇의 정성적 요소 분석에 개선이 필요함을 보여주지만, 해당 항목이 팀원들이 주관적으로 점수를 매겨 평가한 것이므로 성과표 자체의 정확성에도 한계가 있을 수 있다.

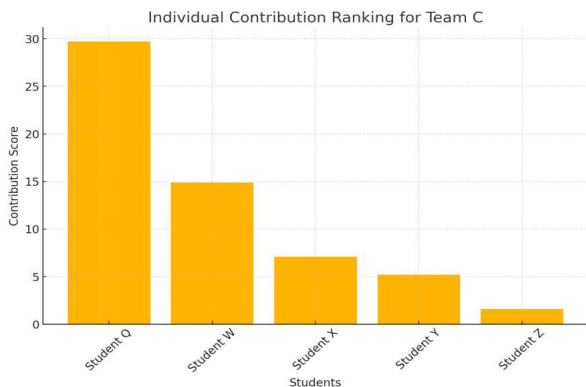


그림 5. 팀 C 내의 개인별 성과 순위  
Fig. 5. Individual Contribution Ranking for Team C

그림 5는 챗봇이 최하위 성과 팀에서의 개인 기여도를 분석하여 순위를 나타낸 그래프이다. 앞서

그림 3에서 챗봇은 팀 K와 팀 J를 최하위 성과 팀으로 분석했지만, 이들 팀은 프로젝트 진행 데이터가 충분하지 않아 개인 기여도 분석에서는 제외되었다. K-Scrum 평가표와 비교한 결과, 학생 Q는 동료 평가에서 4점을 받아 모든 팀원으로부터 가장 우수한 팀원으로 선정되었다. 반면, 학생 Z는 동료 평가에서 -2점을 기록하여 개선이 필요한 팀원으로 평가되었다. 실제 동료 평가에서는 학생 Y가 1점을 받아 2위에 올랐으나, 전체적인 순위 경향성은 챗봇 분석과 유사한 것으로 나타났다. 이는 챗봇이 모든 요소를 고려하여 완벽히 평가하지는 못하지만, 대략적인 경향성을 파악하는 데 유용함을 보여준다.

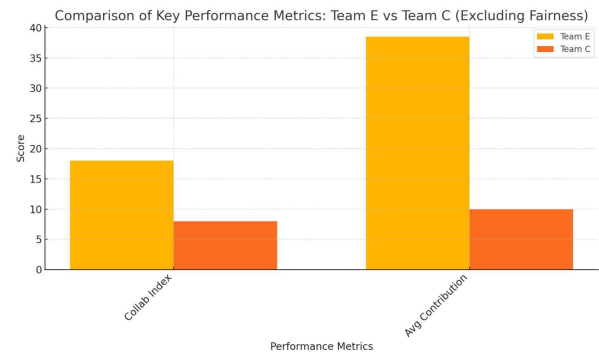


그림 6. 최고 성과 팀과 최하위 성과팀의 차이  
Fig. 6. Top vs. Lowest Performing Team Difference

그림 6은 팀 E와 팀 C의 협업 빈도(Collab Index)와 평균 기여도(Avg Contribution)를 비교한 것이다. 협업 빈도에서 팀 E는 팀 C보다 높은 점수를 기록해 업무에 대한 협력이 더 활발했음을 보여주며, 평균 기여도에서도 팀 E가 팀 C를 크게 앞서 팀원들의 기여도가 더 높았음을 시사한다. 이를 통해 팀 E가 팀 C보다 협업을 통해 팀원들의 평균 기여도를 끌어올려 우수한 성과를 보였음을 알 수 있다.

다음은 챗봇이 최하위 팀으로 분석한 팀 C가 더욱 우수한 프로젝트로 발전하기 위해 제시한 개선 방안이다. 챗봇은 팀 C의 일부 팀원에게 과도한 업무가 집중되는 문제를 해결하기 위해 작업량을 균형 있게 조정하고, 새로운 작업을 할당할 때 기존 작업량을 고려할 것을 권장한다. 또한, 협업이 필요한 작업은 여러 팀원이 함께 수행할 수 있도록 배치하며, 중요한 작업에 우선순위를 두어 신속히 완료하는 것이 성과 향상에 도움이 될 것이라고 조언



한다. 실제로 최고 성과 팀인 팀 E와의 비교에서, 팀 C는 협업 빈도와 평균 기여도에서 큰 차이를 보였다. 이는 챗봇이 단순한 데이터 분석을 넘어 실질적인 프로젝트 관리 지원 도구로서의 잠재력을 갖추고 있음을 보여준다.

#### IV. 결 론

본 연구는 K-Scrum 데이터를 활용해 팀 성과를 분석하고 프로젝트 관리 효율성을 높이기 위한 챗봇을 개발하는 것을 목표로 하였다. 실험 결과, 챗봇은 단순한 데이터 추출을 넘어 각 팀의 상황에 맞는 개선 방안과 해결책을 구체적으로 제시하는데 유용한 도구로 작동하였다. 이를 통해 기존의 수동적 성과 분석 방식에서 벗어나, 데이터 기반의 자동화된 의사결정 지원 도구로서 챗봇의 가능성을 실증적으로 확인할 수 있었다.

챗봇은 성과 평가와 문제 진단에 있어 높은 유사성을 보였으나, 일부 정성적 평가 항목에서 심층적 분석이 부족하고 맞춤형 해결책 제시가 다소 단순한 수준에 머무르는 한계를 보였다. 또한, LLM 기반 챗봇이 직면한 할루시네이션 문제는 분석 결과의 신뢰성에 영향을 준다[11]. 향후 연구에서는 RAG(Retrieval-Augmented Generation) 모델을 도입하여 실시간 외부 데이터를 참조함으로써, 챗봇의 성능과 신뢰성을 높이고, 다양한 데이터와 프로젝트 도메인에 적용하여 프로젝트 관리 도구로서의 활용도를 극대화할 수 있을 것이다.

#### 참 고 문 헌

- [1] C. W. Butler, L. R. Vijayarathy, and N. Roberts, "Managing software development projects for success: Aligning plan-and agility-based approaches to project complexity and project dynamism," *Project Management Journal*, vol. 51, no. 3, pp. 262-277, 2020.
- [2] P. Ochoa Pacheco, D. Coello-Montecel, M. Tello, V. Lasio, and A. Armijos, "How do project managers' competencies impact project success? A systematic literature review," *Plos one*, vol. 18, no. 12, p. e0295417, 2023.
- [3] V. Pantović, D. Vidojević, S. Vujičić, S. Sofijanić, and M. Jovanović-Milenković, "Data-Driven decision making for sustainable IT project management excellence," *Sustainability*, vol. 16, no. 7, p. 3014, 2024.
- [4] J. Fisher, D. Koning, and A. Ludwigsen, "Utilizing Atlassian JIRA for large-scale software development management," Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), techreport, 2013.
- [5] A. Kaur, "App Review: Trello," *Journal of Hospital Librarianship*, vol. 18, no. 1, pp. 95-101, 2018.
- [6] M. P. Robillard, "Turnover-induced knowledge loss in practice," in *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2021, pp. 1292-1302.
- [7] S. K. Devineni, "From chaos to clarity: revolutionizing industries with ai for enhanced trust, efficiency, and innovation," *Journal of Scientific and Engineering Research*, vol. 11, no. 4, pp. 116-128, 2024.
- [8] I. Taboada, A. Daneshpajouh, N. Toledo, and T. de Vass, "Artificial intelligence enabled project management: a systematic literature review," *Applied Sciences*, vol. 13, no. 8, p. 5014, 2023.
- [9] G. Calderini, S. Jaf, and K. McGarry, "A literature survey of recent advances in chatbots," *Information*, vol. 13, no. 1, p. 41, 2022.
- [10] S. Jo, S. Park, and G. Kwon, "Student-Friendly Scrum Framework for Undergraduate Education," *The Journal of Korean Institute of Information Technology*, vol. 20, pp. 199-206, 2022.
- [11] L. Huang et al., "A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions," *arXiv preprint arXiv:2311.05232*, 2023.

# LLM을 활용한 대화형 음악 추천 시스템

김기홍\*, 최상민\*\*

## Utilizing LLM for Conversational Music Recommendation System

Gi-Hong Kim\*, and Sang-Min Choi\*\*

### 요약

추천 시스템은 사용자의 선호를 파악해 개인화된 추천을 제공하는 것을 목표로 한다. 대화형 추천 시스템은 실시간으로 사용자와 상호작용을 해 보다 개인화된 추천 결과를 얻을 수 있게 한다. 최근 대형 언어 모델(LLM)의 발전으로 인해 LLM을 활용한 대화형 추천 시스템 연구가 활발히 진행되고 있다. 하지만 이러한 연구들은 주로 영화 데이터셋을 기반으로 성능을 평가한다. 본 연구에서는 LLM의 음악 추천 성능을 분석하기 위해 음악 추천 대화 데이터셋인 Conversational Playlist Curation Dataset(CPCD)를 전처리하여 GPT-4o-mini를 파인 튜닝한다. 모델의 성능 평가를 위해 테스트 데이터에 Zero-shot prompting을 적용해 모델의 입력으로 사용하였다.

### Abstract

Recommendation systems aim to understand user's preferences and provide personalized recommendations. Conversational recommendation systems allow users to interact with the system in real time to get more personalized recommendation results. Recently, research on conversational recommendation systems using Large Language Model(LLM) has been actively conducted. However, these studies typically use movie datasets to evaluate their system. In this study, to analyze the music recommendation performance of LLM, we preprocess the Conversational Playlist Curation Dataset(CPCD), a music recommendation conversation dataset, to fine-tune GPT-4o-mini. Zero-shot prompting was applied to the test data and used as input to the model for performance evaluation.

### Key words

conversational recommendation, music recommendation, large language model

## I. 서론

대화형 추천 시스템(Conversational Recommender System)은 사용자와의 대화를 통해

직접적으로 사용자 선호도에 대한 정보를 얻을 수 있다. 또한, 사용자의 실시간 피드백을 반영해 추천 결과를 개선할 수 있어 사용자에게 맞춤형 결과를 제공할 수 있다.

ChatGPT와 같은 대형 언어 모델(LLM)의 발전으로

\*경상국립대학교 컴퓨터공학과 학사과정, rlarlghd20@gnu.ac.kr

\*\*경상국립대학교 컴퓨터공학과 조교수, jerassi@gnu.ac.kr(교신저자)

로 인해 다양한 분야에서 LLM을 활용하고 있다. LLM의 대화 맥락 파악 능력과 추론 능력은 사용자의 선호도를 추론할 수 있다는 점에서 대화형 추천 시스템에도 활용되고 있다.

많은 대화형 추천 시스템은 Inspired 데이터셋, ReDIAL 데이터셋과 같이 영화 추천에 대한 데이터셋을 사용한다. 본 논문은 대화형 음악 추천을 위해 음악 추천 대화 데이터셋인 Conversational Playlist Curation Dataset(CPCD)[3]를 전처리 한 후, Gpt-4o-mini 모델을 Fine-Tuning 해 LLM의 음악 추천 성능을 확인하는 실험을 진행한다.

## II. 관련 연구

GPT와 BERT와 같은 사전 학습된 언어 모델은 음악과 관련된 충분한 정보가 있어 대화형 추천 시스템으로써의 역할을 할 수 있다[6]. 또한, LLM은 사용자의 선호를 이해할 수 있으며, 전통적인 추천 시스템과 비교했을 때 만족스러운 결과를 나타냈다[4,5]. 추가로, LLM은 추가적인 검색 시스템이나 파인튜닝 없이 zero-shot 또는 few-shot 프롬프팅으로도 충분한 추천 결과를 얻을 수 있다 [1,2,4,5]. 이를 바탕으로 본 논문에서는 추천 결과를 얻기 위해 zero-shot 프롬프팅을 적용하고, 파인 튜닝이 추천 결과에 어떤 영향을 미치는지 실험을 통해 분석한다.

## III. 연구 방법

Conversational Playlist Curation Dataset(CPCD) 데이터셋은 유저의 요구에 대하여 시스템이 그에 맞는 음악을 추천하고, 유저는 추천된 목록 중 자신이 좋아하는 노래를 플레이리스트에 담는 과정을 대화로 담았다. 해당 데이터셋은 이런 대화 데이터 917개로 이루어져 있으며, 하나의 대화 안에는 평균적으로 5.7개의 턴이 있다. 턴마다 유저는 평균적으로 4.8개의 노래에 대해서 평가를 한다. 좋아한다고 평가하는 노래와 싫어한다고 평가하는 노래의 비는 8:2이다. 데이터셋은 Turns, tracks, goal playlist 3가지 영역으로 이루어져 있다. Turns는 대화를 담고

있고, tracks는 유저가 평가한 노래에 대한 메타데이터를 담고 있고, goal playlist는 user가 선호한다고 평가한 노래를 모은 것이다. 표 1은 Turns 영역의 하나의 턴에 대한 구성요소를 나타낸다. 표 2는 데이터셋의 대화 예시를 보여준다. 공간의 제약으로 인해 두 번의 턴과 3개의 결과만 표시하였고 search queries 와 search results는 생략하였다.

필드명	내용
user query	사용자 요구사항 입력
system response	응답(추천결과 X)
search queries	시스템이 검색을 하기 위한 query
search results	검색한 결과
liked results	시스템의 검색 결과중 사용자가 좋다고 한 노래
disliked results	시스템의 검색 결과중 사용자가 싫다고 한 노래

표 1. CPCD 데이터셋의 하나의 턴에 대한 구성

user query	Definitely talking about pre trilogy Green Day
system response	Here is some green day and pearl jam
liked results	Dissident - Pearl Jam Christie Rd. - Green Day Corduroy - Pearl Jam
disliked results	Once - Pearl Jam Scattered - Green Day Welcome to Paradise - Green Day
user query	Excellent picks! How about some music from other artists like Soundgarden or Nirvana
system response	Here is some soundgarden and nirvana
liked results	Lithium - Nirvana Rusty Cage - Soundgarden Heart-Shaped Box - Nirvana
disliked results	Limo Wreck - Soundgarden Overfloater - Soundgarden Aneurysm - Nirvana

표 2. CPCD 데이터셋 대화 예시

system	You are a system that recommends music based on the conversation between the User
user	Definitely talking about pre trilogy Green Day
assistant	Here is some green day and pearl jam 1. Dissident - Pearl Jam 2. Christie Rd. - Green Day 3. Corduroy - Pearl Jam
user	Excellent picks! How about some music from other artists like Soundgarden or Nirvana
assistant	Here is some soundgarden and nirvana 1.Lithium -Nirvana 2.Rusty Cage - Soundgarden 3.Heart-Shaped Box - Nirvana

표 3. 전처리 이후 대화 예시(파인 튜닝 학습 데이터)

해당 구성은 GPT 파인튜닝과 프롬프트 입력으로 사용하기 적합한 형태로 전처리할 필요가 있다. 파인튜닝의 목적이 사용자의 입력에 대해 더 좋은 추천 결과를 얻기 위함이므로 사용자가 좋아하는 노래들인 liked results를 system response의 끝에 이어 붙였다. Search queries, search results, disliked results는 system response에 넣을 시 입력 프롬프트가 길어지고, 노이즈가 될 수 있어서 활용하지 않았다. 전처리한 결과는 표3과 같다.

#### IV. 실험 및 결과

전처리한 데이터를 GPT-4o-mini를 파인튜닝 할 때 두 가지 방법을 사용하였다. 첫 번째는 Multi-turn 형식을 유지하고 assistant 응답 중 추천 결과가 있을 때만 assistant의 응답을 학습하는 방법이다(표 4에 multi-turn 표기). 두 번째는 마지막 assistant의 응답을 제외한 대화 전체를 사용자의 입력으로 하고 그에 대한 적절한 응답으로 마지막 assistant 응답을 학습하는 방법이다(표 4에 whole로 표기).

모델의 평가를 위해 모델에 입력되는 데이터는 작업 설명과 출력 형식을 명시한 지시어와 마지막

assistant의 응답을 제외한 대화이다. 평가 지표는 recall@k를 사용하였다.

모델	Multi-turn			Whole		
	R@5	R@3	R@1	R@5	R@3	R@1
Base	0.11	0.09	0.04	0.12	0.10	0.05
Fine tuned	0.09	0.05	0.02	0.08	0.07	0.03

표 4. 실험 결과

마지막 assistant의 응답을 제외한 대화 전체를 하나의 prompt로 GPT-4o-mini에 입력했을 때 가장 높은 Recall을 보였다. 또한, 파인튜닝을 진행한 모델이 base모델보다 추천 성능이 떨어진 것을 확인할 수 있었다. 이는 모델이 파인튜닝 과정에서 데이터셋의 대화 맥락에 과적합 되었다고 할 수 있다.

#### V. 결론

본 논문은 LLM 음악 추천에 대한 성능을 평가하기 위해 CPCD 데이터셋을 전처리하여 파인 튜닝해 실험을 진행하였다. 실험 결과 파인 튜닝한 모델이 base모델보다 성능이 낮게 나온 것을 확인할 수 있었다. 추후 연구에서 CPCD데이터셋의 대화 맥락의 활용 방식을 개선해 파인튜닝 성능을 개선하도록 하겠다.

#### 참 고 문 헌

- [1]He, Zhankui, et al. "Large language models as zero-shot conversational recommenders." *Proceedings of the 32nd ACM international conference on information and knowledge management*. 2023.
- [2] Lyu, Hanjia, et al. "Llm-rec: Personalized recommendation via prompting large language models." *arXiv preprint arXiv:2307.15780* (2023).
- [3]Chaganty, Arun Tejasvi, et al. "Beyond single

- items: Exploring user preferences in item sets with the conversational playlist curation dataset." *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2023.
- [4]Sanner, Scott, et al. "Large language models are competitive near cold-start recommenders for language-and item-based preferences." *Proceedings of the 17th ACM conference on recommender systems*. 2023.
- [5]Kang, Wang-Cheng, et al. "Do llms understand user preferences? evaluating llms on user rating prediction." *arXiv preprint arXiv:2305.06474* (2023).
- [6]Penha, Gustavo, and Claudia Hauff. "What does bert know about books, movies and music? probing bert for conversational recommendation." *Proceedings of the 14th ACM conference on recommender systems*. 2020.



# GAN 알고리즘 기반 딥페이크의 동향과 긍정적 사례 분석

김태희\*, 최미령\*, 최현식\*, 최승호\*\*, 이석훈\*†

## Analysis of Deepfake Trends and Positive Cases based on GAN Algorithm

Taehee Kim\*, Miryeong Choi\*, Hyeonsik Choi\*, Seungho Choi\*\*, Sukhoon Lee\*†

### 요 약

본 연구는 GAN 알고리즘의 개요와 GAN의 딥페이크 응용과 함께 딥페이크의 긍정적인 사례를 함께 분석하고 있다. GAN은 딥러닝의 원리를 이용하여 가상의 이미지를 생성하는 알고리즘으로 이미지 생성뿐만 아니라 음성 및 언어 처리 등 다양한 분야에서 활용되고 다양한 변형 모델들로 발전해 딥페이크 생성에서 주요 기술로 자리하고 있다. 딥페이크의 긍정적 사례를 분석하여, 추후 딥페이크의 발전이 어느 분야에서 긍정적인 영향을 줄 수 있을 것인가를 예상한다.

### Abstract

This study analyzes the GAN algorithm's outline and GAN's deepfake application, as well as the positive cases of deepfakes. GAN is an algorithm that generates virtual images using deep learning principles and is used in various fields such as image generation as well as voice and language processing, and has developed into various modified models, becoming a major technology in deepfake generation. By analyzing the positive cases of deepfakes, it is expected in which areas the development of deepfakes in the future can have a positive impact.

### Key words

Deepfake, GAN, Deepfake positive uses

## 1. 서 론

딥페이크(deepfake)는 2017년에 처음 등장하였으며, ‘딥러닝(deep learning)’과 ‘가짜(fake)’의 합성어로 인공지능 기술을 이용하여 현실에는 존재하지 않거나 발생한 적 없는 사건을 이미지, 동영상, 오디오 등을 합성하여 만들어 내는 기술이다[1]. 이러한 딥

페이크를 생성하는 기술에는 다양한 접근방식이 있지만 주로 생성적 적대 신경망(GAN)이라는 인공지능 모델을 활용하여 만들어진다.

최근 딥페이크가 발전함에 따라 이와 관련하여 여러 범죄가 확산되었다. 특히 성범죄와 연관이 크게 되어져 딥페이크에 대한 사회적 인식이 부정적으로 자리잡혀가고 있다.

\* 국립군산대학교 소프트웨어학과

† 교신저자

\*\* 국립군산대학교 소프트웨어융합공학과

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음 (2023-0-00065)

그러나 딥페이크에 대한 사회적 인식과는 다르게 이 기술은 긍정적인 방향으로도 사용될 수 있는 양면성을 지닌 기술이다.

따라서 본 논문에는 GAN을 통해 만들어진 딥페이크의 동향과 긍정적 사례를 분석하고자 한다.

## II. GAN 알고리즘 개요 및 딥페이크의 응용

이 장에서는 딥페이크에 흔히 사용되는 딥러닝 모델인 GAN과 이를 활용한 딥페이크 응용에 관해 설명한다.

GAN은 딥러닝의 원리를 이용하여 현실에는 없는 이미지를 생성하는 기술로, 생성자와 판별자는 무작위 가중치로 초기화되며, 생성자는 판별자를 속이기 위해 진짜 같은 데이터를 생성하도록 학습하며, 판별자는 생성자가 생성한 가짜 데이터를 정확하게 판별할 수 있도록 학습한다[2]. 이 과정을 반복하면서 두 모델은 점점 더 높은 성능을 가지게 된다.

GAN은 다양한 변형 모델로 발전해 왔으며[3], DCGAN, StyleGAN 등 고해상도 이미지 생성이나 합성에 강점을 보여 딥페이크 응용에 널리 쓰인다.

GAN 알고리즘은 이미지 생성뿐만 아니라 음성 및 언어 처리, 악성코드 탐지 등 많은 분야에서 활용되고 있어, 딥페이크 생성 및 탐지에서도 활용되고 있다.

표 1. GAN의 변형 모델의 특징과 응용  
table 1. Features and applications of GAN's variant model

모델	특징	응용
DCGAN [4]	CNN을 활용 이미지 생성에 강점	기본 이미지 생성
StyleGAN [4, 5, 6]	고해상도 이미지 생성 이미지 속성 조절 가능	가상 인물, 게임 캐릭터 등 생성
CycleGAN [7]	유사한 이미지 변환 보장	얼굴 표현 변경 등

## III. 딥페이크의 긍정적 활용 사례 분석

딥페이크는 주로 부정적으로 인식되지만, 긍정적으로 활용한 사례도 존재한다.

### 3.1. 영화 및 엔터테인먼트 분야

이 분야에서는 GAN 기반 딥페이크 기술이 활용되며, 고해상도 이미지 생성에 강점이 있는 변형 모델인 StyleGAN과 같은 변형 모델이 주로 사용된다. 이를 통해 영화에서는 배우의 얼굴을 자연스럽게 합성하거나, 대역을 실제 배우처럼 보이게 만들어 제작비를 절감하는 동시에 영화를 관람하는 관객에게는 현실감 있고 일관성 있는 시각효과를 제공할 수 있다[8].

대표적인 사례로는 스타워즈라는 영화에서 이미 고인이 된 캐리 피셔의 얼굴을 복원하는데 GAN과 여러 가지 기술을 결합한 AI 합성 기술을 통해 자연스러운 합성을 하였다[9]. 이렇듯 할리우드 영화에서는 GAN을 기반으로 배우의 젊은 시절의 모습을 복원하거나 고인이 된 배우의 모습을 복원하는데 많이 사용한다.

### 3.2. 문화 예술 분야

딥페이크는 문화 예술 분야에서 작품을 보는 관객에게 새로운 경험을 줄 수도 있습니다. 예를 들어, 모나리자가 직접 자신에 대해 설명을 해준다면, 실제 명화들이 살아 움직이거나, 명화 주인공의 얼굴이 자신의 얼굴로 변한다면 전시를 즐기는 관람객들은 참여형 전시를 즐길 수 있어 색다른 경험을 제공받을 수 있다[8].

실제로 플로리다주 세인트 피터즈버그에 있는 달리 박물관에서 2019년에 전시회 Dalí Lives를 선보여 이미 세상에서 사라진 역사적 인물 살바도르 달리의 모습을 재현하였다[10]. 이에 따라 관객은 그와 함께 대화하는 것과 같은 경험을 할 수 있었다.

### 3.3. 교육 분야

이 기술을 활용하여 역사 교육 분야에서도 긍정적인 영향을 미칠 수 있다. 대부분의 교육과정은 글을 읽어 교육하는 것이 일반적이지만 이 기술을 활용하여 역사적 인물과의 소통을 통한 교육을 할 수 있어, 교육적인 측면에서 뛰어난 활용성을 보인다.

실제로 미국의 교육기관인 유다시티(Udacity)에서 온라인 교육과정에 딥페이크를 활용하여 학생이 직접 텍스트를 읽는 것이 아닌 콘텐츠나 오디오 나레



이선에서 자동으로 강의 비디오를 생성하는 기술을 조사 중이다. 또한 음성 합성 회사인 CereProc에서는 미국의 존 케네디 전 대통령의 오디오를 분석하여 그의 마지막 연설을 만들었다[8].

이렇듯 이 기술을 활용하여 역사 교육과 온라인 교육에서 사용하게 되면 학습자들은 교육에 있어 지루함을 느끼지 않고 효율적인 교육이 가능해진다.

#### IV. 결 론

GAN은 이미지뿐만 아니라 음성 처리 등 다양한 분야에서 활용되어 딥페이크 생성의 주요 기술로 자리하고 있으며, GAN의 변형 모델들은 현재까지도 딥페이크에 폭넓게 응용되고 있다. 향후 더욱 정교한 딥페이크를 생성하기 위해서는 GAN뿐만 아니라 다른 기술과의 결합을 통해 자연스러운 합성이 가능해지면, 딥페이크가 그리운 사람들을 다시 만날 수 있도록 도움을 주고 영상 매체뿐만 아니라 과거 실종자들의 현재 모습을 구현하거나, 문화 예술, 교육 분야 등 다양한 분야에서 대중들에게 긍정적인 영향을 줄 수 있을 것으로 예상된다.

#### 참 고 문 헌

[1] B. U. Mahmud and A. Sharmin, "Deep insights of deepfake technology: A review", arXiv preprint arXiv:2105.00192, May. 2021.

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, C. Aaron, and Y. Bengio, "Generative Adversarial Nets". Advances in neural information processing systems 27, 2014

[3] J. Asan, I. Ekaputri, C. Natalie, and K. Purwandari, "Exploring Generative Adversarial Networks (GANs) for Deepfake Detection: A Systematic Literature Review", IEEE In 2023 International Workshop on Artificial Intelligence and Image Processing (IWAIP), pp. 189-194, Dec. 2023

[4] B. Jason, "A Tour of Generative Adversarial Network Models", MachinLearningMastery in Generative Adversarial Networks, Jul. 12. 2019:

<https://machinelearningmastery.com/tour-of-generative-adversarial-network-models/>. [accessed: Nov. 04, 2024]

[5] J. Vijaya, A. A. Kazi, K. G. Mishra, and A. Praveen, "Generation And Detection of Deepfakes using Generative Adversarial Networks (GANs) and Affine Transformation", IEEE In 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1-6, Jul. 2023

[6] A. H. Bermano, R. Gal, Y. Alaluf, R. Mokady, Y. Nitzan, O. Tov, O. Patashnik, and D. Cohen-Or, "State-of-the-Art in the Architecture, Methods and Applications of StyleGAN", In Computer Graphics Forum, Vol. 41, No. 2, pp. 591-611, May. 2022

[7] B. Jason, "A Gentle Introduction to CycleGAN for Image Translation", MachinLearningMastery in Generative Adversarial Networks, August. 17. 2019: <https://machinelearningmastery.com/what-is-cyclegan/>. [accessed: Nov. 04. 2024]

[8] C. Millie, "What Are The Positive Applications of Deepfakes?", Jumpstart Magazine, Jun. 9. 2022: <https://www.jumpstartmag.com/what-are-the-positive-applications-of-deepfakes/>. [accessed: Nov. 04. 2024]

[9] L. Vejay, M. Adine, and H. Zach, Fenwick, New York and Santa Monica, USA., "Artificial intelligence: deepfakes in the entertainment industry", WIPO Magazine, Jun. 2022: [https://www.wipo.int/wipo\\_magazine/en/2022/02/article\\_0003.html/](https://www.wipo.int/wipo_magazine/en/2022/02/article_0003.html/). [accessed: Nov. 04. 2024]

[10] dal 'ilives (via artificial intelligence), 11. May. 2019,[Online]: <https://thedali.org/exhibit/dali-lives/>. [accessed: Nov. 04. 2024]

# 웹 크롤링 기반 청년 지원 정책 홍보 및 지원 시스템

서연정\*, 김가인, 소혜진, 신은빛, 허용도\*\*

## Web crawling-based youth support policy promotion and support system

Yeon Jeong Seo\*, Ga In Kim\*, Hye Jin So\*, Eun Bit Sin\*, Yong Do Her\*\*

### 요 약

대한민국 정부는 청년들을 대상으로 취업 지원, 생계 지원, 구직 활동 등 다양한 정책을 시행하여 여러 방면에서 청년들의 삶을 돕고 있다. 그러나 이러한 다양한 정책을 홍보하고 지원받을 수 있도록 해주는 웹사이트들의 비효율적인 정보 제공으로 인해 청년들의 정책 사이트 이용률이 그다지 높지 않은 상황이다. 이 논문에서는 이러한 웹사이트들의 비효율적 정보 제공을 보완하기 위해 웹 크롤링을 사용하여 청년들에게 중요한 정보를 효율적으로 수집하여 볼 수 있도록 하였으며, 그 결과를 활용하여 청년 정책 제도에 쉽게 지원할 수 있도록 하였다. 또한, 사용자가 방문 및 이용하였던 청년 정책 지원에 대한 다양한 기록을 데이터베이스(DB)와 연동하여 과거의 사용 이력을 쉽게 조회하고 재이용할 수 있도록 하였다.

### Abstract

The Korean government has implemented various policies for young people, such as employment support, livelihood support, and job search activities, to help young people's lives in various ways. However, due to the inefficient provision of information on websites that enable them to promote and receive support for these various policies, the use of policy sites by young people is not very high. In this paper, in order to supplement the inefficient information provision of these websites, web crawling was used to efficiently collect and view important information for young people, and the results were used to easily support the youth policy system. In addition, various records of youth policy support visited and used by users were linked to the database (DB) so that past usage history could be easily inquired and reused.

### Key words

Youth Support Policy, Web Crawling, Database, Firebase, Python Selenium, CSS(Cascading Style Sheets)

## 1. 서 론

현재 대한민국 정부는 만 19세에서 34세 청년들

\* 건양대학교 의료 IT 공학과 Email : jjeng2love@naver.com

\*\*건양대학교 의료 IT 공학과(교신저자) Email : ydher@konyang.ac.kr.

※ 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다.

을 대상으로 취업 지원, 생계 지원, 구직 활동 등 다양한 정책을 시행하여 여러 방면에서 청년들의 삶을 돕고 있다. 하지만 이러한 노력에도 불구하고 정책 이용률은 50%에 미치지 못하고 있다[1,2]. 이러한 현상에는 여러 요인이 있지만, 주된 원인은 정보 제공의 비효율성을 꼽을 수 있다. 청년들이 정책을 신청하려면 여러 사이트를 통해 정보를 검색해야 하는데, 이 과정에서 많은 시간을 소비하게 된다. 또, 대부분의 정책 신청 사이트는 지자체에 의해 운영되고 있기 때문에 정책 외의 정보도 함께 제공하고 있어 원하는 정보만 선별하여 파악하기가 쉽지 않다.

이에 본 연구에서는 웹 크롤링 기술을 활용하여 이러한 문제를 해결하고자 한다. 본 연구에서 개발하는 청년 정책 통합 시스템은 웹 크롤링을 이용하여 정책 정보를 자동으로 탐색하고 수집한 후, 필요한 내용만 효율적으로 제공한다. 또한, 데이터베이스(DB) 연동을 통해 사용자가 방문했던 정책을 자동으로 기록하여, 이후 사용자의 필요에 따라 쉽게 찾아보고 재방문 할 수 있는 기능을 제공한다. 이를 통해 사용자가 정책 신청 과정에서 겪는 불편함을 줄여 청년들의 다양한 정책 참여를 돕고, 나아가 정책 참여율을 높이는데 기여하고자 한다.

## II. 웹 크롤링, 데이터베이스(DB) 시스템 설계

본 연구에서 효율적인 정보 제공을 위해 먼저 웹 사이트 정보 수집과 데이터베이스(DB) 환경 구축을 진행하였다. 정보 수집에는 웹상의 정보를 자동으로 탐색하고 수집하는 웹 크롤링 기술을 활용하였으며 [3,4], 이를 위해 파이썬과 HTTP 요청 및 응답을 통해 HTML 구조를 분석하고 파싱할 수 있는 BeautifulSoup와 requests 라이브러리를 사용하였다. 이들로 가져오기 어려운 정보는 Selenium 라이브러리를 추가로 활용하여 동적 요소를 포함한 데이터를 수집할 수 있도록 하였다. DB 환경 구축에는 Firebase와 Cloud Firestore를 사용하여, 실시간 데이터 업데이트와 자동 확장 기능을 통해 편의성과 효율성을 갖출 수 있도록 하였다[5]. 그 다음, DB를 통한 로그인 기능을 구현하여 사용자별로 고유한

컬렉션 구조를 생성하고, 사용자별로 구분된 정보를 저장할 수 있도록 하였다. 크롤링을 통해 수집한 정책 정보(정책 이름, 기간, 분야, 지역 정보, 상세 링크)는 텍스트 형태로 저장되어 리스트 형식으로 출력된다. 출력된 정보는 사용자의 조회에 따라 각 사용자별 고유한 컬렉션 구조에 저장되어, 사용자가 방문한 정책 정보를 쉽게 다시 볼 수 있도록 하였다. 컬렉션 구조에는 정책 제목, 방문 시간, 웹 주소를 포함하고 있기 때문에 사용자가 이전에 방문한 정책 정보를 손쉽게 조회하고 필요할 때 재방문할 수 있도록 하였다.

<그림 1>은 개발한 프로그램의 전체적인 흐름도를 간단하게 도식화한 것이다.

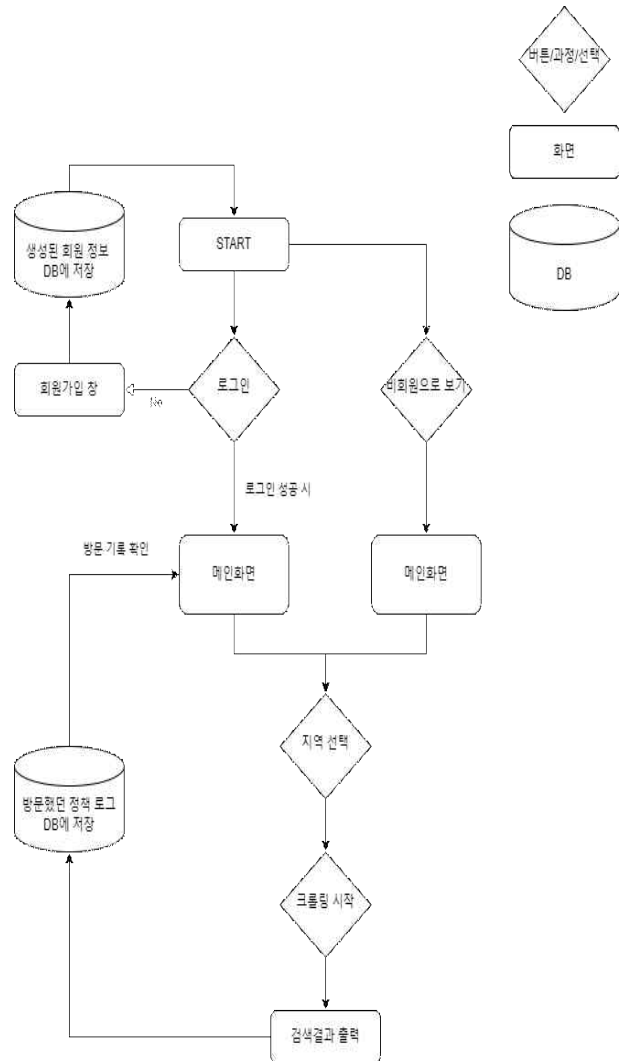


그림 1. 플로우 차트  
Fig. 1. Flow Chart

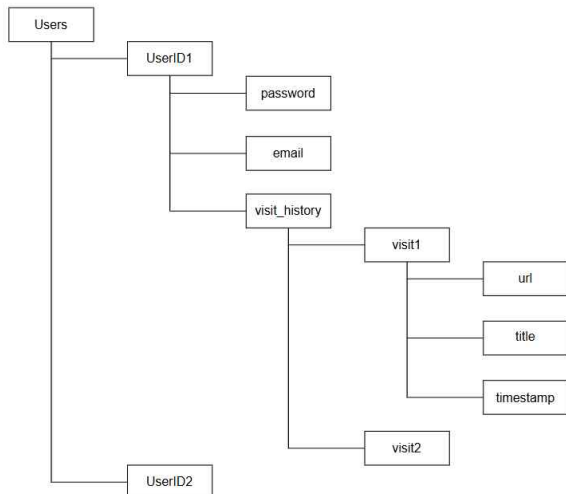


그림 2. DB 구조 설계도  
Fig. 2. Database Schema Design

<그림 2>는 회원가입 내용과 회원의 방문 기록을 관리하기 위해 적절한 데이터 구조를 설계한 설계도이며, 내용은 ‘Users’이라는 최상위컬렉션으로 만들고 그 안에 각 사용자의 ID를 문서명으로 사용한다. 예를 들어 사용자의 아이디가 ‘UserID1’이라면 문서명은 ‘UserID1’이 된다. 각 문서에 저장된 필드에는 사용자 정보(회원가입 내용)과 visit\_history가 저장된다. visit\_history는 각 사용자 문서 하위에 존재하는 또 다른 서브컬렉션으로 설정하였고, 방문 기록을 문서 형태로 저장하였다. 각 문서는 방문한 사이트의 url과 제목, 타임스탬프 등이 포함된다

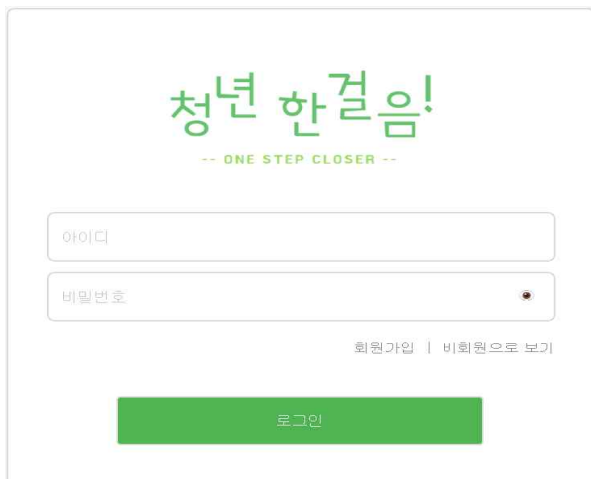


그림 3. 로그인 화면  
Fig. 3. Login

<그림 3>은 시스템을 시작하는 첫 번째 화면으로 로그인 기능을 제공하는 화면이다. 자신의 아이디와 비밀번호를 작성한 후 로그인 버튼을 통해 메인 화면으로 접근할 수 있고, 회원이 아닐 경우 회원가입 버튼을 클릭해 회원가입을 하거나, 비회원으로 보기를 클릭해 메인 화면에 접근할 수 있다. 단, 비회원으로 접근 시 정책 검색, 방문 등의 기본적인 기능은 사용 가능하지만 자신의 개인 아이디가 없으므로 방문기록 서비스는 제공하지 않는다.

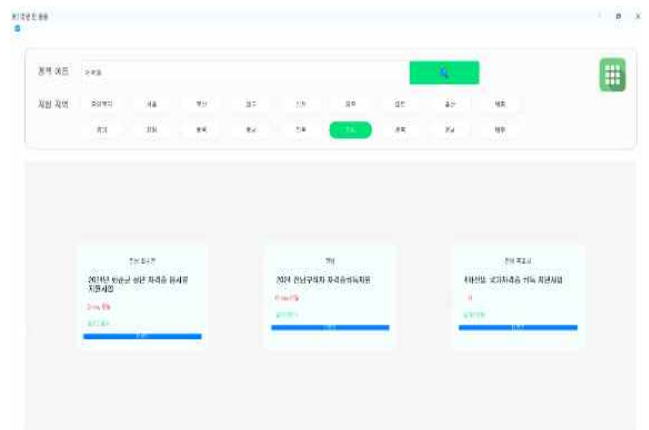


그림 4. 회원 화면  
Fig. 4. Member screen

<그림 4>는 로그인에 성공하면 보이는 회원 화면이다. 회원 화면에서는 크롤링을 통해 출력된 정책을 확인할 수 있다. 사용자가 지역을 선택하게 되면 지역에 연결된 사이트에서 정책 이름, 기간, 분야, 지역 정보, 상세 정보 링크를 수집하고 카드 박스 형태로 결과를 출력한다. 카드 박스에는 정책 이름, 기간, 분야 순으로 나타나고 상단 중앙에는 선택한 지역이 뜬다. 더보기 버튼에는 상세 정보 링크가 연동되어 있어 클릭하면 상세 정보를 확인할 수 있는 페이지로 연결된다. 또, 메인 화면에서는 자신이 방문했던 정책을 확인할 수 있는 방문 기록 버튼이 활성화 되어 있어 클릭하면 방문 기록 화면으로 넘어가게 된다.

<그림 5>는 회원가입을 하지 않은 비회원이 보는 화면으로 전반적인 기능은 회원 화면과 동일하지만 비회원 화면은 자신이 방문했던 정책을 확인할 수 있는 방문 기록 확인 버튼이 활성화 되어있지 않다.

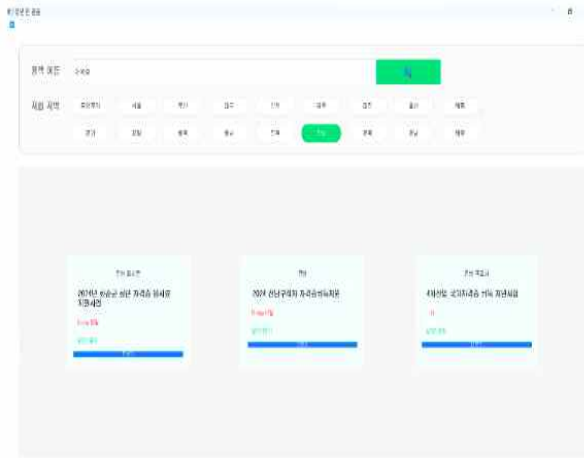


그림 5. 비회원 화면  
Fig. 5. Nonmember screen

쌓인 목록 형태로 항목들을 배치하였기에 구분선으로 항목간의 구분을 명확하게 하고, 스크롤바를 이용해 가독성이 향상된 편리한 탐색을 할 수 있도록 한다. 기록된 데이터를 손쉽게 관리하게 위해 우측 상단에 전체 삭제버튼을 배치하였으며 항목별 삭제버튼이 있어 원하는 항목을 선택하여 편리하게 삭제할 수 있도록 하였다.

### III. 결 론

본 논문에서는 웹 크롤링을 통해 중요한 정보만 자동으로 수집하고 탐색하여 정렬된 형태로 출력한다. 또, 데이터베이스(DB) 연동을 통해 사용자가 방문했던 정책 기록을 제공하여 방문 정보를 쉽게 조회하고 재방문할 수 있도록 한다. 이 시스템은 정책 신청 과정에서 겪는 불편함을 줄여 원하는 정보만 효율적으로 이용할 수 있어 청년들의 다양한 정책 참여를 돕고, 나아가 정책 참여율을 높이는데 기여할 수 있을 것이다.

향후 유지보수를 통해 생기는 문제점들을 꾸준히 보완하고, 크롤링 속도를 개선하여 더 빠르게 정보를 수집할 수 있도록 연구할 계획이다. 또, 다양한 사용자층 확보를 위해 웹이 아닌 앱에서도 시스템을 이용할 수 있도록 변환 작업을 진행할 예정이다.

### 참 고 문 헌

- [1] 청년정책·이슈 토크아보기 설문조사 결과보고서 pp. 25~30, 2024
- [2] 청년센터 이용 경험 49.3% “적극적인 홍보 필요해”, October 2023
- [3] A Study on Web Crawling - Focused on its working mechanism and legal responsibility pp. 4, February 2023
- [4] 만들면서 배우는 파이썬과 40개의 작품들, 2022
- [5] 백견불여일타 파이어베이스, 2019

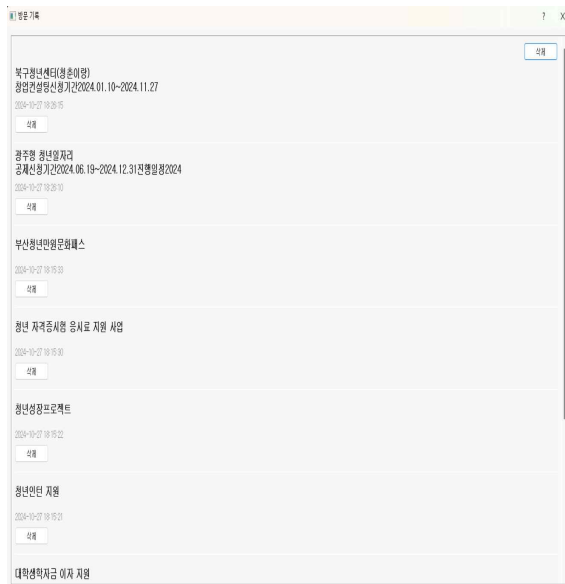


그림 6. 방문 기록 화면  
Fig. 6. Visit history screen

<그림 6>은 개인 아이디를 가지고 있는 회원인 사용자의 방문기록을 확인하고 관리할 수 있는 화면이다. 방문한 정책 사이트의 식별이 편리하도록 정책 제목과, 특정 시간에 방문한 사이트를 추적할 수 있도록 날짜와 시간(시-분-초)을 확인할 수 있도록 하였다. 또한 이 페이지를 나가지 않고도 손쉽게 원하는 사이트에 재방문할 수 있도록 url를 연결해왔기 때문에 방문 기록에 있는 정책 항목을 클릭하면 해당 사이트로 넘어갈 수 있도록 하였다. 스타이

# 효율적 협업을 위한 다기능 통합 플랫폼 Slur

김재영\*, 김태현\*, 유승모\*, 조진혁\*, 김태형\*\*

## Multifunctional Integrated Platform Slur for Efficient Collaboration

Jaeyoung Kim\*, Taehyeon Kim\*, Seungmo Yu\*, Jinhyeok Cho\*, and Teahyong Kim\*\*

### 요 약

팬데믹 이후 원격 근무 환경의 확산으로 효율적인 협업툴의 필요성이 높아지고 있다. 그러나 현재의 협업툴들은 다양한 기능을 제공하지만, 여러 도구를 혼합하여 사용함에 따라 업무 집중도 저하와 리소스 낭비가 발생하고 있다. 본 논문에서는 이러한 문제를 해결하기 위해 여러 협업툴의 주요 기능을 통합한 새로운 플랫폼을 제안한다. IndexedDB, HTTP 캐싱, Protobuf 등의 기술을 활용하여 리닝 커브를 줄이고 사용자 간 원활한 협업을 지원한다. 또한, 플랫폼의 통합 기능은 사용자 경험을 개선하며, 작업 과정에서의 단순함과 효율성을 강화한다. 실험 결과, 제안된 플랫폼이 업무 생산성과 협업 효율성을 향상시키는 데 효과적임을 확인하였다.

### Abstract

The expansion of remote work environments following the pandemic has heightened the need for efficient collaboration tools. However, while current collaboration tools offer various features, the combined use of multiple tools leads to reduced work focus and resource wastage. This paper proposes a new platform that integrates the key functions of various collaboration tools to address these issues, the platform reduces the learning curve and supports seamless collaboration among users. Additionally, the platform's integrated features enhance the user experience, streamlining workflows and improving efficiency. Experimental results confirm that the proposed platform effectively enhances work productivity and collaboration efficiency.

### Key words

collaboration tools, online collaboration, remote work, team communication, IndexedDB, Protocol Buffers, Http Caching

## 1. 서 론

팬데믹 이후 원격 및 하이브리드 근무가 보편화되면서 팀 간 원활한 소통과 협업을 지원하는 온라인 협업 도구의 중요성이 급격히 증가하고 있다.

Slack의 통계인 그림 1에 따르면, 사용자의 79%가 협업 플랫폼이 업무에서 매우 중요한 역할을 한다고 답했으며, IT 의사결정권자의 78%가 팬데믹 이후에도 Slack, Microsoft Teams와 같은 협업 솔루션을 지속 사용할 의향이 있다고 밝혔다[1]. 이러한

\* 국립 금오공과대학교 컴퓨터공학과 학부생, {pape2307, kth5352, dbtmdah12}@gmail.com, whwlsgr1419@naver.com

\*\* 국립금오공과대학교 컴퓨터공학과 교수 (교신저자), taehyong@kumoh.ac.kr

도구들은 이메일을 대체하며, 팀 간 실시간 소통 및 비동기적 협업을 촉진해 효율성과 생산성을 높이는 데 중요한 역할을 하고 있다. 실제로, 회의의 약 40%가 비동기적 Slack 스레드로 대체 가능하다는 의견이 있을 만큼 기존의 회의나 이메일에 대한 의존도도 감소하는 추세이다.

그러나 협업툴의 종류가 지나치게 많고 기능이 각기 달라지면서, 하나의 팀이 여러 도구를 동시에 사용하는 경우가 늘고 있다. 예를 들어, zenkit에서 추천하는 협업툴만 해도 35개가 넘고[2], 각 툴마다 특화된 기능을 제공하는데, 이는 사용자가 여러 도구를 익히고 전환하는 데 많은 러닝 커브 비용과 리소스 낭비를 발생시킨다[3][4]. 특히, 다수의 툴을 혼합하여 사용해야 하는 상황에서 직원들이 각 도구의 인터페이스와 기능에 적응하는 데 드는 시간과 에너지가 업무 효율성을 저해하고 집중도를 떨어뜨리는 요소로 작용한다. rocket.chat의 조사인 그림 2에 따르면 직원의 65%가 여러 플랫폼을 오가는 대신 하나의 통합된 플랫폼을 선호한다고 응답한 바 있다[5].

이에 따라 본 논문은 Jira, Notion, Slack과 같은 기존 협업툴들이 제공하는 기능을 통합하여 사용자로 하여금 도구 간 전환의 번거로움과 리소스 낭비를 줄이고, 하나의 플랫폼에서 효율적인 협업을 가능하게 하는 새로운 협업툴 “Slur“를 제안하고자 한다.

## II. 제안 기법

본 플랫폼은 성능 개선과 사용자 경험 향상을 위해 여러 데이터 관리 및 통신 기법을 사용한다. 클라이언트 측 IndexedDB 캐싱을 통해 데이터 접근 시간을 줄이고, HTTP 캐싱을 활용하여 네트워크 대역폭 사용을 최소화하며, Protobuf를 통한 데이터 직렬화 최적화를 통해 효율적인 데이터 전송을 지원한다. 이러한 기술들을 통해 자원 소비를 줄이고, 데이터 접근과 전송을 보다 원활하게 한다. 아래에서는 각 기법이 Slur 플랫폼에서 어떻게 적용되는지 구체적으로 설명한다.

### 2.1 IndexedDB를 활용한 데이터베이스 캐싱

IndexedDB는 브라우저 내에서 비동기적으로 대량의 데이터를 저장할 수 있는 클라이언트 측 데이터베이스로, Slur에서는 프로젝트 정보, 사용자 상태 정보, 빈번히 조회되는 설정 데이터를 IndexedDB에 캐싱한다. 이렇게 저장된 데이터는 요청 시 빠르게 클라이언트 측에서 접근할 수 있게 되어 네트워크 요청 없이도 데이터 접근 시간을 크게 단축할 수 있다.

특히 프로젝트 정보나 설정 데이터는 자주 변경되는 값이 아니므로, IndexedDB에 데이터를 한번 저장해두면 반복 요청을 줄여 서버 부하를 완화할 수 있다. 결과적으로 클라이언트의 네트워크 의존도가 낮아져, 네트워크 지연 시간이 30~50%가량 감소하며, 로컬에서의 즉각적인 데이터 접근으로 사용자 경험이 크게 개선된다.

### 2.2 HTTP 캐싱을 통한 네트워크 대역폭 절감

Slur는 데이터가 최신 상태일 필요가 없는 경우, HTTP 캐싱을 활용해 서버에서 자주 갱신되지 않는 리소스를 캐싱한다. 캐시 제어 헤더(Cache-Control)와 ETag를 이용하여 변경되지 않은 데이터를 캐싱하고, 일정 시간이 지나지 않으면 새로운 요청을 보내지 않도록 한다. 이러한 HTTP 캐싱 전략을 통해 서버 요청 횟수를 줄이고, 네트워크 대역폭을 최대 40% 절감할 수 있다.

특히 이미지나 스타일 시트, 자바스크립트 같은 정적 리소스에 HTTP 캐싱을 적용하면, 페이지 로딩 시간이 개선되며 서버의 리소스 소모를 줄일 수 있다.

### 2.3 Protocol Buffers를 활용한 프로토콜 최적화

Slur 프로젝트는 데이터 전송 효율을 극대화하기 위해 Google의 Protocol Buffers (Protobuf)를 사용하여 서버와 클라이언트 간 데이터를 직렬화한다. Protobuf는 JSON이나 XML에 비해 훨씬 간결하고 효율적인 이진 형식으로 데이터를 직렬화하기 때문에, 데이터 크기가 평균 50~90% 감소하며 네트워크 비용이 절감된다.

Protobuf는 이진 형태로 직렬화되므로 데이터 전송 속도가 빠를 뿐 아니라 서버와 클라이언트 모두에서 빠르게 디코딩할 수 있다. 이로 인해 처리 속도가 약 20-40% 향상되며, CPU와 메모리 사용량을 최적화하여 서버의 자원 소모를 줄일 수 있다.

또한 Protobuf는 메시지 구조를 고정된 스키마로 관리하므로 데이터의 일관성을 유지할 수 있어, 실시간 통신 시 발생할 수 있는 데이터 손실이나 전송 오류를 줄인다. 이를 통해 안정성과 성능을 동시에 확보하여, Slur의 실시간 통신 시스템에서 높은 응답성과 신뢰성을 제공할 수 있다.

### III. 구현 방법

본 논문에서 제안하는 시스템은 문서 작성, 이슈 트래킹, 코드 리뷰, 실시간 채팅 등의 기능을 하나의 플랫폼에서 제공함으로써 개발자들의 업무 효율을 높이고자 한다.

#### 3.1 문서 작성 페이지

해당 페이지는 문서 작성 기능을 제공하며, 각 문서는 제목과 본문 영역으로 구분된다. 초기 상태에서는 제목 영역과 본문의 루트 블록이 존재하며, 본문은 여러 개의 블록으로 구성된다. 블록은 사용자가 마크다운 형식으로 텍스트, 이미지, 목록, 표 등을 입력할 수 있는 입력 창으로 구성되어 있다. 각 블록은 드래그 앤 드롭 방식으로 이동할 수 있어 블록 간 순서를 자유롭게 조정할 수 있다. 또한 블록을 채팅 입력창으로 드래그하여 쉽게 공유할 수 있는 기능을 제공한다.



그림 1. 문서 작성 페이지  
Fig. 1. Create a document Page

사용자 편의를 위해 다양한 인터페이스 기능을 지원한다. 백스페이스 키를 통해 블록을 삭제하거나, 블록의 내용을 이전 블록에 복제할 수 있으며, 엔터 키 입력을 통해 새로운 블록을 생성할 수 있다. 또한 키보드 방향키를 사용한 블록 간 커서 이동이 가능하여 문서 작성 중에도 직관적이고 효율적인 편집이 가능하도록 설계되었다

#### 3.2 이슈 트래킹 페이지

이슈를 관리하기 위해 ISSUE TASK 페이지를 제공한다. 이 페이지에서는 이슈의 우선순위, 이름, 태그, 내용을 표시한다. 우선순위는 1순위부터 3순위까지 있으며, 태그는 새 기능, 개선, 버그의 세 가지가 있다. 우측 상단의 + 버튼을 클릭하여 새로운 이슈를 추가할 수 있다.

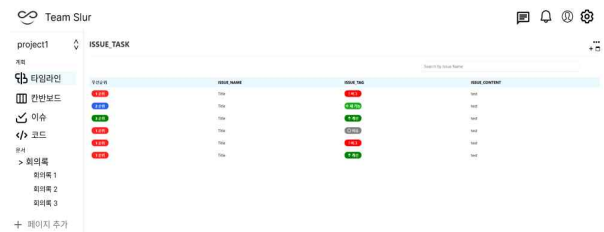


그림 2. 이슈 트래킹 페이지  
Fig. 2. Issue Tracking Page



그림 3. 이슈 칸반 페이지  
Fig. 3. Issue Kanban Page

또한 우측 상단의 칸반보드 전환 버튼을 눌러 화면을 칸반보드 형식으로 변경할 수 있다. 칸반보드



에서는 컨펌, 진행 중, 완료의 세 가지 상태로 이슈를 분류하며, 드래그 앤 드롭으로 이슈의 상태를 변경할 수 있다. 단, 컨펌 상태에서 완료 상태로 바로 옮길 수는 없으며, 반드시 진행 중 상태를 거쳐야 한다.

### 3.3 코드 이슈 페이지

코드와 관련된 이슈를 관리하기 위해 코드 이슈 페이지를 제공한다. 각 이슈는 개별 카드 형태로 리스트업되어 표시되며, 이슈의 타입, 제목, 간략한 설명, 그리고 코드 요약 또는 관련 내용이 포함된다. 페이지 상단에는 검색 영역이 있어 이슈 제목을 통해 특정 이슈를 검색할 수 있다. 이슈 카드를 클릭하면 해당 이슈의 상세 페이지로 이동한다.

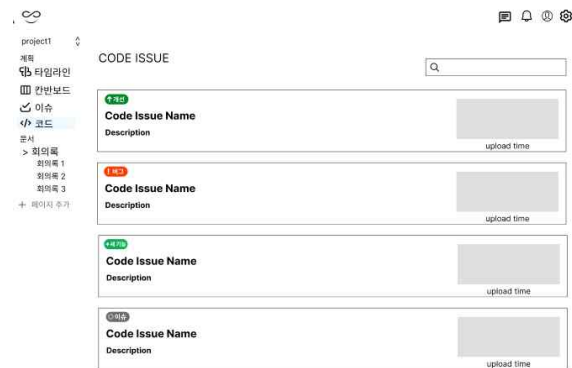


그림 4. 코드 이슈 페이지  
Fig. 4. Code Issue Page

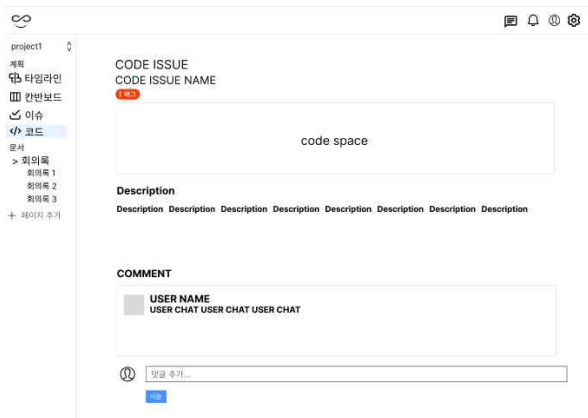


그림 5. 코드 이슈 상세 페이지  
Fig. 5. Code Issue Details Page

코드 이슈 상세 페이지는 이슈에 대한 상세 정보를 출력한다. 페이지 상단에 이슈 생성자가 기입한 코드 본문, 관련 설명을 표출한다. 하단에는 프로젝트의 팀원이 남긴 코멘트가 보여지며, 또한 남길 수 있다.

### 3.4 채팅 기능

채팅 서비스는 프로젝트 참여자 간 실시간 소통을 지원하여 작업 진행 상황을 공유하고 문제 발생 시 빠르게 대응할 수 있도록 설계되었다. 메신저 버튼을 클릭하면 참여자가 속한 채팅방 리스트가 나열되며, 개별 채팅과 그룹 채팅 모두 가능하다.

작업 중에도 실시간 소통을 지원하기 위해 고정 버튼을 제공한다. 이 버튼을 사용하면 채팅창을 화면 오른쪽에 고정시킬 수 있어 다른 서비스를 사용하면서도 채팅을 지속적으로 확인할 수 있다. 예를 들어, 문서 편집 서비스와 채팅을 동시에 사용할 수 있어 편집 작업 중 필요한 정보를 실시간으로 주고받을 수 있다. 또한 드래그 앤 드롭 기능을 통해 마우스만으로 텍스트 및 이미지를 간편하게 복사하고 붙여넣을 수 있어 작업 흐름을 방해하지 않고도 업무를 처리할 수 있다.

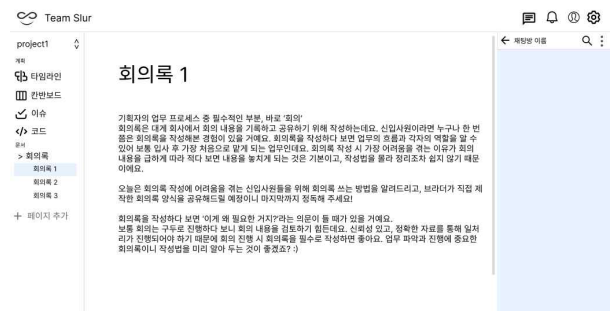


그림 6. 채팅 윈도우  
Fig. 6. Chat window

이러한 기능들은 대화 기록이 저장되어 후속 작업 시 참고 자료로 활용 가능하며, 프로젝트 참여자들이 실시간 소통과 다기능 작업을 병행할 수 있도록 지원한다.

### 3.5 타임라인

해당 페이지에서는 각 이슈를 분류하여 나타낸다. 이슈의 분류에 따라 영역을 나눠 해당 분류에 해당하는 이슈를 리스트업 한다. 리스트의 이슈를 클릭할 시 이슈의 내용을 수정하거나 삭제할 수 있다. 각 리스트의 하단 ' + ' 버튼을 눌러 해당 분류에 속하는 이슈를 생성할 수 있다.



그림 7. 타임라인 - MAIN PLAN 페이지  
Fig. 7. Timeline - MAIN PLAN Page

우측 상단의 버튼을 눌러 모든 이슈를 간트 차트 형식으로 확인 할 수 있다. 각 행은 이슈의 제목, 열은 기간을 표현한다. 이슈의 FRONT\_WORK, END\_WORK에 따라 화살표로 이슈들을 잇는다. 각 이슈 블럭을 클릭할 시 해당 이슈를 수정, 삭제할 수 있다.



그림 8. 타임라인 간트 페이지  
Fig. 8. Timeline Gantt page

### IV. 결론

본 논문에서는 기존의 여러 협업 도구들이 제공하던 개별적인 기능들을 하나의 플랫폼으로 통합함으로써 사용자들이 보다 효율적이고 일관된 작업 환경을 경험할 수 있도록 하였다. 다양한 기능을 한 곳에서 제공함으로써 도구 간 전환의 번거로움을 줄이고 팀 간 소통과 협업의 원활함을 개선하였다. 이는 조직의 생산성과 협업의 질을 높이는 데 중요한 역할을 할 것으로 기대된다.

향후 이 통합된 플랫폼은 사용자 피드백을 반영하여 더욱 발전할 수 있을 것이며, 변화하는 업무 환경에 맞춘 새로운 기능들을 추가함으로써 더욱 강력한 협업 도구로 자리 잡을 것으로 기대된다.

### 참 고 문 헌

- [1] Slack Technologies, "Reinventing Work: New Imperatives for the Future of Working," 2020.
- [2] Zenkit, "35 Best Online Collaboration Tools for Teams in 2024", 2024.
- [3] GetFlow, "Team collaboration tools: benefits and limitations", 2020.
- [4] Projinsights "The Pain of Using Multiple Collaboration Tools: A Challenge in Managing Multiple Projects", 2024.
- [5] Rocket.chat "As Remote Workforces Grow, The Future of Workplace Communication is Evolving", 2021.

# YOLOv5 및 K-prototype 클러스터링 기반 개인 의류 날씨 맞춤형 추천 시스템

박지홍\*, 류다니엘\*, 김주영\*, 차도훈\*, 장용환\*, 이은경\*, 김건우\*\*

## Weather-Aware Outfit Recommendation System Using YOLOv5 and K-prototype Clustering Based on Personal Clothing Data

Ji-Hong Park\*, Daniel Ryu\*, Ju-Young Kim\*, Do-Hun Cha\*, Yong-Hwan Jang\*,  
Eun-Kyung Lee\*, and Gun-Woo Kim\*\*

### 요약

패스트 패션은 최신 유행을 즉각 반영한 디자인과 빠른 상품 회전율이 특징이나, 이로 인한 의류 폐기물 증가는 환경 오염 문제를 심화시킨다. 본 논문에서는 개인이 보유한 의류를 효율적으로 활용하기 위해 YOLOv5를 이용한 의류 속성 판별과 K-prototype 클러스터링 기반 추천을 결합한 날씨 맞춤형 의상 추천 시스템을 제안한다. YOLOv5 모델은 의류의 카테고리, 색상, 길이, 소재, 프린트, 스타일을 판별하며, K-prototype 클러스터링은 이러한 의류 속성과 날씨 정보를 결합하여 최적의 의상을 추천한다. 실험 결과 사용자의 80%가 추천 결과에 만족을 보였으며, 이를 통해 불필요한 의류 소비를 줄이고 의상 선택의 어려움을 해소할 수 있음을 확인하였다.

### Abstract

Fast fashion, characterized by instantly fashionable designs and rapid product turnover, exacerbates environmental pollution through increased clothing waste. This paper proposes a weather-aware outfit recommendation system that combines YOLOv5-based clothing attribute detection and K-prototype clustering-based recommendation to efficiently utilize personally owned clothing. The YOLOv5 model identifies clothing attributes including category, color, length, material, print, and style, while K-prototype clustering combines these attributes with weather information to recommend optimal outfits. Experimental results showed an 80% user satisfaction rate with the recommendations, demonstrating the system's effectiveness in reducing unnecessary clothing consumption and alleviating outfit selection difficulties.

### Key words

clothing recommendation, recommendation system, clustering, you only look once, weather info

---

\* 경상국립대학교 컴퓨터공학부,

\*\* 경상국립대학교 컴퓨터공학과 교수, gunwoo.kim@gnu.ac.kr(교신저자)

※ 본 논문은 2023년도 정부(교육부)의 지원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2021R1G1A1006381)

## 1. 서 론

패스트 패션은 최신 유행을 즉각 반영한 디자인, 저렴한 가격, 빠른 상품 회전율이 특징이며, 소비자의 즉각적인 요구에 따라 빠르게 소비되는 경향이 있다. 지난 20년간 세계적으로 섬유 제조 및 소비는 거의 두 배로 증가했으며, 글로벌 패션 시장은 2020년 1조 5,000억 달러에서 2025년 약 2조 2,500억 달러로 성장할 것으로 예상된다[1]. 패션 산업은 매년 약 790억 입방미터의 물을 사용하며(이는 전 세계 물 소비량의 약 20%), 약 17억 톤의 CO2를 배출하고(전 세계 총 CO2 배출량의 약 10%), 9,200만 톤의 섬유 폐기물을 발생시킨다[2]. 이처럼 패스트 패션으로 인해 폐기되는 의류의 양이 증가하면서 토양과 대기 오염뿐만 아니라 생태계 파괴와 같은 심각한 환경 문제를 초래할 수 있다.

본 논문에서는 날씨에 어울리는 의상을 사용자가 기존의 소유한 옷을 효율적으로 활용해 추천하는 시스템을 제시한다. 이를 통해 불필요한 소비를 줄여 옷 고민과 패스트 패션의 문제를 해결하고 환경 오염을 최소화할 수 있다.

## II. 시스템 설계 및 구현

### 2.1 시스템 구조

본 논문에서 제안하는 의상 추천 시스템의 구조는 다음과 같다. 사용자가 애플리케이션을 통해 개인 옷장에 있는 의상을 촬영하면, clothes detector가 이를 분석하여 카테고리, 색상, 길이, 소재, 프린트, 스타일의 총 6가지 속성을 판별한다.

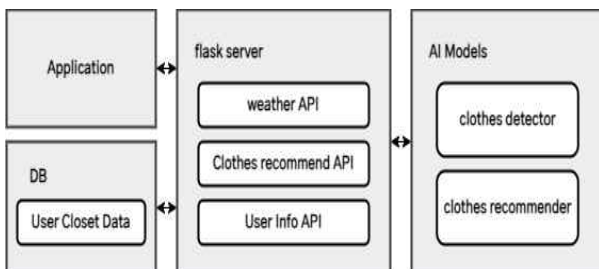


그림 1. 추천 시스템 구조

Fig. 1. Recommendation system structure

이를 기반으로 개인 옷장 데이터베이스를 구축하며, 사용자의 성별, 나이, 스타일 등의 추가 정보를 입력받는다. Clothes recommender는 이러한 사용자 정보, 개인 옷장 데이터, 날씨 정보를 활용하여 개인 옷장에 있는 의상 중 날씨에 적합한 의상을 추천한다.

### 2.2 데이터 수집 및 전처리

의상 추천 모델과 의상 속성 판별 모델의 학습을 위해 K-Fashion 데이터셋과 의류 판매 사이트의 크롤링 데이터를 사용하였다. K-Fashion 데이터셋에서는 카테고리 21개, 색상 14개, 길이 3개, 소재 18개, 프린트 7개, 스타일 10개의 속성을 활용해 데이터를 분류하였다. 의류 판매 사이트에서 크롤링한 데이터는 해당 의상을 착용한 날짜와 기상청 데이터를 연계하여 의상과 날씨 데이터를 매칭하였다.

### 2.3 의상 속성 판별 모델

의상 속성 판별 모델은 의상 카테고리를 판별한 후, 해당 의상의 색상, 기장, 소재, 프린트, 스타일을 분류하는 구조로 설계했다. K-Fashion 데이터를 사용해 의상 카테고리 판별 모델을 학습시킨다. 판별 모델의 성능은 표 1과 같다.

표 1. YOLOv5m 판별 모델 성능

Table 1. YOLOv5m detection model performance

precision	recall	mAP 0.5	mAP 0.95
0.8008	0.8160	0.8600	0.7386

의상 속성 분류 모델은 색상, 기장, 소재, 프린트, 스타일의 분류를 학습했다. YOLOv5m 분류 모델을 사용한 의상 속성 판별 모델의 성능은 표 2와 같다.

표 2. YOLOv5m 분류 모델 성능

Table 2. YOLOv5m classification model performance

분류	loss	acc_top1	acc_top5
color	1.0756	0.7725	0.9923
length	0.9298	0.6469	-
material	1.4504	0.6276	0.9653
print	0.5489	0.9620	0.9986
style	1.7236	0.4676	0.9367

카테고리 판별 모델의 결과 이미지에서 바운딩 박스를 crop하여 분류 모델의 입력으로 사용했다. 학습된 의상 속성 판별 모델을 활용해 크롤링한 데이터셋의 의상 속성을 판별하고, 이를 바탕으로 클러스터링에 사용할 데이터를 확보하였다.

## 2.4 의상 추천 모델

의상 추천 모델을 학습하기 위해 K-prototype 알고리즘을 사용하였다. K-prototype 알고리즘은 K-means와 K-modes 알고리즘을 결합한 방식으로, 연속형과 범주형 데이터가 혼합된 데이터를 클러스터링할 때 사용된다. 범주형 데이터 간의 차이는 식 (1)을 통해 계산되며, 유클리디안 거리와 해밍 거리 간의 균형을 맞추기 위해 가중치  $\gamma$ 를 곱한다. 연속형 데이터인 날씨 데이터와 범주형 데이터인 의상 속성 데이터를 사용하여 데이터 포인트와 군집 중심 간의 거리를 식 (2)로 정의한다. 클러스터링의 최적 k 값을 결정하기 위해 엘보우 방법과 실루엣 계수를 사용해, k를 40으로 설정했다.

$$\delta(X_i, C_j) = \begin{cases} 0, & \text{if } X_i = C_j \\ 1, & \text{if } X_i \neq C_j \end{cases} \quad (1)$$

$$D(X_i, C_j) = \sum_{cont} (X_i - C_j)^2 + \gamma \sum_{cate} \delta(X_i, C_j) \quad (2)$$

클러스터링 결과를 기반으로 각 클러스터에 속한 의상 데이터들의 속성 개수를 계산한다. 사용자의 정보를 입력받아 클러스터링을 수행한 후, 선택된 클러스터의 속성 비율에 맞춰 랜덤하게 의상 속성을 선택하고, 해당 속성에 맞는 의상을 개인 옷장에서 찾아 추천한다.

## 2.5 데이터베이스 및 서버

데이터베이스와 서버는 Firestore를 사용하여 구현하였다. 클라이언트는 user\_id로 구분하고, 각 사용자에게 옷장 데이터셋을 할당하였다. 기상청 초단기 예보 API와 초단기 실황 API를 통해 실시간 날씨 데이터를 받아 옷 추천 모델에 전달한다.

## 2.6 애플리케이션

사용자 개인 옷장 정보와 실시간 날씨 데이터를 기반으로 상의와 하의를 추천한다. 애플리케이션 동작 화면은 그림 2와 같다.

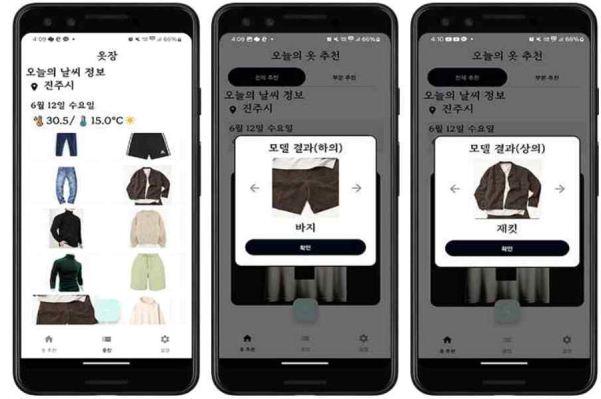


그림 2. 스마트폰 애플리케이션 동작 화면  
Fig 2. Smartphone application behavior screen

## III. 실험 및 평가

본 연구에서 제안한 추천 시스템 성능을 평가하기 위해 대학생 20명에 대해 추천 만족도 조사를 진행했다. 만족도 조사 결과는 그림 3과 같다.

Recommendation Satisfaction

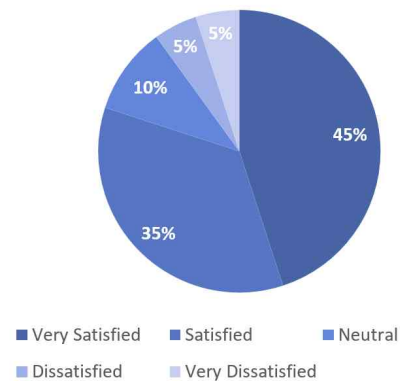


그림 3. 의상 추천 만족도 조사  
Fig. 3. Outfit recommendation satisfaction survey

만족도 조사 결과, "매우 만족" 9명과 "만족" 7명으로 전체 응답자의 80%가 만족한다는 의견을 제시했다. 이를 통해 본 시스템의 추천 성능이 우수하다는 것을 확인할 수 있다.

#### IV. 결 론

본 연구에서 제안한 추천 시스템을 사용하면 기존에 소유한 옷을 활용하여 날씨에 적합한 의상을 추천받을 수 있다. 이를 통해 불필요한 소비를 줄이고, 패스트 패션 문제와 옷 선택에 따른 스트레스를 완화할 수 있다.

#### 참 고 문 헌

- [1] Papamichael, I., Chatziparaskeva, G., Pedreno, J. N., Voukkali, I., Candel, M. B. A., & Zorpas, A. A. (2022). Building a new mind set in tomorrow fashion development through circular strategy models in the framework of waste management. *Current Opinion in Green and Sustainable Chemistry*, 36, 100638.
- [2] Nikolina, S. A. J. N. (2019). Environmental impact of the textile and clothing industry: What consumers need to know.

# 사용자 편의성을 고려한 연합학습 모니터링 시스템 구현

김동현\*, 김재엽\*, 이민지\*, 이상민\*, 서영덕\*\*, 이석훈\*†

## Implementation of a Monitoring System Considering the Convenience of Federated Learning

Donghyun Kim\*, Jaeyeab Kim\*, Minji Lee\*, Sangmin Lee\*, Young-Duk Seo\*\* and Sukhoon Lee\*†

### 요약

최근 인공지능과 머신러닝 기술의 발전과 함께 연합학습이 주목받고 있다. 연합학습은 데이터 보안 문제를 극복하며 모델을 개발할 수 있다는 장점이 있어 활발히 연구가 진행되고 있다. 연합학습 시스템을 효율적으로 운영하기 위해서는 연합학습 과정을 실시간으로 모니터링하고 클라이언트 환경 변화에 대해 빠르게 파악하는 것이 중요하다. 본 논문에서는 사용자 편의성을 고려한 연합학습 모니터링 시스템을 제안하고 구현한다. 사용자는 제안 시스템을 통해 클라이언트의 상태와 학습 진행 상황을 실시간으로 모니터링하고 제어할 수 있다.

### Abstract

With the advancement of AI and machine learning, federated learning has gained attention for its ability to develop models while ensuring data security. Efficiently operating a federated learning system requires real-time monitoring and quick adaptation to client environment changes. This paper proposes and implements a user-friendly federated learning monitoring system that allows users to monitor and control client status and training progress in real-time, improving the efficiency of federated learning.

### Key words

Federated Learning, Monitoring System, Artificial Intelligence,

## 1. 서론

최근 인공지능과 머신러닝 기술의 발전과 함께 연합학습이 주목받고 있다. 연합학습은 중앙 서버에

데이터를 직접 전송하지 않고 개별 클라이언트가 로컬에서 모델을 학습한 후 학습된 파라미터만 전송한다[1]. 이러한 특징으로 의료, 금융, 웨어러블 기기 등 다양한 산업 분야에서 데이터를 보호하며

\* 국립군산대학교 소프트웨어학과

† 교신저자

\*\* 인하대학교 컴퓨터공학과

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065).

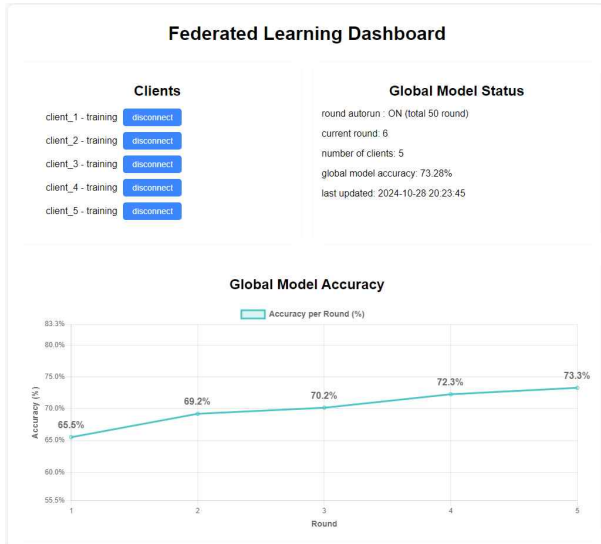


그림 1. 연합학습 모니터링 시스템  
Fig. 1. Federated learning monitoring system

모델을 개발할 수 있다는 장점 때문에 활발히 연구가 진행되고 있다[2]. 이러한 연합학습 시스템을 효율적으로 운영하기 위해서는 분산된 환경의 여러 클라이언트와 중앙 서버 간의 통신과 글로벌 모델의 업데이트 과정을 실시간으로 모니터링하고 클라이언트의 환경 변화에 대해 빠르게 파악하는 것이 중요하다[3]. 기존의 연합학습 플랫폼에서는 이러한 모니터링 시스템이 부재하거나 일부 기능에만 국한되어 있어 학습 진행 상황을 종합적으로 파악하기 어렵다.

본 논문은 사용자 편의성을 고려한 연합학습 모니터링 시스템을 제안한다. 제안 시스템에서는 연합학습 진행 과정에서 클라이언트의 상태 정보와 글로벌 모델 업데이트 진행 상황에 대한 정보를 실시간으로 수집하고 시각화할 수 있다. 이를 통해 사용자는 연합학습의 각 단계에서 발생할 수 있는 문제를 빠르게 파악하고 대응할 수 있다.

## II. 연합학습 모니터링 시스템

웹 인터페이스를 통해 사용자는 참여 클라이언트의 상태와 글로벌 모델의 상태를 실시간으로 모니터링하고 연합학습 진행 상황을 제어할 수 있다. 연합학습을 진행하면서 클라이언트는 진행 상황에 따른 현재 상태를 서버에게 전송하며 서버는 클라이

언트에서 보내준 상태를 수신하면 이를 웹을 통해 시각화한다. 그림 1은 제안하는 연합학습 모니터링 시스템 메인 화면을 보인다.

좌측 상단에 참여한 각 클라이언트의 상태를 확인할 수 있다. 클라이언트의 상태는 join, ready, training, update, finish로 표현되며 각각 join은 처음 연합학습 시스템에 참여한 상태, ready는 서버로부터 파라미터를 전송받아 로컬 모델 학습이 준비된 상태, training은 로컬 모델이 학습 중인 상태, update는 업데이트된 글로벌 모델 파라미터를 전송받아 로컬 모델에 적용하는 상태, finish는 로컬 모델 학습을 완료하고 로컬 파라미터를 서버로 보낸 상태를 나타낸다.

우측 상단은 현재 학습 중인 글로벌 모델의 상태에 대해 보여주며 round autorun은 연합학습의 자동 진행 유무를 나타내고 current round는 현재 진행된 연합학습 라운드 수, number of clients는 연합학습의 참여 중인 클라이언트 수, global model accuracy는 현재 라운드 글로벌 모델의 정확도를 나타내고 last updated는 마지막으로 진행된 라운드의 시간을 보여준다. 또한 하단에 라운드별 글로벌 모델 정확도를 실시간으로 확인할 수 있다.

그림 2는 클라이언트가 연합학습 중 탈퇴한 상황에서의 모니터링 시스템을 보인다. 클라이언트가 학습 중 탈퇴하면 Global Model Status에서 number of clients에 변경 사항이 바로 반영되고 Clients의 클라이언트 목록에서 해당 클라이언트가 삭제된다.

그림 3은 연합학습 진행 중 추가로 클라이언트가 참여한 상황을 보인다. 추가로 참여한 클라이언트는 초기 파라미터를 통해 로컬 모델을 학습하고 서버로 파라미터를 전송한 다음 라운드에 참여하기 위해 대기한다. 추가로 참여했을 때도 number of clients에 변경된 클라이언트 수가 바로 반영되고 Clients의 클라이언트 리스트에도 클라이언트에 대한 정보가 추가된다.

## III. 결 론

본 논문에서는 사용자 편의성을 고려한 연합학습 모니터링 시스템을 제안하였다. 제안한 모니터링 시



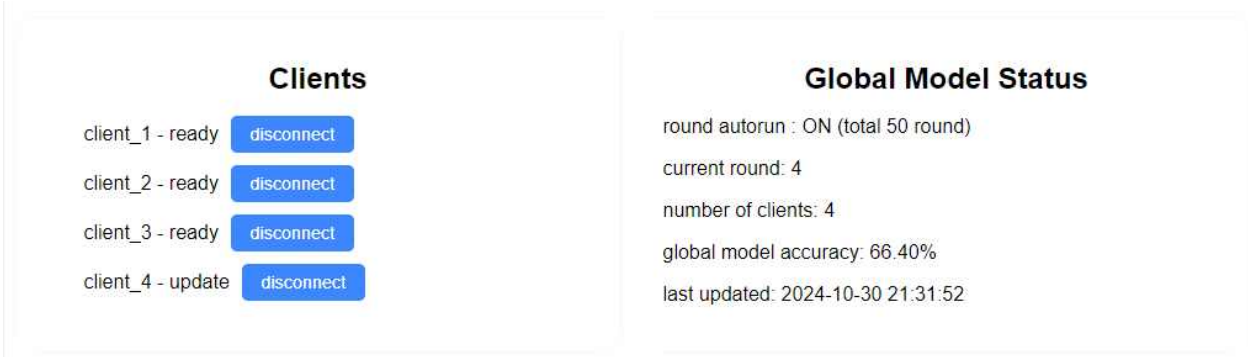


그림 2. 학습중 클라이언트의 탈퇴  
Fig. 2. Client dropout during training

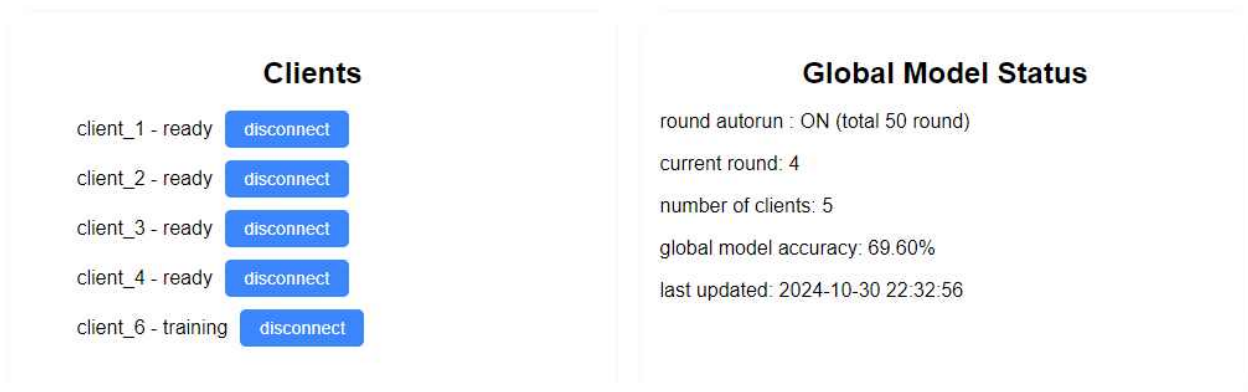


그림 3. 연합학습 진행 중 클라이언트의 추가 참여  
Fig. 3. Client addition during federated learning

시스템은 클라이언트의 상태와 글로벌 모델의 학습 진행 상황을 실시간으로 시각화하고 관리할 수 있도록 설계되었다. 사용자들은 웹 인터페이스를 통해 각 클라이언트의 상태를 쉽게 파악할 수 있으며, 학습 과정 중 발생할 수 있는 문제에 신속히 대응할 수 있다. 이를 통해 연합학습의 각 단계에서 운영 효율성을 높이고, 클라이언트의 환경 변화에도 유연하게 대처할 수 있다.

### 참 고 문 헌

[1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," Proc. 20th Int. Conf. Artif. Intell. Stat. (AISTATS), 2017.

[2] T. H. Rafi, S. Y. Lim, and D.-K. Chae, "A Survey on Federated Learning and Its Business

Applications," Commun. Korean Inst. Inf. Sci. Eng., Vol. 42, No. 9, pp. 21-27, 2024.

[3] J. Moon, S. Yang, and K. Lee, "FedOps: A Platform of Federated Learning Operations With Heterogeneity Management," IEEE Access, Vol. 12, pp. 4301-4314, 2024.

[4] J. Kim, H. Jin, D. Kim, and S. Shin, "Federated Learning GUI Monitoring Tool," Proc. Korean Inst. Inf. Sci. Eng. (KIISE) Conf., pp. 1883-1885, Dec. 20, 2023.

# 횡단보도 사고 예방을 위한 횡단감지 정보제공 시스템 설계 및 구현

이원준\*, 박규원\*, 박규동\*, 문정준\*, 장시웅\*\*

## Design and Implementation of Crosswalk Detection Information Providing System for Preventing Pedestrian Accidents

Won-Jun Lee\*, Gyu-won Park\*, Gyu-Dong Park\*, Jeong-Jun Mun\*, and Si-Woong Jang\*\*

### 요 약

이 연구에서는 실제 도로 환경에 적용할 수 있는 시스템을 설계하고 시뮬레이션한다. 설계된 시스템은 딥러닝 모델로 횡단보도 영상에서 보행자를 식별하고 노면의 조명을 밝혀 보행자 진입 여부와 진행 방향을 전달한다. 이 시스템은 노면 조명 장치를 통해 효과적으로 정보를 전달할 수 있으며, 운전자 사각지대를 대신 감시함으로써 횡단보도의 사고율을 낮출 수 있다.

### Abstract

This study designs and simulates a system applicable to real-road environments. The proposed system uses a deep learning model to identify pedestrians in crosswalk footage and utilizes road surface lighting to communicate pedestrian entry and movement direction. By effectively delivering information through road surface lighting devices and monitoring drivers' blind spots, this system aims to reduce crosswalk accident rates.

### Key words

Crosswalk safety, Pedestrian detection, Road surface lighting, Deep learning model, Driver blind spot monitoring

## I. 서 론

신호 횡단보도에서 사고 비율은 전체 중 65.3%로 무신호 횡단보도보다 높다. 특히 신호 횡단보도에서 치사율은 신호가 없는 횡단보도보다 1.7배 높은 것으로 조사되었고, 사고 원인은 무단횡단, 시간 내에 횡단 미완료, 운전자 부주의, 충분한 제동거리 미확보가 있다[1, 2]. 기존 횡단보도 사고 예방책은 운전

자의 능력에 과하게 의존한다. 이전 연구 및 스마트 횡단보도는 보행자의 주의를 환기하는 방식에 국한되거나 상황별로 적합한 정보를 전달하지 못했다. 이 연구에서는 운전자의 부담을 줄이기 위해 딥러닝 모델을 이용한 제어 시스템을 제안한다.

## II. 관련연구

\* 동의대학교 컴퓨터공학과 학부생, dnjswns648@naver.com, \*\* 동의대학교 컴퓨터공학과 교수, swjang@deu.ac.kr

※ 본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신사업의 결과입니다.(2023RIS-007)

횡단보도 진입로 점자블록에 이미지 처리를 적용하여 보행자의 진입을 검출하는 방식이 제안되었다 [2]. 그러나, 오인식 가능성이 크며, 다중 보행자 인식 어려움 등의 한계로 상용화가 어렵다. 이외에도 횡단보도 사고 예방을 위해 조명을 설치하여 야간에 시야를 제공하는 방안이 적용되고 있지만, 여전히 운전자의 감각에 전적으로 의존하며 사각을 완전히 보완하지 못한다.

노면 표시는 측주식과 문형식 표지판보다 주시율과 시선유도율이 높다[3]. 선행 연구에서 운전자가 가장 인식하기 쉬운 색상은 백색인 것으로 나타났다[4]. 그리고 무채색 배경에 테두리가 강조되는 표지 디자인이 가시성, 색상 대조가 가장 뛰어났다[5]. 시스템의 설계 목표는 위 특성을 반영하여 노면의 검은색과 강한 대비를 이루는 백색광을 이용한다.

### III. 시스템 설계

횡단보도 상황을 관측하기 위해 카메라를 이용한다. 촬영되는 영상은 실시간으로 객체 탐지 모델에 입력된다. 이를 위한 이미지 처리와 딥러닝 모델의 가동은 엣지 디바이스에서 이루어진다. 사용된 엣지 디바이스는 NVIDIA 사의 Jetson Nano development kit이다. 모델의 출력으로 획득한 보행자 위치를 횡단보도 경계와 비교하여 진입했다고 판단된 경우, 조명을 제어하는 MCU에 신호를 전달한다.

MCU가 신호를 받으면 현재 신호등의 점등 상태에 따라 다양한 방식으로 조명을 작동시킨다.

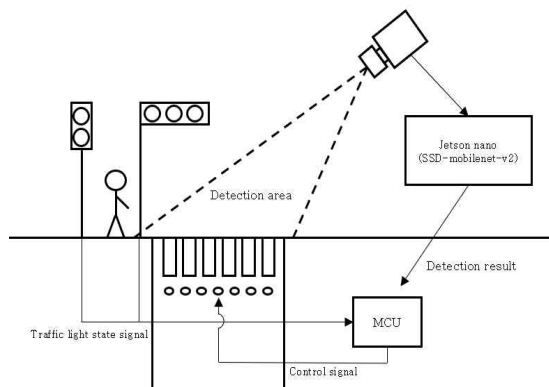


그림1. 시스템 구성과 데이터 흐름  
Fig1. System Configuration and Data Flow

표 2. 시뮬레이션 번호와 시스템 기대 동작  
Table2. Simulation Number and Expected Behavior

시뮬레이션 번호	기대 동작
T-001	차선을 다라 설치한 LED 점등
T-002	각 그룹 하향 순차 점등
T-003	각 그룹 상향 순차 점등

조명역할의 청색 LED를 노면에 매설하고 형광 필름을 부착해 백색광을 출력한다. LED는 차선과 중앙선, 횡단보도로 분리된 4개 부분에 각각 배치하였다. LED 그룹은 보행자 신호가 점멸하거나 모든 신호등이 적색일 때, 횡단보도에 보행자가 진입하면 진행 방향으로 순차 점등을 한다. 그림4에서 상향, 하향은 카메라로 획득한 영상을 기준으로 한다. 차량 녹색 신호일 경우 빠른 속도로 접근하는 차량이 보행자 유무를 멀리서 확인할 수 있도록 차선 LED가 켜진다.

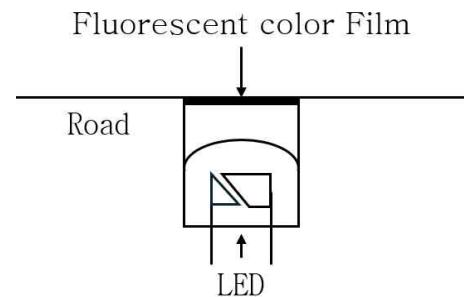


그림2. 조명 장치  
Fig2. Lighting Signal Device

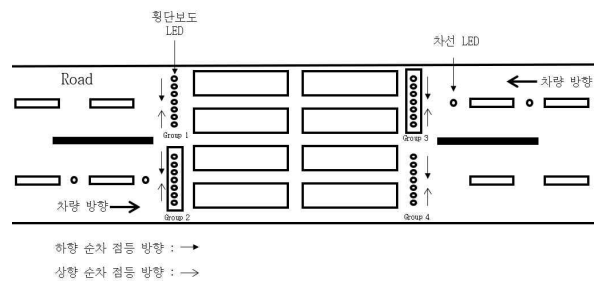


그림3. 조명의 설치 및 동작  
Fig3. Installation and Operation of the Signal Device

딥러닝 모델은 SSD-Mobilenet-v2를 사용하였다. SSD-Mobilenet-v2는 모바일 기기에서 동작하기 적절

한 CNN기반의 딥러닝 모델이다. 시뮬레이션을 위해 준비된 인형을 탐지하도록 모델을 학습시켰다. 카메라의 주간과 야간 모드로 인형을 여러 배경과 각도에서 촬영하여 획득하고 1000장을 학습용 데이터로 선택하였다. 학습된 모델을 TensorFlow lite 모델로 변환하여 엣지 디바이스에 저장하였다.

신호장치를 제어하는 MCU는 아두이노 uno를 사용하였다. TensorFlow 라이브러리와 Pyserial 라이브러리를 이용해 보행자 탐지, MCU와 통신을 수행하는 프로그램을 작성하였다.

#### IV. 시뮬레이션

이 연구에서는 세 가지 경우의 횡단보도 상황을 시뮬레이션하였다. 표1, 표2는 시뮬레이션을 위한 상황의 번호와 그에 대한 설명, 그리고 상황에 따른 기대 동작을 나타낸다.

표 1. 시뮬레이션 번호와 상황 묘사  
Table1. Simulation Number and Situation Description

시뮬레이션 번호	상황 설명		
	자동차 주행 신호	보행자 신호	횡단 상황
T-001	녹색등	적색등	무단횡단
T-002	적색등	녹색 점멸등, 적색등	하향횡단
T-003	적색등	녹색 점멸등, 적색등	상향횡단

#### V. 결론

이 연구에서 운전자가 정확하고 빠르게 횡단 보도 상황을 파악할 수 있도록 해줄 효과적인 신호 전달 방법을 제안하였다. 신호장치를 제어할 수단으로 딥러닝 모델과 엣지 디바이스를 사용하였다. 특히 엣지 디바이스를 사용할 경우, 크기가 작아 실제 도로에 적용할 때 유용하다. 이 연구에서는 시뮬레이션을 위해 인형 이미지를 학습시켰으나, 이미 훈련된 모델을 적용하면 사람을 인식하는 시스템으로 전환할 수 있다.

#### 참 고 문 헌

- [1] 고병곤, 조준한, 전제호, 김인석, “횡단보도 보행 교통사고 유형 및 특성 분석.” 교통 기술과 정책 12.4 (2015):40-47
- [2] 최상범, 윤여수, 이원영, 박현준, "점자블록 인식을 이용한 횡단보도 보행자 안전 시스템." 한국정보통신학회 종합학술대회 논문집 23.2 (2019): 351-354.
- [3] 임채홍, 신소명, 김도훈, 임준범, 김동기. (2020). 주행 중 운전자의 교통안전표지 및 노면표시 인식도 분석. 대한교통학회지, 38(4), 281-291, 10.7470/jkst.2020.38.4.281
- [4] 도현철. "도로환경에서의 운전자 시각 행태 특성에 관한 연구." 국내석사학위논문 한밭대학교, 2015. 대전
- [5] 나은경, 박도윤, 이다현, 이지현, "공공 안내 표지판의 색상 시인성에 관한 연구 - 국립아시아 문화전당 안내 표지판을 예제로 한 색상 시인성에 관한 연구 -." 디자인학연구 35.1 (2022): 297-310.

# XAI 기반 침입 탐지 시스템 - SHAP를 활용한 IDS 투명성 향상 연구

신은성\*, 김성빈\*, 윤영준\*, 정현준\*\*

## XAI-Based Intrusion Detection System - Research on Enhancing IDS Transparency Using SHAP

Eunsung Shin\*, Seongbin Kim\*, Youngjoon Yoon\*, and Hyunjun Jung\*\*

### 요 약

최근 사이버 보안 위협이 증가하면서 머신러닝 기반 침입 탐지 시스템(IDS)이 활발히 연구되고 있으나, 대부분의 모델이 블랙박스 특성을 가져 의사결정 과정의 해석이 어렵다는 한계가 존재한다. 이 연구는 설명 가능한 인공지능(XAI) 기법인 SHAP(SHapley Additive exPlanations)를 활용하여 IDS의 의사결정 과정을 해석 가능하게 하는 방법을 제안한다. CICIDS2017 데이터셋을 사용하여 XGBoost 기반의 분류기를 구현하였으며, 99.89%의 정확도를 달성하면서도 모델의 투명성을 확보하였다. 이 연구에서는 SHAP 기법을 통해 각 특성의 중요도를 분석하고, 공격 유형별 주요 특성을 식별하여 효과적인 보안 대응 전략 수립에 기여하고자 한다.

### Abstract

Recent increases in cyber security threats have led to active research on machine learning-based Intrusion Detection Systems (IDS). However, most models have limitations in that they possess black-box characteristics, making it difficult to interpret their decision-making processes. This research proposes a method for making the decision-making process of IDS interpretable using SHAP (SHapley Additive exPlanations), an explainable artificial intelligence (XAI) technique. An XGBoost-based classifier was implemented using the CICIDS2017 dataset, achieving 99.89% accuracy while ensuring model transparency. Through SHAP, this research analyzes the significance of each feature and identifies key characteristics for each attack type, thereby contributing to the development of effective security response strategies.

### Key words

ids, xai, shap, explainable, ai, machine learning

---

\* 국립군산대학교 소프트웨어학과 학사과정, gjfepfm@kunsan.ac.kr, napo987@kunsan.ac.kr, 2201067@kunsan.ac.kr

\*\* 군산대학교 소프트웨어학부 교수(교신저자), junghj85@kunsan.ac.kr

※이 연구는 정부(과학기술정보통신부)의 재원으로 한국 연구재단의 지원을 받아 수행되고 있습니다.(No. NRF-2022R1G1A1008493).

※본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065).

## I. 서론

최근 사이버 보안 경제는 급속도로 성장했다. 이러한 경제적 성장은 사이버 보안 위협의 증가로 이어져 2023년 사이버 공격 건수는 전년 대비 30% 증가했다. 공격 방식도 복잡화, 지능화 되었다. 이러한 고도화된 위협 환경에서 네트워크 침입 탐지 시스템(IDS)이 중요해졌다[1]. 최근 머신러닝 기술을 활용한 IDS가 활발히 개발되고 있으나, 대부분의 모델이 블랙박스 특성을 가지고 있어 의사결정 과정의 해석이 어렵다.

블랙박스 모델은 예측 결과의 논리적 근거를 이해하거나 설명하기 어렵다. 이는 사용자가 모델이 왜 특정 결정을 내렸는지 파악하지 못하며, 책임 소재를 명확히 규명하기 어렵다[2]. 이러한 문제는 신뢰성, 설명 가능성, 책임성을 요하는 전문 분야에서 나타난다. 또한 시스템 개발자가 블랙박스 모델의 내부 작동 방식을 완전히 이해하지 못하기 때문에 모델을 수정하거나 성능을 개선하는 것이 어렵다. 또한, 규제 기관이 공정성과 데이터 보호 규정을 준수하는지 확인하는 것이 어렵다. 이에 유럽 연합은 개인정보 보호 규정인 GDPR을 제정했다. GDPR은 법적 영향을 미치거나 중대한 영향을 미치는 자동화 처리에 기반한 결정의 대상이 되지 않을 권리를 규정한다.

최근 머신러닝 기반 IDS 관련 연구들도 이러한 한계점을 가진다. 머신러닝 기반 IDS 모델들은 높은 정확도를 보이지만, 의사결정 과정을 이해하기 어려워 보안 전문가들이 모델의 판단 근거를 파악하기 어렵다. 기존 연구들은 모델의 성능 지표에만 집중하여 오답지가 발생했을 때 그 원인을 분석하고 개선하기 어렵다. 마지막으로, 학습된 패턴에만 의존하여 새로운 유형의 공격에 취약하며, 모델의 판단 근거를 알 수 없어 대응 전략 수립에 어렵다[3].

이 연구에서는 이러한 한계점들을 극복하기 위해 설명 가능한 인공지능(XAI) 기법인 SHAP(SHapley Additive exPlanations)를 활용한 IDS를 제안한다. 구체적으로 이 연구의 목적은 다음과 같다. 첫째, SHAP를 활용하여 모델의 의사결정 과정을 해석 가

능한 IDS를 구현한다[4]. 둘째, XGBoost 분류기를 사용하여 IDS의 성능을 평가하고, SHAP 값을 통해 특성 중요도를 분석한다. 셋째, 각 공격 유형별 주요 특성을 분석하여 효과적인 보안 대응 전략 수립에 기여한다.

## II. 관련 연구

### 2.1 XAI 기반 Network 침입공격 탐지 및 공격 유형별 대표적 특징분석에 관한 연구

해당 연구의 설계도에서는 NSL-KDD 데이터셋의 표준 규격인 41개의 특성(6개의 Binary features, 32개의 Continuous features, 3개의 Categorical features)을 포함한 데이터를 XGBoost 분류기의 입력으로 사용할 수 있도록 제안한다[5]. 이때 데이터는 Binary features는 원본 값을 유지하고, Categorical features는 One-Hot-Encoding을 적용하며, Continuous features는 Min-Max Scaling을 통해 전처리된다. 또한 모델의 판단 근거를 파악하기 위해 XAI 기법인 SHAP을 활용하여 각 레이블(DoS, Probe, R2L, U2R)별 대표적 특성들을 선정하고, 이를 피어슨 상관계수로 검증하여 모델의 의사결정 과정을 해석한다. 제안된 시스템은 74.55%의 정확도를 달성하였으며, 각 공격 유형별 대표 특성을 성공적으로 도출하여 AI 모델의 예측 결과에 대한 신뢰성 있는 해석을 제공하였다.

### 2.2 설명 가능한 인공지능(XAI)을 활용한 침입탐지 신뢰성 강화 연구

이 연구에서는 XGBoost 기반 침입탐지 모델의 예측 결과를 해석 가능하게 하기 위해 XAI 기법인 SHAP을 활용하였다[6]. PKDD2007 데이터셋을 이용하여 각 특성의 중요도를 시각화하고, query\_char\_plus\_cnt 등 특성이 특정 공격 유형에 미치는 영향을 분석하였다. 이를 통해 보안 전문가가 예측 근거를 이해할 수 있도록 하여 모델의 블랙박스 문제를 완화하고 침입탐지 시스템의 투명성을 높이는 데 기여하였다.

### III. XAI 기반 침입 탐지 시스템 - IDS 투명성 향상 연구

#### 3.1 시스템 아키텍처

이 연구에서 제안하는 XAI 기반 침입 탐지 시스템은 네트워크 트래픽에서 발생하는 다양한 유형의 침입을 탐지하고 의사결정 과정을 해석 가능하게 제공하여 사용자에게 신뢰성과 투명성을 갖춘 정보를 제공하는 시스템이다. 시스템의 핵심은 XGBoost 분류기를 활용하여 침입을 탐지하고, SHAP를 통해 모델의 판단 근거를 분석하는 데 있다.

네트워크 트래픽 데이터가 시스템에 입력되면 먼저 데이터 전처리 과정을 거친다. 전처리된 데이터는 XGBoost 분류기에 입력되어 정상 트래픽을 포함한 다양한 공격 유형으로 분류된다. 동시에 SHAP 분석이 수행되어 각 특성의 기여도가 산출되며, 이를 통해 모델이 어떤 특성을 중점적으로 고려했는지에 대한 인사이트를 제공한다.

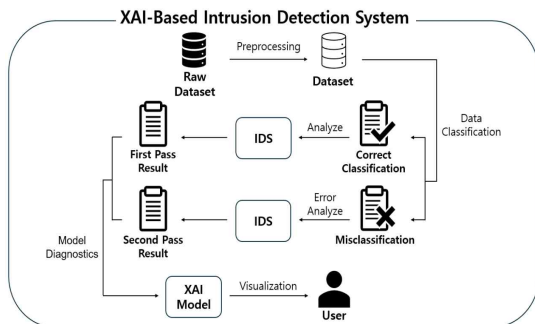


그림 789. XAI-IDS 시스템 아키텍처  
Fig. 1. Architecture of XAI-based Intrusion Detection System

#### 3.2 데이터셋 전처리 과정

데이터셋 전처리는 모델의 학습 성능과 예측 신뢰성을 향상시키기 위한 필수적인 단계로, 데이터의 품질을 개선하여 정확한 학습 결과를 얻는 데 중요한 역할을 한다. 이 연구에서는 CICIDS2017 데이터셋을 사용하며, 데이터 품질 개선을 위해 우선적으로 결측치와 무한값을 제거하는 과정을 수행한다.

결측치가 제거된 데이터는 Min-Max 정규화를 통해 스케일링되며, 이는 상이한 범위를 가진 특성들을 동일한 척도로 조정하여 모델의 학습 공정성과 정확성을 확보한다. 또한 공격 유형을 나타내는 레이블은 정수형으로 인코딩되어 모델 입력 데이터로 사용된다.

#### 3.3 클래스별 특성 중요도 분석 함수

모델의 의사결정 과정을 해석하기 위해 SHAP을 활용한 특성 중요도 분석을 수행한다. SHAP은 각 특성이 모델의 예측에 미치는 영향을 정량적으로 평가하여 전체적인 특성 중요도를 파악하게 한다. 이 과정에서 SHAP Feature Importance Plot을 통해 각 특성의 기여도를 시각적으로 표현하고, 공격 유형별로 주요 특성들을 도출하여 특성별 차이를 면밀히 분석한다. 특히 Dependence Plot을 활용하여 특정 특성값의 변화가 모델 예측에 미치는 영향을 구체적으로 분석할 수 있다. 이러한 분석은 모델이 특정 공격을 탐지할 때 중점적으로 고려하는 네트워크 특성의 중요성을 심층적으로 이해하는 데 기여한다.

표 1. 클래스별 특성 중요도 분석 함수

Table 1. Class-wise Feature Importance Analysis Function

```
important_features = set()
y_encoded_temp = self.label_encoder.fit_transform(y)

for i, class_name in
enumerate(self.label_encoder.classes_):
    class_mask = (y_encoded_temp == i)
    class_data = X[class_mask]
    if class_name in ['BENIGN', 'DDoS']:
        n_features = int(len(numeric_cols) * 0.95)
    else:
        n_features = int(len(numeric_cols) * 0.85)

    class_variances = class_data.var()
    class_important =
class_variances.nlargest(n_features).index

    important_features.update(class_important)
X = X[list(important_features)]
print(f"\nSelected features considering class-specific
importance: {len(important_features)}")
progress.update(1)
```

## IV 실험 및 비교 평가

### 4.1 실험 환경

이 연구에서 사용된 하드웨어 및 소프트웨어 환경은 표 2와 같다. 실험의 일관성을 확보하기 위해 모든 실험은 동일한 환경에서 5회 반복 수행되었으며, CICIDS2017 데이터셋을 8:2 비율로 훈련 데이터와 테스트 데이터로 분할하여 사용하였다.

표 2. 실험 환경 구성

Table 2. Experimental Environment Configuration

Operating System	Ubuntu 22.04.1
Framework	Python 3.12
Library	XGBoost 2.1.2
	SHAP 0.46.0
CPU	Xeon(R) Silver 4210R
GPU	NVIDIA RTX A6000x2
Memory	128 GB

### 4.2 실험 결과

이 연구에서 제안한 XAI-IDS의 성능을 평가하기 위해 CICIDS2017 데이터셋을 사용하여 실험을 진행하였다. XGBoost 분류기는 99.89%의 정확도를 달성하였다.

표 3. 클래스별 성능 평가 결과

Table 3. Class-wise Performance Metrics of XAI-IDS

Class	Precision	Recall	F1-Score
BENIGN	0.958	0.936	0.947
DDoS	0.946	0.963	0.954
DoS	0.951	0.946	0.948
PortScan	0.965	0.951	0.958

표 3은 클래스별 성능 평가 결과로, 각 공격 유형의 Precision, Recall, F1-Score를 보여준다. 가장 높은 Precision은 PortScan의 0.965, Recall은 DDoS의 0.963으로 나타났다. 반면, 가장 낮은 Precision은 DDoS의 0.946, Recall은 BENIGN의 0.936을 기록하였다. 모든 클래스에서 Precision과 Recall이 0.936 이상으로, 제안된 모델의 주요 공격 유형에 대한 효과적인 분류 성능을 확인할 수 있다.

이러한 분류 성능의 근거를 파악하기 위해, 모델의 의사결정 과정을 해석하기 위한 SHAP 분석을 수행한 결과, 각 공격 유형별로 중요한 특성들이 식별되었다. 그림 2는 SHAP Feature Importance Plot을 통해 도출된 전체 특성들의 중요도를 보여준다.

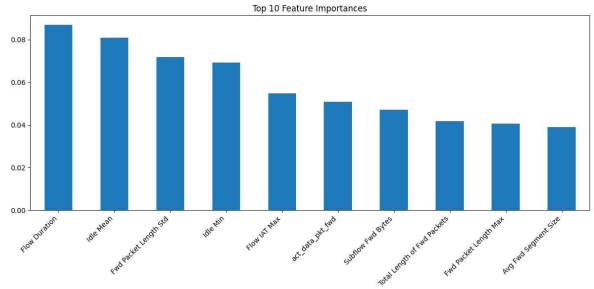


그림 790. SHAP 특성 중요도 그래프  
Fig. 2. SHAP Feature Importance Plot

SHAP 특성 중요도 그래프는 각 특성이 모델의 예측 과정에 미치는 영향도를 정량적으로 분석한다. 이 연구에서는 Flow Duration, Idle Mean, Fwd Packer Length Std 가 주요 특성으로 식별되었다. 이러한 특성별 영향도 분석은 침입 탐지 시스템의 분류 근거를 제시하여 모델의 신뢰성을 검증한다. 또한 블랙박스 특성을 가진 기존 머신러닝 기반 IDS와 달리, 특성 중요도 분석을 통해 모델의 예측 결과에 대한 해석이 가능하며 이는 보안 정책 수립과 시스템 취약성 분석의 기준으로 활용된다.

특성들의 세부적인 영향을 분석하기 위해, SHAP Dependence Plot을 통해 주요 특성들의 영향력을 분석하였으며, 그 결과는 그림 3과 같다.

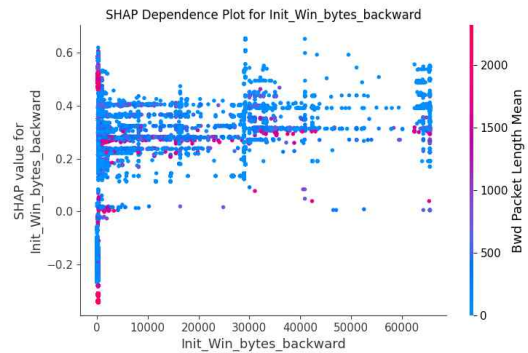


그림 791. SHAP 특성 의존도 분석 그래프  
Fig. 3. SHAP Dependence Plot



그림 3은 SHAP Dependence Plot로 각 특성 간의 상호작용과 예측 결과에 대한 영향도를 시각화한다. 중요도가 높은 특성중 특성 영향도를 가장 잘 보여주는 두 개의 특성인 Init\_Win\_bytes\_backward와 Bwd Packet Length Mean의 관계를 분석하면, Init\_Win\_bytes\_backward 값이 0에 가까울 때 SHAP 값이 -0.2에서 0.6 사이로 넓게 분포되어 있다. 이는 윈도우 크기가 작을 때 패킷 길이의 영향력이 다양하게 나타남을 의미한다. 또한 Init\_Win\_bytes\_backward가 30000 부근에서 SHAP 값이 0.4-0.6으로 집중되는 현상은 특정 임계값에서 모델의 판단에 강한 영향을 미치는 것으로 해석된다. 이러한 특성 간 의존도 분석을 통해 네트워크 패킷의 윈도우 크기와 패킷 길이가 모델의 분류 결정에 미치는 복합적 영향을 파악할 수 있다. 이러한 실험 결과는 제안된 시스템이 향상된 탐지 정확도를 보이면서도, XAI를 통해 의사결정 과정의 투명성을 확보할 수 있음을 입증한다.

## V 결론 및 향후 과제

이 연구는 XAI 기반의 침입 탐지 시스템을 제안하여 모델의 의사결정 과정에 대한 투명성을 확보하고, 99.89%의 탐지 정확도를 달성하였다. 특히 SHAP 분석을 통해 각 공격 유형별 주요 특성을 식별함으로써, 효과적인 보안 전략 수립을 위한 기반을 마련하였다. 이러한 연구 결과는 사이버 보안 분야에서 인공지능의 설명 가능성과 신뢰성을 모두 확보할 수 있음을 입증하였다는 점에서 의의가 있다.

향후 연구 방향으로는 첫째, 스트리밍 데이터 처리 및 실시간 SHAP 계산을 위한 경량화된 모델 개발이 필요하다. 둘째, 온라인 학습과 강화학습을 통한 모델의 적응력 강화가 요구된다. 셋째, LIME, DeepLIFT 등 다양한 XAI 기법과의 비교 분석을 통해 해석력을 향상시킬 필요가 있다. 넷째, 보안 전문가들을 위한 대화형 대시보드 개발이 필요하다. 마지막으로, Random Forest, Deep Neural Network 등 다양한 모델과의 비교 연구를 통해 최적의 모델을 선정하는 연구가 진행되어야 할 것이다.

## 참 고 문 헌

- [1] K. Scarfone and P. Mell, "Guide to Intrusion Detection and Prevention Systems (IDPS)," National Institute of Standards and Technology, NIST Special Publication pp. 800-94, 2007.
- [2] Erico Tjoa and Cuntai Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," IEEE, Transactions on Neural Networks and Learning Systems, vol. 32, no. 11, pp. 4793-4813, Nov. 2021
- [3] C. Zednik, "Solving the black box problem: A normative framework for explainable artificial intelligence," Philosophy & Technology, vol. 34, no. 2, pp. 265-288, 2021. [Online]. Available: <https://doi.org/10.1007/s13347-019-00382-7>.
- [4] P. J. Phillips et al., "Four principles of explainable artificial intelligence," National Institute of Standards and Technology, NIST Interagency/Internal Report (NISTIR) pp. 8312, September 2021.
- [5] Yongsoo Lee, Ye-eun Lee, and Tae-jin Lee, "A Study on Network Intrusion Attack Detection and Analysis of Representative Features by Attack Type," Journal of D-Culture Archives, vol. 5, no. 1, pp. 193-204, 2022, doi: 10.23089/jdca.2022.5.1.011.
- [6] Ilok Jung, Woobin Choi, and Suchul Kim, "The Enhancement of intrusion detection reliability using Explainable Artificial Intelligence(XAI)," KCGSA, vol. 22, no. 3, pp. 101-110, 2022



# EEG 기반 개인 식별을 위한 DTW-N 알고리즘

이가흔\*, 백종화\*, 임철기\*\*, 전성찬\*\*\*, 이성환\*\*\*\*, 서현\*

## DTW-N Algorithm for EEG-Based Personal Identification

GaHeun Lee\*, Jonghwa Paik\*, Cheolki Im\*\*, Sung Chan Jun\*\*\*, Sunghan Lee\*\*\*\*, and Hyeon seo\*

### 요약

뇌파(Electroencephalogram, EEG) 신호는 다양한 연령이나 정신 상태의 사람들에게서 얻을 수 있으며, 데이터 저장 시스템이 손상되더라도 개인의 실제 신원을 파악하기 어려워 높은 보안성을 제공한다. 기존 연구에 따르면, EEG를 활용하여 동적 시간 왜곡(Dynamic Time Warping, DTW)과 유클리드 거리(Euclidean Distance) 기반의 패턴 매칭을 통해 개인 식별을 수행한 경우, ED가 더 높은 정확도를 보였다. 이에 본 연구는 EEG 특징 분포의 정량화 문제를 해결하기 위해, 신호의 정규화와 시간 보정을 반영한 DTW-N 알고리즘을 제안한다. 계산 결과, DTW-N 알고리즘은 ED 및 일반적인 DTW 방법보다 피험자 간 신호 차이를 효과적으로 분류했으며, 시간 보정 값이 0.005에서 0.007 사이일 때 신호 차이에 영향을 미치는 것으로 확인했다. 이는 DTW-N 알고리즘이 EEG 기반 개인 식별에 있어 신호 차이를 효과적으로 분류하는 데 기여할 수 있음을 시사한다. 본 연구팀은 시간 보정 값의 최적화 범위를 적용하고, 여러 피험자 정보를 기반으로 향후 개인 식별에 적용할 계획이다.

### Abstract

Electroencephalogram (EEG) signals can be obtained from individuals of various ages and mental states, and they provide high security as it is difficult to identify an individual's actual identity even if the data storage system is compromised. According to existing studies, when using EEG for personal identification through pattern matching based on Dynamic Time Warping (DTW) and Euclidean Distance (ED), the ED method demonstrated higher accuracy. In response, this study proposes the DTW-N algorithm, incorporating signal normalization and time correction, to address the quantification of EEG feature distribution. The results indicate that the DTW-N algorithm more effectively classifies signal differences between subjects than ED and DTW, with the time correction value impacting signal differences when set between 0.005 and 0.007. This suggests that the DTW-N algorithm can contribute to effectively classifying signal differences in EEG-based personal identification. Our research team plans to apply the optimized range of time correction values and extend the application to personal identification based on multiple subject data in the future.

### Key words

Dynamic Time Warping-Normalization, User Identification, Similarity Assessment

---

\* 경상국립대학교 IT공과대학 컴퓨터공학과 lgh1217@gnu.ac.kr, qorwhdghk12@naver.com, hseo0612@gnu.ac.kr (교신저자)

\*\* 경상국립대학교 기초과학연구소 chelgi47@gmail.com

\*\*\* 광주과학기술원 AI대학원 scjun@gist.ac.kr

\*\*\*\* 한림대학교 뇌혈관질환 선도연구센터 sh.lee@halym.ac.kr

※ 본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (RS-2023-00280241).

## I. 서 론

생체 인증 시스템은 전통적인 패스워드 기반 인증 시스템에 비해 위조, 도난, 망각으로 인한 문제의 발생 가능성이 낮고 사용이 간편하여 다양한 분야에서 널리 활용되고 있다. 생체 데이터는 시간이 지나도 변하지 않는 지문이나 망막과 같은 신체적 데이터와, 시간이 지남에 따라 변화하지만 개인 특성을 지니는 생체 전기 신호로 분류할 수 있다. 이러한 생체 전기 신호에는 뇌파, 근전도, 심전도, 안구 전도가 포함된다. EEG 신호는 다양한 연령이나 정신 상태의 사람들에게서 얻을 수 있으며, 심지어 식물인간 상태나 혼수 상태에서도 가능하다. 얼굴 사진이나 지문과 달리 EEG 데이터는 수집이 어렵고, 데이터 저장 시스템이 손상되더라도 개인의 실제 신원을 파악하기 어렵기 때문에 보안성이 높은 장점을 가진다 [1-4].

기존 EEG 기반 개인 식별 연구에서 Bai et al.[5]는 시각 유발 전위를 사용하고 유전자 알고리즘, Fisher Discriminant Ratio 등을 이용했다. Phung et al.[6]는 빠른 개인 식별을 위해 Shannon Entropy를 특징으로 사용했으며, Ruiz-Blondet et al.[7]는 ERP를 기반으로 한 개인 식별 시스템을 발표했다. Gui et al.[4]은 사용자 식별을 위한 패턴 매칭 기법으로 DTW와 ED를 함께 이용하였다. 참가자들에서 단어를 보여주고 조용히 읽는 상태의 EEG 데이터를 사용했으며 유클리드 거리 기법은 80% 이상의 정확도를, DTW 방법은 약 68%의 정확도를 보였다.

유클리드 거리의 경우, 두 신호 간에 시간 지연이나 불일치가 있는 경우 낮은 유사성 값을 산출하게 되어, 두 신호 간의 내재된 유사 특성을 제대로 반영하기 어렵다. 하지만 기존 Gui et al.[4] 연구에서는 유클리드 거리가 DTW 방식보다 높은 정확도를 보였으며, 이러한 배경에는 EEG 특징 분포를 정량화하는 문제가 남아 있다고 판단했다. 이에 본 연구팀은 피험자의 각 신호마다 진폭이 다를 수 있기 때문에 신호를 정규화하는 과정의 필요성과 DTW 기법의 경우 피험자 내에서 시간 축이 다를 수 있으므로 시간 보정 값의 필요성을 제안한다. 본 연구에서는 일반적인 ED, DTW 알고리즘과, DTW에 정

규화와 시간 보정을 적용한 알고리즘(DTW-N)이 신호 차이에 미치는 영향을 분석하고자 한다. 이를 통해 EEG 기반 개인 식별의 지표로서 DTW-N의 가치를 평가한다.

## II. 연구 방법

본 연구에서는 청각 자극에 대한 실제 발화를 수집한 데이터 중 두 명의 피실험자 데이터를 이용했다. Cz 채널의 EEG 신호를 활용하였고 ED, DTW, DWT-N에 대한 분석을 수행했다.

$dist_{mn}(i, j)$ 는 subject  $m$ 과 subject  $n$ 의 신호에서  $i$ 번째와  $j$ 번째 데이터 포인트 간의 유클리드 거리를 나타내며 다음과 같이 정의한다.

$$dist_{mn}(i, j) = \sqrt{(i-j)^2 + (x_m(i) - x_n(j))^2} \quad (1)$$

$x_m(i)$ 는  $i$ 번째 시간 포인트에서 subject  $m$ 에게서 획득한 신호를 의미한다.

distance matrix를 이용하여 cost matrix는 다음과 같이 정의된다.

$$cost_{mn}(i, j) = dist_{mn}(i, j) + \min \begin{cases} cost_{mn}(i-1, j-1) \\ cost_{mn}(i-1, j) \\ cost_{mn}(i, j-1) \end{cases} \quad (2)$$

ED는 distance matrix를 사용하며, DTW는 cost matrix를 사용해 계산한다.

DTW-N 알고리즘 계산을 위한 distance matrix의 경우 신호의 정규화와 시간에 대한 보정을 다음과 같이 적용한다.

$$dist_{mn}(i, j) = \sqrt{a(i-j)^2 + (\tilde{x}_m(i) - \tilde{x}_n(j))^2} \quad (3)$$

$\tilde{x}_m(i)$ 는  $x_m(i)$ 를 -1에서 1 사이로 정규화한 값이며,  $a$ 는 시간 보정 값으로 0.001부터 0.01까지 0.0001 간격으로 설정하였다. 이후 DTW와 동일하게 cost matrix로 계산한다.

ED, DTW, DTW-N을 비교를 위해 각 알고리즘에 의해 계산된 행렬의 평균과 분산을 사용해 측정 지표( $\alpha_A$ )를 계산하였다.

$$\alpha_A = \frac{|mean(A_{11}) - mean(A_{12})|}{var(A_{11})} + \frac{|mean(A_{22}) - mean(A_{12})|}{var(A_{22})} \quad (4)$$

$A_{mn} = \{ED, DTW, DTW-N\}$ 이며 subject m과 subject n 사이에서 세 가지 알고리즘 중 하나의 방식으로 계산한 결과이다.  $\alpha_A$ 의 분모는 동일 피실험자 데이터와 다른 두 피험자 간의 비교에서 얻는 결과들의 평균 차이로, 피실험자 간의 차이가 동일 피실험자 내의 차이에 비해 얼마나 큰지를 나타낸다. 이를 동일 피실험자의 분산으로 나누어 두 피실험자 간의 신호 차이를 정량화한다.

### III. 연구 결과 및 분석

표 1은 ED, DTW, DTW-N의  $\alpha$  값을 계산한 결과이다.  $\alpha_{ED}$ 의 값이 가장 작았으며, 신호의 정규화가 적용된 DTW-N의 방식이 일반적인 DTW 보다 큰  $\alpha$  값을 보였다. 즉, 신호의 정규화 과정이 두 피험자 간의 차이를 분석하는 데 기여함을 의미한다.

표 1. ED, DTW, DTW-N에 따른  $\alpha$  값  
Table 1.  $\alpha$  values according to ED, DTW, and DTW-N

$\alpha_{method}$	value
$\alpha_{ED}$	6.2591e-04
$\alpha_{DTW}$	2.0940e-04
$\alpha_{DTW-N} (a=1)$	0.0053

DTW-N 값에서 시간 보정의 효과를 보기 위해 a를 0.001에서 0.1까지의 사용했다 (그림 1). 시간 보정의 효과는 0.001에서 0.006까지는  $\alpha_{DTW-N}$ 이 지속적으로 증가했지만, 그 이상부터는 감소 폭이 미세했다. 시간 보정 a 값이 0.006에서 최대  $\alpha_{DTW-N}$  값을 관측했으며, 0.005에서 0.007 범위 내에서 정확도는 약간의 변동이 있었지만 전반적으로 높은 성능을 유지하였다. 이는 시간 보정 값이 신호 간의 차이를 분석할 때 영향을 미치는 것으로 보인다.

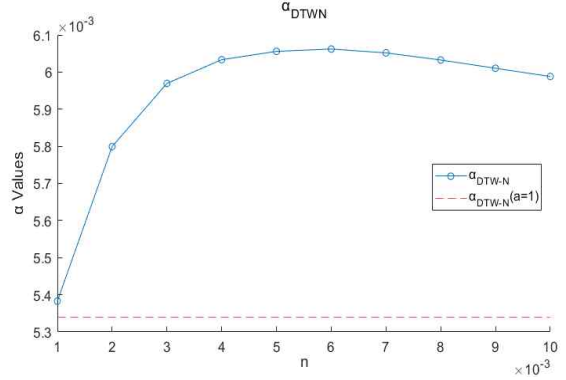


그림 1. 시간 보정 값에 따른 DTW-N의 값  
Fig. 1. Values of DTW-N according to time correction

### IV. 결론

본 연구를 통해 정규화와 시간 보정 값이 적용된 DTW-N이 일반적인 EW, DTW 방식보다 EEG 기반 개인 식별에서 차이를 보는데 기여하는 것을 확인하였다. 향후 연구에서는 이 시간 보정 값의 최적화 범위를 적용하고, 피험자 수와 채널 수를 확장하여 개인 식별에 적용하고자 한다.

### 참고 문헌

- [1] Tatar, Ahmet Burak. "Biometric identification system using EEG signals." *Neural Computing and Applications* 35.1 (2023): 1009-1023.
- [2] Gui, Qiong, Zhanpeng Jin, and Wenyao Xu. "Exploring EEG-based biometrics for user identification and authentication." *2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2014.
- [3] Palaniappan, Ramaswamy, and Danilo P. Mandic. "Biometrics from brain electrical activity: A machine learning approach." *IEEE transactions on pattern analysis and machine intelligence* 29.4 (2007): 738-742.
- [4] Gui, Qiong, et al. "Towards EEG biometrics: Pattern matching approaches for user identification." *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015)*. IEEE, 2015.

- [5] Min, Byoung-Kyong, et al. "Individual identification using cognitive electroencephalographic neurodynamics." IEEE Transactions on Information Forensics and Security 12.9 (2017): 2159-2167.
- [5] Bai, Yanru, Zhiguo Zhang, and Dong Ming. "Feature selection and channel optimization for biometric identification based on visual evoked potentials." 2014 19th International Conference on Digital Signal Processing. IEEE, 2014.
- [6] Phung, Dinh Q., et al. "Using Shannon Entropy as EEG Signal Feature for Fast Person Identification." ESANN. Vol. 4. No. 1. 2014.
- [7] Ruiz-Blondet, Maria V., Zhanpeng Jin, and Sarah Laszlo. "CEREBRE: A novel method for very high accuracy event-related potential biometric identification." IEEE Transactions on Information Forensics and Security 11.7 (2016): 1618-1629.

# 악보 인식(OMR)을 위한 전처리 알고리즘

서승기\*, 김성석\*\*, 김재현\*\*\*

## Preprocessing algorithms for Optical Music Recognition(OMR)

Seunggi Seo\*, Sungsook Kim\*\*, Jaehyun Kim\*\*\*

### 요약

최근 생성형 인공지능로 인해 다양한 분야에서 혁신이 되고 있다. 생성형 인공지능의 경우 내부 파라미터의 개수가 수십 조개에 이를 것으로 추측된다. 당연히 엄청나게 많은 학습 데이터와 이 데이터를 효율적으로 처리할 수 있는 컴퓨팅 인프라를 필요로 한다. 본 논문에서는 악보인식 알고리즘 개발에 관심을 가지고 있다. 하지만, 이전과 달리 적은 데이터와 상대적으로 매우 작은 모델을 이용하여 떨어지지 않은 인식률을 보일 수 있는 접근 방법을 제시한다. 즉, 악보를 입력받으면 전처리 단계에서 음표의 머리, 줄기, 꼬리에 해당하는 획을 분리하는 단계를 수행함으로써 인식 모델을 경량화시키고자 한다.

### Abstract

Recent advancements in generative artificial intelligence have led to innovations across various fields. In the case of generative AI, it is estimated that the number of internal parameters may reach several trillion, necessitating vast amounts of training data and high computational power. In this paper, we focus on the domain of music score recognition. However, unlike previous approaches, we propose a method that achieves competitive recognition performance using a relatively small model with limited data. Specifically, by segmenting the note head, stem, and tail during the preprocessing stage, we aim to streamline the recognition model and reduce its complexity.

### Key words

Optical Music Recognition, preprocessing, stroke classification, CNN, lightweight model

## 1. 서론

최근 향상된 인공지능 기술은 다양한 분야에서 혁신을 일으키고 있다. 이미지, 텍스트, 음악 등 다양한 형태의 데이터를 기반으로 새로운 서비스 생성하는 데 있어 그 잠재력이 주목받고 있다. 음악에서도 인공지능 기술은 작곡, 편곡, 악보 생성 등에서 활발히 활용되고 있으며, 악보 인식(OMR) 역시

주요한 응용 중 하나이다. 하지만, 악보의 복잡한 구조와 다양한 기호로 인해 OMR 시스템의 정확성과 처리 속도에 많은 도전 과제가 따른다. 특히, 악보 이미지의 해상도가 낮을 경우, OMR 시스템은 노이즈 및 왜곡으로 인해 기호 인식에서 성능이 저하될 수 있다.

본 논문에서는 상대적으로 경량 OMR 시스템을

\* 서경대학교 소프트웨어학과, child1188@naver.com, \*\* 서경대학교 소프트웨어학과 sskim03@skuniv.ac.kr 교신저자

\*\*\* 서경대학교 전자컴퓨터공학과, statsr@skuniv.ac.kr

개발하고자 한다. 이를 위해 악보로부터 음표의 머리, 줄기, 꼬리로 구분할 수 있는 전처리 알고리즘을 먼저 제안한다.

## II. 제안한 전처리 알고리즘

기존의 OMR의 방식은 스캔한 악보 이미지를 분석하여 음표, 쉼표, 마디선 등 음악 기호를 자동으로 인식하고 디지털 데이터로 변환한다. 일반적으로, 악보 이미지를 전처리한 후 악보의 보표선과 기호들을 분리한 후, 패턴 인식을 통해 음표나 기호를 인식하는 방식을 사용한다. 음표는 머리, 줄기, 꼬리 등의 세부 요소로 나누기보다는 하나의 통합된 기호로 다루어 인식하는 경우가 많다. 하지만 이 방식은 간단한 악보에서는 효과적이지만 복잡한 악보에서는 성능이 좋지 않을 수 있다.

본 논문에서는 음표 인식을 위한 새로운 전처리 알고리즘을 제안하며, 이는 다음과 같이 4단계로 수행한다: 악보에서 오선 제거 - 오선 제거 과정에서 끊긴 음표/쉼표 이미지 보정 - 음표의 줄기 찾기 - 줄기를 제외한 나머지 머리와 꼬리, 쉼표 구분.

예를 들어, 그림 1-(a)의 원본 악보 이미지에 대하여 4단계의 전처리 작업을 수행하면 그림 1-(b)와 같은 획득이 생성된다.

위에서 언급한 4단계를 그림 2의 원본 악보를 예를 들어 구체적으로 설명하면 다음과 같다.

### ① 악보에서 오선 인식 후, 오선 제거

이미지에서 가로 방향으로 쪽 연결된 5개의 선을 찾아, 각 직선의 y좌표를 저장한다. 이 후, 그림 2에서 5개의 오선을 지운다.

### ② 끊긴 음표, 쉼표 다듬기

그림 3과 같이 오선 삭제로 인해 끊어진 음표나 쉼표 등을 보정하기 위해, 제거된 5개 오선의 y축 좌표값과 각 기호의 연결 여부를 결정한다. 만약 삭제된 오선과 겹친다면, 그 오선을 기준으로 전과 후의 이미지를 이용하여 그림 4와 같이 선을 연결하였다.

### ③ 음표의 줄기 인식

줄기는 무조건 오선의 선 두개 이상을 지난다는 특징을 활용하여, y좌표로 연속된 좌표의 길이가 오

선 두개의 길이보다 긴지를 판명하여, 줄기를 인식한다.



(a) 원본 악보 이미지 (b) 음표 구분 작업  
그림 1. 악보 이미지 전처리 과정

Fig. 1. Preprocessing of music score images



그림 2. 악보 원본 이미지  
Fig. 2. Original image of sheet music



그림 3. 오선 지우기  
Fig. 3. Erase staff



그림 4. 끊어진 음표 연결 작업  
Fig. 4. Reconnect broken notes

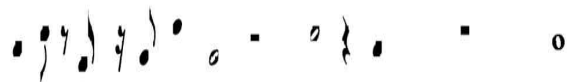


그림 5. 줄기 인식 및 지우기  
Fig. 5. Stem recognition and erasing

### ④ 줄기를 제외한 나머지 머리와 꼬리, 쉼표 구분하기

악보에서 줄기를 지우면, 그림 5와 같이 모든 머리와 꼬리는 분리가 되고, 악보 이미지에 머리, 꼬리, 쉼표만 남길 수 있다. 그렇게 남겨진 이미지에서 BFS를 이용하여 인접한 점을 찾아 하나씩 분리를 해준다.

이와 같이 전처리가 종료되면, 전처리를 통해 추출된 정보로부터 머리/꼬리/쉼표에 해당하는 후보들이 결정된다. 기존의 OMR 학습 방식에서는 음표의



종류를 모두 학습을 시킨다. 하지만, 본 논문은 머리의 종류 2가지 꼬리 종류 1가지를 인식시켜 총 3가지 종류로 모든 음표를 인식을 하고자한다. 또한, 이러한 후보를 단순한 CNN 기반 모델에 적용하여 인식을 수행하고자 한다. CNN은 내부에 수백 개의 은닉층을 가질 수 있으며, 은닉층의 깊이가 깊어질수록, 입력 이미지의 크기가 커질수록, 출력 대상의 개수가 많아질수록 내부 파라미터의 개수가 매우 커지게 된다[3]. 하지만, 본 논문은 경량 모델을 목표로 하므로, 은닉층이 3계층으로 구성된 간단한 CNN을 활용한다.

입력으로 주어진 머리/꼬리/섬표 중에서 가장 높은 확률을 가진 후보를 식별한 후, 이를 결합하면 음표와 섬표를 인식할 수 있다. 그 인식결과들을 Dynamic Programming을 이용하여, 머리와 꼬리의 조합을 통해, 4분음표, 2분음표 등으로 합쳐준다. 이것을 통해, 무슨 음표인지를 판단을 하여, 박자를 인식할수있다.

또한, 머리로 인식되어진 부분을 그전에 추출한 오선을 통해 오선과의 근접성을 확인함으로써, 인식된 음표의 음이 무엇인지 판단을 하도록 한다.

따라서, CNN으로 분류한 결과값을 DP를 통해 음표와 섬표를 결합하여, 인식한 이미지의 박자를 판단을 하고, CNN으로 분류한 결과값중 머리만 따로 가지고 와서, 오선지와의 근접성을 통해 인식한 이미지의 음을 판단을 한다.

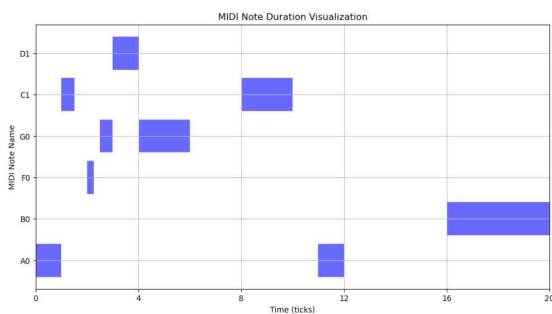


그림 6. 결과값 MIDI 그래프  
Fig. 6. Result MIDI Graph

그렇게 경량화한 CNN을 통해 분석한 결과로, 음표와 음을 판단을 하여, 악보 이미지를 다시 MIDI로 다시 넣어줄 수있다. 그렇게 MIDI에 들어간 결과는 그림 6에서 보는 것과 같이 어떤 음이 몇 박자동안 나오고 있는지, 몇박자를 쉬고 있는지를 확

인할 수 있다.

### III. 결 론

본 논문에서는 이미지로부터 패턴 인식에 널리 활용되는 CNN 기반 모델을 이용하여 경량화 OMR 모델을 개발하고자 한다. CNN 기반 모델의 성능을 높이기 위해서는 더 복잡한 모델이 개발되어야 한다. 하지만, 본 논문에서는 입력 이미지에 대하여 먼저 적절한 전처리 알고리즘을 수행함으로써 경량 모델을 개발하고자 하였다.

악보 이미지를 입력값으로 주면, 악보 이미지를 MIDI로 변환을 시키고 이러한 과정에서 이미지와 결과를 비교했을때, 거의 차이가 없게 나오는 것을 관측하였다.

본 논문에서는 머리의 종류와 꼬리의 개수를 통해 몇분음표인지를 인식하도록 하였고, 좋은 결과를 볼수있었다. 이 방법으로 모든 음표에 대한 데이터를 학습시키는 것이 아닌, 머리와 꼬리 인식으로 가능한 것으로 보이고, 온음표에서 16음표뿐만 아니라 32분음표 등도 가능할 것으로 예측할수있다.

현재 전처리 과정에서는 음표, 섬표, 마디선만 사용을 하였지만, 악보에는 더 많은 기호들이 있다. 지속적으로 전처리 가능한 기호의 폭을 넓여 모델을 개발하고자한다.

### 참 고 문 헌

- [1] SY Lee, HK Park, JH Kim and S.S Kim, "Development of online Hangul-handwritten recognition App", Proc. Korean Institute of Information Technology Fall Conference, pp. 216-218, 2022.
- [2] Haekwang Kim, "Score Image Retrieval to Inaccurate OMR performance", Journal of Broadcast Engineering, 26(7), pp. 838-843
- [3] LeCun Yann and Y. Bengio, "Convolutional networks for images, speech, and time series", The handbook of brain theory and neural networks 3361.10, 1995.
- [4] SG Seo, DO Yu, S.S Kim, SO Yang, "Preprocessing algorithm for light-weight Korean handwriting characters recognition model"

# 트리플렛 손실 함수를 활용한 CNN 기반 친족 관계 예측

이세연\*, 김주영\*, 박지홍\*, 김건우\*\*

## CNN-Based Kinship Prediction Model with Triplet Loss Function

Se-Yeon Lee\*, Ju-Young Kim\*, Ji-Hong Park\*, and Gun-Woo Kim\*\*

### 요약

친족 관계 예측은 얼굴 이미지 간의 유사성을 분석하여 혈연 관계를 판단하는 작업이다. 기존의 CNN 모델은 단순히 이진 분류를 통해 친족 관계를 예측하여 얼굴 특징 벡터 간의 거리 학습에 한계가 있었다. 본 연구에서는 이러한 한계를 극복하기 위해 트리플렛 손실 함수를 CNN에 결합한 새로운 모델을 제안한다. 제안된 모델은 부모-자녀 이미지 쌍의 거리는 최소화하고 부모-기타 인물 이미지 쌍의 거리는 최대화하도록 학습된다. ResNet-50, DenseNet-121, EfficientNet-B7 세 가지 CNN 구조와의 비교 실험 결과, EfficientNet-B7 기반의 제안 모델이 가장 우수한 성능을 보였으며, 기존 모델 대비 F1 Score가 4.28% 향상되었다. 이는 트리플렛 손실 함수의 적용이 친족 관계의 특징을 더 효과적으로 학습할 수 있음을 보여준다.

### Abstract

Kinship prediction is the task of determining blood relationships by analyzing similarities between facial images. Conventional CNN models have limitations in learning feature vector distances as they predict kinship relationships through simple binary classification. This study proposes a new model that combines triplet loss function with CNN to overcome these limitations. The proposed model is trained to minimize the distance between parent-child image pairs while maximizing the distance between parent-unrelated person image pairs. Comparative experiments with three CNN architectures - ResNet-50, DenseNet-121, and EfficientNet-B7 - showed that the EfficientNet-B7-based proposed model achieved the best performance, with a 4.28% improvement in F1 Score compared to the existing model. This demonstrates that the application of triplet loss function can more effectively learn kinship features.

### Key words

Kinship Prediction, Facial Feature Extraction, Triplet Loss, Convolutional Neural Networks

---

\* 경상국립대학교 컴퓨터과학부, 2023014356@gnu.ac.kr, wndudwkd003@gnu.ac.kr, hong\_0002@gnu.ac.kr

\*\* 경상국립대학교 교신저자 컴퓨터과학부, gunwoo.kim@gnu.ac.kr

※ 본 논문은 2023년도 정부(교육부)의 지원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2021R1G1A1006381)

## I. 서론

현대 사회에서 이산가족과 실종아동 문제는 현대 사회에서 해결해야 할 중요한 과제이다. 이산가족은 전쟁, 자연재해, 경제적 이유 등으로 인해 가족이 분리되는 것이다. 이는 개인의 정서적, 정신적 건강에 부정적 영향을 끼친다. 실종아동 문제 또한 매년 수많은 아동이 실종되며 심각한 사회적 이슈로 자리 잡고 있다. 이 문제들을 해결하기 위해 우리는 그들의 친족 관계를 정확하게 파악하는 것은 필수적인 단계이다.

딥러닝 기반 합성곱 신경망(CNN)은 여러 층으로 구성되어 있어 입력 이미지의 구조를 확인하고, 특징을 추출한다. CNN은 주로 이미지 인식 및 처리에 사용되며 친족 예측과 같은 이진 분류 문제에서는 주로 이진 교차 엔트로피 손실(BCE)을 사용하여 모델의 예측 확률과 실제 레이블 간의 차이를 측정한다. 하지만 단순히 예측 확률을 비교하는 기존 CNN은 부모-자녀 이미지 간에 유사성과 부모-기타 인물 이미지의 차이를 학습하는 데 한계가 존재한다.

본 연구에서는 이미지 관계를 더욱 명확하게 반영함으로써 기존 CNN의 한계점을 극복하기 위해 부모, 자녀, 기타 인물에 각각 CNN을 적용해 특징을 추출한다. 더하여 트리플렛 손실(Triplet Loss)을 활용하는 접근 방식을 채택하여 친족 관계 예측이 앞서 밝힌 사회적 문제들을 해결할 수 있는 발판이 되고자 한다.

## II. 관련 연구

### 2.1 ResNet-50

잔차 학습(Residual learning)이란 입력과 출력 간의 차이를 학습하는 방법이다. 일반적인 신경망에서는 층이 깊어질수록 각 층에서 계산된 기울기가 작아져 기울기 소실 문제를 발생시킨다. 반면 잔차 학습에서는 각 층의 출력에 입력을 더해줘서 층이 깊어져도 초기 입력 정보가 보존되어 기울기가 소실되는 것을 방지한다. 그림 1은 잔차 학습 블록의 구성을 나타낸다. 식(1)은 그림 1을 설명하는 수식이다.

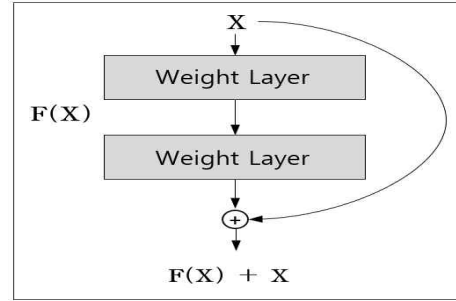


그림 1. 잔차 학습 블록  
Fig. 1. Residual learning block

$$H(x) = F(x) + x \quad (1)$$

### 2.2 EfficientNet-B7

EfficientNet은 깊이(Depth), 너비(Width), 해상도(Resolution) 세 가지를 동시에 조정하는 방법을 사용한다. EfficientNet-B7은 기존의 단일 차원만 확정하는 GPipe 모델과 비교하여 8.4배 더 적은 파라미터와 6.1배 더 빠른 속도를 보였다[3]. 이는 복합 스케일링 방법이 단일 스케일링 방법보다 효율적인 특징 학습한다는 점을 보여준다. 식(2)는 깊이, 너비, 해상도를 포함한 입력 데이터 X와 뉴런의 출력을 결정하는 활성화 함수 F를 설명한다. 인덱스 i에 따라 활성화 함수가 각기 다를 수 있다는 것을 뜻한다[3].

$$N(d, w, r) = \bigodot_{i=1 \dots s} F_i^{d \circ L_i}(X_{< r \circ H_i, r \circ W_i, w \circ C_i >}) \quad (2)$$

### 2.3 DenseNet 121

DenseNet은 모든 레이어의 출력이 다음 레이어에 전달되는 Dense Block 구조를 갖는다. 이러한 구조는 모델이 다양한 정보를 학습할 수 있도록 도와주고 원래 입력 신호와 각 레이어의 출력 신호가 효과적으로 보존되어 네트워크가 깊어지더라도 기울기 소실 문제가 완화된다. 식 (3)은 DenseNet의 기본 수식을 나타내며, 식에서  $[x_0, x_1, \dots]$ 는 각 레이어에서 만들어진 특징 맵의 연쇄(concatenation)를 의미한다.

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (3)$$

### III. 친족 예측 분류를 위한 딥러닝 모델 제안

본 연구에서는 부모, 자녀, 기타 인물에 각각 CNN을 적용하여 트리플렛 로스를 결합한 모델 방법을 제안한다. 이 모델은 기존 CNN 모델이 가지고 있지 않은 이미지 간의 관계를 파악하는 방식을 보여줌으로써 기존 CNN 모델의 문제점을 해결한다. 그림 2는 비교 실험 결과로 가장 뛰어난 성능을 보인 EfficientNet-B7 구조도이다. 그림 3은 Triplet loss가 우리 모델에서 작동하는 모습을 구조도로 나타낸 것이다. 1) 사전 훈련된 EfficientNet-B7 모델의 입력으로 이미지가 전달된다. 2) EfficientNet-B7의 7개의 Block을 거친다. 3) 각 이미지에 대한 특징 벡터가 생성된다. 4) 아빠 특징과 엄마 특징을 합쳐 부모 특징을 만든 다음, 트리플렛 손실을 계산하기 위해 부모 특징을 Anchor, 자녀 특징을 Positive, 기타 인물 특징을 Negative로 정의한다. 6) 부모-자녀 (Anchor - Positive) 거리는 최소화하고, 부모-기타 인물(Anchor-Negative) 거리는 최대화하며 훈련한다.

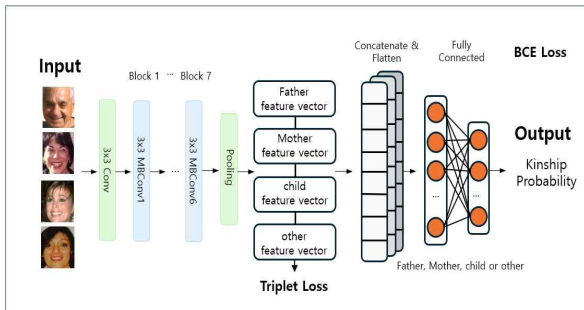


그림 2. EfficientNet-B7 구조에 트리플렛 손실 함수를 추가한 친족 예측 분류 모델 아키텍처

Fig. 2. EfficientNet-B7-Based Kinship Prediction Classification Model Architecture with Triplet Loss

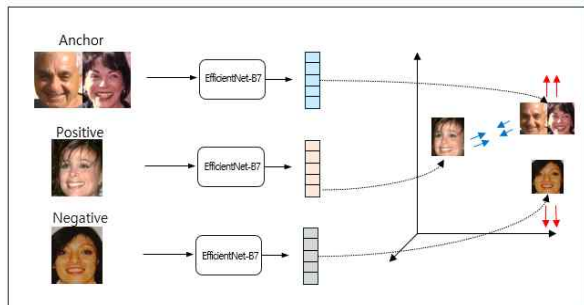


그림 3. 트리플렛 손실 함수 작동 방식

Fig. 3. Triplet Loss Function Operation Mechanism

### IV. 실험 설계

#### 4.1 데이터 세트 구성과 분할

본 연구의 데이터 세트는 부모와 자녀 그리고 다른 사람의 이미지를 포함한 단일 폴더를 생성하였다. 전체의 80%는 훈련 데이터, 나머지 20%는 테스트 데이터로 사용했다. 이때 무작위로 분할 하여 데이터의 다양성을 보장하고자 했다.

#### 4.2 정규화

이미지 전처리 단계에서 모든 데이터 세트의 이미지를 224x224 크기로 조정하였으며, 모델 학습의 일관성을 위해 정규화를 수행하였다. 정규화 과정에서는 각 이미지의 픽셀값을 [0, 1] 범위로 변환 후, 평균 [0.5, 0.5, 0.5]와 표준 편차[0.5, 0.5, 0.5]를 사용하여 스케일링을 진행했다.

#### 4.3 하이퍼파라미터 설정 및 최적화

파라미터 최적화를 위해 AdamW 옵티마이저를 사용했다. 안정적인 수렴을 위해 학습률은 0.0002로 설정하여 초기 학습 속도를 적절히 유지하고자 하였다. 또한 ReduceLROnPlateau 스케줄러를 설정하여 5 에포크 동안 검증 손실이 개선되지 않을 경우, 학습률을 0.5배로 감소시키도록 하였다. 이는 더 이상 개선되지 않는 지역 최적점 문제를 피할 수 있다.

### V. 실험 결과

단순 CNN 기반 모델들의 Recall 점수가 크게 변화한 것을 볼 수 있다. ResNet-50 같은 경우 Recall 점수가 약 27.33%P 증가했다. EfficientNet-B7이 가장 성능이 좋게 나왔으며 본 연구에서 제안한 모델의 F1 Score가 기존 EfficientNet-B7보다 약 4.28% 향상되었다.

표 1은 ResNet-50 비교 실험 결과이다. 표 2는 DenseNet-121, 표 3은 EfficientNet-B7의 비교 실험 결과다. 각각 성능지표에서 최고 성능의 값을 나타낸다. 모든 실험에서 제안하는 모델이 기존 모델보다 개선된 성능을 보여주는 것을 확인할 수 있다.

표 1. ResNet-50 실험 데이터 세트

Table 1. Experimental data sets ResNet-50

평가 지표	ResNet-50	ResNet-50 Triplet Loss
Accuracy	0.5135	<b>0.5434</b>
Precision	0.5439	<b>0.5667</b>
Recall	0.5465	<b>0.8198</b>
F1- Score	0.6709	<b>0.6766</b>
ROC AUC	0.5662	<b>0.5754</b>

표 2. DenseNet-121 실험 데이터 세트

Table 2. Experimental data sets DenseNet-121

평가 지표	DenseNet-121	DenseNet-121 Triplet Loss
Accuracy	0.5403	<b>0.6645</b>
Precision	0.5844	<b>0.8421</b>
Recall	0.5341	<b>0.6397</b>
F1- Score	0.6351	<b>0.6838</b>
ROC AUC	0.5756	<b>0.7223</b>

표 3. EfficientNet-B7 실험 데이터 세트

Table 3. Experimental data sets EfficientNet-B7

평가 지표	EfficientNet-B7	EfficientNet-B7 Triplet Loss
Accuracy	0.6428	<b>0.6645</b>
Precision	0.5886	<b>0.6086</b>
Recall	0.6086	<b>0.8881</b>
F1- Score	0.6695	<b>0.6997</b>
ROC AUC	0.6003	<b>0.731</b>

## 참 고 문 헌

- [1] F. Schroff, D. Kalenichenko, and J. Philbin, "Deep metric learning using triplet network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 470-477, 2015.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", arXiv preprint arXiv:1512.03385, December 2015.
- [3] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", Proceedings of the 36th International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.
- [4] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks", 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

## VI. 결 론

본 연구는 친족 관계 예측을 위한 딥러닝 모델을 개발하고, 여러 성능지표를 통해 비교 분석을 진행했다. 기존의 CNN 모델들이 친족 관계 예측에 보이는 한계점을 극복하기 위해, 부모, 자녀, 기타 인물에 각각 CNN을 적용하고, Triplet Loss를 활용하여 범주 간의 유사성을 학습하는 방식을 제안했다.

연구 결과, 본 논문에서 제안한 모델은 정확도, 정밀도, 재현율, F1 Score, ROC AUC 등 모든 성능 지표에서 단순 CNN 기반 모델보다 더 나은 성능을 보였으며, 이는 본 모델의 구조와 Triplet Loss를 활용한 학습 방식이 기존 모델보다 친족 관계와 같은 복잡한 유사성 구조를 학습하는데 매우 적합하다는 것을 입증하였다.

# 수질 데이터 결측치 보간법에 따른 AI 모델 성능 비교 연구

이상민\*, 김동현\*, 김재엽\*, 이민지\*, 최승호\*\*, 이석훈\*†

## The Impact of Imputation Methods on AI Model Performance in Water Quality Data

Sangmin Lee\*, Donghyun Kim\*, Jaeyeab Kim\*, Minji Lee\*, Seungho Choi\*\*, and Sukhoon Lee\*†

### 요약

본 연구는 수질 데이터에서 발생하는 결측치를 보완하는 다양한 보간법이 AI 예측 모델 성능에 미치는 영향을 비교 평가하고자 한다. 수질 데이터는 환경적 요인으로 인해 결측치 발생이 일어날 수 있고, 이러한 결측치는 예측 정확도에 큰 영향을 미칠 수 있다. 본 연구에서는 선형 보간법, 평균 보간법, 스플라인 보간법, 이동 평균 보간법을 적용하여 결측치를 대체한 후, 시계열 AI 모델 GRU를 통해 보간법에 따른 예측 성능 차이를 분석한다. NSE 지표를 활용하여 각 보간법의 결과를 정량적으로 평가하였다. 실험 결과, 선형 보간법과 이동 평균 보간법이 우수한 성능, 다음으로 스플라인 보간법, 평균 보간법 순으로 성능이 측정되었다.

### Abstract

This study evaluates the impact of various interpolation methods on AI prediction model performance in addressing missing values in water quality data. Missing values, often due to environmental factors, may significantly affect prediction accuracy. Here, missing values are replaced using linear interpolation, mean imputation, spline interpolation, and moving average imputation, with differences in prediction performance analyzed through the GRU time-series model. Each method's results are evaluated quantitatively using the NSE metric, showing that linear and moving average methods perform best, followed by spline and mean imputation.

### Key words

missing value, interpolation

## 1. 서론

최근 데이터는 다양한 분야에서 활발히 생산되고 있으며, 기존보다 그 종류와 정보량이 크게 증가하

\* 국립군산대학교 소프트웨어학과

† 교신저자

\*\* 국립군산대학교 소프트웨어융합공학과

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음 (2023-0-00065).

고 있습니다. 특히, 인공지능 기술의 급속한 발전과 다양한 AI 모델의 출현은 데이터 품질이 AI 성능에 중요한 역할을 하게 만들고 있습니다. 데이터가 결측되거나 품질이 낮을 경우, 인공지능 모델의 성능에도 직접적인 영향을 미칠 것이다[1]-[3].

본 연구는 수질 데이터를 대상으로 다양한 결측치 보간법에 따른 인공지능 모델 성능을 비교하는 것을 목적으로 합니다. 수질 데이터는 다양한 환경요인에 따라 결측치가 발생할 가능성이 높으며, 이로 인해 예측 모델의 성능이 큰 영향을 받을 수 있습니다. 본 연구에서는 선형 보간법, 평균 보간법, 스플라인 보간법, 이동 평균 보간법을 사용하여 결측치를 대체하고, GRU 모델을 통해 각각의 보간법이 모델 성능에 미치는 영향을 비교 평가하고자 합니다.

## II. 학습 데이터셋

본 연구에서 사용된 데이터는 AI 허브에서 제공하는 “수질 측정 및 오염원 데이터”의 자동측정망 데이터를 활용하였다. 자동측정망 데이터는 ymdh, stn\_name, site\_id 등으로 구성되며 59개의 컬럼으로 총 11,686,248건의 데이터이며 2013 ~ 2020년까지의 데이터이다[4].

수질 데이터셋의 정제는 다음과 같은 순서로 진행된다. 1차적으로 물환경측정망 운영계획 고시 등 관련 법령에 근거하여 1차적으로 데이터가 정제되고 2차적으로 중위 절대편차를 사용해 신뢰구간을 이용한 이상치 판단 기법으로 95%의 신뢰도에 따라 신뢰구간 안에 포함되지 않으면 이상치로 제거하였다[5].

AI 모델의 입력 데이터로는 수온, pH, DO(용존산소량), EC(전기전도도), TOC(총유기탄소), T-N(총질소), T-P(총인), CHL-a(클로로필-a) 총 8개 항목을 사용하였다. 이 중 DO, TOC, T-N, T-P, CHL-a 5개 항목에 대해 예측을 수행하였다.

## III. 결측치 보간법

본 연구에서 사용한 결측치 처리를 위해 사용한

보간법은 선형 보간법, 평균 보간법, 스플라인 보간법, 이동 평균 보간법을 사용했으며, 먼저 선형 보간법은 두 데이터 점 사이의 직선을 이용해 결측값을 채우는 방법이며, 결측값이 존재하는 구간에서 앞뒤 값의 비율을 사용해 값을 추정한다. 이동 평균 보간법은 특정 기간의 데이터들의 평균값을 이용하여 결측값을 추정하는 방법이며, 시간에 따른 가중치를 부여하여 계산한다. 스플라인 보간법은 구간별로 다항식을 이용하여 곡선을 맞추는 방법으로, 결측값 전후의 여러 관측값들을 이용하여 결측값을 추정한다. 평균 보간법은 결측값을 동일한 열 또는 데이터의 평균값으로 대체하는 단순한 방법이다.

## IV. 실험 및 평가

연구에서는 AI 모델의 결측치 보간법에 따른 성능 변화를 분석하기 위해, 결측치가 없는 수질 데이터에 임의로 결측치를 생성하여 실험을 진행하였다. 결측치 생성 비율은 10%에서 시작하여 10% 단위로 증가시키며 최대 30%까지 설정하였다. 각 결측치 비율에서 다양한 보간법을 적용하여 결측치를 대체하고, AI 모델의 예측 성능을 평가한다. 보간법으로는 선형 보간법, 평균 보간법, 스플라인 보간법, 이동 평균 보간법을 적용하였다.

연구에서는 시계열 데이터를 처리하기에 적합한 GRU 모델을 사용하였다. 각 모델은 학습 및 평가를 위해 훈련 데이터와 테스트 데이터로 분리하였으며, 동일한 조건으로 결측치 보간법의 성능 차이를 비교한다.

각 모델의 성능을 평가하기 위해 NSE(Nash-Sutcliffe Efficiency)라는 지표를 사용했다. NSE는 -∞에서 1사이의 값을 가지며, 성능이 0보다 작다면 모델의 예측값이 관측값의 평균값으로 대체하는 것보다 못함을 의미하며 1에 가까울수록 모델의 예측값이 관측값을 정확히 예측한다는 의미이다.

표 1은 보간법에 따른 예측 모델 성능 비교 결과를 보여준다. 수질 데이터에 10%부터 30%까지의 인위적 결측치를 생성하여 네 가지 보간법과 AI 모델의 성능을 비교 분석하였다. 선형 보간법과 이동 평균 보간법이 전반적으로 가장 우수한 성능을 보

표 1. 성능 비교 결과

Table 1. Performance comparison results

보간법	결측치	DO NSE	TOC NSE	T-N NSE	T-P NSE	CHL-a NSE
선형보간법	10%	0.9575	0.9059	0.9712	0.9293	0.9400
선형보간법	20%	0.9606	0.9017	0.9719	0.9372	0.9374
선형보간법	30%	0.9605	0.9053	0.9735	0.9421	0.9267
평균 보간법	10%	0.7514	0.7271	0.7691	0.7322	-0.3242
평균 보간법	20%	0.5373	0.5534	0.5780	0.5276	-0.6141
평균 보간법	30%	0.4335	0.4405	0.4504	0.3239	-0.6881
스플라인보간법	10%	0.9107	0.6667	0.7884	0.7151	0.7010
스플라인보간법	20%	0.8925	0.5942	0.6440	0.5723	0.7219
스플라인보간법	30%	0.8853	0.4642	0.5408	0.5176	0.7813
이동평균보간법	10%	0.9661	0.9612	0.9769	0.9610	0.9113
이동평균보간법	20%	0.9645	0.9637	0.9776	0.9585	0.8902
이동평균보간법	30%	0.9643	0.9626	0.9774	0.9629	0.9108

였다. 결측치 비율 증가에도 안정적인 일관된 NSE값을 유지하였다. 스플라인 보간법은 선형 보간법과 이동 평균 보간법에 비해 다소 낮은 성능을 보였으며 DO 항목에서만 비교적 안정적인 성능을 유지하였다. 마지막으로 평균 보간법은 가장 낮은 성능을 보였고 결측치 비율 증가에 따른 성능 저하가 가장 뚜렷하게 나타났다. 특히 CHL-a 항목은 NSE 값이 음수를 나타내며 심각한 성능 저하를 보였다. 이는 평균 보간법이 시계열 데이터 특성을 반영하지 못한다는 것을 알 수 있다. 수질 데이터의 결측치 처리에 있어 선형 보간법과 이동 평균 보간법이 가장 효과적임을 보여준다.

## V. 결론

본 논문에서는 다양한 결측치 보간 방법이 수질 데이터 기반 AI 예측 모델의 성능에 미치는 영향을 비교 분석하였다. 실험 결과 AI 모델 성능은 보간법에 따라 큰 차이를 보였다. 선형 보간법과 이동 평균 보간법이 우수한 성능을 보였으며, 다음으로 스플라인 보간법, 평균 보간법 순으로 성능이 측정되었다.

## 참고 문헌

- [1] Park Kyung Eun, Park Jeong Ha, Lee Jin Yong, and Lee Sang Bok, "Method and Application Case of Quality Verification of File Data in Structured Format(csv) Using ISO/IEC 25024 Quality Characteristics", Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp. 1699-1700, 2024-06-19.
- [2] Korea Data Industry Promotion Agency. "2023 Data Industry White Paper." pp. 76-81
- [3] S. Ha and S. Shin, "Data Quality Standardization Trends for Artificial Intelligence", OSIA Standards & Technology Review Journal, vol. 35, no. 1, pp. 15-19, 2022.
- [4] <https://www.aihub.or.kr/> [Accessed: Oct. 30, 2024]
- [5] Kim, D.S., Choi, H.D., Roh, J.D., Kim, S.I., Lee, S.J., Kim, J.S., Kim, Y.B., "Artificial Intelligence Data Construction and Utilization Guidelines v2.4.1 - Environmental Pollution AI Data (Water Quality Measurement and Pollution Source Data)", 2021.



# LLM기반 리뷰 데이터 생성을 통한 교차 도메인 추천 시스템 구축

강지희\*, 최상민\*\*

## Building a Cross-Domain Recommendation System through LLM-Based Review Data Generation

Jihee Kang\*, and Sang-Min Choi\*\*

### 요 약

본 논문에서는 Large Language Model(LLM)을 이용한 리뷰 데이터 생성 방식을 활용하여 Cross-Domain Recommendation system(CDRs)의 성능을 개선하는 방법을 제안한다. 추천 시스템은 협업 필터링과 같은 알고리즘을 기반으로 발전해왔지만, cold start와 데이터 희소성 문제는 여전히 주요한 과제로 남아 있다. 특히, CDRs에서도 도메인 간 데이터 불균형으로 인해 성능 저하가 발생할 수 있다. 이를 개선하기 위해, 본 연구는 LLM을 기반으로 특정 도메인의 리뷰데이터를 학습한 후, 다른 도메인에서 리뷰 데이터를 생성하는 방법론을 제시한다. 생성된 리뷰 데이터를 Dual-Target Cross-Domain Recommendation(DTCDR) 모델에 적용하여 CDRs의 성능을 평가하고, 생성한 리뷰 데이터와 기존 데이터 간 성능을 비교한다. Amazon Review Data 2023과 MovieLens 100K를 데이터셋으로 사용해 향후 다양한 연구에서 LLM을 통한 데이터 생성이 CDR의 데이터 부족 문제 해결에 활용 가능성을 갖는다는 점을 보여준다.

### Abstract

In this paper, we propose a method to improve the performance of cross-domain Recommendation systems(CDRs) by utilizing the review data generation using the Large Language Model(LLM). Recommendation systems have been developed based on algorithms such as collaborative filtering, but cold start and data sparsity issues remain major challenges. In particular, performance degradation can occur even in CDRs due to data imbalance between domains. To improve this, this study proposes a methodology to generate review data in another domain after learning review data in a specific domain based on LLM. The generated review data is applied to the Dual-Target Cross-Domain Recommendation (DTCDR) model to evaluate the performance of CDRs and compare the generated review data with the existing data. Using Amazon Review Data 2023 and MovieLens 100K as datasets, it shows that data generation through LLM has the potential to be utilized to solve the data shortage problem of CDRs in various future studies.

### Key words

cross-domain, recommender system, large language model, review data

---

\* 경상국립대학교 IT공과대학 컴퓨터공학과, jhk3261@gnu.ac.kr

\*\* 경상국립대학교 IT공과대학 컴퓨터공학과, jerassi@gnu.ac.kr (교신저자)

## I. 서론

추천 시스템은 사용자에게 맞춤형 콘텐츠나 제품을 제공하기 위해 다양한 산업에서 폭넓게 활용되고 있다[1]. 대표적인 알고리즘인 협업 필터링은 사용자와 아이템 간의 상호작용을 기반으로 추천을 제공하지만, cold start와 데이터 희소성 문제라는 한계를 지닌다[2,3]. 이러한 문제를 해결하려는 방법의 하나로 Cross-Domain Recommendation(CDR)이 제안되었으며, 이는 두 개 이상의 도메인 간 데이터를 활용하여 성능을 개선하는 방법으로 알려져 있다[3].

리뷰 데이터도 데이터 희소성 문제를 해결하는데 중요한 역할을 하며, 사용자의 의견을 반영해 추천의 정확도와 설명 가능성을 높인다. 연구에 따르면, 리뷰 데이터를 활용한 추천 시스템이 평점 기반 시스템보다 성능이 우수하며, 이를 통해 사용자 선호도를 더욱 깊이 이해할 수 있다[5].

이에 본 연구에서는 LLM을 활용해 타깃 도메인의 리뷰 데이터를 학습하고, 소스 도메인의 리뷰 데이터를 생성하는 방법을 제안한다. 생성된 리뷰 데이터를 추천 모델에 적용하여 기존 리뷰 데이터와 유사한 성능을 보여준다. 이를 통해 LLM을 활용한 리뷰 데이터 생성이 추천 시스템의 성능을 유지하는 동시에, 향후 다양한 CDR 연구에 유의미한 데이터를 제공할 가능성을 시사한다.

## II. 방법론

### 2.1 LLM 기반 리뷰데이터 생성

본 논문에서는 타깃 도메인의 리뷰 데이터를 학습한 LLM을 통해 소스 도메인의 리뷰 데이터를 생성하는 방식을 제안한다. 구체적인 절차는 다음과 같다.

소스 도메인의 리뷰 텍스트 필드는 비우고, 나머지 필드(user, item, rating)만 유지한다. 이후 특정 도메인의 리뷰 데이터를 바탕으로 GPT-4o 기반 LLM을 사용해 소스 도메인의 리뷰 데이터를 생성하며, 평점과 아이템 메타데이터 정보를 활용한 프

롬프트 엔지니어링으로 자연스러운 텍스트를 생성할 수 있도록 한다. 생성된 리뷰 데이터를 소스 도메인의 기존 데이터에 통합하고, 생성 텍스트가 타깃 도메인의 리뷰와 유사한 형태와 의미를 유지하는지 검토한다. 표 1은 생성된 리뷰 데이터와 기존 리뷰 데이터 간 유사도를 비교한 결과이다.

표 1. 리뷰 데이터 유사도  
Table 1. Review Data Similarity

Metric	Average Result
Precision	0.84
Recall	0.85
F1 score	0.84
Cosine Similarity (BERT Embedding)	0.71

### 2.2 DTCDR모델

본 연구는 CDR 성능 평가를 위해 Dual-Target Cross-Domain Recommendation(DTCDR) 모델을 활용한다[6]. DTCDR 모델은 Multi-Task Learning(MTL)을 통해 두 도메인 간 공통 사용자 데이터를 공유하여 추천 성능을 향상시킨다. DTCDR은 소스 도메인의 정보가 타깃 도메인 추천 성능을 높이는 동시에, 타깃 도메인 정보도 소스 도메인에 활용될 수 있도록 한다. DTCDR 모델 구조는 그림 1에 나타나 있다.

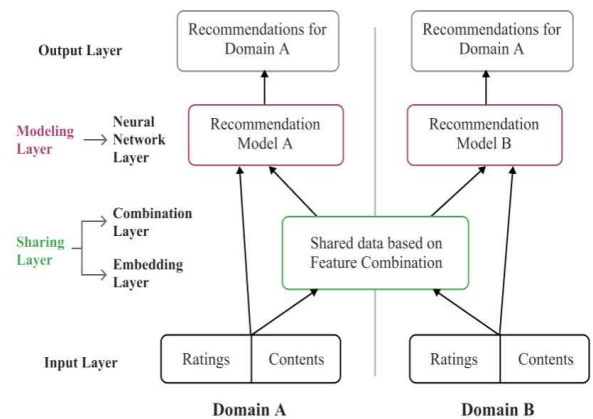


그림 1. DTCDR 모델 구조  
Fig. 1. DTCDR model structure

### III. 실험 및 결과

#### 3.1 데이터 수집

본 논문에서는 Amazon의 Review Data 2023의 Movie, Book과 MovieLens 100K의 Movie Dataset을 이용하여 실험을 진행한다. Amazon의 두 도메인의 Dataset이 매우 크기 때문에 실험의 효율성을 위해 일부 데이터를 추출하여 사용한다. 이후 각 도메인에서 필요한 필드를 필터링하고 불필요한 데이터 및 결측치를 제거하거나 다른 값으로 대체하여 데이터 품질 향상을 위해 전처리를 진행한다.

표 2. 실험 데이터 세트

Table 2. Experimental data sets

Datasets	Amazon Review		MovieLens
Domains	Movie	Book	Movie
#User	5,950	6,793	943
#Item	11,001	13,773	1,683
#Reviews	15,000	15,000	100,000

#### 3.2 성능 평가

추천 성능 평가를 위해 본 연구에서는 Recall, MRR, NDCG, Hit 3개의 평가지표를 활용한다. Recall은 모든 아이템 중에서 얼마나 많은 아이템을 추천 시스템이 찾아냈는지를 나타낸다. MRR은 추천된 아이템의 순위에 따라 성능을 측정한다. NDCG는 추천된 아이템의 순위에 따라 추천의 질을 평가하며, 상위에 있을수록 높은 점수를 받는다. Hit Rate는 추천된 목록 안에 사용자가 실제로 선호하는 아이템이 포함되어 있는지를 측정한다.

#### 3.3 실험 결과

본 연구에서는 LLM을 활용하여 생성된 리뷰 데이터와 기존 리뷰 데이터를 비교하여 CDR의 성능을 평가하였다. 이를 위해 Amazon Movie와 MovieLens Movie 도메인 간의 추천(a), 그리고 Amazon Movie와 Book 도메인 간의 추천(b) 두 가지 실험을 수행하였으며, 각 성능 지표를 기준으로

기존 데이터와 생성 데이터의 성능을 분석하였다. 실험 결과는 그림 2와 같다.

두 가지 실험 모두 생성된 데이터가 기존 데이터와 유사한 성능을 나타냈다. (a)에서는 생성 데이터가 기존 데이터와 거의 동일하거나 소폭 개선된 성능을 보였고, (b)에서는 두 데이터 간의 성능 차이가 미미하게 나타났다. 이러한 결과는 LLM을 활용한 리뷰 데이터 생성이 기존 데이터의 대체 또는 보완 역할을 충분히 할 수 있음을 보여준다.

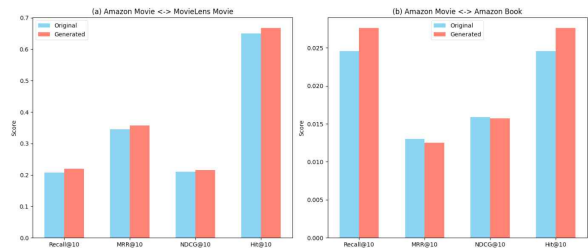


그림 2. 기존 데이터와 생성 데이터를 사용한 성능 비교  
Fig 2. Performance comparison using existing and generated data

### IV. 결 론

본 연구는 LLM을 활용해 생성한 리뷰 데이터를 통해 교차 도메인 추천 시스템(CDR)에서의 추천 성능을 개선하는 방안을 제안한다. 실험 결과, 생성한 리뷰 데이터가 기존 리뷰 데이터와 유사한 성능을 나타낸다는 것을 보여주었다. 이는 LLM을 활용한 리뷰 데이터 생성이 기존 데이터에 비해 실질적인 성능 차이를 보이지 않음을 확인함으로써 향후 다양한 CDR 연구에서 생성 데이터의 활용 가능성을 제시한다. 특히, LLM을 통해 생성한 데이터가 추천 성능에 미치는 영향을 기존 데이터와 비교하여 정량적으로 분석하는 후속 연구도 필요하다.

또한, 생성된 리뷰 데이터를 감정 분석과 연계함으로써 사용자 감정과 선호도를 반영한 정교한 추천 시스템을 구축하여 추천의 정확도와 사용자 만족도를 높일 수 있을 것이다. 더불어, 데이터가 부족한 상황에서 LLM을 활용한 데이터 증강을 통해 추천 시스템의 데이터 다양성과 품질을 높이고, cold start 문제를 완화하는 데에도 기여할 수 있을 것으로 기대된다.

참 고 문 헌

- [1] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowledge Data Eng.*, vol. 17, no. 6, pp. 734-749, Jun. 2005.
- [2] Y. Koren, S. Rendle, and R. Bell, "Advances in collaborative filtering," in *Recommender Systems Handbook*, 2nd ed., Springer, 2021, pp. 91-142.
- [3] T. Zang, Y. Zhu, H. Liu, R. Zhang, and J. Yu, "A survey on cross-domain recommendation: Taxonomies, methods, and future directions," *ACM Trans. Inf. Syst.*, vol. 41, no. 2, pp. 1-39, Feb. 2022.
- [5] Y. Choi, J. Choi, T. Ko, H. Byun, and C. K. Kim, "Review-Based Domain Disentanglement without Duplicate Users or Contexts for Cross-Domain Recommendation," in *Proc. 31st ACM Int. Conf. Information and Knowledge Management*, Oct. 2022, pp. 293-303.
- [6] F. Zhu, C. Chen, Y. Wang, G. Liu, and X. Zheng, "DTCDR: A framework for dual-target cross-domain recommendation," in *Proc. 28th ACM Int. Conf. Information and Knowledge Management*, Nov. 2019, pp. 1533-1542.

# 이더리움 온·오프체인 데이터 기반 주소 식별 방법

강성준\*, 서정훈\*, 정현준\*\*

## Method for Identifying Ethereum Addresses Based on On-Chain and Off-Chain Data

Seongjun Kang\*, Jeonghun Seo\*, Hyunjun Jung\*\*

### 요약

최근 가상자산 해킹 사건의 급증으로 블록체인 주소 식별과 라벨링에 대한 관심이 높아지고 있다. 기존 온체인 데이터만으로는 주소의 실제 성격 파악에 한계가 있어 오프체인 데이터의 활용이 필수적이다. 이 논문에서는 온체인과 오프체인 데이터를 결합하여 이더리움 주소를 분류하는 방법을 제안한다. 실시간으로 이더리움 Geth를 통해 블록 데이터를 수집하고, 수집한 데이터를 변환한 후 CatBoost 모델을 통해 주소를 분류한 뒤, Etherscan 메타데이터 API로 검증한다. 검증된 주소는 라벨링 테이블에 저장되며, 정보가 없는 주소는 임시 테이블에 저장된다. 이 방법은 주소 식별의 정확성을 높여 가상화폐 생태계의 안전성과 신뢰성 향상에 기여할 것으로 기대된다.

### Abstract

Due to the recent surge in cryptocurrency hacking incidents, there is increasing interest in the identification and labeling of blockchain addresses. There are limitations in accurately understanding the actual nature of addresses using only existing on-chain data, making the utilization of off-chain data essential. In this paper, we propose a system that classifies Ethereum addresses by combining on-chain and off-chain data. Real-time block data is collected through Ethereum Geth, and after transforming the collected data, addresses are classified using a CatBoost model, followed by verification using the Etherscan metadata API. Verified addresses are stored in a labeling table, while addresses without information are stored in a temporary table. This system is expected to improve the accuracy of address identification, contributing to enhanced safety and reliability in the cryptocurrency ecosystem.

### Key words

onchain, ethereum, blockchain, crypto, label, offchain

---

- 2024 한국정보기술학회 하계 종합학술대회 논문집 -

- 1 -

\* 국립군산대학교 소프트웨어학과 학사과정, wnswns1946@kloint.co.kr, didam00@naver.com

\*\* 군산대학교 소프트웨어학부 교수(교신저자), junghj85@kunsan.ac.kr

※이 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되고 있습니다.(No. NRF-2022R1G1A1008493).

※본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065)

※ 지원기관표기(사사표기)

## I. 서론

최근 가상자산 해킹 사건이 빈번하게 발생하면서 블록체인 주소의 식별과 라벨링 중요성이 부각되고 있다. 블록체인은 거래의 투명성을 제공하지만, 익명성과 탈중앙화 특성으로 인해 악의적인 활동을 추적하고 방지하는 데 어려움이 있다. 이러한 상황에서 블록체인 주소의 실제 성격과 의도를 정확히 파악하기 위해서는 온체인 데이터와 오프체인 데이터의 활용이 필수적이다[1].

온체인 데이터는 블록체인 네트워크 내에서 발생하는 거래 내역, 잔액, 스마트 컨트랙트 상호작용 등의 정보를 포함한다. 이 데이터는 거래의 흐름과 패턴을 분석하는 데 유용하지만, 주소 소유자의 신원이나 그들이 수행하는 활동의 맥락을 제공하는 데는 한계가 있다. 온체인 데이터만으로는 특정 주소가 어떤 목적을 가지고 있는지, 또는 어떤 기관이나 개인과 연관되어 있는지를 판단하기 어렵다.

오프체인 데이터는 블록체인 외부에서 생성되지만, 블록체인 네트워크의 비즈니스 모델과 밀접하게 연관된 정보를 의미한다. 이는 단순히 모든 외부 데이터를 포함하는 것이 아니라, 블록체인 생태계와 관련된 소셜 미디어 활동, 공식 웹사이트, 뉴스 기사, 그리고 각종 메타데이터 등을 포함한다. 오프체인에서 발생하는 범죄 사실이나 의심스러운 활동에 대한 정보는 블록체인 주소의 위험성을 평가하는 데 필수적이다. 이러한 정보는 온체인 데이터로는 감지하기 어려운 패턴이나 이상 징후를 포착하는 데 도움이 된다.

온체인과 오프체인 데이터를 통합하여 분석하는 것은 블록체인 주소의 실제 성격을 보다 정확하게 파악하고, 악의적인 활동을 효과적으로 감지하고 방지하기 위한 핵심적인 접근 방식이다. 온체인 데이터의 정량적 정보와 오프체인 데이터의 정성적 정보를 결합함으로써 주소의 활동 패턴뿐만 아니라 그 배후의 의도와 연관성을 이해할 수 있다. 이는 블록체인 생태계의 안전성과 신뢰성을 향상시키는 데 중요한 역할을 한다.

이 논문에서는 온체인과 오프체인 데이터를 통합하여 이더리움 주소를 식별하는 방법을 제안한다.

이 방법은 실시간으로 이더리움 Geth를 통해 블록 데이터를 수집하고, 수집된 주소의 온체인 특징 정보를 추출하여 머신러닝 모델에 입력함으로써 해당 주소를 'crime', 'exchange', 'normal'로 분류한다. 또한 Etherscan 메타데이터 API를 활용하여 각 주소에 대한 태그, 레이블, 설명, 관련 웹사이트 및 소셜 미디어 링크 등의 오프체인 메타데이터를 수집하여 주소의 신뢰성을 검증한다. 검증된 주소는 라벨링 테이블에 저장되며, 정보가 부족한 주소는 임시 테이블에 보관된다.

## II. 관련 연구

### 2.1 다크웹 오프체인 데이터를 이용한 다계층 비트코인 클러스터링 기법

이 연구에서는 다크웹 등에서 불법 거래를 위한 지불 수단으로 사용되는 비트코인의 악용 문제를 해결하기 위해 새로운 클러스터링 휴리스틱을 제안한다[2]. 기존 휴리스틱 기법에서는 온체인 데이터만을 사용하여 동일한 지갑의 주소들을 클러스터링하는데 한계가 있었기 때문에, 오프체인 데이터인 Silk Road를 매칭시켜 클러스터링 정확도를 높였다. 이 논문에서는 기존의 방법론을 바탕으로 Etherscan API를 통해 온·오프체인 데이터를 통합하여 신뢰성 있는 이더리움 주소 식별 방법을 제시한다.

### 2.2 가상화폐 조사를 위한 블록체인 기반의 온체인 데이터 분석 기법

이 연구는 블록체인 기반의 온체인 데이터 분석 기법을 통해 가상화폐 관련 범죄를 효과적으로 조사하고자 한다[3]. 이를 위해 시간 특성 방식, 그래프 분석, 이상 징후 탐지 등의 디지털 포렌식 기법을 적용하여, Pagerank 알고리즘을 활용해 거래소 및 불법 행위와 연관된 가상화폐 주소를 식별하였다. 이 연구를 통해 가상화폐 디지털 포렌식 분야에서 체계적인 분석 방법의 확립과 추가적인 실증 연구가 필요함을 제시하고 있다.

### III. 이더리움 온·오프체인 데이터 기반 주소 라벨링 식별 방법

#### 3.1 이더리움 주소 식별 방법

이 논문에서는 이더리움 주소의 라벨링과 식별을 위한 방법을 제안하며, 온체인과 오프체인 데이터를 결합하여 각 주소를 정확하게 파악할 수 있도록 하였다. 라벨링 식별 방법은 데이터 처리, 특징 추출, 예측, 검증의 네 단계로 구성된다. 먼저 이더리움 온체인 데이터는 이더리움 메인넷에서 실시간으로 수집된다. Geth를 사용하여 블록 데이터를 추출하고, 각 주소의 거래 내역과 스마트 컨트랙트 상호작용 등의 다양한 특징을 수집한다. 이를 통해 온체인 데이터를 정제하여 기계 학습 모델의 피처 값으로 활용된다. 오프체인 데이터는 Etherscan 메타데이터 API를 통해 수집되며, 주소와 관련된 태그, 레이블, 설명 등을 포함한다. 이를 통해 온체인 데이터로는 탐지하기 어려운 정보를 보완하고 주소의 신뢰성을 평가할 수 있다. 본 연구에서는 온체인과 오프체인 데이터를 통합하여 주소를 'crime', 'exchange', 'normal'로 분류하는 모델을 개발하였다. 모델은 주소의 활동 패턴을 분석하여 해당 주소의 라벨링을 추정하며, 결과는 온체인 예측 데이터로 저장된다. 이 과정에서 온체인과 오프체인 데이터의 조합을 통해 주소의 활동 특성 및 의도를 파악할 수 있도록 설계하였다. 예측된 주소는 Etherscan을 통해 검증 절차를 거치며, 검증된 주소는 라벨링 테이블에 저장된다. 이 테이블은 주소와 관련된 추가적인 메타데이터를 포함하여 향후 주소 식별과 분석에 활용된다. 반면 정

보가 부족한 주소는 임시 테이블에 저장되며, 추가 데이터 수집을 통해 점진적으로 검증된다. 검증 과정은 주기적으로 수행되어 최신 데이터가 반영되도록 설계하였다.

#### 3.2 이더리움 온체인 데이터 피처 설계

이더리움 메인넷에서 수집한 온체인 데이터를 기반으로 주소 라벨링을 위한 특징을 설계하였다. 논문 'Detection of illicit accounts over the Ethereum blockchain' [4]에서는 총 42개의 특징을 사용하여 이더리움 불법 계정을 탐지하였지만, 본 논문에서는 모델의 성능과 효율성을 고려하여 특징 중요도가 높은 16개의 특징을 선정하였다. 선정된 16개의 특징은 거래 빈도, 전송된 거래의 평균 가치, 활동이 있었던 일수 등 주소의 다양한 활동 패턴을 반영한다.

#### 3.3 이더리움 오프체인 데이터 수집

오프체인 데이터를 수집하기 위해 Etherscan API를 활용하였다. 그림 2는 Etherscan API를 통해 수집한 이더리움 주소의 메타데이터 예시를 보여준다. 예를 들어, 주소 0x098...2f96은 "Ronin Bridge Exploiter"로 식별되어 특정 해킹 사건과 연관된 주소임을 알 수 있다. 이러한 메타데이터는 주소의 신뢰성을 평가하는 데 유용하며, 온체인 데이터로는 파악할 수 없는 정보를 보완한다. 수집된 오프체인 메타데이터는 온체인 분석 결과와 결합하여 주소의 라벨링과 식별에 활용되었다.

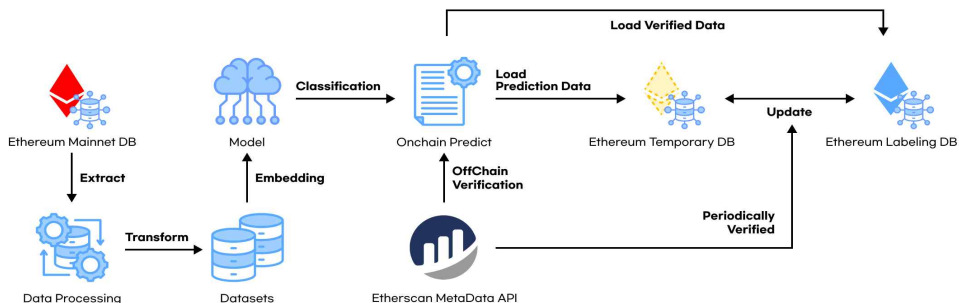


그림 1. 이더리움 온 오프체인 주소 식별 방법  
Fig. 1. Identification Method for Ethereum On-Chain and Off-Chain Addresses

### Ronin Bridge Exploiter

Address	Dx098b716b8aaf21512996dc57eb0615e2383e2196
Name Tag	Ronin Bridge Exploiter
Internal Name Tag	
URL	N/A
Short Description	N/A
Reputation Score	2
Last Updated	2022-03-30 06:35:47

This address is reported to be involved in a hack targeting the Ronin bridge.

#### Labels

Exploit CFAC Sanctions Lists

그림 2. 이더스캔 API를 통한 이더리움 주소 식별  
Fig. 2. Identifying Ethereum addresses with the Etherscan API

## IV. 실험 및 평가

### 4.1 실험 환경

표 1. 실험 데이터

Table 1. Experimental data

Label	Count
Crime	11,027
Normal	11,112
Exchange	11,180
Total	33,319

표 2. 소프트웨어 환경

Table 2. Software environment

Software	Specification
Model	catboost 1.2.7
API	Etherscan.io
DataBase	PostgreSQL 15.8
MLOps	MLflow 2.17.2

이 논문에서 CatBoost 모델 학습에 사용한 데이터는 범죄자 주소, 거래소 주소, 일반 주소를 Kaggle [5] 및 ChainWatcher [6]에서 수집하여 총 33,319개의 데이터셋을 사용하였다. 실험을 위해 사용한 머신러닝 모델은 CatBoost를 사용하였고, 오픈체인 데이터를 가져오기 위해 Etherscan.io API를 활용하였다. 또한 데이터를 저장하기 위해 PostgreSQL DB를 사용하였고, 모델 실험 관리 및 배포를 위해 MLflow를 활용하였다.

### 4.1 실험 결과

이 논문에서 제안한 이더리움 온-오프체인 데이터 기반 주소 라벨링 및 식별 방법의 효과를 검증하기 위해 그림 3과 같이 MLflow를 활용하여 CatBoost 모델을 학습하고 배포하였다. 이 모델은 이더리움 메인넷에서 실시간으로 수집된 온체인 주소 데이터를 입력받아 해당 주소의 유형을 예측한다. 예측된 주소는 임시 라벨링 테이블에 저장되며, 이는 오프체인 정보가 확보되기 전까지 예측된 분류 결과를 바탕으로 이후 분석과 검증 과정에서 활용된다. 임시 라벨링 테이블에 저장된 주소는 Etherscan 메타데이터 API를 통해 추가적인 오프체인 정보를 수집하여 해당 주소의 실제 정보를 바탕으로 식별한다.

OnChain Data Predict

#### Details

Created at	2024-11-03 18:20:04
Created by	User
Experiment ID	0
Status	Finished
Run ID	afb4dc3fc0e4e1787e1917fcd94703
Duration	1.3min
Datasets used	-
Tags	Add
Source	node.py
Logged models	catboost
Registered models	-

그림 3. ML Flow를 활용한 catboost 모델 학습  
Fig. 3. Training catboost models with ML Flow

address	predict_label	create_	updat_
0x2dff572eba84612d2808b...	normal	2024-11-02 19:...	2024-11-02
0x5a42ff5d40158edf62780...	normal	2024-11-02 19:...	2024-11-02
0xb221dd0255dd666f0b8c6...	normal	2024-11-02 19:...	2024-11-02
0x90d612fb10218cbb694ac...	normal	2024-11-02 19:...	2024-11-02
0xb9198fe9ea91771e15a21...	normal	2024-11-02 19:...	2024-11-02
0xe2b0cf515bfb4f0d7c2e8...	normal	2024-11-02 19:...	2024-11-02
0x49ea35be11e8e3b057aaa...	normal	2024-11-02 19:...	2024-11-02
0x00454b404aadf8ed84162...	normal	2024-11-02 19:...	2024-11-02
0xd4988fe020fc431724413...	normal	2024-11-02 19:...	2024-11-02
0xe822271fc5db5269b44dc...	normal	2024-11-02 19:...	2024-11-02

그림 4. 이더리움 임시 라벨링 테이블  
Fig. 4. EThereum Temporary labeling table

이를 통해 각 주소가 거래소, 악의적인 활동, 일



반 사용자 등 어떤 카테고리에 속하는지 정확하게 파악할 수 있다. 식별된 주소는 그림 5와 같이 이더리움 라벨링 테이블에 저장되며, 온체인 예측 결과와 오프체인 검증 정보를 통합하여 보다 신뢰성 있는 데이터베이스를 구축한다. 이러한 과정을 통해 주소 식별의 정확성을 향상시키고, 가상자산 생태계의 신뢰성을 높일 수 있다.

id	address	Label	offchain	created	updated
38077	0x67ed4c80f8e18715c...	crime	Fake_Phishing4141	2024-11-02 19:1.	2024-11-03 18:3.
38078	0xe70f867e7f682a89f...	crime	Fake_Phishing4144	2024-11-02 19:1.	2024-11-03 18:3.
38079	0xb5cd83d28f8182842...	crime	Fake_Phishing4145	2024-11-02 19:1.	2024-11-03 18:3.
38080	0x3c1fa09c487c765a6...	crime	Fake_Phishing4146	2024-11-02 19:1.	2024-11-03 18:3.
38081	0x43cf35acc5ec49ba8...	crime	Fake_Phishing415	2024-11-02 19:1.	2024-11-03 18:3.
38082	0x3ad072778446db054...	crime	Fake_Phishing4150	2024-11-02 19:1.	2024-11-03 18:3.
38083	0xad93fef882517ac61...	crime	Fake_Phishing4151	2024-11-02 19:1.	2024-11-03 18:3.
38084	0x3063494a00b27f994...	crime	Fake_Phishing4153	2024-11-02 19:1.	2024-11-03 18:3.
38085	0x19dd963b241374e9...	crime	Fake_Phishing4154	2024-11-02 19:1.	2024-11-03 18:3.
38086	0x516980e321482b51...	crime	Fake_Phishing4159	2024-11-02 19:1.	2024-11-03 18:3.

그림 5. 이더리움 라벨링 테이블  
Fig. 5. Ethereum labeling table

## V. 결론

이 논문에서는 이더리움 온체인과 오프체인 데이터를 통합하여 이더리움 주소를 효과적으로 라벨링하고 식별하는 방법을 제안하였다. 온체인 예측 결과와 오프체인 검증 정보를 통합하여 보다 신뢰성 있는 데이터베이스를 구축함으로써 주소 식별의 정확성을 향상시켰다. 이를 통해 가상자산 해킹과 같은 악의적인 활동을 사전에 탐지하고 방지할 수 있어, 가상자산 생태계의 신뢰성과 안전성을 높일 수 있을 것으로 기대된다. 향후 연구에서는 Graph Neural Networks(GNN), Graph Convolutional Networks(GCN) 모델을 활용하여 분석의 정밀도를 높이고, Airflow를 이용한 배치 기반 데이터 수집 및 정제, Kubeflow를 통한 머신러닝 학습 자동화, Superset을 이용한 시각화를 포함한 MLOps 프레임워크를 도입하여 시스템의 자동화를 위한 연구로 확장할 것이다.

## 참고 문헌

[1] Gertenbach, W. P., Botha, J. G., and Leenen, L., "A proposed high-level methodology on how OSIN

T is applied in blockchain investigations", Proceedings of the 19th International Conference on Cyber Warfare and Security, University of Johannesburg, Johannesburg, 26-27 March 2024.

- [2] Jin-hee Lee, Min-jae Kim, and Junbeom Hur, "Multi-Layer Bitcoin Clustering through Off-Chain Data of Darkweb", Journal of the Korea Institute of Information Security & Cryptology, Vol. 31, No. 4, p p. 715-729, 2021.
- [3] Seungwon Jeong and Byung Wan Suh, "A Study of Blockchain-based On-Chain Data Analytics Techniques for Cryptocurrency Investigations", Journal of Digital Forensics , Vol. 17, No. 3, pp. 93-105, 20 23.
- [4] Farrugia, S., Ellul, J., and Azzopardi, G., "Detection of illicit accounts over the Ethereum blockchain", Expert Systems with Applications, Vol. 150, Article No. 113318, 2020.
- [5] Kaggle "Ethereum Fraud Detection Dataset", <https://www.kaggle.com/datasets/vagifa/ethereum-fraud-detection-dataset> [accessed: Oct. 19, 2024]
- [6] ChainWatcher "Sonne Finance Hack", <https://chainwatcher.xyz/sonne> [accessed: Oct. 19, 2024]

# Loss 함수에 따른 자율주행 인공지능 모델의 특수 상황 객체 탐지 성능 분석

김정현\*, 김진호\*\*, 김한솔\*+

## Analysis of Object Detection Performance in Special Scenarios of Autonomous Driving AI Models Depending on the Loss Function

Jeonghyeon Kim\*, Jinho Kim\*\*, and Han Sol Kim\*+

### 요 약

자율주행을 위해 실제 세계에서 수집한 데이터는 객체들 사이의 이미지 개수가 불균형한 분포를 띤다. 이를 기반으로 인공지능 모델을 학습시키면 특수 상황에 대한 객체 탐지 성능이 저하되고 모델의 정확도가 일반적인 상황으로 편향되는 문제가 발생한다. 본 논문에서는 이를 해결하기 위해 손실 함수 재조정 기법을 적용한 모델의 객체 탐지 성능을 분석하였다. 특수 상황 이미지 데이터를 추가한 데이터셋을 활용하여 transformer 기반 객체 인식 모델인 RT-DETR 모델을 학습시킨 결과, 특수 상황 클래스에 대한 mAP가 20.4% 향상됨을 확인하였다.

### Abstract

The real-world data collected for autonomous driving exhibits an imbalanced distribution in the number of images across different objects. Training an AI model based on such data can lead to degraded object detection performance in rare situations and bias the model's accuracy towards more common scenarios. In this paper, we analyze the object detection performance of a model enhanced with a loss function reweighting technique to address this issue. Using a dataset augmented with images of specific situations, we trained the RT-DETR model and observed a 20.4% improvement in mAP for the special situation classes.

### Key words

autonomous driving, long tail distribution, loss re-weighting, object detection, deep learning

## 1. 서 론

실제 세계에서 수집된 데이터는 대개 불균형한 분포를 보인다. 자율주행 시스템을 위한 학습 데이

+ 교신저자

\* 단국대학교 전자전기공학과, {jeongh, hansol}@dankook.ac.kr,

\*\* 단국대학교 소프트웨어학과, sunder990@dankook.ac.kr

※이 연구는 2023년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (과제번호 : RS-2023-00251621).

터셋에는 일반적으로 자동차나 보행자와 같은 주요 클래스에 대해 방대한 양의 데이터가 존재하지만, 포트홀(pothole)이나 변색된 표지판에 관한 데이터와 같이 특수 상황에 해당하는 클래스는 극히 드물게 수집된다. 이로 인해 클래스 간 데이터의 양적 불균형이 심화되며, 이러한 불균형은 모델 학습 시 특정 클래스에 대한 예측 성능을 저하시킨다.

일반적으로 전체 데이터셋의 분포에서 많은 양을 차지하는 클래스를 헤드(head) 클래스, 적은 양의 데이터를 가진 클래스를 테일(tail) 클래스라고 한다. 분류기(classifier)의 가중치 업데이트 시 헤드 클래스에 대한 손실 함수 미분 값의 반영이 파라미터 업데이트의 대부분을 차지하게 된다. 이로 인해 테일 클래스에 대한 손실 함수 미분 값이 거의 무시되어 분류기가 테일 클래스에 속하는 이미지를 잘 분류할 수 없게 되는 현상이 발생한다.

이 같은 현상을 완화하기 위해 손실 함수 재조정 기법 [1]이 제안되었는데, 손실 함수 재조정 기법이란 데이터의 손실 함수 미분 값 계산 시 헤드 클래스 데이터의 미분 값을 감소시키고, 테일 클래스의 미분 값을 증가시키는 학습 기법을 말한다. 이와 관련하여 최근 다양한 연구가 수행되었는데도 불구하고, 자율주행 상황에서의 인식 정확도 개선 정도에 관한 분석은 현재까지 미진한 실정이다.

표 1. 실험 데이터 세트의 구성

Class	Images	Instances
Car	4,665	20,138
Van	1,500	2,058
Pedestrian	1,238	3,143
Cyclist	808	1,152
Truck	710	230
Misc	541	677
Tram	252	358
Person_sitting	67	150
Prothole	50	155

앞선 분석에 착안하여 본 논문에서는 KITTI (Karlsruhe Institute of Technology and Toyota Technological Institute) [2] 데이터셋에 포트홀(pothole) 이미지를 50장 추가하여 제작한 자체 데이터

셋을 활용하여, 실시간 객체 탐지 모델인 RT-DETR (Real-Time DEtection TRansformer) [3]을 훈련한다. 훈련 시 분류기 학습 문제에 일반적으로 활용되는 BCE (Binary Cross Entropy) 손실 함수와 손실 함수 재조정 기법이 적용된 focal 손실 함수 [4]를 이용한다. 실험을 통해 손실 함수 재조정 기법이 적용된 focal 손실 함수가 편향 문제를 효과적으로 완화할 수 있음을 확인한다.

## II. Loss 함수 분석

### 1. BCE 손실 함수

BCE 손실 함수는 클래스가 두 개만 존재하는 이진 분류 문제에 특화된 손실 함수이다. BCE 손실 함수는 정답 값과 모델의 예측 값 사이의 불일치를 손실로 측정한다. BCE 손실 함수는 다음과 같이 정의된다.

$$BCE(p) = -(y \log(p) + (1 - y) \log(1 - p)),$$

여기서  $y$ 는 정답 값이고,  $p$ 는 모델이 이미지를 타겟 클래스로 분류할 확률이다. 모델이 잘못된 예측을 하게 되면 손실 값이 커지며, 학습은 이 손실 값을 감소시키도록 진행된다. 이를 통해서 모델의 예측 값이 정답 값과 최대한 유사하게 되도록 모델의 파라미터를 조정하게 된다.

### 2. Focal 손실 함수

본 논문에서는 테일 클래스에 속하는 데이터의 예측 성능을 높이기 위해서 Focal 손실 함수를 사용했다. Focal 손실 함수는 다음과 같이 정의된다.

$$FL(p_t) = -\alpha_t (1 - p_t)^\gamma \log(p_t),$$

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise,} \end{cases}$$

여기에서  $p_t$ 는 타겟 클래스에 대한 모델의 예측 값,  $\gamma$ 는 손실 값 감소 계수이다. 또한  $\alpha_t$ 는 타겟 클래스와 그 외 클래스의 비율을 조정하는 하이퍼 파라미터(hyper parameter)이다.

Focal 손실 함수는 예측하기 쉬운 데이터에 대해서는 낮은 손실을 갖도록 하고, 예측하기 어려운 데이터에 대해서는 높은 손실을 갖도록 하여서 손실 값의 가중치를 조절한다. 이를 통해서 모델의 파라미

터 업데이트 시 발생하는 헤드 클래스와 테일 클래스의 미분 값 반영 차이를 완화하고 자율주행 인공지능 모델이 특수 상황을 더욱 정확하게 인식하도록 훈련시킬 수 있다.

### III. 실험 결과

#### 1. 학습을 위한 데이터셋 구성

본 논문에서 자율주행 인공지능 모델이 객체를 탐지할 때 특수한 상황을 가정하기 위해 KITTI 데이터셋에 포트홀 이미지 50장을 추가하였다.

표 1은 데이터셋의 이미지 수와 개체(instance) 수를 나타내고 있다. KITTI 데이터셋은 자율주행을 위해 만들어진 데이터셋으로 7,481장의 훈련 이미지(training image)와 자동차, 보행자, 트럭 등 8개의 클래스를 가지고 있다. 각 클래스마다 개체의 수는 최대 20,138개이고 최소는 150이다. 따라서 추가된 포트홀까지 총 9개의 클래스 중 Person\_sitting 클래스와 포트홀 클래스를 특수한 상황이라 가정하고 성능을 비교한다.

#### 2. 성능 분석

모델의 객체 탐지 성능 분석을 위해 RT-DETR 모델을 기반으로 손실 함수를 바꿔가며 학습을 시킨 뒤 person\_sitting 클래스와 포트홀 클래스 위주로 mAP (mean Average Precision) 값을 측정하여 성능을 비교하였다. 실험에서 Focal 손실 함수의 하이퍼파라미터인  $\gamma$ 는 2.0으로,  $\alpha$ 는 0.25로 설정했다.

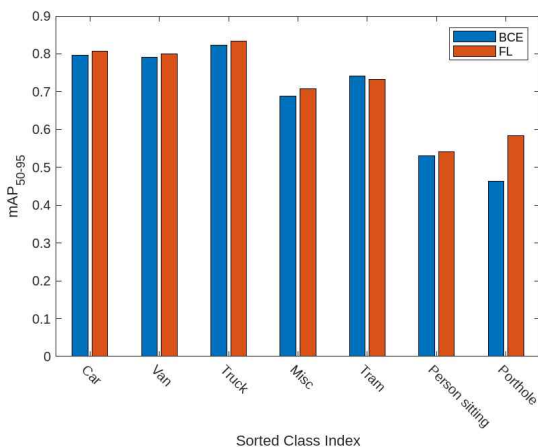


그림 1. 클래스 별 mAP<sub>50-95</sub> 측정 결과

그림 1을 보면 BCE 손실 함수로 학습했을 때, 헤드 클래스에 정확도가 편향되어 포트홀과 같은 테일 클래스에 대해서는 성능이 안 좋은 것을 볼 수 있었다. 하지만 focal 손실 함수를 이용하면 편향 정도가 완화된 것을 볼 수 있다.

표 2는 BCE 손실 함수와 focal 손실 함수의 person\_sitting과 포트홀의 mAP를 나열한 것이다. Focal 손실 함수는 mAP<sub>50-95</sub> 값이 BCE 손실 함수보다 11.9% 높은 것을 확인했다. 이 뿐만 아니라 전체 클래스에 대해서도 2.1%이 높아졌다. 이 같은 결과로 클래스 별 이미지 분포가 불균형한 데이터셋을 통해 딥러닝 모델을 학습시키는 경우 모델의 예측 확률 값을 이용해 손실 함수 값을 재조정 하는 것이 클래스 별 정확도 편향을 줄이고 전체 mAP 상승에 기여하는 것을 알 수 있었다.

표 2. Loss에 따른 mAP<sub>50-95</sub> 비교

Class	BCE	FL
all	0.661	<b>0.682</b>
Person_sitting	0.531	<b>0.541</b>
Pothole	0.464	<b>0.583</b>

### IV. 결 론

본 논문에서는 자율주행 딥러닝 모델이 불균형한 분포의 데이터셋을 학습할 때 발생하는 특수 상황 객체 탐지 성능 저하 문제를 방지하기 위해 손실 함수 재조정 기법을 분석하였다. 이를 위해 자율주행 데이터셋인 KITTI 데이터셋에 특수 상황을 위한 pothole 50장을 추가하고, 이를 기반으로 RT-DETR 객체 탐지 모델을 학습시켰다. 결과적으로 BCE로 학습을 시킨 모델의 mAP보다 FL로 학습을 시킨 모델이 약 20.4% 향상된 수치를 확인할 수 있었다. 이를 통해 불균형한 데이터로 인한 편향 학습 문제를 손실 함수 재조정으로 효과적으로 해결할 수 있음을 확인하였다. 추후 연구에서는 손실 함수뿐만 아니라 데이터 불균형을 해소하는 다양한 방법을 적용하여 특수 상황에 대한 성능을 더욱 보완하고자 한다.

참 고 문 헌

- [1] Y. Zhang, B. Kang, B. Hooi, S. Yan, and J. Feng, "Deep long-tailed learning: a survey.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 9, pp. 10795-10816, 2023.
- [2] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset.", *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231-1237, 2013.
- [3] Y. Zhao, W. Lv, S. Xu, J. Wei, G. Wang, Q. Dang, Y. Liu, and J. Chen, "DETRs beat YOLOs on real-time object detection", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16965-16974, 2024.
- [4] Tsung-Yi Lin, P. Goyal, R. Girshick, K. He, P. Dollár, "Focal loss for dense object detection", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2980-2988, 2017.

# 연령 맞춤형 키오스크 UI 개선 방안 연구

최현정\*, 김장원†

## A Study on Age-Specific UI Enhancements for Kiosks

Hyeonjeong Choi\*, and Jangwon Gim†

### 요약

최근 키오스크는 다양한 분야에서 사용이 확대되고 있으며, 특히 비대면 주문 서비스의 필수적 도구로 자리 잡고 있다. 그러나 연령대별로 사용자의 불편 요소가 존재하며, 이를 해결하기 위한 맞춤형 UI가 요구된다. 본 논문에서는 청년층과 중장년층을 위한 시간 압박 완화 기능과 노년층을 위한 직관적 내비게이션 기능 등 각 연령대의 특성을 반영한 UI 개선 방안을 제시한다. 향후 얼굴 인식을 통한 연령 맞춤형 UI 적용으로 사용자 친화적인 키오스크 환경 조성이 가능할 것으로 기대할 수 있다.

### Abstract

Kiosks are increasingly essential for Contactless ordering across sectors, yet age-related usability issues persist. This study proposes tailored UI solutions: reduced time pressure for younger users and intuitive navigation for older users. By addressing age-specific needs, the enhancements aim to improve user experience. Future integration of age-based facial recognition could further enhance accessibility and satisfaction. Recent kiosks have expanded across various fields, particularly as essential tools for Contactless ordering services.

### Key words

kiosk, user interface, user experience

### 1. 서론

2023년 농림축산식품부의 외식업체경영실태조사에 따르면 전국 3,000개 외식 업체를 대상으로 조사 통계를 진행한 결과, 서울권 92.0%, 수도권 88.8%의

\*국립군산대학교 소프트웨어학과 (2100135@kunsan.ac.kr)

†국립군산대학교 소프트웨어학과 (교신저자) (jwgim@kunsan.ac.kr)

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065).

업체에서 키오스크를 사용한다. 키오스크 UI 연구 중 노년 대상의 UI 개선 연구는 특정 대상에 치우쳐 사용자 전체를 고려한 연구가 아님을 제시한다. 본 논문은 연령별 키오스크 UI 사용 불편함을 검토하고 연령 맞춤형 UI를 제안한다.

## II. 연령별 키오스크 UI 사용 불편함

### 2.1 청년과 중장년의 키오스크 UI 사용 불편함

청년과 중장년이 키오스크 사용에 겪는 불편함은 UI 기능과 복잡한 주문, 시간적 압박이 있다. 첫째, 사용자는 일부 UI 기능으로 불편함을 겪는다. ‘뒤로 가기’를 누르면 이전 과정이 아닌 첫 과정으로 돌아간다. 주문이 초기화되어 처음부터 주문해야 한다[3]. 둘째, 복잡한 주문은 메뉴와 옵션, 결제 수단 선택의 폭이 다양해서 생긴다. 사용자는 다양한 메뉴와 여러 옵션을 선택해야 한다. 이후에 할인 쿠폰, 결제 수단, 포인트 적립 등의 복잡한 결제 과정을 거친다. 셋째, 시간 압박은 두 가지 상황으로 불편함을 겪는다. 정해진 시간에 메뉴를 빨리 선택해야 한다는 부담감과 사용자 뒤에 대기자가 있으면 시간적 압박이 나타난다. 복잡한 주문에 시간이 지체되면 시간적 압박의 결과로 불편함을 겪는다.

### 2.2 노년의 키오스크 UI 사용 불편함

노년이 키오스크 사용에 겪는 불편함은 복잡한 화면 구성 및 주문 과정, 정보 부족으로 작업 과정을 확인할 수 없는 불편함이다[4]. 첫째, 복잡한 화면 구성 및 주문 과정은 전체적인 의미를 파악하는 과정에서 겪는다. 복잡한 주문 과정은 진행 중 실수를 범할 수도 있다는 부담감으로 불편함을 겪는다. 둘째, 정보 부족으로 주문진행 과정 파악에 불편함을 겪는다.

## III. 연령 맞춤형 UI 개선

### 3.1 공통 개선 사항

연령별 사용 불편 분석으로 연령에 따른 UI 개선 방향을 모색한다. 연령 불문하고 공통으로 종합한 문제는 주문 과정이 복잡하다는 것이다. 주문 과정에서 UI를 간소화해 단순하면서도 직관적인 정보 전달을 개선 사항으로 제안한다.

표 1. 키오스크 UI 기본 요소

Table 1. Basic elements of Kiosk

1. 60도 내외의 시야각 규격을 맞춘다.
2. 기능 구현에 맞는 명칭을 사용한다.
3. 버튼 윤곽선을 그려 영역을 구분한다.
4. 버튼과 아이콘을 함께 배치한다.
5. 결제 수단 일부를 통합한다.
6. 주문 내비게이션으로 사용한다.
7. 주문 단계는 ‘포장 여부-메뉴선택-옵션 여부-옵션 선택-결제’와 같이 5단계로 간소화한다.

표 1에서 키오스크 UI 기본 요소를 정리했다. 연령별로 겪는 불편함의 공통 사항과 정보 전달에 필요한 사항을 추출해 기본 요소로 참고했다. 표 1에서 시야각은 화면 구성을 한 번에 인지할 수 있는 화면 구조 사항으로 참고했다. 이외 기본 요소는 키오스크 사용 문제에서 공통으로 경험한 주문 및 결제 과정이 복잡한 점을 해결하기 위해 제시했다.

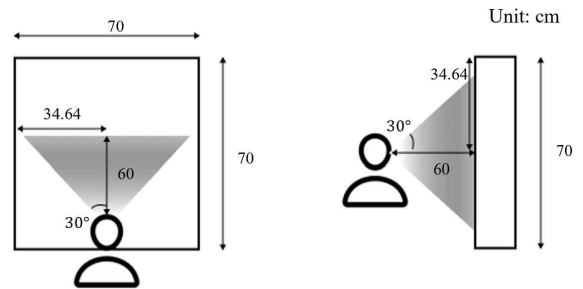


그림 1. 시야각을 고려한 UI 전체 화면 크기  
Fig. 1. Full screen size of UI considering viewing angle

그림 1은 시야각과 평균 거리를 고려해 UI 전체 화면 크기를 계산한 그림이다. 인간은 사방으로 각각 30도, 총 60도 내외의 시야각[5]을 가지고 있으므로 해당 규격에 맞는 UI를 구성했다. 시야각과 함께 키오스크와 사용자 간 평균 거리를 기준으로 이상적인 UI 크기를 설정한다. 평균 거리는 실내 키오스크 기준 거리 약 45~60cm에서 최대 거리인 약 60cm를 기준으로 했고, 가로와 세로가 각각

69.28cm로 측정되어 70cm로 정의할 수 있다.

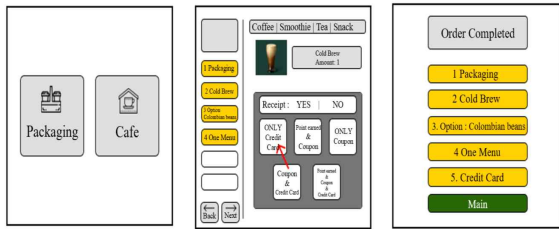


그림 2. 공통 요소로 구성한 UI 예시  
Fig. 2. UI example with common elements

그림 2는 표 1에서 정리한 UI 기본 요소를 바탕으로 구성한 UI 예시이다. 첫 화면에서 버튼과 아이콘을 함께 배치해 버튼의 직관적인 정보를 제공한다. 버튼마다 윤곽선을 그려 터치할 영역을 한정한다. 두 번째 화면에서는 기능에 맞는 명칭과 결제 수단 통합, 주문 내비게이션을 확인할 수 있다. 좌측 하단에 ‘Back’ 과 ‘Next’ 버튼의 명칭은 이전 과정과 이후 과정으로 넘어간다. 결제 수단을 다섯 가지로 분류해 ‘카드 결제’, ‘쿠폰’, ‘포인트 적립’을 일부 통합했다. ‘쿠폰 사용 후 카드 결제’, ‘포인트 적립 후 카드 결제’ 등 한번의 선택으로 결제가 이루어지는 과정이다. 좌측에 주문 과정 5단계를 시각화해 주문 과정 단계가 끝날 때마다 색으로 해당 단계가 완료된 것을 확인한다.

### 3.2 청년과 중장년을 고려한 UI 개선 사항



그림 3. 청년과 중장년을 고려한 UI 예시  
Fig. 3. UI Example Considering Youth and Middle-Aged

그림 3은 청년과 중장년이 겪는 불편함을 고려해 개선한 UI 예시이다. 해당 UI는 시간 압박에 초점을 맞춰 개선한다. 사용자마다 주문 소요 시간을 할당

하고 과정 중에 시간이 끝나가면 시간 연장 기능을 제공한다. 한 번 연장에 30초 연장으로 제한한다. UI 화면 상단의 시간을 시각적으로 보여줘 사용자가 주문 시간이 얼마나 남았는지 확인한다. 대기자는 UI에 표시된 시간을 바탕으로 시간이 얼마나 남았는지 정보를 제공받는다.

### 3.3 노년을 고려한 UI 개선 사항

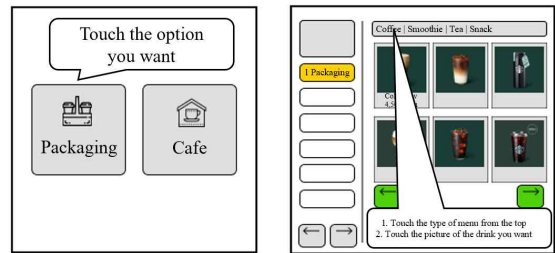


그림 4. 노년을 고려한 UI 예시  
Fig. 4. UI Example Considering Advanced age

노년 맞춤형 UI에는 복잡한 주문 과정에 소극적이라는 점에 집중해 개선 사항을 제시한다. 키오스크 사용법을 설명하는 내비게이션을 이용해 각 단계에서 터치를 유도하고 내비게이션으로 설명하며 주문 과정에 도움을 받을 수 있도록 제안한다.

## IV. 결론

연령별로 키오스크 UI 사용 불편을 분석하고 연령별로 고려해야 할 UI 개선 사항을 제안했다. 향후 연구에서는 키오스크에 개선 사항을 응용한 UI와 얼굴 인식을 적용한다. 카메라로 사용자의 얼굴을 분석해 연령을 추정하고 모드 전환으로 연령 맞춤형 UI를 제시한다. 얼굴 인식의 정확도 구사를 위해 한국인 얼굴 이미지와 테스트용 이미지를 사용해 학습한다. 얼굴과 연령이 일치하지 않는 오차 범위는 사용자의 얼굴 이미지를 학습 자료로 사용해 오차 범위를 줄여나간다.

## 참고문헌

- [1] 김현경. (2021). 국내 키오스크 접근성 실태와 나아가야 할 방향. IE 매거진, 28(1), 31-35.
- [2] A. S. OH. "Interface Design for Service



- Improvement of Unmanned Ordering Device to the Digital Underprivileged ”, Journal of the Korea Institute of Information and Communication Engineering, vol. 26, pp. 1,592-1,598, November 2022
- [3] D. W. KIM. ‘[Reporting] "Oh! I forgot to save points"...Using kiosks, MZ is also difficult’, The JoongAng, vol. 2024-06, pp. 194-195, June 2024.
- [4] M. S. Hwang, "An usability requirements elicitation of kiosk system to improvement digital literacy of seniors", Journal of The Korea Society of Information Technology Policy & Management (ITPM), vol. 15, pp. 3391-3396, December 2023.
- [5] M. S. Kwak and C. S. Kim, "A Study on the User Analysis and Design Improvement Direction of Kiosks", Journal of The Korea Convergence Society, vol. 13, pp. 171-183, April 2022.



# 어휘적 모호성 해결을 위한 문맥기반 한국어 대화 주제 분류 모델

반수경\*, 김건우\*\*

## Context-Based Korean Conversation Topic Classification Model for Vocabulary Ambiguity Resolution

Su-Gyeong Ban\* and Gun-Woo Kim\*\*

### 요 약

최근 인공지능 대화 관련 서비스가 개발되면서 대화의 맥락과 주제를 파악하는 것이 중요해지고 있다. 본 논문은 한국어 문법 구조와 어휘적 중의성을 고려할 수 있는 분류 모델을 제안한다. WSD(Contextual Word Sense Disambiguation)로 사전 학습된 KoBERT 모델과 Skip-gram 모델을 사용한 Word2Vec 알고리즘을 활용하여 단어 벡터를 학습하고, 이를 통해 특정 단어의 주변 단어들을 예측하여 단어 간의 의미적 관계를 파악한다. 학습에는 622,545개의 데이터가 사용되었으며, 불용어 제거 및 형식 교정을 위한 전처리가 진행되었다. 실험 결과 기존의 양방향 모델인 BiLSTM과 BiGRU 보다 우수한 성능을 보였다. 이는 사용자들의 대화 주제와 연결시킬 수 있는 마케팅, 대화형 AI 서비스 발전 등 다양한 분야에 기여할 수 있음을 시사한다.

### Abstract

With the recent development of artificial intelligence conversation-related services, understanding the context and topic of conversation has become more important. In this paper, we propose a classification model that can consider the Korean grammar structure and lexical importance. We learn word vectors using the KoBERT model pre-trained from WSD(Contextual Word Sense Disambiguation) and the Word2Vec algorithm using the skip-gram model, and through this, we predict surrounding words to understand the semantic relationships between words. 622,545 data were used for training, and preprocessing was performed for unnecessary term removal and format modification. Experimental results show superior performance over the existing two-way models BiLSTM and BiGRU. This suggests that it can contribute to various fields such as marketing and development of interactive AI services that can be linked to users' conversational topics.

### Key words

Deep-learning, KoBERT, Word2Vec, SkipGram, NLP, Classification, Bidirectional processing

---

\* 경상국립대학교 컴퓨터과학부, sugyeong0425@gnu.ac.kr

\*\* 경상국립대학교 교신저자 컴퓨터과학부, gunwoo.kim@gnu.ac.kr

※ 본 논문은 2023년도 정부(교육부)의 제원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2021R1G1A1006381)

## I. 서 론

최근 인공지능 대화 시스템, 음성 제어 시스템 등 자연어 처리관련 서비스가 지속적으로 개발되고 있다. 그 중 인공지능 대화의 맥락과 주제를 파악하는 것이 중요하다. 자연어를 대상으로 하는 의도 분류는 많이 진행되고 있지만, 일상 대화를 대상으로 주제를 분류하는 연구는 많지 않다.[1] 과거에는 SVM과 LSTM과 같은 전통 기계 학습 모델이나 딥러닝 모델을 사용하여 자연어를 처리한 반면, 최근에는 양방향 문맥 이해가 가능한 전이 학습 모델을 많이 사용한다. 또한, 한국어에는 2개 이상의 의미로 해석할 수 있는 동음이의어와 다의어가 존재하며 이는 감정을 나타내는 어휘에서도 다수 존재한다.[2] 이러한 다의어에 대한 문제를 해결하기 위해 다중 프로토타입 단어 임베딩을 학습할 수 있는 SkipGram 모델을 일반적인 아키텍처로 사용한다.[3]

본 연구에서는 어휘적 모호성을 해소한 분류 모델을 제안하기 위해 양방향 문맥 이해가 가능하고 문장 내의 단어 간 관계를 학습할 수 있는 SkipGram 모델을 사용한 Word2Vec 알고리즘을 사용하며 한국어 문법 구조, 단어 의미 등을 더 잘 처리할 수 있도록 조정 및 사전 학습된 KoBERT 모델을 사용하여 안부, 날씨, 쇼핑, 취미, TV, 정치경제 총 6가지의 일상 대화를 주제로 분류를 진행하였다.

## II. 본 론

### 2.1 모델 설명

본 연구에서는 한국어 텍스트의 의미적 표현을 학습하기 위해 KoBERT 모델과 Word2Vec 모델을 활용하였다. KoBERT 모델은 Transformer 기반의 모델인 BERT를 한국어 데이터에 맞게 학습시킨 모델로 문장 내의 각 단어 간의 관계를 양방향으로 이해할 수 있는 MLM(Masked Language Model) 방식과 NSP(Next Sentence Prediction) 방식을 사용하여 문맥을 이해한다는 강점을 지닌다. Word2Vec 모델은 단어를 벡터로 표현하여 단어 간의 의미적 관계를 이해할 수 있도록 만든 단어 처리 기법으로 단어 간

의 의미적 유사성을 수치화하여 자연어 처리 작업에 효과적으로 사용할 수 있게 해준다. 이러한 Word2Vec는 SkipGram 모델을 통해 주변 단어를 예측하는 방식으로 학습된다. SkipGram 모델은 주어진 중심 단어를 기반으로 그 주변의 단어들을 예측하는 구조로 이루어져 있어 단어 간의 관계를 벡터 공간에서 잘 포착할 수 있다. 의미적으로 유사한 단어들끼리 가까운 벡터를 가짐으로써 단어 간의 유사성을 수치적으로 표현할 수 있다.

### 2.2 연구 과정

연구를 위해 주어진 한국어 대화 데이터는 불용어 제거 및 형식 교정 등의 간단한 전처리 과정을 거쳤다. 우선 데이터의 노이즈를 줄이기 위해 사분위수를 이용해 이상치를 제거하였다. 그 다음 정규 표현식을 사용하여 영문자, 한글, 숫자, 일부 문자부호만 남기고 ‘/’, ‘\n’과 같이 문장을 구분하는 특정한 형식을 공백으로 대체하였다. 불필요한 문자 삭제로 생긴 비정상적인 공백 형식은 PyKoSpacing을 사용하여 교정하였으며 한국어 대화체 전용 맞춤법 교정 모델인 ET5 Text2Text를 사용하여 맞춤법을 교정하였다.

SkipGram 모델을 사용하여 단어 임베딩을 학습하기 위해 먼저 텍스트 임베딩을 진행하였다. KoBERT의 CLS 토큰을 텍스트 임베딩으로 사용하여 텍스트의 의미를 벡터로 변환하였다. 이 과정에서 텍스트는 Tokenizer를 통해 토큰화한 후, BERT 모델에 입력되어 마지막 히든 상태의 첫 번째 토큰 CLS의 임베딩이 추출되며 각 추출된 임베딩은 텍스트의 의미를 대표하는 벡터로 활용되었다. 이 때 모델을 평가 모드로 전환하고 그래디언트 계산을 비활성화하여 속도를 빠르게 하였다. 그 후 의미적으로 유사한 텍스트들을 시각적으로 확인하기 위해 KMeans 클러스터링을 적용하여 비슷한 의미를 가진 텍스트들을 그룹화하였다. 그림 1은 BERT 모델의 아키텍처이며 그림 2는 KMeans 클러스터링을 적용한 텍스트 임베딩의 결과를 시각화한 것이다.

텍스트 그룹화가 완료된 후 Word2Vec 모델을 사용하여 단어 벡터를 학습하여 단어 간의 의미적 관계를 분석하였다. Word2Vec는 SkipGram 모델을 통해 주변

단어를 예측하는 방식으로 학습되었으며, 이 때 단어 벡터의 차원 수는 100 윈도우는 5로 설정하여 주변 단어를 5개까지 고려하도록 하였다. 또한, min count를 1로 설정하여 최소 빈도 수가 1인 단어만 포함하도록 하였다. 생성된 모델은 각 단어를 고차원 벡터로 변환하며, 이 벡터는 문맥 정보를 반영한 단어의 의미적 특징을 담고 있다. 학습된 단어 벡터는 비슷한 의미를 가진 단어들이 벡터 공간에서 가까운 거리에 위치하도록 조정된다.

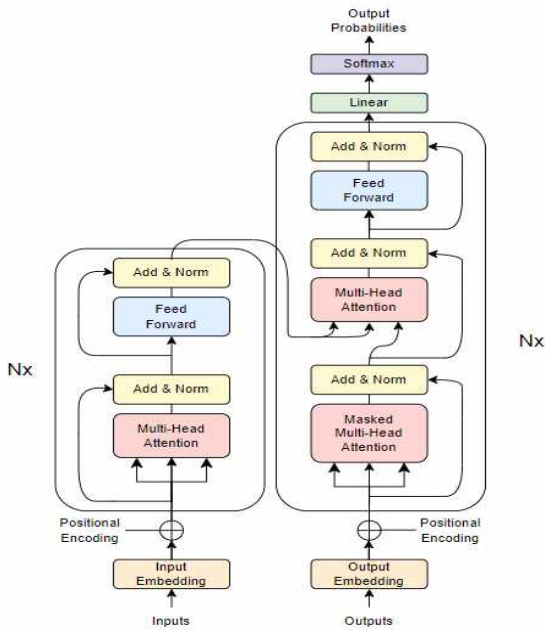


그림 1. BERT 모델 아키텍처  
Fig. 1. BERT Model Architecture

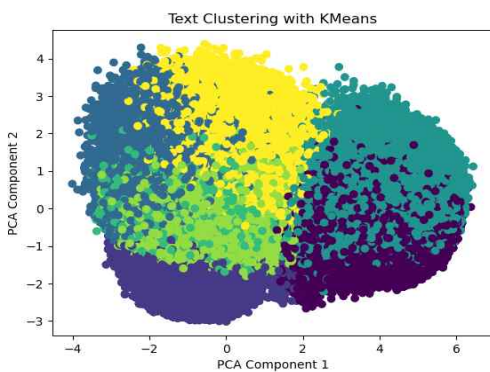


그림 2. KMeans 클러스터링 결과  
Fig. 2. KMeas Clustering Result

### 2.3 연구 결과

학습된 KoBERT 모델은 Accuracy와 precision, recall, f1-score를 통해 성능을 확인하였고 실루엣 점수와 엘

보우 방법을 통해 클러스터링의 성능을 확인하였다. 표 1은 3가지 모델의 성능을 비교한 표이다.

표 1. 성능 결과 비교

Table. 1. Performance Results Comparison

	Accuracy	Precision	Recall	F1-score
KoBERT	0.847	0.852	0.845	0.838
BiLSTM	0.793	0.785	0.790	0.787
BiGRU	0.785	0.770	0.780	0.765

확인 결과 KoBERT는 다른 양방향 모델인 BiLSTM과 BiGRU 보다 더 나은 성능을 보였으며 클러스터링은 실루엣 점수로 대략 0.5 이상의 성능을 보였다.

### III. 결 론

본 연구에서는 한국어 대화의 6가지 주제에 대한 분류를 진행하였다. Word2Vec과 KoBERT 모델을 결합하여 연구를 진행하였으며 텍스트 데이터의 임베딩을 추출하고 이를 기반으로 텍스트 클러스터링과 단어 간의 의미적 유사성을 분석하였다. 또한, Word2Vec의 SkipGram 모델을 통해 학습된 단어 벡터는 단어 간의 의미적 유사성을 표현한다. 이는 단어 간의 의미적 유사성을 수치화 작업함으로써 기존의 분류 방법보다 개선된 결과를 보여준다.

향후 연구에서는 클러스터링과 단어 임베딩의 성능을 더욱 정교하게 개선하고 이를 기반으로 다양한 텍스트 분류 작업을 진행할 예정이다.

### 참 고 문 헌

- [1] 박규현, and 정영섭. "KoBERT 를 사용한 한국어 일상 주제 분류." 한국정보과학회 학술발표논문집. 1735-1737. 2021.
- [2] Ho-yeon Ki and Kyung-shik Shin. "Emotion Analysis Using a Bidirectional LSTM for Word Sense Disambiguation." The Korean Journal of BigData, vol. 5, no. 1, pp. 197-208, 2020.
- [3] LIU, Pengfei; QIU, Xipeng; HUANG, Xuanjing. Learning Context-Sensitive Word Embeddings with Neural Tensor Skip-Gram Model. In: IJCAI. 2015. p. 1284-1290.

# 뇌질환 진단과 병적 특성 차팅 및 시각화 기술 개발

권규빈\*, 정다솜\*, 안소현\*, 오예지\*, 김민지\*, 이정수\*\*

## Development of Autonomous Diagnosis, Charting, and Visualization Techniques for Brain Diseases

Gyubin Kwon\*, Dasom Jeong\*, Sohyeon Ahn\*, Yeaji Oh\*, Minji Kim\*, and Jungsoo Lee\*\*

### 요 약

뇌질환 진단과 치료에 관한 임상 소견을 보조하기 위해 인공지능과 3D 프린팅 기술을 활용하여 뇌영상 기반 뇌질환 진단과 병적 특성 분석 및 시각화가 가능한 시스템을 개발하였다. 본 시스템은 더불어 진단 결과와 병적 특성을 활용해 자율적으로 차팅이 가능하고 생성형 언어 모델을 통한 질의응답 기능으로 차팅에 기반한 추가적인 질환 정보를 사용자 맞춤으로 제공한다.

### Abstract

The proposed system was developed to support clinical insights into the diagnosis and treatment of brain diseases using artificial intelligence and 3D printing technology, enabling brain imaging-based diagnosis, pathological feature analysis, and visualization of brain disorders. This system allows for autonomous charting based on diagnostic results and pathological characteristics, and it provides personalized additional disease information through a query-response function using a language model.

### Key words

brain diseases, diagnosis, charting, visualization, artificial intelligence, 3D printing, language model

## I. 서 론

초고령화 사회에 진입하면서 뇌질환 유병률이 증가하고 있으며, 특히 뇌종양, 뇌졸중, 알츠하이머병 등 주요 뇌질환은 빠르고 정확한 진단과 치료가 요구된다. 기존에 뇌영상을 활용하여 진단과 병적 특성을 분석하는 과정은 시간 소모가 크고, 의료진의

숙련도에 따라 결과의 일관성이 떨어지는 한계가 있다.

자동화 시스템 개발은 진단과 병적 특성 분석에 소요되는 시간을 줄이고 정확도와 일관성을 높일 수 있어 기존 의료 영상 분석 한계를 극복하고 열악한 지역의 의료 공백을 메우는 데 도움을 줄 수 있다.

\* 국립금오공과대학교 메디컬IT융합공학과, Email: gyubink01@gmail.com

\*\* 국립금오공과대학교 메디컬IT융합공학과 교수, Email: jungsoo0319@gmail.com, 교신저자

※ 지원기관: 대구경북혁신대학 금오공대 RIS 사업단

또한 기존에는 환자의 병적 진행 상태를 추적하고 치료 계획을 수립하는 과정과 치료에 따른 경과를 확인하기 위해 병적 특성을 뇌영상으로 확인하고 있다. 병적 특성을 3D 프린팅으로 시각화함으로써 의료진은 다양한 감각으로 더 정확한 임상 소견을 가질 수 있고 임상 소견에 대해 환자와 더 쉬운 공유가 가능할 것이다.

본 연구에서는 뇌질환 증상이 의심되는 대상자의 뇌영상을 입력하면 인공지능 기술을 활용하여 뇌종양, 뇌졸중, 알츠하이머병을 진단하고 진단 결과에 따른 병적 특성 추출을 통해 자율 차팅과 3D 프린팅이 가능한 시스템을 개발했다. 또한 진단 결과, 병적 특성에 따라 의료진과 환자는 상황에 필요한 지식을 얻기 위해 문헌, 전문가의 도움이 필요한데 이를 위해 생성형 언어 모델 기반 질의응답 기능을 더하여 차팅에 기반한 추가적인 질환 정보를 사용자 맞춤으로 제공한다.

## II. 시스템 설계 및 구현

### 2.1 시스템 설계

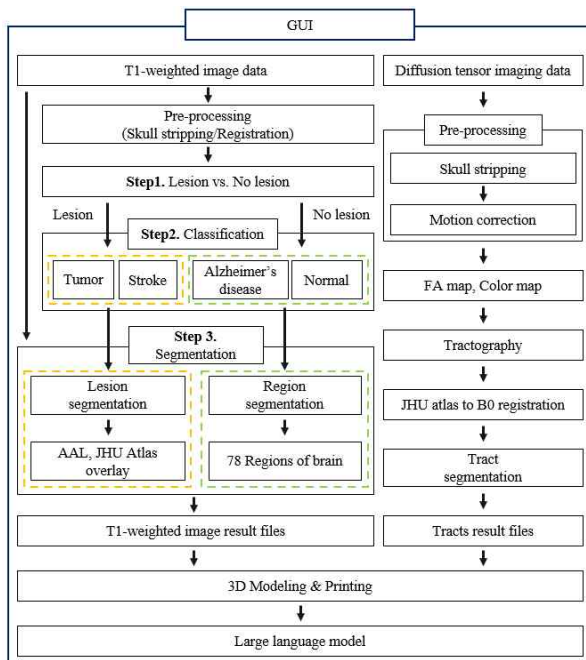


그림 1. T1과 DTI 데이터 처리를 위한 시스템  
Fig. 1. System architecture for T1 and DTI data processing

본 논문에서는 Magnetic Resonance Imaging (MRI) 영상 중 T1-weighted image (T1) 영상과 Diffusion Tensor Imaging (DTI) 영상 데이터 처리를 위한 시스템 아키텍처를 제안한다. 그림 1은 T1 영상을 활용한 자동 분석과 DTI 영상을 통한 신경로 추적 (tractography) 기능을 포함하는 그래픽 사용자 인터페이스 (Graphical User Interface, GUI)의 내부 구조를 보여준다. 제안된 시스템은 뇌질환 증상이 의심되는 대상자의 T1 영상을 통해 뇌질환 환자와 정상인을 분류하고, 뇌질환 환자 중 뇌종양, 뇌졸중, 알츠하이머병 세 가지 뇌질환을 자동으로 진단하며, 병변 탐지 및 침범 영역 파악 기능을 제공한다. 또한, 뇌 신경로 추적 및 신경로 분할 작업을 수행하여 병변 부위를 보다 효과적으로 식별하고 정교한 분석을 가능하게 한다. 생성된 출력 파일은 3D 모델링 및 프린팅 데이터로 활용되며, 질환 정보는 JSON 형식으로 저장되어 생성형 언어 모델에 전송된다. 이를 통해 차팅 결과에 대한 사용자 맞춤 질의응답 기능을 지원한다.

### 2.2 시스템 구현

그림 2(a)는 해부학적 T1 영상의 처리 과정의 개요이다. 이 과정에서는 T1 영상이 입력되면, 자동으로 두개골 제거 (skull stripping) 및 MNI (Montreal Neurological Institute) standard로의 정합 과정을 포함한 전처리 단계를 거친다. 정확한 진단을 위해 계층적인 세 가지 모델을 가진다. 첫 번째 상위층 모델은 병변의 유무를 분류한다. 하위층 모델에 해당하는 두 번째 모델은 병변이 있는 경우 뇌종양과 뇌졸중을 분류하고, 세 번째 모델은 병변이 없는 경우 알츠하이머병과 정상인을 분류한다.

진단 결과를 기반으로, 병변이 있는 경우 해당 뇌의 병변을 분할하고 그 크기를 파악한다. 이때 MNI standard로 정합하면서 생성된 matrix 파일을 역변환하여 회백질 영역에 병변 침범을 식별하기 위한 AAL (Automated Anatomical Labeling) 아틀라스 [1]와 백질 영역에 병변 침범을 식별하기 위한 JHU (Johns Hopkins University) 아틀라스 [2]를 활용하여 병변 위치와 영역별 손상 수준을 파악한다.

병변이 없는 경우 78개의 뇌 영역을 분할하고, 알츠하이머병의 병적 특성을 잘 드러내는 6개의 주요 영역 (hippocampus, amygdala, caudate, entorhinal cortex, fusiform gyrus, and parahippocampal gyrus) 은 필수적으로 3D 모델링을 진행하기 위해 별도의 파일로 저장한다[3].

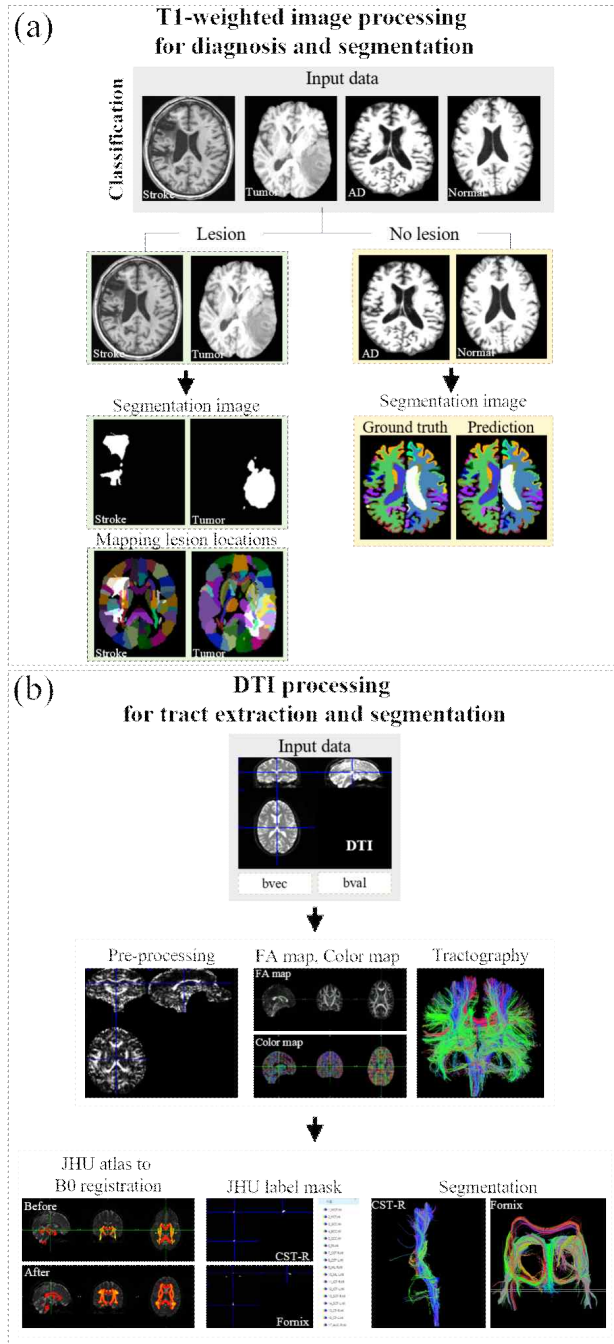


그림 2. T1 및 신경 영상 데이터 처리 과정  
 Fig. 2. Processing workflow : (a) T1-weighted image processing for diagnosis and segmentation, (b) DTI processing for tract extraction and segmentation

그림 2(b)는 DTI에 적용된 신경로 추출 과정 개요이다. 이 과정에서 bvec 및 bval 파일을 포함한 DTI 영상이 입력되면, 두개골 제거와 움직임 보정 (motion correction)과 같은 전처리 단계를 거쳐 분석에 적합한 상태로 준비한다. 이를 기반으로 분할 비등방도 (Fractional Anisotropy, FA) map, 그리고 FA color map이 생성되어 뇌 구조를 시각화한다. 이후, 전체 뇌 영역에서 신경로 추적이 수행되어 뇌 백질 내 신경로의 전반적인 모습을 추출한다. 이어서, B0 native 공간에 정합된 JHU 아틀라스의 레이블을 이용하여 환측과 건측을 포함한 45개의 주요 신경로가 분할된다. 분할된 45개의 데이터의 추출 신경로의 평균 길이, 곡률, 부피, FA, MD (Mean Diffusivity), AD (Axial Diffusivity), 그리고 RD (Radial Diffusivity) 정보 값을 저장한다. 이 중 신경로 손상 수준 정보는 환측과 건측의 FA 값을 활용하고, 이 정보는 차팅을 위해 생성형 언어 모델에 전송된다.

이러한 과정은 영역 및 병변의 크기와 위치, 영역 및 신경로 손상도를 정확하게 평가할 수 있게 하여 뇌질환 환자의 진단 정확도를 높이고, 향후 치료 전략을 수립하는 데 유용한 정보를 제공한다.

3D 모델링 작업은 분할된 영상 정보를 기반으로 다음과 같은 절차를 통해 수행하였다.

영역 및 병변 분할의 경우, MRI 영상의 복셀 정보를 바탕으로 해상도 조정, 메시화, 스무딩 등을 사용했고, 2배로 업샘플링하였다. 모델링 결과물의 표면이 거친 상태이므로 Laplacian smoothing을 사용하여 표면을 부드럽게 출력되도록 작업하였다.

신경로의 경우, 각 섬유 경로를 원기둥 형태로 모델링하고 이 원기둥을 섬유 경로의 각 시작점과 끝점을 잇는 형태로 생성하였다. 이후 DTI 영상으로부터 계산된 방향 벡터와 길이에 기반하여 기존에 설정된 z 좌표를 조정하고, 이를 통해 원기둥의 전체 좌표를 생성 및 연결하였다.

최종적으로, 모델링이 끝난 후 3D 프린팅이 가능하도록 STL 파일로 변환하였다.

뇌질환 환자의 병적 특성 분석 모델의 실행 절차를 직관적으로 보기 위한 GUI를 그림 3과 같이 설계하였다. GUI는 그림과 같이 진행된다.



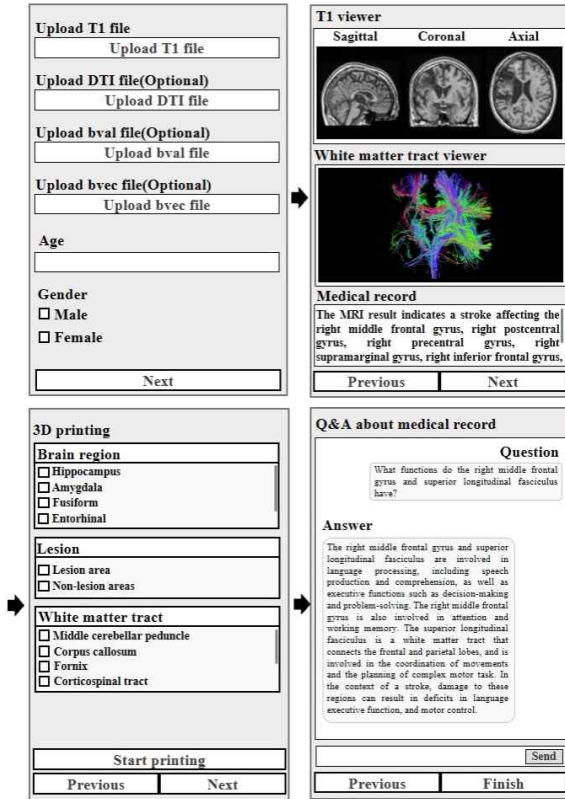


그림 3. AI 기반 뇌질환 특성 차팅 및 시각화 GUI  
Fig. 3. AI-based GUI for charting and visualizing brain disease characteristics

### III. 성능 평가 및 결과

표 1. 진단 및 분할 성능 평가  
Table 1. Performance of classification and segmentation

Process	Type	Acc	Dice
Step 1. Classification	Lesion vs. No lesion	1.00	-
	Tumor vs. Stroke	0.91	-
Step 2. Classification	Alzheimer's disease vs. Normal	0.76	-
	Lesion	Stroke	0.51
Step 3. Segmentation	Alzheimer's disease	-	0.81
	Region	-	0.55
	Normal	-	0.86

본 논문에 사용된 데이터셋은 ATLAS V2.0의 뇌졸중 환자 200명, BraTS2020의 뇌종양 환자 200명, ADNI (Alzheimer's Disease Neuroimaging Initiative)의 알츠하이머 환자 199명 및 정상인 200명으로 구성되며, 각 집단 간 중복된 피험자는 없다.

표 1은 진단 및 분할 과정의 성능 평가 결과를 제시한다. 진단 과정에서는 세 가지 간단한 모델을 각각 적용하여 정확도 (Accuracy, Acc)를 기준으로 성능을 측정하였다. 분할 과정에는 일반적인 이미지 분할에 활용되는 3D U-Net 대신, 다양한 크기의 필터를 병렬로 적용하는 Inception 모듈을 통합한 3D Inception U-Net 모델을 사용하였다[4]. 이를 통해 효과적으로 특징을 추출하고, 정밀한 병변 분할을 구현하였다. 성능은 Dice 계수를 통해 평가하였다.

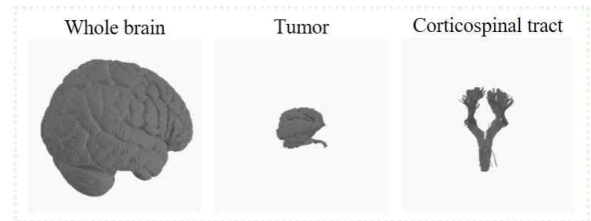


그림 4. STL 파일 형식의 3D 뇌 모델링 결과  
Fig. 4. 3D-printed models of the brain, tumor, and corticospinal tract

생성된 STL 파일을 확인한 결과는 그림 4와 같다. MRI 데이터 복셀 크기를 반영하여 실제 대상자의 뇌, 종양, 신경로의 크기와 유사한 것을 확인하였고 3D 프린팅 품질이 양호한 수준의 높은 해상도로 모델링이 되었다.

Llama 3.1-8B [5]를 사용하여 프롬프트를 구성하고 인공지능 모델을 통해 나온 진단 결과를 기반으로 차팅과 질의응답을 진행한다. 언어 모델을 통한 질의응답으로 사용자에게 해당 질환에 대한 추가적인 정보를 제공한다.

뇌졸중 환자의 뇌영상 데이터를 입력하면 진단, 병적 특성 분석 과정을 스스로 거치고 아래와 같은 차팅 결과를 얻었다.

“The subject is a 62-year-old female, and the diagnostic finding indicates signs of a stroke.

The lesion is located in the right hemisphere with a volume of 147.227 mL.

The affected gray matter regions due to the lesion are as follows: right middle frontal gyrus (59% damage), right postcentral gyrus (59% damage), right precentral gyrus (56% damage), and right supramarginal gyrus (72% damage).

Additionally, the affected white matter region is the right superior longitudinal fasciculus (67% damage). The tract integrity of the affected superior longitudinal fasciculus is 0.432, and the tract integrity of the unaffected superior longitudinal fasciculus is 0.683.”

뇌졸중 진단 결과를 기반으로 진행한 질의응답은 다음과 같다.

Q: What functions do the right middle frontal gyrus and superior longitudinal fasciculus have?

A: The right middle frontal gyrus and superior longitudinal fasciculus are involved in language processing, including speech production and comprehension, as well as executive functions such as decision-making and problem-solving. The right middle frontal gyrus is also involved in attention and working memory. The superior longitudinal fasciculus is a white matter tract that connects the frontal and parietal lobes, and is involved in the coordination of movements and the planning of complex motor task. In the context of a stroke, damage to these regions can result in deficits in language executive function, and motor control.

#### IV. 결 론

본 논문에서는 뇌영상 데이터를 입력하면 인공지능 및 3D 프린팅 기술을 기반으로 뇌질환 진단과 병적 특성 차팅 및 시각화 기술을 제안하였다. 제안된 방법은 T1, DTI 영상을 이용하여 뇌질환을 진단하고, 병변과 뇌영역 및 신경로를 분할하여 병변의 크기와 위치, 영역 및 신경로 손상도를 자동으로 추정하고 차팅한다. 그리고 진단 결과에 따라 병적 특성들은 모델링을 통해 시각화가 가능하고 3D 프린팅이 가능하다. 마지막으로 생성형 언어 모델을 이용한 질의응답 시스템을 통해 진단 결과와 관련된 유용하고 의미 있는 정보를 사용자 맞춤형으로 제공한다. 본 시스템은 의료진이 진단 보조 장치로서 신속하고 일관성 있는 진단을 수행할 수 있

도록 지원하며 치료 계획 수립에 도움을 줄 수 있다. 또한, 병적 특성의 3D 프린팅을 통해 환자에게 추가적인 이해 요소를 제공함으로써 진단 결과와 경과 추적 및 치료 효과에 대한 이해도를 높일 수 있을 것이다.

추후 연구에서는 진단 및 분할 성능을 향상시키기 위해 알고리즘 최적화를 지속적으로 진행할 계획이다. 더불어, 현재 개발 중인 GUI를 웹 기반 플랫폼으로 확장하여, 사용자들의 접근성을 높이고 결과를 실시간으로 확인할 수 있도록 할 예정이다.

#### 참 고 문 헌

- [1] N. Tzourio-Mazoyer; B. Landeau; D. Papathanassiou; F. Crivello; O. Etard; N. Delcroix; Bernard Mazoyer & M. Joliot (January 2002). "Automated Anatomical Labeling of activations in SPM using a Macroscopic Anatomical Parcellation of the MNI MRI single-subject brain". *NeuroImage*. 15 (1): 273-289.
- [2] Johns Hopkins University. (2021). Multi-atlas segmentation: Applications in neurodegenerative disease research. *Frontiers in Neuroscience*. Retrieved from
- [3] Braak, H., and Braak, E. (1997). Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol. Aging* 18, 351-357.
- [4] Pun, N. S., & Agarwal, S. (2020). Inception U-Net architecture for semantic segmentation to identify nuclei in microscopy cell images. *\*ACM Transactions on Multimedia Computing, Communications, and Applications\**, 16(1), Article No. 12, 1-15.
- [5] Llama Team, AI@Meta, "The Llama 3 Herd of Models", arXiv preprint arXiv:2407.21783, 2024.

# 메타데이터 기반 아두이노 코드 생성 기법

김재엽\*, 김동현\*, 이상민\*, 이민지\*, 이의종\*\*, 이석훈\*†

## Metadata-Based Arduino Code Generation Techniques

Jaeyeob Kim\*, Donghyun Kim\*, Sangmin Lee\*, Minji Lee\*, Euijong Lee\*\*, and Sukhoon Lee\*†

### 요약

최근 IoT 환경은 아두이노와 같은 저비용 마이크로컨트롤러를 활용하여 제어하고 모니터링에 널리 활용되고 있다. 하지만, 아두이노 기반 디바이스의 제어와 데이터 수집을 위한 소프트웨어 개발은 전문 지식이 필요로 한다. 이를 해결하고자 본 연구에서는 메타데이터 정의를 통해 LLM 기반의 IoT 디바이스별 맞춤형 코드 생성을 진행하고 일관성과 정확성을 향상시키는 기법을 제안한다. 제안 방안을 위해 코드 생성에 필요한 메타데이터를 정의하고 기존 일반 프롬프트와 성능을 비교 평가하여 메타데이터 기반 프롬프트 방식이 IoT 디바이스 제어에 있어 효율적임을 보인다.

### Abstract

Recently, IoT environment is widely used for control and monitoring using low-cost microcontrollers such as Arduino. However, the development of software for control and data collection of Arduino-based devices requires expertise. To solve this problem, this study proposes a method to generate customized code for each IoT device based on LLM through metadata definition and improve consistency and accuracy. For the proposed method, the metadata-based prompt method is shown to be efficient in IoT device control by defining metadata required for code generation and comparing and evaluating performance with existing general prompts.

### Key words

ChatGPT, Large Language Models, Metadata, IoT

## 1. 서론

최근 IoT(Internet of Things) 환경은 다양한 디바이스가 네트워크로 연결되어 상호작용하며 데이터를 수집하고 이를 바탕으로 특정 작업을 수행할 수

있는 체계를 지향하고 있다[1]. 이러한 환경에서 아두이노(Arduino)와 같은 저비용의 마이크로컨트롤러는 IoT 디바이스의 제어와 모니터링에 널리 활용되고 있으며, 그 중요성은 계속해서 증가하고 있다[2]. 하지만 아두이노 기반 디바이스의 제어와 데이터

\* 국립군산대학교 소프트웨어학과

† 교신저자

\*\* 충북대학교 소프트웨어학부

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065).

표 1. 코드 생성을 위한 메타데이터 정의 및 예시

Table 1. Metadata definitions and examples for code generation

key	data type	description	example
actuator_pin	integer	액추에이터 핀 번호를 기술한다.	9,10,11
sensor_pin	string	센서의 핀 번호를 기술한다.	A0
commands	string	액추에이터 제어 명령어를 기술한다.	ON, OFF, SET_COLOR
initial_value	string	액추에이터의 초기 설정 값을 기술한다.	RED
delay	integer	명령어 실행 후 지연 시간을 기술한다.	5000(ms)
purpose	string	디바이스의 목적을 기술한다.	색상 제어
library	string	아두이노 코드에 필요한 라이브러리를 선택한다.	

수집을 위한 소프트웨어 개발은 여전히 많은 전문 지식과 시간을 필요로 한다.

LLM(Large Language Model)은 대규모 데이터 학습을 통해 다양한 언어 작업을 수행할 수 있으며, 특히 코드 생성 및 제어 명령어 작성에 있어 높은 성능을 보여준다[3]. 그러나 기존의 일반적인 프롬프트 기반 접근 방식은 IoT 디바이스의 고유 특성을 반영하지 못해 일관성 문제와 코드의 기능 정확성에 한계가 발생하는 경우가 많다. 이러한 문제는 특히 필요한 특정 정보가 부족하거나 디바이스 고유의 정보가 반영되지 않을 때 더욱 뚜렷하다.

따라서, 다양한 IoT 디바이스의 고유한 특성을 반영한 메타데이터 정의 기반의 접근이 필요하며, 이를 통해 코드 생성 과정의 효율성 및 일관성을 크게 향상시킬 수 있을 것으로 예상된다. 본 연구에서는 메타데이터 정의 기반 코드 생성 방안을 제안한다. 이를 통해 일반적인 프롬프트 기반의 접근 방식에서 발생하는 문제를 해결하고, 보다 효율적이고 정확한 자동화 코드를 생성할 수 있는 가능성을 탐구하고자 한다.

## II. 제안 기법

### 2.1 코드 생성을 위한 메타데이터 정의

표 1은 본 연구에서 정의한 메타데이터이다. 각 IoT 디바이스의 고유 특성을 반영한 메타데이터를 저장하여 코드 생성의 기준이 될 수 있도록 한다. 사용자는 자신이 구현한 디바이스에 맞게 메타데이터를 등록하고 그에 맞는 코드를 생성할 수 있다.

### 2.2 제안 기법의 구현 및 결과

본 연구에서 제안한 메타데이터를 기반으로 코드를 생성하고자 GPT-3.5 API를 사용한다. 또한, 메타데이터의 필요성을 알아보기 위해 GPT-3.5 API의 프롬프트를 메타데이터 기반의 코드 생성에 맞게 수정하고 기존 GPT-3.5와의 코드 생성 비교 평가를 진행한다. 제안 시스템의 구현을 위해 GPT-3.5 API를 이용하여 코드를 생성하는 과정은 그림 1과 같다.

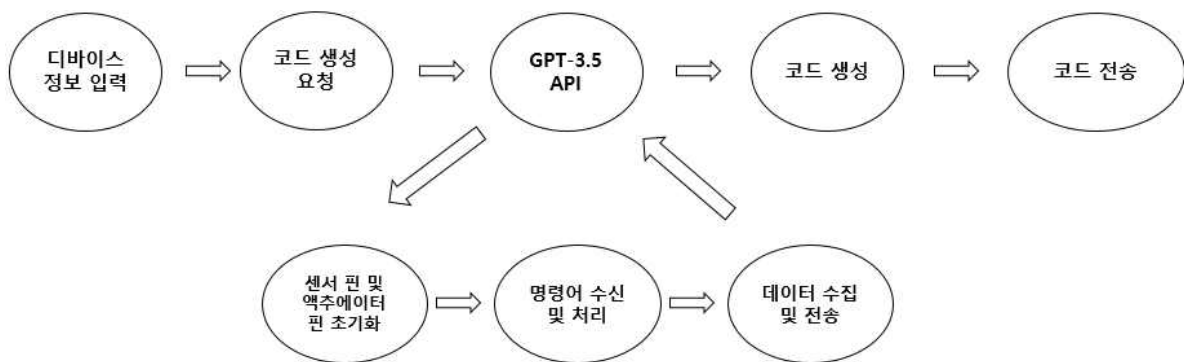


그림 1. GPT-3.5 API를 이용한 코드 생성 과정

Fig. 1. Code generation process using GPT-3.5 API

표 2. 제안 시스템과 일반 프롬프트와의 비교 평가

Table 2. Evaluation of a comparison between the proposal system and the general prompt

프롬프트 종류	기능 일치성 평균	핀 번호 정확성 평균	출력 형식 일치성 평균	코드 동작률 (%)	평균 코드 생성 시간 (초)
메타데이터 기반 프롬프트	10	10	10	90%	4.56
일반 프롬프트	9.4	6.5	5.6	60%	4.98

표 3. 일반 프롬프트와 메타데이터 프롬프트 예시

Table 3. An example of general prompts and metadata prompts

프롬프트	프롬프트 내용
일반 프롬프트	RGBLED와 MQ-7 센서가 연결되어 있으며 RGBLED를 켜고 끄는 코드를 작성하고 MQ-7 센서로부터 측정된 값을 읽어 출력하는 코드를 생성하세요.
메타데이터 기반 프롬프트	<pre>deviceData:{device_id:1, device_name:RGBLED&amp;MQ7, metadata: {actuator_pin:9 10 11, sensor_pin:A0, commands:ON SET_COLOR OFF, delay:5000ms}}</pre> 주어진 메타데이터 기반 코드를 작성하세요. setup함수에서 시리얼 통신을 설정하고, loop 함수에서는 commands 명령을 처리하는 코드를 작성합니다. 액추에이터 상태와 센서 데이터 값을 시리얼 통신을 통해 delay에 맞게 json 형식으로 전송하세요.

디바이스의 정보를 입력하면 각 디바이스에 맞는 코드를 정확하게 생성할 수 있도록 설계되었으며 메타데이터 기반의 코드를 생성할 수 있도록 한다. 표 2는 제안 시스템을 통해 생성된 코드와 일반 프롬프트를 통해 생성된 코드의 비교 평가이다. 3가지 디바이스를 번갈아가며 똑같은 프롬프트를 입력하여 30회 진행하였으며, 이 과정에서 사용자 의도에 맞게 코드를 작성하였는지를 평가한다. 평가 요소는 불일치, 부분일치, 완전일치로 순서대로 0, 5, 10의 점수로 평가하였다. 사용자가 원하는 기능을 구현하였는지에 대한 기능 일치성과 사용자의 핀 번호에 대한 일치성, 센서값과 액추에이터 상태를 출력하는 형식에 대한 일치성을 평가하고 이에 대한 코드 동작률과 코드 생성 시간을 평가한다. 메타데이터 기반 프롬프트는 일반 프롬프트에 비해 우수한 성능

을 나타낸다. 메타데이터 기반 프롬프트는 사용자가 요구하는 기능을 정확히 반영하고 핀 번호와 출력 형식을 일관성 있게 생성하는 것을 보여준다. 이를 통해 디바이스 제어 코드 생성의 효율성을 높일 수 있음을 보인다.

### III. 결 론

본 연구에서는 메타데이터 기반 아두이노 코드 생성 기법을 통해 일반 프롬프트 방식과 비교하여 성능을 평가하였다. 실험 결과, 메타데이터를 활용한 프롬프트 방식은 기능 일치성, 명령어 인식 정확도, 핀 번호 정확성, 출력 형식 일치성 등 다양한 지표에서 일반 프롬프트 방식에 비해 우수한 성능을 나타냈다.

### 참 고 문 헌

- [1] J. Lee, "Technology Trends of Smart Energy Field using AIoT", The Korean Institute of Electrical Engineers, Vol. 72, No. 12, pp. 10-16, 2023.
- [2] J. Kim, J. Yoon, S. Choi, and M. Ryu, "IoT Platform Development Trends and Development Directions", The Journal of The Korean Institute of Communication Sciences, Vol. 30, No. 8, pp. 29-39, 2013.
- [3] S. Yeo, Y. Ma, H. Jun, T. Kim, and S. Kim, "Survey on Recent Large Language Models for Code Generation", KIISE Transactions on Computing Practices, Vol. 30, No. 8, pp. 372-381, 2024.

# 자동차 법률 문서 해석에 특화된 데이터 전처리 및 프롬프트 엔지니어링 방법론 연구

안성규\*, 문벼리\*\*, 박상범\*\*\*, 김세희\*\*\*\*, 김경외\*\*\*\*\*

## A Study on Data Preprocessing and Prompt Engineering Methodologies Specialized for the Interpretation of Automobile Law Documents

Byeori Moon\*, Seonggyu Ahn\*\*, Sangbeom Park\*\*\*, Sehee Kim\*\*\*\*, and Keungoui Kim\*\*\*\*\*

### 요 약

이 연구는 자동차 개발자가 법률과 각종 규제를 검토해야 하는 사항에 대하여 비전문가인 개발자가 더 효율적으로 법률 규제 검토 업무를 수행할 수 있도록 여러 데이터 포맷이 가진 장점을 적극 활용한 새로운 전처리 방법론을 고안한다. 그렇게 구성된 새로운 계층적 전처리 방법론이 더욱 모델에게 잘 인식될 수 있도록 Parent-Child Chunking 기법을 추가한 RAG 서비스 구조를 구축하여 개발자가 더 쉽고 빠르게 법률 업무를 수행할 수 있다고 예상된다.

### Abstract

This study devises a new preprocessing methodology that utilizes the advantages of multiple data formats to make the legal and regulatory review work more efficient for non-expert developers who need to review laws and regulations for automotive developers. The new hierarchical preprocessing methodology is expected to make the legal work easier and faster for developers by building a RAG service structure that adds a parent-child chunking technique to make it more recognizable to the model.

### Key words

Automobile Law, Braking System, Prompt Engineering, RAG, Parent-Child chunking

### 1. 서 론

자동차 개발자는 자동차 설계 및 개발 과정에서

법률을 담당하는 전문가와 지속적인 소통과 협력을 통해 다양한 법규와 규제를 고려하여 자동차를 설계해야 한다[1]. 자동차 관련 법규의 경우 여러 환

\*한동대학교 ICT창업학부 학사과정, 22000245@handong.ac.kr

\*\*한동대학교 생명과학부 학사과정, 21800418@handong.ac.kr

\*\*\*한동대학교 법학부 학사과정, youfred@handong.ac.kr

\*\*\*\*한동대학교 ICT창업학부 학사과정, kimsehee7414@handong.ac.kr

\*\*\*\*\*한동대학교 AI융합교육원 교수(교신저자), awekim@handong.edu

경적 또는 사회경제적인 이유로 법률 개정이 빈번히 일어나고, 고려해야 하는 법규의 범위와 대상이 다양하다. 때문에 개발자들이 위 사항을 설계 과정 중에 지속적으로 확인하는 것이 매우 중요하다. 만약 법률 개정에 민감하게 반응하지 못해 설계를 조정하거나, 신차 출시 후 법률 위반 사례가 발견될 경우 엄청난 비용 손실로 이어질 수 있기 때문에 자동차 설계 과정 중에 수행되는 법률 검토는 매우 중요한 부분이다. 그러나 이러한 높은 중요성에도 불구하고, 비 법률 전문가인 자동차 개발자들이 법률 자문과 검토를 직접 수행하기는 어려우며, 유관 부서와의 시의적절한 소통에도 어려움이 발생하고 있다.

본 연구에서는 자동차 설계 및 개발 과정에서 발생하는 법률 검토의 비효율성 문제[2]를 해결하기 위해, 개발자가 법률 및 규제 관련 질문에 직접 답을 얻을 수 있는 생성형 AI 모델 기반의 챗봇 서비스 구조를 제안하고자 한다. 이를 위해 본 연구는 자동차의 모든 범용 법률을 다루기보다는, 부품 선정 시 국토교통부의 리콜 건수 분석을 통해 부품의 위험성을 평가하고, 자율주행 등 미래 기술의 안전 시스템과 통합이 필수적인 제동장치를 최종 대상으로 선정하였다[3]. 또한, 법률 규제는 각국이 선도국의 법률을 기준으로 삼는 특성을 반영하여, 현재 자동차 선진국들이 모두 참여하는 유일한 글로벌 협의체인 UNECE 법률 문서를 채택하였다[4]. 이를 통해 본 연구는 법률 자문에 대한 접근성을 높이고, 개발자가 보다 효율적으로 법률 요구사항을 반영할 수 있는 환경을 제공하고자 한다.

## II. 본문

### 2.1 검색 증강 생성 (RAG)

LLM이 가지고 있는 정보는 학습 데이터로 제한되기 때문에 학습이 완료된 시점 이후에 생성된 정보에 대해서는 LLM이 관련한 질의에 적절히 답변할 수 없다[5]. 더욱이 확률에 기반하여 동작하는

특성에 의하여 LLM이 사용자에게 잘못된 정보를 제공하는, 이른바 “환각” 현상이 발생할 가능성도 존재한다[5]. 검색 증강 생성, 이른바 RAG는 앞서 언급한 LLM의 한계를 극복하기 위해 추가적인 데이터의 학습 없이 LLM을 원본 문서를 저장한 데이터베이스와 연동하여 사용자의 질의에 답변하기 위한 근거를 검색하고 이를 기반으로 보다 정확한 답변을 제공하는 방법론이다[6]. RAG에서 사용자 답변에 근거가 되는 문서들은 검색 단위인 chunk로 분할된다. 각 chunk들은 수치화된 벡터의 형태로 저장되는데 이를 “임베딩”이라고 한다. 임베딩된 벡터는 서로 간의 유사도를 구할 수 있으며 이를 이용해 사용자의 질의와 가장 유사한, 즉 가장 높은 관련성이 있는 문서 근거를 찾을 수 있다. 사용자 질의와 관련한 근거 문서 내용을 모델에게 함께 제공함으로써 LLM은 더욱 정확한 답변을 사용자에게 제공할 수 있다.

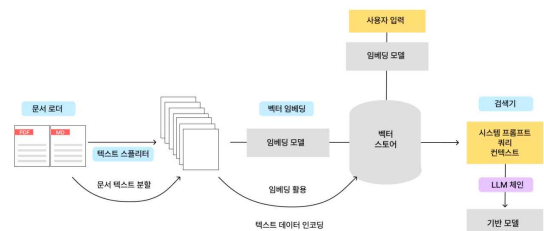


그림 1. RAG 프로세스  
Fig. 1. RAG process

### 2.2 데이터 전처리

기존 RAG 방법론에서는 원본 문서에서 추출한 텍스트를 고정된 토큰 수로 분할하여 청크를 생성한다. 이에 따라 문장 중간에서 분할이 시작되어 일부 요소가 탈락한 불완전한 문장이 chunk에 포함될 수 있으며 앞뒤 문맥에 대한 손실로 해석을 위한 충분한 정보가 chunk에 포함되지 않을 가능성이 있다. 또한 PDF 문서의 경우 텍스트 추출 과정에서 일부 문자가 부정확하게 인식되거나 누락되는 문제가 발생할 수 있다.

본 연구가 다루는 UNECE 문서를 포함한 대부분

의 법률 및 규정 문서는 의미 단위로 구분되는 조문 체계를 가지고 있다. 상위 조문의 내용은 하위 조문을 해석하는 데 주요한 전제가 되기에 문서의 상하위 체계 구조에 대한 이해는 조문의 더 정확한 해석을 위해 필수적이다. 그러나 기존 방식으로 생성된 chunk는 이러한 구조 정보를 온전히 반영할 수 없으며 앞뒤 문맥을 포함한 필요 정보가 누락되어 있어 법률 및 규정의 전달에 적합하지 않다. 본 연구에서는 법률 조문 체계를 반영하고 해석에 필요한 정보의 압축적 전달을 목표로 문서 chunking 단위이자 데이터 구성 포맷인 계층적 전처리 방법론을 구성하였다.

계층적 전처리 방법론은 앞서 설정한 목표를 달성하기 위해 상용되는 데이터 포맷을 참고하여 구성되었다. JSON (JavaScript Object Notation)은 컴퓨터 언어인 Javascript의 객체 단위를 표현하는 데이터 포맷이다. JSON의 포맷은 객체 내부에 다른 객체를 포함할 수 있으며 이를 이용하여 객체 간 상하 관계를 효과적으로 표현할 수 있다[7]. 이러한 특성을 이용하여 상위 조문 JSON 객체 안에 하위 조문 JSON 객체를 포함하는 형태로 포맷을 구성함으로써 법규 및 규정 문서의 체계 구조를 반영하였다. 특정 조문을 온전히 해석하기 위해 필요한 정보를 식별하여 하나의 JSON 조문 객체에 압축적으로 전달하였다.

```

{
  "Chapter": 5
  "Title": "Specifications"
  "Description": [
    "This chapter specifies the requirements that must be satisfied in the brake system."
  ],
  "Article 5.1.": {
    "Description": [
      "Parking brake system",
      "If a parking brake system is fitted, it shall hold the vehicle stationary",
      "The parking brake system shall: [Item]"
    ]
    "Item": [
      "(a) Have a control which is separate from the service brake system controls;"
      "(b) Be held in the locked position by solely mechanical means."
    ]
  },
  "Article 5.2.": {
    "Description": [.....],
    "Paragraph 1.2.1.": {.....},
    "Paragraph 1.2.2.": {.....}
  }
}

```

그림 2. JSON구조  
Fig. 2. JSON format

본 연구가 다루는 UNECE 문서를 포함한 많은 수의 법률 및 규정 문서는 조문을 구성하는 문자 텍스트뿐만 아니라 이미지, 수식, 표를 포함한 다양한 형태의 자료를 포함하고 있다. 이미지는 대부분 텍스트에 대한 보충적 설명을 위해 사용되기에 이미지 데이터가 생략되어도 남은 텍스트 설명으로 해당 내용에 대한 보완이 가능하다. 그러나 수식과 표는 텍스트 설명이 가지고 있지 않은 고유한 정보를 포함하고 있는 경우가 많다. 따라서 조문의 정확한 해석을 위해서는 관련된 수식과 표가 가지고 있는 정보에 대한 이해가 필수적이나 일반적인 문자 텍스트로는 수식과 표가 가지고 있는 의미를 온전히 표현하기 어렵다는 문제가 있다.

수식의 경우는 문서에서 수식이 포함된 조문을 확인하고 해당 수식을 LaTeX 포맷으로 구성하여 조문 JSON 객체에 삽입하였다. LaTeX는 TeX 언어를 기반으로 한 문서 작성 도구로 수식 작성이 용이하여 수학 및 공학 등 이공계 학술 작업물에 활발히 사용되고 있다[8]. 표의 경우는 마찬가지로 표가 포함된 조문을 확인하고 해당 표를 Markdown 포맷으로 구성하여 조문 JSON 객체에 삽입하였다. Markdown은 표를 표현하기 위한 고유 문법이 존재하며 셀 병합 및 정렬 방법 등 세부적인 요소 또한 표현할 수 있다.

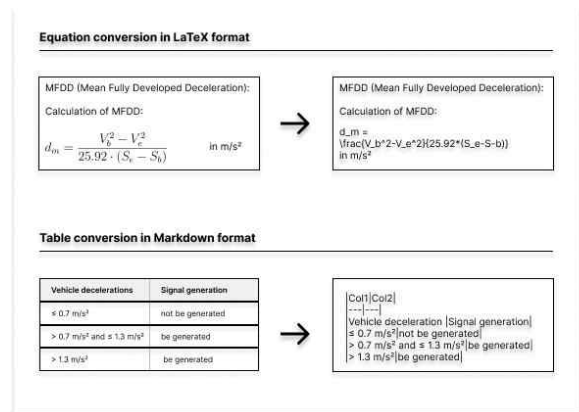


그림 3. 수식, 표 변환 예시  
Fig. 3. Formula and table conversion examples

계층적 전처리 방법론은 상기 목표한 법조문 체계의 반영과 압축적 정보 전달을 위해 고안되었지



만, 다음과 같은 이유로 chunk 검색을 위한 해상도 문제가 발생한다. 1) 문맥 정보의 보존을 위한 Article 단위 chunking으로 단일 chunk 간 토큰 크기가 균일하지 않다. 2) 상위 조항을 공유하고 있는 chunk 들의 경우 서로 중복되는 내용을 포함하고 있다. 3) 문서 출처 등 메타 데이터의 삽입이 모든 chunk에 공통으로 이루어진다. 계층적 전처리 방법론의 구조적 장점을 보존하되 이러한 검색 해상도 문제를 해결하기 위하여 본 연구에서는 Parent-Child chunking을 적용하여 데이터 형식을 구성하였다.

Parent-Child chunking의 핵심 아이디어는 벡터 검색을 위한 Child chunk와 답변을 위해 근거가 되는 Parent chunk의 분리이다. Child chunk는 포함된 내용은 적지만 이에 따라 의미상으로 서로 간 뚜렷하게 구분되는 특징이 있다. Parent chunk는 앞뒤 문맥을 포함한 풍부한 내용을 가지고 있지만 서로 간의 의미적 해상도는 비교적 떨어지는 경향성을 보인다. 사용 질의에 대하여 최초로 Child chunk로 높은 해상도로 유사도 검색을 진행하고 검색된 Child chunk를 포함하고 있는 Parent chunk를 확인 및 모델에게 전달하여 풍부한 맥락 정보를 활용할 수 있다.

본 연구에서는 문서를 문장 단위로 분할하여 child chunk로 구성하였으며 개별 문장에 고유한 ID를 부여하였다. Parent chunk는 앞서 계층적 전처리 방법론을 적용한 포맷으로 구성되었으며 각 parent chunk에 포함되는 child chunk의 ID에 대한 정보를 가지고 있다. 최초 사용자 질문 시 벡터 유사도 검색을 이용하여 child chunk를 획득하고 ID를 기반으로 이에 대응하는 계층적 전처리 방법론 형식의 parent chunk를 확인 및 이를 모델에게 전달한다.

본 연구에서는 계층적 전처리 방법론의 작동을 기존 방법론과 비교하기 위해 다음과 같이 4개의 형식으로 원본 문서를 가공하였다.

표 1. 문서 가공 표  
Table 1. Document Processing Table

	가공 방법	분할 방식
단순 텍스트	없음	재귀분할
Markdown	마크 다운으로 전환	재귀분할
JSON	계층적 전처리 방법론으로 전환	의미분할

PC-JSON	Legal JSON + Parent-Child 구조 적용	의미분할
---------	---------------------------------	------

\*4가지 모두 기본 chunk 사이즈는 500토큰이며 오버랩 단위도 100토큰이다.

### 2.3 벡터 임베딩 및 유사도 검색

사용자 질의와 개별 chunk 간의 유사도를 효과적으로 비교하기 위해서 텍스트 데이터를 숫자로 구성된 벡터로 전환하는 임베딩 과정이 필요하다. 임베딩은 오픈소스 모델로 진행하였으며 검색 효율을 위해 상용 벡터 스토어 프레임워크를 이용하여 저장하였다. 유사도 검색은 코사인 유사도를 사용하여 진행하며 해당 방법은 개별 chunk 크기와 상관없이 적용 가능하다는 장점이 있다. 사용자 질문 입력 시, 해당 질문을 바탕으로 벡터 저장소에서 유사도 검색을 시행하고 유사도가 가장 높은 상위 K개의 문서를 획득한다. Parent-Child chunking이 적용된 데이터의 경우는 벡터 임베딩 및 유사도 검색을 이용하여 사용자 질의와 가장 유사한 K개의 child chunk를 획득하고 연동된 ID를 바탕으로 검색된 child chunk를 포함하는 parent chunk를 확인하고 모델에게 전달한다.

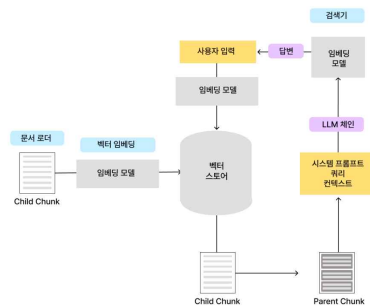


그림 4. PC JSON 도식  
Fig. 4. PC JSON process

### 2.4 성능 평가

모델의 성능 평가는 우리의 새로운 데이터 전처리 방법론과 기존의 방법론 간의 성능 차이를 확인하기 위하여 진행되었으며, RAG 모델의 성능을 평가하기 위한 법률 문서 기반의 질의를 준비하여 해

당 질문에 답변하는 정도를 수치화하여 성능지표로써 활용하고자 하였다[9].

모델이 제공한 답변의 정답 여부를 객관적으로 확인하기 위하여 “예” 혹은 “아니오”로 대답이 가능한 단답형 형식 또는 4개의 선지 중 옳은 선지를 고르는 형식으로, 문제로 구성하였다. 이 과정에서 표나 수식, 각종 법률 용어의 개념과 특정 상황에서 시나리오를 고려하였으며, 최종적으로 구성된 질문은 UNECE Regulation 78, 139, 152를 총합하여 51개 문항이고 단답형 11개 문항, 객관식 12개 문항, 참/거짓을 판단하는 문제 32개 문항으로 이루어져 있다.

표 2. 실험 결과 데이터 표

Table 2. Experimental result data table

데이터	정답	오답	정답률
PDF	36	15	36/51
Markdown	39	12	39/51
JSON	32	19	32/51
PC-JSON	44	7	44/51

### III. 결 과

앞서 제안한 전처리 방법론별로 각기 다른 RAG 프로세스를 구축하였고, 평가 질문지를 통해 데이터 타입별 평가를 진행하였다. PDF와 Markdown은 각기 36, 32개의 정답을 맞췄으며, JSON 형식의 데이터는 다른 데이터 형태에 비해 해상도가 크게 떨어지는 문제로 가장 정답률이 낮았다. 그리고 이를 보완한 PC-JSON 형식의 RAG 프로세스는 해상도 문제와 PDF, Markdown의 데이터 타입에서 발생한 문제를 해결하고 최종적으로 가장 높은 정답률을 보였다.

### IV. 결 론

본 논문에서는 법률 RAG 모델에 상하 구조가 있는 데이터의 전처리 방법과 법을 쉽게 해석할 수 있게 하는 프롬프트를 제안하였다. 제안된 방법은 수식과 테이블을 적절한 형태로 변환시킨 JSON 포맷에 Parent-Child Chunking을 적용한 데이터로 RAG Process를 진행하고 법률 데이터에 맞는 프롬프트로

기존의 방법보다 개선된 업무 과정을 보여준다. 이를 통해 제동장치와 관련한 법규 외 다른 법규에 적용할 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] Office of Public Affairs, "Takata Corporation Pleads Guilty, Sentenced to Pay \$1 Billion in Criminal Penalties for Airbag Scheme," February 27, 2017.
- [2] Yiquan Wu, Siying Zhou, Yifei Liu, Weiming Lu, Xiaozhong Liu, Yating Zhang, Changlong Sun, Fei Wu, Kun Kuang, "Precedent-Enhanced Legal Judgment Prediction with LLM and Domain-Model Collaboration," October 13, 2023.
- [3] S. B. Choi and S. K. Jung, "Trends and Market Analysis of Electric Braking Systems," Korea Institute of Science and Technology Information, pp. 60-62, November 2017.
- [4] H. W. Lee, "A Study on the Improvement of Motor Vehicle Safety Standards," Ajou University, pp. 5-6, December 2010.
- [5] Liu, H., Ning, R., Teng, Z., Liu, J., Zhou, Q., & Zhang, Y., "Evaluating the logical reasoning ability of chatgpt and gpt-4", arXiv preprint arXiv:2304.03439, pp. 6-7, April 2023
- [6] Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, Haofen Wang.. "Retrieval-Augmented Generation for Large Language Models: A Survey", arXiv preprint arXiv:2312.10997, pp. 2-4, December 2023.
- [7] Xin, J., Afrasiabi, C., Lelong, S., Adesara, J., Tsueng, G., Su, A. I., & Wu, C., "Cross-linking BioThings APIs through JSON-LD to facilitate knowledge exploration", BMC bioinformatics, pp 1-7, February 2018
- [8] Lemenkov, V. and Lemenkova, P., "Using TeX markup language for 3D and 2D geological plotting", Foundations of Computing and Decision

Sciences, 46(1), pp 43-69, March 2021

- [9] Gauthier Guinet, Behrooz Omidvar-Tehrani, Anoop Deoras, Laurent Callot “Automated Evaluation of Retrieval-Augmented Language Models with Task-Specific Exam Generation” arXiv preprint arXiv:2405.13622

# 피드백 루프 탈출하기: LLMs 기반 사용자 피드백 증강을 통한 추천 다양성 증진 및 인기편향 완화 방안

정혜진\*, 최상민\*\*

## Escaping the Feedback Loop: Enhancing Recommendation Diversity and Mitigating Popularity Bias through LLM-based User Feedback Augmentation

Hye-Jin Jeong\* and Sang-Min Choi\*\*

### 요 약

추천 시스템에서 사용자 피드백을 통해 사용자의 선호도를 파악하는 것은 매우 중요하다. 하지만 사용자 피드백은 종종 편향되는 경향이 있다. 편향된 데이터는 사용자의 선호도를 왜곡함으로써 소수의 인기 있는 항목의 노출만을 증가시키고 그 결과 피드백 루프가 형성된다. 수집된 피드백의 편향을 해결하기 위해 여러 방법이 제안되었으나, 여전히 일부 방법은 학습단계만을 조정함으로써 데이터 단계에서의 편향을 완전히 제거하지 못하고 있다. 데이터의 편향성으로 인해 발생하는 문제를 해결하기 위해 우리는 사용자가 아직 발견하지 못한 항목에 대한 피드백을 LLMs(Large Language Models)을 통해 증강하는 추천 방법을 제안한다. 먼저 피드백의 편향 정도를 측정하고 우리의 방법론이 실제 추천 과정에서 효과적인지 분석하였다. 그 결과 LLMs을 통한 피드백 증강이 데이터의 원본성을 최대한 유지하면서 추천 결과의 다양성을 높이고 데이터 단계에서의 편향을 줄이는데 기여함을 확인할 수 있다.

### Abstract

The In Recommender systems, capturing user preferences through feedback is essential. However, user feedback often exhibits bias, distorting true user's true preferences and disproportionately increasing the visibility of a few popular items, ultimately forming a feedback loop. While various methods have been proposed to mitigate bias in collected feedback, some approaches focus only on adjustments in the training phase, leaving bias at the data level insufficiently addressed. To Address issues stemming from data bias, we propose a recommender method that uses a large language models(LLMs) to augment feedback for items that users have not yet discovered. We first measure the degree of bias in user feedback and then analyze the effectiveness of our methodology within the recommender process. Our results demonstrate that LLM-based feedback augmentation effectively reduces bias at the data level, enhancing recommendation diversity while preserving data authenticity.

### Key words

recommender system, popularity bias, feedback loops, LLMs, data augmentation

---

\* 경상국립대학교 학사과정, cucumber\_love@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과 조교수, jerassi@gnu.ac.kr (교신저자)

※ 본 논문은 교육부와 한국연구재단의 지원으로 지원 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0), 정부(과학기술정보통신부) 재원의 한국연구재단 지원(RS-2022-00165785)과 2023년도 교육부의 지원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다(2021RIS-003).

## I. 서론

전자상거래가 전 세계적으로 확산되면서 우리는 오프라인 매장에서 볼 수 없었던 다양한 제품들을 온라인에서 쉽게 접할 수 있게 되었다. Anderson[2]은 디지털화가 소수집단을 겨냥한 틈새 제품이 더 큰 시장 점유율을 차지할 수 있도록 할 것이라고 주장하였다. 하지만 현실에서는 여전히 극소수의 인기 제품들이 지배적인 롱테일 현상을 확인할 수 있다[13].

이러한 편향된 데이터로 학습된 추천 시스템은 인기 있는 극소수의 아이템만 반복적으로 노출시키고 이것이 다시 사용자에게 소비됨으로써 피드백 루프가 형성된다[15]. 결과적으로 이러한 현상이 반복되는 시스템에서 수집된 사용자 피드백이 사용자의 선호도를 반영한 것인지에 대한 의문을 갖게 된다.

따라서 본 연구에서는 데이터 단계에서의 편향을 완화하기 위해 LLMs을 활용하여 데이터를 증강한 후 생성된 텍스트의 신뢰성을 검증하기 위해 자기 반영 과정을 수행한다. 실험 결과를 통해 이러한 방법론이 추천되는 아이템의 다양성이 증가하고 편향의 영향을 줄일 수 있음을 발견한다.

## II. 관련 연구

추천시스템에서 데이터 편향은 추천의 성능과 시스템의 신뢰도를 떨어뜨린다. 뿐만 아니라 편향성이 강한 데이터는 궁극적으로 피드백 루프라는 또 다른 문제로 이어진다[15]. 따라서 연구자들이 사용자 피드백의 편향을 측정하고 완화하려 노력하고 있다 [1, 3, 8, 9]. 그러나 기존의 편향 지표는 표준화되어 있지 않아 객관성이 불명확하며, 모델을 수정하여 편향을 제거하려는 접근은 계산 비용을 증가시킨다는 문제가 있다[11]. 설령 사용자의 선호도 편향을 식별하더라도 단순히 편향성을 제거하는 방법은 데이터의 원본성을 손실할 수 있는 위험이 있어 신중한 접근이 필요하다.

이러한 문제를 해결하기 위해 기존 피드백의 긍정적 및 부정적 분포 균형을 계산하여 사용자 피드백을 증강하는 연구도 있다[4]. 이에 반해 전통적인

데이터 증강 방법은 종종 잡음과 방해물 초래할 수 있다는 점을 강조하며, 이를 방지하면서 효과적으로 데이터를 증강할 수 있는 다중 행동 데이터 증강 접근 방식을 제안하는 연구도 있다[5].

최근 LLMs의 등장 이후, 텍스트 생성 기능을 활용하여 데이터를 증강하는 사례들이 나타났다[7]. 예를 들어, [6]에서는 LLMs의 요약기능을 활용하여 각 문장의 단어를 의미가 비슷한 다른 단어로 대체하여 의료 대화 데이터를 증강하였다. 하지만 LLMs이 생성하는 출력의 신뢰성에 대한 우려, 특히 입력과 관련이 없거나 잘못된 정보를 포함하는 환각(hallucination) 문제는 여전히 해결해야 할 문제로 남아있다[12]. 따라서 LLMs이 생성한 텍스트의 신뢰성을 검증하기 위해 자기 반영 단계를 도입하는 등의 방법론을 제시한 연구가 있다[14]. 이를 통해 생성된 텍스트의 품질을 높이고 신뢰성을 강화할 수 있다.

## III. 방법론

본 논문에서는 MovieLens(ml-25m)의 Tag Genome 데이터셋[10]을 사용하여 사용자 피드백의 편향을 측정한다. 태그는 사용자에게 의해 영화에 부여되는 주요 속성으로, 'friendship', 'based on a book' 또는 'spaghetti western' 등과 같이 영화의 주요 특성을 핵심 키워드를 사용하여 나타낸다. 본 실험에서는 이러한 특성을 반영하여 태그를 아이템의 메타 데이터로 활용한다.

### 3.1 사용자 피드백에 내재된 편향 조사

먼저 영화 별 보유한 태그의 수를 추출하여 그 분포를 그림 1.과 같이 나타내었다. 그림에서 x축은 상호작용 수를 기반으로 정렬한 영화의 랭킹, y축은 해당 랭킹 영화의 태그 수를 나타낸다. 이는 태그 정보가 소수의 인기있는 아이템에만 집중되어 있음을 보여준다. 정확한 편향성 측정을 위해 영화 별 보유한 태그수의 롱테일 지수를 계산한 결과, 상위 20%의 아이템이 전체 태그 빈도의 96.07%를 차지하는 매우 극단적인 롱테일 분포를 보이고 있음을 알 수 있었다.

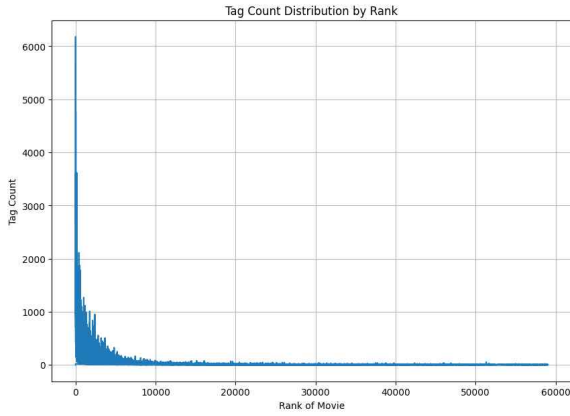


그림 1. 영화랭킹별 보유 태그수 분포  
Fig. 1. Tag Distribution by Movie Ranking

### 3.2 편향된 데이터에서 대푯값 설정하기

본 논문에서는 그림 1.에서 과도하게 많은 태그 수를 보유한 상위 랭크의 영화들을 이상치로 결정한다. 이상치가 갖는 태그를 제거하는 것이 편향을 완화하는 데 효과적일지라도, 태그가 없는 아이템의 표현을 학습하는데는 여전히 어려움이 존재하기 때문에 데이터 증강을 통한 편향 완화가 필요하다. 먼저, 이상치를 고려한 데이터의 대푯값을 추출하여 증강할 태그 수를 결정한다. 이상치는 식 (1)과 같이 IQR(사분위수 범위)을 사용하여 도출된다.

$$X_{non-outliers} = x_i \in X \mid Q1 - 1.5 \times IQR \leq x_i \leq Q3 + 1.5 \times IQR \quad (1)$$

그런 다음, 우리는 이상치 값의 평균과 정상 데이터의 평균을 계산하여 각 값에 가중치를 적용함으로써 태그 수의 대푯값을 설정하였다.

$$N_{non-outliers} = |X_{non-outliers}| \quad (2)$$

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (N = \text{전체 데이터의 수}) \quad (3)$$

$$\mu_{weighted} = \frac{N_{non-outliers} \times \mu_{non-outliers} + N_{outliers} \times \mu_{outliers}}{N} \quad (4)$$

실제 추천과정에서 어떤 대푯값이 더 나은 성능

을 갖는지 비교하기 위해 이상치를 제외한 데이터의 평균, 전체 데이터의 평균, 그리고 가중평균(식 (2)-(4))을 각각 사용하였다.

위 3가지의 대푯값으로 증강하여 추가 실험을 진행하였을 때 각각의 차이는 크지 않았지만 가중평균을 사용한 증강 데이터셋의 추천의 다양성이 제일 높아 본 실험에서는 해당 데이터 셋을 사용하였다.

### 3.3 LLM을 활용한 데이터 증강과 자가 검증

LLM은 특히 텍스트 형태의 데이터 증강에 뛰어난 성능을 보여준다. 그러나, 2장에서 언급했듯이 '환각'의 위험이 있으므로, 출력의 신뢰성을 보장하기 위해 엄격한 지침이 필요하다.

The screenshot shows a chat interface with a blue speech bubble from 'You' and a grey speech bubble from 'LLMs'. The 'You' bubble contains a prompt asking the LLM to act as an expert in labeling movies with appropriate tags, to create tags for movies with fewer than 7 tags, and to provide criteria for generating tags. The 'LLMs' bubble shows a table of movie IDs and their generated tags. A second 'You' bubble asks for an explanation of the reason for generating tags for a specific movie. The 'LLMs' bubble provides an analysis of the new tag 'woman' added to movie 77, comparing it to the original tags and explaining the reason for its addition based on the movie's plot.

그림 2. 데이터 증강과 자가검증에서 사용된 프롬프트  
Fig. 2. Prompts for Data Augmentation and Self-Verification

우선, 우리는 IMDb에서 영화 요약 및 리뷰를 포함한 텍스트를 크롤링하였다. 그런 다음, 3.2절에서 설정한 대푯값과 증강 지침을 포함하는 프롬프트를 그림 2.와 같이 생성한 후 GPT-4o-mini 모델을 사용한다. 첫 번째 단계는 LLMs에서 수행해야 할 작

업의 범위를 제한하기 위해 페르소나 변수를 할당하는 것이다. 페르소라란 모델에게 특수 상황에 대한 역할을 부여하여 원할한 응답을 끌어내기 위한 방법으로 본 실험 프롬프트의 'You are an expert in labling moviles with appropriate tags.'부분에 해당한다. 다음으로 생성할 태그 수, 생성 규칙과 함께 태그를 생성할 때 참조할 수 있는 기존 데이터의 태그 목록이 제공된다.

LLMs이 생성한 초기 텍스트는 환각이 발생하는지 확인하기 위해 'Explain the reason~'과 같은 자기 반영 과정을 거친다. 새로운 태그가 업데이트된 일정 수의 영화를 무작위로 선택하고, 각 영화에 대해 새로운 태그를 생성하는 과정과 이유를 설명하도록 요청받는 과정을 거침으로써 신뢰성이 검증된 태그는 추천에서 아이템의 메타데이터로 활용된다.

#### IV. 실험 결과

본 실험에서는 아래의 세가지 데이터셋을 사용하여 비교실험을 진행하였다.

- 원본 데이터셋(Raw)
- 가중 평균을 기반으로 증강된 데이터셋(Ours)
- 이상치를 제거한 데이터셋(Remove)

이를 추천 모델에 적용한 결과 각각의 차이는 미미했지만 가중평균을 기반으로 한 데이터셋이 가장 높은 다양성 점수를 기록한 것을 그림 3.에서 확인할 수 있다.

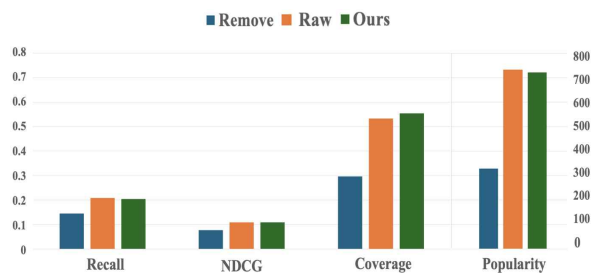


그림 3. 3개의 증강데이터를 사용한 실험 결과  
Fig. 3. Results with 3 Augmented Data Sets

각각의 방법론을 비교하기 위해 Recbole 라이브러리에서 제공하는 Recall, NDCG, Coverage, Popularity 지표를 사용하였다[16]. 우리의 방법(Ours)

은 원본 데이터(Raw)에 비해 Recall 과 NDCG에서 다소 낮은 정확도를 보이지만 Popularity 지표는 더 낮게 나타나는 것으로 보아 우리의 방법론이 인기 편향의 문제점을 부분적으로 해소한 것으로 볼 수 있다. 또한 우리의 방법론에서 Coverage가 더 높아, 원본 데이터셋을 사용하는 것보다 추천의 다양성이 향상되었음을 확인할 수 있다. 반면, 이상치를 제거한 데이터셋은 다른 데이터셋에 비해 정확도가 크게 떨어지는데, 이는 이상치 제거 과정에서 이용자 정보를 포착할 수 있는 많은 표현들이 손실되었을 것으로 예측할 수 있다.

#### V. 결 론

편향된 데이터는 피드백루프를 심화시켜 추천시스템의 정확도와 신뢰도를 떨어뜨린다. 이러한 문제를 해결하기 위해 본 논문에서는 추천시스템에 사용되는 영화 데이터의 편향을 측정하고 LLM을 활용하여 부족한 데이터를 증강하였다.

실험 결과, 이러한 방법론이 데이터의 인기 편향을 성공적으로 완화하고 추천 결과의 다양성을 향상시켰음을 입증하였다. 추가적으로, 이상치 제거가 추천 성능에 미치는 영향을 분석하기 위한 보충 실험을 수행하였다. 이러한 추가 실험을 통해 이상치 제거가 정확도와 다양성 간의 트레이드오프에 어떻게 영향을 미치는지를 이해할 수 있었다.

#### 참 고 문 헌

- [1] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. 2019. Managing popularity bias in recommender systems with personalized re-ranking. arXiv preprint arXiv:1901.07555 (2019).
- [2] Chris Anderson. 2008. *The long tail: why future of business is selling less of more*. Hyperion, New York.
- [3] Calmon Flavio, et al. "Optimized pre-processing for discrimination prevention." *Advances in neural information processing systems* 30 (2017).
- [4] Chen, Lei, et al. "Improving recommendation

- fairness via data augmentation." *Proceedings of the ACM Web Conference 2023*.
- [5] Jing Xiao, Weike Pan, and Zhong Ming. 2024. A Generic Behavior-Aware Data Augmentation Framework for Sequential Recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1578-1588.
- [6] Haixing Dai, Zhengliang Liu, Wenxiong Liao, Xiaoke Huang, Yihan Cao, Zihao Wu, Lin Zhao, Shaochen Xu, Wei Liu, Ninghao Liu, et al. 2023. Auggpt: Leveraging chatgpt for text data augmentation. *arXiv preprint arXiv:2302.13007* (2023).
- [7] Bosheng Ding, Chengwei Qin, Ruochen Zhao, Tianze Luo, Xinze Li, Guizhen Chen, Wenhan Xia, Junjie Hu, Anh Tuan Luu, and Shafiq Joty. 2024. Data augmentation using llms: Data perspectives, learning paradigms and challenges. *arXiv preprint arXiv:2403.02990* (2024).
- [8] Jesse Vig, Shilad Sen, and John Riedl. 2012. The tag genome: Encoding community knowledge to support novel interaction. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 3 (2012), 1-44.
- [9] Sahin Cem Geyik, Stuart Ambler, and Krishnaram Kenthapadi. 2019. Fairness-aware ranking in search & recommendation systems with application to linkedin talent search. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining*. 2221-2231.
- [10] F. Maxwell Harper and Joseph A. Konstan. 2015. The MovieLens Datasets: History and Context. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 5, 4 (2015), 19:1 -19:19. <https://doi.org/10.1145/2827872>
- [11] Neil C Thompson, Kristjan Greenewald, Keeheon Lee, and Gabriel F Manso. 2021. Deep learning's diminishing returns: The cost of improvement is becoming unsustainable. *Ieee Spectrum* 58, 10 (2021), 50-55.
- [12] Huang, Lei, et al. "A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions." *arXiv preprint arXiv:2311.05232* (2023)
- [13] Yoon-Joo Park and Alexander Tuzhilin. 2008. The long tail of recommender systems and how to leverage it. In *Proceedings of the 2008 ACM conference on Recommender systems*. 11-18.
- [14] Ziwei Ji, Tiezheng Yu, Yan Xu, Nayeon Lee, Etsuko Ishii, and Pascale Fung. 2023. Towards mitigating hallucination in large language models via self-reflection. *arXiv preprint arXiv:2310.06271* (2023).
- [15] Masoud Mansoury, Himan Abdollahpouri, Mykola Pechenizkiy, Bamshad Mobasher, and Robin Burke. 2020. Feedback loop and bias amplification in recommender systems. In *Proceedings of the 29th ACM international conference on information & knowledge management*. 2145-2148.
- [16] Wayne Xin Zhao, Shanlei Mu, Yupeng Hou, Zihan Lin, Yushuo Chen, Xingyu Pan, Kaiyuan Li, Yujie Lu, Hui Wang, Changxin Tian, et al. 2021. Recbole: Towards a unified, comprehensive and efficient framework for recommendation algorithms. In *proceedings of the 30th acm international conference on information & knowledge management*. 4653-4664.



# 폐기된 모바일 기기를 활용한 실시간 알고리즘 기반 낙상 감지 시스템

박익범\*, 차승훈\*, 신연순\*\*

## Real-Time Algorithm-Based Fall Detection System Using Repurposed Mobile Devices

Ik-Bum Park\*, Seung-Hoon Cha\*, and Youn-Soon Shin\*\*

### 요약

본 연구는 고령화 사회에서 증가하는 낙상 사고 문제에 대응하기 위해 폐기된 모바일 기기를 활용한 실시간 낙상 감지 시스템을 제안한다. 제안된 시스템은 모바일 기기에 내장된 카메라와 Apple Vision Framework를 사용하여 사용자의 다양한 동작을 분석하고 낙상 여부를 판단한다. 얼굴과 골반의 변위 및 신체 비율 변화를 기반으로 낙상을 감지하는 알고리즘을 설계하여, 즉각적인 낙상 상황 감지와 알림 기능을 구현하였다. 본 시스템은 접근성과 이동성 측면에서 우수하며, 환경 보호에도 기여할 수 있다.

### Abstract

This study proposes a real-time fall detection system using repurposed mobile devices to address the growing issue of fall incidents in an aging society. The proposed system leverages the built-in camera and Apple's Vision Framework on mobile devices to analyze various user movements and determine falls. A fall detection algorithm, based on facial and pelvic displacement and body ratio changes, was designed to enable immediate fall detection and alerting. This system demonstrates high accessibility and mobility, along with potential contributions to environmental protection.

### Key words

Fall detection, Real-time detection, Elderly safety, Mobile devices, Pose Estimation, Environmental protection

## I. 서론

### 1.1 서론

고령화 사회의 가속화와 함께 노인 인구의 증가

로 낙상 사고가 주요 사회적 문제로 떠오르고 있다. 특히 낙상은 노인들에게 치명적인 부상과 사망을 유발하는 주된 원인으로, 낙상 이후 즉각적인 대처

\*동국대학교 공과대학 컴퓨터공학전공{pibum, papa021326}@naver.com,

\*\*동국대학교 AI소프트웨어융합학부, ysshin@dongguk.edu(교신저자)

※ "본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음"(2023-0-00049)

※ "본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음(IITP-2024-2020-0-01789)"

가 이루어지지 않으면 신체적 손상뿐만 아니라 심리적 공포와 경제적 부담까지 초래할 수 있다. 이를 해결하기 위해 본 연구에서는 모바일 기기와 낙상 감지 알고리즘을 결합한 실시간 낙상 감지 시스템을 개발하고자 한다. 본 시스템은 모바일 기기에 내장된 카메라를 통해 사용자 동작을 추적하며, 낙상 여부를 판단하는 알고리즘을 적용하여 고령자 및 낙상 위험이 높은 사용자에게 안전망을 제공하는 것을 목표로 한다. 특히 폐 모바일 기기를 활용함으로써 접근성과 이동성은 물론 자원 절약과 환경 보호에도 기여함을 목표로 한다. 또한 실시간 감지가 가능하여 사용자에게 신뢰성 있는 안전 시스템을 제공한다.

## 1.2 선행연구

낙상 감지 연구는 크게 비전 기반 방법과 웨어러블 장치 기반 방법으로 나눌 수 있다. 비전 기반 방법은 카메라를 활용하여 사용자의 동작을 추적하고 분석하여 낙상을 감지하는 방식이다. 최근 OpenPose와 같은 스켈레톤 추출 알고리즘을 사용하여 인체 관절 데이터를 실시간으로 수집하고 사용자의 골반 하강 속도, 몸 중심선과 지면 간의 각도 등을 바탕으로 낙상을 감지하는 연구들이 주목 받고 있다 [1, 2].

반면, 웨어러블 장치 기반의 낙상 감지는 가속도 계나 자이로스코프와 같이 센서를 통해 낙상 시의 급격한 속도 변화를 감지하는 방식이다. 하체의 기울기 센서를 이용하여 걷기 중 각도 변화와 낙상 후의 각도 상태를 측정하여 낙상을 감지하는 등의 연구가 진행되었다 [2]. 이와 같은 방식은 장치 착용이 불편하거나 배터리 지속 시간이 부족할 경우 장기간 착용이 어렵고, 이를 일상적으로 사용하는 데에 제약이 있다.

본 연구는 이러한 기존 연구들을 바탕으로, 폐기된 모바일 기기를 활용한 실시간 낙상감지 시스템을 제안하고자 한다. 이는 기존의 웨어러블 기기 방식이 가진 제약을 보완할 수 있고, 포즈 추출만 가능하다면 3가지의 낙상 감지 알고리즘을 통해 실시간으로 낙상 여부를 판단하여 낙상 후 보호자에게

알림을 보내주어 즉각적인 대처가 가능할 수 있도록 할 수 있다.

## II. 낙상 감지 시스템

### 2.1 낙상 감지 시스템

본 논문에서는 낙상 감지 알고리즘을 개발하여 인간의 다양한 행동을 분석하고, 그 중 위험 상태와 낙상 상태를 구분하고자 한다. 제안하는 시스템은 그림 1과 같다. 이를 위해 그림 2와 같이 걷기, 뛰기, 쪼그려 앉기, 낙상 등 여러 행동을 구분하고, 각 행동에서 머리와 골반의 Y축 변위를 측정하여 분석하였다. 이 과정에서 특정 임계치를 초과하는 변위가 발생하면 이를 위험 상태로 간주하고, 실시간으로 감지하는 시스템을 구현하였다. 위험 상태가 감지되면 보호자에게 낙상 영상을 제공하기 위해 영상 녹화를 시작한다. 위험 상태에서 신체 비율의 변화량이 임계치를 넘어 5초 이상 유지되면, 이를 낙상 상태로 간주하여 영상 녹화를 종료하고 보호자에게 낙상 알림과 함께 낙상 영상을 제공한다.

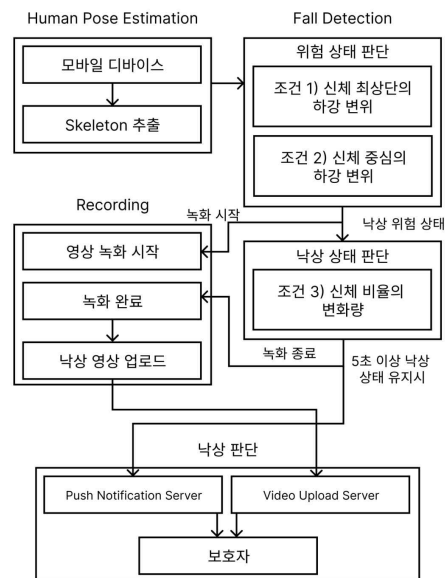


그림 1. 낙상 감지 시스템  
Fig. 1. Fall Detection System



그림 2. 앉기, 걷기, 뛰기, 낙상 이미지  
Fig. 2. Sitting, Walking, Running, Fall Image

## 2.2 낙상 감지 알고리즘

본 연구에서는 머리, 골반 좌표의 변위와 신체의 전체적인 비율 변화를 바탕으로 세 가지 알고리즘을 통해 낙상을 감지하였다.

초당 30프레임의 이미지를 Apple의 Vision Framework[3]를 이용하여 주요 관절 포인트를 추론하였고, 아래와 같은 세 가지 알고리즘을 이용하여 낙상 여부를 판단하였다.



그림 3. 얼굴 및 골반 포인트  
Fig. 3. Face and pelvis point

얼굴 위치 기반 낙상 감지 알고리즘은 그림 3의 좌측과 같이 머리 위치의 변화를 분석하여 낙상을 감지한다. 얼굴의 주요 포인트인 왼쪽 귀, 오른쪽 귀, 그리고 코의 y좌표를 기반으로 중심 위치를 계산한다. 왼쪽 귀의 y좌표를  $y_l$ , 오른쪽 귀의 y좌표를  $y_r$ , 코의 y좌표를  $y_n$  이라고 할 때 중심 위치  $C_y$ 를 수식 (1)과 같이 정의할 수 있다.

$$C_y = \text{avg}(y_l, y_r, y_n)$$

(1)실시간으로 수집되는 해당 데이터는 최근 0.1초 동안의 중심 위치 변화를 수식 (2)를 통해 계산하여 변위를 측정한다.

$$\Delta C_y(t) = C_y(t) - C_y(t - 0.1)$$

(2)이때,  $|\Delta C_y(t)|$ 가 일정한 임계값을 초과하는 경우 낙상으로 판단한다. 골반 위치 기반 감지 알고리즘은 양쪽 골반의 y좌표를 이용하여 중간 위치를

계산한다. 양쪽 골반의 y좌표  $y_{lp}$ 와  $y_{rp}$ 를 사용하여 중간 위치  $P_y$ 를 수식 (3)과 같이 계산한다.

$$P_y = \frac{y_{lp} + y_{rp}}{2} \quad (3)$$

실시간으로 수집되는 해당 데이터는 최근 0.1초 동안의 중심 위치 변화를 수식 (4)를 통해 계산하여 변위를 측정한다.

$$\Delta P_y(t) = P_y(t) - P_y(t - 0.1) \quad (4)$$

이때,  $|\Delta P_y(t)|$ 가 일정한 임계값을 초과하는 경우 낙상으로 판단한다. 수식 (5)와 같이, 얼굴 위치 기반 감지 알고리즘과 골반 위치 기반 감지 알고리즘의 임계치를 모두 초과할 경우 이를 위험 상태로 판단하고, 신체 비율 기반 감지 알고리즘을 적용하여 낙상 여부를 판단한다.

$$|\Delta C_y(t)| > T_{face} \text{ and } |\Delta P_y(t)| > T_{pelvis} \quad (5)$$

신체 비율 기반 감지 알고리즘은 신체의 전체적인 비율 변화를 활용하여 낙상을 감지한다.

$$W = x_{max} - x_{min}, H = y_{max} - y_{min} \quad (6)$$

수식 (6)은 신체 관절 좌표 집합의 최댓값  $x_{max}, y_{max}$ 와 최솟값  $x_{min}, y_{min}$ 를 사용해 경계 상자를 설정한다.

$$R = \frac{W}{H} \quad (7)$$

수식 (7)과 같이 R이 임계치(1.0)를 초과할 경우 사용자가 수평으로 누운 상태로 간주하여 낙상으로 판단한다. 특히, 이 상태가 5초 이상 지속될 때 보호자에게 알림을 전송하여 즉각적인 대응이 가능하게 한다. 해당 상황은 그림 4와 같다.

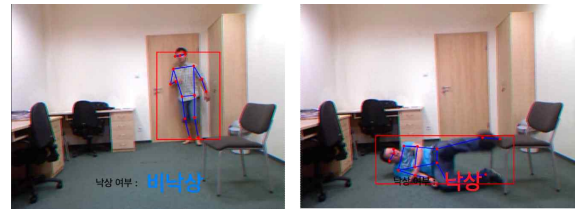


그림 4. 신체 비율 기반 낙상 감지  
Fig. 4. Fall detection based on body ratio

3가지 알고리즘의 올바른 임계치를 설정하기 위해, UR Fall 데이터셋[4]의 62개 영상과 4명의 실험 지원자가 제공한 50개의 녹화 영상을 활용하여 낙상과 비 낙상의 변위 차이를 분석하였다.

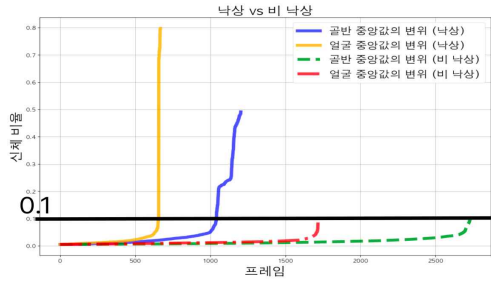


그림 5. 변위 최고값 그래프

Fig. 5. Graph of maximum displacement

그림 5에서와 같이 비 낙상 시 변위의 최대값이 0.1 이하로 나타나므로, 얼굴 및 골반 변위를 이용한 위험 상태 감지 알고리즘의 임계값을 0.15로 설정하였다. 또한, 신체 비율을 통한 낙상 확정 알고리즘의 임계값은 그림 6에서 볼 수 있듯이 신체 비율이 1보다 커질 때 낙상으로 판단하도록 설정하였다.

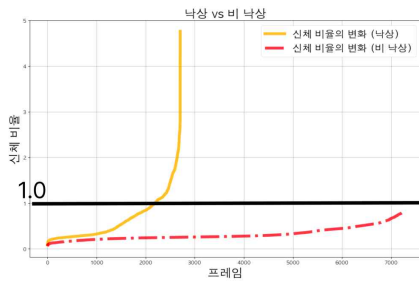


그림 6. 신체 비율 최고값 그래프

Fig. 6. Graph of Maximum Body ratio

위에서 설명한 알고리즘을 바탕으로 구현된 모바일 애플리케이션을 이용하여 미리 촬영된 영상 112개를 이용하여 낙상 여부만을 판단하였다.

행동별로 분류된 영상에 대해 낙상 여부를 판별한 실험 결과는 표 1에 제시되어 있다.

표 1. 낙상 탐지 결과

Table 1. Experimental Environment

행동	성공한 영상 수	실패한 영상 수	총 영상 수
걷기	11	1	12
뛰기	9	1	10
쪼그려 앉기	10	2	12
허리 굽히기	12	1	13
누웠다가 일어나기	14	0	14
낙상	49	2	51

실험 결과, 대부분의 영상에서 임계치를 통해 낙상 여부를 정확히 추론할 수 있었다. 그러나 포즈 추론 과정에서 발생하는 이상치가 포함된 경우에는 정확한 검증이 어려워 탐지에 실패하는 경우도 있었다. 이 문제는 보호자의 휴대폰으로 전송된 영상을 통해 해당 상황을 재확인함으로써 큰 문제없이 보완할 수 있을 것으로 판단된다.

## 2.3 낙상 감지 애플리케이션

낙상 감지 애플리케이션은 낙상을 탐지하기 위해 카메라와 알고리즘이 구동되는 애플리케이션과, 낙상 상황 발생 시 보호자가 영상을 확인하고 낙상 여부를 판단할 수 있는 보호자용 애플리케이션으로 구분된다.

### 낙상 감지 애플리케이션

낙상 감지 애플리케이션은 카메라와 알고리즘을 통해 사용자의 낙상을 감지한다. 감지 알고리즘은 세 단계로 구성되며, 먼저 얼굴과 골반의 변위를 기반으로 위험 상태를 탐지한다. 얼굴 및 골반 변위가 사전 설정된 임계치를 초과할 경우 이를 위험 상태로 간주하고, 이 시점부터 영상 녹화를 시작한다. 이후, 위험 상태에서 그림 4와 같이 신체 비율 기반 알고리즘의 임계치도 초과하며 이 상태가 5초간 지속될 경우, 영상 녹화를 종료하고 해당 상황을 낙상으로 판단한다. 이를 통해 낙상 발생 사실을 보호자에게 그림 7과 같은 푸시 알림으로 전송한다.

### 보호자 애플리케이션

보호자용 애플리케이션은 보호자가 알림을 수신한 후, 그림 7과 같이 녹화된 영상을 확인하여 실제로 낙상이 발생했는지를 판단할 수 있는 기능을 제공한다. 이를 통해 낙상 탐지의 정확도를 높이며, 실제 낙상 시 보호자가 신속하게 대응할 수 있도록 지원한다.

## III 결론

본 연구에서는 낙상 감지 시스템을 구현하여 위험 상태와 낙상 상태를 효과적으로 구분하는 방법을 제시하였다. 이를 위해 걷기, 뛰기, 쪼그려 앉기,

허리 굽히기, 누웠다가 일어나기, 낙상 등의 다양한 행동을 구분하고, 각 행동에서 머리와 골반의 Y축 변위 및 신체 비율 변화를 분석하여 낙상을 감지할 수 있었다.

연구 결과, 대부분의 낙상 상황을 성공적으로 감지할 수 있었으며, 보호자가 낙상 알림을 받고 실제 위험 상황인지를 추가로 판단할 수 있게 함으로써 시스템의 정확도를 높였다. 또한, 보호자는 낙상 발생 시 신속히 대처할 수 있어, 고령자나 낙상 위험이 높은 사용자에게 중요한 안전망을 제공하여 실제 활용 가능성을 높인다.

추가적으로, 이 시스템은 표준화된 포즈 추출이 가능한 모든 모바일 기기에서 사용할 수 있어, 폐기된 모바일 기기를 재활용하여 자원 절약과 환경 보호에도 기여할 수 있다.

### 참 고 문 헌

- [1] Chen, W., Jiang, Z., Guo, H., & Ni, X. (2020). Fall detection based on key points of human-skeleton using openpose. *Symmetry*, 12(5), 744.
- [2] Sun, J., Wang, Z., Pei, B., Tao, S., & Chen, L. (2015, August). Fall detection using plantar inclinometer sensor. In 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom) (pp. 1692-1697). IEEE.
- [3] Apple. (n.d.). DetectHumanBodyPoseRequest. <https://developer.apple.com/documentation/vision/detecthumanbodyposerequest>
- [4] Kwolek, B., & Kepski, M. (2014). Human fall detection on embedded platform using depth maps and wireless accelerometer. *Computer methods and programs in biomedicine*, 117(3), 489-501.

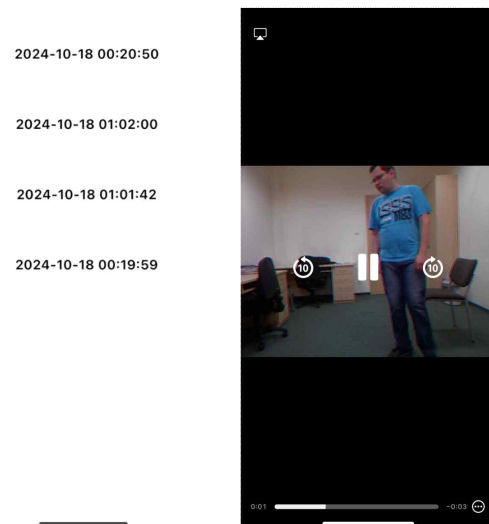
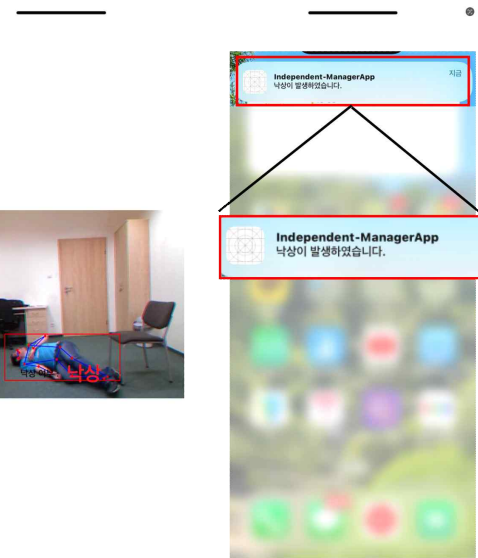


그림 7. 낙상 감지 애플리케이션 작동 이미지  
Fig. 7. Fall detection application operation image

# Performance Comparison of Deep Learning Models for Skin Disease Classification using the HAM10000 Dataset

C. Jaehyo\*, P. John\*, O. Chibuike\*, and Yang. X\*

## HAM10000 데이터셋을 활용한 피부 질환 분류용 딥러닝 모델 성능 평가 및 비교

최재효\*, 박준\*, 옥팔라 치부이케\*, 양시아오펑\*

### Abstract

The Skin diseases, particularly life-threatening conditions like melanoma, require accurate and prompt diagnosis to ensure effective treatment outcomes. In this study, we compared the performance of 12 deep learning models—ConvNeXtTiny, ConvNeXtSmall, ConvNeXtBase, EfficientNetV2S, EfficientNetV2M, EfficientNetV2L, InceptionV3, MobileNet, MobileNetV2, ResNet50, ResNet50V2, and Xception—on the classification of seven types of skin lesions using the HAM10000 dataset. The results revealed that EfficientNetV2S achieved the best performance, with a top accuracy of 91.45% and the lowest loss of 0.3834. Xception followed closely, recording an accuracy of 89.76% with a loss of 0.4076. This study demonstrates the exceptional capability of deep learning models in skin disease classification, suggesting their potential to facilitate the early detection of critical conditions like melanoma. These findings highlight the importance of AI-based diagnostic tools in improving medical outcomes by enabling timely and accurate disease identification.

### 요 약

피부 질환, 특히 흑색종과 같은 생명을 위협할 수 있는 질환은 효과적인 치료를 위해 정확하고 신속한 진단이 필요하다. 본 연구에서는 HAM10000 데이터셋을 사용하여 7가지 유형의 피부 병변 분류에 대해 ConvNeXtTiny, ConvNeXtSmall, ConvNeXtBase, EfficientNetV2S, EfficientNetV2M, EfficientNetV2L, InceptionV3, MobileNet, MobileNetV2, ResNet50, ResNet50V2, Xception의 12가지 딥러닝 모델 성능을 비교했다. 그 결과 EfficientNetV2S가 가장 우수한 성능을 보였으며 최고 정확도 91.45%와 최저 손실 0.3834를 기록했다. Xception 모델은 정확도 89.76%와 손실 0.4076으로 두 번째로 높은 성능을 보였다. 본 연구는 딥러닝 모델이 피부 질환 분류에서 뛰어난 성능을 나타냈음을 입증하며, 흑색종과 같은 중대한 피부암 질환의 조기 발견에 기여할 가능성을 시사한다. 이는 AI 기반 진단 도구가 신속하고 정확한 진단을 통해 의료 기술을 향상시키는 데에 있어 중요한 역할을 할 수 있음을 나타낸다.

### Key words

Classification, Convolutional Neural Network (CNN), Model Comparison, Skin Disease

\* Department of ICT Convergence, Handong Global University, {pickpie, ParkJohn1027}@handong.ac.kr, yxp233@handong.edu, \*Dept. of Human Ecology & Technology, Handong Global University, myokpala@handong.ac.kr

※ This work was supported by the ESG, ICT, Start-up(EIS), Handong Global University, under Talent training project.

## 1. INTRODUCTION

Skin cancer is one of the most prevalent cancers, with more people diagnosed with it each year than all other cancers combined. [1] The incidence of melanoma of the skin among men has steadily increased from 39,080 cases in 2009 to 59,170 cases in 2024. [2] This trend highlights that, despite rapid advancements in medical and scientific fields, life-threatening diseases continue to rise. Traditionally, medical professionals have diagnosed skin diseases based on superficial symptoms unless deeper pathological conditions were suspected. For critical conditions such as skin cancer, the accuracy of early diagnosis plays a crucial role in determining treatment outcomes. It is essential to differentiate these conditions from others and perform histopathological examinations when necessary. To detect subtle early symptoms that are difficult to identify through visual inspection, more advanced AI-based diagnostic tools are required. With the advancement of deep learning technologies, research on AI-based classification and diagnosis of skin diseases continues to progress, and AI is expected to play a significant role in early diagnosis. However, deep learning models vary widely in their architectures, resulting in differences in data processing methods and performance. Consequently, the choice of model can significantly impact the outcomes when applied to medical tasks such as skin disease classification. In this context, our research aims to directly compare and analyze the effectiveness of various models in classifying skin diseases. We will evaluate 12 models, including traditional Convolutional Neural Network (CNN) architectures such as ResNet50 and InceptionV3, as well as other architectures like ConvNeXt and Xception, to determine which model is most suitable for the classification of skin cancer.

### MobileNetV2

Sandler et al. [3] introduce MobileNetV2, an

optimized neural network for mobile and resource-limited environments. The model enhances the original MobileNet with inverted residuals and linear bottlenecks, improving computational efficiency without compromising accuracy. By using depthwise separable convolutions, it reduces computation while maintaining high performance in tasks like image classification and object detection. Additionally, MobileNetV2 features SSDLite, a lightweight version of the SSD object detection framework, further optimizing it for mobile applications. This architecture strikes a balance between efficiency and expressiveness, making it well-suited for real-world scenarios on devices with limited resources.

### EfficientNetV2

Tan et al. [4] propose EfficientNetV2, a model that optimizes both training speed and parameter efficiency by combining Neural Architecture Search (NAS) with advanced scaling techniques. To address training bottlenecks present in the original EfficientNet, this model incorporates new operations like Fused-MBConv within its search space. Additionally, it introduces a progressive learning strategy, starting with small images and weak regularization, gradually increasing image size and regularization strength to enhance training efficiency. EfficientNetV2 demonstrates superior performance across datasets, including ImageNet, achieving up to 11 times faster training and 6.8 times fewer parameters compared to previous models. It achieves 87.3% top-1 accuracy after pretraining on ImageNet21k. The model outperforms traditional vision transformers (ViT) in terms of speed and efficiency, improving its adaptability across various hardware platforms.

### Xception

Chollet et al. [5] introduce Xception, an extension of the Inception architecture that maximizes performance and efficiency through depthwise

separable convolution and pointwise convolution. The pointwise convolution (1x1 convolution) optimizes inter-channel interaction, enhancing feature representation while maintaining computational efficiency. This design simplifies Inception's parallel paths into a single pathway, allowing the network to process cross-channel and spatial correlations independently. The architecture comprises 36 convolutional layers and residual connections, ensuring stable learning and high performance. Xception achieves similar or better performance than Inception V3 with fewer parameters and reduced computational cost, as demonstrated on datasets such as ImageNet.

## ConvNeXt

Liu et al. [6] propose ConvNeXt, a pure convolutional neural network (ConvNet) architecture that builds on improvements to the original ResNet while incorporating design elements from Vision Transformers (ViT). The model leverages optimizations such as depthwise convolution to maintain a simple structure while delivering high performance. ConvNeXt achieves comparable or superior results to Swin Transformer on datasets such as ImageNet-1K and COCO, with particular strengths in inference speed and memory efficiency.

Compared to ViT, ConvNeXt relies on traditional convolution operations, eliminating the need for complex global attention mechanisms. While ViT processes images as patches, ConvNeXt employs a sliding window approach for efficient feature extraction. Additionally, ConvNeXt is more memory-efficient and easier to train than ViT, with the ability to adapt to various resolutions and tasks. This study demonstrates that ConvNets can remain competitive without the use of Transformers, providing a powerful alternative that delivers both high efficiency and strong performance without complex architectures.

## II. METHODOLOGY

### Dataset Description

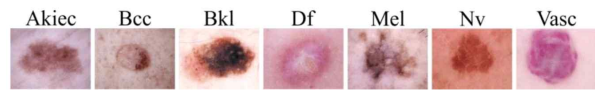


Fig. 1. Seven skin lesions classes in the HAM10000 dataset

The Human Against Machine with 10000 training Images (HAM10000) dataset, published by Philipp Tschandl, was used in this study. [7] This dataset contains 10,015 dermoscopic images representing seven types of pigmented skin lesions, including both benign and malignant conditions. The seven diseases are Actinic keratoses / Bowen's disease (akiec), which are precancerous lesions and early-stage skin cancers induced by ultraviolet radiation; Basal cell carcinoma (bcc), a type of skin cancer originating from basal cells with a low risk of metastasis; Benign keratosis-like lesions (bkl), which include seborrheic keratosis, solar lentigo, and other benign conditions; Dermatofibroma (df), a benign lesion that can occur as a result of minor trauma; Melanocytic nevi (nv), benign melanocytic moles; Melanoma (mel), a malignant melanocytic tumor that is curable if detected early; and Vascular lesions (vasc), which encompass various vascular-related conditions such as angiomas and abscesses. The images are provided in .jpg format. Collected over 20 years from diverse sources and patient populations, these images offer high diversity and reliability for application in real clinical settings.

Notably, more than half of the cases were confirmed by histopathological diagnosis, while the remaining cases were validated through follow-up observation, expert consensus, or advanced imaging techniques such as in-vivo microscopy. The dataset is accessible via the International Skin Imaging Collaboration (ISIC) archive and is widely used to evaluate and compare the performance of deep learning models for skin lesion classification. It is



particularly valuable for the development of AI tools that enable the early and accurate diagnosis of life-threatening skin cancers, such as melanoma.

### Data pre-processing

Before starting the training process, we resized the input images to 224x224 for models such as EfficientNetV2 and ConvNeXt. This size corresponds to the pretrained configuration of most models and allows for efficient training in terms of GPU memory usage. For Xception, the images were resized to 299x299. In the case of Inception, ResNet, and MobileNet, we did not specify a fixed input size; instead, the models were trained to process the input images according to their original dimensions.

### Evaluation Metrics

Our research focuses on comparing the effectiveness of different models in classifying skin diseases. To achieve this, we evaluate each model based on key performance metrics: accuracy, highest accuracy achieved during training, and loss. These metrics allow us to determine which model performs best and is most suitable for this classification task.

### Training Environment

The model training was conducted using Anaconda3, Jupyter, and TensorFlow. All models were trained for 50 epochs with a batch size of 16. Early stopping was applied during training to prevent overfitting.

### Experiment

The models compared using the HAM10000 dataset include ConvNeXtTiny, ConvNeXtSmall, ConvNeXtBase, EfficientNetV2S, EfficientNetV2M, EfficientNetV2L, InceptionV3, MobileNet, MobileNetV2, ResNet50, ResNet50V2, and Xception. The goal of this research is to identify which of these models is most suitable for skin disease classification.

All models used for training are available through Keras, and the performance metrics calculated post-training include accuracy, highest accuracy, and loss.

## III. RESULT

We conducted the training of all 12 models, and the results are presented in Table 1. Based on these results, we aim to present the findings for the two models with the highest accuracy.

Model	Accuracy	Highest Accuracy	Loss
ConvNeXtTiny	80.93	84.47	0.5277
ConvNeXtSmall	81.55	86.73	0.5097
ConvNeXtBase	65.67	66.85	1.1541
EfficientNetV2S	86.50	91.45	0.3834
EfficientNetV2M	74.05	78.64	0.7718
EfficientNetV2L	76.27	80.31	0.6624
InceptionV3	73.74	77.73	0.7409
MobileNet	66.99	68.56	0.9511
MobileNetV2	81.56	87.87	0.5328
ResNet50	72.87	75.66	0.7517
ResNet50V2	70.77	72.39	0.7742
Xception	84.72	89.76	0.4076

Fig. 2. MODEL PERFORMANCE METRICS

The EfficientNetV2S model demonstrated the best performance among the 12 models in classifying skin diseases. It achieved an accuracy of 86.50%, with the highest accuracy reaching 91.45%, indicating its effectiveness in precise classification. Additionally, it recorded the lowest loss at 0.3834, reflecting superior performance compared to other models.

The Xception model followed as the second-best performer, with an accuracy of 84.72% and a peak accuracy of 89.76%. Although slightly lower than EfficientNetV2S, it still achieved impressive results. Xception also had the second-lowest loss at 0.4076.

### Limitation

The first limitation is that the dataset size is not large enough. If each class in the dataset contained a sufficient number of samples during the training process, the model would be better equipped to

classify new lesion images accurately. In particular, research related to human diseases, such as skin lesions, demands extremely high accuracy, as even 5-10% misclassification could lead to medical errors. Therefore, a sufficiently large dataset is essential to ensure reliable results and prevent such risks. The second limitation is the lack of optimization efforts to enhance the models' accuracy. As seen in the results section, models like MobileNet exhibited significantly lower accuracy compared to others after training. With proper fine-tuning, these models could have demonstrated better performance. Even EfficientNetV2S and Xception, which achieved high accuracy, could potentially deliver even better results if fine-tuned to optimize them for the specific task.

## Conclusion

In conclusion, this study compared and analyzed 12 deep learning models using the HAM10000 dataset, with EfficientNetV2S and Xception demonstrating the highest performance in classifying skin diseases. Notably, EfficientNetV2S achieved the highest accuracy of 91.45% with a loss of 0.3834, confirming its effectiveness as the most suitable classification model.

These findings highlight the potential of deep learning models to excel in skin disease classification and make significant contributions to the development of AI-based diagnostic systems. Such tools could play a critical role in improving treatment outcomes by enabling the early detection of life-threatening diseases like skin cancer. Furthermore, the development of AI solutions compatible with mobile and embedded environments will empower general users to perform preliminary skin assessments with ease.

By improving accessibility to medical care and facilitating the early diagnosis of serious conditions, we believe this research will contribute to the advancement of AI-based diagnostic systems and the improvement of public health outcomes.

## Reference

- [1] A. Jabbar, R. Ullah, M. Ullah, and G. S. Choi, "Skin cancer classification using intelligent data augmentation based on transfer learning," *Medical Image Analysis*, vol. 74, p. 102294, 2021. [Online]. Available: <https://doi.org/10.1016/j.media.2021.102294>
- [2] Estimated number of new cancer cases among men in the U.S. in 2009 and 2024, by cancer type\* [Graph], American Cancer Society, January 17, 2024. [Online]. Available: <https://www.statista.com/statistics/268496/number-of-new-cancer-casesamong-men-in-the-us/>
- [3] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C., "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 4510-4520. Available: <https://arxiv.org/abs/1801.04381>.
- [4] Tan, M., and Le, Q. V., "EfficientNetV2: Smaller models and faster training," in *Proc. 38th Int. Conf. Machine Learning (ICML)*, 2021, pp. 10096-10106. Available: <https://arxiv.org/abs/2104.00298> <https://arxiv.org/abs/1512.00567>.
- [5] Chollet, F., "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1251-1258. Available: <https://arxiv.org/abs/1610.02357>.
- [6] Liu, Z., et al., "A ConvNet for the 2020s," *arXiv preprint arXiv:2201.03545*, 2022. Available: <https://arxiv.org/abs/2201.03545>.
- [7] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *\*Sci. Data\**, vol. 5, no. 1, p. 180161, Aug. 2018

# LLM과 언리얼 엔진 기반의 감정 표현 NPC 시스템 구현

소훈\* ,정현준\*\*

## Implementation of an Emotional Expression NPC System Based on LLM and Unreal Engine

Hoon So\* and Hyunjun Jung\*\*

### 요약

대다수의 게임 구조에서 NPC는 개발자가 설정한 상호작용 대사와 선택 옵션을 통해 플레이어와 상호작용하므로, 유저의 다양한 선호와 기대를 충분히 반영하기 어렵다. 이 논문에서는 이러한 문제를 해결하기 위해 Gemini API와 프롬프트 엔지니어링을 활용한 감정 표현 NPC 시스템을 제안한다. 제안된 시스템은 유저의 텍스트 입력을 실시간으로 감정 분석하여 NPC가 맥락에 맞는 맞춤형 응답을 생성하고, 이를 언리얼 엔진 5의 State Machine과 몽타주 시스템을 통해 5가지 기본 감정(기쁨, 슬픔, 화남, 두려움, 중립)을 자연스럽게 표현한다. 또한, 이 시스템은 NPC가 감정적으로 공감하면서도 실질적인 조언과 해결책을 제공할 수 있는 상호작용 방식을 가능하게 하여 게임 몰입도를 높인다. 향후 연구에서는 ChatGPT와 같은 다른 LLM 모델과의 비교를 통해 최적의 모델을 탐색하고, 더 복합적인 감정 표현과 멀티모달 상호작용이 가능한 시스템으로 확장할 계획이다.

### Abstract

In most game structures, NPCs interact with players through developer-defined dialogue and choice options, which makes it difficult to fully reflect the diverse preferences and expectations of users. This paper proposes an emotion-expressive NPC system using the Gemini API and prompt engineering to address this issue. The proposed system analyzes the user's text input in real-time to generate contextually appropriate responses for the NPC, which then expresses one of five basic emotions (joy, sadness, anger, fear, and neutrality) through Unreal Engine 5's State Machine and montage system. Additionally, this system allows the NPC to offer empathetic and constructive advice, enhancing engagement and immersion in the game. Future research will compare different LLM models, such as ChatGPT, to explore the optimal model and expand the system to enable more complex emotional expressions and multimodal interactions.

### Key words

unreal engine, npc, emotion analysis, llm, animation system, prompt engineering

\* 군산대학교 소프트웨어학부 학사과정, hoonso20@kunsan.ac.kr

\*\* 군산대학교 소프트웨어학부 교수(교신저자), junghj85@kunsan.ac.kr

※이 연구는 정부(과학기술정보통신부)의 재원으로 한국 연구재단의 지원을 받아 수행되고 있습니다.(No. NRF-2022R1G1A1008493).

※본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음(2023-0-00065)

## I. 서론

대다수의 게임 구조에서 NPC는 개발자가 설정해 둔 상호작용 대사와 선택 옵션을 통해 플레이어와 상호작용한다. 이러한 방식은 게임의 스토리와 캐릭터성을 강화하지만, 제한된 선택지로 인해 플레이어의 다양한 선호와 기대를 충분히 반영하기 어렵다 [1]. 이 논문에서는 생성형 AI 기반 NPC와 플레이어 간의 맞춤형 상호작용을 유도하는 방식을 제안한다.

Large Language Model(LLM) 중에서도 Google의 Gemini API를 선택한 이유는 방대한 데이터베이스와 통합된 정보 소스를 통해 더욱 정확한 감정 분석과 맥락에 맞는 응답을 생성할 수 있기 때문이다 [2]. 또한, 프롬프트 엔지니어링을 활용하여 사용자 감정을 보다 깊이 이해하고, NPC의 캐릭터 특성과 상황에 맞는 맞춤형 응답을 생성할 수 있다. 이 시스템은 Gemini API를 통해 유저의 텍스트 입력을 실시간으로 감정 분석하며, 분석 결과를 기반으로 NPC가 텍스트 응답과 연동된 감정 표현 애니메이션을 수행하도록 설계되었다.

감정 표현의 시각화를 위해 사용된 언리얼 엔진 5는 에픽 게임즈에서 개발한 강력한 그래픽 엔진으로, 실시간 렌더링과 고품질 애니메이션 처리가 가능하다. 특히 언리얼 엔진의 State Machine은 복잡한 상태 전환을 효율적으로 관리할 수 있으며, 몽타주 시스템은 여러 애니메이션을 자연스럽게 블렌딩하여 부드러운 동작 전환을 가능하게 한다[3]. 이러한 기능들을 활용하여 기쁨, 슬픔, 화남, 두려움, 중립의 5가지 기본 감정에 따른 자연스러운 동작 전환을 구현하였다. 또한 블루프린트 시스템을 통해 복잡한 감정 표현 로직을 직관적으로 구현할 수 있다.

이 논문의 주요 기여점은 다음과 같다. 첫째, Gemini API와 프롬프트 엔지니어링을 통해 NPC가 플레이어의 감정을 깊이 있게 이해하고 맥락에 맞는 응답을 생성할 수 있도록 설계하였다. 둘째, 감정 분석 결과를 시각적 피드백으로 확장하여 실시간 애니메이션을 구현하였다. 이는 특히 RPG나 어드벤처 게임에서 플레이어의 몰입도를 높이고 개인 맞춤형 게임 경험을 가능하게 한다.

## II. 관련 연구

### 2.1 LLM 기반의 NPC와의 상호작용 연구

최근 게임 산업에서는 LLM을 활용하여 NPC와의 상호작용을 더욱 자연스럽게 만들려는 연구가 활발히 진행되고 있다[4]. ChatGPT나 Gemini와 같은 생성형 AI 모델 기반의 NPC는 텍스트 입력을 통해 유저와 자연스러운 대화가 가능하며 문맥을 이해하고 적절한 반응을 생성할 수 있다. 최근 연구들은 LLM의 자연어 처리 능력과 게임 엔진의 애니메이션 시스템을 결합하여 NPC가 유저의 감정을 이해하고 이에 맞는 시각적 피드백을 제공하는 방향으로 발전하고 있다.

### 2.2 감정 표현을 위한 비언어적 애니메이션 시스템

2D 게임 캐릭터의 감정 표현을 위한 비언어적 커뮤니케이션 시스템에 대한 연구에서는 캐릭터의 감정 표현이 플레이어와의 유대감 형성과 게임 몰입도 향상에 중요한 역할을 한다고 설명한다. 이 논문에서는 이러한 연구를 바탕으로 언리얼 엔진 5의 State Machine과 몽타주 시스템을 활용하여 기쁨, 슬픔, 화남, 두려움, 중립의 5가지 기본 감정을 정의하고[5], 각 감정 상태가 자연스럽게 전환되도록 구현하였다. 특히 각 감정 표현 후 자동으로 중립 상태로 복귀하는 시스템을 설계하여 NPC의 안정적인 감정 표현이 가능하도록 하였다.

## III. 제안하는 감정 시각화 NPC 시스템

### 3.1 시스템 구조

이 논문에서 제안하는 감정 시각화 NPC 시스템은 언리얼 엔진 5를 기반으로 플레이어와 NPC 간의 자연스러운 상호작용을 구현한다. 시스템의 전체 구조는 사용자 상호작용을 담당하는 UI 시스템과 감정 분석 및 응답 생성을 수행하는 Gemini API 통

신 부분 그리고 NPC의 감정 표현을 제어하는 애니메이션 시스템으로 구성된다. 플레이어는 Chat Interface를 통해 NPC와 상호작용을 시작하며, 입력된 텍스트는 Gemini API로 전송되어 감정 분석과 응답 생성이 이루어진다. API 응답의 감정은 다섯 가지 기본 감정(기쁨 슬픔 화남 두려움 중립) 중 하나로 분류되며, 이는 State Machine을 통해 관리된다. 분류된 감정에 따라 해당 애니메이션이 몽타주 시스템을 통해 재생되며, 동시에 API의 텍스트 응답은 Chat Interface를 통해 플레이어에게 표시된다.

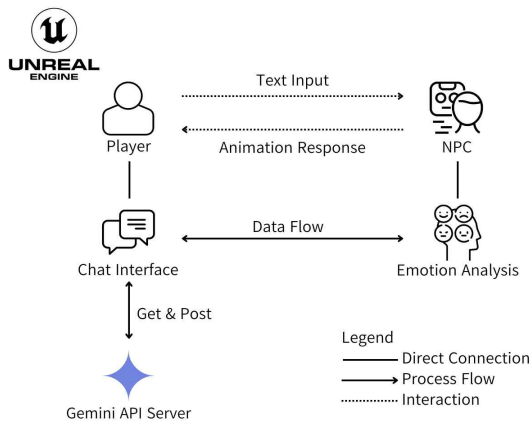


그림 1. 감정 표현 NPC 시스템 구조도  
Fig. 1. Emotion-Expressive NPC System Architecture

### 3.2 감정 분석 모듈

감정 분석 모듈은 플레이어의 입력에 대한 NPC의 감정적 반응을 결정하는 핵심 요소이다. 이 시스템에서는 Gemini API를 활용한 텍스트 분석과 언리얼 엔진의 State Machine을 통한 감정 상태 관리가 유기적으로 연동되어 있다.

#### 3.2.1 감정 상태 정의

표 1. 감정 상태 분류

Table. 1. Emotion State Classification

```

class NPCState
  IDLE    : Default system state
  NEUTRAL : Neutral emotion
  HAPPY   : Joy emotion
  SAD     : Sadness emotion
  ANGRY   : Anger emotion
  FEAR    : Fear emotion
    
```

#### 3.2.2 API 통신 및 감정 분석

Gemini API는 단순히 텍스트의 감정을 분류하는 것을 넘어, 상담 심리학적 접근 방식을 통해 사용자의 감정에 대한 깊은 이해를 제공한다. 예를 들어, 사용자가 분노를 표현할 경우 단순히 '화남' 상태로 분류하는 것이 아니라, 분노의 근본적인 원인을 파악하고 이에 대한 공감과 함께 건설적인 해결 방안을 제시한다. 이는 기존 NPC 시스템의 단순 응답 방식과 차별화되는 핵심 특징이다.

표 2. 감정 분석 및 상담을 위한 API 통신 알고리즘  
Table. 2. API Communication Algorithm for Emotion Analysis and Counseling

```

Function SendMessageToGeminiAPI(UserMessage) {
  Request->URL = APIServer
  Request->Method = "POST"
  Request->Header = "Content-Type: application/json"
  SystemPrompt = {
    role: "supportive_npc",
    instructions: "
      1. Emotion Classification:
      Select from: Joy, Sadness, Anger, Fear, Neutral
      2. Understanding Emotional Cause
      3. Show Empathy
      4. Propose Solutions
    "
  }
  JsonData {
    message = UserInput
    content = SystemPrompt
  }
  Request->Content = JsonData
  Request->OnComplete = ProcessResponse
  Request->Send()
}
    
```

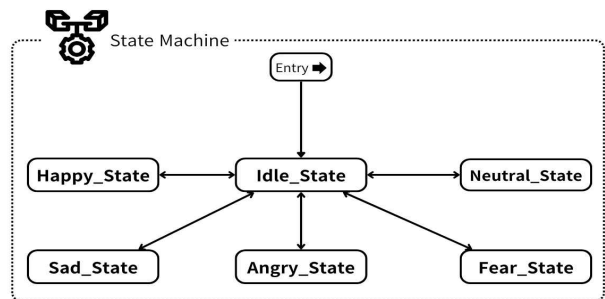


그림 2. NPC 감정 상태 머신  
Fig. 2. NPC Emotion State Machine

State Machine은 NPC의 감정 상태를 결정하고 전환하는 논리적 관리자 역할을 수행한다. 새로운 감

정이 감지되면 현재 상태에서의 전환 가능성을 평가하고, 전환이 결정되면 이를 애니메이션 제어 모듈에 전달한다. 애니메이션 제어 모듈은 이러한 감정 상태를 실제 시각적 표현으로 구현하는 실행자 역할을 하며, 감정 상태에 따라 적절한 애니메이션을 재생한다. 이처럼 State Machine과 애니메이션 제어 모듈이 분리되어 동작함으로써, 감정 상태 관리와 시각적 표현이 독립적으로 최적화될 수 있다.

### 3.3 애니메이션 제어 모듈

애니메이션 제어 모듈은 언리얼 엔진의 몽타주 시스템을 활용하여 사전 정의된 6가지 애니메이션을 관리한다. 각 감정 상태에 맞는 애니메이션과 블렌딩 시간을 적용하여 애니메이션 간의 자연스러운 전환을 구현한다. 이를 통해 급격한 동작 변화를 방지하며 현재 재생 중인 애니메이션에서 다른 감정 상태로의 매끄러운 전환을 보장한다.

표 3. 감정 애니메이션 제어 알고리즘  
Table. 3. Emotion Animation Control Algorithm

```

Function SetEmotionState(NewEmotion) {
    if (AnimInstance valid)
        SelectMontage {
            IDLE      : IdleMontage
            NEUTRAL   : NeutralMontage
            HAPPY     : HappyMontage
            SAD       : SadMontage
            ANGRY    : AngryMontage
            FEAR     : FearMontage
        }
    if (Montage)
        PlayAnimation()
        ReturnToIdle()
}
    
```

## IV. 실험 결과 및 평가

### 4.1 실험 환경

이 논문에서 구현한 LLM 기반 감정 표현 NPC 시스템의 개발 환경은 표 4과 같다. 언리얼 엔진 5.1.1 버전을 사용했으며, Google Gemini API를 활용하여 텍스트 분석 및 응답 생성을 구현했다. 캐릭터 모델은 MC\_Sample Character Asset을 사용했다.

표 4. 소프트웨어 환경  
Table 4. Software environment

Software	Version
Unreal Engine	5.1.1
API	Google Gemini API
3D Model	MC_Sample Character Asset

### 4.2 실험 결과

이 논문에서 구현한 감정 표현 NPC 시스템의 효과성을 평가하기 위해 두 가지 방식으로 실험을 진행하였다. 단순 텍스트 입력에 대한 Gemini API의 기본 응답을 테스트한 결과, NPC는 사용자의 메시지에 대해 일반적인 수준의 응답을 생성했다. 이 경우 사용자의 감정을 깊이 있게 이해하거나 NPC의 특성을 반영한 응답을 생성하는 데 한계를 보였다. 감정 이해 기반 프롬프트 테스트 결과 NPC가 사용자의 감정을 이해하고 자신의 역할에 맞는 응답을 생성하도록 프롬프트를 설계한 결과, 응답의 품질이 기본 응답 대비 약 32% 향상되었다.



그림 3. NPC의 Happy 감정 상태 전환 과정  
Fig. 3. NPC's Transition Process to Happy Emotional State

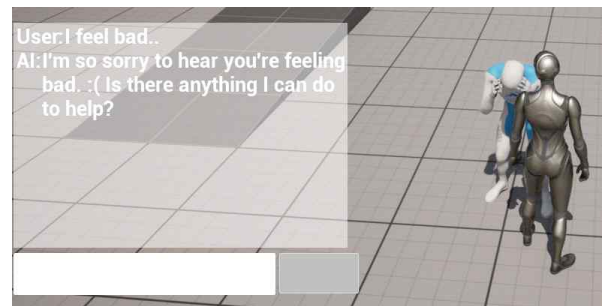


그림 4. NPC의 Sad 감정 상태 전환 과정  
Fig. 4. NPC's Transition Process to Sad Emotional State

시스템 성능 측면에서는 API 응답 시간이 평균 0.78초로 실시간 대화에 적합했으며, 0.25초의 블렌딩 시간 적용으로 부드러운 동작 전환을 구현하였다.

표 5. 감정 이해 성능 평가  
Table 5. NPC System Performance Analysis Results

Feature	Basic	Prompt
Context Understanding	24.04%	56.92%
Response Quality	19.13%	63.86%

표 6. NPC 시스템 성능 분석 결과  
Table 6. NPC System Performance Analysis Results

Feature	Specification
API Response Time	0.78s
Blend Time	0.25s

## V. 결론

이 논문에서는 LLM 기반의 실시간 감정 분석과 언리얼 엔진의 애니메이션 시스템을 결합하여 사용자와 자연스럽게 상호작용하는 NPC 시스템을 구현하였다. Gemini API를 활용하여 사용자의 자유로운 텍스트 입력에 대해 맥락에 맞는 응답을 생성하고, 이를 시각적 감정 표현과 연동하여 더욱 몰입감 있는 상호작용을 가능하게 하였다. 특히 감정 이해를 위한 프롬프트 엔지니어링을 통해 NPC의 응답 품질을 크게 향상시켰다. 실험 결과, 맥락 이해도가 24.04%에서 56.92%로, 응답 품질이 19.13%에서 63.86%로 향상되었으며, NPC 캐릭터성 유지와 감정 인식의 정확도도 크게 개선되었다. 또한 구현된 시스템은 5가지 기본 감정을 State Machine으로 관리하고 몽타주 시스템을 통해 0.25초의 블렌딩 시간을 적용하여 자연스러운 애니메이션 전환을 구현함으로써, 텍스트 기반 대화의 한계를 넘어선 풍부한 비언어적 커뮤니케이션을 제공한다.

향후 연구에서는 현재의 5가지 기본 감정을 넘어 더 복잡한 감정 표현이 가능한 시스템으로 확장하고, NPC의 개성과 역할에 따라 차별화된 응답을 생성할 수 있도록 프롬프트를 발전시킬 계획이다. 또한, ChatGPT와 같은 다른 LLM 모델을 비교 분석하여 감정 분석과 맞춤형 응답 생성에 가장 적합한

모델을 선정하고자 한다. 이를 통해 다양한 LLM의 특성을 반영한 맞춤형 AI 시스템을 구현할 수 있는 가능성을 탐색할 예정이다. 더불어 음성 입력과 표정 인식을 추가하여 멀티모달 상호작용으로 확장하는 것을 고려하고 있다. 이 연구는 게임 개발에서 LLM과 같은 최신 AI 기술을 활용하여 더욱 몰입감 있는 게임 경험을 제공할 수 있는 가능성을 실험을 통해 입증하였다.

## 참고 문헌

- [1] Cho, D. H., Oh, S. J., Sung, M. Y., & Jun, K. K. (2007). NPC Control by Hybrid Architecture of Finite State Machine and Inference Engine?. 한국 HCI 학회: 학술대회논문집, 168-173.
- [2] N. Rane, S. . Choudhary, and J. Rane, "Gemini v ersus ChatGPT: applications, performance, architecture, capabilities, and implementation", J. Appl. Artif. Intell., vol. 5, no. 1, pp. 69-93, Mar. 2024.
- [3] Lee, Myoun-Jae, "A Proposal on Game Engine Behavior Tree," Journal of Digital Convergence, vol. 14, no. 8, pp. 415-421, Aug. 2016.
- [4] [3] Byong-Kwon, Lee, Moo-Young, Kim, Ye-chan, Park, and San-Hwa Lee, "A Study on Methods of Utilizing Generative AI in Game Engines", Proceedings of the Korean Society of Computer Information Conference , 제주, pp. 715-717, 2024-07-11.
- [4] Chang, Edward Y. "Behavioral Emotion Analysis Model for Large Language Models." 2024 IEEE 7th International Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, 2024.
- [5] Jae Hun Bang and Sungyoung Lee, "Call Speech Emotion Recognition for Emotion based Services", Journal of KISS : Software and Applications, Vol. 41, No. 3, pp. 208-213, 2014.

# 효율적인 항만 운영 시스템을 위한 기계학습 기반 항만 내 선박 대기시간 예측

최민화\*, 윤웅창\*\*

## Machine learning-based port ship waiting time prediction for efficient port operation system

Min-hwa Choi\* and Woongchang Yoon\*\*

### 요 약

항만 내 선박의 체선 현상은 선박의 채산성과 국가 물류 경쟁력에 중요한 문제이며, 이를 해결하기 위한 항만 시설 확충은 오랜 시간과 비용이 소요된다. 본 논문에서는 효율적인 항만 운영 시스템을 위해 기계학습을 활용하여 항만 내 선박 대기시간 예측 알고리즘을 제안한다. 정박지와 접안지 사이의 거리, 선종, 실제 정박 시각, 도착 항의 소속 국가, 재화중량톤수 등 7가지 중요 요인을 기반으로 다양한 기계학습 모델을 사용하여 선박 대기시간을 예측 및 비교하였다. 본 연구를 통해 항만 내 선박 대기시간 예측 모델을 개발함으로써, 항만 운영 시스템이 효율적으로 개선되고 항만의 서비스 만족도를 향상시킬 수 있을 것으로 예상된다.

### Abstract

The phenomenon of ship congestion in ports is a significant issue for the profitability of ships and the national logistics competitiveness. Expanding port facilities to address this problem is time-consuming and costly. In this paper, we propose a ship waiting time prediction algorithm within ports using machine learning for an efficient port operation system. We predicted and compared ship waiting times using various machine learning models based on ten different factors, including the distance between berth and berth, ship type, actual berth time, country of arrival at port, goods weight tonnage. Through this study, it is anticipated that by developing a ship waiting time prediction model, the port operation system can be improved efficiently, thereby enhancing the satisfaction with port services.

### Key words

Ship waiting time, Prediction, Regression, Machine Learning, Hyper-parameter Tuning

## 1. 서 론

선박의 채산성은 항만 내에서의 정체 현상 없이 신속하게 입출항하여 예정된 항해를 완료함으로써 확보된다. 선박의 항만 내에서의 정체되는 체선 현

상은 선박의 채산성과 직결되는 문제이다 [1, 2]. 국가의 물류 경쟁력을 확보하기 위해서는 체선 현상 해결이 필수적이나, 이를 개선하기 위한 항만 시설 확충은 시간이 오래 걸리고 비용이 많이 든다 [3]. 이전 연구에서는 부산 신항의 컨테이너 반출입

\* 경상국립대학교 컴퓨터공학부, 220161105@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과/바이오의료빅데이터학과, neobits@gnu.ac.kr



정보를 기반으로 심층학습 모델인 RNN, LSTM 및 GRU를 활용하여 반출입 차량의 평균 대기시간을 예측하는 모형을 제시하였다 [4]. 해당 연구에서는 부산 신항의 실제 컨테이너 반출입 데이터를 활용하여 RNN, LSTM, GRU 모델의 성능을 검증하였고, 반출입 차량의 대기시간 예측에 효과적임을 보였다.

다른 연구에서는 웨이퍼 로트의 대기시간을 예측하기 위해 심층학습 기반의 예측 모델을 제안하였다 [5]. 이 모델은 공정 시작 시점의 제공 정보 (work-in-process)와 디스패칭 룰 (dispatching rule)을 기반으로 대기시간에 영향을 미치는 예측 변수를 최적화하였다. 유전 알고리즘을 통해 최적의 조합이 도출되었다.

본 논문은 항만 내 선박의 대기시간을 예측하는 알고리즘을 제안하여 항만 내 체선 현상을 더 효율적으로 개선하는 데 도움이 되고자 하였다. 이 알고리즘은 정박지와 접안지 사이의 거리, 선종, 실제 정박 시각, 도착 항의 소속 국가, 재화중량톤수 등 7가지의 여러 요인들을 입력값으로 받아 학습된 기계학습 모델을 통해 예측된 대기시간 정보를 제공하였다. 다양한 기계학습 모델을 평가 지표를 통해 성능을 비교함으로써 어떤 모델이 선박 대기시간 예측에 가장 적합한지 실험하였다. 또한, 선별된 모델의 하이퍼파라미터를 조정하여 모델 최적화 및 성능을 향상시키는데 중점을 두었다.

## II. 실험 방법

### 2.1 실험의 개요

본 논문에서는 항만에서 선박의 대기시간을 예측하기 위해 Linear Regression (LR), Random Forest Regressor (RFR), XGBoost Regressor (XGBR), LightGBM Regressor (LGBR), CatBoost Regressor (CBR) 등 다양한 모델들을 활용하여 비교실험을 진행하였다. 정박지와 접안지 사이의 거리, 선종, 실제 정박 시각, 도착 항의 소속 국가, 재화중량톤수 등 7가지의 요인들을 입력 값으로 받아 학습된 기계학습 모델을 통해 예측된 대기시간 정보를 제공한다. 다음은 제안하는 알고리즘의 흐름이다 (그림 1).

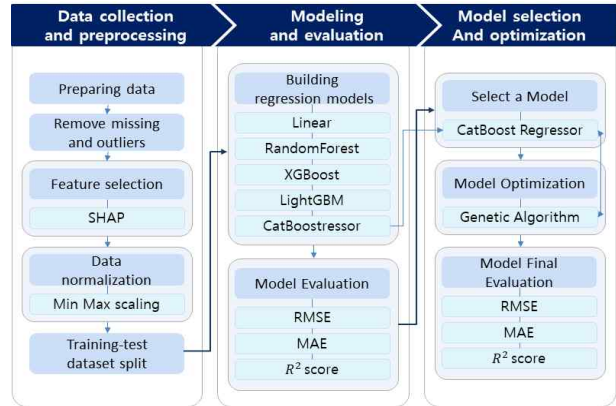


그림 1. 항구 내 선박 대기시간 예측 알고리즘 흐름도  
Fig. 1. Flowchart of the algorithm for predicting ship waiting time in port

### 2.2 데이터 세트 전처리와 feature selection

본 논문에서는 AI Challenge인 DACON에서 수집한, 총 391,939개의 개별 운행 기록에 관한 정보를 수집한 데이터 세트를 사용하였다 [6]. 해당 데이터 세트는 선박의 폭, 도착 항의 소속 국가, 항구명, 선박의 국적, 소유주, 정박지와 접안지 사이의 거리 등 총 23개의 항목 정보를 포함하고 있다.

전처리 과정을 통해서 164,631개의 결측치와 중복 항목 등을 제거하였으며, 인코딩을 통해서 범주형 자료를 연속형 자료로 변환하였다. 대기시간 열에 대하여 IQR (Inter Quantile Range)을 계산하여 이상치를 정의하고 총 105,906개의 이상치를 제거하였으며, Min-max 정규화를 수행하였다.

Feature Importance는 SHAP (Shapley Additive ex Planations) 기법을 사용하여 상위 7개의 요인들 (정박지와 접안지 사이의 거리, 선종, 실제 정박 시각, 도착 항의 소속 국가, 재화중량톤수, 접안지 폴리곤 크기, 도착 항의 항구명)을 독립변수로 선택하였다(표 1).

표 1. SHAP값 기준 상위 7개의 주요 요인들

Table 1. Top 7 key features based on SHAP values

순위	Feature names	SHAP
1	dist (정박지와 접안지 사이 거리)	14.035
2	ship_type_category (선종)	2.818
3	ata (실제 정박 시각)	1.143
4	ari_co (도착항의 소속국가)	0.882
5	dead weight (재화중량톤수)	0.849
6	port size (접안지 폴리곤 크기)	0.783
7	ari_po (도착항의 항구명)	0.728

이후, 범주형 변수를 제외한 이들 상위 7개 요인에 대한 기초 통계를 상세히 분석하여 요인별 데이터 분포와 특성을 명확히 파악할 수 있도록 정리하였다 (표 2).

표 2. 선택된 주요 요인들의 분포  
Table 2. Distribution for selected key features

Statistic	dist (km)	dead weight (ton)	port size (km <sup>2</sup> )	ata (minute)
Min	0.000	100.000	0.000	0.000
Max	199.980	404000.000	0.002	1439.000
Mean	19.328	63577.506	0.001	689.169
Median	6.929	47400.000	0.001	659.000
Std	30.930	63515.531	0.001	421.338

### 2.3 데이터 세트의 구분과 모델 성능 평가

Training 데이터 세트와 test 데이터 세트는 전체 121,402건의 데이터를 training 데이터 97,121건과 test 데이터 24,281건으로 각각 구분하여 8:2의 비율로 분리하여 사용하였다. 모델의 일반화 성능을 더 정확히 평가하고, 과적합 (Over-fitting)을 방지하기 위해 5-fold cross validation을 사용하여 평가를 진행하였다. 모델의 성능 평가에는 RMSE와 MAE, R<sup>2</sup> 값을 사용하였다.

## III. 결 과

### 3.1 다양한 기계학습 모델을 활용한 항만대기 시간 예측 및 비교

우리는 먼저 LR, RFR, XGBR, LGBR, CBR 총 5가지 모델의 성능 평가를 RMSE, MAE, R<sup>2</sup> 값의 비교 및 분석을 통해서 수행하였다 (그림 2).

모델 중 RFR의 Training과 Test의 RMSE와 MSE, R<sup>2</sup> 값의 차이는 각각 10.318, 5.589, 0.426로 학습이 제대로 이루어지지 않았음을 확인하였다 (그림2와 표 3). RFR을 제외한 나머지 4가지 모델 (LR, XGBR, LGBR, CBR)은 Training과 Test의 RMSE, MAE 값의 차이는 약 1 정도의 차이가 있었고, R<sup>2</sup> 값의 차이는 0.009, 0.082, 0.027, 0.048이었다. 그중 LGBR의 Training과 Test에서의 RMSE, MAE 값의

차이가 0.276, 0.132로 가장 낮았고, 학습이 전반적으로 안정적이고 잘 이루어지고 있음을 확인하였다 (그림2와 표3).

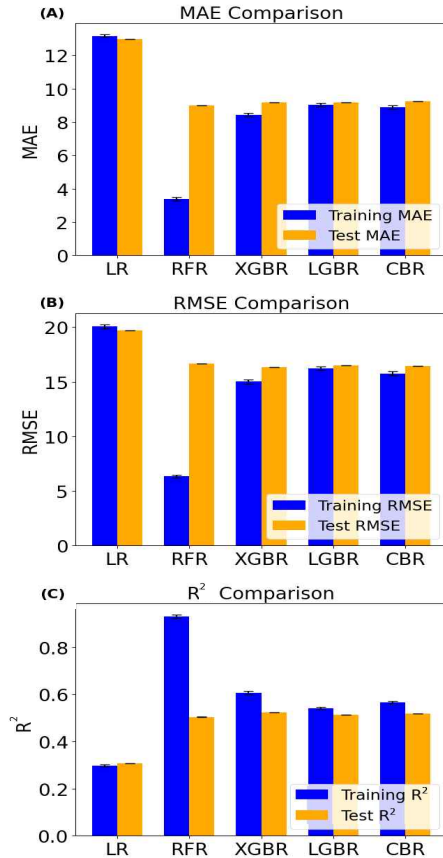


그림 2. 회귀 모델별 Training 및 Test 데이터 세트의 성능평가 결과

Fig. 2. Performance evaluation results for Training and Test datasets by regression model

표 3. 회귀 모델별 Training 및 Test에서의 성능평가 결과  
Table 3. Performance evaluation results for training and test by regression model

Model	Data set	RMSE	MAE	R <sup>2</sup>
Linear regression	Training	20.049	13.165	0.296
	Test	19.703	12.988	0.305
RandomForest regressor	Training	6.319	3.395	0.930
	Test	16.637	8.984	0.504
XGBoost regressor	Training	15.012	8.444	0.605
	Test	16.321	9.178	0.523
LightGBM regressor	Training	16.213	9.031	0.540
	Test	16.489	9.163	0.513
CatBoost regressor	Training	15.758	8.892	0.565
	Test	16.432	9.247	0.517

추가로 LR, RFR, XGBR, LGBR, CBR 총 5가지 모델의 RMSE 값의 분포를 비교 및 분석하였다 (그림 3).

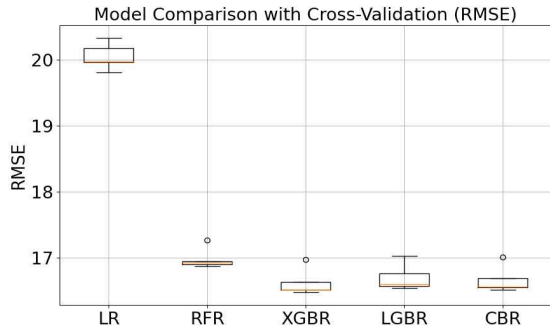


그림 3. 5가지 회귀 모델별 RMSE 결과  
Fig. 3. RMSE results by five regression models

위의 결과를 확인할 때, XGBR의 RMSE가 16.619로 가장 성능이 높게 평가되었지만 LGBR과 LR을 제외한 모든 모델의 RMSE 값에 이상치가 존재함을 확인할 수 있었다 (그림 3). 때문에, RMSE 값이 16.701로 LR보다 다음으로 성능이 좋은 LGBR을 선택하였다. 표준편차는 RFR을 제외한 4가지 모델 (LR, XGBR, LGBR, CBR) 모두 약 0.18로 비슷했다 (표 4).

표 4. 회귀 모델별 test의 RMSE 평균 및 표준편차  
Table 4. RMSE mean and standard deviation of test by regression model

Model	RMSE 평균	표준편차
Linear regression	20.049	0.182
RandomForest regressor	16.976	0.147
XGBoost regressor	16.619	0.182
LightGBM regressor	16.701	0.180
CatBoost regressor	16.665	0.181

### 3.2 하이퍼파라미터 튜닝

다양한 최적화 기법 중에서 좋은 성능을 보장하는 유전 알고리즘 (GA, Genetic algorithm)을 사용하여 LGBR의 하이퍼파라미터를 튜닝하였다. 트리 최대 리프 수, 트리 최대 깊이, 학습률, 트리 개수, 자식 노드의 최소 가중치 합, 각 트리 학습 시 사용하는 데이터 비율 및 컬럼 비율을 튜닝하여 최적화된 값을 확인하였다 (표 5).

표 5. LightGBM Regressor의 튜닝된 하이퍼파라미터  
Table 5. Tuned hyperparameters of the LightGBM Regressor

Hyper Parameter	Before	After
num_leaves	31	84.121
max_depth	-1	91
learning_rate	0.1	0.072
n_estimators	100	337
min_child_weight	0.001	7.706
subsample	1.0	0.907
colsample_bytree	1.0	0.640

LGBR 모델의 하이퍼파라미터 튜닝 전후의 성능을 비교한 결과, Training에서 RMSE는 16.213에서 14.760로, MAE는 9.031에서 8.607로 감소하였으며,  $R^2$  값은 0.540에서 0.619로 증가하였다. Test 단계에서도 RMSE가 16.489에서 16.307로 감소하고,  $R^2$  값은 0.513에서 0.524로 증가한 것을 확인할 수 있었다 (표 6). 또한, 튜닝 후의 Training 데이터 세트와 Test 데이터 세트에서의 RMSE와 MAE 결과를 비교해 보았을 때, 큰 차이를 보이지 않아서 Over-fitting은 발생하지 않았음을 확인하였다.

표 6. LightGBM Regressor 하이퍼파라미터 튜닝 전후 성능 비교  
Table 6. Performance comparison of LightGBM Regressor before and after hyper-parameter tuning

Parameter Tuning	Data set	RMSE	MAE	$R^2$
Before	Training	16.213	9.031	0.540
	Test	16.489	9.163	0.513
After	Training	14.760	8.607	0.619
	Test	16.307	9.455	0.524

## IV. 결 론

정박 대기시간은 항만의 서비스 품질을 결정하는 중요한 요소이며, 우리는 항만의 체선 현상을 개선하고자 항만 내 선박의 대기시간을 예측하는 알고리즘을 제안하였다. 본 연구에서는 정박 대기시간 예측의 정밀도 향상을 목표로 모델들을 비교한 후 LGBR 모델을 선택하였으며, 최적화 과정을 거쳐 예측 모델의 성능을 개선하였다.

이후에는 해당 연구를 실무에 더욱 효과적으로 적용할 수 있도록 입력값을 받아 예측 결과를 제공

하는 웹 기반 서비스를 개발할 계획이다. 이를 통해 사용자들이 실시간으로 대기시간을 예측하고 체선에 대한 대응책을 즉각 마련할 수 있는 환경을 제공할 것이다. 이러한 시스템은 항만 관리자와 운항자 간의 소통을 원활하게 하여 전체적인 운영 효율성을 높이고, 항만 서비스의 품질을 한층 더 강화할 수 있을 것이라 생각된다.

결과적으로 본 연구를 통해 예측된 선박 대기시간을 기준으로 하여 체선 현상에 미리 대비할 수 있으며, 선박 스케줄링을 효율적으로 개선할 수 있을 것이다. 또한, 본 연구를 통한 예측 결과는 항만 관리자 뿐만 아니라 각 선박 운항자에게도 제공되어 항만의 서비스 만족도를 향상시킬 수 있을 것으로 기대된다.

[6] DACON, (2024), [train.csv], Retrieved from <https://dacon.io/competitions/official/236158/data>.

## 참 고 문 헌

- [1] Kim, JH, Kim, SW, Nam, BA, Choi, JW, Jang, DW, Kim, KH and Hong, "Arrival Pattern Analysis of Gate In-Out Trucks in Container Terminal", Korean Institute of Industrial Engineers, 2015 Fall Conference of Korean Institute of Industrial Engineers, pp. 1097-1102, 2015.
- [2] Vasiliasukas A.V., Barysiene J., Dragan D., "An economic evaluation model of the logistic system based on container transportation", Transport, 23 (4), pp. 311-315, 2008.
- [3] K. H. Sa and S. B. Choi, "Empirical Study on the Congestion of Bulk Ships- Focus on the Queuing Time for Berthing", Shipping and Logistics Research vol 25, no. 2, pp. 263-284, August 2009.
- [4] Y. I. Kim, "A Study on the Prediction of Gate In-Out Truck Waiting Time in the Container Terminal", J . Navig. P ort Res. vol. 46, no.4, pp. 344-350, August 2020.
- [5] S. C. Park, "Queue-Time Prediction for a Wafer Lot with Time-Constraints", Korean Journal of Computational Design and Engineering, vol. 25, no. 4, pp. 343-349, December 2020.

# 전이학습 모델에 대한 적대적 공격

정지홍\*, 유동옥\*, 김성석\*\*, 양순옥\*\*\*

## Adversarial Attacks on Transfer Learning Model

Jihong Jeong\*, Dongok Yu\*, Sungsook Kim\*\*, and Sun Ok Yang\*\*\*

### 요 약

인공지능 기술이 발전함에 따라 보안에 대한 중요성도 함께 높아지고 있다. 인공지능 보안에서 활발히 사용하는 적대적 공격 기법인 FGSM과 PGD는 주로 이미지 분야에서 활발히 연구되었다. 본 연구에서는 이러한 기법을 이미지가 아닌 시계열 데이터에 적용하고자 한다. 즉 미세먼지 관련 시계열 데이터를 기반으로 예측 모델에 보안 기법을 적용해본다. 이러한 데이터는 적대적 훈련을 통한 강건한 모델을 만들기 위한 기반이 된다. 이를 위하여 FGSM과 PGD를 이용하여 전이학습 후 실제 예측값과 비교하여 성능을 평가하였다.

### Abstract

As artificial intelligence technology advances, the importance of security is also increasingly emphasized. FGSM and PGD, which are adversarial attack techniques actively used in AI security, have been extensively studied in the field of image processing. In this study, we aim to apply these techniques to time series data instead of images. The generated data will then be applied to a predictive model for transfer learning. Such data serves as the foundation for building robust models through adversarial training. To this end, we employed FGSM and PGD for transfer learning and evaluated performance by comparing the predicted values with the actual results.

### Key words

LSTM , prediction model , Indoor/Outdoor Air Pollution , Adversarial Attack , FGSM , PGD

## 1. 서 론

현재 우리 일상에서 여러 인공지능 기반 서비스가 널리 활용되고 있다. 이러한 인공지능 기술은 주로 데이터로 학습하면서 특정 서비스에 적합한 모델을 생성한다. 만약 학습 데이터가 부정확하거나 결함이 있으면 올바른 서비스가 제공되지 않을 수 있다. 예를 들면, 학습과정에서 데이터 보안에 문제가 발생하거나 악의적인 공격을 받을 경우에는 서비스 품질에 큰 피해가 갈 것이다. 따라서 인공지능

관련 기술의 적용범위가 확대됨에 따라 인공지능의 보안의 중요성도 함께 커지고 있다.

인공지능 기술을 공격하는 방법 중 적대적 공격이 주요한 방식이며, 그 중에서 적대적 예제를 생성하는 방법이 대표적이다. 이와 관련하여 초기에는 주로 이미지 데이터를 다루는 분야에서 활발하게 연구되었다. 컴퓨터 비전 분야의 경우, 입력 데이터 값이 아주 미세하게 변조되더라도 모델의 분류 정확도가 떨어지게 된다[1, 2].

이에 반해, 시계열 데이터를 다루는 분야는 보안

\* 서경대학교 소프트웨어학과, [jeongjihong0303@naver.com, dongok218@skuniv.ac.kr]

\*\* 서경대학교 소프트웨어학과, sskim03@skuniv.ac.kr, 교신저자,

\*\*\* 가천대학교 교양학부, sovang@gachon.ac.kr

과 관련된 연구가 상대적으로 활발하지 않다. 그러므로 본 연구에서는 시계열 데이터에 대해 동일한 적대적 공격 실험을 하고자 한다. 결과적으로는 적대적 공격에도 정확한 예측을 수행하는 강건한 모델 개발을 최종 목표로 한다.

## II. 예측 모델에 대한 적대적 공격

### 2-1 시계열 데이터 기반 예제 모델

[3] 연구에서 학교의 실외 미세먼지를 예측하는 모델을 이용하여, 데이터가 적게 가지고 있는 새로운 학교에 전이학습을 시킬 수 있음을 보였다. 이러한 가정 아래에 전이학습을 시키는 과정에서 여러 적대적 예제를 생성한 뒤 시계열을 다루는 모델에 적용한다. 적용 후 결과들을 평가지표인 RMSE(Root Mean Square Error, 평균 제곱근 오차)를 이용하여 비교하려고 한다.

사용하는 시계열 데이터는 시간 단위이며, 입력값으로 주는 인자는 Air Korea와 기상청의 데이터를 사용하였다. 예측하는 값으로는 학교의 실외 미세먼지 값이다.

표 1. 데이터 세트  
Table 1. Data sets

학교 분류	데이터 수	입력인자 수	예측 값
기존 학교	9,930	10	pm10
새로운 학교	96	10	pm10

적대적 공격이란 인공지능 모델을 목표로 하여 공격하는 기법이다. 적대적 예제 공격이란 모델에게 주어지는 데이터에 미세한 변화를 주어 잘못된 의사결정을 내리게 하는 기법이다. 이 논문에서는 선행된 연구에서 개발한 학교 미세먼지 예측 모델을 목표로 설정했으며, 적대적 예제 공격기법 중에서 손실함수의 기울기를 이용하는 FGSM과 PGD 기법을 적용하였다[4].

### 2-2 FGSM을 이용한 적대적 예제 생성

개발한 모델에서 학교와 가장 가까운 위치의 AirKorea 측정소 및 기상청 측정소의 데이터 집합

( $x$ )을 입력값으로 사용하였다. 그에 대한 예측값( $y$ )은 학교 운동장(실외)의 미세먼지 농도 PM10으로 설정하였다. 손실함수에서 기울기 변화의 부호를 구한다( $sign(\nabla_x J(\theta, x, y))$ ). 그리고 공격 강도를 조절하는  $\epsilon$ 을 곱하여  $x$ 에 더한다. 위와 같은 과정을 거치면 원래 예측과 다른 값인  $x'$ 을 얻게 된다. 식으로 다시 표현하면 아래와 같다.

$$x' = x + \epsilon \cdot sign(\nabla_x J(\theta, x, y)) \quad (1)$$

실험을 위하여  $\epsilon$ 을 각각 0.5, 0.1, 0.05, 0.01, 0.005, 0.001로 설정해 적대적 예제를 생성하였다.

### 2-3 PGD를 이용한 적대적 예제 생성

PGD는 FGSM과 같이 손실함수의 기울기를 이용한다. 하지만 FGSM은 기울기 계산이 한 번만 이루어진다면, PGD는 여러 번 계산이 이루어진다.

$$x_{t+1}^{adv} = x_t^{adv} + \alpha \cdot sign(\nabla_x L(f_\theta(x_t^{adv}), y)) \quad (2)$$

계산된 기울기 값이 일정 범위를 벗어나지 않게 조절하기 위해서 투영함수를 사용한다. 이로 인해 FGSM보다 정교한 예제를 생성할 수 있다(식 3).

$$x_{t+1}^{adv} = Proj_{B_\epsilon}(x)(x_{t+1}^{adv}) \quad (3)$$

위의 과정을  $t$ 번 반복을 하게 된다(식 2). FGSM과 마찬가지로  $\epsilon$ 값으로 0.5, 0.1, 0.05, 0.01, 0.005, 0.001를 설정하였다. 기울기를 업데이트하는 크기인  $\alpha$ 는 0.1, 0.01, 0.001로 설정하여 생성하였다.

### 2-4 적대적 예제 적용 및 결과

A 학교를 대상으로 훈련한 모델을 B 학교에 전이학습을 수행할 것이다. B 학교는 A 학교보다 상대적으로 데이터가 적은 학교이다. B 학교 데이터를 훈련 및 테스트를 위해서 75:25로 나누었다.

훈련데이터를 대상으로 적대적 예제를 생성하였

다. 그리고 테스트 데이터로 예측을 수행하였다. 예측 결과의 평가 지표로 RMSE를 사용했다.

전이학습을 시킨 모델의 예측 결과를  $\epsilon$ 에 따른 FGSM의 결과,  $\alpha$ 와  $\epsilon$ 에 따른 PGD 결과를 비교할 것이다.

FGSM의 경우에는  $\epsilon$ 을 0.5로 설정한 경우 RMSE 값이 50.12가 나왔다. 이에 비해  $\epsilon$ 을 0.001과 0.005로 설정한 한 경우 각각 4.82, 5.89가 나왔다. 이는 원본 데이터로 예측을 수행한 값과 유사하게 측정되었다. 즉, 낮은  $\epsilon$ 값을 준 경우 적대적 예제 공격의 효과가 없음을 보여준다(그림 1 참고( $\epsilon$ 값은 순서대로 0.5, 0.1, 0.05, 0.01, 0.005, 0.001)).



그림 1. 원본과 FGSM에 대한 결과  
Fig. 1. Results for the original and FGSM

PGD의 경우,  $\alpha$ 값과  $\epsilon$ 값을 각각 0.1, 0.5로 준 경우 RMSE값이 70.51로 가장 높게 측정되었다. 하지만 같은  $\alpha$ 값을 주어도  $\epsilon$ 을 준 경우에는 실제 예측값과 유사하게 측정되었다. 이와 반대되게  $\alpha$ 값을 낮게,  $\epsilon$ 값을 높게 준 경우에도 실제 예측값과 유사하게 측정되었다(그림 2참고).

즉,  $\alpha$ 과  $\epsilon$ 값을 크기 설정할수록 RMSE가 높게 나왔음을 보여준다. 두 개의 값중 하나라도 낮게 주는 경우 RMSE값도 낮게 나옴을 보여준다.



그림 2. 원본과 PGD에 대한 결과  
Fig. 2. Results for the original and PGD

### III. 결 론

본 논문에서는 전이학습 과정 중 시계열 데이터를 다루는 예측 모델에 대해서 적대적 예제 공격이 적용됨을 관찰하였다. 단,  $\alpha$ 값과  $\epsilon$ 값을 너무 높게 설정하여도 실제 예측값과 많이 차이 남을 볼 수 있었다.

추후 이러한 전이학습 예측 모델을 더욱 강건하게 개발하고자 한다. 그러기 위해서는 FGSM과 PGD를 활용한 적대적 훈련을 하고자 한다. 단, 그러기 위해서는 적절한  $\alpha$ 값과  $\epsilon$ 값을 설정하여 효율적인 훈련이 이루어져야 한다.

### 참 고 문 헌

- [1] K. R. Mopuri, A. Ganeshan, and V. B. Radhakrishnan, "Generalizable data-free objective for crafting universal adversarial perturbations," IEEE transactions on pattern analysis and machine intelligence, 2018.
- [2] M. Zhang, K. T. Ma, J. Lim, Q. Zhao, and J. Feng, "Anticipating where people will look using adversarial networks," IEEE transactions on pattern analysis and machine intelligence, 2018.
- [3] 한규한, 정지홍, 김재현, 김성석, "학교 실내 미세먼지 예측 모델 개발", 한국정보기술학회 하계종합학술대회, p. 168-171, 2023.
- [4] X. Ma, Y. Niu, L. Gu, Y. Wang, Y. Zhao, J. Bailey, F. Lu, "Understanding Adversarial Attacks on Deep Learning-Based Medical Image Analysis Systems," ArXiv, 2014.

# 편향 완화를 통한 추천 시스템의 Unexpectedness 극대화: 개인화 추천 모델과 인기 추천 모델 간의 비교

정시은\*, 최상민\*\*

## Maximizing Unexpectedness in Recommendation Systems through Bias Mitigation: A comparison between personalized recommendation and Popularity Model

### 요 약

본 연구는 추천 시스템의 편향된 문제를 완화하고 Serendipity의 평가지표인 Unexpectedness 증대를 목표로 한다. 개인화 추천 모델은 사용자에게 예측 가능한 항목만을 반복적으로 추천하는 경향이 있어, 추천 항목의 다양성이 떨어진다. 이를 해결하기 위해 본 연구에서는 PM(Popularity Model)과 PR(Personalized Recommendation) 간의 추천 항목을 비교하여 Unexpectedness를 평가한다. PM은 전체 사용자들 사이에서 인기가 높은 항목을 추천하는 모델로, PR과 대조적으로 다양성을 제공할 수 있다. Unexpectedness 기반의 성능 평가 결과, PM과 PR 간의 비교가 PR과 PR 간의 비교보다 더 나은 편향성 개선 효과를 나타내었으며, Unexpectedness 범위가 더욱 확대되었음을 확인한다. 이를 통해 개인화된 추천 시스템이 발생시키는 편향을 완화하고 Serendipity를 증대시킬 수 있는 가능성을 입증한다.

### Abstract

This research aims to mitigate the problem of bias in recommendation systems and increase Serendipity's metric, Unexpectedness. PR, a personalized recommendation model, tends to repeatedly recommend only predictable items to users, resulting in a lack of diversity in recommendations. To solve this problem, we evaluate Unexpectedness by comparing recommendations between Popularity Model (PM) and Personalized Recommendation (PR). PM is a model that recommends items that are popular among all users, which can provide diversity in contrast to PR. The performance evaluation based on Unexpectedness shows that the comparison between PM and PR provides better bias reduction than the comparison between PR and PR, and the range of Unexpectedness is further expanded. This demonstrates the potential of personalized recommendation systems to mitigate bias and increase Serendipity.

### Key words

Recommender System, Personalized Recommendation, Popularity Model, Serendipity, Unexpectedness, Filter Bubble

### 1. 서 론

추천 시스템은 사용자의 과거 행동과 선호도를

분석하여 개인화된 추천을 효율적으로 제공하지만, 지나치게 개인화 된 추천은 특정 유형의 콘텐츠를 반복 추천함으로써 사용자에게 편향된 경험을 제공

\* 경상국립대학교 컴퓨터공학부, 220161105@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과/바이오의료빅데이터학과, neobits@gnu.ac.kr



하고 새로운 항목의 발견을 제한한다.[2] 이러한 문제를 해결하기 위해 Serendipity의 중요성이 대두되며, 이는 예상치 못한 유용한 콘텐츠 추천을 통해 다양한 경험을 제공하는 개념이다.[3] 본 연구는 이러한 편향된 추천을 보완하고 Serendipity의 평가지표 중 하나인 Unexpectedness를 극대화하기 위해 인기 기반 추천 모델인 PM과 Matrix Factorization 기반 개인화 추천 모델인 PR을 비교한다.[1] 본 논문에서 Unexpectedness의 평가지표는 PR과 PR, PR과 PM 간의 대칭 차집합으로 정의한다. 실험 결과, PR과 PM 간의 Unexpectedness는 99.62로 PR 간의 비교보다 8.86 더 높은 결과를 나타냈다. 이는 개인화된 추천만으로는 Serendipity를 충분히 확보하기 어렵다는 점을 시사하며, 다양한 추천 전략을 활용하는 것이 더 나은 사용자 경험을 제공할 수 있음을 보여준다.

## II. 제안 방법론

### 2.1 추천 시스템 모델

#### 2.1.1 PR(Personalized Recommendation)

PR 모델은 Matrix Factorization 기반의 개인화 추천 방식으로, 사용자의 과거 행동과 항목의 특성을 분석해 사용자와 항목 간의 잠재 요인을 학습하고 이를 바탕으로 개인화된 추천을 제공한다. 이는 사용자 평가 데이터를 활용해 아직 평가하지 않은 항목의 평점을 예측하고, 높은 순서대로 추천한다.[4]

#### 2.1.2 PM(Popularity Model)

PM 모델은 사용자가 평가하지 않은 항목 중 평균 평점이 높고 평가 개수가 100개 이상인 항목을 추천한다. 전체 사용자들의 평가 데이터를 기반으로 평균 평점을 계산하여 상위 N개의 항목을 추천 리스트로 구성한다.[1] 본 연구는 PR과 PM 간 비교로 Unexpectedness를 극대화하고, PR 간 비교로 추천 시스템의 편향성을 분석한다.

$$\bar{r}_i = \frac{1}{|u_i|} \sum_{u \in u_i} r_{ui} \quad (1)$$

위 수식에서  $\bar{r}_i$ 는 항목  $i$ 의 평균 평점을,  $|u_i|$ 는 해당 항목을 평가한 사용자의 수를 나타낸다. 이를 통해 추천 시스템에서 항목의 인기도를 평가한다.

$$I_u^{PM} = \text{argtop}_N(\bar{r}_i) \quad \text{for } i \in I_u^{\text{unrated}} \quad (2)$$

위 수식에서  $\text{argtop}_N(\bar{r}_i)$ 는 평균 평점  $\bar{r}_i$ 가 높은 상위 N개의 항목을 선택하는 연산으로, 이를 통해 사용자가 아직 평가하지 않은 인기 항목들을 추천하는데 활용한다.

### 2.2 평가지표

#### 2.2.1 Unexpectedness

Unexpectedness는 추천 시스템에서 Serendipity를 평가하는 중요한 지표로, 사용자가 예상하지 못한 항목을 추천받는 정도를 측정한다.[1] 기존의 추천 시스템은 지나친 개인화로 편향된 추천을 유발해 새로운 콘텐츠 발견을 제한하고 다양성을 저해할 수 있다.[3] Unexpectedness는 이러한 편향을 완화하여, 추천 항목이 사용자의 기존 선호나 행동 패턴과 얼마나 다른지를 평가함으로써 더 다양한 경험을 제공한다.[1] PR과 PM 모델의 추천 항목 유사도는 다음과 같은 수식으로 정의된다.[4]

$$I_u^{\text{diff}} = \left\{ j \in I_u^A \mid \max_{i \in I_u^B} \text{Sim}(i, j) < \tau \right\} \quad (3)$$

수식(3)은 A와 B로 표현된 두 추천 목록 간의 대칭 차집합을 나타내며, 본 연구에서 Unexpectedness를 평가하는 지표로 사용된다. 여기서 A와 B는 PR 간 비교 시 서로 다른 두 사용자의 PR 모델을, PM-PR 비교 시 PM과 PR 모델을 의미한다.

여기서  $\tau$ 는 유사도의 임계값으로, 추천 목록 간의 코사인 유사도가 임계값보다 작을 때 해당 항목들을 Unexpectedness로 정의하며, 본 논문에서는 임계값  $\tau$ 를 0.7로 설정한다.

#### 2.2.2 Serendipity

Serendipity는 추천 시스템에서 사용자가 예상하지

못한 유용한 항목을 발견하게 함으로써 사용자 만족도를 높이는 경험을 의미한다. Serendipity가 높을수록 반복적인 추천의 편향 문제를 완화하고, 더 다양한 콘텐츠를 제공해 사용자 경험을 풍부하게 한다.[1,3] 본 연구는 Serendipity의 평가지표인 Unexpectedness를 높이고, 이를 통해 사용자가 새로운 콘텐츠를 발견할 기회를 확장하고 Serendipity를 극대화할 가능성을 확인하고자 한다.

### III. 실험

#### 3.1 데이터셋

표 1. 실험 데이터 세트

Table 1. Experimental data sets

Dataset	File used
MovieLens Beliefs Dataset 2024	movies.csv
	user_rating_history.csv

데이터셋으로는 GroupLens의 MovieLens Beliefs Dataset 2024를 사용한다.

#### 3.2 실험 결과

표 2. 추천 모델의 성능

Table 2. Performance of PR Model

Accuracy measurement method	Results
MSE	0.96
RMSE	0.98
MAE	0.69
NDCG@5	0.74

표 3. Unexpectedness 결과

Table 3. Unexpectedness result

	PR과 PM	PR과 PR
Unexpectedness	99.62	90.76

표 2는 개인화 추천 모델(PR)의 성능으로, 해당 모델이 사용자의 선호도를 잘 반영해 높은 예측 정확도를 보여준다. 표 3은 Unexpectedness의 실험 결과로, PR과 PM 간의 Unexpectedness가 PR간의

Unexpectedness보다 8.86 더 높게 나타나, 인기 항목 추천 모델과의 비교가 더 큰 Serendipity를 제공할 가능성을 나타낸다.

### IV. 결론

본 연구는 추천 시스템의 편향을 완화하고 Serendipity의 평가지표인 Unexpectedness를 극대화하는 방법을 탐구한다. 기존 개인화 추천은 동일한 콘텐츠를 반복 추천해 새로운 발견 기회를 제한한다.[3] 본 연구에서는 이를 해결하기 위해 PM과 PR 간 추천 항목을 비교해 편향을 완화하고 Serendipity를 극대화할 가능성을 확인했다. 그러나 본 연구는 Serendipity의 또 다른 평가 요소인 Useful에 대한 데이터가 부족해, Unexpectedness에 중점을 둔 분석을 진행했다. 향후 연구에서는 Useful 데이터를 포함해 두 요소를 모두 고려하여 Serendipity를 더욱 정교하게 평가하고, 이를 통해 사용자 만족도를 극대화하는 연구가 진행되어야 할 것이다.

### 참고 문헌

- [1] Ge, M., Delgado-Battenfeld, C., & Jannach, D. (2010, September). Beyond accuracy: evaluating recommender systems by coverage and serendipity. In Proceedings of the fourth ACM conference on Recommender systems (pp. 257-260).
- [2] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8), 30-37.
- [3] Fu, Z., Niu, X., & Maher, M. L. (2023). Deep learning models for serendipity recommendations: a survey and new perspectives. *ACM Computing Surveys*, 56(1), 1-26.
- [4] Rahutomo, F., Kitasuka, T., & Aritsugi, M. (2012, October). Semantic cosine similarity. In The 7th international student conference on advanced science and technology ICAST (Vol. 4, No. 1, p. 1). South Korea: University of Seoul.

# 텍스트 프롬프트를 활용한 블라인드 폭우 얼굴 영상 복원

최연우\*, 김철환\*, 손창환\*\*

## Blind Heavy Rain Face Image Restoration Using Text-Prompt

Yeon-U Choi\*, Cheol-Hwan Kim\*, and Chang-Hwan Son\*\*

### 요약

폭우 환경에서 얼굴 인식 성능을 개선하기 위해 텍스트 프롬프트를 활용한 멀티모달 기반 얼굴 영상 복원 기법을 제안한다. 기존 블라인드 생성형 복원 기술은 해상도 개선에 중점을 두어 기상 조건에 따른 색상과 텍스처 왜곡이 발생하며 본 연구는 사전 학습된 텍스트 프롬프트를 사용해 이를 개선하고자 한다. 실험 결과, 제안 모델은 텍스처 및 색상 왜곡을 줄이고 디테일 표현을 강화함을 확인하였다.

### Abstract

To improve face recognition performance in heavy rain environments, we propose a multimodal-based face image restoration technique that utilizes text prompts. Conventional blind generative restoration techniques focus on improving resolution, but color and texture distortions occur depending on weather conditions. This study attempts to improve this by using pre-trained text prompts. Experimental results confirm that the proposed model reduces texture and color distortion and enhances detail expression.

### Key words

shape recovery, shape from focus, artificial neural network, depth estimation, focus measure

## 1. 서론

현재 지능형 감시카메라는 이상행동 탐지, 침입 경고, 객체 수 계산, 용의자 얼굴 인식 등 고차원 정보 처리를 수행하고 있지만 실외 설치 시 기상 조건에 취약하다는 단점이 존재한다. 안정적인 영상 분석을 위해서는 열화된 영상의 화질 복원이 필수적이다. 기존 블라인드 얼굴 복원은 일반적으로 맑

은 날씨에 초점을 맞춘 초해상화 기술에 집중한다. 그러나 이를 폭우 환경에서 동일한 모델 적용 시 색상 및 텍스처 왜곡이 발생해 결함을 완전히 제거하지 못하는 한계가 있다. 이를 개선하기 위해 기존 블라인드 초해상화 기술을 폭우 환경에 적합하도록 고도화하고 사전 학습된 텍스트 프롬프트를 활용한 멀티모달 기반 접근을 제안한다. 실험을 통해 기존 방식보다 제안한 방법이 우수한 결과를 보였다.

\*국립군산대학교 소프트웨어학과 학사과정

\*\*국립군산대학교 소프트웨어학과 교수(교신저자), cson@kunsan.ac.kr

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065)

## II. 관련 연구

### 2.1 블라인드 얼굴 영상 복원 기법

블라인드 얼굴 영상 복원은 얼굴 영상의 눈, 코, 입 같은 구조적 요소를 정밀하게 재현하는 기술이다. 최근 생성형 모델의 발전에 따라 원본 데이터의 분포를 학습하여 복원하는 방식이 주류가 되고 있다. 대표적인 생성형 모델인 StyleGAN[2]은 랜덤 벡터에 스타일 변환을 적용해 최종적으로 원하는 영상을 생성한다. StyleGAN을 활용한 얼굴 영상 복원 모델로는 PSFRGAN[3]이 있다. 이 모델은 파싱 맵과 저해상도 영상에서 스타일 벡터를 추정한 뒤 이를 입력 랜덤 벡터에 적용하여 고해상도 영상으로 복원하는 방식이다.

### 2.2 텍스트-영상 멀티모달 모델

최근 텍스트와 영상 데이터를 통합적으로 이해하고 분석하기 위해 두 데이터를 공통 임베딩 공간에서 매핑하는 사전 학습 모델이 각광받고 있다. 이 중 대표적인 CLIP 모델[4]은 대조 학습(Contrastive Learning)을 통해 텍스트와 영상 쌍의 유사도를 학습하여 두 정보를 효과적으로 정렬한다. CLIP은 사전 학습 단계에서 방대한 데이터 쌍을 학습해 별도의 추가 학습 없이도 활용할 수 있는 범용성과 높은 성능을 제공한다.

## III. 제안한 텍스트 활용 폭우 얼굴 영상 복원

### 3.1 모델 아키텍처

그림 1은 제안한 모델의 아키텍처를 보여준다. 본 모델은 기존 PSFRGAN과 달리, 텍스트-영상 정렬 기반 통계적 특징 변환(TIA-SFT) 블록을 새롭게 추가한 점이 핵심이다. TIA-SFT 블록은 입력된 텍스트 정보에서 대상의 통계적 특성인 평균과 표준편차를 추출하여, 기존 SFT 블록과 결합해 최종 통계 분포를 결정한다. 기존 SFT 모듈은 폭우 환경 영상의 통계적 특성을 정확히 반영하기 어려운 한계가 있으므로, 추가적인 텍스트 정보로 이를 보완하여 PSFRGAN의 성능을 향상시키고자 한다.

### 3.2 TIA-SFT 블록

그림 2은 TIA-SFT 블록의 상세 구조로 이 블록은 학습과 테스트 단계에서 서로 다른 경로로 동작한다. 학습 단계에서는 상단 경로를 따라 원본 고해상도 영상을 사전 학습된 이미지 인코더에 입력해 임베딩 공간으로 매핑하고, 이어서 임베딩 벡터는 완전연계층을 통해 타겟의 표준편차와 평균을 추출한다. 테스트 단계에서는 원본 고해상도 영상 대신 텍스트 정보를 입력 정보로 이용하는데 사전 학습된 텍스트 인코더에 들어가 이미지와 동일한 흐름으로 이어지게 된다.

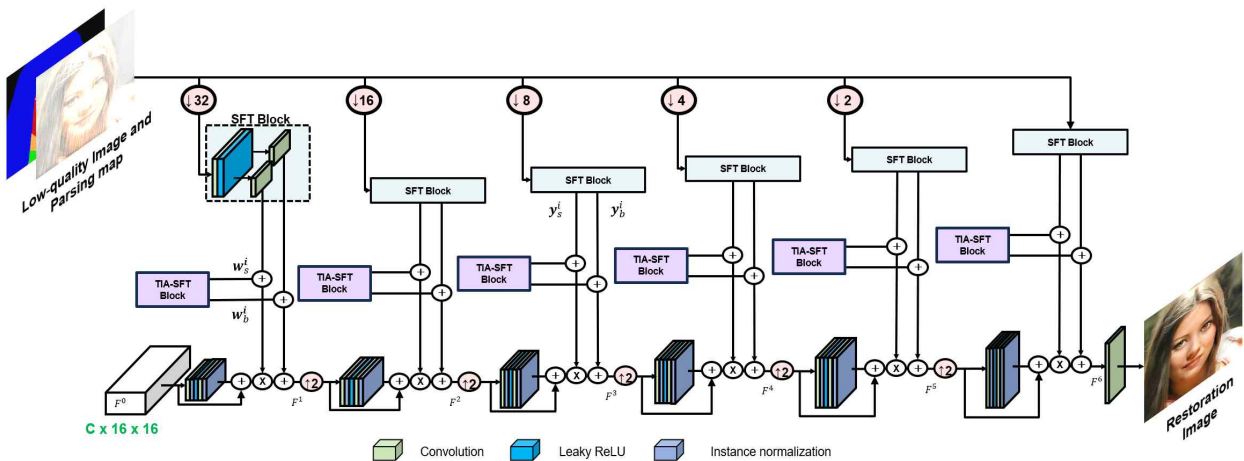


그림 1. 제안한 멀티모달 텍스트-이미지 정렬 기반 블라인드 폭우 얼굴 영상 생성.  
Fig. 1. Blind heavy rain face image generation based on multimodal text-image alignment.

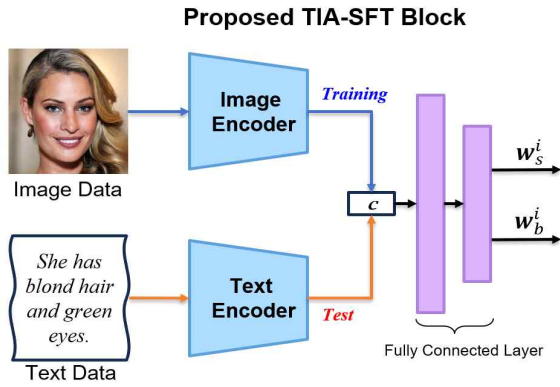


그림 2. 제안한 TIA-SFT 블록.  
Fig. 2. Proposed TIA-SFT block.

$$[w_s^i, w_b^i] = f_2(f_1(c)) \quad (2)$$

수식 (2)의  $f$ 는 완전연계계층으로 각각 첫 번째, 두 번째로 통과하는 층,  $c$ 는 텍스트-영상의 임베딩 벡터,  $[w_s^i, w_b^i]$ 는 TIA-SFT의 출력인 표준편차와 평균을 나타낸다. 학습 단계에서 고해상도 영상의 통계치를 추출하고, 테스트 단계에서 텍스트 정보를 통해 추정해 통계적 특징 변환 기능을 강화한다.

### 3.3 통계적 특징 변환

기존 SFT 블록과 새롭게 제안된 TIA-SFT 블록에서 얻은 평균과 표준편차는 푸전 과정을 통해 최종

타겟의 통계 분포를 형성한다.

$$F^i = (w_s^i + y_s^i) \frac{\Phi_{up}(F^{i-1}) - \mu(\Phi_{up}(F^{i-1}))}{\sigma(\Phi_{up}(F^{i-1}))} + (w_b^i + y_b^i) \quad (3)$$

수식 (3)은 통계적 특징 변환은 사전 학습된 텍스트 프롬프트 모델에서 추출된  $w_s^i$ 와  $w_b^i$ 를 활용한다. 즉, 기존 폭우 환경의 얼굴 영상에서 얻은 와 통계치에 열화되지 않은 텍스트나 원본 고해상도 영상에서 얻은 통계치를 추가함으로써 통계적 특징 변환을 강화하여 복원 영상의 품질을 더욱 향상할 수 있다.

## IV. 결 론

### 4.1 실험 환경

제안한 모델의 학습과 테스트에는 MM-CelebA-H Q 얼굴 영상, 파싱 맵, 그리고 텍스트 데이터셋을 사용했으며, 총 70,000장의 이미지 중 18,000장은 훈련에 100장은 테스트에 활용했다. 학습 설정은 배치 크기 4, 100 에폭, 학습률 0.0002이며, 아담 옵티마이저와 RTX A6000 GPU를 사용했다. 손실 함수는 PSFRGAN 모델과 동일하다.

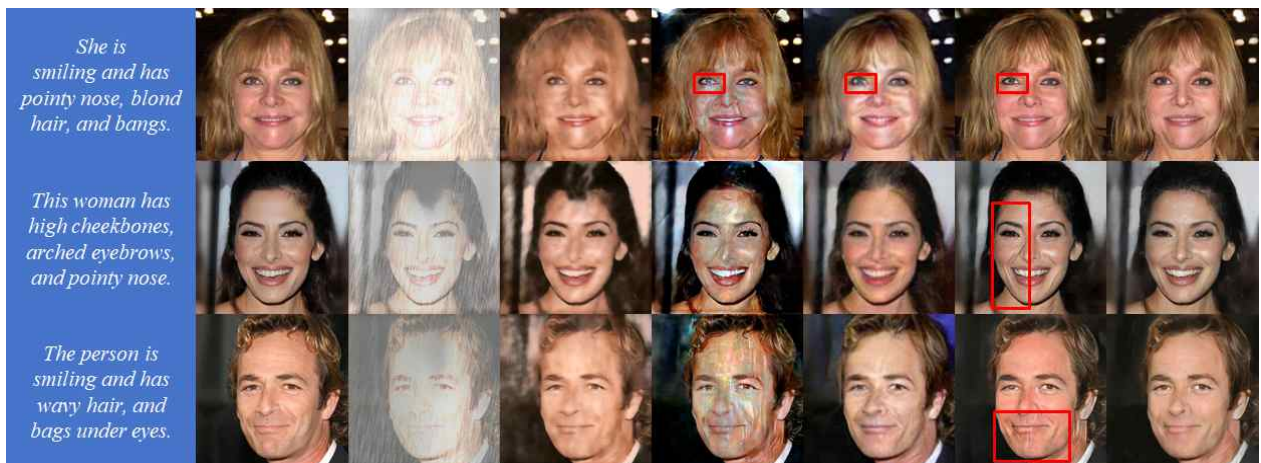


그림 3. 실험결과; 입력 텍스트 예시 (첫 열), 입력 폭우 영상 (두 번째 열), 폭우 제거 기법 (세 번째 열), HiFaceGAN (네 번째 열), FSRNet (다섯 번째 열), PSFRGAN (여섯 번째 열), 제안한 모델 (마지막 열).

Fig. 3. Experimental results; input text examples (first column), input heavy rain images (second column), heavy rain removal (third column), HiFaceGAN (fourth column), FSRNet (fifth column), PSFRGAN (sixth column), and proposed model (last column).

표 1. 정량적 화질 평가.

Table. 1. Quantitative image quality evaluation.

	PSNR ↑	SSIM ↑	LPIPS ↓	FID ↓
Heavy-Rain Removal [5]	18.0903	<b>0.6403</b>	0.4088	144.1401
HiFaceGAN [6]	19.1665	0.4862	0.3709	185.9966
FSRNet [7]	20.6047	0.6280	0.4457	135.6126
PSFRGAN [3]	21.6870	0.6020	0.2749	81.5439
<b>Proposed Model</b>	<b>21.7378</b>	0.6066	<b>0.2740</b>	<b>79.4575</b>

#### 4.2 정성적 화질 평가

제안한 모델의 성능을 평가하기 위해 기존 폭우 제거 모델[5], HiFaceGAN[6], FSRNet[7], PSFRGAN [3]과 비교하였다. 폭우 제거 모델은 열화된 영상에서 빗물 요소를 제거하고 다른 모델들은 얼굴 영상의 초해상화에 특화되어 있다. 그림 3에는 실험 결과가 시각화되어 있으며 첫 번째와 두 번째 열은 입력된 텍스트와 원본 얼굴 영상을 나타낸다. 세 번째 열의 얼굴 영상은 블라인드 열화 모델을 적용하여 다양한 해상도와 빗줄기 효과를 반영한 폭우 환경 영상을 보여준다. 결과에서 폭우 제거, HiFaceGAN, FSRNet 모델은 고해상도 복원에서 부족하여 텍스처 왜곡, 색상 저하, 완전하지 않은 빗줄기 제거 등 한계를 드러냈다. 반면, PSFRGAN과 제안한 모델은 빗줄기 제거와 선명도에서 더 나은 성능을 보였다. 특히, PSFRGAN과 비교했을 때 제안 모델은 얼굴 구성요소의 선명도와 빨간 박스 영역에서 머리카락 디테일을 개선하고 텍스처와 색상 왜곡을 줄여 보다 자연스러운 결과를 제공했다.

#### 4.2 정량적 화질 평가

정량적 평가를 위해 PSNR(Peak Signal-to-Noise Ratio), SSIM(Structural Similarity), LPIPS(Learned Perceptual Image Patch Similarity), FID(Frechet Inception Distance) 지표를 사용하였다. PSNR과 SSIM은 원본 영상이 필요한 레퍼런스 기반 평가 지표로 값이 높을수록, LPIPS와 FID는 원본 영상 없이도 인지적 화질을 측정하며 값이 낮을수록 화질이 우수함을 나타낸다. 표 1에 나타난 결과에서 제안한 모델은

PSNR, LPIPS, FID 지표에서 가장 뛰어난 성능을 보였으며, SSIM에서는 폭우 제거 모델이 우수했다. 그러나 PSFRGAN보다 모든 지표에서 개선된 성능을 확인할 수 있다. 이는 TIA-SFT 블록이 영상 복원에 효과적임을 입증한다.

#### References

- [1] C.-H. Son and S.-Y. Jeong, "Blind heavy rain face image restoration using statistical feature transformation for each facial component," The Journal of Korean Institute of Information Technology, vol. 21, no. 3, pp. 107-115, 2023. 10.14801/jkiit.2023.21.3.107.
- [2] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 8110-8119, 2020. <https://doi.org/10.48550/arXiv.1912.04958>.
- [3] C. Chen, X. Li, L. Yang, and X. Lin, "Progressive semantic-aware style transformation for blind face restoration," In Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 11896-11905, Virtual, 2021. <https://doi.org/10.48550/arXiv.2009.08709>.
- [4] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, et al., "Learning transferable visual models from natural language supervision," Proc. 38th Int. Conf. Machine Learning, pp. 8748-8763, Feb. 2021. <https://doi.org/10.48550/arXiv.2103.00020>.

- [5] R. Li, L. F. Cheong, and R. T. Tan, "Heavy rain image restoration: Integrating physics model and conditional adversarial learning," In Proc. IEEE International Conference on Computer Vision and Pattern Recognition, Long Beach, USA, pp. 1633-1642, June 2019. <https://doi.org/10.1109/CVPR.2019.00173>.
- [6] L. Yang, C. Liu, P. Wang, S. Wang, P. Ren, S. Ma, and W. Gao, "Hifacegan: Face renovation via collaborative suppression and replenishment," In Proc. ACM International Conference on Multimedia, Seattle, USA, pp. 1551-1560, Oct. 2020. <https://doi.org/10.1145/3394171.3413965>.
- [7] Y. Chen, Y. Tai, X. Liu, C. Shen, J. Yang, "FSRNet: End-to-end learning face super-resolution with facial priors," In Proc. IEEE International Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 2018, pp. 2492-2501. [10.1109/CVPR.2018.00264](https://doi.org/10.1109/CVPR.2018.00264).

# 전북특별자치도 건축물 노후화도에 따른 에너지 소모량 분석을 위한 시각화 시스템

최현식\*, 이민지\*, 서정훈\*, 최미령\*, 김태희\*, 이상민\*, 김재엽\*, 김동현\*, 이석훈\*†

## Visualization System for Analyzing Energy Consumption according to the Aging Level of Buildings in Jeonbuk-Do

Hyeonsik Choi\*, Minji Lee\*, Jeonghoon Seo\*, Miryeong Choi\*, Taehee Kim\*, Sangmin Lee\*,  
Jaeyeob Kim\*, Donghyun Kim\*, Sukhoon Lee\*†

### 요 약

본 논문은 전북 특별 자치도가 제공하는 공공데이터를 바탕으로 노후화도에 따른 에너지 사용량의 변화량을 볼 수 있는 시각화 시스템을 제안한다. 국토교통부 자료에 따르면 전북특별자치도의 노후화 건물 비율은 57.9%로 17개 시도 중 6번째로 높다. 그러나 기존 에너지 사용량 분석 데이터들은 노후화도와 건축물의 에너지 사용량에 미치는 영향을 반영하지 못하고 있으며, 건축물에 노후화도가 에너지 사용량에 미치는 영향을 분석하기 위한 도구 또한 부족하다. 이에 본 논문이 제안하는 시스템은 전북특별자치도의 에너지의 사용량과 건축물 데이터를 통합해 OpenLayers와 GeoServer를 이용한 시각화 웹 시스템을 통해 건축물 별 데이터와 에너지 사용량 데이터를 제공한다. 노후화도에 따른 에너지 사용량의 추이를 볼 수 있는 그래프와 대시보드를 제공한다.

### Abstract

This paper proposes a visualization system that allows users to see the change in energy usage according to the degree of aging based on public data provided by the Jeonbuk-Do. According to Ministry of Land, Infrastructure, and Transport data, the percentage of aging buildings in Jeonbuk is 57.9%, the 6th highest among 17 provinces. However, the existing energy usage analysis data does not reflect the impact of aging on the energy usage of buildings, and there is a lack of tools to analyze the impact of aging on the energy usage of buildings. Therefore, the system proposed in this paper integrates energy usage and building data of Jeonbuk-Do and provides building-specific data and energy usage data through a visualization web system using OpenLayers and GeoServer. It provides graphs and dashboards to view the trend of energy usage according to the degree of aging.

### Key words

aging building, public data, building energy usage, GIS

\* 국립군산대학교 소프트웨어학과

† 교신저자

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 "SW중심대학사업" 지원을 받아 수행되었음 (2023-0-00065)



표 1. 건축물 GIS 데이터셋

Table 1. A dataset of building GIS

법정동명	번지	건축물 구조명	건축물면적	허가날자	건폐율	PNU
전북특별자치도 남원...	129	일반목구조	51.44	1905-01-01	23.92	5219038029...
전북특별자치도 완주...	1141-2	일반목구조	118.38	2024-04-01	30.5792	5271025626...
전북특별자치도 남원...	6-1	강파이프구조	2218.5	2013-03-06	0.0	5219038028...
전북특별자치도 남원...	6-1	강파이프구조	2218.5	2013-03-06	0.0	5219038028...
전북특별자치도 남원...	6-1	강파이프구조	2218.5	2013-03-06	0.0	5219038028...

## I. 서론

국도교통부에서는 매년 ‘전국 건축물 현황’을 발행하고 있으며, 전북특별자치도는 57.9%로 전국 17개 시도 중 6위에 해당한다. 노후화도는 건축물의 에너지 사용량에 영향을 끼치는 요인 중 하나이다.

그러나 기존 데이터들은 건물의 이용 형태, 물리적 특징, 지리적 배치 등에만 초점을 맞춰져 있다.

그러나 노후도 또한 에너지 사용량에 영향을 끼치는 요소인 만큼 그에 따른 건축물의 노후화도와 에너지 사용량의 상관관계를 분석하기 위한 도구가 필요하다.

이에 본 논문은 건축물의 노후화도에 따른 에너지 상관관계 분석을 위해 공공데이터 포털에서 제공하는 건축물 GIS 데이터와 전기 에너지 사용량 데이터를 통합한 데이터를 기반으로 지리정보시스템(GIS)을 구축했다. 다. 이를 통해 건축물의 노후화와 에너지 사용량 간의 상관관계를 분석하고 문제 해결의 기반이 되는 것을 목표로 한다.

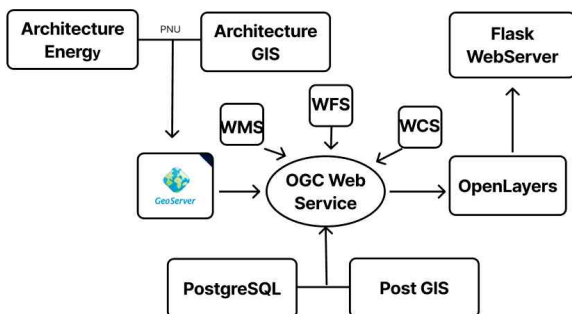


그림 1. 제안 시스템 구조  
Fig 1. A proposed system architecture

## II. 제안 시스템 설계

그림 1은 제안 시스템의 구조를 보인다. 제안 시스템은 건축물 GIS 데이터셋과 건축물 에너지 데이터를 통합하고 통합된 데이터셋을 GeoServer와 연결한 후 이를 OpenLayers를 통해 Flask 기반 웹 서버로 시각화 서비스를 제공한다. 이때 PostgreSQL과 PostGIS가 사용된다.

### 2.1 데이터 전처리

지도에 올릴 건축물 폴리곤 출력에 사용될 데이터셋인 건축물 GIS 데이터와 에너지 데이터를 PostgreSQL을 사용하여 DB의 테이블에 넣고 각각 PNU(필지 고유번호)를 추출해 새로운 컬럼을 생성했다. 이를 통합을 위한 키로 두고 에너지 사용량 데이터 또한 PNU 추출을 통해 이것을 키로 두고 진행했다. 표 1과 표 2는 각각 건축물 GIS와 건축물 에너지 사용량 데이터셋의 예를 보인다.

표 2. 건축물 에너지 사용량 데이터셋

Table 1. A dataset of building energy

달	PNU	주소	사용량
201707	4514025024...	전라북도 익산시 힘...	753.0
201707	4513014100...	전라북도 군산시 수...	134.0
201707	4511310700...	전라북도 전주시 덕...	188.0
201706	4519011800...	전라북도 남원시 고...	213.0
201706	4514010400...	전라북도 익산시 중...	101.0

## 2.2 시각화 모듈

Flask를 이용한 시각화 시스템 구축을 목표로 하고 통합된 데이터를 지오서버와 연동했다. 이를 WMS를 통해 OpenLayers로 전달하고 지도 레이어를 생성했다.

건축물의 노후화도에 따라 지도에 건축물 폴리곤 색상을 채도를 통해 표현했으며, 오래된 건물을 진한 색상으로 나타내고 정보가 없는 건축물은 보라색으로 표현하여 노후화 정도와 정보가 없음을 직관적으로 파악할 수 있게 했다. 또한 네비게이션 바를 통해 건물의 클릭시 좌표에 해당하는 건물의 에너지 사용량과, 건축물의 대장 정보를 보여주며 이 또한 사용량에 따라 채도로 표현되게 해 직관적으로 정보를 전달할 수 있게 했다.

## III. 구현 결과

건축물의 나이에 따라 붉은색의 채도를 낮춰 표현했다. 데이터가 없는 건축물은 보라색으로 표기했으며, 좌측에 네비게이션 바를 두어 건축물의 에너지 사용량의 변화를 보여주는 그래프와 건축물의 데이터를 제공한다.

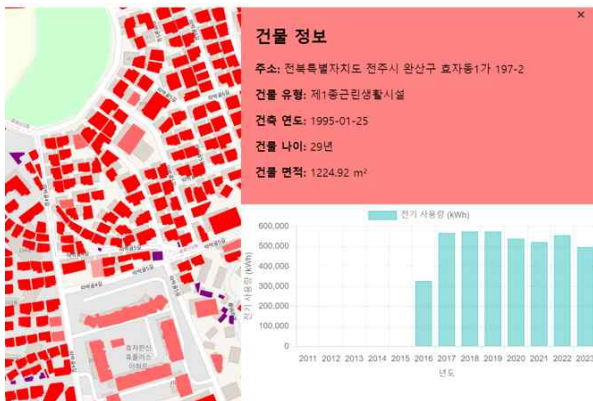


그림 2. 구현 결과 스크린샷  
Fig 2. A screenshot of implementation result

## VI. 결 론

본 논문에서는 GeoSever와 OpenLayers를 이용한 건축물의 노후화도에 따른 에너지사용량 시각화 시스템을 제안하였다. 데이터셋 간의 기준점인 PNU를

기준으로 통합을 진행하여 노후화도에 따라 건축물을 채도로 구별하게 하였고 에너지 사용량의 변화량을 볼 수 있도록 그래프를 제공하였다. 이를 Flask를 통해 웹으로 시스템을 구축하였다.

향후 연구에서는 에너지 사용량에 따른 탄소 배출량과 건축물 재질에 따른 에너지 변화량 등 여러 요인들을 추가하여 더 다양한 기능을 제공하는 시각화 도구가 될 수 있을 것으로 기대한다.

## 참 고 문 헌

- [1] S. Y. Lim, S. H. Kim, S. M. Oh, and J. Y. Kim, "A study on the energy evaluation and management system improvement of old public buildings," Proceedings of the 2024 Spring Conference of the Korean Society of Energy Engineering, The Korean Society of Energy Engineering, p. 182, May 2024.
- [2] H. J. Seo and N. H. Byun, "A study on the support for renovating aging residential buildings - A focus on private building green remodeling initiatives," Journal of the Architectural Institute of Korea, vol. 40, no. 4, pp. 187-194, Apr. 2024.
- [3] S. Y. Kim, S. I. Moon, J. W. Oh, M. R. Hur, and J. H. Oh, "A study on the evaluation of energy loss in residential buildings for urban carbon neutrality," Proceedings of the 2024 Spring Conference of the Korean Solar Energy Society, The Korean Solar Energy Society, p. 183, Apr. 2024.
- [4] J. Yang, J. Kim, and J. Kim, "Evaluation of insulation performance of old public buildings according to aging," Proceedings of the Kongju National University Conference, Kongju National University, p. 3, 2024.

# 한국어 NLP에서 형태소 기반 토큰라이저와 형태소 기반 서브워드 토큰라이저의 성능 비교

정찬우\*, 정현준\*\*

## Performance comparison of morpheme-based tokenizer and morpheme-based subword tokenizer in Korean NLP

Chanwoo Jeong\*, and Hyunjun Jung\*\*

### 요 약

이 논문은 한국어 자연어 처리(NLP)에서 형태소 기반 토큰라이저(Okt)와 형태소 기반 서브워드 토큰라이저(KLUE BERT의 Morpheme-based Subword)의 성능을 비교 분석하는 데 중점을 둔다. 각 토큰라이저가 한국어 데이터셋에서 문법적 정밀성, 문맥 보존, 처리 안정성 등의 측면에서 어떠한 차이를 보이는지 실험을 통해 평가하며 데이터셋의 특성과 복잡도에 따라 두 토큰라이저의 성능이 어떻게 달라지는지 분석한다. 이를 통해 작업별로 적합한 토큰라이저의 선택이 한국어 자연어 처리 성능에 미치는 영향을 확인하고 향후 한국어 NLP 모델 구축 시 효과적인 토큰라이저 선택에 필요한 기준을 제시하고자 한다.

### Abstract

This paper focuses on comparing the performance of a morpheme-based tokenizer (Okt) and a morpheme-based subword tokenizer (Morpheme-based Subword of KLUE BERT) in Korean natural language processing (NLP). We experimentally evaluate the differences between each tokenizer in terms of grammatical precision, context preservation, and processing stability on a Korean dataset, and analyze how the performance of the two tokenizers varies depending on the characteristics and complexity of the dataset. Through this, we verify the impact of selecting an appropriate tokenizer for each task on the performance of Korean natural language processing, and suggest criteria for selecting an effective tokenizer when building a Korean NLP model in the future.

### Key words

morpheme-based tokenizer, Morpheme-based Subword tokenizer, Korean natural language

---

\* 국립군산대학교 소프트웨어학과 학사과정, jcw0690@gmail.com

\*\* 군산대학교 소프트웨어학부 교수(교신저자), junghj85@kunsan.ac.kr

※이 연구는 정부(과학기술정보통신부)의 재원으로 한국 연구재단의 지원을 받아 수행되고 있습니다.(No. NRF-2022R1G1A1008493).

※본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065).

## I. 서론

자연어 처리(NLP) 분야에서 대규모 언어 모델(LLM)의 발전은 다양한 언어적 과제를 해결하는 데 진전을 가져왔다. 특히 GPT, BERT와 같은 모델이 여러 언어에서 우수한 성능을 보인다. 현재 대부분의 LLM은 영어 기반으로 개발되었기 때문에 한국어와 같은 고유 문법 구조를 가진 언어를 다루는데 한계가 있다[5]. 한국어는 교착어로서 문장 내 조사와 어미 변형이 중요한 역할을 하며 이러한 특성 때문에 단순한 띄어쓰기 기반 단어 분할로는 문법적 정밀성을 확보하기 어렵다[6]. LLM의 한국어 자연어 처리 성능을 높이기 위해서는 여러 접근 방식이 있는데 한국어의 복잡한 문법적 특성과 문맥적 의미를 모델이 제대로 이해하기 위해서는 토크나이즈의 역할이 중요하다[3]. 토크나이즈란 텍스트 데이터를 처리 가능한 형태로 변환하는 중요한 전처리 도구로 입력 문장을 단어 또는 그 이하의 단위로 분할하여 모델이 의미 있는 패턴을 학습할 수 있도록 돕는 역할을 한다. LLM을 이용한 자연어처리를 할 때 텍스트 데이터는 일련의 단계를 거쳐 모델의 입력으로 사용된다. 입력된 텍스트는 토크나이즈를 통해 단어, 서브워드, 형태소 등의 단위로 분할된다. 이러한 토큰들은 임베딩 레이어를 통해 벡터 형태로 변환되어 대규모 언어 모델(LLM)의 입력으로 사용되고 LLM은 이 벡터들을 처리하여 최종 결과를 도출한다.

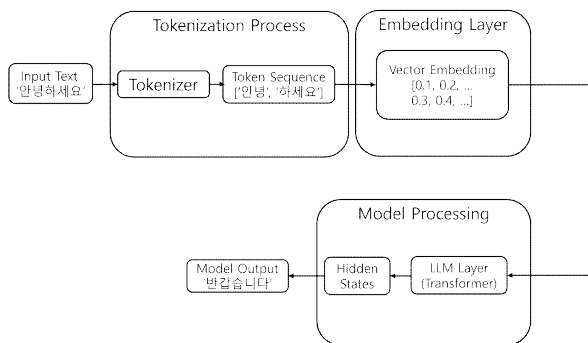


그림 1. LLM 토크나이즈 파이프라인  
Fig 1. LLM Tokenizer Pipeline

이러한 과정에서 토크나이즈의 종류와 방식에 따라 입력 텍스트의 정보 손실 정도나 문법적 일관성

이 크게 달라지며 이는 곧 NLP 모델의 성능에 직접적인 영향을 미친다[3]. 이처럼 좋은 성능을 내기 위해서는 한국어에 특화된 토크나이즈의 역할이 중요하다. 토크나이즈는 크게 형태소 기반 토크나이즈와 서브워드 기반 토크나이즈로 나누어진다. 형태소 기반 토크나이즈인 Okt나 MeCab-ko는 조사와 어미 등을 분리하여 문법적 정확성을 높이는 데 유리하고 감정 분석과 같은 문맥적 정보가 중요한 작업에 적합하지만 비표준 언어와 사전에 정의되지 않은 단어(OOV)에 약점을 보인다[3]. 반면에 SentencePiece나 WordPiece같은 서브워드 기반 토크나이즈는 단어를 더 작은 단위로 분할하여 어휘 확장성과 미등록어(OOV) 문제를 해결하는 성능이 뛰어나지만, 학습된 토큰의 수가 성능에 영향을 미쳐 토큰 수를 잘못 설정할 경우 분류 성능이 저하될 수 있다[4]. Morpheme-based Subword 토크나이즈는 형태소와 서브워드 방식을 결합하여 문맥 보존과 어휘 처리의 유연성을 동시에 제공한다[1]. 형태소 기반 토크나이즈(Okt)와 형태소 기반 서브워드 방식의 KLUE 토크나이즈의 성능을 작업별로 분석하고 기존의 형태소 기반 토크나이즈 대비 형태소 기반 서브워드 토크나이즈의 개선된 점을 분석함으로써 한국어 데이터셋 분석에 최적화된 토크나이즈 선택에 기여한다.

## II. 관련 연구

### 2.1 한국어 NLP 작업에서의 토크나이즈 성능 비교 연구

이 논문에서는 한국어 자연어 처리(NLP)에서 두 가지 방식의 토크나이즈의 성능을 비교하고 분석했다. 형태소 기반 토크나이즈와 서브워드 기반 토크나이즈가 각각의 NLP 작업에서 다른 성능 특성을 보인다는 것을 실험적으로 입증하였다. 특히 토크나이즈의 선택에 따라 입력 텍스트의 정보 손실과 문법적 일관성이 달라지며 이는 곧 모델의 성능에 직접적인 영향을 미친다는 점을 강조한다.

## 2.2 KLUE: 한국어 자연어 이해 평가를 위한 벤치마크와 형태소 기반 서브워드 토큰라이저

이 논문에서는 기존의 서브워드 토큰라이저에 한국어의 고유한 문법적 특성을 반영한 형태소 기반 서브워드 토큰라이저를 제안했다. 이를 통해 한국어 문법과 문맥적 의미를 보다 효과적으로 처리할 수 있음을 입증했다. 형태소 기반 서브워드 토큰라이저는 형태소 수준에서 태깅, 감지, 생성 작업에서 우수한 성능을 보인다.

### III. 형태소 기반 토큰라이저와 형태소 기반 서브워드 토큰라이저의 성능 평가 방법

#### 3.1 토큰라이저 비교

본 논문에서는 형태소 기반 토큰라이저 Okt와 형태소 기반 서브워드 토큰라이저 KLUE 토큰라이저의 성능을 비교하였다. Okt 토큰라이저는 한국어의 형태소를 기반으로 문장을 의미 단위로 분리하여 한국어 문법 구조에 적합한 세밀한 분석을 가능하게 한다. 이를 통해 각 형태소가 가지는 문법적 역할을 파악하며 문장 내 의미 단위를 유지한다. KLUE의 형태소 기반 서브워드 토큰라이저는 형태소 분석과 서브워드 토큰화를 결합한 방식으로 한국어의 언어적 특성을 반영하면서도 미등록어(OOV) 처리에 장점을 가져 다양한 문맥에서 유연하게 사용할 수 있다.

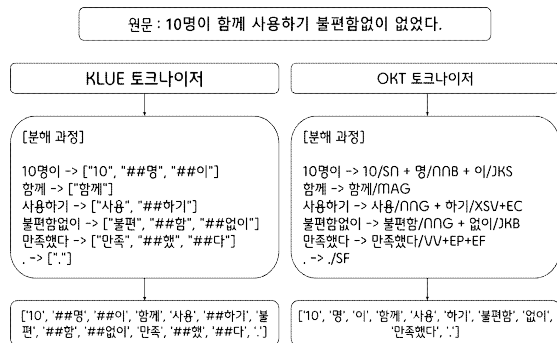


그림 2. 토큰화 예시

Fig 2. Tokenization Example

## 3.2 데이터셋과 평가 방법

실험은 KorQuAD, NSMC, KLUE NLI의 3가지 데이터셋으로 구성했다.

KorQuAD는 한국어 질의응답(QA) 작업을 위한 데이터셋으로 문맥 속에서 질문에 대한 정확한 답변을 추출하는 작업을 다룬다. QA 작업의 성능 평가는 F1 Score와 EM(Exact Match) Score를 사용하여 모델이 질문과 문맥에서 정확하게 답변을 추출하는지 측정한다.

NSMC는 긍정과 부정을 분류하는 감정 분석 데이터셋으로 영화 리뷰를 기반으로 작성되었다. 감정 분석 작업에서는 문장의 전반적인 흐름에서 감정 표현이 잘 반영되는지를 평가한다. NSMC 데이터셋의 성능 평가는 Accuracy와 F1 Score를 사용하여 긍정, 부정 예측의 정확성과 정밀도를 측정한다.

KLUE NLI는 두 문장 간의 관계를 추론하여 의미와 문맥적 연결성을 평가하는 데이터셋이다. 자연어 추론(NLI) 작업에서는 문맥 보존이 중요한 요소로 작용하며 두 문장이 맥락상 어떻게 연결되는지 평가한다. 이 작업의 성능 평가는 Pearson Correlation과 Spearman Correlation을 통해 예측된 유사도와 실제 유사도 간의 상관관계를 측정한다.

## IV. 실험 결과 및 분석

### 4.1 실험 환경

실험 환경은 다음의 포치럼 구성했다. 보다 정확한 성능 평가를 위해 CUDA환경에서 실험을 진행했다.

표 1. 실험 환경

Table 1. experimental environment

Feature	Specification
CPU	i9-13900K
GPU	NVIDIA RTX A6000
OS	Window11 pro
CUDA	12.4
Memory	32GB

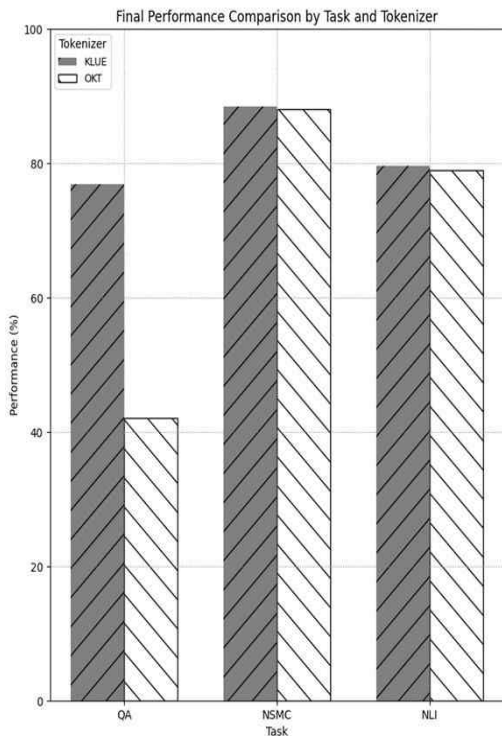
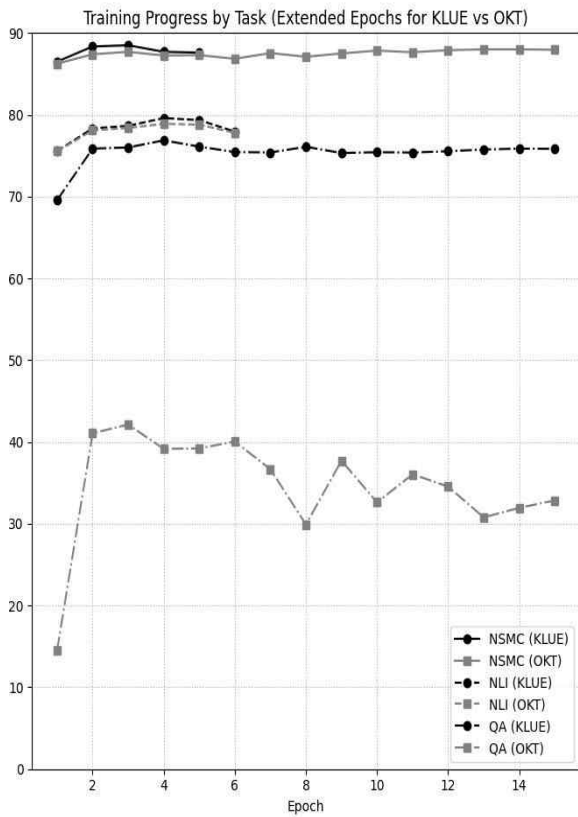


그림 3. 성능 평가 지표 및 결과  
Fig 3. Performance metrics and results

## 4.2 파인튜닝 환경 구성

OkT와 KLUE의 토크나이즈를 각각 동일한 BERT 기반 모델에 파인튜닝한다. 모델 구조와 하이퍼 파라미터는 동일하게 설정해 토크나이즈의 차이가 작업 성능에 미치는 영향을 분석할 수 있도록 한다. 시퀀스의 길이는 데이터가 긴 QA는 256토큰, 데이터의 길이가 비교적 짧은 NSMC, NLI는 128토큰으로 설정했다. 학습 횟수는 예폭당 625번 15번의 예폭동안 9,375번의 학습을 진행하게 되고 학습 도중 연속적인 성능 감소를 과적합의 신호로 판단해 학습을 조기 종료한다.

표 2. 파인튜닝 파라미터  
Table 2. Finetuning Parameters

max_train_samples	20000
max_eval_samples	2000
batch_size	32
epoch	15
max_length	256/128
early_stopping_patience	2

## 4.3 비교 평가 결과

그림 3을 보면 질의응답(QA) 태스크에서 KLUE 토크나이즈가 OKT 토크나이즈보다 높은 성능을 보였다. KLUE 토크나이즈를 사용한 모델은 F1 Score 76.87%와 EM Score 68.15%를 기록했는데 OKT 토크나이즈는 F1 Score 42.13%, EM Score 21.05%로 큰 차이를 보였다. 특히 OKT 토크나이즈의 경우 전체 데이터의 4.75%만을 유효하게 처리할 수 있었다. 긴 문맥 이해가 필요한 QA 태스크에서 형태소 기반 토크나이즈의 한계를 보인다.

감성분석(NSMC) 태스크에서는 두 토크나이즈가 적은 차이를 보였다. KLUE 토크나이즈는 Accuracy 88.15%, F1 Score 88.39%를 기록했으며, OKT 토크나이즈는 Accuracy 87.95%, F1 Score 88.09%를 기록했다. 이는 짧은 문장의 감성 판별에는 두 토크나이즈 모두 효과적임을 보여준다. 또한 KLUE 토크나이즈가 3번째 예폭에서 최고 성능에 도달한 반면, OKT는 13번째 예폭에서 최고 성능을 보여 학습

효율성에 차이가 있었고 KLUE 토크나이저는 5에폭에서 조기 종료 되었지만 OKT 토크나이저는 15 에폭까지 진행했다.

자연어 추론(NLI) 태스크에서는 중간 정도의 성능 차이를 보였다. KLUE 토크나이저는 Accuracy 79.60%, F1 Score 79.60%를 기록했으며, OKT 토크나이저는 Accuracy 78.90%, F1 Score 78.88%를 기록했다. 두 토크나이저 모두 4번째 에폭에서 최고 성능에 도달했으며 6번째 에폭에서 조기 종료되어 유사한 학습 패턴을 보였다.

이러한 결과는 태스크의 복잡도가 증가할수록 형태소 기반 서브워드 토크나이저(KLUE)의 장점이 더욱 부각됨을 보여준다. 특히 문맥 이해와 긴 시퀀스 처리가 필요한 태스크에서 KLUE 토크나이저의 성능이 뛰어났다.

## V. 결 론

이 논문에서는 형태소 기반 토크나이저(OkT)와 Morpheme-based Subword 토크나이저(KLUE)의 성능을 KorQuAD, NSMC, KLUE NLI 3가지의 한국어 데이터셋에서 비교했다. 실험 결과 OkT는 조사와 어미를 포함한 문법적 정밀성이 중요한 감정 분석에서 유사한 성능을 보였고 KLUE 토크나이저는 문맥을 잘 반영하고 복합어와 문맥 처리가 중요한 QA 및 NLI 작업에서 높은 성능을 보였다. 이로써 작업 특성에 따른 토크나이저의 중요성을 확인할 수 있었다.

향후 연구에서는 더 다양한 한국어 NLP 데이터셋과 하이브리드 토크나이저 방식을 적용하여 한국어 특성에 최적화된 토크나이저를 연구하고자 한다. 이를 통해 한국어 NLP 분야에서 LLM의 성능을 향상시킬 수 있는 가이드를 제공할 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1] Park, S. (2021). KLUE: Korean Language Understanding Evaluation
- [2] Ham, J., Choe, Y. J., Park, K., Choi, I., & Soh,

- H. (2020). KorNLI and KorSTS: New benchmark datasets for Korean natural language understanding.
- [3] 이소람. (2023). Analysis of Tokenizers for Various Korean NLP Tasks (Doctoral dissertation, 서울대학교 대학원).
- [4] Choo, S., & Kim, W. (2023). A study on the evaluation of tokenizer performance in natural language processing. Applied Artificial Intelligence, 37(1), 2175112.
- [5] Joo, H. Y., Oh, H., & Yang, J. (2023). A Survey on Open Source based Large Language Models. The Journal of Korea Institute of Information, Electronics, and Communication Technology, 16(4), 193-202.
- [6] Park, S. U. (2021). Analysis of the status of natural language processing technology based on deep learning. The Journal of Bigdata, 6(1), 63-81.

# 소규모 데이터셋에서의 가중 평균 기반 하이브리드 추천 시스템 성능 검증

김민성\*, 최원석\*, 조운식\*\*

## Performance Verification of Weighted Average-Based Hybrid Recommendation System in Small-Scale Datasets

Minseong Kim\*, Wonseok Choi\*, Yoonsik Cho\*\*

### 요약

디지털 스트리밍 서비스 이용이 폭발적으로 증가함에 따라 사용자의 선호도에 맞춘 추천 시스템 또한 큰 주목을 받고 있다. 본 연구는 소규모 데이터셋 환경에서 추천 시스템의 성능을 검증해보고자 협업 필터링, 콘텐츠 기반 필터링, 하이브리드 모델을 비교 분석하였다. 이를 위해 음악 선호도 조사와 음악 특징 추출을 바탕으로 소규모 데이터셋을 구축해 협업 필터링과 콘텐츠 기반 필터링 각각의 신경망 모델을 구현하고, Soft Ensemble 기법을 적용한 하이브리드 모델을 통해 추천 성능을 검증하였다. 실험 결과, 가중 평균을 적용한 하이브리드 모델이 소규모 데이터셋에서도 최적의 성능을 보였으며, 하이퍼파라미터 실험을 통해 협업 필터링과 콘텐츠 기반 필터링 간의 최적의 가중 평균 비율을 도출하였다.

### Abstract

As the use of digital streaming services has exploded, the recommendation system tailored to users' tastes is also attracting a lot of attention. This study compared and analyzed collaborative filtering, content-based filtering, and hybrid models to verify the performance of recommended systems in a small-scale datasets environment. To this end, small-scale datasets were established based on music preference surveys and music feature extraction, implementing neural network models for collaborative filtering and content-based filtering, and verifying recommended performance through hybrid models with Soft Ensemble techniques. As a result of the experiment, the hybrid model with weighted average showed optimal performance even in small-scale datasets, and through hyperparameter experiments, the optimal weighted average ratio between collaborative filtering and content-based filtering was derived.

### Key words

Collaborative Filtering, Content-based Filtering, Hybrid Recommendation System, Weighted Average

---

\*중앙대학교 AI학과 학부과정, {maruniverse, jsandjk}@cau.ac.kr

\*\*중앙대학교 AI학과 교수, yoonsik@cau.ac.kr (교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2021-0-01341)과 한국연구재단의 지원(No. RS-2024-00419201)을 받아 수행된 연구임



## I. 서론

음악, 영화, 드라마 스트리밍 서비스의 폭발적인 인기로 힘입어 방대한 정보 속 아이템을 추천해주는 추천 시스템 또한 큰 주목을 받고 있다. 추천 시스템은 사용자의 선호도를 파악해 맞춤형 아이템을 제공하는데 목적이 있으며, 특히 데이터가 방대해질수록 사용자 경험을 개선하는데 큰 역할을 한다. 이러한 추천 시스템은 크게 협업 필터링(Collaborative Filtering) [1, 2, 3]과 콘텐츠 기반 필터링(Content-based Filtering) [4, 5]으로 나뉜다.

협업 필터링은 사용자 간의 상호작용을 기반으로 추천을 수행하여 다양한 아이템 추천이 가능하다. 하지만 데이터가 부족한 초기 단계에서 데이터가 존재하지 않아 추천을 할 수 없는 상황인 콜드 스타트(Cold Start) 문제가 발생하고, 계산량이 비교적 많은 알고리즘이기 때문에 사용자 데이터가 많으면 오히려 계산 효율성이 저하되는 한계가 있다.

콘텐츠 기반 필터링은 아이템 자체의 특성을 바탕으로 추천을 수행하기 때문에 사용자 데이터가 적은 환경이나 새로운 사용자 유입이 많을 때 용이하다. 하지만 사용자가 이미 선호한 아이템과 유사한 콘텐츠만을 추천하기 때문에 추천의 다양성 측면에서 한계를 가지고 있다.

이러한 단점들을 극복하기 위해 하이브리드 추천 시스템(Hybrid Recommender System) [6]이 많이 연구되고 있다. 하이브리드 추천 시스템은 협업 필터링과 콘텐츠 기반 필터링을 결합해 서로의 단점을 보완하여 소규모 데이터셋 환경에서 더 나은 성능을 기대할 수 있다. 하지만, 하이브리드 방식이 협업 필터링이나 콘텐츠 기반 필터링 방식과 비교하여 실제로 소규모 데이터셋 환경에서 우수한 성능을 발휘하는지에 대한 검증이 부족한 상황이다.

본 연구는 소규모 데이터셋 환경에서 협업 필터링, 콘텐츠 기반 필터링, 그리고 하이브리드 추천 시스템의 성능을 비교하고, 하이브리드 방식이 소규모 데이터셋 환경에서 효과적인지 검증하고자 한다. 이를 위해 가중 평균(Weighted Average) 방식을 적용한 Soft Ensemble 기반의 하이브리드 모델을 설계하고, 가중치에 따른 추천 성능을 분석해보고자 한다.

## II. 본론

### 2.1 협업 필터링 구조

본 연구에서는 협업 필터링을 신경망 모델로 구현한다. 사용자 임베딩과 아이템 임베딩을 학습하고, 두 임베딩 간의 내적(dot product)값에 대해 Sigmoid 함수를 통해 0과 1사이의 값으로 반환하여 이진 분류를 할 수 있도록 모델 구조를 설계한다.

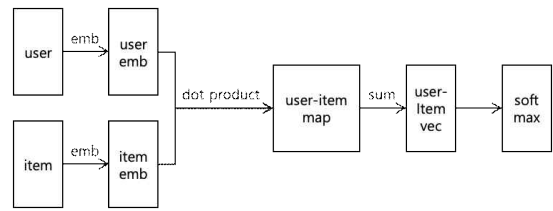


그림 1. 협업 필터링 신경망 모델 구조

### 2.2 콘텐츠 기반 필터링 구조

본 연구에서는 콘텐츠 기반 필터링을 신경망 모델로 구현한다. 사용자 임베딩과 아이템의 특징을 함께 입력받아 3개의 Fully connected layers들을 거친 후, 마지막 레이어에서 Sigmoid 함수를 통해 0과 1사이의 값으로 반환하여 이진 분류를 할 수 있도록 모델 구조를 설계한다.

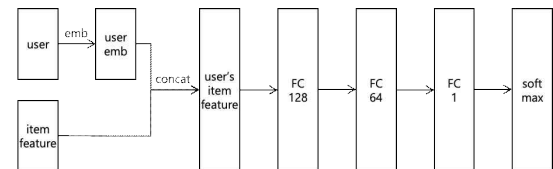


그림 2. 콘텐츠 기반 필터링 신경망 모델 구조

### 2.3 하이브리드 추천 시스템 구조

본 연구에서는 하이브리드 추천 시스템을 Soft Ensemble 모델로 구현한다. 협업 필터링 모델과 콘텐츠 기반 필터링 모델의 예측값을 가중 평균으로 결합하는 방식이다. 최종 예측값  $\hat{y}$ 은 협업 필터링 예측값  $\hat{y}_{CF}$ 과 콘텐츠 기반 필터링 예측값  $\hat{y}_{CB}$ 을 가중치  $\alpha$ 로 가중 평균하여 도출된다.

$$\hat{y} = \alpha \hat{y}_{CF} + (1 - \alpha) \hat{y}_{CB} \quad (1)$$

### III. 실험

#### 3.1 음악 선호도 데이터셋 수집

20대 대학생을 상대로 한국 최신 가요 30곡에 대한 선호도 조사를 진행했다. 음악을 들으면서 좋아요/싫어요(like/dislike)를 선택할 수 있는 설문조사 웹사이트를 배포하여 30곡에 대한 22명의 음악 선호도 데이터를 수집했다.

#### 3.2 음악 특징 추출

콘텐츠 기반 필터링에 사용하기 위해, 음악 30곡 각각에 대한 크로마(Chroma), RMS(Root Mean Square), 스펙트럼 중심(Spectral Centroid), 스펙트럼 대역폭(Spectral Bandwidth), 스펙트럼 롤오프(Spectral Rolloff), 영 교차율(Zero Crossing Rate), 고조파(Harmonic), 템포(Tempo), 멜-주파수 캡스트럼 계수(MFCC)를 추출하고 각 특징의 평균과 분산을 구해 음악 특징 데이터셋을 구축했다.

#### 3.3 실험 세팅

본 연구는 앞서 소개한 음악 선호도 및 음악 특징 데이터셋을 전처리하고, 훈련 및 테스트 데이터셋을 고정하여 각 모델을 평가한다. 데이터셋의 80%는 훈련 세트로서 학습에 활용되고, 나머지 20%는 테스트 세트로서 예측 성능을 평가하는데 활용된다. 협업 필터링 모델 학습에는 음악 선호도 데이터셋의 훈련 세트를 사용하고, 콘텐츠 기반 필터링과 하이브리드 모델 학습에는 음악 선호도 및 음악 특징 데이터셋의 훈련 세트를 사용한다. 하이브리드 모델에는 가중 평균을 위한 가중치 하이퍼파라미터  $\alpha = 0.05$ 로 설정했다.

#### 3.4 주요 결과

표 1은 음악 선호도 데이터셋에 대한 협업 필터링 모델(CF), 콘텐츠 기반 필터링(CB), 하이브리드 모델(Hybrid)의 성능을 나타낸다. 실험 결과, 하이브리드 모델이 가장 좋은 성능을 보였다. 협업 필터링과 콘텐츠 기반 필터링을 결합하여 사용자 선호도 예측의 정확도를 향상시킬 수 있었다. 다만, 정밀도(Precision)는 하이브리드 모델 보다 콘텐츠 기

반 필터링 모델이 더 높은 성능을 보이는데, 이는 모델이 추천한 아이템 중 사용자가 선호하는 아이템의 비중이 더 많다는 것을 의미하지만, 하이브리드 모델과 정밀도 차이가 크지 않고, 하이브리드 모델이 사용자가 선호할만한 아이템을 찾아주는 비율인 재현율이 더 높아 우수함을 알 수 있다.

표 1. 세 가지 모델의 성능 비교

Model	Music Preferences Dataset		
	Precision	Recall	F1-score
CF	0.561	0.372	0.448
CB	<b>0.697</b>	0.802	0.746
Hybrid	0.689	<b>0.849</b>	<b>0.760</b>

#### 3.5 Hyperparameter Sensitivity

표 2는 가중 평균의 가중치  $\alpha$  값에 따른 하이브리드 모델의 성능을 보여준다. 실험 결과,  $\alpha = 0.05$ 일 때 최적으로 사용자의 선호도를 예측할 수 있었다. 반면,  $\alpha$ 가 증가함에 따라 성능이 좋지 않았던 협업 필터링의 영향력이 증가해 사용자 선호도 예측 정확성이 감소하는 경향을 보였다. 이는 협업 필터링과 콘텐츠 기반 필터링을 결합할 때 콘텐츠 기반 필터링의 영향력을 크게 두되, 약간의 유저 및 아이템간의 상관관계를 고려하여 사용자 선호도 예측 정확도를 개선할 수 있음을 의미한다.

표 2.  $\alpha$ 에 따른 하이브리드 모델의 성능 비교

$\alpha$	Music Preferences Dataset		
	Precision	Recall	F1-score
0.01	0.683	0.802	0.738
0.05	<b>0.689</b>	<b>0.849</b>	<b>0.760</b>
0.1	0.683	0.826	0.747
0.5	0.646	0.488	0.556
0.9	0.569	0.384	0.458

### IV. 결론

본 연구에서는 설문 및 추출을 통해 구축한 음악 선호도 데이터셋, 음악 특징 데이터셋에서 단일 협업 필터링 모델, 콘텐츠 기반 모델보다 Soft Ensemble을 통해 두 모델을 결합한 하이브리드 모델이 추천 성능을 향상시킬 수 있음을 모델간 성능 비교 실험 및 하이퍼파라미터에 따른 하이브리드

모델 성능 비교 실험을 통해 검증하였다. 연구 결과, 추천시스템을 사용해야 하는 상황에서 적은 양의 데이터만이 존재할 때 제시 모델이 유용하게 활용될 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems", *Computer*, Vol. 42, No. 8, pp. 30-37, August 2009.
- [2] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets", *Proceedings of the Eighth IEEE International Conference on Data Mining*, pp. 263-272, December 2008.
- [3] B. MARLIN, "Collaborative Filtering: A Machine Learning Perspective", University of Toronto, pp. 2239-2239, 2004.
- [4] R. van Meteren and M. van Someren, "Using Content-Based Filtering for Recommendation", University of Amsterdam, 2000.
- [5] J. Son and S. Kim, "Content-based Filtering for Recommendation Systems Using Multiattribute Networks", *Expert Systems with Applications*, Vol. 89, pp. 404-412, December 2017.
- [6] G. Geetha, M. Safa, C. Fancy, and D. Saranya, "A Hybrid Approach Using Collaborative Filtering and Content-Based Filtering for Recommender System", *Journal of Physics: Conference Series*, vol. 1000, National Conference on Mathematical Techniques and its Applications (NCMTA 18), January 2018



# YOLO 기반 2차원 영상에서의 물체 거리 측정 방법 비교 분석

정윤걸\*, 최민정\*\*, 이영학\*\*\*

## Comparative Analysis of Object Distance Measurement Methods in 2D Images using YOLO

Yun-Geol Jung\*, Min-Jeong Choi\*\*, and Yeung\_Hak Lee\*\*\*

### 요 약

본 연구는 과수원의 유해조류를 효과적으로 관리하기 위한 반응형 방조장치 개발의 일환으로, 2D 영상에서 물체의 거리를 정확히 측정하는 방법을 탐구한다. 이를 위해 YOLOv8 모델과 focal length 기반의 거리 측정 기법을 활용하여, 다양한 기준 길이에 따른 거리 측정의 정확성을 비교 분석한다. 연구의 주요 내용은 물체 감지 및 거리 측정 방법 개발 그리고 다양한 기준 길이(높이, 너비, 대각선)를 적용한 거리 측정 방법의 성능을 비교하며, 실제 환경(교내 도로)에서의 거리 측정 정확도를 검증하였다. 실험은 교내 도로 환경에서 자동차를 대상으로 진행되었으며, 결과 분석을 통해 몇 가지 주요 발견을 확인하였다. 첫째, 모든 측정 방법이 허용 가능한 오차 범위 내에서 유의미한 거리 측정이 가능함을 입증하였다. 둘째, 대각선을 기준으로 한 방식이 물체의 방향 변화에 대해 높은 일관성을 보였다. 셋째, 높이와 너비를 기준으로 한 방식은 물체의 각도 변화에 따라 약간의 오차가 발생하는 것을 확인하였다.

### Abstract

This study develops a method for accurate distance measurement in 2D images using YOLOv8 and focal length-based techniques, aimed at creating a responsive bird deterrent system for orchards. Experiments conducted with automobiles in a campus environment compared various reference length approaches (diagonal, height, and width). Results showed all methods provide reliable distance estimates, with the diagonal-based approach demonstrating the highest consistency across object orientations. This research contributes to improving real-time distance estimation in orchard bird control systems and highlights the importance of appropriate reference length selection for specific environments.

### Key words

focal length, 2D image, distance, yolov8

---

\* 국립안동대학교 컴퓨터공학과, ykjung0917@kakao.com,

\*\* 국립안동대학교 컴퓨터공학과, chlalswjd4@naver.com, \*\*\*국립안동대학교 SW융합교육원, yhyi@anu.ac.kr(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)

## I. 서론

과수원에서 유해 조류로 인한 피해는 수확량 감소와 경제적 손실로 이어져, 이를 해결하기 위한 방조 장치의 필요성이 크다. 기존의 방조 장치는 일관된 패턴으로 작동하여 조류의 빠른 환경 적응으로 대응이 어렵지만, 최근 컴퓨터 비전 기술의 발전으로 실시간 조류 탐지가 가능해졌다. 특히 YOLO(You Only Look Once) 모델은 정확성과 속도에서 우수해, 유해 조류 탐지 및 퇴치 시스템에 적합하다.

본 연구는 YOLO 기반의 반응성 방조 장치에서 2D 영상으로 새의 거리를 측정하는 방법을 제안한다. 이를 위해 focal length 공식을 활용해 물체의 거리 측정 방법을 검증하고, 새 대신 자동차를 사용한 실험을 통해 높이, 너비, 대각선을 기준으로 한 거리 측정 방식의 정확성을 비교 분석할 것이다.

## II. 관련 연구

2D 이미지에서 물체의 거리를 측정하는 방법은 단일 카메라로 제한된 정보를 활용해야 하기 때문에 도전적인 문제로 여겨진다. 특히, 3D 정보를 얻을 수 없는 2D 영상에서는 물체의 크기나 위치를 기반으로 거리를 추정하는 것이 필수적이다. 이러한 문제를 해결하기 위해 focal length와 물체의 실제 크기를 활용하는 방식이 널리 연구되었다. [1] 이 방식은 단일 카메라로 2D 이미지를 획득한 뒤, 물체의 픽셀 상 크기와 카메라의 고정된 초점 거리를 이용해 거리를 계산한다.

유해 조류 탐지 시스템에서는 주로 3D 정보를 얻기 어려운 과수원 환경에서 단일 카메라를 활용한 접근법이 제시되고 있다. 특히 인공지능 기반 탐지 모델인 YOLO(You Only Look Once)는 2D 이미지에서 실시간으로 물체의 바운딩 박스를 생성하는 기능을 제공하며, 바운딩 박스의 크기를 이용해 물체와 카메라 간 거리를 추정할 수 있다. 기존 연구들에서는 YOLO 모델을 통해 조류나 해충을 실시간으로 탐지한 사례가 있으며, 본 연구에서는 이를 활용해 조류의 높이, 너비, 대각선을 기준으로 거리를 추정하고, 이 방식을 비교 분석하고자 한다.

## III. 이론적 배경

### 3.1 Focal Length

2D 영상에서 물체와 카메라 간의 거리를 측정하는 데 있어 가장 일반적으로 사용되는 방법 중 하나는 카메라의 focal length(초점 거리)와 물체의 실제 크기를 활용하는 방식이다. 이 방식은 카메라가 물체의 크기를 픽셀 단위로 측정한 값을 사용하여 거리를 계산한다. 기본적으로 focal length 공식은 다음과 같이 정의된다.

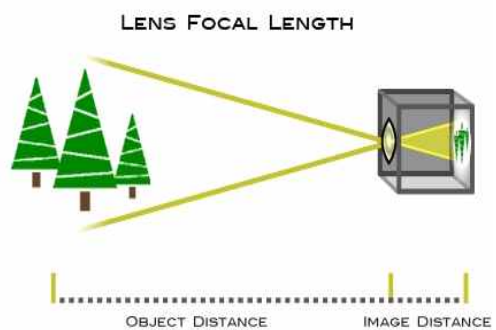


그림 1. 렌즈 초점거리의 예시 [2]  
Fig. 1. Example of Lens Focal Length [2]

그림 1에서 image distance가 focal length를 의미하며, 본 연구에서 구해야 할 값은 Object Distance이다.

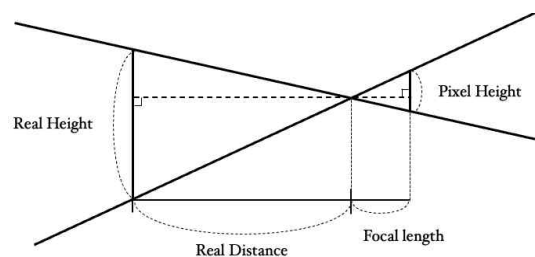


그림 3 Focal length 공식 유도 그림  
Fig. 2. Geometric Derivation of the Focal Length Formula

그림 2와 같이 두 변을 공유하는 직각삼각형은 맞꼭지각을 형성하여 두 삼각형은 닮음 상태를 취하고 있다. 이로써 각 삼각형의 길이 비율이 같다는 것을 알 수 있고, 따라서 Real Height(실제크기), Focal length(초점 거리), Pixel Height(픽셀 높이)를

알면 Real Distance(실제 거리)를 구할 수 있다.

$$\frac{RealHeight}{RealDistance} = \frac{PixelHeight}{Focal\ length}$$

$$RealDistance = \frac{Focal\ length \times Real\ Distance}{Pixel\ Height}$$

삼각형의 성질을 활용하여 첫 번째 식을 유도하고 이를 변형하여 Real Distance(실제 거리)를 구하는 두 번째 식으로 만들어 활용할 수 있다.

### 3.2 YOLO(You Only Look Once) 모델

YOLO 모델은 실시간 물체 탐지를 위해 널리 사용되는 딥러닝 기반의 모델이다. YOLO는 입력된 이미지에서 여러 물체를 한 번의 신경망 통과로 탐지하는 특징을 가지고 있어, 매우 빠른 속도와 높은 정확성을 자랑한다. YOLO는 이미지에서 물체를 탐지할 때, 바운딩 박스(bounding box)를 통해 물체의 위치를 표현하며, 바운딩 박스는 물체의 대각선, 높이, 너비 정보를 제공한다. 이 바운딩 박스의 크기는 본 연구에서 focal length 공식을 적용해 물체와 카메라 간 거리를 계산하는 데 중요한 역할을 한다.

YOLO 모델은 물체의 중심 좌표와 크기 정보를 제공하기 때문에, 본 연구에서는 탐지된 물체의 바운딩 박스를 활용하여 물체와 카메라 간 거리를 추정한다. 특히, 본 연구는 물체의 대각선, 높이, 너비를 각각 사용하여 거리를 계산하고, 이를 비교 분석하여 가장 정확한 방식을 도출하고자 한다.

## IV. 실험

### 4.1 실험 환경

본 실험은 교내 왕복 2차선 도로 중 한쪽 차선을 이용하여 진행되었다. 실험 구간의 시작점부터 일정한 거리 간격(처음 3m, 이후 5m 간격)으로 지면에 선을 그어 기준을 설정하였다. 이 기준선은 자동차가 정지하거나 통과할 때의 정확한 위치를 확인하기 위해 사용된다. 자동차는 실험 구간을 따라 서서히 주행하며, 자동차의 거리가 변화할 때마다 카메라로 영상을 촬영하였다.

실험은 햇빛이 일정한 낮 시간대에 진행되어 조명 조건이 최대한 일정하게 유지되었고, 실험 구간에 다른 물체가 방해하지 않도록 제한된 환경에서 촬영이 이루어졌다. 카메라는 도로의 한쪽 인도에 고정된 위치에 설치되었으며, 영상에서 자동차가 도로의 중앙선을 따라 다가오는 장면이 기록되었다.

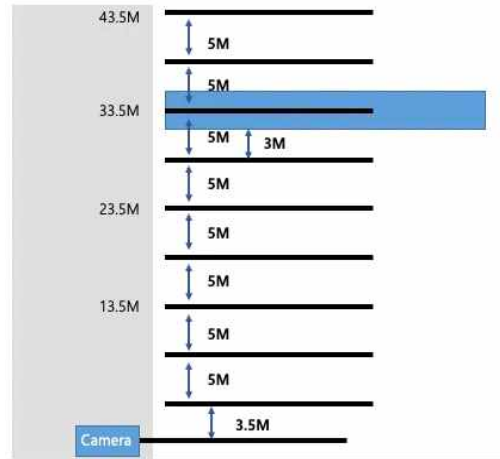


그림 4. 도로 표시선 간 거리  
Fig. 3. Distance Between Road Markings

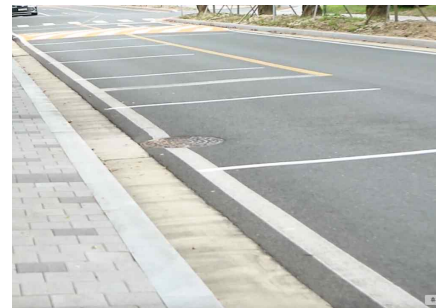


그림 4. 도로 위 표시선 예시  
Fig. 4. Examples of Road Markings

### 4.2 실험 장비

실험에 사용된 장비는 다음과 같다.

표 1. 실험 장비 세부 사항  
Table 1. Experiments Environment

카메라	Canon EOS M3 (focal length 29mm)
소프트웨어	python 3.12.3 , YOLOv8(pytorch)
pc 하드웨어	m1 MacBookPro
실험 대상	자동차 (TOYOTA AVALON)
기타	삼각대, 테이프 등

카메라는 29mm 초점 거리(focal length)를 가진 Canon EOS M3를 사용하였으며, 고정된 위치에서 도로의 차량 움직임을 2D 영상으로 기록하였다. YOLOv8 모델은 실시간 물체 탐지와 바운딩 박스 정보를 추출하기 위해 사용되었다.

### 4.3 실험 방법

실험 방법은 다음과 같다.

i. 기준 설정: 실험 구간의 시작점부터 일정한 거리(처음 3m, 이후 5m 간격)로 지면에 선을 그어 기준을 만들었다. 각각의 기준선은 자동차가 해당 지점에 도달할 때 촬영된 영상을 통해 거리를 확인하는 데 사용되었다.

ii. 자동차 주행 및 영상 촬영: 자동차는 출발지에서 기준점(카메라)으로 이동, 각 기준선 지점에 도달하는 모습을 영상에 담는다. 영상은 고정된 카메라 위치에서 동일한 화각으로 촬영되었다.

iii. YOLOv8.pt 모델을 통한 바운딩 박스 생성: 촬영된 영상에서 YOLOv8.pt 모델을 사용하여 자동차의 바운딩 박스를 추출하였다. 바운딩 박스는 자동차의 높이, 너비를 기준으로 생성되었으며, 이 정보는 focal length 공식을 적용해 물체와 카메라 간 거리를 계산하는 데 사용되었다.

iv. 거리 계산: focal length 의 공식을 활용하여 각 길이(높이, 너비, 대각선)를 적용하여 자동차의 거리를 계산한다.

v. 결과 분석: 각 방식(높이, 너비, 대각선)으로 계산된 거리를 실제 측정된 거리와 비교하고 분석하였다. 이를 통해 측정 방식의 차이점을 알 수 있다.

## V. 실험 결과

본 연구에서는 2D 영상 기반으로 focal length 공식을 활용하여 자동차와 카메라 간의 거리를 계산하는 방법을 실험적으로 분석했다. 높이, 너비, 대각선을 기준으로 한 거리 계산 결과를 실제 거리와 비교하여 각 방식의 정확도를 평가하였다. 실험 결과, 각 방식마다 고유한 장단점이 존재함을 확인할 수 있었다.

### 1) 높이를 기준으로 한 거리 계산

높이를 기준으로 한 방식은 물체의 높이가 변화

거나 자동차의 서스펜션이 눌리는 등의 높이 변화가 있을 때 다소 값이 편차가 큰 값을 보여준다. 이는 오차범위 내의 값이지만, 실험 대상의 성질에 따라 한계점으로 작용할 수 있다.

### 2) 너비를 기준으로 한 거리 계산

너비를 기준으로 한 계산은 상대적으로 정확한 값을 보여주었다. 이는 자동차라는 실험대상의 가로 너비값의 변화가 상대적으로 적기 때문이다. 하지만 고정된 카메라를 빗겨지나가는 실험 특성상 가까워질수록 객체를 대각선 방향으로 촬영하기 때문에 바운딩 박스의 너비와 값의 편차가 발생하였다. 하지만 이 역시 오차범위 내의 값으로 상황에 따라 무의미하다고 볼 수 있다.



그림 5. 실험 결과 1(실제 거리 약 40m)  
Fig. 5. Experimental resultt 1(Real Distance: about 40m)



그림 6. 실험 결과 2(실제 거리 약 28m)  
Fig. 6. Experiment resultt 2(Real Distance: about 28m)

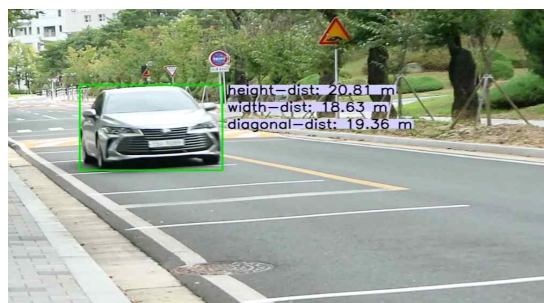


그림 7. 실험 결과 3(실제 거리 약 19m)  
Fig. 7. Experiment resultt 3(Real Distance: about 19m)



### 3) 대각선을 기준으로 한 거리 계산

대각선을 기준으로 계산한 방식은 물체의 전체적인 크기를 반영하여 비교적 안정적인 거리 추정값을 제공했다. 특히, 물체가 흔들리거나 예상과 다르게 움직일 때 대각선 계산은 다른 방식에 비해 오차율이 낮았다. 이는 대각선이 물체의 너비와 높이를 모두 포함하기 때문에, 물체의 크기 변화가 일정하지 않더라도 전체적인 크기 변동을 반영할 수 있기 때문이다.

## VI. 결 론

본 연구는 2D 영상에서 YOLOv8 모델을 활용하여 focal length를 통한 실제 거리 계산 방법을 제안하고, 물체의 높이, 너비, 대각선을 기준으로 한 세 가지 거리 측정 방법의 성능을 비교 분석하였다. 실험 결과, focal length를 통해 도출된 거리 측정값은 전반적으로 유의미한 정확도를 보였으며, 세 가지 측정 방법 모두 허용 가능한 오차 범위 내에서 일정 수준 이상의 정확한 거리 추정이 가능함을 확인하였다.

각 측정 방법의 성능은 특정 환경 조건에 따라 다소 차이를 보였다. 물체가 카메라와 수직 또는 정면에 가까운 경우, 높이나 너비를 기준으로 한 방식이 더 정확한 결과를 제공하는 경향을 보였다. 반면, 물체의 방향이 일정하지 않거나 변화가 심한 환경에서는 대각선을 기준으로 한 방식이 더 일관된 결과를 제공하였다. 이러한 결과는 각 방법이 서로 다른 상황에서 고유한 장점을 가지고 있음을 시사한다.

본 연구의 결과는 과수원 유해조류 방조 장치와 같은 실시간 거리 측정이 필요한 시스템 설계에 중요한 시사점을 제공한다. 특히, 대상 물체의 특성과 환경적 변수를 고려하여 최적의 거리 측정 방법을 선택하는 것이 시스템의 성능을 크게 향상시킬 수 있음을 보여준다. 예를 들어, 조류의 비행 패턴이 불규칙하거나 빠르게 변화하는 환경에서는 대각선 기반 측정 방식이 더 효과적일 수 있으며, 조류가 주로 일정한 방향으로 접근하는 환경에서는 높이나 너비 기반 측정 방식이 유용할 수 있다.

향후 연구에서는 다양한 기후, 조도, 속도, 객체의 불규칙적인 움직임 등을 고려한 종합적인 성능을 평가하고, 이를 보완하기 위한 보정 알고리즘을 개발하여 다양한 조건에서도 실제 거리가 측정되는지 평가가 필요하다.

## 참 고 문 헌

- [1] 김승기, 고영민, 배철균, 김대진, "단일 카메라를 이용한 3D 깊이 추정 방법," 방송공학회논문지, 제24권, 제2호, pp.281-291, 2019.
- [2] Nikon. (n.d.). Understanding focal length. Nikon USA. Retrieved October 10, 2024, from <https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/understanding-focal-length>(accessed by October 11, 2024)

# 인공지능을 사용한 가정용 전력 사용량 예측

신민혁\*, 김민호\*, 이세훈\*, 이상기\*, 강병권\*\*

## Predicting Home Power Usage Using Artificial Intelligence

Min Hyuck Shin\*, Min Ho Kim\*, Seh Un Lee\*, Sang Gi Lee\*, Beong-Gwon Kang\*\*

### 요약

본 연구에서는 2010년부터 2018년까지의 전력 데이터, 인구수, 평균 전력 요금을 이용해 LSTM(Long Short-Term Memory) 모델로 전력 소비량을 예측하는 방법을 제안하였다. 분석 결과, 전력 데이터와 인구수 간의 높은 상관관계를 확인하였으며, 이를 바탕으로 전력 소비량 예측에 필요한 입력 변수로 선택하여 모델을 학습시켰다. LSTM 모델은 시계열 데이터 처리에 효과적이며, 높은 예측 정확도를 달성하였다.

### Abstract

This study proposes a method for predicting electricity consumption using an LSTM (Long Short-Term Memory) model, leveraging electricity data, population statistics, and average electricity rates from 2010 to 2018. The analysis confirms a strong correlation among these variables, which were selected as input features for the model. The LSTM model, effective in processing time series data, achieved high prediction accuracy. The findings are expected to contribute to electricity consumption forecasting and support policy development in the energy sector.

### Key words

LSTM, electricity consumption, time series prediction, correlation analysis, forecasting

## 1. 서론

최근 소득증가, 기후변화 등 많은 요인으로 에너지 소비는 급격한 증가추세에 있다. 특히 전력소비는 계절적 요인 없이도 지속적으로 증가하고 있으며, 그중에서도 가정용 전력소비는 에너지 다소비 계층의 증가에 따라 꾸준히 증가하고 있다[1]. 이러한 현상은 지난해 러시아-우크라이나 전쟁으로 인해 에너지 시장이 매우 불안정해졌기 때문에[2], 전

력 소비 예측의 중요성을 더욱 부각시키며, 정확한 예측은 에너지 관리에 필수적인 요소 중 하나가 된다. 가정용 전력 소비량에 중요한 영향을 미치는 요소는 인구수와 평균 전력 요금의 변동이다. 최근 이러한 요소들을 인공지능 알고리즘에 적용하여 전력 사용량을 예측하는 모델들이 다수 출시되고 있다. 본 연구에서는 과거의 데이터 요소를 기반으로 미래 전력 사용량을 예측하는 인공지능을 제안한다. 특히, 기존의 인공지능과는 다르게 시계열을 다룰

\*순천향대학교 정보통신학과 학부과정 smh3140ko@sch.ac.kr

\*\*순천향대학교 정보통신공학과 교수 bgkang@sch.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 연구 결과로 수행되었음  
“(2021-0-01399)

수 있는 모델인 LSTM(Long Short-Term Memory)을 활용하여 적은 변수와 전처리로 정확도를 높이고, 모델 훈련 소요시간을 줄여 효율성 또한 향상시키는 것에 중점을 두었다.

## II. 본론

### II - 1. 데이터 수집 및 전처리

본 연구논문에서 사용된 데이터는 전력거래량, 인구수, 평균 전력 요금으로, 전력데이터 개방 포털 시스템에서 수집되었다. 2010년부터 2018년까지의 데이터는 학습데이터로 활용되었고, 2019년부터 2021년까지의 데이터는 테스트데이터로 활용되었다. 수집된 데이터에 대한 상관관계 분석을 수행한 결과, 가장 높은 상관관계를 나타내는 변수는 인구수로 확인되었으며, 이를 본 연구의 예측 변수로 포함하였다. [그림 1]은 이러한 상관관계 분석 결과를 시각적으로 나타내고 있다.

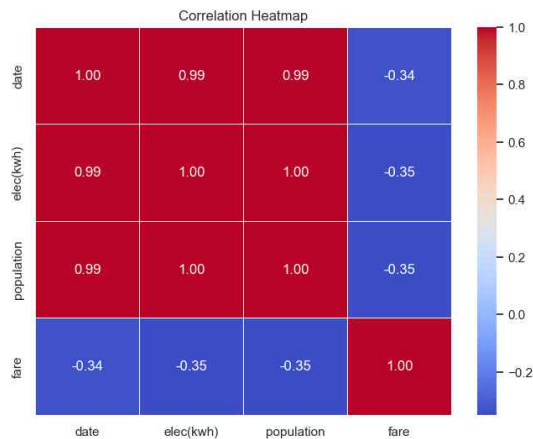


그림 1. 히트맵  
Fig. 1. Heatmap

데이터 간의 상관관계를 분석하고, 데이터를 서로 합친 후 전력값과 인구수 변수에 대한 scaling 전처리를 진행하였다. scaling 알고리즘은 정규화를 적용시켜 주었으며, feature\_range는 0.1로 처리하였다. 이 경우, 단위가 큰 전력과 인구수는 [표1.] 과 같이 0과 1사이의 값으로 조정되며 모델 학습 효과가 있기 때문에 효율성을 높일 수 있는 전처리중

하나이다.

표 1. scaling 된 데이터의 비교  
Table 1. Comparison scalped data

날짜	전력소비량	
	전	후
2010/01	12610307	0.00528171
2010/02	12624529	0.00979318
2010/03	12636677	0.01693697
2010/04	12655913	0.02336666
2010/05	12673226	0.03097943

### II - 2. 모델 선정 및 학습

전처리된 데이터를 학습하기에 앞서 데이터에 적합한 모델을 선정하였다. 본 논문에서 사용되는 데이터는 시계열 데이터이므로, 그에 맞게 가장 보편적으로 시계열 데이터에 사용할 수 있는 LSTM(Long Short-Term Memory) 모델과 구글에서 2017년에 발표한 모델인 트랜스포머(transformer) 모델을 사용하며 비교한다. MAPE(Mean Absolute Percentage Error)는 회귀모델의 학습 정도를 확인할 때 사용되는 평가 지표이다. MAPE는 다음과 같은 공식으로 정의된다.

$$MAPE = \frac{100}{n} \times \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \quad (1)$$

식(1)에서  $Y_i$ 는 실제 값이며  $\hat{Y}_i$ 는 예측값이다.

다음으로 손실률(Loss)은 실제 정답과 모델이 예측한 값 사이의 차이이다. 손실이 커질수록 오류가 커진다고 볼 수 있다. 본 논문에서 활용된 모델의 손실함수는 MSE(Mean Squared Error)를 사용하였다. MSE(Mean Squared Error, 평균제곱오차)는 이름에서 알 수 있듯이 오차를 제곱한 값의 평균이다. 이는 아래의 수식으로 정의된다.

$$MSE = \frac{1}{N} \sum_i^N (pred_i - target_i)^2 \quad (2)$$

식(2)에서 pred는 예측값이며 target은 실제값이다.

표 2. Transformer와 LSTM의 지표

Table 2. Transformer and LSTM Indicators

	LSTM	Transformer
MAPE	0.26%	0.0002
Loss	0.94%	0.0103

표2에서 실험 결과 각 모델의 정확도와 손실률을 최종적으로 나타낸다. 위 실험 결과 전체적인 결과는 Transformer보다 LSTM의 성능이 탁월함을 알 수 있고 최종적인 모델은 LSTM을 사용하여 구현하는 것을 제안한다. 결정된 LSTM 모델의 파라미터를 조정하며 [표 3.] 과 같이 실험을 진행하였다. [표 3.]의 MAPE는 각각의 기준으로 모델을 적용했을 때 결과이다.

표 3. 훈련 결과

Table 3. Result

실험	epochs	batch size	optimizer	MAPE
1	100	8	adam	0.35
2	100	8	SGD	0.06
3	100	16	adam	0.67
4	100	16	SGD	1.21
5	200	8	adam	0.38
6	200	8	SGD	0.08
7	200	16	adam	0.33
8	200	16	SGD	0.25

위 실험의 결과, 실험 2 가 가장 좋은 결과를 나타냄을 알 수 있었다. 이를 기반으로 일반화된 모델을 만들었고, 그에 따른 테스트 데이터인 2019년부터 2021년의 데이터를 예측하였다. 그에 따른 그래프는 [그림 2.]에 나타난다.

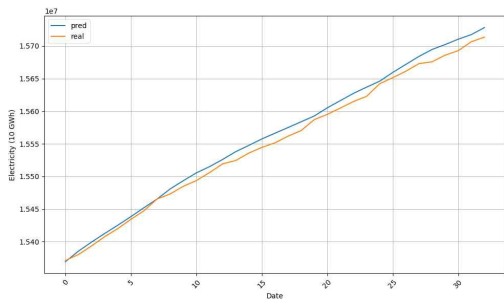


그림 2. 예측데이터와 실제데이터

Fig 2. Predicted and Real Data

### III. 결론 및 향후 과제

본 연구에서는 2010년부터 2018년까지 축적된 가정용 전력 소비 데이터를 학습용으로 사용하여, LSTM의 우수한 시계열 예측 성능을 통해 2019년 이후의 전력 소비량을 높은 정확도의 예측 결과를 얻을 수 있었으며, 이는 본 모델이 전력 소비 패턴을 성공적으로 학습했음을 보여준다. 이러한 예측 결과는 가정용 전력 수요 변동을 미리 파악함으로써 에너지 자원의 효율적 배분과 미래 전력 수급 정책에 기여할 수 있는 가능성을 가졌다. 특히 사계절이 존재하는 우리나라의 계절적 변동성이나 장기적인 트렌드를 예측하는데 있어 LSTM 모델의 우수한 가능성과 적합성을 확인하였다. 결과적으로 이러한 모델을 활용함으로써 향후 미래의 전력 소비를 예측·대비할 수 있다는 가능성을 보인다.

본 연구는 가정용 전력 소비 패턴에 한정하여 모델을 학습했으므로, 대한민국 전체 전력 소비 패턴을 포괄하는 데 한계가 있다. 보다 완전한 예측을 위해서는 상업용 및 산업용을 포함한 다양한 계약 유형별 전력 소비 패턴을 학습한 모델이 추가적으로 필요하다. 또한, 본 연구에서 구축된 모델은 인구수 및 과거 전력 소비 패턴에 크게 의존하고 있어, 전쟁이나 질병과 같은 예외적 사건에 대한 예측 탄력성이 낮다는 한계를 가진다. 이러한 한계는 비정상적 상황에서 모델의 예측 성능 저하를 야기할 수 있다. 더불어, 본 연구에서 사용된 모델은 월 단위 예측을 수행하므로, 전력 소모의 실시간 변동에 대한 즉각적 대응이 어려운 한계가 있다. 향후 연구에서는 예측 단위를 주 단위 또는 일 단위로 세분화하고 외부 충격 변수들을 추가함으로써 전력 소모 패턴의 미세한 변동을 보다 정확히 반영하고자 한다. 이를 통해 피크 전력을 기저 전력으로 대체하거나 피크 전력 부족에 사전 대비하는 효과를 기대할 수 있으며, 궁극적으로 전력 수급의 안정성 제고와 효율적인 에너지 관리 전략 수립에 기여할 수 있을 것으로 사료된다.

참 고 문 헌

- [1] 김민경. (2013). “서울시 가정용 전력소비의 변화요인과 저감방안. “ 정책리포트,(149), 1-20.
- [2] 김창식, 김영민. (2023). 에너지위기 대응과 국내 수요관리 정책 현황. 전기저널,, 22-27.

# YOLO 시리즈 비교 분석

김기욱\*, 이영학\*\*

## Comparative Analysis of YOLO Series

Gi-Uk Kim\* and Yeung-Hak Lee\*\*

### 요 약

YOLO 시리즈는 실시간 객체 탐지 기술의 지속적인 혁신을 통해 성능과 효율성을 크게 향상시켜 왔다. 본 연구는 YOLOv8 이후 시리즈에 대해 분석한다. YOLOv8은 객체감지 성능을 높이기 위해 개선된 손실 함수와 동적 어텐션 메커니즘을 도입했고, YOLOv9은 GELAN과 PGI 같은 최신 기법을 사용하여 높은 정확도를 유지하면서 연산량을 줄였다. YOLOv10은 효율성에 중점을 두어 더 빠른 추론 속도와 뛰어난 객체 인식 성능을 제공했으며, YOLOv11은 고급 데이터 증강 기법을 통해 작은 객체 인식 성능을 더욱 강화했다. 이러한 발전은 자율주행, 드론 감시 등 다양한 산업 분야에서 실시간 객체 탐지의 활용 가능성을 확장하고, 앞으로도 기술혁신을 이끌어갈 것으로 기대된다.

### Abstract

The YOLO series has greatly improved performance and efficiency through continuous innovation in real-time object detection technology. This study analyzes the series after YOLOv8. YOLOv8 introduced an improved loss function and dynamic attention mechanism to improve object detection performance, and YOLOv9 reduced the amount of computation while maintaining high accuracy by using the latest techniques such as GELAN and PGI. YOLOv10 focused on efficiency, providing faster inference speed and excellent object recognition performance, and YOLOv11 further enhanced small object recognition performance through advanced data augmentation techniques. These advancements are expected to expand the usability of real-time object detection in various industrial fields such as autonomous driving and drone surveillance, and continue to lead technological innovation in the future.

### Key words

yolo series, artificial intelligent, GELAN, PGI

## 1. 서 론

YOLO 시리즈는 실시간 객체 탐지 분야에서 선도적인 역할을 해왔으며, 버전이 거듭될수록 성능과

효율성이 크게 향상되었다. 각 버전은 기존의 한계를 극복하고 새로운 기술을 통합하며 발전해 왔다.

YOLOv8은 피쳐 추출 성능을 최적화하고 다양한 스케일의 객체 감지 성능을 높이기 위해 개선된

\* 국립안동대학교 사학과, giuk030505@naver.com,

\*\* 국립안동대학교 SW융합교육원, yhyi@anu.ac.kr(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)

손실함수와 동적 어텐션 메커니즘을 도입했다.

YOLOv9는 YOLOv8의 성과를 바탕으로 Generalized Efficient Layer Aggregation Network (GELAN)와 Programmable Gradient Information(PGI) 같은 최신 기법을 채택해 연산량을 줄이면서도 더 높은 정확도를 구현했다.

YOLOv10은 효율성 중심의 모델 설계를 통해 더 빠른 추론 속도와 뛰어난 객체인식 성능을 자랑했다.

마지막으로, YOLOv11은 YOLOv10의 개선 사항을 바탕으로 작은 객체 인식 성능을 한층 더 강화하고, 빠른 속도를 제공한다. 이와 같은 발전은 객체 탐지의 정확성과 속도를 높이며, 여러 산업 분야에서 실시간 응용의 가능성을 확장하고 있다.

## II. 본 론

YOLO 시리즈는 실시간 객체 탐지 분야에서 지속적인 혁신을 통해 중요한 발전을 이루어왔다. YOLOv8은 CSPDarknet53의 수정 버전을 사용한다. 5번의 다운샘플링 과정을 통해 5개의 특징 맵을 생성합니다. 모델의 피쳐 추출 성능을 최적화하기 위해 CSP 모듈 대신 2Cf모듈을 채택했다. 객체 감지 성능 향상을 위해 기존의 CloU 대신 EloU와 WIoU와 같은 개선된 손실 함수를 적용하였으며 복잡한 배경과 작은 객체가 많은 환경에서 높은 성능을 발휘할 수 있도록 BiFormer라는 동적 회소 어텐션 메커니즘을 도입해 중요한 정보를 선택하고 불필요한 계산 비용을 줄인다. 또한 다양한 스케일의 객체를 감지하기 위해 FFNB(Feature Fusion Network Block)를 사용해 두 개의 추가 탐지 스케일을 제공하고, 얇은 정보와 깊은 정보를 융합해 감지 성능을 극대화한다. 계산 비용 절감을 위해 PConv(부분 합성곱) 등을 활용하여 효율적인 특징 추출을 실현했다[1].

YOLOv9는 YOLOv8의 성과를 기반으로 Generalized Efficient Layer Aggregation Network(GELAN)와 Programmable Gradient Information(PGI) 같은 최신 기법을 적용하여 연산량을 줄이면서도 높은 정확도와 빠른 속도를 구현했으며 적은 매개변수로 높은 성능을 달성할 수 있다. 또한 얇은 모델에서도 효과적인 학습을 가능하게 하여 경량 모델의 성능을 향상

시켰다. 이 버전은 다양한 응용 분야에 맞춘 모델 변형을 제공하여 모바일 기기에서의 사용성을 강화했다. 이러한 혁신은 객체 감지뿐만 아니라 다양한 컴퓨터 비전 작업에서의 활용을 가능하게 했다[2].

YOLOv10은 일대다 할당과 일대일 할당을 결합하는 이중 레이블 할당 전략을 사용하여 NMS(No NMS Learning) 없는 학습과 효율성-정확도 중심 모델 설계를 통해 실시간 객체 탐지 분야에서 새로운 표준을 제시한다. 또한, 분류에 필요한 계산량을 줄여 전체 모델의 효율성을 높이고 다운샘플링 과정에서 계산비용을 줄이고 정보 손실을 최소화한다. 그리고 모델의 각 부분에 적합한 블록을 설계하여 불필요한 계산을 줄이며 수용영역을 넓혀 모델의 표현력을 향상시켰고 전역 정보를 효과적으로 활용하여 성능을 향상시킴으로써 계산량을 줄이고 정확도를 높여 모델 성능을 최적화 했다[3].

마지막으로, YOLOv11은 뛰어난 정확도를 입증하고 복잡한 배경과 다양한 객체 크기에 대한 견고한 성능을 보여주어 실제 환경에서의 활용 가능성을 높였다. 또한 YOLOv11은 실시간 객체탐지에 필수적인 빠른 속도를 제공하여 자율 주행, 로봇 공학 등 실시간 처리가 요구되는 분야에서 유용하게 활용될 수 있다. 그리고 다양한 크기와 종류의 객체를 정확하게 탐지할 수 있는 능력이 강화되어 작은 객체나 밀집된 객체에 대한 탐지능력을 향상시켰으며 농업, 의료 등 다양한 분야에서 활용 가능하다. YOLOv11은 안정적인 모델 아키텍처를 기반으로 하여 다양한 환경변화에도 견고한 성능을 유지하며 메모리 사용량을 최적화하여 다양한 환경에서 실행될 수 있도록 설계되었다[4].

## III. 결 론

YOLO 시리즈는 실시간 객체 탐지 기술에서 지속적으로 혁신을 이뤄내며, 각 버전이 발전할수록 0 성능과 효율성이 크게 향상되었다. YOLOv8은 효율적인 특징 추출을 위해 계산 비용을 절감하는 기술을 적용했으며, YOLOv9는 GELAN과 PGI 같은 최신 기법을 적용하여 높은 정확도를 유지하면서도 연산량을 줄였다. YOLOv10은 효율성에 중점을 두

어 빠른 추론 속도와 뛰어난 객체 인식 성능을 제공하며, YOLOv11은 작은 객체 인식 성능을 한층 강화했다.

이러한 발전들은 객체 탐지의 정확성과 속도를 크게 향상시킬 뿐만 아니라, 다양한 산업 분야에서 실시간 응용의 가능성을 넓히는 데 기여하고 있다. 특히, 자율주행차, 드론 감시, 보안 시스템 등 다양한 분야에서 YOLO 시리즈의 활용 가능성이 더욱 확대되고 있다. 앞으로도 YOLO 시리즈는 객체 탐지 기술의 혁신을 지속적으로 이끌어갈 것으로 기대되며, 이로 인해 실시간 객체 탐지 기술의 새로운 지평이 열릴 것이다. 이러한 혁신은 연구자와 개발자에게 새로운 기회를 제공하고, 더욱 정교하고 효과적인 시스템 개발에 기여할 것이다.

## 참 고 문 헌

- [1] G. Wang, Y. Chen, P. An, H. Hong, J. Hu and T. Huang, "UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios", AUG. 2023.
- [2] C. Y. Wang, I. H. Yeh, H. Y. M. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information", Feb. 2024.
- [3] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, G. Ding, "YOLOv10: Real-Time End-to-End Object Detection", arXiv:2405.14458, 23 May 2024.
- [4] R. Sapkota, Z. Mengb, M. Churuvijaa, X. Dub, Z. Ma, M. Karkee, "Comprehensive Performance Evaluation of YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments", DOI:10.48550/arXiv.2407.12040, July 2024.



# DCRNN 및 LogSparse Attention 기반 수화 인식 트랜스포머 모델

김나은\*, 최해영\*\*, 강창구\*\*\*

## DCRNN and LogSparse Attention-Based Sign Language Recognition Transformer Model

Naeun Kim\*, HaeYeong Choe\*\*, and Changgu Kang\*\*\*

### 요 약

본 연구에서는 DCRNN과 LogSparse Attention을 활용한 Transformer 모델을 통해 수화 인식 성능을 향상시키는 방안을 제안한다. 제안된 모델은 그래프 구조를 활용하여 손가락 모양을 고려하고, 왼손과 오른손의 특징을 독립적으로 추출하여 이를 효과적으로 통합함으로써 인식 성능을 개선하고자 한다. 본 연구에서는 양손과 거리 노드 간의 관계를 반영하여 DCRNN을 수행한 후, 각각의 손에 대해 LogSparse Attention을 적용하여 특징을 추출한다. 계층별 k-fold 교차 검증을 통해 k를 5로 설정하고 10 epochs 동안 학습을 진행하였다. 최종적으로 테스트 데이터에서 F1 스코어 0.94를 기록하여 높은 성능을 보여주었다. 이 연구는 DCRNN과 LogSparse Attention을 결합한 Transformer 모델이 다양한 수화 동작의 복잡한 인식을 효과적으로 개선할 수 있음을 보여준다.

### Abstract

This study proposes a method to enhance sign language recognition performance using a Transformer model that incorporates DCRNN and LogSparse Attention. The proposed model utilizes a graph structure to consider finger shapes and aims to improve recognition performance by independently extracting features from both the left and right hands and effectively integrating them. In this research, DCRNN is applied to reflect the relationship between both hands and the distance node, followed by the application of LogSparse Attention to extract features for each hand. A hierarchical k-fold cross-validation was conducted with k=5 and training was performed over 10 epochs. Ultimately, the model achieved an F1 score of 0.94 on the test data, demonstrating high performance. This research shows that the Transformer model combining DCRNN and LogSparse Attention can effectively improve the recognition of complex sign language movements.

### Key words

Sign language, DCRNN, Log sparse transformer

---

\* 경상국립대학교 컴퓨터공학과, nelove288@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, chyong118@gnu.ac.kr

\*\*\* 경상국립대학교 컴퓨터공학과, cgk@gnu.ac.kr(교신저자)

## 1. 서 론

현대 사회에서 사람들은 일상적으로 타인과 의사소통을 하며 살아가고, 이 과정에서 가장 주된 수단으로 음성언어가 사용된다. 그러나 청각 장애인에게는 이러한 음성언어가 의사소통의 장벽으로 작용하며, 그로 인해 수화와 같은 비음성 언어가 중요한 대안이 된다. 최근 IT 기술의 발전과 함께 수화 인식 기술에 대한 연구가 활발히 이루어지고 있으며, 이러한 연구들은 주로 딥러닝을 활용한 모델들을 기반으로 하고 있다.[1][2] 특히, 순환 신경망 계열의 모델들은 일정 수준의 성과를 거두었으나, 복잡한 손동작을 정확하게 인식하는 데에는 한계가 있었다.

본 연구에서는 이러한 한계를 극복하기 위해서 DCRNN(Diffusion Convolutional Recurrent Neural Network) 모델[3]과 LogSparse Attention[4]을 활용한 트랜스포머 기반 모델을 제안하여 수화 인식 성능을 개선하고자 한다. 기존 방법들과 달리, 본 연구는 왼손과 오른손의 특징을 각각 독립적으로 추출한 후, 그래프 형태로 각 특징을 할당하여 보다 정교한 특징을 학습한다. 이후, 독립적으로 추출된 두 손의 특징을 통합하는 과정을 통해 복잡한 손동작까지도 효과적으로 인식할 수 있도록 모델을 설계하였다.

이와 같은 접근 방식은 수화 동작의 복잡성을 처리하는 데 있어 높은 잠재력을 지니며, 궁극적으로 청각 장애인과 비장애인 간의 원활한 의사소통을 돕는 기술 발전에 기여하고자 한다.

## II. 수화 인식을 위한 트랜스포머 모델

### 2.1 전처리

본 연구에서는 림모션(Leap Motion) 장치를 통해 획득한 총 174개의 특징 값을 전처리하여 학습 및 평가에 활용하였다. 먼저, 손바닥, 손가락 관절, 손가락 끝의 쿼터니언 값 168개는 0에서 1 사이의 값으로, 그대로 모델 학습의 입력 특징으로 사용하였다. 또한, 손바닥의 위치값 6개는 두 손 사이의 거

리를 계산하는 데 사용되었으며, 두 손 사이의 거리는 유클리디안 거리(Euclidean distance)를 이용하여 산출되었다. 이 계산된 거리를 새로운 특징으로 추가한 반면, 기존의 위치값 6개는 제거하였다. 최종적으로 총 169개의 특징 값을 모델의 입력으로 사용하였으며, 이 값들은 그래프 형태의 데이터를 구성하기 위해 손 및 거리와 일치하는 각각의 노드에 할당하였다. 이러한 전처리 과정을 통해 모델의 인식률을 향상시키고자 하였다.

### 2.2 모델 구성 및 아키텍처

DCRNN 모델과 LogSparse Attention을 활용하여 수화 인식을 위한 트랜스포머 모델을 구성하였다. 모델의 전체적인 구성은 그림 1과 같다.

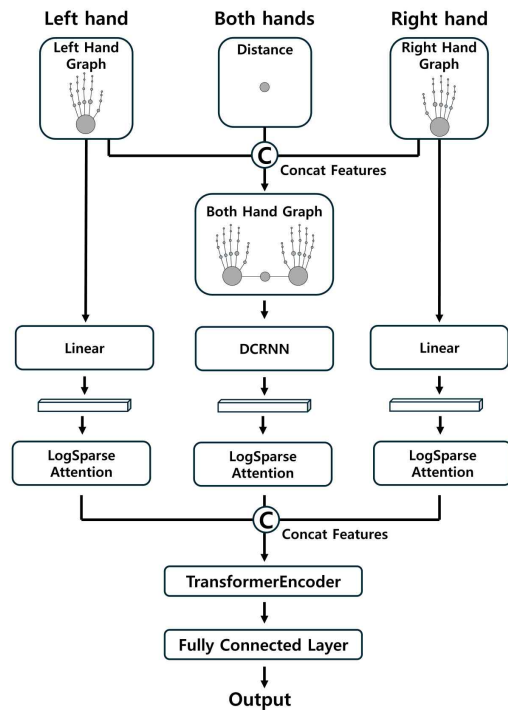


그림 1. 트랜스포머 모델 구성  
Fig. 1. Transformer Model Configuration

본 모델은 크게 세 가지 모듈로 구성된다. 첫 번째 모듈은 양손의 그래프 구조를 입력 데이터로 사용하는 DCRNN이며, 두 번째 모듈은 각 특징 벡터에 LogSparse Attention을 적용하는 모듈이다. 세 번째 모듈은 이들 특징 벡터를 연결하여 Transformer

Encoder에 전달하는 과정으로 구성된다. 입력 데이터인 왼손과 오른손의 그래프 구조를 기반으로 양손의 상호작용 그래프를 생성하여, 수화 인식에서 중요한 특징을 학습한다.

먼저, DCRNN은 그래프 구조 상에서 시간적 종속성을 학습할 수 있는 네트워크로, 손의 움직임 패턴을 시계열적으로 추출하는 데 효과적이다. 양손 및 거리 그래프를 DCRNN에 입력하여 시간에 따른 두 손의 상호작용을 학습하고, 양손의 움직임을 통해 수화 인식에 필요한 중요한 정보를 얻는다.

두 번째로, 왼손과 오른손의 특징 벡터를 Linear 레이어를 거쳐 DCRNN 출력과 동일한 차원으로 맞춘 후, 왼손, 오른손, 그리고 상호작용 특징 벡터의 세 가지를 LogSparse Attention 메커니즘에 입력한다. LogSparse Attention은 입력 시퀀스 길이가 길어도 중요한 부분에 더 집중할 수 있도록 설계된 메커니즘으로, 이를 통해 모델이 수화 동작에서 핵심적인 손동작을 효과적으로 인식할 수 있게 한다.

마지막 단계에서는, DCRNN과 LogSparse Attention을 통해 추출된 각 특징 벡터를 연결하여 Transformer Encoder에 전달한다. Transformer Encoder는 다층 구조로 되어 있어 시간적 종속성과 복잡한 공간적 패턴을 학습하는 데 강점을 지닌다. Encoder의 출력을 통해 수화 동작 시퀀스에서 중요한 특징을 학습하며, 최종적으로 각 시퀀스의 마지막 시간 단계에서의 출력 벡터를 완전 연결 레이어에 전달하여 수화 동작을 분류한다.

### 2.3 학습 및 평가

학습은 에포크 10, 배치 사이즈 64 학습률 0.001로 설정하여 학습 및 검증 그리고 평가까지 진행하였다. 학습 중 과적합을 방지하기 위해 드롭아웃은 0.5로, 계층별 k-fold 교차 검증의 k는 5로 적용하여서 총 에포크 50이 될 수 있도록 하였다. 그림 2와 그림 3은 각각 학습 및 검증의 손실과 정확도의 결과를 나타낸다.

모델의 평가를 위해 림모션 장치를 활용하여 수화별 20개를 수집한 후 총 160개로 평가를 진행하였다. 평가의 F1스코어는 0.94으로 높은 성능을 보였다. 테스트의 혼동행렬은 그림 4와 같다.

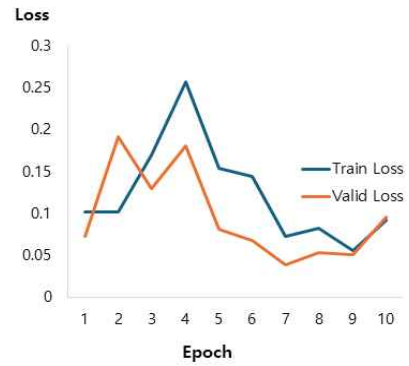


그림 2. DCRNN & LogSparse Attention Transformer 모델에 대한 손실 그래프  
Fig. 2. Loss Graphs for the DCRNN & LogSparse Attention Transformer Model

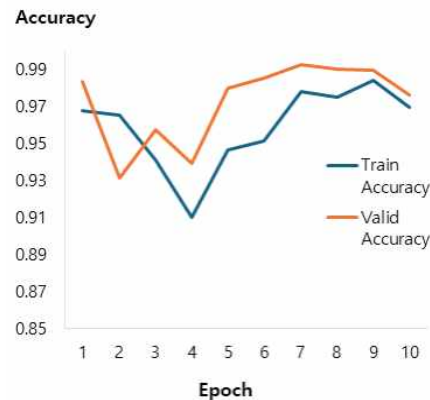


그림 3. DCRNN & LogSparse Attention Transformer 모델에 대한 정확도 그래프  
Fig. 3. Accuracy Graphs for the DCRNN & LogSparse Attention Transformer Model

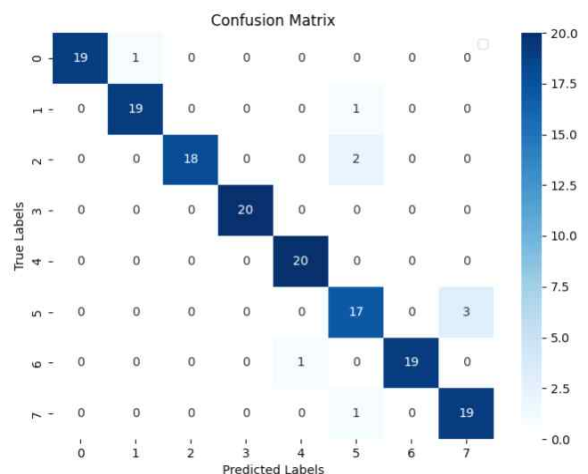


그림 4. DCRNN & LogSparse Attention Transformer 모델의 혼동행렬  
Fig. 4. Confusion Matrix for the DCRNN & LogSparse Attention Transformer Model

### III. 결 론

본 연구에서는 DCRNN & LogSparse Attention Transformer 모델을 활용하여 수화 인식을 개선하였다. 왼손과 오른손의 특징을 독립적으로 추출하고 양손을 그래프 형식으로 구성하여 복잡한 손동작을 효과적으로 인식하였으며, F1 스코어 0.94을 기록하여 높은 성능을 보였다. 이 모델은 다양한 수화 인식에 기여할 것으로 기대된다.

### 참 고 문 헌

- [1] Liu, T., Zhou, W., and Li, H. "Sign language recognition with long short-term memory," in 2016 IEEE International Conference on Image Processing (ICIP), IEEE, 2016.
- [2] Mittal, A., et al. "A modified LSTM model for continuous sign language recognition using leap motion," IEEE Sensors Journal, Vol. 19, No. 16, pp. 7056-7063, August, 2019.
- [3] Li, Y., et al. "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," arXiv preprint, arXiv:1707.01926, 2017.
- [4] Li, S., et al. "Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting," Advances in Neural Information Processing Systems, Vol. 32, 2019.

# YOLO v8과 YOLO v10을 사용한 사용자 표정 분석 알고리즘 성능 비교

김태형\*, 김주현\*, 백경호\*\*, 강병권\*\*\*

## Comparative Performance Analysis of Emotion Recognition Algorithm using YOLO v8 and YOLO v10

Tae-Hyeong Kim\*, Ju-Hyun Kim\*, Kyung-Ho Paik\*\*, and Byeong-Gwon Kang\*\*\*

### 요약

본 연구에서는 다양한 환경에서 사용자의 표정을 분석하는 표정 인식 알고리즘을 YOLO v8과 YOLO v10의 성능을 비교를 제시하였다. 구현한 알고리즘은 YOLO v8과 YOLO v10 모델을 사용하여 사용자의 표정을 파악한다. 목표는 시각 데이터를 기반으로 인간의 표정을 정확하게 인식하는 데 있어 두 모델의 성능을 평가하고 비교하는 것이다. 사용자 표정에 맞는 데이터 추출을 통해 높은 정확도와 처리 속도를 보여준다. 모델 훈련과 정확도 결과는 결론에서 제시된다.

### Abstract

This study presents a performance comparison of emotion recognition algorithms using YOLO v8 and YOLO v10 in various environments. The implemented algorithm utilizes YOLO v8 and YOLO v10 models to recognize user emotions. The goal is to evaluate and compare the performance of the two models in accurately recognizing human emotions based on visual data. The system demonstrates high accuracy and processing speed through emotion-based data extraction. The model training and accuracy results are presented in the conclusion.

### Key words

Performance comparison, Emotion recognition, YOLO v8, YOLO v10, High accuracy

### 1. 서론

최근 들어 인간과 컴퓨터 사이의 상호작용이 급증함에 따라 여러 분야에서 정확하고 효율적인 표정 인식 시스템이 중요한 역할을 하고 있다. 또한,

이미지 인식과 분류 문제들을 해결하는 딥러닝이 큰 발전을 하고 있다. 이러한 필요성을 충족하기 위해 YOLO(You Only Look Once) 와 같은 딥러닝 기반 객체 탐지 모델을 사용한다. YOLO는 탐지와

\* 순천향대학교 정보통신공학과 학부과정, dony615@naver.com, wnguss@naver.com

\*\* 순천향대학교 전자공학과 겸임교수, pkhtow@naver.com

\*\*\* 순천향대학교 정보통신공학과 교수, bgkang@sch.ac.kr

※ 이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0012724, 2024년 산업혁신인재성장지원사업)

분류 작업을 단일 신경망에서 수행하는 단일 통합 모델, 실시간 객체 탐지, 위치 및 분류 예측 등의 역할을 하고 있다.

개인 맞춤 추천 프로그램이 많아지는 가운데 사용자의 표정을 통해 기분을 인지하고 분석한 표정 인식 프로그램을 만들었다. 그 과정 중 YOLO v8을 이용해 표정 데이터 학습을 진행했다. YOLO는 높은 속도를 위해 정확도가 떨어지는 단점이 있다. YOLO v8보다 더 좋은 성능을 가지는 YOLO v10과 성능을 비교해 정확도가 얼마나 더 향상되는지 비교해보고자 이 연구를 진행하게 되었다.

본 연구는 알고리즘의 정확도를 향상하기 위해 최신 YOLO 모델인 YOLO v8과 YOLO v10의 성능을 비교하고자 한다. 두 모델의 성능을 평가하여 개인 맞춤 프로그램에서 중요한 정확도를 비교 분석한다. 또한, 비교 분석을 통해서 표정 인식의 높은 정확도와 효율성을 달성하기 위해 시각적 단서로부터 의미 있는 데이터를 추출하는 과정의 중요성을 강조한다. 모델 학습과 성능 평가 결과를 통해 두 모델이 표정 인식에서 어떤 성능을 발휘하는지 확인하고, 이러한 결과가 향후 표정 기반 시스템의 발전에 이바지할 가능성을 제시하고자 한다.

## II. YOLO v8과 YOLO v10을 활용한 표정 인식 성능 비교 분석

본 논문에서는 YOLO v8과 YOLO v10을 활용하여 표정 인식의 성능을 비교 분석한다. 딥러닝 기반 객체 탐지 알고리즘인 YOLO v8을 사용하여 표정에 대한 데이터들을 학습했다.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

위 수식은 모델 성능 평가를 위한 정확도 (Accuracy)를 계산한 수식이다.

[표 1]을 보면 다양한 에포크(epochs)와 배치 크기(Batch Size) 설정을 사용하여 모델을 학습했다. 정확도 결과는 다음과 같다. 에포크 100, 배치 크기 16의 경우 정확도는 0.612로, 모든 실험 설정 중 가

장 높은 성능을 보였다. 이를 통해, 충분한 학습을 반복해 데이터를 효과적으로 학습할 수 있음을 보여준다. 대조적으로, 배치 크기를 32로 늘린 경우, 정확도는 0.542로 감소하는 것을 볼 수 있다. 이것은 배치 크기가 커질수록 모델이 세부적인 패턴을 인식하는 데 어려움을 겪을 수 있음을 내포한다. 대체로, 에포크 수가 증가할수록 정확도가 향상되는 경향이 있었으며, 배치 크기가 작을수록 모델이 더 정밀하게 학습할 수 있었다. 하이퍼파라미터 설정은 실시간 표정 인식 시스템에서 정확도를 높이는 최적의 설정으로 보인다.

표 1. YOLO v8을 사용한 Epoch와 Batch 값에 따른 표정 인식률

Table 1. Expression recognition rate according to Epoch and Batch values Using YOLO v8

YOLO V8		
Epoch	Batch	accuracy
100	16	0.612
100	32	0.542
70	16	0.486
70	32	0.494

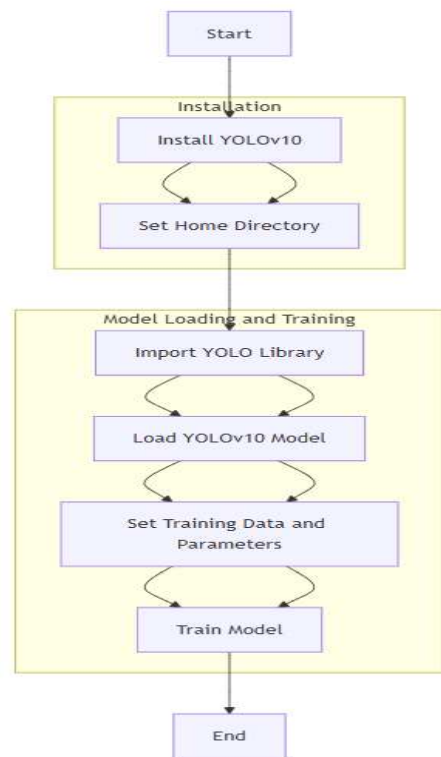


그림 1. YOLO v10에 대한 학습 흐름도  
Fig 1. Learning Flowchart for YOLO v10

[그림 1]은 YOLO v10을 사용하여 학습을 진행하는 흐름도이다. 이 흐름도는 YOLO v10 모델 설치부터 학습까지의 과정을 나타낸다. 설치 단계에서는 YOLO v10 설치와 홈 디렉터리 설정을 포함하며, 이후 모델 로드와 학습 단계에서는 YOLO 라이브러리를 가져오고 데이터를 설정하여 모델을 학습한다.

표 2. YOLO v10을 사용한 Epoch와 Batch 값에 따른 표정 인식률

Table 2. Using YOLO v10 Expression recognition rate according to Epoch and Batch values

YOLO V10		
Epoch	Batch	accuracy
100	16	0.544
100	32	0.642
200	32	0.661
200	16	0.673

위의 [표 2]는 YOLO v10을 사용한 정확도를 나타낸다. 학습한 결과, 에포크 100, 배치 크기 16의 경우 정확도가 0.544가 나왔다. 이후 배치 크기를 32로 증가시켜 정확도가 0.642로 향상된 것을 알 수 있다. 이를 통해 배치 크기의 증가가 더욱 향상된 정확도를 제공할 수 있다는 것을 확인했다. 그러므로 에포크를 더 좋은 성능을 위해 에포크를 200으로 늘리고 배치 크기를 유지하여 학습한 결과, 정확도는 0.661이 나왔다. 배치 크기가 작을수록 모델이 데이터를 더 정밀하게 학습하고, 과적합을 줄이는데 도움이 될 수 있기에 배치 크기를 16으로 줄여 학습했다. 그 결과 에포크 200, 배치 크기 16의 설정이 더 나은 정확도 성능을 보였다.

[그림 2]는 혼동 행렬(Confusion Matrix)과 정밀도-신뢰도 곡선(Precision-Confidence Curve)을 나타낸다. 혼동 행렬을 보면 Happy 클래스에서 130건, Anger 클래스에서는 86건의 정확한 예측이 있었음을 보여주고, Fear 클래스에서는 16건의 잘못된 예측이 있었음을 보여준다. 오차는 주로 감정 간 유사성으로 인해 발생할 수 있으며, 이를 통해 모델의 주요 혼동 부분을 파악할 수 있다.

정밀도-신뢰도 곡선을 보면, X축은 예측의 신뢰도(0에서 1 사이의 값)이며, Y축은 해당 신뢰도에서의 정밀도를 나타낸다. 굵은 파란색 곡선은 모든 클

래스에 대한 종합 정밀도를 나타내며, 0.973에서 신뢰도가 1.0인 상태로 종료된다. happy 클래스는 신뢰도 구간의 대부분에서 높은 정밀도를 유지한다. neutral과 sad 클래스는 신뢰도가 낮은 구간에서 정밀도가 상대적으로 낮게 유지되는 경향을 보인다. 정밀도-신뢰도 곡선은 모델이 각 감정 클래스에 대해 얼마나 정확한지와 신뢰 수준(confidence) 간의 관계를 보여준다.

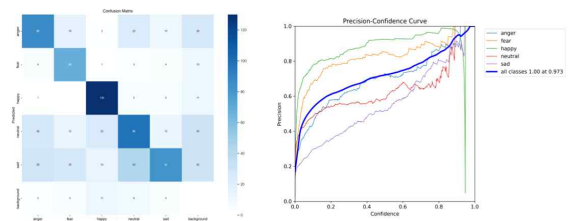


그림 2. YOLO v10 모델의 이상적 성능 분석: 혼동 행렬과 정밀도-신뢰도 곡선

Fig2. Ideal Performance Analysis of YOLO v10 Model: Confusion Matrix and Precision-Reliability Curve

### III. 결 론

본 연구의 결과는 표정 인식 시스템 설계 시 모델의 목적에 따라 최적의 하이퍼파라미터 설정을 선택하는 것이 중요함을 보여준다. 향후 연구에서는 데이터 증강 기법과 추가적인 딥러닝 모델 도입을 통해 성능을 더욱 개선하고, 다양한 환경에서의 표정 인식 능력을 향상할 수 있을 것으로 예상된다.

### 참 고 문 헌

[1] 이승빈, 조동현, 양승진, 김희범, and 강병권. "YOLOv8 알고리즘을 활용한 얼굴 추적 지능형 키오스크." Proceedings of KIIT Conference 2023.11 (2023): 755-757.

[2] 박원규, et al. "객체 인식을 활용한 향수 추천 알고리즘에 관한 연구." 한국통신학회 학술대회논문집 (2024): 1199-1200.

# YOLOv9 모델을 활용한 실시간 화재 감지 시스템 개발

황보재윤\*, 이현진\*, 김기태\*, 권성수\*\*, 전용준\*, 배지훈\*\*\*

## Development of a Real-Time Fire Detection System Using the YOLOv9 Model

Jaeyun Hwangbo\*, Hyunjin Lee\*, Kitae Kim\*, Sungsoo Kwon\*\*, Yong-Jun Jon\*, Ji-Hoon Bea\*\*\*

### 요 약

화재는 예측이 어려워 인명과 재산에 큰 피해를 줄 수 있는 중요한 문제로, 신속한 감지가 필수적이다. 기존 화재 감지 시스템은 고정형 센서와 카메라에 의존하여 설치 비용이 높고 넓은 지역 감지에 한계가 있었다. 이러한 한계를 극복하기 위해 본 연구에서는 YOLOv9 사전 학습 모델을 재학습하여 실시간 화재 경고와 데이터베이스 기능을 갖춘 화재 감지 시스템을 구현하였다. 실험 결과, YOLOv9 모델은 YOLOv8 대비 적은 파라미터로도 높은 정확도를 보였으며, 향후 저사양 임베디드 장치에서의 경량화를 통해 유지 비용 절감을 목표로 한다.

### Abstract

Fire is an unpredictable hazard that can cause significant harm to lives and property, making rapid detection essential. Traditional fire detection systems rely on fixed sensors and cameras, resulting in high installation costs and limited coverage in extensive areas. To address these limitations, this study re-trained the YOLOv9 pre-trained model to implement a real-time fire detection system equipped with alert and database functionalities. Experimental results showed that YOLOv9 achieved high accuracy with fewer parameters than YOLOv8. Future work aims to reduce maintenance costs by optimizing the model for low-spec embedded devices.

### Key words

Real-time fire detection system, YOLOv9, YOLOv8, Object detection, Fire alarm system

---

\* 대구가톨릭대학교 소프트웨어융합대학 AI빅데이터공학과  
\*\* 대구가톨릭대학교 일반대학원 AI빅데이터공학과  
\*\*\* 한국교원대학교 컴퓨터교육과  
※ 교신저자: 한국교원대학교 배지훈(baejh@knue.ac.kr)  
※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음 (2019-0-01056).



## 1. 서론

화재는 예측이 어려워 발생 시 큰 위험을 초래할 수 있는 중요한 사회적 문제 중 하나이다. 이러한 문제는 인명과 재산에 심각한 피해를 줄 수 있어, 이에 대한 효율적인 대응이 필수적이다[1]. 특히 초기 화재 진압이 지연되면 피해 규모는 기하급수적으로 증가하므로, 신속하고 정확한 화재 감지 기술의 개발은 매우 중요하다.

기존 화재 감지 시스템은 주로 고정형 센서와 카메라를 이용해 특정 공간을 모니터링하는 방식에 의존해 왔다[2]. 이러한 시스템은 설치 비용이 많이 들며, 넓은 지역을 효과적으로 감지하는 데 한계가 있다. 또한, 카메라와 연동된 중앙 제어 시스템은 인적 개입을 필요로 하여 처리 속도가 지연되는 문제가 발생할 수 있다.

따라서 본 연구에서는 설치 비용을 줄이고 넓은 지역에서 실시간 화재 감지를 구현하기 위해, 수집된 화재 데이터를 활용해 객체 탐지 모델 You only look once version 9 (YOLOv9) 사전 학습 모델을 재학습하고, 이를 기반으로 실시간 알림 및 데이터베이스 기능을 포함하는 화재 감지 시스템을 제안하고자 한다.

## II. YOLOv9 기반 화재 감지 모델 구현

본 연구에서는 화재 발생 전 연기 발생 순간을 촬영한 영상에서 추출한 이미지들로 구성된 AI Hub의 화재 예측 데이터셋을 활용하였다[3]. 이 데이터셋은 화재 초기 징후를 포착한 총 16,000장의 이미지를 포함하며, 그림 1과 같이 Bounding box 방식으로 라벨링되었다. 화재 감지를 위한 클래스는 'fire'로 지정하였으며, 라벨링된 데이터는 모델 학습을 위해 7:2:1 비율로 훈련, 검증, 테스트 데이터로 분할하였다.

앞서 언급한 방식으로 전처리된 데이터들은 화재 감지를 위해 실시간 객체 탐지 모델의 대표적인 프레임워크 중 하나인 YOLO 시리즈 중 이미지에서 여러 객체를 실시간으로 탐지할 수 있는 고성능 모델인 YOLOv9[4] 모델 학습에 활용된다.



(a) 화재 원본 데이터셋



(b) 라벨링한 데이터셋

그림 1. 화재 데이터셋 및 라벨링  
Fig. 1. Dataset and labeling for fire

YOLOv9는 이미지와 동영상 내 다양한 객체를 신속하고 정확하게 탐지할 수 있는 최신 딥러닝 모델로, 이전 YOLO 모델들의 장점을 계승하면서도 특히 작고 복잡한 객체에 대한 탐지 성능이 뛰어나다는 특징이 있다. 이 모델은 이미지를 입력받아 여러 개의 Bounding box를 생성하고, 각 Bounding box에 객체의 종류와 신뢰도를 예측하는 방식으로 작동한다. 이를 통해 실시간으로 화재를 탐지하고 추적하는 데 유용하다는 장점이 있다.

본 연구에서는 화재 감지 모델의 성능 비교를 위해 최신 YOLOv9 모델과 이전 버전인 YOLOv8 모델[5]을 활용하여 화재 감지 모델을 구현하고 성능을 평가하였다. 신뢰성 있는 비교를 위해 각 모델의 사전 학습 모델 중 가장 많은 매개변수를 포함하는 YOLOv9-e와 YOLOv8-x 모델을 재학습하여 평가를 진행하였으며, 표 1에 두 모델의 성능 비교 결과를 제시하였다.

표 1에서 확인할 수 있듯이, YOLOv8-x 모델이 YOLOv9-e 모델보다 매개변수가 많음에도 불구하고, mAP50과 mAP50-95에서 YOLOv9-e 모델이 각각 0.94와 0.56의 더 높은 성능을 나타냈다. 여기서 mAP50은 예측된 경계 상자와 실제 경계 상자의 일

치 정도를 50% 이상으로 평가한 평균 정확도를 의미하며, mAP50-95는 50%에서 95%까지 5% 간격으로 일치 정도를 평가한 평균 정확도를 나타낸다. 두 모델의 공정한 비교를 위해, 동일한 조건에서 모델을 학습시켰으며, Epoch는 160, Batch Size는 30, 손실 함수는 Stochastic gradient descent(SGD)로 설정하여 실험을 수행하였다.

표 1. YOLOv9-e과 YOLO8-x 모델 성능 비교  
Table 1. YOLOv9-e and YOLO8-x model performance comparison

Model	Parameters	mAP50	mAP50-95
YOLOv8-x	682,000	0.92	0.54
<b>YOLOv9-e</b>	<b>581,000</b>	<b>0.94</b>	<b>0.56</b>



그림 2. YOLOv9-e 모델 화재 감지 예측 결과  
Fig. 2. YOLOv9-e model fire detection prediction results

그림 2는 화재 감지를 위해 재학습된 YOLOv9-e 모델의 학습 결과를 보여준다. 그림에서 볼 수 있듯이, YOLOv9-e 모델은 산불 발생 위치를 신속하게 포착하며 정확하게 화재를 감지하는 성능을 나타내어, 화재 탐지 모델로서의 우수한 활용 가능성을 확인할 수 있다.

#### IV. 실시간 화재 감지 시스템 설계 및 구현

본 연구에서는 실시간으로 화재를 감지하고 효과적으로 대응하기 위해 그림 3과 같이 실시간 화재 감지 시스템을 설계하였다. 이 시스템은 영상 입력 데이터를 처리하여 일정 기간 화재가 지속해서 감지될 경우 경고 메시지를 전송하고, 감지된 데이터를 MySQL 기반 데이터베이스에 저장하는 방식으로 작동한다.

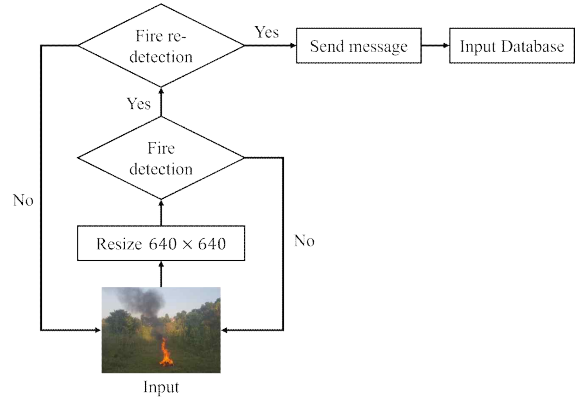


그림 4. 실시간 화재 감지 시스템 아키텍처  
Fig. 4. Real time fire detection system architecture

먼저, 화재 관련 영상이 입력되면 YOLOv9 기반 화재 감지 모델에 맞게 데이터 크기를 640x640으로 조정한다. 이후 조정된 데이터는 YOLOv9 모델을 통해 화재 여부를 판별하며, 화재가 감지된 경우 오탐을 방지하기 위해 10 프레임 연속으로 화재가 감지되는지 재검증 과정을 거친다. 10 프레임 동안 연속으로 화재가 감지되면 Short Message Service(SMS) API로 실시간 화재 감지 메시지가 본 시스템 사용자에게 SMS 형식으로 전송되고, 화재 감지 좌표와 감지 시간 등의 관련 정보가 그림 5와 같이 MySQL 기반 데이터베이스에 저장되도록 설정하였다. 이러한 정보는 이후 화재 발생 패턴 분석, 시스템 성능 개선, 화재 발생 위치의 정확성 확인 등의 목적으로 활용되며, 장기적으로는 더 정교한 화재 예측 모델을 개발하기 위한 중요한 데이터로 사용될 수 있다.

id	top	left	bottom	right	start_time	finish_time
1	859	688	883	859	2024-10-16 18:03:29	2024-10-16 18:03:30
2	859	688	883	859	2024-10-16 18:03:29	2024-10-16 18:03:30
3	1061	679	896	1061	2024-10-16 18:03:42	2024-10-16 18:03:43
4	1060	663	895	1060	2024-10-16 18:04:01	2024-10-16 18:04:02
5	1060	663	895	1060	2024-10-16 18:04:01	2024-10-16 18:04:02
6	1060	670	896	1060	2024-10-16 18:04:20	2024-10-16 18:04:20
7	1061	690	894	1061	2024-10-16 18:04:31	2024-10-16 18:04:31
8	1061	690	894	1061	2024-10-16 18:04:31	2024-10-16 18:04:31
9	859	688	883	859	2024-10-16 18:07:32	2024-10-16 18:07:32
10	1061	679	896	1061	2024-10-16 18:07:44	2024-10-16 18:07:45
11	1060	663	895	1060	2024-10-16 18:08:03	2024-10-16 18:08:03
12	1060	670	896	1060	2024-10-16 18:08:21	2024-10-16 18:08:21
13	1061	690	894	1061	2024-10-16 18:08:32	2024-10-16 18:08:32

그림 5. 데이터베이스 구성  
Fig. 5. Database Configuration

#### V. 결론

본 연구에서는 조기 화재 감지를 위해 YOLOv9

사전 학습 모델을 기반으로 화재 감지 모델을 구현하였으며, 이를 통해 YOLOv8의 가장 높은 파라미터를 가진 사전 학습 모델과 비교했을 때, 더 낮은 파라미터 수를 유지하면서도 더 우수한 성능을 보이는 것을 확인하였다. 또한, 구현된 모델을 활용하여 실시간 경고 메시지 전송 및 데이터베이스 연동 기능을 포함한 신뢰성 높은 실시간 화재 감지 시스템을 제안하였다. 향후 연구에서는 YOLOv9 모델의 성능을 개선하고, 유지 비용 절감을 위해 저사양 임베디드 장치에 적용할 수 있도록 경량화 연구를 진행할 계획이다.

### 참 고 문 헌

- [1] S. Kwon, T. Kim, D. Kim, D. Roh, S. Lee, J. L and J-H. Bae, "Implementation of Real-time Fire Detection and Speech Transformation Lightweight Model Based on Raspberry Pi", Proceeding of KIIT Conference, pp. 326-328. Jeju, 2022.
- [2] Abdusalomov, Akmalbek, N. Baratov, A. Kutlimuratov, and T. K. Whangbo, "An Improvement of the Fire Detection and Classification Method Using YOLOv3 for Surveillance Systems," Sensors, 21(19), 6519, 2021.
- [3] AI Hub, "Fire occurrence prediction images", AI Hub Dataset, <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=176> [accessed: Oct. 10, 2024]
- [4] Ultralytics YOLOv9 Documentation, "Ultralytics," <https://docs.ultralytics.com/ko/models/yolov9> [accessed: Oct. 30, 2024]
- [5] L. Ramos, E. Casas, E. Bendek, C. Romero and F. Rivas-Echeverría, "Hyperparameter optimization of YOLOv8 for smoke and wildfire detection: Implications for agricultural and environmental safety," Artificial Intelligence in Agriculture, vol. 12, pp. 109-126, 2024.

# 머신러닝을 활용한 의류 추천 시스템

김민호\*, 전배준\*, 최재정\*, 강병권\*\*

## Clothing Recommendation System using Machine Learning

Min Ho Kim\*, Bae Jun Jeon\*, Jae Jeong Choi\*, and Byeong-Gwon Kang\*\*

### 요약

머신러닝 알고리즘을 활용하여 온라인 의류 쇼핑에서 사용자에게 최적의 의류 사이즈를 추천하는 시스템을 제안한다. 연구에서는 다양한 모델을 비교하여 정확도와 정밀도가 높은 모델을 선정하고, 하이퍼파라미터 최적화를 통해 성능을 개선하였다. 최종적으로, 추천 시스템이 온라인 의류 쇼핑에서 자주 발생하는 사이즈 불일치 문제를 해결하고자 하며, 이를 통해 반품률을 줄이고 소비자 만족도를 높이는 데 기여할 수 있음을 보인다. 또한, 연구는 판매자의 비용 절감과 온라인 의류 시장의 효율성 향상에도 긍정적인 영향을 미칠 것으로 기대된다.

### Abstract

This study proposes a system that utilizes machine learning algorithms to recommend the optimal clothing size to users in online shopping. The research compares various models to select those with high accuracy and precision, enhancing performance through hyperparameter optimization. Ultimately, the recommendation system aims to address frequent size mismatch issues in online clothing shopping, thereby reducing return rates and improving consumer satisfaction. Additionally, this study is expected to contribute positively to cost reduction for sellers and increased efficiency in the online apparel market.

### Key words

Clothing Size Recommendation, Machine Learning Algorithms, Online Shopping, Hyperparameter Optimization

## 1. 서론

엔데믹 이후 온라인 시장 활성화로 올해 온라인 의류 거래액이 20조 원을 돌파했다. 국가통계포털에 따르면, 2023년 1~8월 의류 온라인 거래액은 전년

대비 4.4% 증가한 12조 4천억 원에 달했다. 특히 MZ세대는 월평균 10만 원을 의류 구매에 지출하며, 상당 부분을 온라인에서 소비한다. 그러나 온라인 반품률은 약 20%로 오프라인의 3배 이상이며, 특히 신발과 의류의 반품률은 30%에 가깝다. 이는 사이

\* 순천향대학교 정보통신공학과 학부과정 kiminho9946@gmail.com

\*\* 순천향대학교 정보통신공학과 교수 bgkang@sch.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 연구 결과로 수행되었음 (2021-0-01399)

즈 선택의 어려움에서 비롯되며[1], 판매 정보만으로 적절한 사이즈를 선택하는 데 한계가 있기 때문이다. 본 논문은 이를 해결하기 위해 AI-hub에서 얻은 데이터를 활용해 최적 사이즈를 추천하는 시스템을 제안한다. 다양한 머신 러닝 모델을 적용하여 최적의 성능을 보이는 모델을 선정하고 분석한다.

## II. 본론

### II-1 데이터 전처리

본 연구에서는 의류 구매 데이터를 효과적으로 분석하고 모델링하기 위해 체계적인 데이터 전처리 과정을 수행한다. 데이터 전처리는 데이터의 품질을 높이고 분석 및 모델링의 정확도와 신뢰성을 향상시키기 위한 필수 단계로, 이를 통해 결측치와 이상치를 식별하고 적절히 처리하여 데이터 왜곡을 최소화하고자 한다. 결측치의 경우, 각 변수의 특성과 중요도를 고려하여 제거하거나 대체하는 방법을 사용한다. 예를 들어, 중요 변수의 결측치는 평균값이나 중앙값으로 대체해 데이터의 일관성을 유지한다. 이상치 처리는 각 변수에 대해 상위 25%와 하위 75% 값을 계산하고, 사분위 범위(IQR)을 통해 상한선과 하한선을 설정하는 방식으로 이루어진다. 이렇게 정의된 상한선과 하한선을 벗어나는 데이터 포인트는 이상치로 간주해 적절히 처리한다. 변수 간의 관계를 분석하기 위해 피어슨 상관 계수 히트맵을 사용하여 변수 간의 상관관계를 시각적으로 이해할 수 있도록 한다.

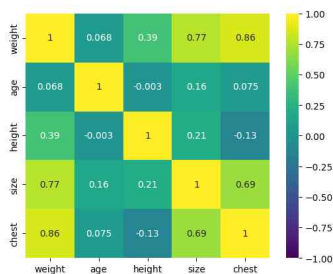


그림 1. 훈련 결과 플로팅  
Fig. 1. Train result floating

이 히트맵을 통해 모델링을 위한 중요한 통찰을 얻었으며, 그 결과는 [그림 1]에 제시되어 있다. 위와 같은 전처리 과정을 통해 데이터의 품질을 개선하고 분석 및 모델링을 위한 기반을 강화한다. 이는 모델 성능을 최적화하고 의류 사이즈 추천 시스템의 정확도를 높이는 데 기여한다.

### II-2 모델 별 성능비교

본 연구는 다양한 모델을 활용해 최적의 성능을 가진 모델을 선정하고, 하이퍼파라미터 튜닝을 통해 성능을 극대화하고자 한다. 실험에 사용된 모델은 랜덤 포레스트, Decision Tree, Extra Trees Classifier, Gradient Boosting Classifier, K-Nearest Neighbors, CatBoost, LightGBM 등이며, 테스트 비율은 30%로 설정하여 실험을 진행했다. [표 1]은 (식 1)을 통해 산출한 정확도(Accuracy), 정밀도(Precision), F1-Score 등의 성능 지표 결과를 보여준다.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1 - score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}$$

표 1. 실험 데이터 세트

Table 1. Experimental data sets

Machine Learning Model	Accuracy	Precision	F1-Score
random forest	0.9062	0.9030	0.9014
decision tree	0.9584	0.9584	0.9584
K-Nearest Neighbors	0.9533	0.9533	0.9585
Gradient Boosting Classifier	0.9562	0.0589	0.9562
catboost	0.9585	0.0589	0.9585
LightGBM	0.8992	0.0589	0.8992

[표 1]에서 확인할 수 있듯이, CatBoost, Extra

Trees Classifier, Gradient Boosting Classifier, Light GBM 모델의 F1-Score가 상대적으로 낮은 것을 볼 수 있으며, 이는 모델별 특성으로 인해 불균형이 있을 가능성을 시사한다. 또한, 각 모델의 성능을 비교한 결과, Decision Tree 모델이 가장 우수한 성능을 보였다. 이에 따라, Decision Tree의 하이퍼 파라미터를 튜닝 하여 더 나은 성능을 도출하고자 한다.

### II-3 의사결정 트리 튜닝

하이퍼파라미터 튜닝에는 여러 방법이 있으나, 그중 그리드 탐색(Grid Search)과 랜덤 탐색(Random Search)이 자주 사용된다. 그리드 탐색은 사전 정의된 값들의 모든 조합을 탐색해 최적의 조합을 찾는 방식이고, 랜덤 탐색은 범위 내에서 무작위로 조합을 선택하여 성능이 높은 설정을 찾는다. 본 연구에서는 의사결정 트리 모델의 최적 하이퍼 파라미터를 도출하기 위해 그리드 탐색 방법을 활용한다. 한 번에 여러 매개변수를 동시에 조정하면 성능에 미치는 영향을 명확히 평가하기 어렵다. 따라서 본 연구에서는 매개변수를 하나씩 조정하는 방식을 택했다. 이 방법은 데이터 특성에 맞는 최적화에 기여한다. 튜닝 결과, 최적 성능을 위해 max\_features는 제한 없이 설정하고, max\_depth는 8, min\_samples\_leaf는 1, min\_samples\_split은 2로 설정하는 것이 효과적임을 확인했다. 또한 엔트로피 기반 기준을 사용해 정보 이득이 높은 분할을 찾도록 하여 모델 성능을 개선하였다.

표 2. 튜닝에 따른 성능 비교  
Table 2. Experimental data sets

decision tree	Accuracy	Precision	F1-Score
Before tuning	0.9584	0.9584	0.9584
After tuning	0.9933	0.9933	0.9931

튜닝 전후의 성능 차이는 위 [표 2]에 제시되어 있으며, 이는 하이퍼파라미터 튜닝과 엔트로피 활용이 의사결정 트리 성능 최적화에 중요한 역할을 한다는 점을 시사한다.

### III. 결 론

본 연구에서는 머신러닝 알고리즘을 통해 사용자에게 적합한 의류 사이즈를 추천하는 시스템을 제안하였다. 다양한 모델을 평가하여 최적의 성능을 보이는 모델을 선정하고, 하이퍼파라미터 최적화를 통해 성능을 향상시켰다. 이를 통해 본 연구는 기존 연구[2]보다 높은 정확도를 달성할 수 있었다. 그러나 데이터 세트의 차이에 따른 결과의 일반화 가능성에 대해서는 신중히 접근할 필요가 있다.

본 연구에서 제안한 사이즈 추천 시스템은 온라인 의류 쇼핑에서 발생하는 사이즈 불일치 문제를 줄이고, 소비자가 적합한 사이즈를 쉽게 찾도록 하여 반품률 감소와 고객 만족도 향상에 기여할 것으로 기대된다. 또한 이는 판매자의 비용 절감 효과에도 긍정적인 영향을 미칠 수 있다. 본 연구는 의류 사이즈 추천 시스템 개발을 위한 중요한 기초 자료가 될 뿐만 아니라, 관련 연구와 응용 분야에서의 혁신에도 기여할 것이다. 이러한 연구 성과는 소비자 중심의 서비스 개선과 업계의 지속 가능한 성장을 촉진하는 데 의미가 있다.

### 참 고 문 헌

- [1] Y. Chen and Z. Yang, "The behavioral analysis of choice difficulty states during clothing online shopping," *International Journal of Clothing Science and Technology*, vol. 33, no. 4, pp. 577-589, 2021.
- [2] D. Lee, "Personalized Size Recommender System for Online Apparel Shopping: A Collaborative Filtering Approach," *Journal of the Korea Society of Computer and Information*, vol. 28, no. 8, pp. 39-48, 2023.

# AI 기반 CAPTCHA 보안 우회 가능성

박서인\*, 박지은\*, 왕태수\*, 장시웅\*\*

## AI-based CAPTCHA security bypassability

Seo-in Park\*, Ji-eun Park\*, Tae-su Wang\*, and Si-Woong Jang\*\*

### 요약

CAPTCHA는 웹사이트에서 자동화된 봇 공격을 방어하기 위한 주요 보안 수단이다. 그러나 최근 AI 기술의 발전으로 CAPTCHA 시스템이 상대적으로 쉽게 우회될 수 있는 가능성이 제기되고 있다. 본 논문에서는 EfficientNetV2-RNN 기반의 CAPTCHA 우회 방법을 실험하고, 그 결과를 통해 CAPTCHA 시스템의 보안적 취약성을 확인하였다. 이러한 결과는 기존 CAPTCHA 시스템이 AI 기반의 공격에 의해 쉽게 위협받을 수 있음을 시사하며, 보안 대책의 필요성을 제기한다.

### Abstract

CAPTCHA is the main security tool for defending against automated bot attacks on websites. However, with the recent development of AI technology, the possibility that the CAPTCHA system can be bypassed relatively easily has been raised. In this paper, we experimented with the CAPTCHA bypass method based on EfficientNetV2-RNN and confirmed the security vulnerability of the CAPTCHA system through the results. These results suggest that the existing CAPTCHA system can be easily threatened by AI-based attacks and raise the need for security measures.

### Key words

CAPTCHA, AI, EfficientNet, RNN

## 1. 서론

CAPTCHA(Completely Automated Public Turing test to tell Computers and Humans Apart)는 웹사이트에서 자동화된 봇으로부터 보호하기 위해 고안된 시스템으로, 주로 사용자가 직접 입력해야 하는 문자 또는 이미지 문제를 사용하여 인간과

봇을 구별한다. 이는 오랜 시간 동안 웹 보안의 핵심 요소로 활용되어 왔다. 그러나 최근 AI 기술이 급속도로 발전하면서 CAPTCHA 시스템의 보안이 크게 위협받고 있다. 딥러닝, 특히 EfficientNet과 같은 첨단 모델들이 CAPTCHA의 구조를 이해하고 이를 우회하는 방식으로 발전하고 있다. 본 논문에서는 EfficientNetV2-S와 RNN 기반 모델의 딥러닝

\* 동의대학교 컴퓨터공학과, 20193104@office.deu.ac.kr, 20213855@office.deu.ac.kr, tswang@office.deu.ac.kr

\*\* 동의대학교 스마트자동차연구소 연구소장, swjang@deu.ac.kr(교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업임(IITP-2024-RS-2020-II201791, 50%). 또한 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학협력 선도대학 육성사업(LINC 3.0)의 연구결과입니다(50%).

모델을 사용하여 CAPTCHA를 자동으로 풀어내는 방법을 실험하고, 이를 통해 CAPTCHA 시스템의 취약성을 확인한다.

## II. 관련 연구

CAPTCHA가 자동화 프로그램에 의해 쉽게 해독될 수 있는 문제를 해결하기 위해, 문자의 왜곡 강도와 이미지 배경의 간섭 수준을 조절한 연구와[1], SVM을 사용하여 배경과 문자를 분리하고, CNN을 통해 문자를 인식하는 방식으로 CAPTCHA의 취약성을 실험한 연구가 있다[2].

## III. AI를 이용한 CAPTCHA 우회 실험

본 연구에서는 이미지 기반 CAPTCHA를 효과적으로 우회하기 위해 RNN 기반 모델을 설계하고, 이를 통해 문자 인식 과정을 자동화하는 실험을 수행하였다. 모델은 EfficientNetV2-S[3]를 활용하여 이미지에서 특징을 추출하고, 예측된 문자를 바탕으로 CAPTCHA를 해결하는 방식으로 동작한다. CAPTCHA 이미지는 다양한 형식으로 구성되어 있으며, 해당 모델은 훈련된 데이터셋을 기반으로 성능을 검증하였다.

### 1) 데이터셋 개요

본 시스템에서 사용된 데이터셋은 CAPTCHA 이미지로 구성된 총 6개의 케글 데이터셋 약 24만장을 포함하며, 각 데이터셋은 서로 다른 문자 구성과 글꼴을 포함하고 있다. 데이터셋은 학습(80%)과 검증(10%), 테스트(10%)로 나뉘어져 있으며, 다양한 형태의 CAPTCHA를 처리할 수 있도록 설계되었다. 학습 과정에서 모델은 다양한 문자와 크기의 이미지를 처리하도록 학습되었으며, 최종적으로 문자 인식의 정확도를 높이는 것을 목표로 하였다.

### 2) 실험 방법

이 모델은 먼저 CNN(Convolutional Neural Network)을 통해 이미지에서 주요 특징을 추출하고, 그 후 RNN(Recurrent Neural Network)을 사용해

연속적인 문자 시퀀스를 예측하는 방식으로 구성된다. 분류한 CAPTCHA 데이터셋을 CNN을 대신해 EfficientNetV2-S를 사용하여 특징을 학습한다. EfficientNetV2-S는 일반적인 CNN과 달리 더 작은 파라미터로 높은 성능을 제공하며 CAPTCHA와 같은 시각적 잡음이 많은 이미지에서도 특징을 빠르게 추출할 수 있다. 이 과정에서 모델은 문자와 배경의 시각적 차이를 구분할 수 있도록 시각적 정보를 캡처한다. 추출된 특징은 RNN이 이를 문자 시퀀스로 변환시켜 이미지 내 연속된 CAPTCHA 텍스트를 순서에 맞게 예측한다. 예측된 CAPTCHA 텍스트는 실제 정답과 비교하여 성능을 평가하며, 각 문자에 대한 일치율을 분석하였다.

### 3) 실험 결과 및 검증

실험 결과, 모델은 학습된 CAPTCHA 문자를 높은 정확도로 예측하였으며, 실제 문자와 비교했을 때 일치율이 높았다.

그림 1은 실험 데이터셋이며, 표 1은 실험 데이터를 검증한 결과를 나타낸 것이다. \* 표시가 있는 (D), (E)는 새로운 CAPTCHA 이미지로 검증에 사용됐다. 다양한 형태의 CAPTCHA 이미지에 대한 성능 검증에서도 우수한 결과를 보였으며, 글자 수가 많거나 노이즈가 포함된 CAPTCHA 이미지의 경우 일부 문자가 틀리는 경향이 있었으나 전반적인 인식 성능은 우수하였다. 데이터셋 외의 새로운 CAPTCHA 이미지에서도 모델은 훌륭한 결과를 도출하였다.

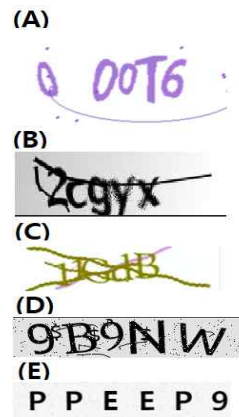


그림 1. 실험 데이터 세트  
Fig. 1. Experimental data sets



표 1. 실험 데이터 세트

Table 1. Experimental data sets

	Label	Prediction
(A)	000T6	000T6
(B)	2cgyx	2cgyx
(C)	11GdB	11GdB
*(D)	9B9NW	9B9NW
*(E)	PPEEP9	PPFE9

#### IV. 결 론

본 연구에서는 CNN 기반의 EfficientNetV2-S 모델과 RNN을 사용해 CAPTCHA 시스템의 취약성을 평가하였다. 실험 결과, 본 모델은 다양한 형태의 CAPTCHA를 높은 정확도로 해독하였으며, AI 기술이 CAPTCHA의 보안을 쉽게 우회할 수 있음을 확인하였다. 이는 기존 CAPTCHA의 보안성이 AI 발전에 의해 위협받고 있음을 시사하며, 보다 강력한 보안 방안을 마련할 필요가 있다.

#### 참 고 문 헌

- [1] SongSeul-gi (2017). Study on A Hybrid CAPTCHA Test Based on Characters and images for Security Enhancement [Master's Thesis, KangwonNationalUniversity]. <http://www.riss.kr/link?id=T14572868>
- [2] Jaehwan Kim, Suah Kim, & Hyoung Joong Kim (2015). Breaking character and natural image based CAPTCHA using feature classification. Journal of the Korea Institute of Information Security & Cryptology, 25(5), 1011-1019.
- [3] Tan, Mingxing, and Quoc Le. "Efficientnet v2: Smaller models and faster training." In international conference on machine learning. PMLR, 2021.

# 머신러닝을 활용한 심전도의 RR간격을 통한 부정맥 예측 시스템

이세훈\*, 유채린\*, 신민혁\*, 김민호\*, 백경호\*\*, 강병권\*\*\*

## Real-Time ECG Signal Analysis and Arrhythmia Prediction System Using Machine Learning

Se-Hun Lee\*, Chae-Rin Yoo\*, Minhyuk Shin\*, Minho Kim\*, Kyung-Ho Paik\*\*  
and Byeong-Gwon Kang\*\*\*

### 요약

본 연구는 유도 심전도를 활용하여 부정맥 예측 시스템 개발을 목표로 한다. 기존에 문제점을 보완하기 위하여 해당 연구는 유도 심전도 데이터를 받아 통계적인 특징을 바탕으로 RR 간격을 분석한다. 이를 기반으로 부정맥을 예측하고 평가 지표를 통하여 높은 정확도를 갖는 부정맥 예측 머신러닝 모델을 제안한다. 결과적으로 본 연구에서 제시하는 부정맥 예측 머신러닝 모델은 86%의 정확도를 가진다. 본 연구는 높은 정확도로 부정맥을 감지할 수 있는 가능성을 제시하며 향후 부정맥의 조기 발견과 예방에 기여할 수 있을 것으로 기대된다.

### Abstract

This study aims to develop an arrhythmia prediction system using an induced electrocardiogram. In order to compensate for the existing problems, the study receives induced electrocardiogram data and analyzes the RR interval based on statistical characteristics. Based on this, we predict arrhythmia and propose an arrhythmia prediction machine learning model with high accuracy through evaluation indicators. As a result, the arrhythmia prediction machine learning model presented in this study has an accuracy of 86%. This study presents the possibility of detecting arrhythmia with high accuracy and is expected to contribute to the early detection and prevention of arrhythmia in the future.

### Key words

Arrhythmia Detection, ECG, RR Interval Analysis, Machine Learning Model, Prediction Accuracy

1. 서론  
부정맥은 심장 박동의 이상으로 발생하므로 심장

의 전기적 활동과 관련된다. 심전도(ECG)는 심장의 탈분극과 재분극시 발생하는 생체전기를 피부에 부

\*순천향대학교 정보통신공학과 학부과정, tpgns127@naver.com

\*\*순천향대학교 전자공학과 겸임교수, pkhtow@naver.com

\*\*\*순천향대학교 정보통신공학과 교수, bgkang@sch.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 연구 결과로 수행되었음 (2021-0-01399)

착된 전극을 통하여 측정된 그래프로서, P파, QRS 군, ST-segment, T파로 구분할 수 있다.[1] 부정맥은 일시적으로 발생하는 경우가 많아 일반적인 검진에서는 발견이 어려울 수 있다. 심전도는 실시간 모니터링이 가능하여 정확한 진단과 부정맥 유형 구분에 정보를 제공한다[2]. 병원에서의 ECG 측정은 3개, 5개, 12개의 Lead를 주로 사용하며, 최근에는 웨어러블 기기를 이용한 1개의 Lead만으로도 ECG 측정이 가능 해졌을 뿐만 아니라 심전도 부정맥 검출이 탁월함을 보였다.[3] 심전도 기기에서 전달되는 데이터의 포맷과 웨어러블 기기에서 출력되는 데이터의 포맷은 서로 상이 할 뿐만 아니라 P파, Q파, R파, S파, T파와 같은 세부적인 정보 확인이 불가하다.[1] 해당 연구는 실시간 심전도 데이터의 통계적인 특성을 활용하여 RR 간격을 계산하고 이를 기반으로 부정맥을 예측한다. 해당 예측 모델의 정확도(Accuracy)를 높이기 위하여 정밀도, 재현율 등 다양한 평가 지표를 사용하여 머신러닝 모델을 여러 각도로 평가하고 개선한다. 본 연구를 통해 실시간 심전도 데이터 수집 및 분석을 통하여 부정맥을 예측함으로써 예방 및 조기 치료에 도움이 되고자 연구를 진행하였다.

## II. 본론

해당 연구에서는 심전도 신호를 분석하여 부정맥을 예측하는 과정에 앞서, 입력받은 1 유도 심전도 신호 데이터를 전처리하는 과정을 거쳤다. 입력받은 신호 데이터에서 필터링 과정을 거쳐 잡음을 줄이고, R 피크를 추출하여 RR간격 값을 계산해 부정맥을 판정하는 세 단계의 순서를 거쳤다. 데이터 전처리 단계 이후에는 랜덤 포레스트 머신러닝 모델을 이용하여 예측된 데이터의 정확도를 계산하였다. 정확도와 정밀도, 재현율 등 추가적인 평가지표를 통해 성능을 다양한 각도로 평가하여 예측값의 정확도를 높일 수 있었다.

### II-1. 심전도 데이터 전처리

본 연구는 부정맥과 정상 데이터를 효과적으로 분석하고 모델링하기 위해 체계적인 데이터 전처리

과정을 수행한다. 데이터파일은 순천향대학교 데이터센터에서 데이터를 받아 전처리를 진행하였다. 정상 데이터셋은 아래 [그림 1]과 같이 R파의 값이 일정한 주기마다 반복되고 있는 것을 확인하였다.

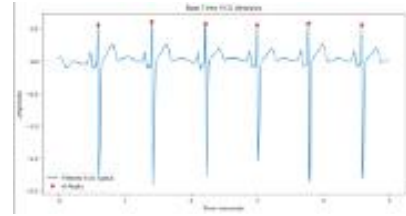


그림 1 정상데이터의 R-peak 검출  
Fig. 1. R-pak detection of normal data

반대로 [그림 2]를 보면 부정맥 데이터셋의 주기가 일정하다가 부정맥인 부분에서 주기가 일정하지 않은 것을 확인하였다.[4] 이를 이용해 부정맥 데이터셋과 정상 데이터셋을 분류해서 저장했다. 그리고 머신러닝의 지도학습을 하기 위해 라벨링을 하여 전처리하였다.

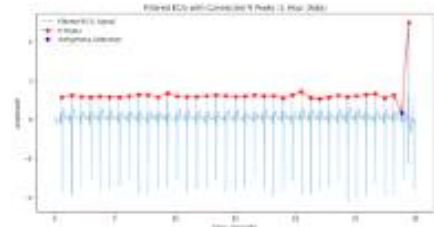


그림 2 부정맥 데이터의 R-peak와 Arrhythmia 검출  
Fig. 2. R-peak detection of Arrhythmia data

### II-1.1. 모델의 성능

머신러닝 결과 [그림 3]을 통해 분류모델의 예측 결과와 실제 라벨을 비교하였다. 또한 [그림 5]와 식(1)을 사용하여 정확하게 분류한 비율과 실수한 비율을 구체적으로 파악하였다.

[그림 4]는 ROC Cuve를 통해 성능지표를 출력 해 본 결과 정밀도는 약 0.803이 나왔고, 재현율은 약 0.767이 나왔다. 머신러닝 결과 모델의 정확도는 86.7%가 나왔다.

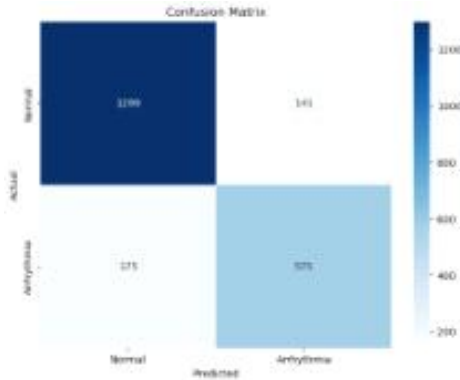


그림 3 Confusion Matrix  
Fig. 3 Confusion Matrix

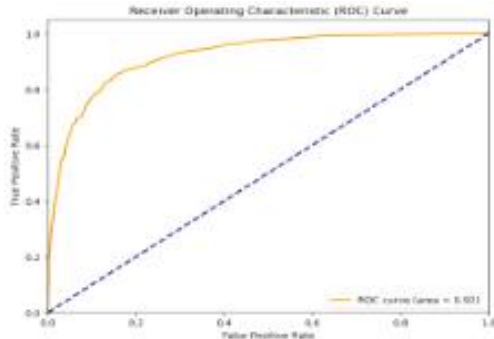


그림 4. ROC Curve  
Fig. 4. ROC Curve

		실제 정답	
		True	False
분류 결과	True	True Positive	False Positive
	False	False Negative	True Negative

그림 5. Confusion Matrix 구성 요소  
Fig. 5. Confusion Matrix

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN}$$

### III. 결 론

ECG를 이용하여 데이터를 받아 통계적인 특징을 바탕으로 RR 간격을 분석하고, 이를 기반으로 부정맥을 예측하고 추가적인 평가 지표를 통하여 높은 정확도를 갖는 부정맥 예측 머신러닝 모델을 제안 하였다. 결과적으로 본 연구에서 제시하는 부정맥 예측 머신러닝 모델은 86%의 정확도를 가진다. 본 연구는 기존의 논문과는 다른 데이터 처리 방식으로 더 높은 정확도로 부정맥을 감지하였다[1]. 향후 부정맥의 조기 발견과 예방에 기여 할 수 있을 것으로 기대 된다. 차후 심전도의 파형을 통해 RR의 간격뿐만 아니라 P파, Q파, R파, S파, T파의 파형을 분석하여 좀 더 정밀하게 부정맥을 파악할 수 있게 추가적인 연구를 진행할 것이다.

### 참 고 문 헌

- [1] 정구영, 윤명중, 유기호. (2009). 실시간 심전도 분석 및 모니터링 시스템 개발. 제어로봇시스템학회 논문지, 15(4), 406-412.
- [2] Kwon S, Lee S-R, Choi E-K, Ahn H-J, Song H-S, Lee Y-S, Oh S. Validation of Adhesive Single-Lead ECG Device Compared with Holter Monitoring among Non-Atrial Fibrillation Patients. Sensors. 2021; 21(9):3122.
- [3] 장주희, 곽진희, and 정재희. “심전도 (ECG) 의 파형 검출과 RRI 를 통한 조기수축 부정맥 판단 시스템.” 한국정보과학회 학술발표논문집 (2021): 1243-1245.
- [4] 배정현, 임승주, 김정주, 박성대, 김정도. (2012). ECG 신호에서 단위패턴간 유사도분석을 이용한 부정맥 분류 알고리즘. 정보처리학회논문지D, 19(1), 105-112.

# HILS 시스템을 활용한 자율주행 경로 제어 알고리즘 연구

조현준\*, 김곤우\*\*

## Autonomous Driving Path Control Algorithm with HILS System

Cho Hyun-jun\*, and Kim Gon-woo\*\*

### 요 약

자율주행 차량의 경로 제어 성능을 평가하기 위해 HILS(Hardware-in-the-Loop Simulation) 시스템을 구축하고, 이를 통해 자율주행 경로 제어 알고리즘을 테스트하고 검증하는 과정을 설명한다. 연구에서는 MORAI 시뮬레이터 가상 환경에서의 경로 추적 성능을 평가하였다. Pure Pursuit 알고리즘을 적용하여 직선 및 곡선 경로에서의 경로 추적 성능을 Original Path와 비교 분석하였다. 본 연구는 자율주행 기술의 안전성과 효율성을 높이는 기초 자료로 활용될 수 있으며, 향후 Xycar와 MORAI 간의 완전한 연동을 통해 실질적인 성능 평가를 제시한다.

### Abstract

This study describes the process of constructing a Hardware-in-the-Loop Simulation (HILS) system to evaluate the path control performance of autonomous vehicles and test and verify autonomous driving path control algorithms. The research assesses path tracking performance in the virtual environment of the MORAI simulator. By applying the Pure Pursuit algorithm, path tracking performance on both straight and curved paths was evaluated and compared with the Original Path. This study provides foundational data to enhance the safety and efficiency of autonomous driving technology and suggests practical performance evaluation through complete integration of Xycar and MORAI in future studies.

### Key words

Autonomous Driving, HILS, Path Control, Xycar, MORAI, Pure Pursuit

## 1. 서 론

자율주행 기술의 발전은 다양한 산업 분야에서 매우 중요한 역할을 하고 있으며, 실제 주행 환경에서 발생할 수 있는 다양한 시나리오를 안전하게 테스트할 수 있는 시스템이 필요하다. HILS (Hardware

-in-the-Loop Simulation) 시스템은 자율주행 차량의 실제 하드웨어와 가상 시뮬레이션 환경을 통합하여 이러한 문제를 해결할 수 있는 방법으로 주목받고 있다[1]. HILS 시스템을 통해 자율주행 차량의 주행 데이터를 실시간으로 반영하고, 가상 환경에서 다양한 주행 상황을 시뮬레이션하여 경로 제어 알고리

\*충북대학교 정보통신공학부, tonycho98@chungbuk.ac.kr

\*\*충북대학교 지능로봇공학과, gwkim@cbnu.ac.kr(교신저자)

즘의 성능을 검증할 수 있다.

본 연구에서는 Xycar와 MORAI 시뮬레이터를 연동하여 HILS 시스템을 구축하고, 특히 Pure Pursuit 알고리즘을 활용하여 가상 환경에서 차량의 경로 추적 알고리즘을 설계하였다. 이는 향후 자율주행 기술의 안전성과 효율성을 높이는 데 기여할 것이다.

## II. Pure Pursuit 알고리즘과 HILS 시스템

본 연구의 목적은 자율주행 차량의 경로 제어 성능을 평가하기 위한 HILS 시스템을 구축하는 것이다. Xycar와 MORAI를 연동한 이 시스템은 가상 환경에서 자율주행 차량이 경로를 추적하고 장애물을 회피하는 성능을 평가하는 데 사용된다. 이를 통해 다양한 주행 시나리오를 안전하게 시뮬레이션할 수 있으며, 실제 환경에서 실험 전에 가상 환경에서 알고리즘의 성능을 미리 검증할 수 있다.

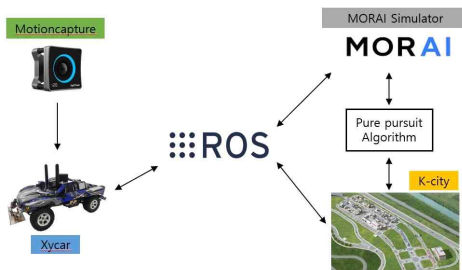


그림 1. HILS 시스템의 주요 구성 요소  
Fig. 1. Key components of the HILS system

표 1. HILS 시스템의 주요 구성 요소

Table 1. Key components of the HILS system

구성 요소	설명
Xycar	자율주행 모형 차량
ROS	차량 제어 및 데이터 통신 시스템
MORAI 시뮬레이터	가상 환경에서의 주행
Pure Pursuit 알고리즘	경로 추적 알고리즘
모션캡처 카메라	모형 차량의 운동 기록

Pure Pursuit 알고리즘[2,3,4]은 자율주행 차량이 주어진 경로를 따라가기 위해 조향각을 계산하는 경로 추적 알고리즘이다. 이 알고리즘은 차량의 현재 위치에서 Lookahead point까지의 거리를 계산하

여 최적의 경로를 추적하도록 설계되었다. 차량의 속도에 따라 Lookahead distance가 동적으로 변화하며, 빠른 속도로 주행할수록 더 먼 지점을 바라보고 주행할 수 있다.

$$\theta: \text{조향각(Steering angle)} \quad (1)$$

$$L: \text{차량의 휠베이스 (Wheelbase)} \quad (2)$$

$$\alpha: \text{차량과 전방 주시 지점 간의 각도} \quad (3)$$

$$L_d: \text{전방 주시 거리 (Lookahead distance)} \quad (4)$$

$$\theta = \arctan\left(\frac{2L \sin(\alpha)}{L_d}\right) \quad (5)$$

위 수식을 기반으로 차량의 조향각이 결정되며, 경로 추적 및 장애물 회피 성능을 극대화할 수 있다.

본 연구에서는 자율주행 모형 차량인 Xycar와 가상 환경 시뮬레이터 MORAI를 사용하여 HILS 시스템을 구축하였다.

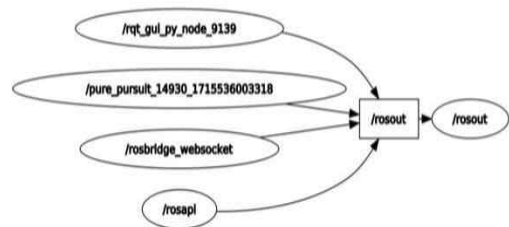


그림 2. 자율주행 차량의 rqt\_graph  
Fig. 2. rqt\_graph of autonomous vehicle

실험은 MORAI 시뮬레이터 상에서 진행되었다. Pure Pursuit 알고리즘을 적용한 차량은 가상 환경에서 직선 경로와 곡선 경로로 나누어서 평가하였다. 경로 추적 성능을 비교하기 위해 Original Path와 Pure Pursuit Path의 차이를 색상으로 구분하여 시각화하였다.

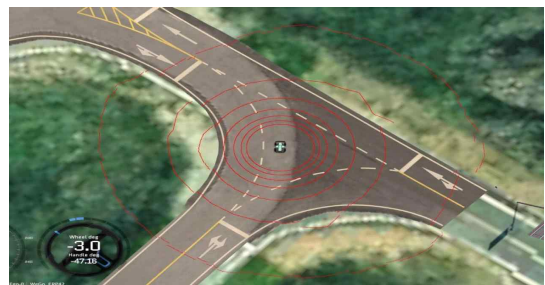


그림 3. MORAI 시뮬레이터 실험  
Fig. 3. Training in MORAI Simulator

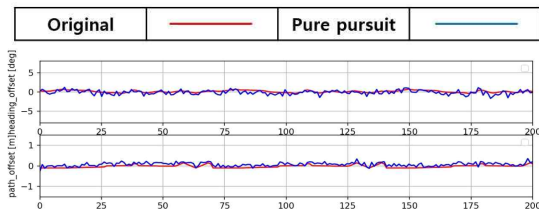


그림 4. 직선 경로에서의 성능 차이  
Fig. 4. Performance Difference in Straight Path

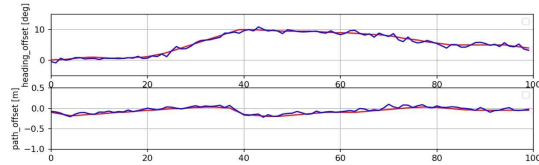


그림 5. 곡선 경로에서의 성능 차이  
Fig. 5. Performance Difference in Curved Path

### III. 결 론

본 논문에서는 자율주행 차량의 경로 추적 성능을 평가하기 위해 HILS 시스템 구축하고 직선 및 곡선 구간에서의 Pure Pursuit 알고리즘 성능을 분석하였다. HILS 시스템을 통해 가상 환경에서 자율주행 차량의 주행 성능을 사전에 검증함으로써, 향후 실제 주행 환경에서의 안전성과 효율성을 높이는 데 기여할 수 있다.

향후 연구에서는 Xycar와 MORAI 시뮬레이터 간의 데이터 연동을 개선하고, 곡선 구간에서 Pure Pursuit 알고리즘의 성능을 높일 수 있는 방안을 모색할 예정이다.

### 참 고 문 헌

[1] M. S. Lee, G. W. Jeon, T. H. Kim and K. S. Yun, "A Study on HILS Environment for the Evaluation of Autonomous Driving Performance", KSAE Fall Conference Proceedings, pp.805-807, 2019.

[2] H. J. Joo and K. B. Lee, "Changes in Path Tracking Performance of Autonomous Vehicle According to Design Feedback Gain and Look-ahead Distance of Pure-pursuit Algorithm", Transactions of KSAE, Vol.29, No.9, pp.839-846,

2021.

[3] H. S. Shin, M. J. Kim and J. H. Kim, "Path Generation and Tracking Algorithm of Auto Valet Parking System Development", Transactions of KSAE, Vol.28, No.6, 2020.

[4] H. G. Lee, T. Y. Kim, D. Y. Yu, S. H. Hwang, "Path-Following Correction Control Algorithm Using State Errors", Transactions of KSAE, Vol.30, No.2, 2022

# BERT 모델을 활용한 텍스트 기반 MBTI 성격 유형 분류

이준학\*, 윤웅창\*\*

## Text-based MBTI personality classification using a BERT model

Junhak Lee\*, and Woongchang Yoon\*\*

### 요약

언어 분석을 통한 개인 성격의 분류는 언어 분석 심리학 연구에서 중요하게 다뤄지는 문제이다. 본 연구에서는 BERT 모델을 활용한 텍스트 기반 MBTI 성격 유형 분류 모델을 제안한다. Reddit에서 수집된 106,067개의 게시글 데이터를 활용하여 16가지 MBTI 유형을 분류하는 모델을 개발하였다. 평가지표로 정확도, 정밀도, 재현율, F1-Score를 사용하여 모델 성능을 평가하였다. 그 결과, 우리가 제안한 모델은 평균 정확도 73.75%, 정밀도 74.79%, 재현율 73.75%, F1-Score 73.76%의 성능을 보였다. 초기 학습 시 일부 데이터 그룹에서 성능 저하를 확인하였으나, 워밍업 기법을 적용하여 훈련 손실 함수 값이 46.69% 감소하고 검증 성능이 22% 이상 향상되는 안정적인 결과를 보였다.

### Abstract

Classification of individual personality through language analysis is a significant issue in the study of language analysis psychology. In this study, we propose a text-based MBTI personality type classification model using the BERT model. The model was developed to classify 16 MBTI types using 106,067 post data collected from Reddit. We evaluated the model's performance using accuracy, precision, recall, and F1-Score as evaluation metrics. As a result, our proposed model demonstrated a performance with an average accuracy of 73.75%, precision of 74.79%, recall of 73.75%, and F1-Score of 73.76%. Although a performance degradation was observed in some data groups during initial training, the application of the warm-up technique resulted in a 46.69% reduction in training loss function value and more than a 22% improvement in validation performance, leading to stable results.

### Key words

MBTI, Personality typing, Natural language processing, Deep learning, Text mining

---

\* 경상국립대학교 컴퓨터공학과, 221161115@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과/바이오의료빅데이터학과, neobits@gnu.ac.kr



## 1. 서론

개인의 성격은 다양한 환경적 요인에 의해 형성되며, 대인관계에서 자신과 타인의 성격을 이해하는 것이 중요하다 [1]. 인간의 성격과 심리는 언어를 통해 감정과 함께 표현되며 최근 인공지능 기술의 발달로 언어 분석을 통한 심리학 연구가 활발하게 이루어지고 있다 [2].

‘MBTI 성격 유형 지표 (Myers-Briggs Type Indicator, 이하 MBTI)’는 심리학적 성격 유형을 구분하는 도구로 ‘psychological types’ 이론을 기반으로 고안된 검사이다 [3]. 해당 검사는 힘을 얻고 사용하는 방식과 정보 인식 방식, 판단 방식, 외부 세계에 대한 경향 등 4가지 부분을 조합하여 총 16가지의 유형으로 판단한다. 검사의 결과는 피검사자가 설문지 형태의 검사 문항에 응답한 값을 종합하여 대표적인 알파벳 4자로 표현된다 [4].

최근에는 딥러닝 모델을 활용한 MBTI 관련 연구가 활발하게 이루어지고 있으며, 대표적으로 사전학습된 임베딩 벡터와 BPE (Byte pair encoding) 알고리즘을 적용한 순환신경망 계열의 모델을 활용하여 자연어 응답 기반의 MBTI 성격 유형 분류 모델을 제안한 연구가 있다 [5]. 또한, 소셜 미디어 데이터의 무의식적인 표현을 활용하여 K-means 클러스터링과 그래디언트 부스팅 알고리즘을 결합한 MBTI 유형을 예측하는 모델이 연구도 있다. 해당 연구에서는 TF-IDF 매트릭스 기법과 하이퍼파라미터를 튜닝을 통해 향상된 성능을 보였다 [6].

이러한 기계학습 모델을 활용한 MBTI 연구들은 주로 설문 기반의 검사 방식을 통해서 데이터를 생성하였으나, 이러한 방법은 시간이 많이 소요되며 피검사자의 주관적 판단에 따라 편향이 발생하는 자기 보고적 검사라는 한계점을 지니고 있었다 [7].

때문에, 텍스트 데이터를 사용하여 기계학습 모델을 활용한 성격 유형 예측 연구가 시작되고 있다. 대표적으로 에세이 텍스트 데이터를 사용하여 MBTI와 Big Five 성격 유형을 XLNet의 순열 언어 모델링 방식을 이용해 양방향 문맥을 학습하여 성격을 예측한 연구가 있다 [8].

BERT (Bidirectional Encoder Representations from

Transformers)는 트랜스포머 기반의 사전 학습된 언어 모델로, 입력 텍스트의 양방향 정보를 동시에 학습할 수 있어 문맥을 더 잘 이해하는 특성을 보이는 장점을 가지고 있다 [9].

본 연구에서는 BERT 모델을 사용하여 피검사자의 SNS와 같은 일상생활에서 편리하게 수집 가능한 데이터를 기반으로 하여 MBTI 성격 유형을 분류하는 모델을 제안한다. 우리는 일상적으로 사용하는 언어 습관을 분석하여 사용함으로써 향상된 객관성과 효율성, 정확성을 기대하였다.

## II. 연구 방법

본 연구에서는 인터넷 커뮤니티 사이트인 Reddit에서 MBTI가 공개된 작성자들의 게시글을 크롤링한 데이터 세트를 Kaggle에서 수집하여 사용하였다 [10]. 데이터 세트는 게시글 (Posts) 106,067개와 게시글 작성자의 MBTI 유형 (type)으로 이루어져 있다 (표 1).

표 1. MBTI 성격 유형 텍스트 데이터 세트  
Table 1. MBTI personality type text data sets

	평균값	중앙값	최대값	최소값	표준편차
Posts	6629	1235	24961	181	48.11
Word	500.85	500	527	479	2.66

NLTK (Natural Language Toolkit)를 사용하여 일반적인 불용어를 제거하고 모든 텍스트를 소문자로 변환하였다. 또한, 정규 표현식을 통해 특수 문자를 제거하고 MBTI 유형을 수치형 레이블로 변환하는 등의 데이터 전처리를 수행하였다.

클래스 불균형 문제를 해결하기 위해 층화 샘플링 (Stratified Sampling)을 수행하여 181개의 인스턴스를 가지는 ESFJ 유형을 제외한 모든 유형은 360개의 인스턴스를 가지도록 하였다.

학습 모델은 사전 학습된 BERT 모델 (bert-base-uncased)을 텍스트 기반의 데이터를 사용하여 전이학습 (Transfer learning)을 통해 미세조정 (fine-tuning)하여 MBTI 성격 유형 분류에 특화된 모델을 얻었다.

학습을 위한 데이터 배치는 PyTorch의 Dataset 클래스를 사용하였다. 게시물은 BERT 토큰나이저를

사용하여 최대 512개의 토큰으로 토큰나이징 하고, 특수 토큰을 사용하여 모든 입력 시퀀스를 같은 길이로 맞추도록 시퀀스 패딩을 수행하였다. 데이터 세트를 10-fold 교차 검증 (Cross-Validation) 방법을 이용하여 학습/테스트 세트로 분할 하였으며 배치의 크기는 28로 설정하였다. 과대 적합 (Over-fitting) 을 방지하기 AdamW 옵티마이저를 사용하였으며, 메모리 제약을 피하기 위해서 4단계의 그라디언트 누적을 활용하고 AMP (Automatic Mixed Precision) 패키지의 GradScaler를 사용하여 혼합 정밀도 학습을 수행하였다 [9].

모델의 성능은 학습 손실, 정확도, 정밀도, 재현율 및 F1-Score를 사용하여 평가하였다. 모델은 폴드 당 10에포크 (Epoch)동안 학습하고 매 에포크당 평가지표를 통해 학습의 진행을 확인하였다.

학습곡선 (learning rate)과 혼동행렬 (confusion matrix) 분석을 통해서 모델의 과대 적합과 과소 적합 (Under-fitting) 발생 여부를 확인하고 학습의 진행상황을 확인하였다.

초기 학습에서 학습률을 낮추어 모델이 점진적으로 패턴을 학습하도록 유도하는 워밍업 (warm-up) 기법을 적용하여 학습 초기에 모델이 수렴하지 못하거나 특정 패턴에 과도하게 편향되지 않도록 하였다. 이를 통하여 학습 후반부로 갈수록 손실 값이 안정적으로 감소하며 각 유형에 대한 예측 정확도가 고르게 나타나도록 개선하고자 하였다.

### III. 연구 결과

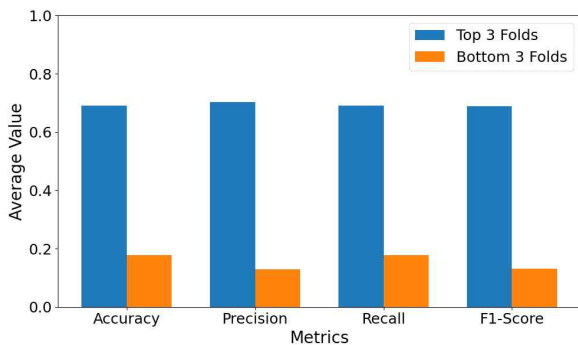


그림 1. 상위 성능 폴드와 하위 성능 폴드 간의 비교  
Fig. 1. The comparison between high-performance folds and low-performance folds

10-fold CV를 통해 분할된 폴드의 학습에서 특정 폴드가 매우 낮은 성능을 보이는 것을 확인하였다 (그림 1). 해당 폴드의 혼동행렬에서 16개의 성격 유형 중 일부 유형만 집중적으로 예측하는 결과를 보여서, 지역 최적해 (local optimum)에 고착된 것을 확인하였다 (그림 2).

이를 해결하기 위하여 학습 초기 단계에서 학습률을 점진적으로 증가시키는 워밍업 기법을 적용하였다. 워밍업 단계를 전체 학습 과정의 10%인 0.1로 설정하여 초기에 낮은 학습률에서 시작해 학습을 안정적으로 진행시키고 결과를 확인하였다. 그 결과, 워밍업 기법 적용 전 기본 모델은 약 27.5개의 예측에 성공하였으나, 워밍업 적용 모델은 평균 31.5개의 예측에 성공한 것을 확인하였다. (그림 3, 4).

위의 결과를 바탕으로 워밍업 기법을 적용하여 모델의 성능이 향상된 것을 확인하였으며, 이후에는 학습 손실, 정확도, 정밀도, 재현율, F1-Score의 학습곡선을 각각 확인하였다.

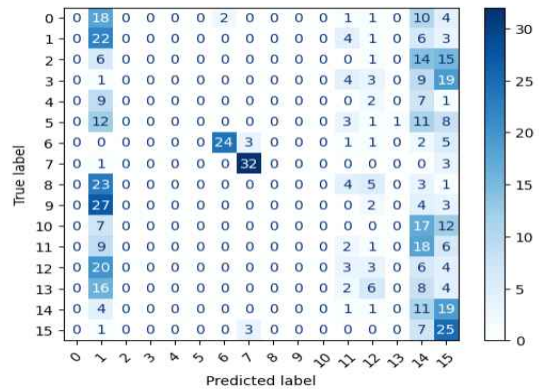


그림 2. 성능이 낮은 폴드의 혼동행렬  
Fig. 2. Confusion matrix of low-performance fold

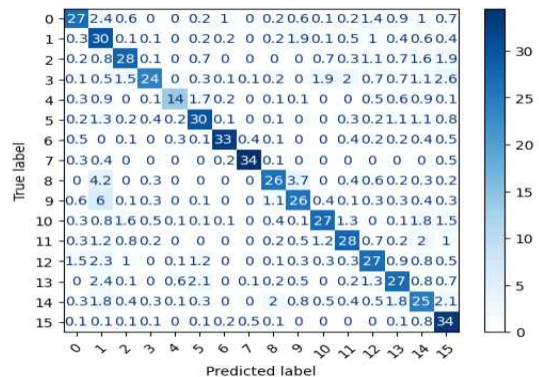


그림 3. 워밍업 비적용 모델의 혼동행렬  
Fig. 3. Confusion matrix of model w/o warm-up applied

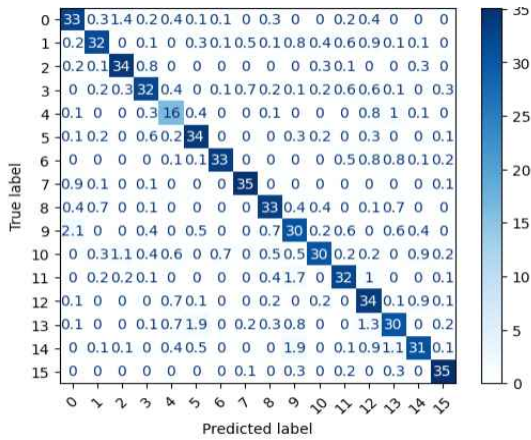


그림 4. 워밍업 적용 모델의 혼동행렬

Fig. 4. Confusion matrix of model with warm-up applied

훈련 손실 함수 값은 2.79로 시작하여 0.47까지 감소한 것을 확인하였으며, 테스트의 경우에는 2.75에서 1.02까지 감소한 것을 확인하였다 (그림 5). 훈련 정확도는 7%에서 시작하여 88.69%까지 상승한 것을 확인하였으며, 테스트의 경우도 73.75%까지 상승한 것을 보였다 (그림 6). 훈련 정밀도, 재현율, F1-Score는 훈련이 진행되면서 지속적으로 상승하여 마지막에는 90%에 가까운 값을 보였다 (그림 7-9). 테스트 정밀도, 재현율, F1-Score 또한 점차 상승하여 최종적으로 각각 74.79%, 73.75%, 73.76%를 보였다 (그림 7-9).

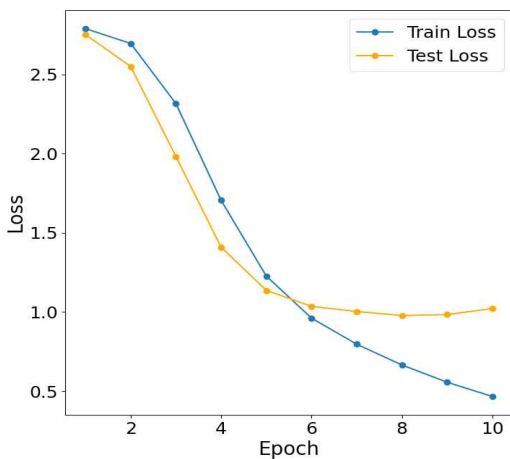


그림 5. Epoch에 따른 손실 함수 변화

Fig. 5. Loss function changes according to epoch

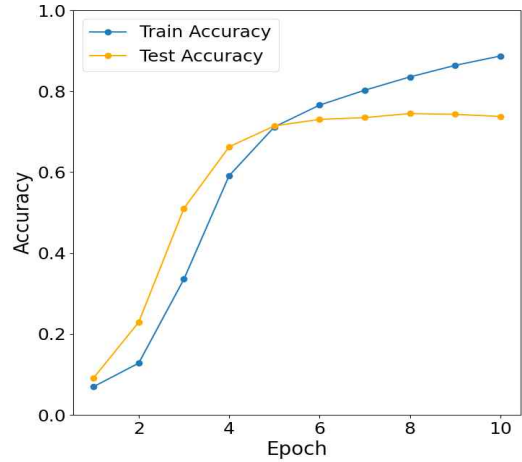


그림 6. Epoch에 따른 정확도 변화

Fig. 6. Accuracy changes according to epoch

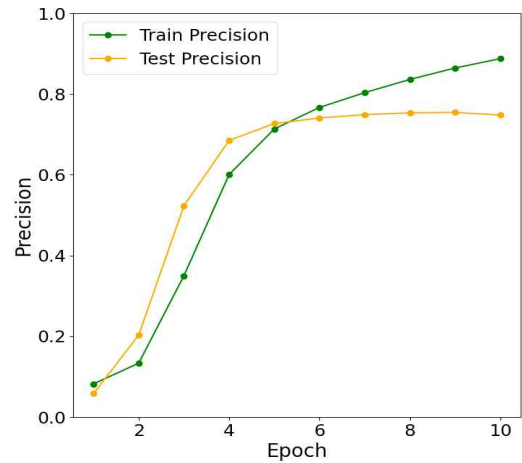


그림 7. Epoch에 따른 정밀도 변화

Fig. 7. Precision changes according to Epoch

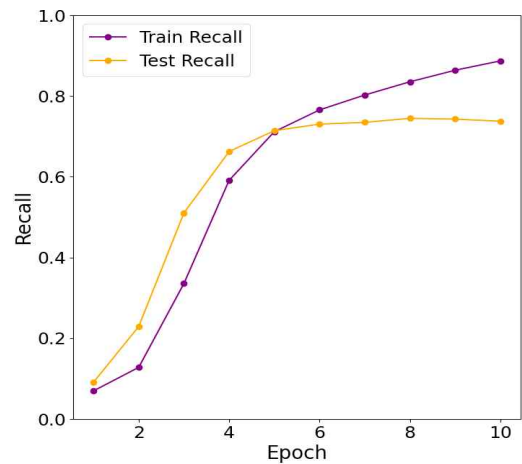


그림 8. Epoch에 따른 재현율 변화

Fig. 8. Recall changes according to epoch

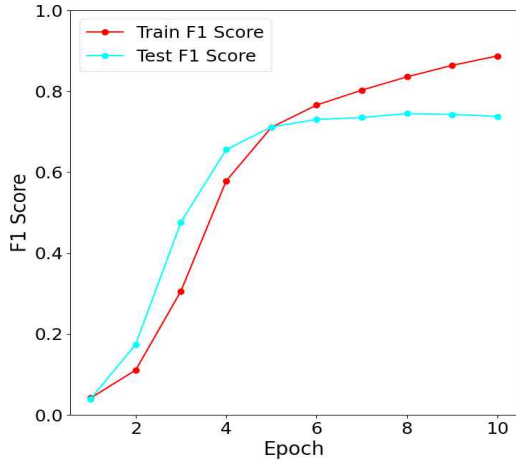


그림 8. Epoch에 따른 F1-Score 변화  
Fig. 8. F1-Score changes according to epoch

웜-업 기법을 적용하기 전과 후의 모델의 성능을 종합적으로 분석한 결과, 웜-업 적용 모델은 기존 모델에 비하여 훈련 손실 함수 값이 26.11%가 감소하였으며 테스트 손실 함수 값 또한 34.5%의 감소를 보였다 (표 2).

이외의 다른 평가지표의 값도 모두 23-24%의 차이를 보여 웜-업 기법을 적용한 모델의 성능이 향상되었음을 확인하였다 (표 2).

표 2. 웜-업 적용 전후 모델 간 성능 비교  
Table 2. Performance comparison between models before and after applying warm-up

	기존 모델		웜-업 적용 모델		적용 전후 성능 변화 (%)	
	훈련	검증	훈련	검증	훈련	검증
손실 함수값	0.632	1.560	0.466	1.022	73.9	65.5
정확도	0.790	0.602	0.886	0.737	112.2	122.5
정밀도	0.744	0.601	0.887	0.747	119.3	124.4
재현율	0.790	0.602	0.886	0.737	112.2	122.5
F1-Score	0.716	0.593	0.887	0.737	123.9	124.3

#### IV. 결 론

본 연구에서는 BERT 모델을 활용한 텍스트 기반 MBTI 성격 유형 분류 예측 모델을 제안하였으며, 웜-업 기법을 적용하여 성능의 향상을 보였다. 이 연구를 통하여 기존의 설문지 기반 성격 검사 방식

과 다른 SNS와 같은 일상생활 속에서 수집 가능한 텍스트 데이터를 통하여 보다 자연스럽게 객관적인 성격 유형 분류 모델을 제시하고자 하였다. 이는 설문 유형의 데이터가 가지는 한계를 극복하고 사용자의 접근성을 높여서 시간과 비용을 절감과 편리성을 추구할 수 있을 것이다.

본 연구는 Reddit에서 수집된 영어 텍스트 데이터에 한정되어 있어 언어적·문화적 다양성을 충분히 반영하지 못한 한계가 있다. 때문에, 향후 연구에서는 다른 언어와 문화권에서의 데이터 세트를 추가하여 모델을 개선하고, 대용량 텍스트 데이터에서도 일관된 성능을 보일 수 있도록 모델의 확장성이 요구된다. 또한, BERT 이외에도 GPT나 XLNet과 같은 다른 자연어 처리 모델을 적용하거나 앙상블 모델을 구축함으로써 성능을 더욱 향상시키는 방안도 생각해 볼 수 있다.

결과적으로 본 연구와 같은 인공지능 기술 기반의 객관적이며 효율적인 성격 유형 분류 연구를 통해서 심리 상담과 교육 분야의 발전을 기대할 수 있다.

#### 참 고 문 헌

- [1] Goldberg, L. R. (1993). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 66-100). Guilford Press.
- [2] Vistorte AOR, Deroncele-Acosta A, Ayala JLM, Barrasa A, López-Granero C and Martí-González M (2024) Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review. *Front. Psychol.* 15:1387089. doi: 10.3389/fpsyg.2024.1387089
- [3] Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd ed.). Consulting Psychologists Press.
- [4] Gardner, W. L., & Martinko, M. J. (1996). Using the Myers-Briggs Type Indicator to study

- managers: A literature review and research agenda. *Journal of Management*, 22(1), 45-83.
- [5] Hayeong Ryu and Youngwan Cho, "Analysis of Personality Types Based on Natural Language Responses Using LSTM", *Journal of Korean Institute of Intelligent Systems*, Vol. 34, No. 4, pp. 356-363, 2024.
- [6] Mushtaq, Zeeshan, Sagar Ashraf, and Nosheen Sabahat. "Predicting MBTI personality type with K-means clustering and gradient boosting." 2020 IEEE 23rd International Multitopic Conference (INMIC). IEEE, 2020.
- [7] Furnham, A., Moutafi, J., & Crump, J. (2003). The relationship between the revised NEO-personality inventory and the Myers-Briggs type indicator. *Social Behavior and Personality: an international journal*, 31(6), 577-584
- [8] Ashok Kumar Jayaraman, Gayathri Ananthakrishnan, Tina Esther Trueman, Erik Cambria, Chapter Four - Text-based personality prediction using XLNet, Editor(s): Preetha Evangeline David, P. Anandhakumar, *Advances in Computers*, Elsevier, Volume 132, 2024.
- [9] Rogers, Anna, Olga Kovaleva, and Anna Rumshisky. "A primer in BERTology: What we know about how BERT works." *Transactions of the Association for Computational Linguistics* 8 (2021): 842-866.
- [10] <https://www.kaggle.com/>

# HMD를 이용한 6축 매니플레이터 원격 제어 구현

이아선\*, 박주환\*\*, 옥승호\*\*\*

## Implementation of Remote Control for a 6-Axis Manipulator Using HMD

A-Sun Lee\*, Joo-Hwan Park\*\* and Seung-Ho Ok\*\*\*

### 요 약

산업용 매니플레이터의 수요가 증가함에 따라 가상환경에서 매니플레이터를 원격 제어하는 디지털트윈 연구가 활발히 진행되고 있다. 디지털트윈은 사용자에게 위험한 작업 환경의 접근을 최소화하며 비용과 시간을 절감할 수 있다. 본 논문에서는 Head Mounted Display(HMD)를 이용해 가상환경에서 6축 매니플레이터를 원격 제어하는 환경을 구현하였다. 사용자가 HMD를 통해 디지털트윈 환경에 접속하고, TCP/IP 통신으로 가상환경의 6축 매니플레이터 엔드이펙터 좌표값을 실제 환경 속 6축 매니플레이터에게 전달함으로써 사용자가 VR 컨트롤러의 움직임으로 6축 매니플레이터를 원격 제어할 수 있다.

### Abstract

As the demand for industrial manipulators increases, research on remotely controlling manipulators in virtual environments through digital twins is actively progressing. Digital twins minimize user exposure to hazardous work environments while saving costs and time. In this paper, we implemented an environment for remotely controlling a 6-axis manipulator in a virtual setting using a Head Mounted Display (HMD). Users connect to the digital twin environment through the HMD, and the end effector coordinate values of the 6-axis manipulator in the virtual environment are transmitted to the actual 6-axis manipulator via TCP/IP communication. This allows users to remotely control the 6-axis manipulator through the movements of the VR controller.

### Key words

digital twin, 6 axis manipulator, head mounted display, virtual reality, unreal engine

## 1. 서 론

산업용 매니플레이터의 수요가 증가함에 따라 가

상환경에서 매니플레이터를 원격 제어하는 디지털트윈 연구가 활발히 진행되고 있다[1,2]. 디지털트윈 기술은 사용자에게 위험한 작업 환경의 접근을 최

\* 동의대학교 로봇자동화공학과, 20222153@office.deu.ac.kr

\*\* 동의대학교 인공지능학과, 20246146@office.deu.ac.kr

\*\*\* 동의대학교 로봇자동화공학과, osh@deu.ac.kr(교신저자)

※ 이 논문은 교육부의 지방대학활성화사업 사업비를 지원받아 작성되었음

※ 본 논문(저서)는 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임

소화하며 비용과 시간을 절약한다[3]. 또한 사용자에게 고정된 시야가 아닌 1인칭, 3인칭 등 다양한 시야에서 작업 환경을 관찰할 수 있는 환경을 제공해 작업의 효율성을 높인다[4].

본 논문은 Head Mounted Display(HMD)를 이용한 6축 매니플레이터 원격 제어를 구현하였다. 사용자는 HMD를 이용해 Unreal Engine으로 구축된 가상 환경에 접속하며 가상환경에서 VR 컨트롤러를 통해 6축 매니플레이터를 원격 제어한다.

## II. 본 론

그림 1은 Unreal Engine 기반 디지털트윈 동작 흐름도이다. 본 논문에서는 Unreal Engine 5.0.3을 사용하여 가상환경을 구현하였고 ZEUS의 6축 매니플레이터와 Oculus3 VR HMD와 컨트롤러를 사용하였다. 가상환경의 VR 컨트롤러 좌표값은 TCP/IP Socket 통신을 이용해 실시간으로 실제 환경의 6축 매니플레이터에게 전달하였다.

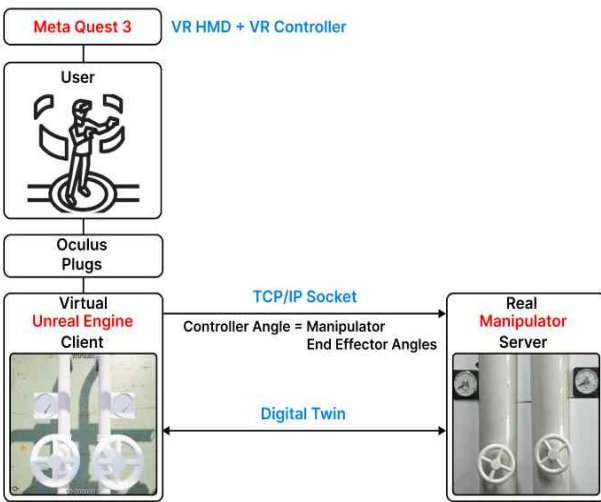


그림 1. Unreal Engine 기반 디지털트윈 동작 흐름도  
Fig. 1. Digital Twin operation flowchart based on Unreal Engine

그림 2는 Unreal Engine을 사용하여 6축 매니플레이터 엔드이펙터의 1인칭 시점 구현 방법을 나타낸다. Unreal Engine의 Component에 가상현실에서 사용자를 의미하는 Pawn을 배치하였다. VR 컨트롤러의 역할인 MotionController에 사용자의 팔움직임을

나타내기 위해 SkeletalMesh를 사용하였으며 팔의 움직임에 따라 Staticmesh의 매니플레이터 엔드이펙터가 구동되도록 하였다.

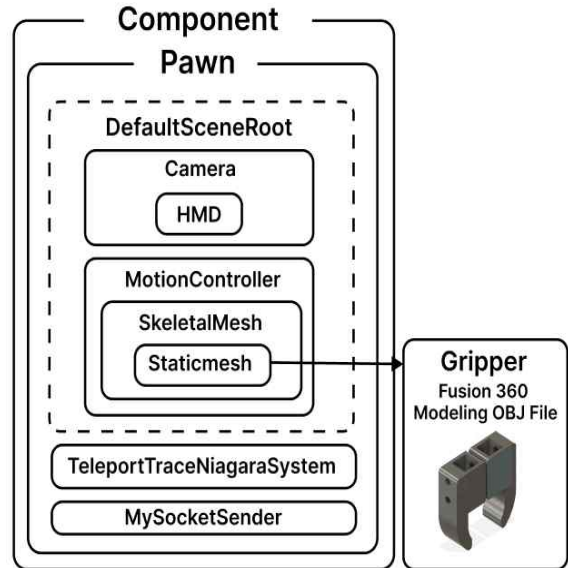


그림 2. Unreal Engine을 사용한 엔드이펙터 1인칭 시점 구현 방법

Fig. 2. Method for implementing first-person perspective of end effector using Unreal Engine

실제 환경에서 6축 매니플레이터의 움직임은 가상환경에서 사용자가 VR 컨트롤러를 움직일 때 변화하는 좌표값을 실시간으로 받아 동작한다. 그림 3은 통신에 사용된 Unreal Engine 블루프린트이다. 블루프린트는 Unreal Engine에서 제공하는 비주얼 스크립팅 시스템이다. 프로그래밍 언어 없이도 여러 노드를 서로 연결하고 관계를 정의함으로써 프로그래밍할 수 있다. 하지만 특정 좌표값인 VR 컨트롤러의 값을 통신하기 위해 C++을 이용해 TCP/IP Socket 통신의 Socket Actor를 생성 후 블루프린트에 사용하였다.

그림 4는 Unreal Engine 기반 가상환경에서 사용자가 관찰할 수 있는 환경이다. 사용자는 HMD를 통해 그림 속 A의 위치인 6축 매니플레이터 엔드이펙터의 1인칭 시점으로 가상환경에서 작업을 수행할 수 있다.

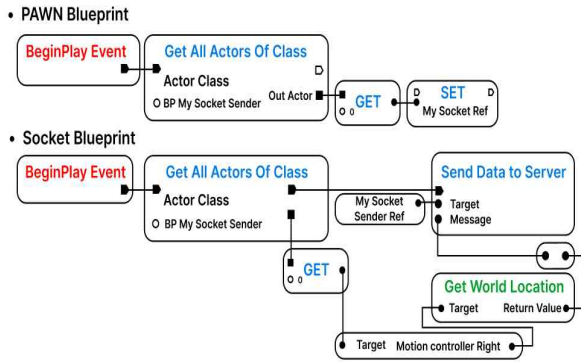


그림 3. Unreal Engine TCP/IP socket 통신 블루프린트 노드 연결 방식

Fig. 3. Unreal Engine TCP/IP socket communication Blueprint Node connection method

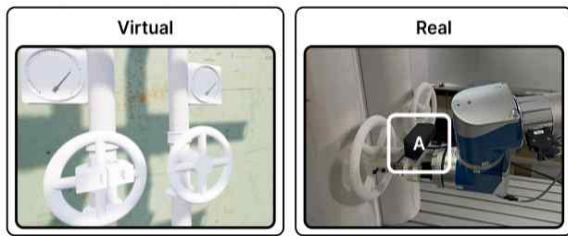


그림 4. 가상환경에서 사용자가 작업하는 시야

Fig. 4. User's view in a virtual environment

### III. 결 론

본 논문에서는 6축 매니퓰레이터 원격 제어를 위한 Unreal Engine 기반 가상환경에서 VR 컨트롤러를 사용해 6축 매니퓰레이터 원격 제어를 구현하였다. HMD를 착용한 사용자는 Unreal Engine 기반 디지털트윈 환경에 접속하여 6축 매니퓰레이터 엔드이펙터의 1인칭 시점에서 작업이 가능한 것을 확인하였다. 따라서 가상환경에서 6축 매니퓰레이터 엔드이펙터 좌표값인 VR 컨트롤러의 좌표값을 실시간으로 실제 환경의 6축 매니퓰레이터에게 통신되며 사용자는 가상환경에서 실제 환경의 6축 매니퓰레이터를 제어할 수 있다. 향후 매니퓰레이터에 카메라를 연동해 실시간으로 작업 현장을 모니터링 가능한 시스템을 연구할 계획이다.

### 참 고 문 헌

[1] J. W. Lee, J. H. Kim, H. Khan, S. J. Abbasi and

M. C. Lee, "Unity-Based Haptic Device Implementation of a Virtual Robot Manipulator Operation System", Journal of Institute of Control, Robotics and Systems, Vol. 28, no. 8, pp.740-748, 2022.

[2] H. Chen, F. Liu, Y. Yang and W. Meng, "MultiVR: Digital Twin and Virtual Reality Based System for Multi-people Remote Control Unmanned Aerial Vehicles," 2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV), pp. 647-652 2022.

[3] F. Oettl, J. Schoeler and J. Schilp, "Development of a method for evaluating the benefits of using a digital twin," 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Maldives, Maldives, pp. 1-7, 2022.

[4] Z. Li, D. Akkil and R. Raisamo, "Gaze Augmented Hand-Based Kinesthetic Interaction: What You See is What You Feel", IEEE Transactions on Haptics, vol. 12, no. 2, pp. 114-127, April-June 2019.



# 객체 추적을 활용한 교차로 지시 위반 차량 검출 알고리즘

박지은\*, 왕태수\*, 박서인\*, 장시웅\*\*

## An Algorithm for Detecting Vehicles Violating Intersection Signals Using Object Tracking

Ji-Eun Park\*, Tae-su Wang\*, Seo-In Park\*, and Si-Woong Jang\*\*

### 요 약

본 논문에서는 교차로 CCTV 영상에서 차선을 인식하지 않고도 차량의 궤적을 추적하여 직진 금지 표시를 위반하는 차량을 검출하는 알고리즘을 제안한다. YOLOv8 모델을 활용해 실시간으로 차량의 바운딩 박스를 인식하고, 중심점을 이용해 궤적을 추적한다. 궤적의 흔들림을 줄이기 위해 지수 평활법을 적용해 중심 좌표를 저장하며, 이 좌표들을 이용해 직선 궤적을 예측한다. 차량의 움직임이 특정 오차 범위 내에서 일정 프레임 동안 유지될 경우 직진 차량으로 판단한다. 이러한 기술은 스마트 교차로 구축과 교통 안전 효율성 증진에 기여할 것이다.

### Abstract

This paper proposes an algorithm for detecting vehicles that violate no-straight-driving signs at intersections by tracking their trajectories without the need for lane recognition in CCTV footage. Utilizing the YOLOv8 model, the algorithm identifies the bounding boxes of vehicles in real time and tracks their trajectories based on the center points. To minimize trajectory fluctuation, exponential smoothing is applied to store the center coordinates, which are then used to predict a straight trajectory. If a vehicle's movement remains consistent within a specific margin of error for a certain number of frames, it is classified as moving straight. This technology will contribute to the development of smart intersections and enhance traffic safety efficiency.

### Key words

Vehicle Detection, Trajectory Tracking, Traffic Safety, No-Straight-Driving Signs

## 1. 서 론

도로 주행 중 직진 금지 표시가 없는 좌회전, 우회전 차로에서 직진하는 것은 불법이 아니다. 그러

---

\* 동의대학교, 20213855@office.dew.ac.kr, tswang@office.deu.ac.kr, 20193104@office.deu.ac.kr

\*\* 동의대학교 스마트자동차연구소 연구소장, swjang@deu.ac.kr(교신저자)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 지역지능화혁신인재양성사업임(IITP-2024-RS-2020-II201791, 50%). 또한 본 연구는 중소벤처기업부의 기술혁신개발사업의 지원에 의한 연구임(RS-2024-00425148, 50%).

나 직진 금지 표시가 있거나, 금지 표시가 없더라도 다른 차량의 통행을 방해한다면 단속 대상이 될 수 있다[1]. 본 연구는 다양한 각도의 CCTV 영상에서 별도의 차선 인식 없이 차량의 궤적을 추적하고, 지시 위반 차량을 검출하는 것을 목표로 한다. 제안하는 알고리즘은 CCTV 각도와 차선의 영향을 받지 않기 때문에 다양한 환경에서 적용할 수 있다는 장점이 있으며, 이러한 기술은 스마트 교차로 구축과 교통 안전 효율성 증진에 기여할 것이다.

## II. 지시 위반 차량 검출 알고리즘

본 연구에서는 차량의 궤적을 추적하고 직진 여부를 판단하기 위해 YOLOv8 모델을 기반으로 알고리즘을 개발하였다. 알고리즘 테스트를 위해 부산교통정보센터에서 제공하는 도로 CCTV 영상을 사용하였다. CCTV 영상은 차량이 회전 중인지, 직진 중인지 명확히 구별되면서 교차로를 넓게 비추는 곳을 선정하였다.

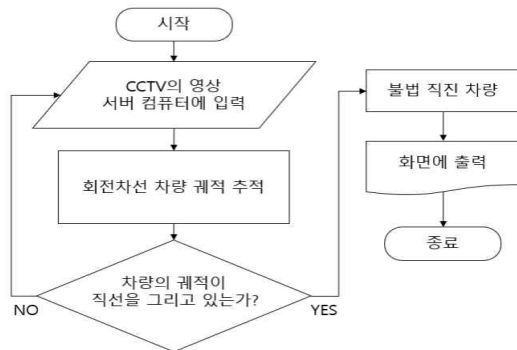


그림 1. 흐름도  
Fig. 1. Flow Chart

그림 1의 알고리즘 흐름은 다음과 같다. CCTV 영상에서 차량이 움직이면 추적을 시작한다. 궤적은 차량의 바운딩 박스의 중심점을 따라 그려지게 되는데, 실시간 인식 및 추적의 특성상 바운딩 박스가 흔들릴 수 있다. 따라서 지수 평활법을 사용하여 전처리 과정을 거친 후 중심 좌표를 저장한다. 이는 현재 좌표와 이전 좌표 간 가중 평균을 계산하는 방식으로, 안정적인 궤적을 생성하는 데 도움을 준다[2].

이렇게 저장된 중심 좌표는 첫 번째 좌표를 시작점, 마지막 좌표를 끝점으로 하는 직선의 기울기를 계산한다. 이때 각 좌표에서 x축의 위치를 바탕으로 해당 직선 위에서 기대되는 y축 값을 예측한다. 기대 y값과 실제 y값의 차이를 비교해 허용 오차를 벗어나지 않고 일정 프레임 이상 주행한다면 해당 차량은 직선 주행 중인 차량으로 판단한다. 직진 중인 차량의 궤적은 영상에 표시된다. 다음의 그림 2, 3은 서로 다른 각도에서 촬영된 교차로 영상에서 알고리즘을 적용한 결과이다.

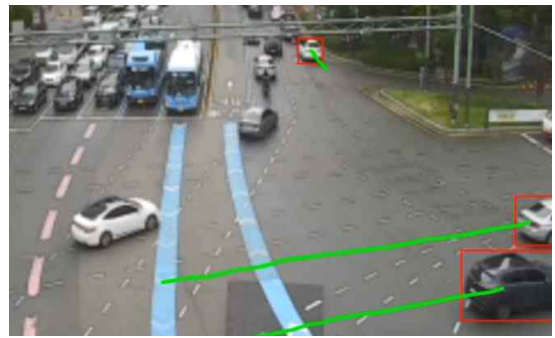


그림 2. 교차로 테스트 사진  
Fig. 2. Intersection Test Picture



그림 3. 도로 테스트 사진  
Fig. 3. Road Test Picture

## III. 결 론

본 연구는 다양한 CCTV 각도와 환경에서 차선 인식 없이 차량의 궤적을 추적하고 직진 금지 표시 위반 여부를 판단하는 알고리즘을 제안하였다. 이는 지능형 교차로 시스템을 구현하는 데 효과적인 도구가 될 수 있으며, 교통 혼잡 완화 및 안전 강화에 기여할 가능성이 크다.

참 고 문 헌

- [1] The Jeon-buk Ilbo Internet Article, "Vehicles that drive straight as left-turn vehicles 'often' violate laws", <https://www.jjan.kr/article/20210530733242>, Jun 2021.
- [2] Tae-Jin Park and Gab-Sig Sim, "A Comparative study on smoothing techniques for performance improvement of LSTM learning model", Journal of The Korea Society of Computer and Information Vol. 28 No. 1, pp. 17-26, Jan 2023.

# 합성 데이터를 활용한 객체 검출 성능 평가

김경훈\*, 박주환\*\*, 옥승호\*\*\*

## Performance Evaluation of Object Detection Using Synthetic Data

Gyeong-hun Kim\*, Joo-Hwan Park\*\* and Seung-Ho Ok\*\*\*

### 요약

합성 데이터는 대규모 실제 데이터 셋 수집의 시간 및 비용 문제를 해결할 수 있는 효과적인 대안으로 주목받고 있다. 본 논문에서는 물류 창고에서 매니플레이터가 다루는 상자 객체에 대한 합성 데이터를 생성하고, 그 개수를 다양하게 조절하여 YOLOv8 모델을 학습하였다. 이후 500장의 실제 데이터만으로 학습한 모델과 성능을 비교하였다. 실험 결과, 합성 데이터의 개수가 증가함에 따라 모델의 mAP(IOU  $\geq$  0.5) 성능이 지속적으로 향상되었으며, 특히 4배(2000장) 이상의 합성 데이터를 활용했을 때 실제 데이터만을 사용한 모델의 성능인 85.2%를 넘어 86.3%의 성능을 달성하였다. 이를 통해 합성 데이터의 효과성을 확인할 수 있었다.

### Abstract

Synthetic data has garnered attention as an effective alternative to address the time and cost challenges of large-scale real-world dataset collection. In this paper, we generated synthetic data for box objects handled by a manipulator in a warehouse environment and adjusted the quantity of this data to train a YOLOv8 model. We then compared its performance with a model trained solely on 500 real images. Experimental results showed that as the amount of synthetic data increased, the model's mAP (IOU  $\geq$  0.5) performance steadily improved. Notably, using synthetic data exceeding four times (2,000 images) the amount of real data achieved an mAP of 86.3%, surpassing the 85.2% performance of the model trained solely on real data. This confirms the effectiveness of synthetic data.

### Key words

synthetic data, object detection, performance evaluation, YOLOv8

## 1. 서론

대량의 실제 데이터 셋을 확보하고 레이블 작업

을 진행하는데 많은 시간과 비용이 소요되는 어려움이 있다[1]. 합성 데이터는 컴퓨터 시뮬레이션과 인공지능 기술을 통해 생성되며, 실제 데이터와 유

\* 동의대학교 로봇자동화공학과, 20202135@office.deu.ac.kr

\*\* 동의대학교 인공지능학과, 20246146@office.deu.ac.kr

\*\*\* 동의대학교 로봇자동화공학과, osh@deu.ac.kr(교신저자)

※ 이 논문은 교육부의 지방대학활성화사업 사업비를 지원받아 작성되었음

※ 본 논문(저서)는 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임

사한 품질을 유지하면서도 다양한 환경에서 대량의 데이터를 효율적으로 생성할 수 있다[2]. 이러한 특성으로 인해 합성 데이터는 컴퓨터 비전과 객체 탐지 분야에서 주목받고 있다[3].

물류 창고 등 산업 현장에서 매니플레이터의 활용도가 증가하며, 상자 분류, 적재, 하역 등 다양한 작업에 사용되고 있다[4]. 본 논문에서는 매니플레이터가 주로 다루는 객체인 상자를 기반으로 합성 데이터를 생성하였다. 생성된 합성 데이터의 개수를 다양하게 조절하여 학습한 YOLOv8 모델[5]과, 실제 데이터로 학습한 YOLOv8 모델의 mAP(IOU  $\geq 0.5$ ) 성능 비교를 통해 합성 데이터의 개수에 따른 성능 차이를 분석하였다.

## II. 본 론

그림 1은 컴퓨터 시뮬레이션 기반 합성 데이터를 생성하는 흐름도이다. 시뮬레이션 도구로는 높은 현실감과 다양한 환경 조건을 설정할 수 있는 NVIDIA 사의 Isaac Sim을 활용하였다. 실제 상자를 AR Code 앱을 이용해 3D 모델로 변환하는 작업으로 시작된다. 생성된 3D 모델은 Isaac Sim에 추가되며, 이를 기반으로 다양한 합성 데이터를 생성하게 된다. 합성 데이터 생성 과정에서 Isaac Sim의 합성 데이터 생성 도구인 Replicator가 활용된다. Replicator 도구는 상자의 포즈, 위치, 카메라의 각도, 조명의 밝기 및 위치, 배경 등을 무작위로 설정하는 Domain Randomization 기법을 통해 다양한 환경을 생성한다. 이를 통해 생성된 합성 데이터는 2D 이미지와 2D 바운딩 박스 정보를 포함하며, YOLOv8 모델의 학습에 사용된다.

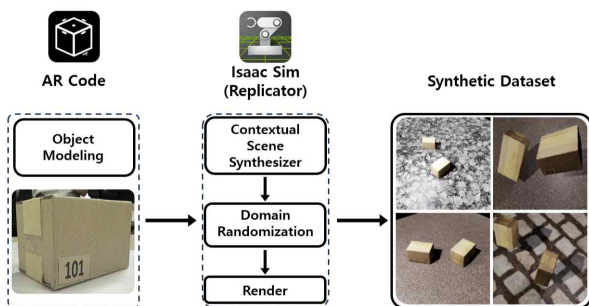


그림 1. 합성 데이터 생성 흐름도  
Fig. 1. Flowchart for synthetic data generation

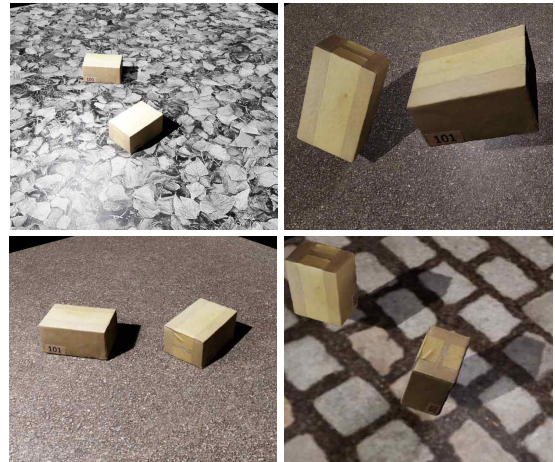


그림 2. 합성 데이터 예시 사진  
Fig. 2. Example of synthetic data image

그림 2는 이러한 방법으로 생성된 합성 데이터 이미지를 보여준다.

실험에서는 실제 데이터 500장을 사용하고, 합성 데이터의 개수를 다양하게 조절하여 YOLOv8 모델을 학습하였다. 합성 데이터는 실제 데이터의 2배, 3배, 4배, 5배, 10배, 20배에 해당하는 1000장, 1500장, 2000장, 2500장, 5000장, 10000장으로 준비하였고, 학습(train) 데이터와 검증(validation) 데이터의 비율은 8:2로 구성하였다. 이후, 그림 3과 같은 학습 데이터에 포함되지 않은 100장의 실제 데이터를 사용하여 mAP(IOU  $\geq 0.5$ ) 성능 평가를 진행하였다.

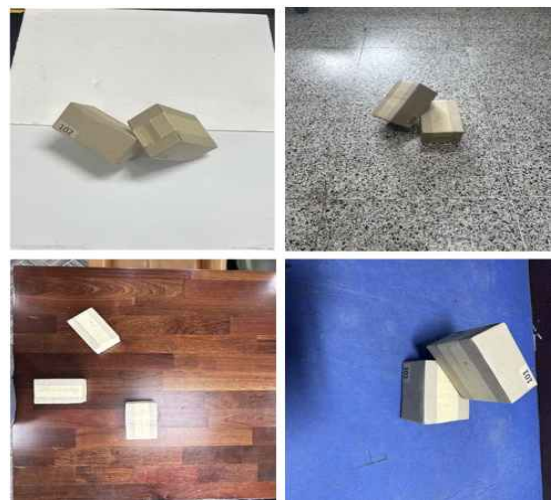


그림 3. 테스트 데이터 예시 사진  
Fig. 3. Example of test data image

그림 4는 mAP(IOU  $\geq$  0.5) 성능 평가 결과이다. 합성 데이터의 개수가 증가할수록 모델 성능은 점진적으로 향상되었다. 실제 데이터만 사용한 모델의 성능은 85.2%였지만, 합성 데이터를 4배(2000장) 사용했을 때는 86.3%의 성능을 보였다. 이후 합성 데이터의 개수를 증가시켜 20배(10000장)까지 사용한 경우 모델의 성능은 89.3%에 도달하며 합성 데이터를 4배(2000장) 사용했을 때 보다 더 높은 성능을 달성하였다.

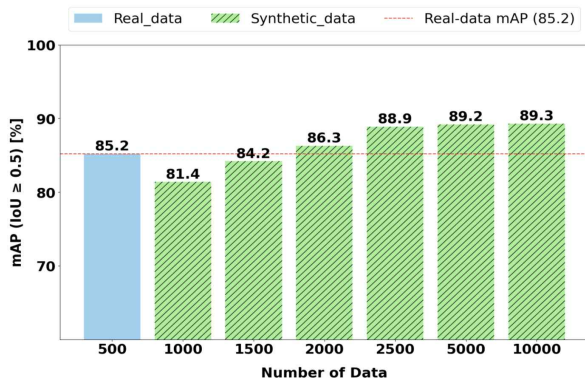


그림 4. 성능 평가 결과

Fig. 4. Performance evaluation results

### III. 결 론

본 논문에서는 물류 창고에서 매니퓰레이터가 다루는 상자 객체에 대한 합성 데이터를 활용하여 객체 검출 성능을 평가하였다. 합성 데이터의 개수를 다양하게 조절하여 YOLOv8 모델을 학습한 후 실제 데이터를 사용한 모델과 성능을 비교하였다. 그 결과, 합성 데이터의 개수가 증가할수록 모델 성능이 개선되었으며, 실제 데이터의 4배(2000장) 이상의 합성 데이터를 사용했을 때 실제 데이터만을 사용한 모델의 성능인 85.2%보다 향상된, 86.3%의 성능을 보였다. 또한, 실제 데이터의 20배(10000장)까지 늘렸을 때 89.3%의 성능을 확인할 수 있었다.

그러나 합성 데이터는 가상으로 생성되기 때문에 실제 환경의 미세한 표면 질감, 반사, 조명 조건과 같은 현실적 차이를 완전히 극복하기는 어렵다. 따라서 향후 연구에서는 실제 환경에서 수집한 데이터를 함께 사용하여 분석 결과의 정확성을 높일 계획이다.

### 참 고 문 헌

- [1] A. Damian et al., "Experimental Results on Synthetic Data Generation in Unreal Engine 5 for Real-World Object Detection", Proc. EMES, pp. 1-4, 2023.
- [2] L. C. Adi and T. M. Cheng, "Analysis of Impact of Synthetic Image Data with Multiple Randomization Strategies on Object Detection Performance", Proc. ECICE, pp. 206-210, October 2022.
- [3] X. Yang, X. Fan, J. Wang and K. Lee, "Image translation based synthetic data generation for industrial object detection and pose estimation", IEEE Robotics and Automation Letters, vol. 7, no. 3, pp. 7201-7208, July 2022.
- [4] F. Yumbra et al., "Design of a Conveyor Belt Manipulator for Reposition of Boxes in Logistics Centers", Proc. CASE, pp. 2225-2230, October 2022.
- [5] M. Safaldin, N. Zaghden and M. Mejdoub, "An improved YOLOv8 to detect moving objects", IEEE Access, vol. 12, pp. 59782-59806, 2024.

# 클라우드 후원을 위한 LLM 기반 오픈소스 프로젝트 분석

권혁원\*, 김재형\*, 고강민\*, 김한나\*, 변영철\*\*

## Analysis of LLM-based Open Source Projects for Crowdfunding

Hyeokwon Kwon\*, Jaehyung Kim\*, Kangmin Ko\*, Hanna Kim\*, and Yung-Cheol Byun\*\*

### 요 약

오픈소스 소프트웨어는 Linux, React, Flutter 등 널리 사용되는 프로젝트들의 근간이 되어 현대 기술 발전의 핵심 역할을 담당하고 있다. 그러나 오픈소스 커뮤니티 내 자원 분배의 불균형으로 인해 다수의 소규모 프로젝트들이 지속적인 성장에 어려움을 겪고 있다. 본 연구에서는 이러한 문제를 해결하기 위하여 오픈소스 프로젝트 개발자와 후원자를 연결하는 플랫폼을 구성하는 핵심 요소로서 오픈소스 분석 및 후원을 연계하는 오픈소스 서포터 시스템에 관하여 연구하였다. 본 연구는 균형 있는 자원 분배를 통해 소규모 프로젝트들의 지속가능성을 제고하는 것을 목표로 한다. 후원자는 포인트 시스템을 통해 프로젝트에 기여할 수 있으며, 개발자는 이를 통해 프로젝트의 지속적인 운영이 가능하다. 이를 위하여 오픈소스 서포터 시스템의 설계와 주요 기능을 설명하고 이를 통해 오픈소스 커뮤니티의 자원 불균형 문제를 해소하고자 한다.

### Abstract

Open-source software underpins widely used projects such as Linux, React, and Flutter, playing a crucial role in modern technological advancements. However, resource allocation imbalances within the open-source community hinder the sustainable growth of many smaller projects. To address this challenge, this study investigates an open-source supporter system that connects project developers and supporters, integrating open-source project analysis with funding mechanisms as core components of a comprehensive platform. The primary objective is to enhance the sustainability of smaller projects through equitable resource distribution. Supporters can contribute to projects via a points-based system, thereby enabling developers to sustain ongoing operations. This paper outlines the design and key functionalities of the open-source supporter system, aiming to mitigate resource disparities within the open-source community.

### Key words

open source, sustainability, resource disparity, support platform, project funding

---

\* 제주대학교 컴퓨터공학과, prince5390@gmail.com, Hiruca3220@gmail.com, aknlz652@naver.com, govmo1282@gmail.com,

\*\* 제주대학교 컴퓨터공학과, ycb@jejunu.ac.kr(교신저자)

※ 본 연구는 중소벤처기업부의 벤처스타트업 아카데미 사업의 지원을 받아 수행된 연구결과임.

## 1. 서 론

오픈소스 소프트웨어[1]는 현대 기술 환경에서 중요한 역할을 하며, Linux, React, Flutter와 같은 많은 기술적 혁신의 기반이 되어 왔다. 오픈소스의 발전은 많은 개발자가 자유롭게 협력하고 지식을 공유하는 방식으로 가능해졌으며, 이를 통해 개발자들이 손쉽게 다양한 툴과 라이브러리를 이용할 수 있게 되었다. 그러나 모든 오픈소스 프로젝트가 동일한 지원을 받지 않는다. 일부 대규모 프로젝트는 대기업이나 재단으로부터 충분한 자원을 제공받아 성장할 수 있었지만, 자원 부족으로 인해 많은 소규모 프로젝트들은 지속 가능한 성장을 이루지 못하고 있다. 이는 오픈소스 생태계 내에서 자원의 불균형을 초래하고, 새로운 프로젝트가 발전할 수 있는 기회를 제한하는 결과를 낳고 있다.

이러한 문제의 해결방안으로서 본 논문에선 오픈소스 프로젝트의 개발자와 후원자를 연결하는 '오픈소스 서포터' 플랫폼을 제안한다. 본 플랫폼은 오픈소스 프로젝트의 개발자(수혜자)와 후원자를 연결하여 소규모 프로젝트들이 필요한 자원을 확보할 수 있도록 돕고, 이를 통해 오픈소스 생태계의 포용성을 높이고자 한다. '오픈소스 서포터'는 후원자들이 포인트 기반 시스템을 통해 특정 프로젝트에 기여하고, 수혜자는 이를 통해 필요한 자원을 지원받는 구조를 갖추고 있다. 이러한 시스템은 후원자들에게는 간편한 기여 방식을, 수혜자들에게는 지속 가능성을 제공하여, 오픈소스 생태계의 자원 불균형 문제를 해소하는 데 기여하고자 한다.

## II. 관련 연구

최근 LLM(Large Language Model)을 활용한 코드 분석 및 프로젝트 평가 연구가 활발히 진행되고 있다. 특히 Open AI사의 GPT 기반 모델들은 코드 이해와 분석에서 높은 성능을 보이고 있다. 오픈소스 프로젝트의 품질 평가는 전통적으로 코드 매트릭스, Commits 빈도, 이슈 해결 속도 등 정량적 지표를 중심으로 이루어져 왔다. 그러나 최근에는 LLM을 활용하여 코드의 구조적 특성뿐만 아니라 문서화

수준, 코딩 컨벤션 준수 여부, API 설계의 일관성 등 질적인 측면까지 종합적으로 평가하는 방법론이 제시되고 있다.

자연어 처리 기술을 활용한 프로젝트 분석은 코드뿐만 아니라 README 파일, 이슈 댓글, 풀 리퀘스트 설명 등 프로젝트의 텍스트 데이터를 포괄적으로 분석할 수 있다는 장점이 있다. 최근에는 GPT-4o의 등장으로 더욱 정교한 의미론적 분석이 가능해졌으며, 코드와 문서 간의 일관성 검증, 보안 취약점 탐지, 사용자 피드백 분석 등 다양한 영역으로 활용이 확장되고 있다.

오픈소스 프로젝트의 가치와 잠재력을 평가하기 위해서는 체계적인 분석 지표가 필요하다. 본 연구에서는 Github 플랫폼에서 제공하는 REST API [2]를 통해 각 레포지토리 별로 수집 가능한 공개 데이터를 기반으로 다음과 같은 분석 지표를 설계하였다.

첫째, 레포지토리 소유자의 개발 역량을 평가하는 지표이다. 전체 Stars 수는 개발자가 만든 프로젝트들의 인지도와 유용성을 나타내며, 총 Commits 수와 Pull Requests 수는 지속적인 개발 활동과 협업 능력을 반영한다. Issues 처리 건수는 문제 해결 능력을, Total Contributions는 오픈소스 생태계 전반에 대한 기여도를 보여준다. 이러한 지표들을 통해 프로젝트 소유자의 전반적인 신뢰도를 평가할 수 있다.

둘째, 프로젝트의 현재 상태를 평가하는 지표이다. Repository Description과 README 문서의 품질은 프로젝트의 목적 명확성과 문서화 수준을 나타낸다. 이는 프로젝트의 접근성과 유지보수 용이성을 판단하는 중요한 기준이 된다. LLM을 활용하여 이러한 텍스트 데이터를 정성적으로 평가함으로써, 프로젝트의 완성도를 측정할 수 있다.

셋째, 프로젝트의 발전 가능성을 예측하는 지표이다. 일 단위로 수집된 1년간의 커밋 이력은 프로젝트의 개발 활동 트렌드를 보여준다. 최근의 커밋 빈도와 패턴 분석을 통해 프로젝트의 활성화 수준과 지속가능성을 평가할 수 있다. 이러한 지표들은 LLM을 통해 종합적으로 분석되며, 각 지표의 중요도는 프로젝트의 특성과 목적에 따라 차등적으로



적용된다. 본 연구에서는 이러한 다차원적 분석을 통해 투자 가치 평가의 객관성과 신뢰성을 확보한다.

오픈소스 프로젝트의 투자 가치를 분석하기 위해서는 앞서 설계한 다양한 지표들을 종합적으로 평가할 수 있는 방법론이 필요하다. 본 연구에서는 GPT-4o와 같은 고성능 LLM을 활용하여 정량적 데이터와 정성적 데이터를 통합 분석하는 방법을 제안한다.

LLM을 활용한 분석은 크게 세 단계로 구성된다. 첫째, 프롬프트 엔지니어링을 통해 수집된 데이터를 LLM이 이해하고 처리할 수 있는 형태로 구조화한다. 개발자의 활동 지표, 프로젝트 현황 데이터, 시계열 정보 등을 명확한 컨텍스트와 함께 제공하여 분석의 정확도를 높인다. 본 연구에서는 아래 [그림 1]과 같은 Open AI사의 ‘AI Assistant’ [3] 기능을 활용하여 System Instructions, Functions를 커스터마이징하여 엔지니어링을 마친 어시스턴트를 생성한 후, 이를 REST API를 통해 호출하여 사용한다.

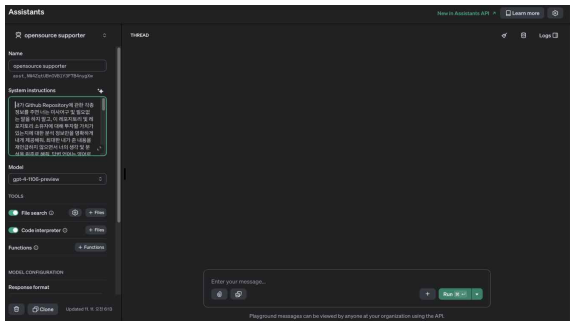


그림 1. AI Assistant 커스터마이징 화면

둘째, LLM은 입력받은 데이터를 바탕으로 프로젝트의 현재 가치와 미래 잠재력을 평가한다. 이 과정에서 README의 품질, 커밋 패턴의 일관성, 개발자의 전문성 등 다양한 요소들을 종합적으로 고려한다. 특히 자연어 처리 능력을 활용하여 레포지토리 문서의 전문성과 완성도를 평가한다.

셋째, 분석 결과를 명확한 투자 권고 의견으로 도출한다. 단순한 수치적 평가를 넘어, 투자 결정의 근거가 되는 핵심 요인들을 구체적으로 제시한다. 이는 투자자가 프로젝트의 가치를 객관적으로 판단할 수 있는 기준을 제공한다. 이러한 LLM 기반 분

석 방법론은 기존의 정량적 지표 중심 평가 방식의 한계를 보완하고, 보다 포괄적이고 심층적인 프로젝트 가치 평가를 가능하게 한다.

### III. 오픈소스 서포터 시스템 구현

본 연구에서 제안하는 오픈소스 서포터 시스템은 LLM 기반 프로젝트 분석 엔진과 포인트 기반 후원 시스템으로 구성된다. 전체 시스템 아키텍처는 [그림 2]와 같이 프론트엔드와 백엔드 서비스로 구분된다. 프론트엔드는 React 기반으로 구현되어 AWS S3를 통해 서비스되며, 백엔드는 Spring Boot 기반의 마이크로서비스 아키텍처로 설계되었다.

프로젝트 분석 엔진은 Github REST API를 통해 프로젝트 및 개발자 데이터를 수집하고, 사전에 튜닝을 마친 LLM 모델에 REST API 호출을 함으로써 이를 분석하고 프로젝트의 투자 가치를 평가한다. 수집된 데이터는 MySQL 데이터베이스에 저장되며, 전체 시스템은 AWS EC2 인스턴스에서 운영된다. 이러한 클라우드 기반 아키텍처를 통해 시스템의 확장성과 안정성을 확보하였다.

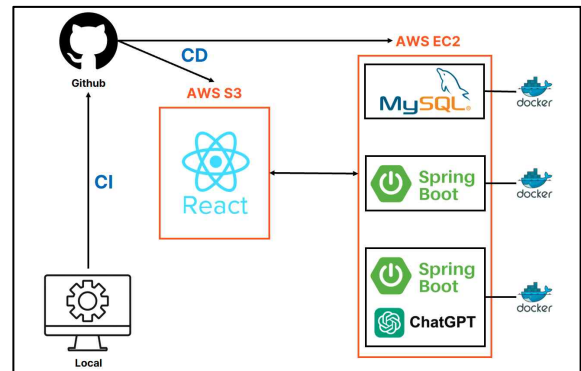


그림 2. 서포터 시스템 아키텍처

오픈소스 서포터 시스템은 [그림 3]과 같이 직관적인 메인페이지를 통해 사용자들에게 서비스의 핵심 가치와 기능을 전달한다. 후원이 필요한 프로젝트들을 쉽게 탐색할 수 있도록 설계되었으며, 프로젝트 분석 결과를 한눈에 파악할 수 있는 인터페이스를 제공한다.

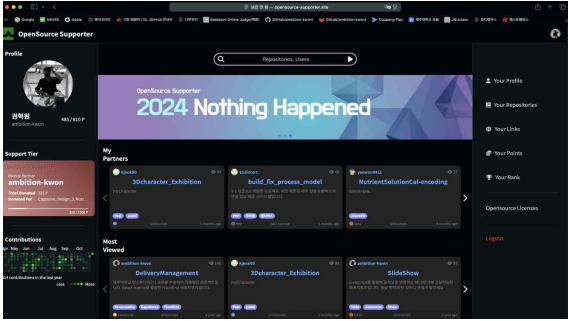


그림 3. 오픈소스 서포터 메인 페이지

프로젝트 분석을 위한 데이터 수집은 Github REST API를 통해 이루어진다. 수집되는 정보는 Stars, Commits, Pull Requests, Issues, Contributions 등 개발자의 전반적인 활동 지표와 함께, README 문서 내용, Repository Description, 주간 커밋 이력 등 특정 레포지토리의 상세 정보를 포함한다. [그림 4-1, 2]는 이러한 데이터를 시각화하여 보여주는 프로젝트별 상세 정보 화면이다.

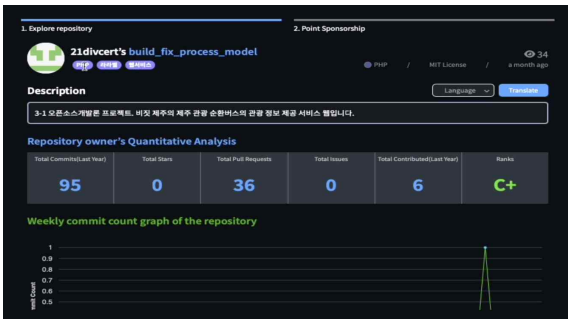


그림 4-1. 레포지토리 상세정보 모달(정량적 분석)

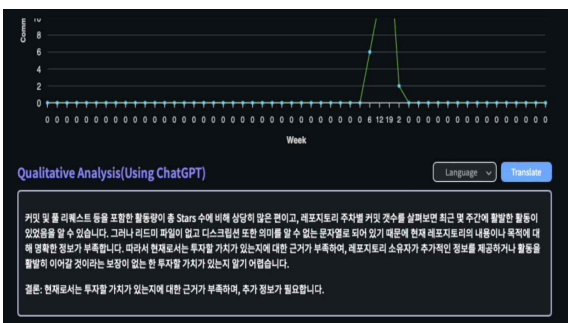


그림 4-2. 레포지토리 상세정보 모달(정성적 분석)

분석 결과를 바탕으로 [그림 5]와 같이 후원자들은 관심 있는 프로젝트에 포인트를 기부할 수 있

며, 이는 프로젝트 개발자들에게 실질적인 지원으로 전환된다.

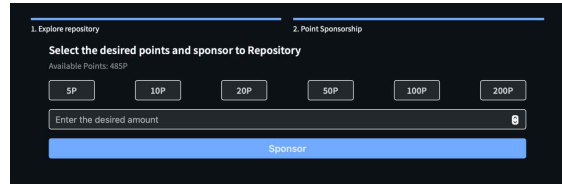


그림 5. 레포지토리 후원 모달

본 플랫폼은 오픈소스 생태계의 자원 불균형 문제 해소를 위한 새로운 접근 방식을 제시하며, LLM 기술을 활용한 객관적인 프로젝트 가치 평가와 효율적인 후원 시스템의 결합을 통해 지속 가능한 오픈소스 개발 환경 구축에 기여하고자 한다.

#### IV. 결 론

본 연구는 LLM을 활용하여 오픈소스 생태계의 자원 불균형 문제를 해소하기 위한 '오픈소스 서포터' 플랫폼을 설계하고 구현하였다. 본 플랫폼은 Github API를 통해 수집된 프로젝트 데이터를 LLM으로 분석하여 투자 가치를 평가하고, 이를 바탕으로 후원자와 개발자를 연결하는 포인트 기반 후원 시스템을 제공한다.

파일럿 테스트 결과, LLM 기반의 프로젝트 분석은 기존의 정량적 평가 방식을 보완하여 프로젝트의 잠재력과 가치를 다각적으로 평가할 수 있음을 입증하였다. 특히 깃허브 스타트와 레포지토리 평가 시스템은 프로젝트의 가치를 효과적으로 전달하는 것으로 나타났으며, '오픈소스 서포터' 발전 가능성에 대한 높은 긍정적 평가는 본 플랫폼이 오픈소스 생태계의 지속 가능한 발전에 실질적으로 기여할 수 있음을 시사한다.

LLM을 활용한 프로젝트 가치 평가와 포인트 기반 후원 시스템의 결합을 통해, 오픈소스 프로젝트의 지속 가능한 발전을 위한 실질적인 방안을 제시하였다는 점에서 의의가 있다. 향후 연구에서는 실제 서비스 운영을 통한 장기적 효과 검증과 함께, 프로젝트 분석 지표의 고도화 및 후원 시스템의 개선 및 정량적, 정성적 지표를 활용한 레포지토리 추

천 시스템을 개발할 계획이다.

## 참 고 문 헌

- [1] Red Hat, "What is open source?," Red Hat Enterprise,  
<https://www.redhat.com/ko/topics/open-source/what-is-open-source>, 2024/11/10
- [2] Github, 리포지토리에 대한 REST API 엔드포인트, Github Docs,  
<https://docs.github.com/ko/rest/repos/repos?apiVersion=2022-11-28#get-a-repository>, 2024.11.10.
- [3] OpenAI, "Assistants API Overview," OpenAI Platform Documentation,  
<https://platform.openai.com/docs/assistants/overview>, 2024.11.10
- [4] Matt Asay, "칼럼 | 오픈소스 지속가능성 문제... 기업 간 후원이 답이다," CIO Korea,  
<https://www.cio.com/article/3519187>, 2024.11.10
- [5] Bill Doerrfeld, "오픈소스, 지속 가능한 자금 조달 방법을 상상하다" ITWorld Korea,  
<https://www.itworld.co.kr/news/352475>, 2024.11.10
- [6] Dries Buytaert, "더 강력한 오픈소스 프로젝트를 위한 3가지 제안, "공정성, 혜택, 균형"" ITWorld Korea,  
<https://www.itworld.co.kr/news/136023>, 2024.11.10

# 심전도 R점 검출 기술을 이용한 개선된 태깅 환경 개발

백승민\*, 조은총\*, 문승현\*, 반성범\*\*, 박은상\*\*

## Development of Upgraded Tagging Environment Utilizing a ECG R Detection Technique

Baek Seung-Min\*, Cho Eun Chong\*, Moon Seung Hyun\*, Pan Sumbum\*\*, and Bak EunSang\*\*

### 요 약

본 논문에서는 심전도 신호의 특이점 중의 하나인 R 점 (R peak)을 검출하는 새로운 알고리즘을 제안하고, 이를 태깅도구에 적용하여 태깅 환경을 개선시킨다. 제안된 알고리즘은 중앙값 기반의 적응형 문턱값 설정, 주기 기반 검출, 그리고 평균값 기반의 단계적 적용을 통해 R 점의 검출 정확도를 향상시킨다. 실험 결과를 보면 제안한 알고리즘은 이전보다 더 높은 정확도를 보였으며, 수동 태깅 도구와 결합하여 반자동 태깅 도구로 전환함으로써 태깅 수행 시간을 크게 단축시켰다.

### Abstract

This paper proposed a new algorithm for detecting R-peaks in electrocardiogram (ECG) signals, creating an enhanced tagging environment. The proposed algorithm improves R-peak detection accuracy through adaptive thresholding based on median values, period-based detection, and stepwise adaptations using average values. Experimental results demonstrate that the algorithm achieved higher accuracy than the existing method. Applying the algorithm to a manual tagging tool to advance it into a semi-automatic tagging tool significantly reduced processing time.

### Keywords

ECG, R-peak detection, Adaptive Thresholding, Period-based Detection, Tagging Tool

## 1. 서 론

심전도 신호(Electro-cardiograph : ECG)는 심장의 활동으로 인해 생기는 전기적 신호로서 비침습적인 방법으로 심장의 활동 상태를 인지할 수 있기 때문에 심장 질환을 일차적으로 진단하는 데 널리 사용되고 있다. 심전도 신호의 R 점은 심전도 신호내에서 상대적으로 가장 큰 진폭을 가지므로 심박수

측정에 주로 사용되며, R-R interval을 통해 심박수를 계산할 수 있다. 그러나 전력선 간섭, 기저선 출력임 (drift)와 같은 잡음 및 왜곡이 존재하기 때문에 R 점 검출은 단순하지 않은 작업이다

기존의 심전도 신호의 R 점 검출 방법은 크게 수동으로 사람이 직접 찾는 방법과 특정 알고리즘을 활용하여 반자동으로 찾는 방법이 있다. 각 신호마다 R 점 검출을 수동으로 찾는 것은 매우 비효율

\* 조선대학교 전자공학과, Email: baegs4854@gmail.com, dmschd0033@gmail.com, moon1128@chosun.kr

\*\* 조선대학교 IT 연구소, Email(교신저자): sbpan@chosun.ac.kr, bakeunsang@chosuna.c.kr

적이기 때문에 반자동 방법을 주로 활용한다.

초기 알고리즘은 신호의 기울기를 극대화하여 R 점을 찾는 푸리에 변환과 같은 주파수 분석 기법이 널리 사용되었다. Pan과 Tompkins의 알고리즘은 필터링과 적응형 임계값을 사용하여 높은 정확도로 R 점을 검출하는 초기 접근 중 대표적인 방법으로 잘 알려져 있다. 그 밖에 머신 러닝, 딥러닝을 기반으로 한 다양한 기법들도 제안되었다.

본 논문에서는 간단한 계산량의 알고리즘을 제안하고, 이것이 태깅 도구에 효율적으로 사용되는 것을 소개하고자 한다.

## II. 제안하는 방법

그림 1은 심전도 신호에서 R 점을 검출하기 위해 단계별로 최적화된 알고리즘을 도식화하였다. 각 단계는 신호의 특성을 분석하고, 정확한 R 점을 검출하는 데 중점을 두었다. 이 과정은 크게 네 단계로 나뉜다.

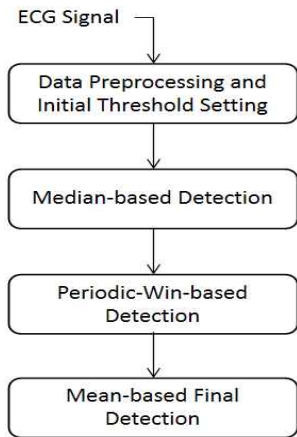


그림. 1. 심전도 신호 검출 순서도

### 1. 전처리 및 중앙값 기반 1차 검출

신호값 상위 1000개를 내림차순으로 정렬하고, 이 값들의 중앙값을 기준으로 문턱값을 정한다. 1차로 R 점을 검출하는 방법은 심전도 신호를 사람의 평균 심박 주기로 구간을 나누고, 그 구간에 있는 신호 중 문턱값 이상인 점들 중에 최대값을 R 점으로 검출한다.

### 2. 중앙값 기반 2차 검출

1차 검출과 달리 두 번째 단계에서는 1차에서 검

출한 R점 (정확히는 R점이 될 수 있는 후보)들의 중앙값을 기반으로 다시 R점을 검출하였다. 즉 1차로 검출된 R점의 중앙값을 기반으로 새로운 문턱값을 만들고, 심전도 신호를 일정 구간으로 분리하여 각 구간에서 문턱값 이상인 값들 중 최대값을 찾아 R 점으로 검출하였다.

### 3. 주기 기반 3차 검출

이전 단계에서 검출된 R 점들을 이용하여 해당 데이터의 실제 심박수를 계산하고 이를 활용하여 R 점의 검출 성능을 개선하였다. 즉 개인별 심박 주기를 활용한 것이다.

### 4. 평균값 기반 4차 검출

네 번째 단계에서는 지금까지 검출된 R점의 평균값을 이용하여 최종 R점을 선정하였다. 검출된 R 점 중에서 평균값 절반 이하인 R점을 제거함으로써, 신호의 잡음이나 잘못된 검출로 인한 오류를 줄이고 최종적인 R점을 검출하는 정확도를 높였다.

## III. 실험 결과

본 논문에서 제안하는 알고리즘의 성능을 평가하기 위해 제안한 알고리즘을 MATLAB을 사용하여 구현하였으며, 심전도 신호는 조선대학교 IT 연구소에서 직접 취득한 14개의 데이터를 사용하였다. 비교 대상으로는 심전도 신호의 최대값을 활용하여 R 점의 후보를 선정한 후, 사람의 평균 심박수에 의한 윈도우 값을 기반으로 R 점을 검출하는 알고리즘을 기존 방법으로 삼았다.

그림 2에서 확인할 수 있듯이, 기존 방법으로는 R 점을 찾지 못하거나 잘못 찾았는데, 개선된 방법에서는 이를 해결하였다.

본 논문에서 제안하는 알고리즘을 적용하기 이전에는 그림 3과 같이 심전도 신호의 R 점을 수동으로 태깅하였고, 각 심전도 데이터 당 약 2분 30초 정도의 태깅시간이 소요되었다. 하지만 제안한 알고리즘을 수동태깅에 적용하여 반자동 방법으로 태깅을 하면, 약 10초가 소요되어 15배 정도의 시간적 이득을 얻을 수 있었다.

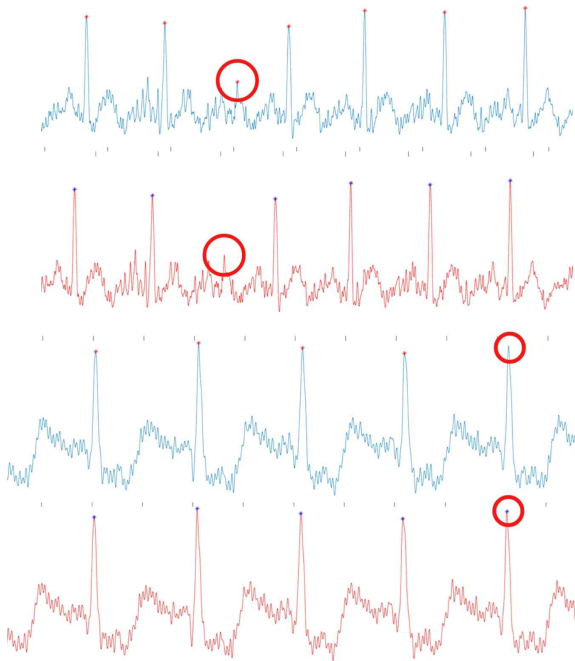


그림 2. T점을 R점으로 잘못 검출한 기존 방법(파랑신호), 기존 문제를 해결한 개선된 방법 (빨강신호)



그림 3. 수동으로 R 점을 태깅하는 모습

#### IV. 결 론

본 논문에서는 심전도 신호에서 R 점을 검출하기 위한 새로운 알고리즘을 제안하였다. 제안된 알고리즘은 기존의 문제점들을 해결하기 위해 중앙값 기반의 적응형 문턱값 설정, 주기 기반 검출, 그리고 평균값 기반의 단계적 검출을 도입하여 보다 높은 정확도를 달성하였다. 실험 결과, 제안된 알고리즘은 기존의 방법에 비해 높은 정확도로 R 점을 검출하였다. 기존 방법에서는 R 피크와 T 피크를 명확하게 구분하지 못하거나 R 피크를 누락하는 문제가 있었지만, 제안된 알고리즘은 이러한 문제점을 개선하였다.

또한, 본 알고리즘을 수동 태깅 도구와 결합하여

반자동 태깅 도구로 전환함으로써, 심전도 데이터의 R 피크를 검출하는 데 소요되는 시간을 크게 단축시킬 수 있었다. 이로써 보다 효율적으로 심전도 데이터를 분석할 수 있게 되었다.

하지만, 심전도 신호의 비이상적 변동성에 의해 R 점보다 T 점이 더 큰 경우와 주기성이 어긋나는 경우가 발생하면, R 점을 정확히 검출하지 못하는 문제점이 있었다. 다양한 데이터를 기반으로 제안된 알고리즘의 성능을 더욱 검증하고, 실시간 신호 처리에 적합한 형태로 개선하는 연구가 필요할 것이다. 이러한 추가 연구를 통해 본 알고리즘이 실제 환경에서 활용될 수 있는 가능성을 높이고, 심전도 신호 분석의 정확도와 태깅 효율성을 향상시킬 수 있을 것으로 기대된다.

#### 참 고 문 헌

- [1] Optimal Threshold Setting Method for R Wave Detection According to The Sampling Frequency of ECG Signals, Ik-sung Cho· Hyeog-soong Kwon, Jul 2017.
- [2] A Critical Review of Various Peak Detection Techniques of ECG Signals, DESH DEEPAK GAUTAM, V.K. GIRI.
- [3] Noise-robust electrocardiogram R-peak detection with adaptive filter and variable threshold, MD Saifur Rahman<sup>1</sup>, Chul-Hyung Choi<sup>1</sup>, Si-Kyung Kim<sup>1</sup>, In-Deok Park, Young-Pil Kim.
- [4] New ECG R-peak detection algorithm for reducing complexity, Hwang La Young, Nam Yun Chan, Kim Seung Chan, Jang Yu Rim, Shin Tae Mi, 2017.

# VR 환경에서 손목 회전을 이용한 문자 입력 방법

김민우\*, 강창구\*\*

## Text Input Method Using Wrist Rotation in a VR Environment

Minwoo Kim\* and Changgu Kang\*\*

### 요 약

IT 기술의 발전으로 VR(Virtual Reality) 장비의 하드웨어와 소프트웨어가 고도화되고, 여러 IT 기업들이 다양한 VR 기기를 출시하고 있다. HMD(Head Mounted Display) 가격 하락과 콘텐츠 제작 증가로 VR 기술은 빠르게 대중화되고 있으며, VR 콘텐츠 개발과 효율적인 상호작용 방법에 대한 연구도 활발히 진행되고 있다. 현재 가상 쿼티 키보드를 활용한 문자 입력은 컨트롤러의 방향에 따라 레이저 포인터로 키를 선택하는 방식이 일반적이다. 그러나 이 방법은 사용자에게 피로감을 주고 긴 문장 입력 시 속도가 느리다는 문제점이 있다. 이러한 문제를 해결하기 위해 본 논문에서는 쿼티 배열을 유지하며 손목 회전을 이용한 문자 입력 방식을 제안한다.

### Abstract

With the advancement of IT technology, the hardware and software of VR (Virtual Reality) devices have become more sophisticated, and various IT companies are launching diverse VR equipment. The decline in prices of HMDs (Head Mounted Displays) and the increase in content production have led to the rapid popularization of VR technology, and research on VR content development and efficient interaction methods is also actively underway. Currently, the input method using a virtual QWERTY keyboard typically involves selecting keys with a laser pointer based on the direction of the controller. However, this method can cause user fatigue and is slow for entering long sentences. To address these issues, this paper proposes a text input method utilizing wrist rotation while maintaining the QWERTY layout

### Key words

Virtual reality, virtual text input, wrist rotation

## 1. 서 론

IT 기술의 발전으로 VR(Virtual Reality) 장비의 하드웨어 성능과 소프트웨어 기술은 고도화되고 있

으며, 애플, 메타, 마이크로소프트, 바이브 등 여러 IT 기업들이 다양한 VR 장비를 출시하고 있다. VR 장비에 대한 선택 폭이 넓어지고, HMD(Head Mounted Display) 가격이 하락하며 다양한 콘텐츠가

\* 경상국립대학교 컴퓨터공학과, rlaalsdn4560@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과, cgk@gnu.ac.kr(교신저자)

제작되면서 VR 콘텐츠는 빠르게 대중화되고 있다.

제조, 의료, 문화 콘텐츠 등 다양한 산업 분야에서 VR 기술을 활용한 콘텐츠 제작이 증가하고 있으며, 동시에 VR 환경에서 효율적인 상호작용 방법에 대한 연구도 활발히 진행되고 있다[1-4].

VR 환경에서의 상호작용에는 가상 환경 내의 3D 객체와 직접적인 상호작용, 그리고 가상 쿼티 키보드를 이용한 문자 입력이 포함한다. 현재 가상 쿼티(QWERTY) 키보드를 활용한 문자 입력 방법으로는 컨트롤러의 방향에 따라 레이저 포인터를 움직여 가상 키보드의 키를 선택하는 방식이 많이 사용되고 있다. 그러나 컨트롤러로 키보드 버튼을 조준하는 과정에서 사용자가 쉽게 피로감을 느낄 수 있으며, 특히 긴 문장을 입력할 때 문자 입력 속도가 느리다[2]. 이와 같은 문제점을 해결하기 위해 가상 키보드의 배열을 변경하거나 컨트롤러의 터치 패드를 활용하는 방법 등의 연구가 진행되었다[3,4]. 하지만 사용자들의 쿼티 배열의 익숙함으로 인해 속도 개선에는 한계점이 있었다. 본 논문에서는 쿼티 배열을 유지하면서 손목 회전을 이용한 문자 입력 방식을 제안한다.

## II. 손목 회전 기반 문자 입력

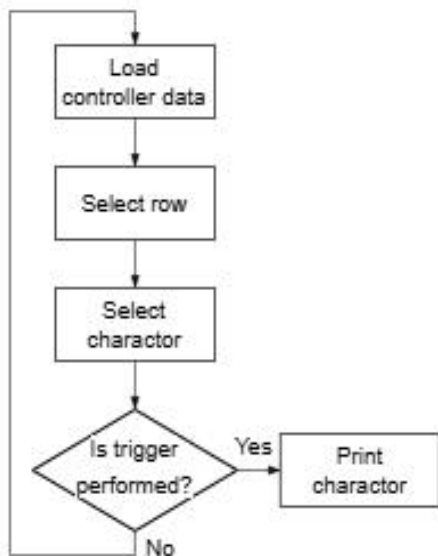


그림 1. 문자 입력 순서도  
Fig. 1. Text Input Flowchart

제안된 방법의 문자 선택은 컨트롤러의 회전각에 따라 결정된다. 적절한 문자가 선택되면 사용자는 트리거를 눌러 해당 문자를 입력할 수 있다. 입력 방식은 물리적 키보드 입력 방식과 같이 쿼티 배열을 좌우로 나누어 왼손과 오른손이 각각 맡은 영역으로 구성되며, 각 영역은 세 개의 행으로 구분되어 다섯 개의 문자를 선택할 수 있게 된다.

행 선택은 컨트롤러의 피치 회전각으로 수행되며, 사용자는 이를 조절하여 원하는 행을 선택할 수 있다. 선택된 행 위에는 인디케이터가 활성화된다. 문자의 선택은 컨트롤러의 롤 회전으로 결정되며, 각 문자는 특정 각도 범위를 갖는다. 사용자가 컨트롤러를 회전할 때 롤 회전각이 해당 문자의 범위에 포함되면 인디케이터가 표시된다. 이후 트리거를 누르면 인디케이터가 가리키는 문자가 입력된다.

손목 회전 기반 문자 입력 방식은 손목의 회전각과 쿼티 배열 가상 키보드를 이용한다. 손목의 회전각 데이터는 VR 컨트롤러의 자이로 센서를 통해 수집되며 피치(Pitch), 요(Yaw), 롤(Roll) 회전축의 각도로 나타낼 수 있다. 본 논문에서는 키보드의 행과 문자 선택을 위해 피치와 롤의 각도만 사용되었다.

각 축의 회전은 독립적인 기능을 수행하기 위해 다른 축에 영향을 미쳐서는 안 된다. 그러나 롤 회전 시 컨트롤러의 롤 회전축과 손목의 회전축이 일치하지 않아 의도치 않은 피치 각도 변화가 발생한다. 그림 2에서 적색 선은 컨트롤러의 롤 회전축, 녹색 선은 손목의 회전축을 나타낸다.

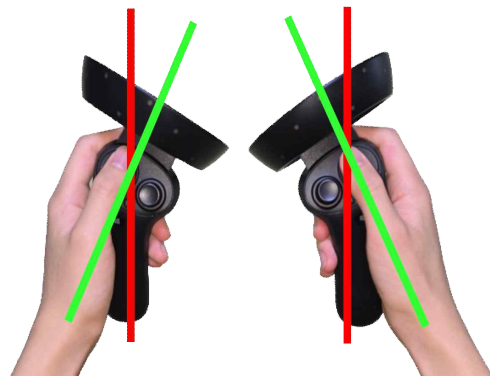


그림 2. 컨트롤러와 손목의 회전축 차이  
Fig. 2. Difference in rotation axis between controller and wrist



두 회전축을 일치시키기 위해 컨트롤러의 회전각 쿼터니언에 보정각 쿼터니언을 곱하여 두 회전을 합성하였으며, 그 결과 피치 각도의 변화를 최소화한다.

그림 3은 손목 회전 기반 가상 키보드의 구성을 나타낸다. 쿼터 배열을 양손에 할당하기 위해 좌우로 분할하고 각 영역을 세 개의 행으로 나누었다. 사용자가 손목의 피치를 조절하면 인디케이터가 상하로 이동하고, 롤을 조절하면 좌우로 이동하여 문자 위에 표시된다. 사용자가 원하는 문자 위에 인디케이터를 위치시켰을 때 트리거를 누르면 그림 4와 같이 해당 문자가 입력된다.

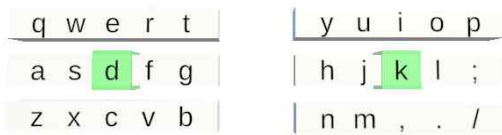


그림 3. 가상 키보드 구성  
Fig. 3. Virtual keyboard layout

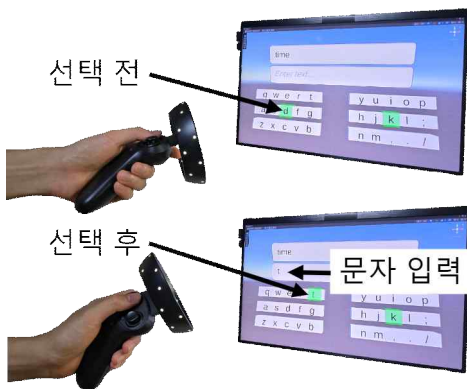


그림 4. 컨트롤러와 손목의 회전축 차이  
Fig. 4. Difference in rotation axis between controller and wrist

### III. 결 론

본 논문에서는 VR 환경에서 손목 회전을 이용한 문자 입력 방식을 제안했다. 향후 연구에서는 제안된 방식과 기존의 레이저 포인트 입력 방식을 비교하여 입력 속도 및 편의성을 분석하고 평가하겠다.

### 참 고 문 헌

- [1] S. M. Yun, C. S. Leem, and S. H. Ban, "A Study on Classification Model Development of Industry-Efficiency XR Technology and case Analysis," *Journal of Service Research and Studies*, vol. 12, no. 4, pp. 50-71, Dec. 2022.
- [2] J. S. Lee, H. S. Lee, and W. J. Choi, "Interactive Interface Design Through VR Hand Tracking," *The Journal of the Korea institute of electronic communication sciences*, vol. 18, no. 1, pp. 213-218, Feb. 2023.
- [3] K. Wang, Y. Yan, H. Zhang, X. Liu, and L. Wang, "Eye-shaped keyboard for dual-hand text entry in virtual reality," *Virtual Reality & Intelligent Hardware*, vol. 5, no. 5, pp. 451-469, Oct. 2023.
- [4] J. Wu, Z. Wang, L. Wang, Y. Duan, and J. Li, "FanPad: A Fan Layout Touchpad Keyboard for Text Entry in VR," *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp.222-232, Mar. 2024.

# SDR기술을 활용한 차량 간 무선통신 디스플레이 서비스

김주영\*, 김태준\*\*

## Wireless communication display service between vehicles using SDR technology

JuYeong Kim\* and Taejoon Kim\*\*

### 요 약

본 연구는 소프트웨어 정의 라디오 (SDR software defined radio) 기술을 기반으로 차량 간 무선통신 서비스를 개발하고 실용성을 검증하는 것을 목표로 한다. SDR의 주파수와 프로토콜 유연성을 활용하여 5.9 GHz V2X 대역 등 다양한 주파수로 신호를 실시간으로 수신하고 워터폴 디스플레이로 시각화하여 교통 정보와 안전 경고를 제공한다. 라즈베리파이를 서버로 활용한 프로토타입은 SDR 기술을 차량 통신에 적용할 수 있음을 확인하였다. 향후 고성능 프로세서를 통해 실시간 경고 기능을 추가하여 스마트 교통 시스템의 가능성을 확장할 계획이다. 본 연구는 SDR 기술이 차량 간 통신의 유연성과 효율성을 높여 안전하고 원활한 교통 환경을 구축에 기여할 수 있음을 시사한다.

### Abstract

This study aims to develop and validate an SDR-based vehicle-to-vehicle communication service. Leveraging SDR's adaptability in frequency, the system receives and visualizes signals in real-time, including the 5.9 GHz V2X band, providing traffic data and safety alerts via a waterfall display. The prototype, utilizing a RaspberryPi as a server, demonstrates the feasibility of SDR technology in vehicular communication. Future enhancements with high-performance processors will expand real-time alert functions, supporting SDR's potential to enhance flexible and efficient vehicle communication for safer and smoother traffic systems.

### Key words

Vehicle-to-Vehicle Communication, Software-Defined Radio (SDR), V2X,  
Real-Time Signal Visualization, Smart Traffic Systems

---

\* 충북대학교 정보통신공학부, rkswkdwhjtjs@naver.com

\*\* 충북대학교 정보통신공학부 교수, ktjcc@chungbuk.ac.kr

## 서 론

기술 혁신은 교통과 차량 산업 발전의 핵심 원동력이 되며, 차량 간 무선통신의 중요성도 높아지고 있다. 복잡한 현대 교통 시스템에서는 차량 간 실시간 정보 공유를 위한 다양한 기술이 연구되고 있으며, 본 연구에서는 소프트웨어 정의 라디오 (SDR) 를 무선통신을 지원하는 기술로 제안한다.

SDR 기반의 차량 간 무선통신 디스플레이 서비스는 실시간으로 다양한 차량 정보와 교통 데이터를 전달하여 운전자에게 시각적 정보를 제공하는 것을 목표로 한다. 이를 통해 교통 혼잡을 예측하고 안전한 운행 환경을 조성하며, 교통 흐름의 최적화에 기여할 수 있다[1]. 예를 들어, 앞 차량의 급정거나 도로 장애물 발생 정보를 즉시 공유하여 운전자가 빠르게 대처하고 교통사고를 예방할 수 있다.

전통적인 통신 시스템의 한계를 극복하고, SDR 기술로 사용하는 통신 주파수와 필요 기능을 실시간으로 조정하여 다양한 시나리오에서 최적화된 통신이 가능하다. 본 논문에서는 SDR 기술을 활용한 차량 간 무선통신 디스플레이 서비스를 구현하고, 그 효과와 이점을 분석하고자 한다. 이를 통해 교통 안전성을 높이고 교통 혼잡을 완화하는 방안을 모색하는 것이 본 연구의 목적이다.

## II. 차량 간 무선통신 서비스 구조

본 장에서는 SDR 기반 차량 간 무선통신 디스플레이 서비스의 아키텍처를 제안한다. 이 아키텍처는 차량 내부 통신 시스템과 중앙 서버 시스템으로 구성되며, 각 시스템은 차량 간 실시간 정보 교환과 데이터 처리를 통해 안전 운행을 지원한다. 차량 내부 통신 시스템은 데이터 수집 및 송수신을 담당한다. 다양한 센서가 차량의 속도, 위치, 가속도, 브레이크 상태를 모니터링하고 SDR 모듈에 데이터를 전달하여 디지털 신호로 변환한다. SDR 모듈은 RF 프론트엔드와 신호 처리 모듈로 구성되어 무선 신호를 송수신하고, 소프트웨어로 주파수와 프로토콜을 유연하게 관리하여 안정적인 통신을 가능하게 한다[2]. 통신 시스템은 차량 간 V2V (vehicle to

vehicle) 및 V2I (vehicle to infrastructure) 통신을 관리하며, 차량 간 애드혹 네트워크를 형성해 주변 정보를 공유하고 중앙 서버와 연결한다[3].

디스플레이 인터페이스는 교통 상황, 경고 메시지 등을 시각적·음성으로 운전자에게 제공해 즉각적 대응을 돕는다. 중앙 서버 시스템은 수집된 데이터를 분석하여 각 차량에 교통 정보와 안전 경고를 제공하며, AI 기반 예측 모델로 교통 혼잡을 예측하고 위험 구간을 파악한다. 네트워크 스위치와 라우터는 주파수 대역과 경로를 최적화하여 데이터 충돌을 방지하고, 인덱스 서버와 로그 관리 모듈은 데이터와 통신 기록을 관리해 시스템 안정성과 복구를 지원한다. 이 아키텍처는 WEB SDR의 유연성과 확장성을 최대한 활용하여 차량 간 실시간 정보 교환을 가능하게 하며, 스마트 교통 시스템으로의 발전 가능성을 높인다.

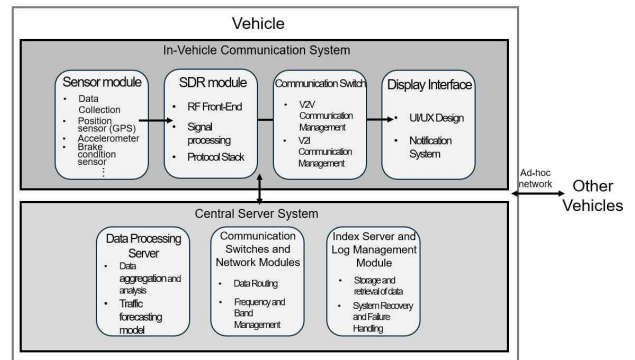


그림 1. 차량 간 무선통신 서비스 아키텍처  
Fig. 1. Inter-Vehicle Wireless Communication Service Architecture

## III. 소프트웨어 구성

WEB SDR 기반 차량 간 무선통신 디스플레이 서비스의 소프트웨어 구성은 신호 처리 모듈, 프로토콜 스택, 그리고 사용자 인터페이스 (UI)로 이루어져 있다.

신호 처리 모듈은 SDR 모듈을 통해 수신된 디지털 신호를 해석하고 필요한 데이터를 추출하는 역할을 수행한다. 이 모듈은 주파수 변환, 신호 변조와 복조, 오류 검출 및 수정 기능을 포함하여 통신 신뢰성을 보장한다. 변조와 복조 과정을 통해 신호

를 전송 주파수에 맞추어 변환하며, 오류 검출 및 교정 알고리즘을 통해 신뢰성 높은 데이터 전송을 가능하게 한다. 이를 통해 밀집된 교통 환경에서도 정확한 데이터 교환이 이루어질 수 있다.

프로토콜 스택은 차량 간 통신을 위한 표준화된 규칙을 제공하여 데이터 송수신의 일관성을 유지한다. 본 시스템은 V2V 및 V2I 통신을 모두 지원하며, 프로토콜 계층 구조를 통해 다양한 통신 환경에서 안정적인 데이터 흐름을 보장한다. 물리 계층에서는 SDR 모듈이 주파수와 전력 제어를 수행하고, 네트워크 계층에서는 라우팅 알고리즘을 통해 최적의 경로로 데이터를 전송하며, 상위 계층에서는 패킷화 및 에러 체크를 통해 데이터의 일관성과 신뢰성을 확보한다[4].

UI는 운전자에게 실시간 정보를 제공하는 중요한 구성 요소로, 차량 내 디스플레이를 통해 교통 상황, 경고 메시지, 경로 안내 등을 시각적·음성적으로 전달한다. UI는 HUD (Head-Up Display)와 대시보드 디스플레이를 통해 운전 중 시야 분산을 최소화하며, 주요 경고 사항은 음성 안내와 결합되어 운전자가 신속하게 대응할 수 있도록 설계되었다. 직관적인 디자인으로 정보 접근성을 높이며, 운전자는 개인 선호에 맞춰 인터페이스를 설정할 수 있어 다양한 상황에서 유연하게 정보를 수용할 수 있다.

#### IV. 하드웨어 구성

하드웨어 구성에서 안테나는 차량 간 무선통신의 신뢰성과 성능을 높이는 데 중요한 역할을 한다. 안테나는 차량이 주행하는 다양한 환경에서 안정적으로 신호를 송수신 할 수 있게 한다. 차량 간 통신에서는 옴니 방향 안테나 (Omni-directional Antenna)와 지향성 안테나 (Directional Antenna)가 사용되며, 환경에 따라 상호 보완적으로 활용된다.

옴니 방향 안테나는 모든 방향으로 신호를 송수신할 수 있는 특징을 가지고 있어, 도심과 같이 차량이 밀집된 구역에서 주변 모든 차량과 통신하기에 적합하다. 이 안테나는 넓은 커버리지를 제공하여 다방향에서 오는 신호를 수신할 수 있으며, 다양한 주행 환경에서 안정적인 통신이 가능하다.

지향성 안테나는 특정 방향으로 신호를 집중 송수신하여 통신 거리와 신호 강도를 극대화한다. 고속도로와 같은 환경에서 원거리 통신이 필요할 때 주로 사용되며, 신호를 특정 방향으로 집중시켜 간섭을 줄이고 통신 효율을 높인다. 이를 통해 주행 속도가 높은 구간에서도 안정적인 데이터 전송이 가능하다.

SDR 모듈은 안테나에서 수신한 아날로그 신호를 디지털 신호로 변환하고, 차량 간 통신이 가능한 형태로 데이터를 처리하는 역할을 담당한다. SDR 모듈은 소프트웨어로 주파수와 통신 프로토콜을 유연하게 조정할 수 있어 다양한 통신 환경에 대응할 수 있다. 이 모듈은 RF 프론트엔드를 통해 주파수를 변환하고, 데이터 전송에 적합한 신호로 변환하여 다른 차량 또는 중앙 서버와 원활한 데이터 교환을 가능하게 한다.

#### V. SDR 기술을 활용한 차량 간 무선통신 시스템 프로토타입 시현

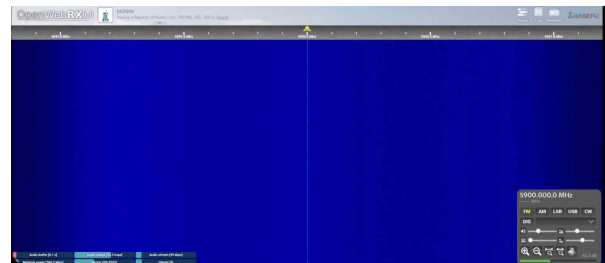


그림 2. 프로토타입 시현 디스플레이  
Fig. 2. Prototype Display

본 연구에서는 WEB SDR 기반 차량 간 무선통신 디스플레이 서비스의 프로토타입을 구축하여 시스템의 기능을 검증하였다. 프로토타입 구현에서는 현실적인 제약으로 인해 서버와 CPU로 라즈베리파이를 사용하였으며[5], 이는 프로토타입의 간단한 데이터 처리와 신호의 시각화를 위함이다. 라즈베리파이는 소형 서버 역할을 수행하여 신호 수신과 기본적인 데이터 흐름을 관리하며, 주파수 변동을 시각화하는 워터폴 (waterfall) 디스플레이 기능을 가능하게 한다.

본 시스템의 SDR 기반 프로토타입은 차량 간 통

신에서 중요한 5.9 GHz 대역을 포함한 다양한 주파수 신호를 수신하고, 신호의 수신 강도와 주파수 스펙트럼을 안드로이드 환경에서 워터폴 디스플레이 형태로 제공하도록 설계되었다. 이를 통해 운전자는 안드로이드 애플리케이션을 통해 신호 상태와 변화를 시각적으로 확인할 수 있으며, SDR 모듈을 통해 수신된 V2X 통신 주파수 대역의 신호 패턴이 실시간으로 디스플레이에 표시된다. 프로토타입에서 제한된 기능을 구현했으나, 이를 통해 SDR 기술을 사용한 신호 모니터링의 가능성을 보여주었으며, 라즈베리파이를 서버로 활용하여 시스템 아키텍처를 검증하였다.

## VI. 결 론

본 연구에서는 WEB SDR 기반 차량 간 무선통신 디스플레이 서비스의 프로토타입을 구축하여 SDR 기술을 활용한 차량 간 통신의 가능성과 실용성을 검토하였다. 제한된 환경에서 라즈베리파이를 서버와 CPU로 사용하여 기본적인 데이터 처리와 주파수 시각화를 구현했으며, 이를 통해 V2X 통신 주파수인 5.9 GHz 대역을 포함한 다양한 주파수 신호를 실시간으로 모니터링할 수 있음을 확인하였다.

이번 연구에서 구축된 프로토타입은 제한적인 데이터 처리 능력과 기능을 가졌으나, 라즈베리파이를 활용하여 SDR 기술을 적용할 수 있는 시스템 아키텍처를 성공적으로 검증하였다. 이를 바탕으로 향후 연구에서는 더 고성능의 서버와 프로세서를 도입하여 교통 정보의 실시간 제공, 경고 메시지 전송 등 스마트 교통 시스템으로서의 기능을 추가하고자 한다. SDR 기술을 통한 실시간 모니터링과 주파수 가변성의 유연성은 차량 간 통신의 신뢰성과 효율성을 증대시킬 수 있으며, 본 연구의 결과는 향후 차량 네트워크 환경을 확장하는 기초 자료로 활용될 수 있을 것이다.

따라서 본 연구의 결과는 V2X 통신에서 SDR 기술을 활용해 차량 간 실시간 통신의 가능성을 확인한 중요한 사례로, 스마트 교통 시스템의 발전을 위한 실질적인 적용 가능성을 제시하였다.

## 참 고 문 헌

- [1] Carl W. Rossler, Emre Ertin and Randolph L. Mosesa, "A software defined radar system for joint communication and sensing", IEEE, pp. 1050 ~ 1055, 2011.
- [2] H. Harada, "A small-size software defined cognitive radio prototype," 2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1-5, 2008.
- [3] J. Liu, J. Wan, B. Zeng, Q. Wang, H. Song and M. Qiu, "A Scalable and Quick-Response Software Defined Vehicular Network Assisted by Mobile Edge Computing", IEEE Communications Magazine, vol. 55, no. 7, pp. 94-100, July 2017,
- [4] I. Vitas, D. Šimunić and P. Knežević, "Evaluation of Software Defined Radio systems for smart home environments," Electronics and Microelectronics (MIPRO), pp. 562-565, 2015
- [5] R. Helbet, V. Monda, A. C. Bechet and P. Bechet, "Low Cost System for Terrestrial Trunked Radio Signals Monitoring Based on Software Defined Radio Technology and Raspberry Pi 4," 2020 International Conference and Exposition on Electrical And Power Engineering (EPE), pp. 438-441, 2020,

# NPU 기반 자세 추정을 통한 반려로봇의 능동적 눈 맞춤 기법

김지원\*, 전채욱\*\*, 옥승호\*\*\*

## Active Eye Contact Method for Companion Robots Using NPU-based Pose Estimation

Jiwon Kim\*, Chae-Wook Jeon\*\* and Seung-Ho Ok\*\*\*

### 요 약

최근 반려로봇의 수요 증가와 함께 사용자와 로봇의 상호작용인 Human Robot Interaction (HRI)의 중요성이 강조되고 있다. 이 중 사용자와 로봇의 비언어적 상호작용인 눈 맞춤 (eye contact)에 관한 연구가 진행되고 있으나, 로봇의 카메라 각도만 조정하는 방법으로는 사용자의 시선 방향을 파악하여 눈 맞춤을 시도하는 데 한계가 있다. 본 논문에서는 Raspberry Pi 5와 Hailo-8L NPU를 활용하여 자세 추정 (Pose estimation)모델을 기반으로 로봇이 사용자의 눈 위치를 실시간으로 파악하고, 능동적으로 눈 맞춤이 가능한 위치로 이동하여 자연스러운 눈 맞춤을 구현하는 방법을 제안한다.

### Abstract

With the recent increase in demand for companion robots, the importance of Human Robot Interaction (HRI) is being emphasized. While research on eye contact, a form of non-verbal interaction between users and robots, is ongoing, methods that only adjust the robot's camera angle have limitations in detecting the user's gaze direction and attempting eye contact. In this paper, we propose a method that implements natural eye contact interaction by utilizing Raspberry Pi 5 and Hailo-8L NPU to detect the user's eye position in real-time based on pose estimation model, and actively moving the robot to a position where eye contact is possible.

### Key words

HRI, pose estimation, companion robot, eye contact, position adjustment

## 1. 서 론

사회 고령화 및 1인 가구 증가로 정서적 지원에 대한 수요가 증가하면서 사용자와 정서적으로 교감

하는 반려로봇의 시장이 확대되고 있다[1-3]. 로봇과 사용자 간의 비접촉 상호작용에서 눈 맞춤은 정서적 지원을 위한 핵심적인 비언어적 상호작용 요소로 작용한다[4,5]. 그러나 기존 로봇들은 고정된 위

\* 동의대학교 로봇자동화공학과, 20222155@g.deu.ac.kr

\*\* 동의대학교 인공지능학과, 20246146@office.deu.ac.kr

\*\*\* 동의대학교 로봇자동화공학과, osh@deu.ac.kr(교신저자)

※ 이 논문은 교육부의 지방대학활성화사업 사업비를 지원받아 작성되었음

※ 본 논문(저서)는 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임

치에서 카메라만을 움직여 사용자와 눈 맞춤을 시도하기 때문에 제한된 동작 범위로 인해 자연스러운 상호작용에는 한계가 있다[6]. 따라서 본 논문에서는 자세 추정 모델을 활용하여 로봇이 사용자의 시선과 위치를 실시간으로 파악하고, 눈 맞춤이 가능한 위치에 능동적으로 이동하는 방법을 제안한다.

## II. 본론

로봇과 사용자의 눈 맞춤을 위해서는 로봇과 사용자가 일직선상에 최소 2m에서 최대 3m 사이의 거리를 두고 위치하여야 한다. 이를 위해 자세 추정에서 양쪽 눈의 키포인트가 인식되어 화면 중앙의 가상의 영역에 위치하고, 사용자를 인식한 바운딩 박스의 면적이 화면 대비 15%일 때 반러로봇이 사용자와 눈 맞춤을 성공한 것으로 정의하였다. 본 논문에서는 로봇이 사용자와 눈 맞춤하기 위한 범위를 그림 1과 같이 설정하였다.

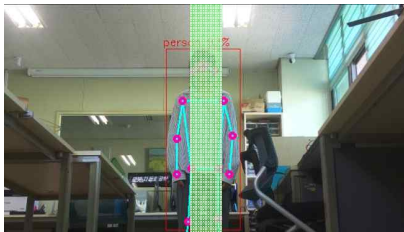


그림 1. 사용자와 눈 맞춤을 위한 설정 범위  
Fig. 1. Eye contact range with user

실험 환경 구성을 위해 NVIDIA Jetbot 프레임을 사용하였으며, 로봇의 하드웨어 구성은 표 1과 같다. 카메라는 로봇의 전방 3m에 있는 사용자와 상호작용하는 것을 목표로 하여 그림 2와 같이 지면으로부터 20cm의 위치에 70도 각도로 설치하였다.

표 1. 로봇 하드웨어 구성  
Table 1. Robot hardware configuration

구성 요소	시스템 구성
Main processor	Raspberry Pi 5
NPU	Hailo 8L
Camera	Raspberry Pi Camera v2

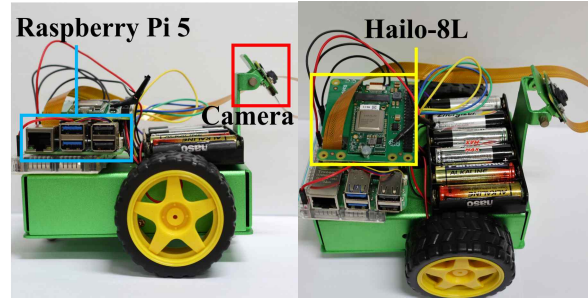


그림 2. 모바일 로봇  
Fig. 2. Mobile robot

사용자의 눈 위치 검출을 위해 YOLOv8-pose 모델을 사용하였다. YOLOv8-pose의 경우, 사용자의 눈을 포함한 17개의 키포인트 위치를 추정한다.

시스템의 동작 흐름도는 그림 3과 같다. 로봇이 사용자와 눈 맞춤을 시도할 때 카메라에 인식된 사용자의 눈 위치와 바운딩 박스를 기준으로 로봇의 위치를 조정한다. 그림 3의 시스템 흐름도에서의 각 오류 유형별 예시는 그림 4-6과 같다.

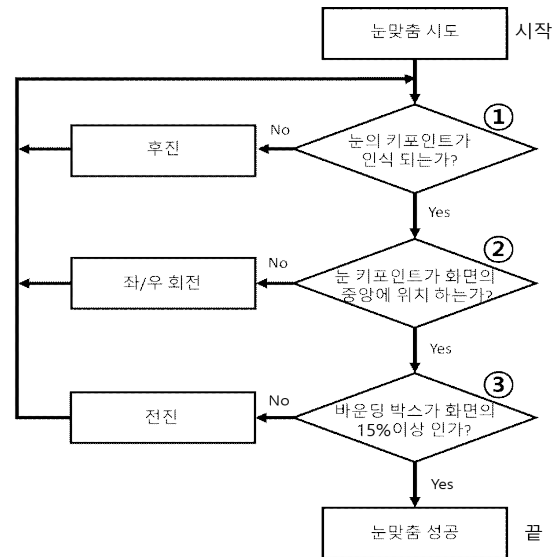


그림 3. 눈 맞춤 기법 흐름도  
Fig. 3. Eye contact method flowchart

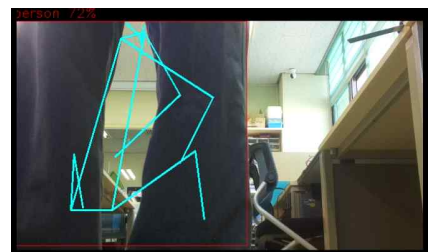


그림 4. 눈이 인식되지 않는 경우  
Fig. 4. Case of eye detection failure

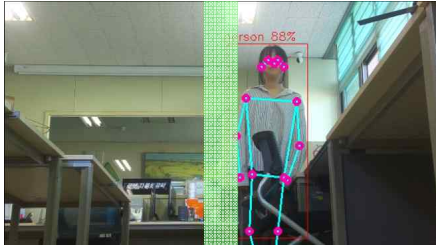


그림 5. 눈이 중앙에 위치하지 않는 경우  
Fig. 5. Case of off-center eye position



그림 6. 바운딩 박스의 크기가 화면의 30% 미만일 경우  
Fig. 6. Case of bounding box size less than 30% of screen

그림 3의 ①의 경우 우선 두 눈의 랜드마크 검출을 통해 로봇과 사용자가 최소거리 내에 있는지 판단한다. 그림 4와 같이 두 눈이 인식되지 않았을 경우 사용자가 로봇과 너무 가까이에 있는 것으로 판단해 로봇이 눈을 인식할 수 있을 때까지 후진한다.

그림 3의 ②에서는 사용자가 로봇과 일직선상에 있는지 판단한다. 이때 그림 5와 같이 인식된 양 눈의 랜드마크가 중앙 영역에 위치한 것이 아니라 한 방향으로 치우쳐 있을 시, 중앙 영역에 위치할 때까지 로봇을 치우친 방향으로 제자리 회전한다.

그림 3의 ③에서는 사용자가 로봇과 눈 맞춤이 가능한 최대거리에서 벗어났는지 판단한다. 그림 6과 같이 바운딩 박스의 면적이 화면 면적의 15% 미만일 경우, 로봇과 사용자의 거리가 멀다고 판단하여 바운딩 박스가 화면의 15%가 될 때까지 로봇을 전진한다.

### III. 결 론

본 논문에서는 자세 추정 모델을 활용하여 로봇이 자율적으로 위치를 조정해 사용자와의 눈 맞춤을 가능하게 하는 방법을 제시하였다. 실험을 통해 사용자가 눈 맞춤 하기 어려운 위치에 있을 때도

로봇이 능동적으로 위치를 보정하여 효과적인 상호작용이 가능하도록 이동함을 확인하였다. 향후 연구에서는 다양한 상황에서 발생할 수 있는 오류를 보정할 수 있는 시스템으로 확장할 예정이다.

### 참 고 문 헌

- [1] C. L. Kok et al, "A Novel Implementation of a Social Robot for Sustainable Human Engagement in Homecare Services for Ageing Populations", *Sensors*, 2024.
- [2] S. Toyoguch et al, "A Human-Centered and Adaptive Robotic System Using Deep Learning and Adaptive Predictive Controllers", *Journal of Robotics and Mechatronics* Vol.35 No.3, pp. 834-843, 2023.
- [3] X. Wang and Y. Wang, "Design Proposal to the Future Companion Intelligent Products", *Ergonomics In Design*, Vol. 77, pp.555-562, 2023
- [4] J. P. Kolengadan et al, "Development of Interactive Visual Recognition Assistant Bot", *ICECONF*, pp. 1-8, 2023.
- [5] J. H. Lee, "Investigating User Classification and Nonverbal Cues for Companion Robot Design", *2022 KSBDA Fall International Conference*, 2022.
- [6] D. Semar and T. Chen, "Automating SmartPhone Automation", *2015 24th Australasian Software Engineering Conference*, pp. 108-117, 2015.



# 객체 인식을 이용한 FPGA 기반 얼굴 추적 시스템

이용문\*, 최민서\*, 박준형\*, 이용환\*\*

## FPGA-based Face Tracking System using Object Recognition

Lee Yongmun\*, Choi Minseo\*, Park Junhyung\*, Lee Yonghwan\*\*

### 요 약

기존에 존재하는 버튼식 선풍기는 일반적으로 가장 많이 보편화되어 있지만, 사용자의 편의성 측면에서는 조금 한계가 있다. 선풍기의 설정을 조정해야할 때 수동으로 해야 하는 불편함에 직면하는데 빈번한 조정이 필요한 환경에서는 사용자의 활동을 방해하는 데까지 이어질 수 있다. 이를 보완하기 위해, PYNQ-Z1 보드를 활용한 얼굴 추적 시스템을 적용하니 사용자의 편의성 측면에서 크게 향상되었다. 또한, 시스템을 소프트웨어적으로만 구현하는 게 아닌 하드웨어적으로도 같이 구현하여 처리 속도를 향상시킨다. 이 작품을 통해, 객체 인식 기술을 전자제품에 적용하면 사용자의 편의성을 극대화할 수 있다는 가능성을 보여준다. 이러한 기술이 보편화됨에 따라 앞으로 TV, 냉장고, 세탁기 등 무수한 가전제품에 적용할 수 있을 것이다. 소형 가전도 프리미엄으로 구매하는 요즘 소비 시장에서 차별화된 경쟁력을 보여줄 것이다.

### Abstract

Existing button-type fans are generally the most common, but there are some limitations in terms of user convenience. When adjusting the settings of the fan, it faces inconvenience of having to do it manually, and in an environment that requires frequent adjustments, it can even lead to interruption of the user's activities. To compensate for this, applying the face tracking system using the PYNQ-Z1 board greatly improved in terms of user convenience. In addition, the system is not only implemented in software, but also in hardware to improve processing speed. Through this work, it shows the possibility that applying object recognition technology to electronic products can maximize user convenience. As these technologies become more common, they will be able to be applied to countless home appliances such as TVs, refrigerators, and washing machines in the future. It will show differentiated competitiveness in the consumer market these days, when small home appliances are also purchased at a premium.

### Key words

face tracking system, software, hardware, processing speed, differentiated competitiveness

---

\* 국립금오공과대학교 반도체시스템전공, 학부생, cocoyi00@naver.com, mincer1121@naver.com, gp6409@naver.com

\*\* 국립금오공과대학교 전자공학부, 교신저자, yhlee@kumoh.ac.kr

※ 이 연구는 정부(교육부-산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0022161, 반도체전공트랙사업)

## 1. 서론

기존 선풍기는 버튼을 누르거나 위치를 옮김으로써 방향을 조정하는 방식이었기 때문에 사용자가 직접 움직여야 하는 불편함이 있었다. 이러한 한계를 극복하기 위해 본 프로젝트에서는 얼굴 인식 기술을 활용한 새로운 선풍기 제어 시스템을 제안한다.

제안하는 시스템은 Webcam을 통해 사용자의 얼굴을 인식하고, 이를 기반으로 선풍기의 방향을 자동으로 제어한다. 얼굴 인식 알고리즘을 통해 사용자의 얼굴을 정확히 인식하고, 인식된 정보를 활용하여 사용자의 상황에 맞는 선풍기 동작을 실현한다. 이를 통해 버튼 조작의 불편함을 해소하고 사용자의 편의성을 향상시킬 수 있다.

다양한 얼굴 인식 알고리즘 중에서 이번 프로젝트에서는 LBP(Local Binary Pattern) 알고리즘을 활용하였다. LBP 알고리즘은 계산 효율성이 높고 조명 변화에 강인한 장점이 있어 실시간 얼굴 인식에 적합하다.

## II. 실시간 얼굴 추적 시스템의 구현

제안된 시스템은 PYNQ-Z1 보드, Webcam, HDMI 모니터 등의 주요 하드웨어 구성 요소로 이루어져 있다. PYNQ-Z1 보드는 Xilinx Zynq-7000 AP SoC를 기반으로 하는 임베디드 시스템 보드로, 시스템의 중추 역할을 한다. 이 보드에서 얼굴 인식 알고리즘이 실행되고 선풍기 제어를 위한 신호가 생성된다.

Webcam은 사용자의 얼굴을 촬영하는 역할을 하며, 촬영된 영상은 PYNQ-Z1 보드로 전송되어 처리된다. 얼굴 인식 결과와 Webcam 영상은 HDMI 모니터를 통해 시각화되어 사용자가 시스템의 동작 상황을 확인할 수 있다.

이렇게 PYNQ-Z1 보드, Webcam, HDMI 모니터가 유기적으로 연동되어 얼굴 인식 기반의 선풍기 제어 시스템을 구축하게 된다. 각 구성 요소는 고유한 역할을 수행하면서 전체적으로 통합된 시스템을 이루고 있다.

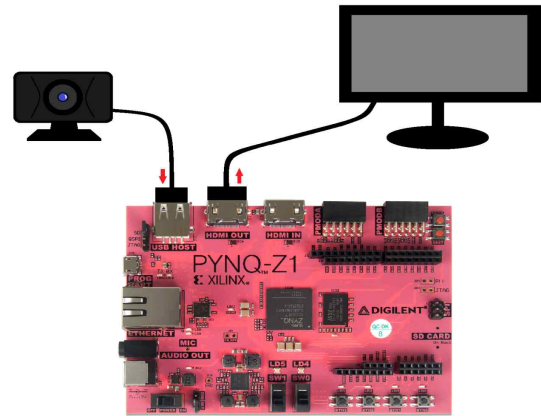


그림 1. 전체 구상도

Fig. 1. Overall Schematic Drawing

이 시스템의 하드웨어와 소프트웨어는 유기적으로 통합되어 작동한다. 하드웨어 측면에서 PYNQ-Z1 보드는 Xilinx Vivado 프로그램을 사용하여 프로그래밍되었으며, 보드 상에서 실행되는 소프트웨어는 Python 언어를 사용하여 Jupyter Notebook 프로그램으로 작성되었다. Webcam은 USB 인터페이스를 통해 PYNQ-Z1 보드와 연결되어 있으며, OpenCV 라이브러리를 활용하여 영상 데이터를 처리한다.

소프트웨어 측면에서 LBP 알고리즘 기반의 얼굴 인식 기능은 OpenCV 등의 파이썬 라이브러리를 사용하여 구현되었다. 이 알고리즘은 입력 영상에서 얼굴 영역을 검출하고 검출된 얼굴 이미지를 분석하여 특징 벡터를 추출한다. 추출된 특징 벡터는 사전에 학습된 얼굴 인식 모델과 비교되어 최종적으로 사용자의 얼굴을 인식하게 된다.

전체 시스템의 동작 흐름은 다음과 같다. Webcam으로부터 입력된 영상 데이터는 PYNQ-Z1 보드로 전송되어 OpenCV 라이브러리에 의해 처리된다. 처리된 영상에서 LBP 알고리즘이 얼굴 영역을 검출하고 인식 과정을 수행한다. 인식된 얼굴 정보는 PYNQ-Z1 보드의 GPIO 핀을 통해 선풍기 제어를 위한 신호로 변환되어 출력된다. 동시에 얼굴 인식 결과와 Webcam 영상은 HDMI 모니터에 시각화되어 사용자에게 제공된다.

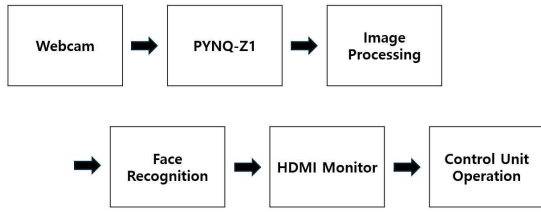


그림 2. 시스템 처리 과정  
Fig. 2. System Processing Process

이 시스템에서 사용된 LBP 알고리즘은 영상 내 특정 픽셀 주변의 화소 값 패턴을 이용하여 텍스처 특징을 추출하는 방법이다. 이 알고리즘은 조명 변화에 강인하고 계산 비용이 낮아 실시간 얼굴 인식에 적합하다.

제안된 시스템의 성능을 평가한 결과, 3~4미터 이내에서 얼굴 인식 정확도는 약 85%로 나타났습니다. 이는 LBP 알고리즘의 조명 변화 강인성과 계산 효율성 덕분입니다. 또한 얼굴 인식 후 선풍기 제어까지의 반응 시간은 평균 1초 이내로 측정되어 실시간 제어가 가능한 수준임이 확인되었다.

기존의 버튼식 선풍기와 비교했을 때, 제안된 시스템은 버튼 조작의 불편함을 해소하고 사용자 중심의 편리한 제어를 제공한다는 장점이 있다. 또한 선풍기의 작동이 사용자의 상황에 맞춰져 에너지 효율성도 높아질 것으로 기대된다. 반면 시스템의 단점으로는 Webcam을 통한 사용자 프라이버시 침해 우려, 실내 조명 변화에 따른 인식 오류 가능성 등이 있다. 향후 프라이버시 보호 기술과 조명 변화 보정 알고리즘을 추가하여 이러한 문제를 해결할 필요가 있다.

### III. 결 론

본 연구에서는 LBP 알고리즘 기반의 얼굴 인식 기술을 활용하여 사용자 편의성이 높은 새로운 선풍기 제어 시스템을 제안하였다. 기존의 버튼식 조작 방식과 달리 제안된 시스템은 사용자의 위치를 자동으로 인식하여 선풍기의 방향을 능동적으로 제어합니다. 이를 통해 사용자의 편의성을 향상시키고 에너지 효율성도 높일 수 있을 것으로 기대된다.

이 프로젝트는 시작일 뿐이다. 객체 인식 기술을 가전제품에 적용하면 사용자에게 높은 편의성을 제

공할 수 있다는 가능성을 보여주는 프로젝트이다. 소형 가전제품도 편의성과 만족도가 높다면 프리미엄으로 구매하는 요즘 소비 시장에서 차별화된 경쟁력을 보여줄 것이다.

### 참 고 문 헌

- [1] opencv: <https://opencv.org/>
- [2] PYNQ documentation: <http://pynq.readthedocs.io/>
- [3] SG - 90 Servo Motor Datasheet

# 하드웨어 가속기를 이용한 객체 감지 도로 반사경

이성준\*, 안인선\*, 김찬혁\*, 이용환\*\*

## Object Detector for Road Condition using Hardware Accelerator

Sung Jun Lee\*, In Seon An\*, Chan Hyeok Kim\*, and Yong Hwan Lee\*\*

### 요약

일반 도로 반사경은 사용자가 직접 거울을 통해 차량이나 보행자의 접근 여부를 확인해야 하며, 어두운 밤이나 먼지, 훼손된 상태에서는 시각적 확인이 어려운 단점이 있다. 본 연구에서는 이러한 문제를 해결하기 위해 도로 반사경을 하드웨어 가속을 시킨 객체 감지 시스템으로 발전시키는 방안을 제안한다. 가속을 시켰기에 더욱 빠르게 객체를 감지할 수 있고 해당 시스템에 LED matrix를 연결하여 어두운 환경에서도 밝고 선명한 이미지를 제공할 수 있어, 운전자와 보행자가 보다 안전하게 상황을 인식할 수 있도록 한다.

### Abstract

The general road reflector mirror has the disadvantage that the user must directly check whether a vehicle or pedestrian has approached through a mirror, and it is difficult to visually check in dark night, dust, or damaged state. In order to solve this problem, this study proposes a method of developing a road reflector mirror into a hardware accelerated object detection system. Because it is accelerated, it can detect objects more quickly and provide bright and clear images even in dark environments by connecting LED matrix to the system, allowing drivers and pedestrians to recognize the situation more safely.

### Key words

object detection, hardware accelerator, real-time detection, computer vision

## 1. 서론

도로 반사경은 도로의 안전성을 높이기 위해 설치되는 장치로 주로 교차로, 골목길, 곡선 도로 등에서 사용된다. 이 반사경은 차량이나 보행자가 오는 방향을 확인할 수 있도록 돕는다. 하지만 기존의 반사경은 먼지나 손상으로 인해 반사 능력이 저하될 수 있어 정기적인 청소와 점검이 필요하고 어두

운 밤에는 반사경만으로는 시각적인 인식이 어려울 수 있다. 이러한 단점을 보완하기 위해 이 논문에서는 라즈베리파이 4로 TensorFlow Lite를 활용하여 사람이나 자동차가 인식되면 LED matrix로 알려주는 시스템을 만들려고 하였다. 우리는 여기에 더해 하드웨어 가속기를 사용하여 더욱 빠르게 객체 감지를 하여 실시간 시스템에 더욱 적합하게 만들 것이다.

\* 국립금오공과대학교 반도체시스템전공, sotoria1123@naver.com, inseon0117@naver.com, cksgur0417@naver.com

\*\* 국립금오공과대학교 반도체시스템전공, E-mail: yhlee@kumoh.ac.kr

※ 이 연구는 정부(교육부-산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0022161, 반도체전공트랙사업)

## II. SSD-MobileNet 객체감지와 하드웨어 가속기

SSD는 Single Shot Multibox Detector의 약자로 backbone에서 추출된 feature map들을 3x3 사이즈의 커널을 사용한 Convolutional Neural Network(CNN)로 연산해 해당 feature map에 있는 객체 정보를 얻는다. 다양한 feature 사이즈에서 객체 탐지를 수행하기 위함이다. 이런 방식으로 객체를 추출한 덕분에 SSD는 상대적으로 낮은 입력 이미지 해상도에서도 좋은 성능과 빠른 속도를 보여줄 수 있었다고 한다. YOLO(You Only Look Once)는 448x448 해상도의 이미지를 사용하지만 SSD는 VGG-16 기반이라 224x224 해상도의 이미지를 사용하기 때문에 작은 이미지 해상도를 가지고도 성능이 더 잘 나온다[1].

MobileNet이란 스마트폰 및 기타 모바일 장치와 같이 리소스가 제한된 환경에서 효율적인 계산을 위해 설계된 경량 심층신경망이다. MobileNet은 Depthwise separable convolution을 활용하여 모델을 경량화했다. Depthwise separable convolution은 필터의 채널과 공간 차원을 분리할 수 있다는 생각에서 비롯됐다. 일반적인 convolution에서는 채널 방향과 공간 방향의 convolution을 동시에 수행하는 반면 Depthwise separable convolution은 공간 방향의 depthwise convolution과 채널 방향의 pointwise convolution을 따로 수행하여 합치는 방식을 사용한다[2].

TensorFlow Lite, SSD, MobileNet은 모두 저전력

및 저성능 환경에서도 효과적으로 사용할 수 있다. 따라서 본 논문에서는 위 방식들을 사용하여 라즈베리파이에서 실시간 객체 감지를 수행하도록 만들기 하였다. 거기에 더해 Google Coral USB Accelerator를 이용하여 하드웨어 가속을 시켜 성능을 더 향상시켰다. 해당 제품은 Edge TPU 칩을 내장하고 있어 작은 크기의 컴퓨터에서도 높은 성능의 AI 작업을 실행할 수 있다. 이러한 하드웨어 가속기의 원리는 병렬처리를 통해 가속을 시키는 것이다. CPU는 여러 작업을 순차적으로 처리하는 데 적합하지만 GPU나 TPU는 수천개의 코어로 대량의 연산을 동시에 처리할 수 있다[3]. 특히 TPU는 Tensor Processing Unit의 약자로 구글이 인공지능(AI) 및 머신러닝(ML) 작업을 가속화하기 위해 개발한 특수 목적의 하드웨어 프로세서이다. TPU는 딥러닝에서 자주 사용되는 대규모 행렬 곱셈을 빠르게 처리할 수 있게 해주기 때문에 연산 시간을 크게 단축할 수 있다[4].

## III. 라즈베리파이4를 이용한 시스템 제작

본 논문에서mobilenet는 그림 2, 그림 3의 라즈베리 파이 4와 Google Coral USB Accelerator를 이용하여 제작을 하였다. 하드웨어 가속을 시킨 것과 시키지 않은 것의 속도 차이를 비교해 보았다. 가속시키지 않았을 때는 FPS가 4 정도 나왔고 가속을 시킨 경우 FPS가 18 정도 나왔다. 이를 통해 가속을 시킨 것이 약 4배 정도 빠르다는 것을 알 수 있다.

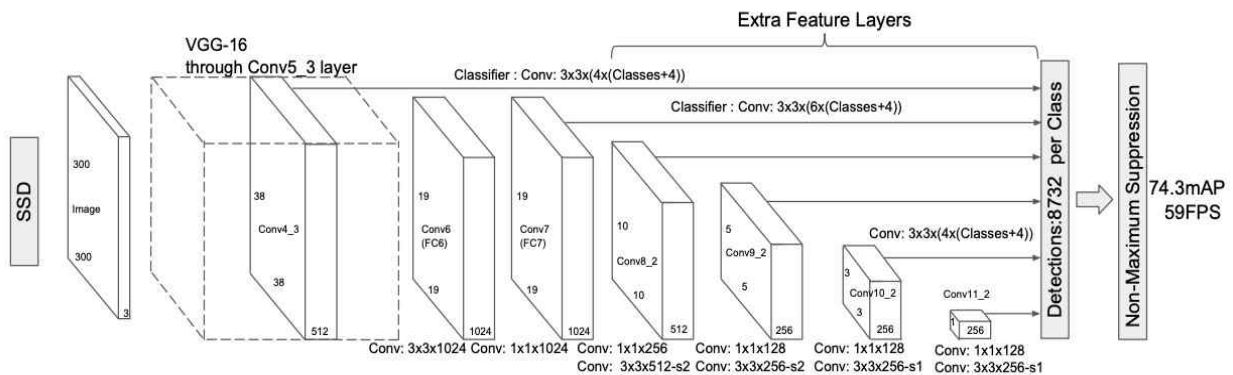


그림 1. SSD 구조  
Fig. 1. SSD structure

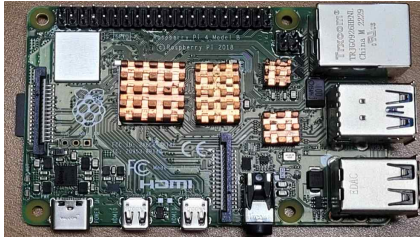


그림 2. 라즈베리파이 4  
Fig. 2. raspberrypi 4



그림 3. Google Coral USB Accelerator  
Fig. 3. Google Coral USB Accelerator

표 1. 가속 전 후 FPS  
Table 1. FPS before and after acceleration

가속 여부	FPS
가속 전	4
가속 후	18



그림 4. 객체 감지 도로 반사경  
Fig. 4. object detection road reflector mirror

#### IV. 결 론

본 논문에서는 Google Coral USB Accelerator를 이용한 객체 인식 가속 방법을 제안하였다. 제안된 방법은 Edge TPU를 활용하여 실시간으로 객체 인식의 소프트웨어 기반 방법에 비해 향상된 성능을 보여준다. 이를 바탕으로 객체인식 도로 반사경은 실사용 환경에서도 충분히 적용 가능성을 확인하였다.

따라서 도로 반사경과 같은 실시간 객체 인식이

중요한 상황에서 제안된 시스템은 신속하고 정확한 반응을 제공하여 교통안전성을 향상시킬 수 있을 것으로 기대된다.

#### 참 고 문 헌

- [1] Liu Wei, Anguelov Dragomir, Erhan Dumitru, Szegedy Christian, Reed Scott, Fu Cheng-Yang, Berg Alexander C . "SSD: Single Shot MultiBox Detector" Lecture Notes in Computer Science : 21-37.
- [2] Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." ArXiv, abs/1704.04861. <https://arxiv.org/abs/1704.04861>.
- [3] Manor Erez, Greenberg Shlomo. "Using HW/SW Codesign for Deep Neural Network Hardware Accelerator Targeting Low-Resources Embedded Processors" IEEE ACCESS : 22274-22287.
- [4] Jouppi N , Young C , Patil N , Patterson D . "Motivation for and Evaluation of the First Tensor Processing Unit" IEEE Micro : 10-19.

# 표정 기반 감정 분석 콘텐츠 추천 서비스 설계

이세민\*, 손민섭\*, 박진현\*\*, 이창우\*\*\*, 정은미\*\*\*\*

## Designing Facial Expressions-Based Emotional Analysis Content Recommendation Service

Semin Lee\*, Minseop Son\*, JinHyun Park\*\*, Changwoo Lee\*\*\*, and Eunmi Jung\*\*\*\*

### 요 약

본 논문에서는 이미지 분석과 감정 인식 기술을 기반으로 한 콘텐츠 추천 서비스를 제안한다. 사용자로부터 실시간 웹캠 사진을 촬영하여 감정 인식 API를 통해 표정을 분석하고, 추출된 감정 데이터를 활용하여 음악, 영화, 책 등 다양한 콘텐츠를 추천한다. 백엔드는 Nest.js로 구축되어 감정 인식 요청 및 각각의 콘텐츠 API와의 연동을 관리하며, 프론트엔드는 Vite.js와 React.js로 구현되어 사용자 인터페이스와 실시간 촬영 기능을 제공한다. 본 시스템은 사용자 감정에 맞는 콘텐츠 추천을 통해 맞춤형 경험을 제공하며, 이를 통해 콘텐츠 소비 만족도를 높일 수 있을 것으로 기대된다.

### Abstract

In this paper, we propose a content recommendation service based on image analysis and emotion recognition technology. It takes a real-time webcam picture from a user and analyzes the expression through the emotion recognition API, and recommends various contents such as music, movies, and books using the extracted emotion data. The backend is built with Nest.js and manages emotion recognition requests and interworking with each content API, and the front end is implemented with Vite.js and React.js to provide a user interface and real-time shooting function. This system provides a customized experience through content recommendations tailored to the user's emotions, and through this, it is expected to increase content consumption satisfaction.

### Key words

Emotion Recognition, Content Recommendation, Image Analysis, Real-time Processing, User Personalization

## 1. 서 론

최근 디지털 콘텐츠의 폭발적인 증가로 인해 사

용자는 다양한 선택지 속에서 자신에게 맞는 콘텐츠를 탐색하는데 어려움을 겪고 있다[1]. 이에 따라 사용자 맞춤형 시스템의 필요성이 점진적으로 강조

\* 국립안동대학교 AI융합학과, dltpals7498@gmail.com, spcle2412@gmail.com

\*\* 국립안동대학교 멀티미디어공학과, pa990422@gmail.com

\*\*\* 국립안동대학교 교육공학과, toplcw2437@naver.com

\*\*\*\* 국립안동대학교 SW융합교육원 emjung@anu.ac.kr

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

되고 있으며, 특히 이미지 분석 및 감정 인식 기술을 활용한 개인화된 추천 시스템이 주목받고 있다 [2]. 감정 인식 기반의 추천 시스템은 사용자의 감정을 실시간으로 분석하여 심리 상태에 맞는 콘텐츠를 제공함으로써 기존의 일반적인 추천 알고리즘과는 차별화된 경험을 제공할 수 있다.

본 논문은 사용자 표정에서 감정을 추출하여 음악, 영화, 드라마, 책 등의 콘텐츠를 추천하는 서비스의 구현을 목표로 한다. 웹캠으로 촬영된 이미지를 통해 사용자 감정을 실시간으로 분석하고, 이를 다양한 콘텐츠 API(예: Spotify API)에 전달하여 감정에 맞는 콘텐츠를 자동으로 추천하는 구조를 갖추었다. 시스템의 백엔드는 Nest.js로 구축하여 감정 인식 요청 및 데이터 통신을 처리하고, 프론트엔드는 Vite.js와 React.js로 구성하여 웹캠 사진 촬영 및 사용자 인터페이스를 구현하였다. 본 시스템은 감정 기반의 콘텐츠 추천 시스템이 사용자의 개인화된 경험을 극대화하고 콘텐츠 소비 만족도를 높일 수 있는 가능성을 탐구한다. 이를 통해 기존 추천 시스템의 한계를 극복하고, 감정 인식 기술이 콘텐츠 추천에서 가지는 실용적 가치를 제시하고자 한다.

## II. 관련 연구

### 2.1 콘텐츠 기반 필터링 알고리즘

콘텐츠 기반 필터링은 추천 시스템에서 사용자의 이전 선호도를 기반으로 유사한 콘텐츠를 제공하는 알고리즘이다. 이 알고리즘은 추천할 항목(도서, 음악 등)의 특성(예: 장르, 주제, 태그)을 분석하여 사용자에게 적합한 항목을 추천한다. 보통 이 방식은 사용자의 선호도를 반영한 "프로필"을 구축하고, 사용자가 좋아하는 콘텐츠와 유사한 특성을 가진 항목을 찾아 추천 목록에 포함하는 구조이다. 콘텐츠의 특성들 간의 유사도를 계산할 때는 코사인 유사도, 피어슨 상관계수 등의 방법이 자주 활용된다.

그림 1의 콘텐츠 기반 필터링 시스템의 아키텍처는 네 가지 주요 단계로 구성된다. 첫 번째로, 데이터 수집 단계에서는 콘텐츠의 주요 특성인 장르, 주제, 키워드 등을 추출하여 메타데이터로 저장한다.

이 과정은 시스템이 콘텐츠의 특성을 학습하고 사용자에게 맞춤형 추천을 할 수 있도록 준비하는 단계이다. 두 번째 단계인 사용자 프로파일링에서는 사용자의 이전 활동과 선호도를 분석해 사용자 프로필을 생성한다. 이 프로필에는 사용자가 선호하는 콘텐츠 특성이 반영되며, 개인화된 추천을 위한 기준으로 작용한다. 세 번째 유사도 계산 단계에서는 사용자 프로필과 추천할 콘텐츠의 특성 간의 유사도를 계산하여, 사용자와 유사한 특성을 지닌 콘텐츠를 식별한다.

이때 코사인 유사도와 피어슨 상관계수 같은 다양한 유사도 측정 방법이 사용된다. 마지막으로, 추천 생성 단계에서는 유사도가 높은 콘텐츠를 추천 목록에 포함시켜 최종 추천 결과를 사용자에게 제공한다. 이러한 단계적 아키텍처를 통해 콘텐츠 기반 필터링 시스템은 사용자가 선호할 가능성이 높은 콘텐츠를 지속적으로 제공할 수 있다[3].

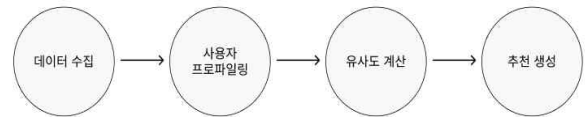


그림 1. 콘텐츠 기반 필터링 알고리즘  
Fig. 1. Content-based filtering algorithms

### 2.2 본 시스템에서의 활용

본 시스템의 초기 추천 알고리즘은 감정을 키워드와 단순 매칭하여 추천하는 방식이었으나, 감정과 무관한 콘텐츠가 추천될 수 있는 한계가 있었다. 이를 해결하기 위해 콘텐츠 기반 필터링을 도입하여, 감정과 관련된 문맥적 의미를 반영한 추천이 이루어지도록 개선했다. 예를 들어, "happy" 감정일 때는 단순히 "happiness"라는 키워드를 기반으로 추천하지 않고, 감정과 연관된 콘텐츠의 특성(긍정적인 주제 등)을 우선적으로 분석하여 관련성을 높였다.

이러한 개선은 BooksService, SpotifyService, TMDbService 등 여러 서비스에 적용되어, 감정과 콘텐츠 간의 연결성을 강화하고 더욱 개인화된 추천을 제공할 수 있게 한다.



### III. 본 론

본 시스템은 사용자 웹캠을 통해 실시간으로 촬영된 이미지를 분석하여 감정에 맞는 음악, 영화, 드라마, 책 등의 콘텐츠를 추천하는 서비스다.

이 시스템은 백엔드와 프론트엔드로 나뉘어 상호 협력하며, 감정 추출과 데이터 통신을 수행한다. 백엔드는 Nest.js로 구축되어 감정 분석 API와 연동된 엔드포인트를 통해 감정 분석 및 추천 데이터를 관리한다. EmotionController는 클라이언트로부터 감정 요청을 받아 EmotionService에서 Python 스크립트를 호출해 감정을 추출하며, 결과를 SpotifyService, TmdbService, BooksService 등과 연결하여 감정에 맞는 장르와 콘텐츠 데이터를 반환한다.

특히 SpotifyService는 감정을 특정 장르로 매핑하는 mapEmotionToGenre 함수를 통해 감정에 따라 음악을 필터링하여 추천한다.

그림 2는 웹캠으로 촬영한 이미지를 분석하여 사용자의 감정에 맞는 책, 드라마, 영화, 음악 등의 콘텐츠를 추천하는 시스템 흐름도이다.

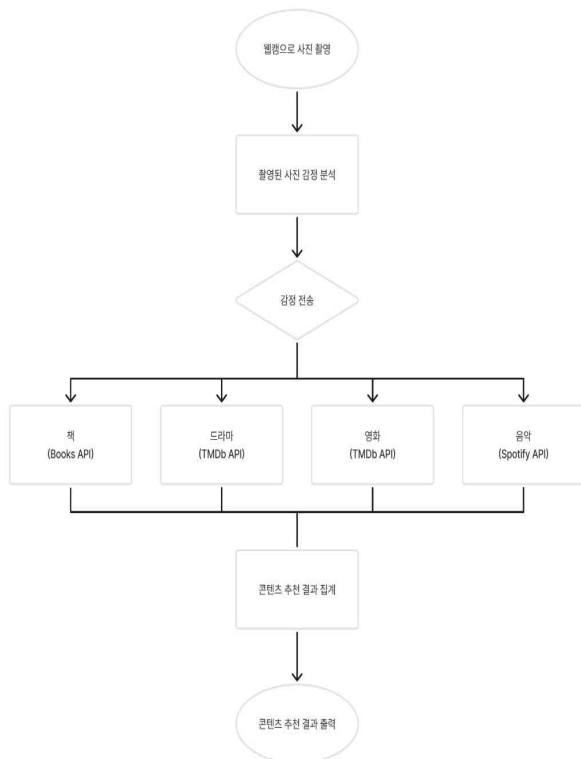


그림 2. 감정 기반 콘텐츠 추천 시스템 흐름도  
Fig. 2. Emotion-Based Content Recommendation System Flowchart

이 시스템은 실시간으로 웹캠을 통해 사용자의 이미지를 촬영하고, 촬영된 이미지는 감정 분석 API에 전달되어 감정이 추출된다. 감정 분석이 완료되면 추출된 감정데이터가 다음 단계로 전달되어, 다양한 콘텐츠 추천 API를 호출하게 된다.

콘텐츠 추천 API는 감정에 맞춰 각각의 분야에서 적합한 콘텐츠를 선택한다. 책 추천은 Books API를 통해, 드라마와 영화 추천은 TMDb API를 통해, 음악 추천은 Spotify API를 통해 이루어진다. 이렇게 추천된 각 분야의 콘텐츠는 사용자에게 적합하게 정제된 후, 최종적으로 추천 결과가 사용자에게 출력된다.

그림 3은 프론트엔드 구현 페이지이다. 프론트엔드는 Vite.js와 React.js를 사용하여 사용자와의 인터페이스와 상호작용을 구현하며, Webcam 컴포넌트를 통해 웹캠 사진 촬영 및 감정 분석 결과 전송을 담당한다. 사진 촬영 후, 이미지는 서버로 전송되어 감정이 분석되며, 분석된 감정 데이터를 Redux 스토어에 저장되어 Main 페이지에 표시된다. Main에는 웹캠 초기화와 페이지를 이동 지원하고, 결과 화면을 호출하는 Standard 컴포넌트에는 감정 결과를 기반으로 선택한 콘텐츠 API로부터 데이터를 불러와 사용자에게 맞춤형 추천 콘텐츠를 제공한다. 이 과정에서 데이터 통신 중 발생하는 오류는 Error Handling 기능을 통해 사용자에게 안내되어 안정적인 서비스를 지원한다.

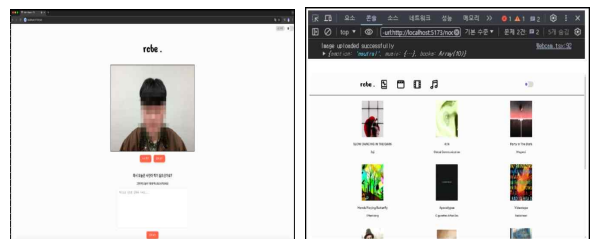


그림 3. 실시간 촬영 화면 및 콘텐츠 추천 페이지  
Fig. 3. Real-Time Capture and Content Recommendation Page

### IV. 결 론

본 논문에서는 사용자 감정을 고려한 콘텐츠 추천을 통해 기존의 정형화된 추천 방식과 차별화된

경험을 제공하며, 특히 감정 기반의 개인화된 콘텐츠 추천이 사용자의 만족도를 높이고 콘텐츠 소비를 보다 풍부하게 할 수 있음을 확인할 수 있었다.

향후 개발에는 감정 분석의 정확도를 높이기 위한 머신러닝 기반의 정교한 분석 모델을 도입하고, 더 다양한 콘텐츠 유형과의 연동을 통해 추천 시스템의 활용 범위를 확대하여 시스템을 확장할 수 있는 기대가 있다.

## 참 고 문 헌

- [1] 이수아, 이소미, 강수진. (2023). 사용자 간 추천을 활용한 OTT 서비스 내 콘텐츠 탐색과 선택 방안 제안. 디자인융복합연구, 22(6), 47-64. 10.31678/SDC103.3
- [2] 오의택, "나만을 위한 취향 맞춤형 사용자 경험", 디지털 인사이트, 2023. 06. 12.
- [3] 장령령. "콘텐츠 기반 필터링을 적용한 논문추천시스템에 대한 연구." 국내박사학위논문 전남대학교, 2016. 광주

# NPU 기반 임베디드 시스템을 위한 포즈 추정 모델 최적화

최경철\*, 전채욱\*\*, 옥승호\*\*\*

## Optimization of a Pose Estimation Model for NPU-Based Embedded Systems

Gyeong-Chul Choi\*, Chae-Wook Jeon\*\*, and Seung-Ho Ok\*\*\*

### 요약

본 논문은 YOLOv8n-pose 모델을 최적화하여 Khadas사의 VIM3 임베디드 보드에 적용하는 방법을 제안한다. 최근 임베디드 시스템에서는 전력 소모와 물리적 크기가 제한된 환경에서 효율적인 딥러닝 모델 연산을 가능하게 하는 AI 전용 가속기인 Neural Processing Unit (NPU)가 주목받고 있다. NPU에 딥러닝 모델을 실행하기 위해 모델의 Head 구조의 출력 부분을 수정하고, 양자화를 적용하여 모델을 최적화하였다. 실험 결과 uint8, int8, int16으로 양자화된 모델 중 추론 속도는 uint8 모델이 23.13ms로 가장 우수하였으며 평균 정밀도는 int16 모델이 0.763으로 가장 높았다. 이를 통해 VIM3에 포즈 추정 모델을 적용할 수 있음을 확인하였다.

### Abstract

This paper proposes a method of optimizing the YOLOv8n-pose model and applying it to Khadas' VIM3 embedded board. Recently, in embedded systems, the neural processing unit (NPU), an AI dedicated accelerator that enables efficient deep learning model operation in environments with limited power consumption and physical size, is attracting attention. In order to execute the deep learning model on the NPU, the output part of the head structure of the model was modified, and the model was optimized by applying quantization. As a result of the experiment, among the quantized models with uint8, int8, and int16, the inference speed was the best at 23.13ms in the uint8 model, and the average precision was the highest at 0.763 in the int16 model. Through this, it was confirmed that the pose estimation model can be applied to the VIM3.

### Key words

npu, embedded board, khadas vim3, yolov8n-pose, model optimization

---

\* 동의대학교 로봇자동화공학과, 20202128@office.deu.ac.kr

\*\* 동의대학교 인공지능학과, 20246146@office.deu.ac.kr

\*\*\* 동의대학교 로봇자동화공학과, osh@deu.ac.kr(교신저자)

※ 이 논문은 교육부의 지방대학활성화사업 사업비를 지원받아 작성되었음

※ 본 논문(저서)는 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임

## 1. 서 론

최근 임베디드 시스템과 같은 전력 소모와 물리적 크기가 제한된 환경에서는 신경망의 높은 성능을 유지하면서 전력 소비를 최소화하는 시스템이 필요하다[1,2]. 따라서 딥러닝 모델 연산을 효율적으로 처리할 수 있는 Neural Processing Unit (NPU)가 주목받고 있다[3,4]. 하지만 딥러닝 모델을 NPU에 적용하기 위해서는 모델의 최적화가 필요하다.

본 논문은 NPU 기반 임베디드 보드 중 Khadas사의 VIM3에 YOLOv8n-pose 모델을 최적화한다. 모델의 Head 구조를 Python 코드를 통해 수정하고 양자화를 진행하여 모델을 최적화한 후, 생략된 모델의 Head 구조의 Non Maximum Suppression (NMS)와 같은 후처리 과정을 Python 코드로 구현한다. 이후 VIM3에서 모델의 추론 속도와 평균 정밀도를 측정한다. 이를 통해 NPU 기반 임베디드 보드에서 포즈 추정 모델을 적용하는 방법을 제시한다.

## II. 본 론

VIM3의 NPU에 모델을 실행하기 위해 ONNX 변환 및 Amlogic사에서 제공하는 Software Developer Kit (SDK)를 사용하여 양자화가 필요하다[7]. YOLOv8n-pose는 Head에서 Neck이 제공한 특징을 처리하여 각 객체 후보에 대한 예측을 생성하고 후처리 과정을 거쳐 출력하게 된다[8]. 그림 1은 YOLOv8n-pose를 VIM3의 NPU에 실행하기 위해 모델을 변환하는 과정을 나타낸다. 그림 1의 ONNX와 같이 ONNX 변환 시 모델의 Head 구조를 수정하여 각 그리드의 모든 원시 출력값을 출력하도록 변환한다. 다음으로 그림 1의 Aml NPU SDK와 같이 Python API 기반 변환 도구를 이용하여 모델을 양자화한다. 양자화된 모델들은 그림 1의 최종 출력값과 같이 각 그리드의 원시 출력값을 1차원으로 출력한다.

그림 2는 그림 1의 최종 출력값을 사용하여 모델의 Head에서 수정된 후처리 과정을 Python 코드로 구현한 흐름도이다. 먼저, 그림 2의 바운딩 박스는 각 셀에서 예측된 좌푯값에 softmax를 적용한 후, 출력 그리드 셀과 원본 이미지 간의 크기 비율

(stride)을 사용하여 실제 이미지상의 좌표로 변환된다. 다음으로 키포인트는 각 셀에서 예측된 좌푯값과 신뢰도에 대해 sigmoid를 적용하여 그리드 셀 좌표와 stride를 바탕으로 실제 이미지상의 좌표로 변환된다. 마지막으로 변환된 좌푯값들은 NMS 알고리즘을 통해 겹치는 좌푯값을 제거한 후 출력된다.

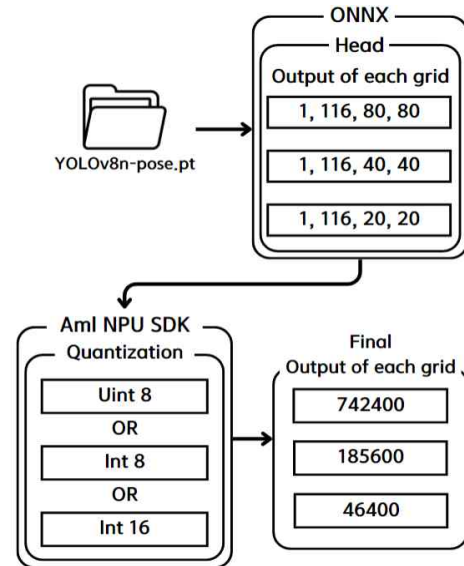


그림 1. NPU 구동을 위한 모델 변환 과정 흐름도  
Fig. 1. Model conversion process flowchart for NPU execution

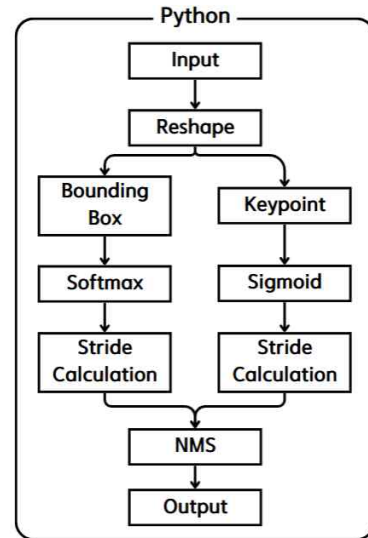


그림 2. 후처리 과정 흐름도  
Fig. 2. Flowchart of post-processing steps

### III. 실험 결과

VIM3에 YOLOv8n-pose를 최적화하기 위한 실험 환경은 표 1과 같이 진행된다. 저전력, 소형 임베디드 환경에서 NPU를 활용 가능한 VIM3를 사용하였고, 구조 해석 및 변환이 용이한 YOLOv8n-pose를 사용하였다[5,6].

표 1. 실험 환경

Table 1. Experimental environments

	Khadas VIM3
Ubuntu	20.04.6 LTS
ONNX	1.8.0
Ultralytics	8.0.86
Torch	1.8.0

성능 평가를 위해 COCO 2017 데이터셋을 사용하여 양자화된 모델의 추론 속도와 평균 정밀도를 측정하였다[9]. 데이터셋의 정답 키포인트 범위에 예측된 키포인트가 포함되면 이를 올바른 예측으로 간주하여 성능 평가를 진행하였다. 표 2는 모델의 추론 속도와 평균 정밀도를 보여준다. 추론 결과는 uint8로 양자화된 모델의 추론 속도가 23.13ms로 가장 우수하였으며, int16으로 양자화된 모델의 평균 정밀도가 0.763으로 가장 높게 출력되었다.

표 2. 모델 추론 결과

Table 2. Model inference results

Model	Inference Speed	Average Precision
uint8	23.13ms	0.758
int8	31.83ms	0.744
int16	60.33ms	0.763

### IV. 결론

본 논문에서는 NPU 기반 임베디드 보드 중 Khadas사의 VIM3에 YOLOv8n-pose를 최적화 및 구현하였다. 그 결과 양자화된 모델 중, 평균 정밀도는 int16 모델이 0.7633으로 가장 높았으며 추론 속도는 uint8 모델이 23.13ms로 가장 빠른 것을 확인하였다. 따라서 NPU 기반 임베디드 보드에 포즈 추정 모델을 적용할 수 있음을 확인하였다. 향후 본

시스템을 활용하여 모바일 방법 로봇에 적용 가능한 행동 감지 시스템에 사용할 예정이다.

### 참고 문헌

- [1] L. Wei et al., "Advances in the Neural Network Quantization: A Comprehensive Review" Appl. Sci., vol. 14, no. 7445, pp. 1-14, 2024.
- [2] S. M. Yoo et al., "Structure of Deep Learning Inference Engines for Embedded Systems" Proc. IEEE ICTC, pp. 920-922, October, 2019.
- [3] L. Chen et al., "Onboard Real-time Object Detection for UAV with Embedded NPU" Proc. 11th IEEE CYBER, pp. 192-197, July, 2021.
- [4] B. H. Phong et al., "Performance evaluation of license plate detection using deep neural networks on NPU VIM3 hardware platform" Proc. ATC, pp. 92-97, October, 2022.
- [5] S. Ali et al., "Decentralized Edge-AI System for Real-Time Syrian License Plate Recognition Using Khadas Vim3" Proc. 6th REEPE, pp. 1-5, February, 2024.
- [6] P. Zheng, Y. Hu, and Y. Chen "Application of Enhancing YOLOv8 Algorithm Using FasterNet Structure in Human Pose Estimation" Proc. 9th ICCS, pp. 207-212, April, 2024.
- [7] S. Lee et al., "NPU-aware Parameter Generation in Quantization ONNX Model" Proc. IEEE ICCE-Asia, pp. 1-3, November, 2021.
- [8] J. Li, "Improving the Application of YOLOv8 in Image Object Detection" Proc. 6th IEEE Access, pp. 668-673, May, 2024.
- [9] S. Xiang, X. Chen, and J. Zhou "An Efficient Method for Boosting Human Pose Estimation" Proc. IEEE BMSB, pp. 1-6, August, 2021.

# 엣지 컴퓨팅 환경을 위한 경량화 컨테이너 런타임 설계 및 성능 검증

김창인\*, 허유정\*\*, 최창범\*\*

## Design and Performance Evaluation of a Lightweight Container Runtime for Edge Computing Environments

Changin Kim\*, Youjeng Heo\*\*, and Changbeom Choi\*\*

### 요 약

본 연구에서는 엣지 환경에 최적화된 경량화 컨테이너 런타임을 제안하여, 빠른 애플리케이션 시작 속도와 효율적인 자원 관리를 목표로 설계하였다. 네임스페이스와 cgroups 설정을 단순화하고, 파일 시스템 및 네트워크 구성을 최적화함으로써 엣지 장치의 제한된 자원 내에서 성능을 극대화하고자 하였다. Raspberry Pi 4B를 활용한 실험을 통해 제안된 경량화 컨테이너 런타임의 성능을 검증하였으며, 이를 통해 엣지 컴퓨팅 환경에서 자원 효율성과 실시간성을 효과적으로 지원할 수 있음을 확인하였다. 본 연구는 엣지 컴퓨팅에 특화된 컨테이너 솔루션의 가능성을 제시하며, 향후 다양한 엣지 환경에 맞춘 최적화 방안으로의 확장을 목표로 한다.

### Abstract

This study proposes a lightweight container runtime optimized for edge environments, aiming to improve application startup speed and resource efficiency. By simplifying namespaces and cgroups configurations and optimizing file system and network settings, the proposed runtime is designed to enhance performance within the limited resources of edge devices. Experimental validation was conducted using a Raspberry Pi 4B, demonstrating that the lightweight container runtime effectively supports the resource efficiency and real-time responsiveness required in edge computing. This research highlights the potential of a container solution tailored for edge computing and aims to extend optimization strategies for various edge environments in future work.

### Key words

Lightweight Container Runtime, Edge Computing Optimization, Resource Efficiency

### 1. 서 론

엣지 컴퓨팅(Edge Computing)은 데이터를 생성하

는 기기 근처에서 처리하는 분산형 컴퓨팅 구조로, 자율 주행, IoT, 스마트 시티와 같은 실시간 응답성이 요구되는 분야에서 광범위하게 활용되고 있다.

\* 국립한밭대학교, 20217140@edu.hanbat.ac.kr,

\*\* 국립한밭대학교, 20211939@edu.hanbat.ac.kr, \*\* 국립한밭대학교, cbchoi@hanbat.ac.kr

이러한 환경에서는 클라우드로 데이터를 전송하여 처리하는 방식보다 신속하고 효율적인 데이터 처리가 필수적이며, 엣지 장치의 제한된 자원으로 높은 성능과 저비용 자원 관리를 수행하는 것이 요구된다. 이에 따라 애플리케이션의 빠른 시작과 안정적인 자원 활용은 엣지 컴퓨팅의 성공적인 구현에 중요한 요소로 작용한다[1],[2].

컨테이너(Container) 기술은 엣지 컴퓨팅에서 경량화된 가상화를 가능하게 하여 자원 소모를 줄이고, 가상 머신에 비해 더 빠른 실행 속도를 제공한다. 이는 엣지 환경에서 필요한 애플리케이션 격리와 유연성을 확보하는 데 적합한 솔루션으로 평가된다. 특히, 컨테이너는 하드웨어와의 결합을 최소화하면서 독립적인 실행 환경을 제공하므로 엣지 디바이스의 소규모 리소스 내에서도 효율적으로 운영될 수 있다. 그러나 널리 사용되는 Docker와 같은 범용 컨테이너 런타임은 클라우드 환경에 최적화되어 있어 엣지 장치에서는 불필요한 자원 소모와 느린 애플리케이션 시작 시간 등의 문제가 발생할 수 있다. 따라서 엣지 환경에서 자원 효율적 관리와 빠른 시작 속도를 제공하는 경량화 컨테이너 런타임에 대한 연구가 필요하다[3],[4].

본 연구에서는 엣지 컴퓨팅의 특수한 요구 사항을 충족하기 위해 최적화된 경량화 컨테이너 런타임을 제안한다. 제안된 런타임은 기존 Docker 런타임 대비 더 빠른 애플리케이션 시작 속도와 효율적인 메모리 사용을 목표로 설계되었다. 연구의 주요 목적은 엣지 환경에서 자원 효율성을 입증하고, 제안된 경량화 컨테이너 런타임이 애플리케이션 시작 속도와 자원 활용 측면에서 더 나은 성능을 제공할 것을 검증하는 것이다. 이를 통해 엣지 컴퓨팅의 특수한 성능 요구를 충족하는 최적화된 컨테이너 기술의 가능성을 탐구하고, 다양한 실사용 환경에서의 응용 가능성을 제시하고자 한다.

## II. 본 론

본 연구는 엣지 컴퓨팅 환경에서 기존 Docker 컨

테이너 런타임의 자원 소모 요인을 분석하고, 이를 해결하기 위해 경량화된 컨테이너 런타임을 설계하여 애플리케이션 시작 속도와 자원 효율성을 높이는 데 중점을 둔다. Docker는 클라우드 환경에서 강력한 격리와 유연성을 제공하지만, 이러한 설계로 인해 엣지 장치에서는 불필요한 자원 소모를 초래할 수 있다. 본 연구에서는 이러한 문제를 해결하기 위해 다음과 같은 최적화 요소를 통해 경량화 컨테이너 런타임을 설계하였다. [그림 1]은 본 연구의 설계를 시각적으로 표현한 도식이다.

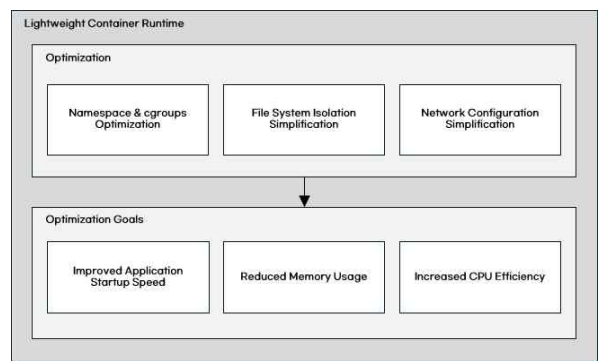


그림 948 경량형 컨테이너 런타임 설계도

### 2.1 네임스페이스 및 cgroups 최적화

Docker는 프로세스, 네트워크 및 사용자 격리를 위해 네임스페이스와 cgroups를 활용하여 독립적인 환경을 제공하지만, 이러한 복잡한 격리는 엣지 환경에서 과도한 자원 소비를 야기한다. 본 연구에서는 사용자 네임스페이스 격리를 단순화하고, 프로세스 네임스페이스에서 기본적인 독립성만을 유지하도록 하여 자원 소모를 줄였다. 또한, cgroups의 자원 할당을 제한하여 불필요한 CPU 및 메모리 소비를 줄였다. 이를 통해 필수적인 자원 격리를 유지하면서도 엣지 장치의 자원 소비를 최소화하였다.

### 2.2 파일 시스템 및 네트워크 설정 단순화

일반적인 Docker 환경에서는 overlay2와 같은 파일 시스템 격리를 사용하여 다층적인 파일 시스템 구조를 유지하지만, 이는 엣지 환경에서 복잡성을 증가시키고 자원 사용을 비효율적으로 만든다. 본 연구에서는 aufs 대신 간단한 bind mount 방식을 채택하여 애플리케이션이 실제 파일 시스템에 직접

접근하도록 하여 파일 시스템 격리를 단순화하였다. 또한, 네트워크 설정에서 다중 네트워크 브리지 및 가상 네트워크 인터페이스를 제거하고, 단일 인터페이스로 옛지 장치에 직접 연결하여 네트워크 설정을 최소화하였다.

### III. 사례 연구

본 연구의 성능 평가는 ARM 기반의 저사양 옛지 장치(Raspberry Pi 4B)에서 진행되었다. 이 장치는 메모리와 CPU 자원이 제한적이기 때문에, 옛지 환경에서 경량화 컨테이너 런타임의 성능 개선 효과를 확인하는 데 적합하다. 실험은 동일한 애플리케이션을 Docker와 경량화 런타임에서 각각 실행하여 시작 속도, 응답 시간, 메모리 사용량, CPU 사용량을 측정하였다. 성능 지표는 다음과 같이 정의하고, 각 지표를 측정하는 도구를 사용하였다.

#### 3.1 성능 지표

실험의 주요 목표는 Docker와 경량화 런타임에서의 애플리케이션 시작 속도, 메모리 사용량, CPU 사용량을 비교하여 옛지 환경에서 자원 효율성을 증명하는 것이다. 이를 위해 다음과 같은 성능 지표를 선정하였다.

애플리케이션 시작 속도는 각 런타임에서 애플리케이션이 시작되는 데 걸리는 시간을 `time` 명령어로 측정하였으며, 메모리 사용량과 CPU 사용량은 `cgroups`의 실시간 모니터링 기능을 통해 각각 MB 및 % 단위로 기록하였다. 성능 지표와 해당 측정 방법은 [표 1]에 요약되어 있다.

표 1 성능 지표와 측정 방법

Performance Metric	Measurement Tool	Unit
Startup Time	<code>time</code> command	sec
Memory Usage	<code>cgroups</code> monitoring feature	MB
CPU Usage	<code>cgroups</code> monitoring feature	%

#### 3.2 실험 절차

실험은 ARM 기반의 Raspberry Pi 4B 장치를 사용하여 진행되었으며, 운영체제로는 Ubuntu Server 20.04를 설치하였다. 이 장치는 1.5GHz CPU와 4GB RAM, 16GB SD 카드를 장착하고 있어, 옛지 환경에서 발생하는 자원 제한 상황을 반영할 수 있는 적절한 실험 환경을 제공한다. 실험은 다음과 같은 절차에 따라 수행되었다.

먼저, Raspberry Pi 4B에 Docker와 경량화 컨테이너 런타임을 동시에 설정하고, 두 런타임에서 동일한 애플리케이션을 실행하여 일관된 성능 평가를 받을 수 있도록 모든 실험 조건을 통제하였다. 이후, Docker 환경에서 기준 성능을 측정하고 이를 토대로 경량화 런타임의 성능을 측정 및 분석하였다. 성능 비교 실험은 다음과 같은 순서로 진행되었다:

첫째, 기준 성능 측정을 위해 Docker에서 애플리케이션을 실행하고, 설정된 성능 지표(애플리케이션 시작 속도, 메모리 사용량, CPU 사용량)를 기준 데이터로 기록하였다.

둘째, 경량화 런타임 성능 측정을 위해 동일한 애플리케이션을 경량화 컨테이너 런타임에서 실행하였으며, 시작 속도는 `time` 명령어로 측정하고, 메모리 및 CPU 사용량은 `cgroups` 모니터링 기능을 통해 실시간으로 기록하였다.

셋째, 데이터 수집 및 분석을 통해 각 성능 지표의 측정 데이터를 평균값으로 산출하였으며, 이를 바탕으로 Docker와 경량화 런타임 간의 성능 차이를 비교하였다. 이를 통해 두 런타임의 자원 효율성을 수치적으로 분석하고, 경량화 컨테이너 런타임의 성능 개선 효과를 정량적으로 평가하였다.

#### 3.3 실험 결과 및 분석

실험 결과, 경량화 컨테이너 런타임은 애플리케이션 시작 속도와 자원 사용량에서 Docker 대비 유의미한 성능 개선을 보였다. [표 2]는 각 성능 지표에서의 평균값과 개선율을 요약한 것이다.



표 2 성능 평가 결과 (평균값)

Performance Metric	Docker	Lightweight Runtime	Improvement (%)
Startup Time (sec)	3.5	2.4	31.4
Memory Usage (MB)	120	90	25
CPU Usage (%)	75%	60	20

경량화 컨테이너 런타임은 Docker 대비 애플리케이션 시작 속도가 평균 30% 이상 개선되었으며, 이는 네임스페이스와 cgroups 설정을 단순화하여 자원 격리를 최소화한 결과이다. 또한, 메모리와 CPU 사용량은 각각 25%와 20% 절감되었으며, 이는 파일 시스템과 네트워크 설정을 단순화한 설계 덕분이다.

이와 같은 결과는 경량화 컨테이너 런타임이 엣지 환경에서 요구되는 실시간 응답성과 자원 절약의 측면에서 기존 Docker의 한계를 극복할 수 있음을 시사한다.

#### IV. 결론

본 연구는 엣지 컴퓨팅 환경에서의 자원 효율성과 빠른 애플리케이션 시작 속도를 실현하기 위한 경량화 컨테이너 런타임의 설계와 성능을 평가하였다. 기존 Docker 런타임이 클라우드 환경에 최적화된 구조로 인해 엣지 장치에서 불필요한 자원 소모를 초래할 수 있다는 문제를 인식하고, 이를 개선하기 위해 네임스페이스와 cgroups 설정을 단순화하고 파일 시스템 및 네트워크 설정을 간소화하였다.

실험 결과, 경량화 컨테이너 런타임은 Docker 대비 애플리케이션 시작 속도에서 약 30%, 메모리 사용량에서 25%, CPU 사용량에서 20% 개선된 성능을 보였다. 이러한 성능 향상은 엣지 환경에서 실시간 응답성과 자원 절약을 요구하는 애플리케이션의 효율적인 실행을 가능하게 하며, 엣지 컴퓨팅에 특화된 컨테이너 솔루션의 잠재력을 확인할 수 있다.

향후 연구는 다양한 엣지 컴퓨팅 환경에서 경량

화 컨테이너 런타임의 성능을 더욱 최적화하고, 특정 애플리케이션 요구사항에 맞춘 맞춤형 자원 관리 방안을 제시하는 방향으로 확장할 것이다.

#### 사 사 문 구

“본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음”(2022-0-01068)

#### 참 고 문 헌

- [1] H. M. Park and T. H. Hwang, "Changes and trends in edge computing technology", The Journal of The Korean Institute of Communication Sciences, vol. 36, no. 2, pp. 41-47, 2019.
- [2] J. H. Hong, K. C. Lee and S. Y. Lee, "Trends in edge computing technology", J. Electron. Telecommun. Trends Anal., vol. 35, no. 6, pp. 78-87, 2020. doi: 10.22648/ETRI.2020.J.350608
- [3] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey and S. Linkman, "Systematic literature reviews in software engineering—a systematic literature review", Inf. Softw. Technol., vol. 51, no. 1, pp. 7-15, 2009.
- [4] K. Petersen, R. Feldt, S. Mujtaba and M. Mattsson, "Systematic mapping studies in software engineering", Proc. Int. Conf. Eval. Assessment Softw. Eng., pp. 68-77, 2008.

# 스마트 차량을 위한 운전자 행동 인식 분류

정현구\*, 강현수\*\*

## Driver Behavior Classification for Smart Vehicle Systems

Jung Hyun Gu\* and Kang Hyun Soo\*\*

### 요약

본 논문에서는 스마트 차량 시스템에서 운전자 행동을 인식하기 위한 분류 모델을 제시하였다. 운전자의 측면 이미지를 기반으로 총 10개의 클래스(정상 운전, 텍스트링, 전화 통화, 라디오 조작, 음료 섭취 등)를 분류하기 위해 전이 학습 기반의 VGG19과 ResNet152 모델을 사용하였다. ResNet152는 VGG19에 비해 더 깊은 네트워크 구조로 인해 학습 시간이 더 오래 걸렸지만, 복잡한 행동 패턴 인식에서 더 높은 분류 정확도를 기록하였다. 이를 통해 운전자 안전을 보장하는 스마트 차량 시스템 구현에 기여할 것으로 기대된다.

### Abstract

This paper presents a classification model for recognizing driver behavior in smart vehicle systems. Based on side-view images of the driver, the model classifies behavior into 10 categories, such as normal driving, texting, phone calls, radio operation, and drinking, using transfer learning-based VGG19 and ResNet152 models. Compared to VGG19, ResNet152 has a deeper network structure, resulting in longer training times, but it achieves higher classification accuracy in recognizing complex behavior patterns. This is expected to contribute to the development of smart vehicle systems that enhance driver safety by ensuring the timely detection of risky behaviors.

### Key words

Driver Behavior Recognition, VGG, ResNet152, Transfer Learning, Classification Model, Deep Learning, Behavior Analysis

## 1. 서론

스마트 차량 기술의 발전으로 인해 자율주행 시스템과 운전자 지원 시스템에 대한 관심이 빠르게 증가하고 있다. 이러한 시스템은 도로 안전성을 향상시키고 운전자와 승객의 안전을 보장하는 데 중요한 역할을 한다. 특히, 운전 중에 발생하는 다양

한 운전자 행동을 실시간으로 인식하고 분석하는 것은 사고 예방과 운전자의 위험 행동에 대한 신속한 대응을 가능하게 하여 안전을 강화할 수 있다.

본 논문은 운전자의 행동을 총 10개의 클래스(정상 운전, 텍스트링, 전화 통화, 음료 섭취, 라디오 조작 등)로 분류하여 운전자 행동을 정확히 인식하고 분류하는 데 초점을 맞춘다. 이러한 행동 인식 분류

\* 충북대학교 정보통신공학부 학부생, wjdgusrn951@naver.com

\*\* 충북대학교 정보통신공학부 교수, hskang@cbnu.ac.kr

※ 이 논문은 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업(2019-0-01183)의 지원을 받아 작성되었음

는 운전자의 위험 행동을 조기에 탐지하고 신속히 대응하는 데 기여하고, 스마트 차량 시스템의 안전성을 더욱 강화할 수 있다.

본 논문에서는 VGG19와 ResNet152 모델을 사용하여 보다 정교한 운전자 행동 인식 시스템 개발에 기여할 수 있는 연구 결과를 제시한다.

## II. 본 론

본 연구는 스마트 차량 시스템에서 운전자 행동을 인식하여 교통사고를 예방하고 운전자의 안전을 목적을 두고 있다. 운전 중 부주의한 행동은 심각한 사고로 이어질 수 있으므로, 이를 정확히 인식하고 조기에 대응하는 시스템 개발이 중요하다. 이를 위해 VGG19와 ResNet152 모델을 사용하여 총 10가지 운전자 행동을 분류하는 실험을 진행하였다.

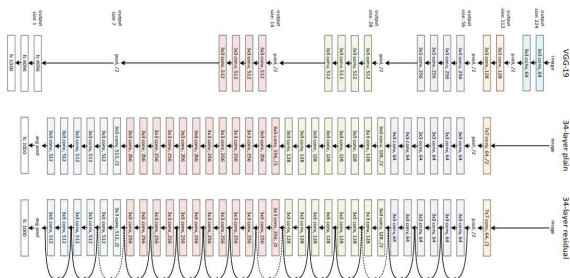


그림 1. VGG19 과 ResNet 34-layer plain과 34-layer residual 구조

Fig. 1. VGG19 and the structure of ResNet 34-layer plain and 34-layer residual

VGG19와 ResNet152는 모두 널리 사용되는 이미지 분류 모델로, 전이 학습을 통해 대규모 데이터셋에서 사전 학습된 가중치를 활용하여 운전자 행동 인식에 적용되었다. VGG19는 19개의 레이어로 구성된 비교적 단순한 구조 덕분에 학습 속도가 빠르다. 반면, ResNet152는 152개의 레이어와 [그림 1]에서 보이는 잔차 연결(Residual Connection) 구조를 사용하여, 더 깊은 네트워크에서도 기울기 소실 문제를 해결하며 복잡한 행동 패턴을 효과적으로 학습할 수 있다.

총 22,424개의 운전자의 측면 이미지 데이터셋을 사용하여 행동을 분류하였다. 이 데이터는 정상 운전, 텍스팅(왼손 및 오른손), 전화 통화(왼손 및 오

른손), 라디오 조작, 음료 섭취, 뒤로 몸을 돌리는 행동, 머리 또는 화장 손질, 승객과 대화 등 총 10가지 운전자 행동을 포함한다. 각 클래스는 운전 중 발생할 수 있는 다양한 행동을 포함하며, 해당 행동들을 인식하는 것을 목표로 한다.

두 모델 모두 Adam 옵티마이저와 categorical crossentropy 손실 함수를 사용하였고, 학습률은 ReduceLROnPlateau 콜백을 통해 자동으로 조정되었다. 총 20 에포크 동안 학습이 진행되었으며, 배치 크기는 16으로 설정되었다.

Categorical crossentropy 손실 함수는 모델이 예측한 클래스 확률과 실제 레이블 간의 차이를 계산하여 손실 값을 도출한다. 이때 사용된 손실 함수의 수식은 다음과 같다.

$$L = - \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (1)$$

여기서 N은 클래스의 총 개수이며,  $y_i$ 는 실제 레이블(원-핫 인코딩으로 표현)이고,  $\hat{y}_i$ 는 모델이 예측한 클래스 확률을 의미한다. 이 수식은 예측된 확률과 실제 값이 일치할수록 손실이 줄어들도록 설계되었으며, 모델이 학습을 통해 예측의 정확도를 높이는 방향으로 손실을 최소화하게 된다.

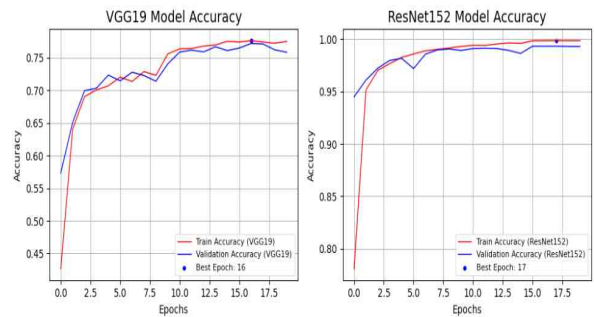


그림 2. VGG19와 ResNet152의 학습 및 검증 정확도 비교

Fig. 2. Comparison of Training and Validation Accuracy for VGG19 and ResNet152

모델의 성능은 정확도(Accuracy)와 손실(Loss)을 기준으로 평가되었다. [그림 2]는 VGG19와 ResNet152의 학습 및 검증 정확도 변화를, [그림 3]은 각 모델의 손실 변화를 보여준다.

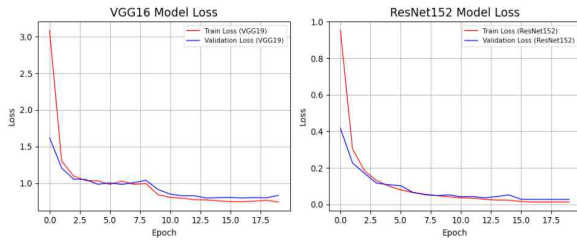


그림 3. VGG19와 ResNet152의 학습 및 검증 손실 비교  
Fig. 3. Comparison of Training and Validation Loss for VGG19 and ResNet152

VGG19는 초기 학습 속도가 빠르고, 빠른 수렴을 보였으나, 복잡한 행동 패턴을 분류하는 데에는 낮은 정확도를 보였다. 반면, ResNet152는 더 깊은 네트워크 구조와 잔차 연결(Residual Connection)을 통해 더욱 복잡한 행동을 인식할 수 있었으며, 검증 정확도에서 VGG19가 보여준 0.7721보다 높은 0.9931이라는 수치를 보여주며 VGG19 보다 우수한 성능을 기록했다.

Grad-CAM 결과 [그림 4]와 같이, 모델이 운전 중 전화 통화와 같은 특정 행동을 인식할 때, 운전자의 얼굴과 손, 핸들과 같은 중요한 영역에 집중하는 것을 확인할 수 있었다. 이를 통해 학습된 모델을 통해 복잡한 행동 패턴을 효과적으로 인식할 수 있음을 시각적으로 확인할 수 있었다.

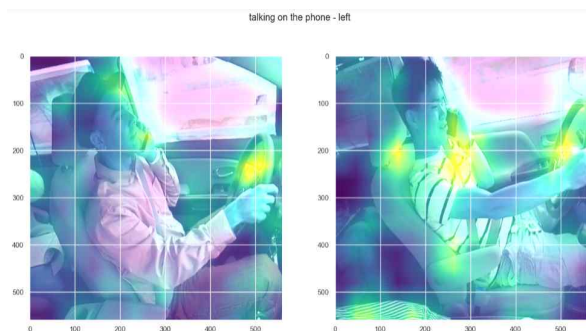


그림 4 . 전화 통화 행동에 대한 Grad-CAM 시각화 결과  
Fig. 4. Grad-CAM Visualization for 'Talking on the Phone - Hand' Action

### III. 결 론

본 논문에서는 스마트 차량 시스템에서 운전자 행동을 인식하는 딥러닝 기반의 분류 모델을 제안하고, 이를 위해 VGG19와 ResNet152 모델을 사용하여 총 10가지 운전자 행동을 분류하였다. 실험 결

과, ResNet152는 더 깊은 네트워크 구조와 잔차 연결(Residual Connection) 통해 복잡한 행동 패턴을 효과적으로 학습할 수 있었으며, [4]와 비교했을 때 높은 검증 정확도 0.9931을 기록하며 VGG19보다 뛰어난 성능을 보였다.

또한, Grad-CAM(Gradient-weighted Class Activation Mapping) 기법을 통해 ResNet152가 행동을 분류할 때 운전자의 얼굴과 손과 같은 중요한 영역에 집중하는 것을 시각화하여 직접 확인할 수 있었다. 이를 통해 ResNet152는 운전자 행동 인식 시스템을 개발하는 데 있어 더 나은 선택지임을 입증하였다.

이번 연구를 통해 스마트 차량 시스템에서 운전자 행동을 실시간으로 인식하고, 이를 통해 교통사고를 예방하는 데 기여할 수 있는 중요한 기초 자료를 제공한다. 앞으로 더 많은 데이터를 활용하거나, 경량화된 모델을 적용하여 실시간 응답 성능을 개선하여 스마트 차량 발전에 기여할 것으로 기대된다.

### 참 고 문 헌

- [1] Birkan Buyukarikan, Erkan Ulker , "Classification of physiological disorders in apples fruit using a hybrid model based on convolutional neural network and machine learning methods", SCIE, vol 34, pp. 16973-16988, May 2022
- [2] HSheldon Mascarenhas, Mukul Agarwal, "A comparison between VGG16, VGG19 and ResNet152 architecture for image classification", Proc. CVPR, pp. 162, Feb 2022.
- [3] Sumaira Manzoor, Eun-Jin Kim, Sung-Hyeon Joo, Sang-Hyeon Bae, Gun-Gyo In, Kyeong-Jin Joo, "Edge Deployment Framework of GuardBot for Optimized Face Mask Recognition With Real-Time Inference Using Deep Learning", IEEE Trans. Pattern Anal. Machine Intell., vol. 10, pp. 11, July 2022.
- [4] Cong Duan, Yipeng Gong, Jiakai Liao, Minghai Zhang, Libo Cao "FRNet: DCNN for Real-Time Distracted Driving Detection Toward Embedded Deployment", IEEE Trans. Pattern Anal. Machine Intell., PP(99):1-14, Sep 2023.

# 이미지 분석 기반 음원 생성 서비스

정현섭\*, 이세민\*\*, 손민섭\*\*, 김신\*\*\*, 박건\*\*\*, 김희선\*\*\*

## Image Analysis-Based Sound Source Generation Service

Hyeonseop Jeong\*, Semin Lee\*\*, Minseop Son\*\*, Shin Kim\*\*\*, Geon Park\*\*\*,  
and Heesum Kim\*\*\*

### 요약

본 논문에서는 이미지 분석을 통해 대상 이미지에 알맞은 음원을 생성하고, 이를 SNS 및 다양한 플랫폼에서 공유할 수 있는 서비스에 대해 다룬다. 이미지의 색상, 분위기, 감정 등을 분석하여 인공지능 기반의 음원을 자동 생성하며, 사용자는 이를 다양한 소셜 미디어 플랫폼에서 손쉽게 공유할 수 있다. 또한, 생성된 음원에 대해 다른 사용자와 피드백을 주고받을 수 있으며, 인스타그램, 유튜브와 같은 여러 플랫폼을 통해 음원을 활용하여 자신의 성향을 나타낼 수 있다.

### Abstract

In this paper, we discuss a service that generates sound sources suitable for target images through image analysis and shares them on SNS and various platforms. It automatically creates artificial intelligence-based sound sources by analyzing the color, mood, and emotion of the image, and users can easily share them on various social media platforms. You can also exchange feedback with other users about the created sound source, and express your tendencies by using the sound source through various platforms such as Instagram and YouTube.

### Key words

Image Analysis, Deep Learning, Image-to-Text, Music Sharing, GCP(Google Cloud Platform)

## 1. 서론

지금의 디지털 이미지와 인공지능 기술의 발전은 예술과 음악 생성 분야에 새로운 가능성을 열어주고 있다[1]. 본 논문에서 다루고자 하는 ‘이미지 분석 기반 음원 생성 서비스’는 사용자가 업로드한

이미지의 색상, 분위기, 감정 등 다양한 시각적 요소를 분석해, 해당 이미지에 어울리는 음악을 자동으로 생성하는 것을 목표로 하는 서비스이다. 이 서비스는 이미지 분석을 넘어 인공지능 기술을 통해 이미지와 음악을 결합한 특별한 사용자 경험을 제공한다.

\* 국립안동대학교 멀티미디어공학과, jhs789654123@gmail.com,

\*\* 국립안동대학교 AI융합학과, dltpals7498@gmail.com, spcle2412@gmail.com,

\*\*\* 국립안동대학교 소프트웨어융합학과, rlatls999@gmail.com, nugtrup@naver.com, hskim@anu.ac.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

최근에 자신의 성격을 파악하거나 표현하는 수단, 예를 들면 MBTI 성격 유형 테스트와 같은 것이 인기를 끌며, 사람들이 자신을 더욱 깊이 이해하고 표현하는 방식에 많은 관심을 보인다. 이러한 흐름 속에서 개인의 정체성이나 감정을 표현하는 도구들이 빠르게 확산하고 있으며, 이는 개인의 취향을 넘어서 사람들 간의 연결성을 형성하는 중요한 요소가 되고 있다[2, 3]. 본 서비스는 이와 같은 트렌드에 맞추어 사용자가 자신의 이미지를 통해 자신을 나타낼 수 있는 배경음악을 제공함으로써 자기표현을 돕고자 한다.

기존의 음악 생성 시스템들은 주로 텍스트 기반의 프롬프트로 음악을 생성하는 것에 비해, 본 서비스는 이미지를 통해 음악을 생성한다는 점에서 차별화된다[4]. 본 연구에서 Google Cloud Platform의 이미지 분석 기술을 활용하여 이미지에서 텍스트 설명을 생성하고, 이를 기반을 감정과 분위기를 분석한 후 Meta의 MusicGen 모델을 활용해 적합한 음악을 생성하는 과정을 거친다. 이를 통해 사용자가 업로드한 이미지에 가장 잘 어울리는 개인 맞춤형 배경음악을 자동으로 생성해 제공하는 방안을 연구한다.

## II. 본 론

본 연구에서는 Meta MusicGen 모델을 기반으로 사용자 맞춤형 음악을 생성하는 시스템을 구축한다. 시스템의 주요 과정은 다음과 같이 구성된다. 사용자가 웹 애플리케이션을 통해 이미지를 업로드하면, 이 이미지는 Google Cloud Storage(GCS)에 저장된다. 이후 업로드된 이미지는 OpenAI API를 사용하여 이미지에 대한 설명(캡셔닝)을 생성하며, 이 설명이 MusicGen 모델의 입력값으로 사용이 된다. 이 과정에서 생성된 설명 텍스트를 기반으로 모델은 이미지에 어울리는 배경음악을 생성하게 된다. 생성된 음악 파일은 다시 GCS에 저장되고, 사용자는 웹 애플리케이션을 통해 이를 접근하고 다운로드할 수 있다.

본 시스템의 개발을 위해 Google Cloud Platform(GCP)의 다양한 서비스를 활용한다. 먼저, 사용자가 이미지를 업로드하면, 이 이미지는 Google Cloud

Storage(GCS)에 저장된다. GCS는 이미지 데이터를 안전하게 저장하고 관리하는 역할을 하며, 이후 이미지를 분석하는데 필요한 데이터를 제공한다. 이미지 업로드 후에는 OpenAI의 API를 사용하여 이미지에 대한 설명(캡셔닝)을 자동으로 생성한다. 이 설명은 이미지의 색상, 분위기, 인물 등을 분석한 결과로, MusicGen 모델의 입력값으로 사용된다. 이를 통해 사용자가 업로드한 이미지에 가장 적합한 음악을 생성할 수 있게 한다.

MusicGen 모델은 Meta에서 개발한 텍스트-음악 변환 모델로, 텍스트 입력을 기반으로 음악을 생성한다. 본 연구에서는 이 모델을 Google Cloud의 Vertex AI를 이용해 배포한다. Vertex AI는 모델의 학습, 배포, 추론을 위한 통합 플랫폼으로, 본 시스템에서는 모델 배포와 추론을 자동화한다. 이를 통해 사용자는 복잡한 설정 없이도 손쉽게 이미지를 업로드하고 음악을 생성할 수 있다.

Vertex AI 파이프라인을 사용하여 이미지 캡셔닝 생성부터 음악 생성, 결과 저장까지의 모든 과정을 자동화한다. 파이프라인을 통해 데이터 입력, 모델 추론, 결과 저장이 일관되게 이루어져 시스템의 안정성과 효율성을 높일 수 있다. 다시 말해서 사용자가 이미지를 업로드하면 파이프라인이 자동으로 실행되어 캡셔닝 생성과 음악생성을 처리하며, 생성된 음악 파일은 GCS에 저장된다.

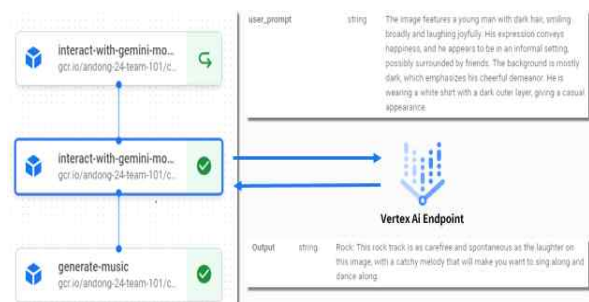


그림 1. 자동화 된 음악생성 파이프라인  
Fig. 1. Automated music creation pipeline

사용자가 웹기반의 인터페이스를 통해 쉽게 접근하고 사용할 수 있게 설계 및 구현이 된다. 사용자가 이미지를 업로드하면, 챗봇 형식의 인터페이스를 통해 이미지에 대한 설명이 자동으로 생성되고, 이를 기반으로 음악이 생성되는 과정을 시각적으로

확인할 수 있다. 챗봇 인터페이스는 자연어 처리를 이용하여 사용자와의 상호작용을 자연스럽게 흥미롭게 만들어준다. 음악 생성이 완료되면, 생성된 음악 파일은 사용자가 웹 애플리케이션을 통해 직접 재생하거나 다운로드 할 수 있도록 제공된다. 이를 통해 사용자는 생성한 음악을 소셜 미디어에 공유하거나 개인적인 용도로 사용할 수 있다.

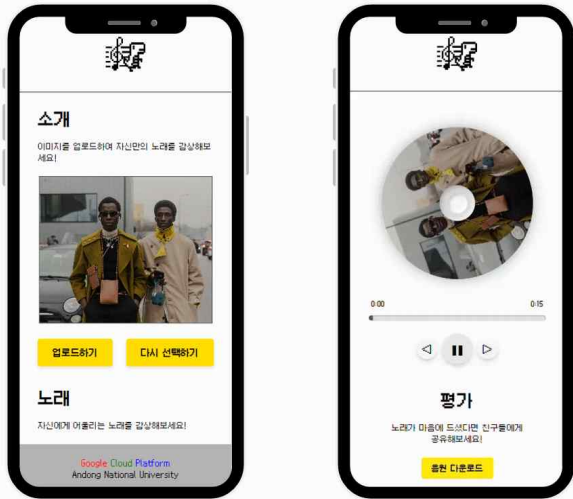


그림 2. 음악 생성 사용자 화면  
Fig. 2. Music creation user screen

시스템의 효율성을 높이기 위해 Google Cloud의 다양한 서비스를 적극적으로 활용한다. 특히 Vertex AI의 파이프라인 기능을 통해 모델 배포와 추론 과정을 자동화하여 개발 시간을 단축하고, 반복적인 작업을 줄인다. 또한, GCS에 저장된 음악 파일에 자동 삭제 기능을 도입하여 스토리지 사용량을 줄이고 운영 비용을 절감한다. 시스템의 확장성 측면에서, 본 연구에서 구축한 구조는 새로운 모델이나 기능을 추가하기에 용이하도록 설계한다. 모델의 업데이트나 변경이 필요할 때도 Vertex AI를 활용해 빠르고 효율적으로 배포할 수 있으며, 이를 통해 사용자는 최신의 기능과 성능을 유지한 상태로 서비스를 이용할 수 있다.

### III. 결 론

본 연구에서는 Meta의 MusicGen 모델을 기반으로 사용자 맞춤형 음악 생성 시스템을 개발한다. 사

용자가 업로드한 이미지를 분석하여 그에 어울리는 배경음악을 자동으로 생성하고, 이를 웹 애플리케이션을 통해 제공하는 시스템을 구축한다. 이를 통해 사용자는 자신을 표현할 수 있는 음악을 간편하게 생성하고, 이를 소셜 미디어에서 공유할 수 있다.

본 시스템은 Google Cloud Platform(GCP)의 Vertex AI와 Google Cloud Storage(GCS)를 적극적으로 활용하여 모델의 배포와 추론을 효율적으로 처리하였으며, 사용자가 복잡한 설정 없이도 손쉽게 음악을 생성할 수 있다. Vertex AI의 파이프라인 기능을 통해 모델 추론 과정을 자동화하고 GCS에 음악 파일을 저장함으로써, 실시간에 가까운 반응 속도와 효율성을 유지할 수 있었다.

이 연구 결과는 개인을 표현할 수 있는 새로운 형태의 AI 콘텐츠 생성 가능성을 보여주며, 향후 다양한 콘텐츠 생성 분야에서의 응용 가능성을 제시한다. 특히, 사용자 맞춤형 AI 기반 서비스의 발전과 예술과 기술의 융합을 통한 창작에 긍정적인 영향을 줄 수 있을 것으로 기대한다.

### 참 고 문 헌

- [1] Park Eun Ji, "An In-Depth Analysis of the Latest Artificial Intelligence Music Generators," The Korean Journal of Arts Studies, No. 44, pp. 5-30, 2024.
- [2] Jang, Nayeon, Joo, Jinyoung, and Shin, Kyulee, "The relationship between Narcissism, Social Media usage motivation and Conspicuous leisure consumption among MZ generation," Korean Journal of Leisure, Recreation & Park, Vol. 46, No. 3, pp. 53-64, 2022.
- [3] Weiqi Su and Kim hyun tai, "Case Analysis and Characteristics of Popular Music Creative Activities Using Artificial Intelligence," The Journal of Humanities and Social science, Vol. 13, No. 2, pp. 1937-1948, 2022.
- [4] Han Man-hyuk, "[Generation AI Guide] Just enter text to complete your own song 'Suno'", Dong-A Ilbo, 2024, 08. 30.

# U-Net 및 U-Net++ 기반 MRI 뇌종양 탐지

김효신\*, 정현구\*, 강현수\*\*

## MRI brain tumor detection based on U-Net and U-Net++

Kim Hyo Sin\*, Jung Hyun Gu\*, and Kang Hyun Soo\*\*

### 요약

본 논문에서는 U-Net과 U-Net++ 모델을 활용하여 MRI 영상에서 뇌종양을 자동으로 탐지하고 분할하는 방법을 제안하였다. 두 모델은 모두 다단계 컨볼루션 신경망(CNN) 구조를 기반으로 설계되었으며, 공개된 MRI 데이터셋을 사용해 학습되었다. 실험 결과, 두 모델 모두 높은 정확도와 민감도로 뇌종양을 성공적으로 탐지했으며, U-Net++은 U-Net에 비해 더 세밀한 예측 성능을 보여주었다. 이는 MRI 기반 뇌종양 탐지의 자동화 가능성을 입증하였으며, 이러한 모델들이 향후 자동화된 진단 도구로서 의료 영상 분석 분야에 기여할 것으로 기대된다.

### Abstract

This paper presents a method for automatic detection and segmentation of brain tumors in MRI images using U-Net and U-Net++ models. Both models are based on a multi-stage convolutional neural network (CNN) architecture and were trained using publicly available MRI datasets. Experimental results demonstrate that both models achieved high accuracy and sensitivity in detecting brain tumors, with U-Net++ showing superior performance in more precise prediction compared to U-Net. This study confirms the potential for automating brain tumor detection in MRI images, and these models are expected to contribute to automated diagnostic tools in medical image analysis in the future.

### Key words

U-Net, U-Net++, Brain Tumor Detection, Deep Learning, Convolutional Neural Networks, Diagnostic Imaging

## 1. 서론

뇌종양은 환자의 생명에 중대한 영향을 미치는 질환으로, 조기 발견과 정확한 진단이 생존율을 높이는 데 매우 중요하다. MRI는 뇌종양 진단에 널리 사용되는 비침습적 기법으로, 종양 부위를 상세하게 시각화할 수 있지만, 수작업으로 진행되는 종양 분

할은 시간과 노력이 많이 들고, 일관된 결과를 내기 어렵다.

딥러닝 기술의 발전으로 의료 영상 분석에서 자동화된 진단 도구의 가능성이 열리고 있다. 특히, U-Net과 같은 컨볼루션 신경망(CNN) 기반 모델은 이미지 분할 작업에서 탁월한 성능을 보여, 종양과 같은 구조를 정확하게 탐지하는 데 유용하다. 최근

\* 충북대학교 정보통신공학부 학부생, gimhyoshin@gmail.com, wjdgusrn951@naver.com

\*\* 충북대학교 정보통신공학부 교수, hskang@cbnu.ac.kr

※ 이 논문은 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업(2019-0-01183)의 지원을 받아 작성되었음



U-Net++ 모델은 U-Net을 개선한 구조로, 개선된 스킵 연결을 통해 분할 경계의 정밀도를 높였으며, 복잡한 종양 구조에서 더 우수한 성능을 보인다.

본 연구에서는 U-Net과 U-Net++을 비교하여 MRI 영상에서 뇌종양을 자동으로 탐지하고 분할하는 프로그램을 개발하였다. 두 모델 모두 공개된 MRI 데이터셋을 활용해 학습되었으며, 각각의 성능을 비교 분석한 결과, U-Net++은 U-Net에 비해 더 세밀한 예측 성능을 보여주었다. 이를 통해 수작업 분할의 비효율성을 개선하고, 자동화된 진단 도구로서의 실현 가능성을 평가하고자 한다. 이 연구는 뇌종양 탐지의 자동화를 위한 새로운 접근법을 제시하며, 향후 의료 영상 분석 분야에서 실질적인 기여를 할 수 있을 것으로 기대된다.

## II. 본 론

본 연구에서는 LGG Segmentation Dataset을 활용하여 뇌종양을 자동으로 탐지하고 분할하는 두 가지 딥러닝 모델인 U-Net과 U-Net++을 비교하였다. 이 데이터셋은 The Cancer Imaging Archive (TCIA)에서 제공된 MRI 영상과 저등급 글리오마(LGG) 환자의 FLAIR 이상 마스크 이미지를 포함하고 있으며, The Cancer Genome Atlas (TCGA)에 속한 110명의 환자 데이터를 기반으로 한다. 이를 통해 두 모델이 종양 부위를 학습할 수 있도록 데이터를 구성하였다.

U-Net과 U-Net++은 모두 컨볼루션 신경망(CNN) 기반의 이미지 분할 모델로, 다운샘플링과 업샘플링 단계를 통해 이미지에서 종양 영역을 예측한다. U-Net은 간단한 구조로 높은 성능을 보이지만, U-Net++은 개선된 스킵 연결 구조를 통해 종양 경계를 더욱 정밀하게 예측할 수 있다 [그림 1].

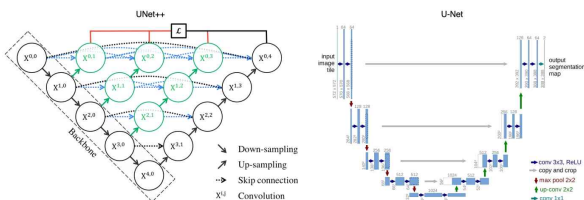


그림 1. U-Net 및 U-Net++ 모델 구조 비교  
Fig. 1. Comparison of U-Net and U-Net++ Model Architectures

MRI 이미지를 입력받고, 3x3 필터와 ReLU 활성화 함수로 주요 특징을 추출한다. 그 다음, 맥스풀링을 통해 해상도를 줄여 중요한 정보를 요약한 후, 업샘플링 단계를 거쳐 종양 영역을 예측한다.

모델 학습에는 Adam 옵티마이저와 Dice 손실 함수를 사용하여 성능을 최적화하였으며, 두 모델 모두 총 100에포크(epoch) 동안 학습되었다. 모델 성능 평가는 검증 데이터셋을 기반으로 Dice 계수와 같은 주요 성능 지표를 사용하여 이루어졌다. 학습 과정에서는 데이터 증강 기법(회전, 확대, 축소 등)을 적용하여 데이터셋의 다양성을 높였다. [그림 3]은 학습 중 손실과 Dice 계수의 변화를 시각적으로 보여준다.

[그림 2]에서 U-Net 모델은 Dice 계수 0.84를 달성하여, 예측된 분할 영역이 실제 마스크와 높은 일치도를 보였다.

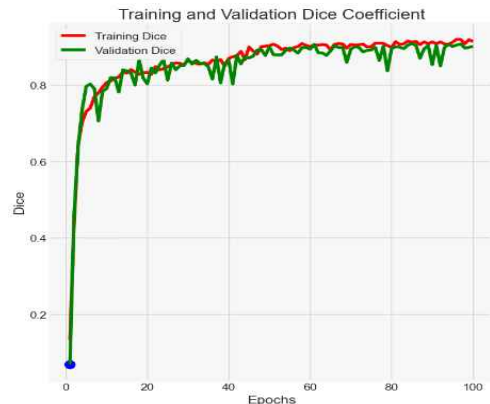


그림 2. epoch에 따른 U-Net 모델의 Dice 계수 변화  
Fig. 2. Changes in the Dice coefficient of the U-Net model over epochs

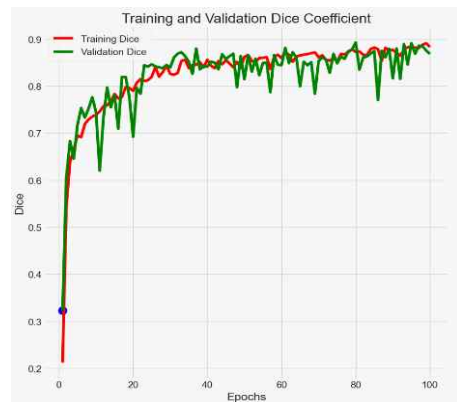


그림 3. epoch에 따른 U-Net++ 모델의 Dice 계수 변화  
Fig. 3. Changes in the Dice coefficient of the U-Net++ model over epochs

반면, [그림 3]에서 U-Net++ 모델은 Dice 계수 0.90를 달성하며 U-Net보다 더 나은 성능을 보여주었다. U-Net++은 더욱 복잡한 종양 경계에서도 우수한 성능을 발휘하여, 기존 U-Net 모델에 비해 향상된 분할 능력을 입증하였다.

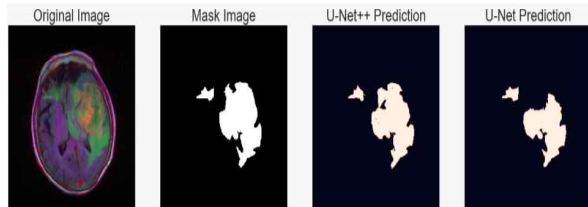


그림 4. MRI 원본 이미지, 실제 마스크 이미지, U-Net++ 및 U-Net 모델의 뇌종양 분할 결과 비교

Fig. 4. Comparison of Brain Tumor Segmentation Results in MRI: Original Image, Ground Truth Mask, U-Net++, and U-Net Predictions

[그림 4]는 실제 MRI 영상과 함께 U-Net 및 U-Net++ 모델이 예측한 종양 분할 결과를 보여준다. 두 모델 모두 종양 부위를 일관되게 탐지하고 분할하였으며, U-Net++ 모델은 U-Net 모델에 비해 종양 경계를 조금 더 세밀하게 예측한 것을 확인할 수 있다. 이를 통해 제안된 모델들의 성능을 시각적으로 비교할 수 있다.

### III. 결 론

본 논문에서는 U-Net과 U-Net++ 기반의 딥러닝 모델을 활용하여 MRI 영상에서 뇌종양을 자동으로 탐지하고 분할하는 방법을 제안하였다. U-Net 모델은 다단계 컨볼루션 신경망(CNN) 구조를 통해 종양 부위를 정확하게 분할하도록 설계되었으며, 실험 결과 높은 정확도와 민감도를 바탕으로 우수한 성능을 입증하였다. 특히, U-Net++ 모델은 Dice 계수 0.90를 달성하며 U-Net보다 더 나은 결과를 도출하였다.

이번 연구는 딥러닝 모델을 활용한 뇌종양 탐지 및 분할의 자동화 가능성을 확인하였으며, 이러한 모델들이 향후 의료 영상 분석에서 자동화된 진단 도구로 활용될 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] Fernando Daniel Hernandez-Gutierrez, Eli Gabriel Avina-Bravom, Daniel F. Zambrano-Gutierrez, Oscar Almanza-Conejo, Mario Alberto Ibarra-Manzano, Jose Ruiz-Pinales, Emmanuel Ovalle-Magallanes, Juan Gabriel Avina-Cervantes, "Brain Tumor Segmentation from Optimal MRI Slices Using a Lightweight U-Net", SCI, vol 12(10) Oct 2024.
- [2] Jiajun Zhu, Rui Zhang, Haifei Zhang, "An MRI brain tumor segmentation method based on improved U-Net", SCIE vol 21 (1), Dec 2023.
- [3] Walsh, J.; Othmani, A.; Jain, M.; Dev, S. Using U-Net network for efficient brain tumor segmentation in MRI images. Healthc. Anal. Feb 2022
- [4] Sourodir Ghosh, Aunkit Chaki & KC Santosh, "Improved U-Net architecture with VGG-16 for brain tumor segmentation". vol 44, pages 703-712, May 2021.
- [5] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, Jianming Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation", SCI, pages 3-11, Jul 2018

# Pseudo-CT 생성 오픈 소스의 데이터셋 적응성 연구

박진주\*, 권수영\*, 서현\*

## Study on the Dataset Adaptability of Pseudo-CT generation Open Source

Jin-Ju Park\*, Sue-Young Kwon\*, and Hyeon Seo\*

### 요 약

CT(Computed Tomography)와 MRI(Magnetic Resonance Imaging)는 의료 진단과 치료 계획, 시뮬레이션에 널리 활용되는 대표적인 영상 기법이다. CT는 환자의 고해상도 기하 구조를 제공하지만, 방사선 노출로 인해 반복 촬영이 어렵다. 이를 대안으로, 최근 인공지능 모델을 활용해 MRI 데이터를 pseudo-Computed Tomography(pCT)로 변환하는 기술이 발표되었다. 구현의 복잡성으로 널리 활용되지 못했으나 최근 오픈 소스가 발표되었다. 본 연구팀은 이 오픈 소스가 다른 공개 MRI 데이터셋에서도 안정적으로 pCT를 생성할 수 있는지 평가하였다. 새로운 MRI 데이터를 pCT로 변환하고, 정확도 계산을 위해 영상 분할을 수행하였다. CT 대비 pCT의 오차를 평균 절대 오차, 평균 제곱 오차, Jaccard와 Dice 계수를 통해 정량화하였다. 또한, 슬라이스 별 오차를 시각화하여 효과적으로 활용할 수 있는 범위를 확인한 결과, Slice 100~150에서 오차가 고르고 낮았다. 연구 결과, pCT는 두개골 구조가 단순하고 경계가 뚜렷한 머리 중간 부분의 MRI 영상을 활용할 때 범용성이 높다는 결론을 내렸다.

### Abstract

CT and MRI are key imaging techniques in medical diagnosis and treatment planning. While CT offers high-resolution anatomical details, it is limited by radiation exposure, which complicates repeated scans. Recently, AI-based methods for generating pseudo-Computed Tomography (pCT) from MRI have emerged, though their complexity has hindered widespread use. A new open-source tool aims to address this, and our study assesses its applicability with other MRI datasets. We converted new MRI data to pCT, performed segmentation, and evaluated errors using mean absolute error, mean squared error, and Jaccard and Dice coefficients. Slices 100-150 showed the lowest errors, indicating that pCT is most effective when using MRI of the middle head section with simple skull structures, improving its general applicability.

### Key words

Pseudo-CT, AI, Open Source, Skull Structure

---

\* 경상국립대학교 IT공과대학 컴퓨터공학과  
{jinju2435, victoria619}@naver.com, hseo0612@gnu.ac.kr (교신저자)

※ 본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (RS-2023-00280241).

## I. 서 론

의료 영상, MRI와 CT는 다양한 질병에 대한 보완적인 진단 정보를 제공한다. CT는 주로 뼈와 같은 단단한 조직 구조를 시각화하는데 유용하며, MRI는 연조직, 특히 뇌 조직의 형태를 상세히 파악하는데 주로 사용된다. 그러나 두개골 영상의 표준으로 간주되고 방사선 밀도를 추정하는 데 사용되는 CT의 경우 방사선 노출이 다른 의료기기보다 높은 수준을 보인다[1]. 따라서 본 연구는 MRI를 이용하여 CT 영상과 유사한 뼈의 정보를 반영할 수 있는 기술인 psedo-CT(pCT)의 활용성을 확인하고자 한다. 이는 특히 반복적인 시뮬레이션이 필요한 상황에서 환자의 방사선 노출을 줄이고 치료의 정확성을 높이는데 기여할 수 있다.

pCT는 MR 이미지를 기반으로 CT 이미지를 합성하는 기술로, 최근 컨볼루션 신경망(CNN)과 같이 여러 인공지능 모델을 활용하여 MR 이미지를 CT 이미지로 변환하는 기술이 발표되었다 [2-4]. 그러나 이 기술은 구현이 복잡하여 연구자들 사이에서 널리 활용되지 못하고 있다. 이러한 문제를 해결하기 위해, 최근 Yaakub[5]의 연구에서는 3D residual U-Net을 활용하여 T1-weighted MRI를 입력으로 PCT를 생성할 수 있는 오픈 소스를 공개했다. 본 연구팀은 공개된 오픈 소스를 활용하여 학습 및 테스트에 사용된 데이터셋 외에 다른 공개 MRI 데이터셋에서도 안정적이고 효과적으로 pCT를 생성할 수 있는지 평가하고자 한다.

pCT의 활용성 확인 방법은 다음과 같다. 먼저, MRI를 바탕으로 pCT를 만든다. 그다음 오차와 유사성 판단을 위해 CT, pCT에서 각각 머리 전체 영역과 두개골 영역에 대한 마스크를 만들고 씌워서 계산한다. 오차와 유사성의 평균값을 계산하고 오차가 크고 작은 영역의 시각화를 통해 적합성을 판단하고 모색할 것이다.

## II. 연구 방법

pCT 제작은 yaakub[5]의 논문에서 공개한 오픈소스를 사용하였다. 코드의 알고리즘은 데이터로부터 직접 학습하는 딥러닝 도구인 합성곱 신경망

(convolutional neural networks, CNN)을 이용하고 있다. yaakub이 CNN으로 만든 pCT의 오차를 확인해 보면, 머리 전체 영역에서는 MAE의 평균이  $109.8 \pm 13.0$ 로 나왔고 두개골 영역에선  $319.3 \pm 31.9$ 로 나왔다. 하지만 범용성 있는 사용을 위해서라면 다른 환경(컴퓨터의 종류, 환자의 다른 정도)에서도 잘 작동하여야 할 것이다.

우리는 다른 환경에서 작동을 알아보기 위해 yaakub이 공개한 코드를 이용하되 다른 공개 MRI 데이터셋으로 실험 과정을 그대로 진행하였다. MRI에서 pCT로 전환하는 과정을 거친 후, 오차의 정도를 알기 위해 평균 절대 오차(mean absolute error, MAE)와 평균 제곱 오차(mean squared error, MSE)를 확인하였다. 또한 유사성 판단을 위해 Jaccard와 Dice 계수도 계산하였다. 이때 단면적인 결과를 피하고자 5명의 MRI를 이용하여 진행하기로 하였다. MRI와 CT는 Thummerer가 공개한 데이터[6]를 이용하였다.

생성한 pCT와 CT 간의 오차 계산 시 머리 전체 영역과 두개골 영역에 대한 mask 계산을 위해 영역 분할이 필요하다. Seg3D software (NIH/NIGMS CIBS, University of Utah, UT, USA)를 이용하여 머리 전체와 두개골 영역을 분할했으며, 이 때 otsu threshold를 이용하였다. 이후엔 마스크 영역별로 오차를 계산하였다.

## III. 연구 결과

5명을 대상으로 머리 전체 영역(표1)과 두개골 영역(표2)에서 CT 대비 pCT의 오차를 계산하였다. 첫 번째 열부분은 대상의 번호를 나타내고 있고, 첫 행의 왼쪽 값부터 각각 MAE, MSE, Jaccard, Dice이다. MAE와 MSE는 둘째 자리에서, Jaccard와 Dice는 셋째 자리에서 반올림해 계산하였으며, 맨 아래 행은 5명의 평균값을 나타낸다.

새로운 공개 MRI 데이터를 사용한 경우, yaakub의 결과 대비 더 큰 오차를 관측하였다. yaakub의 연구 결과에서는 머리 전체 영역에서 평균 절대 오차:  $109.8 \pm 13.0$ 의 수치를 보고했다. 반면, 새로운 공개 MRI는 평균 절대 오차는 약 62% 증가했으며 Jaccard는 0.45, Dice는 0.61로 약한 유사성을 보였다.

표 1. 머리 전체 영역에서의 pCT 정확도

head	MAE	MSE	Jaccard	Dice
1	158.8±45.9	66865.9±28523.6	0.49±0.27	0.61±0.30
2	230.1±124.5	156686.9±122608.4	0.38±0.16	0.52±0.19
3	192.2±60.8	99062.4±48497.6	0.37±0.18	0.51±0.24
4	152.7±57.4	63608.3±35108.6	0.49±0.17	0.77±0.15
5	155.0±59.1	102351.8±63732.1	0.50±0.23	0.63±0.23

표 2. 두개골 영역에서의 pCT 정확도

skull	MAE	MSE	Jaccard	Dice
1	293.88±79.93	135578.67±76.35	0.76±0.27	0.83±0.27
2	385.88±116.20	227449.04±116527.23	0.62±0.17	0.74±0.17
3	338.81±74.80	172428.26±65924.19	0.63±0.26	0.73±0.28
4	249.33±80.30	102351.79±63732.14	0.64±0.18	0.86±0.14
5	292.07±65.07	129739.08±48554.29	0.80±0.20	0.87±0.18

두개골 영역에서는 머리 전체 영역과 달리 yaakub의 연구 결과와 유사한 오차를 확인하였다. yaakub의 연구 결과에서는 두개골 영역에서 평균 절대 오차:  $319.3 \pm 31.9$ , Jaccard:  $0.70 \pm 0.04$ , Dice:  $0.82 \pm 0.03$  의 수치를 보고했다. 새로운 공개 MRI 데이터로 만든 pCT의 경우 평균 절대 오차:  $311.99 \pm 88.26$ , Jaccard:  $0.69 \pm 0.22$ , Dice:  $0.81 \pm 0.21$  의 수치를 확인했다. 전체적인 수치는 유사한 반면 새로운 MRI 데이터를 활용한 경우 높은 분산을 관측하였다.

다음으로 슬라이스별 MAE 분포를 통해 pCT가 CT와 유사한 영역을 분석하였다. 그림 1은 환자 3번의 머리 영역에서 슬라이스별 MAE 값을 보여준다. 슬라이스 인덱스가 작고 큰 부분은 MAE의 값이 너무 크기 때문에 이 영역의 pCT는 그대로 사용하기 어렵다. Slice 100~150에서 비교적 고르고 낮은 MAE를 나타내고 있으며, 모든 대상에서 유사한 양상을 관측했다. 즉, 평균적으로 Slice 100~150에서 pCT가 CT와 비교 시 낮은 오차를 보여주고 있었다.

본 연구팀은 증가한 오차를 보이는 pCT를 의료용으로 안전하게 활용할 수 있는 방안을 모색하기 위해 상대적으로 MAE가 크고 작은 영역을 확인해 보았다 (그림 2). 오차가 작은 영역은 주로 두개골의 구조가 상대적으로 단순하고 머리와 배경의 경계가 뚜렷한 경우였다. 반면, 머리 영역에선 마스크가 흐리게 나오는 후반 부분이, 두개골 영역에선 뼈

가 복잡하게 생긴 초중반 부분과 중후반 부분이 MAE가 높게 나왔다. 따라서 머리의 중간 영역인 Slice 100~150이 의료용으로 사용하기에 가장 적합하다는 결론을 내릴 수 있었다. 이렇게 만들어진 pCT의 경우 초음파 결과를 확인하는 등 다양한 분야에서 사용될 수 있을 것이다.

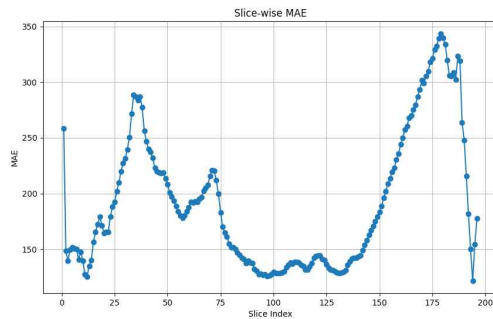


그림 960. 환자 3번의 평균 절대 오차 분포  
Fig. 1. Mean Absolute Error Distribution for Patient 3

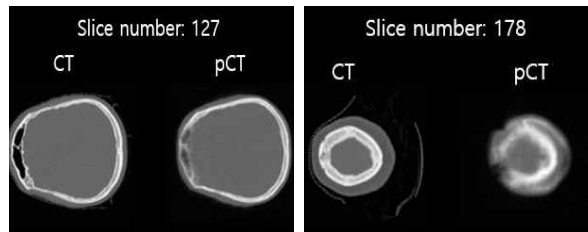


그림 2. CT, pCT의 MAE가 큰 영역(좌)과 작은 영역(우)의 예시

Fig. 2. An example of regions with high MAE (left) and low MAE (right) in CT and pCT slices

#### IV. 결론 및 향후 계획

전통적으로 3차원 CT는 방사선 치료에서 주요 영상 기법으로 사용되어 고해상도의 정교한 기하 구조를 제공하며, 환자의 위치 설정 및 선량 전달 전, 도중, 후의 모니터링에 널리 활용되었다. 그러나 3차원 CT는 방사선 노출로 인해 반복 촬영이 어려운 한계가 있어, MRI만을 사용하는 방사선 치료를 통해 이러한 문제를 줄이고자 인공지능 기반의 pCT가 제안되었다. 인공지능 모델의 구현이 복잡하여 연구자들 사이에서 널리 활용되지 못했으나, 최근 Yaakub의 연구 결과로 pCT 생성 방법이 오픈 소스로 공개되었다. 이에 본 연구팀은 pCT 생성 오픈 소스의 새로운 데이터셋에 대한 적응성을 평가하고자 연구를 수행하였다.

새로운 MRI 데이터로 pCT의 적응성을 평가한 결과, Yaakub의 논문과 비교하여 평균 절대 오차는 머리 전체에서 약 62% 증가했으며, 두개골 영역에서는 유사한 오차를 확인했다. 두개골 영역에서는 유사한 오차가 관측되었지만, 편차가 상대적으로 크게 나타나 범용성을 확인하기 위해 MAE의 최솟값과 최댓값, 그리고 슬라이스별로 시각화했다. 결과적으로, 슬라이스 100~150이 두개골의 구조가 비교적 단순하여 오차가 낮고 고르게 나타나 사용하기에 적합하다는 것을 알게 되었다. 따라서 MRI만을 활용한 방사선 치료를 계획하거나 반복적인 시뮬레이션이 필요한 경우, pCT를 제작하는 오픈 소스를 활용하되 머리의 앞부분과 뒷부분을 제외한, 즉 머리와 배경의 경계가 뚜렷한 중간 부분의 MRI 영역을 사용하면 정확도 높은 두개골 구조를 확인할 수 있을 것이다.

#### 참 고 문 헌

- [1] 김 현중, 김정원, 이수형, 김주연 and 김무영. (2015). 우리나라 종합검진 프로그램의 방사선 노출. *Korean Journal of Health Promotion*, 15(3), 136-140.
- [2] Su P, Guo S, Roys S, Maier F, Bhat H, Melhem ER, et al. Transcranial MR imaging-guided

focused ultrasound interventions using deep learning synthesized CT. *Am J Neuroradiol* 2020;41:1841e8.

- [3] Koh H, Park TY, Chung YA, Lee J, Kim H. Acoustic simulation for transcranial focused ultrasound using GAN-based synthetic CT. *IEEE J Biomed Heal Informatics* 2021;2194:1e11.
- [4] Miscouridou M, Pineda-Pardo JA, Stagg CJ, Treeby BE, Stanziola A. Classical and learned MR to pseudo-CT mappings for accurate transcranial ultrasound simulation. *IEEE Trans Ultrason Ferroelectrics Freq Control* 2022;69:
- [5] Yaakub, Siti N.White, Tristan A.Kerfoot, EricVerhagen, LennartHammers, AlexanderFouragnan, Elsa F. et al. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*, Volume 16, Issue 1, 75 - 78
- [6] Thummerer, A., van der Bijl, E., & Maspero, M. (2023). SynthRAD2023 Grand Challenge validation dataset: synthesizing computed tomography for radiotherapy (1.1) [Data set]. Zenodo.

# LLM-RAG 기반 실시간 개인맞춤 여행지 추천 웹 서비스 설계

정소미\*, 김진현\*\*, 정은미\*\*\*

## Design of a Real-Time Personalized Travel Destination Recommendation Web Service Based on LLM-RAG

So-Mi Jeong\*, Jin-Hyun Kim\*, and Eun-Mi Jung\*\*\*

### 요 약

본 연구는 대규모 언어 모델(LLM)과 RAG(Retrieval-Augmented Generation) 시스템을 활용하여 사용자와 실시간 상호작용을 통해 개인 맞춤형 여행지를 추천하는 웹 서비스를 설계하였다. 제안된 시스템은 단계별 질문 생성, 사용자 응답의 실시간 분석, RAG 기반 추천 엔진, 그리고 상세한 여행지 정보 제공 기능을 포함한다. 이를 통해 기존 추천 시스템이 개인화 수준이 낮고 동적 선호도를 반영하지 못하는 한계를 극복하고, 사용자의 선호도를 정확히 파악하여 최적화된 여행지 추천을 제공한다. 향후 연구에서는 프로토타입 구현과 사용자 평가를 통해 시스템의 효과성을 검증할 예정이다.

### Abstract

This study presents the design of a web-based personalized travel recommendation service that employs Large Language Models (LLMs) in conjunction with a Retrieval-Augmented Generation (RAG) system. The proposed service addresses the limitations of existing recommendation systems by enabling real-time user interactions to accurately capture evolving preferences. The system architecture includes modules for sequential question generation, real-time analysis of user responses, a RAG-based recommendation engine, and detailed travel information provision. By leveraging LLMs and RAG, the service delivers highly personalized travel recommendations and relevant information to enhance user experience and decision-making in travel planning. Future research will focus on implementing a prototype and conducting user evaluations to assess the system's effectiveness and usability.

### Key words

personalized travel recommendation, large language models (LLMs), retrieval-augmented generation (RAG), web service design, user interaction, dynamic preference analysis, artificial intelligence in tourism

---

\* 국립안동대학교 컴퓨터공학과, someee0706@naver.com

\*\* 국립안동대학교 의류학과 wlsqus0325@naver.com

\*\*\* 국립안동대학교 SW융합교육원, enjung@anu.ac.kr(교신저자)

※“본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음”(2019-0-01113)

## 1. 서 론

여행은 삶의 질 향상에 중요한 역할을 하지만, 방대한 정보와 개인의 다양한 선호도로 인해 현대인들은 최적의 여행지를 선택하는 데 어려움을 겪는다. 기존의 여행지 추천 시스템은 개인화 수준이 낮으며, 추천 결과의 다양성이 부족하다는 한계를 지닌다. 대부분의 기존 연구는 협업 필터링(CF)과 내용 기반(CB) 추천 시스템에 초점을 맞추고 있으며, 최근에는 딥러닝을 활용한 하이브리드 추천 시스템이 성능 향상을 입증하고 있다[1]. 그러나 이러한 연구들은 사용자의 과거 데이터에 의존하고 실시간 상호작용을 통한 동적 선호도 파악이 부족하다는 한계가 있다.

본 연구는 이러한 한계를 극복하기 위해 대규모 언어 모델(LLM)과 RAG(Retrieval-Augmented Generation) 시스템을 활용하여 사용자와 실시간으로 상호작용하며 개인 맞춤형 여행지를 추천하는 웹 서비스를 설계하고자 한다. 제안된 시스템은 사용자와의 실시간 대화를 통해 선호도를 정교하게 파악하고, 최적의 여행지와 관련 정보를 제공하는 것을 목표로 한다.

## II. RAG 기반 여행지 추천 웹 서비스 설계

### 2.1 시스템 개요

본 연구에서 제안하는 여행지 추천 웹 서비스는 사용자와의 단계별 상호작용을 통해 개인 맞춤형 여행지를 추천하는 차세대 시스템이다. 사용자는 일련의 선택형 질문에 답변하며, 이를 바탕으로 시스템은 사용자의 선호도를 정교하게 분석하여 최적의 여행지를 추천한다. 또한, 선정된 여행지의 풍경, 액티비티 정보 등 다양한 관련 정보를 제공함으로써 사용자의 의사결정을 효과적으로 지원한다.

그림 1은 본 연구에서 제안된 여행지 추천 웹 서비스의 주요 동작 과정을 시퀀스 다이어그램으로 도식화한 것이다. 사용자가 초기 질문에 응답하면, 질문 생성 모듈(LLM)이 사용자의 응답을 바탕으로 후속 질문을 생성한다. 이 과정이 반복되며 사용자

의 선호도를 점진적으로 파악한다. 수집된 응답은 실시간으로 응답 분석 모듈(LLM)에 전달되어 심층 분석된다. 분석 결과는 추천 엔진(RAG)에 입력되어 개인화된 여행지 추천을 생성하는 데 활용된다. 최종적으로 사용자는 웹 인터페이스를 통해 추천된 여행지의 세부 정보, 이미지 등을 확인할 수 있다.

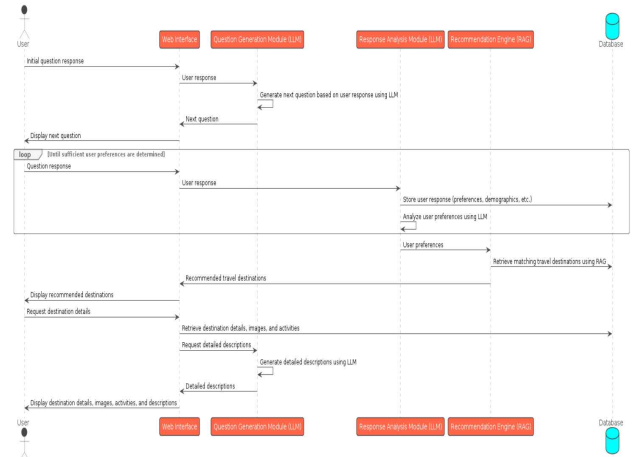


그림 1. 여행지 추천 웹 서비스의 시퀀스 다이어그램  
Fig. 1. Sequence Diagram of the Travel Destination Recommendation Web Service

### 2.2 단계별 질문 생성 및 선택 프로세스

사용자의 여행 성향을 파악하기 위해 LLM을 활용하여 단계별 최적화된 질문을 생성한다. 초기 질문은 사용자의 대략적인 여행 선호도를 파악하기 위해 구성되며, 이후 질문은 사용자의 이전 응답을 바탕으로 동적으로 생성된다. 이 과정에서 GPT-4, Claude와 같은 최신 대규모 언어 모델을 활용하여 세부 선호도를 점진적으로 파악할 수 있도록 설계되어 있다.

질문 생성 알고리즘의 구체적인 과정은 그림 2에 도식화되어 있다. 이 알고리즘은 초기 질문 세트 준비, 사용자 응답 분석, 후속 질문 생성, 그리고 질문 최적화의 네 단계로 구성된다. 초기 질문 세트는 여행 스타일, 예산, 선호 활동 등 기본적인 정보를 수집하는 질문들로 구성된다. 사용자의 응답은 LLM을 통해 분석되어 의미와 감정이 추출된다. 이를 바탕으로 더욱 구체적이고 개인화된 후속 질문들이 생성되며, 최종적으로 관련성이 높고 유용한 정보를 제공할 수 있는 질문들이 선택된다. 이 과정은 반복



적으로 수행되어 사용자의 선호도를 점진적으로 정교화한다.

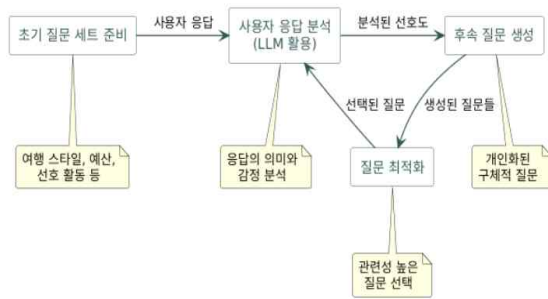


그림 2. 질문 생성 알고리즘 프로세스  
Fig. 2. Question generation algorithm process

### 2.3 사용자 응답의 실시간 분석 및 처리

사용자가 제공한 응답은 고도화된 자연어 처리 기법을 통해 전처리 및 벡터화 과정을 거친 후, 감성 분석과 의미론적 분석을 통해 심층적으로 분석된다. 감성 분석에는 BERT 기반 모델이 활용되어 응답에서 긍정적, 부정적 및 중립적인 감정을 추출한다. 또한, 의미론적 분석은 Word2Vec과 같은 최신 임베딩 기술을 활용하여 사용자 응답에서 주요 키워드와 개념을 추출하고, 이를 사전에 구축된 여행 관련 온톨로지와 매핑하여 사용자의 선호도를 구조화된 형태로 표현한다. 이렇게 도출된 분석 결과는 RAG 시스템의 입력으로 활용되어 개인화된 여행지 추천의 정확도를 크게 향상시킨다.

### 2.4 RAG 시스템 통합을 통한 개인화된 여행지 추천

RAG 시스템은 사용자 선호도와 방대한 여행지 데이터를 효과적으로 매칭하여 최적의 여행지를 추천한다. 이 과정에서 Dense Passage Retrieval(DPR) 기법이 활용되어 관련성 높은 여행지 정보를 효율적으로 검색한다. 검색된 정보는 재순위화 과정을 거쳐 최종 추천 결과로 도출된다. 이러한 과정에 강력한 사전학습 언어모델을 적용함으로써 추천의 정확도와 다양성이 크게 향상된다.

RAG 시스템의 작동 원리는 크게 세 단계로 구분된다. 첫째, 검색 단계에서는 사용자 선호도 벡터를 쿼리로 사용하여 DPR를 통해 관련 여행지 정보를 검색한다. 둘째, 생성 단계에서는 검색된 정보와 사용자 선호도를 입력으로 받아 LLM이 개인화된

여행지 추천문을 생성한다. 마지막으로, 재순위화 단계에서는 생성된 추천문의 관련성과 품질을 종합적으로 평가하여 최종 추천 순서를 결정한다.

### 2.5 추천 여행지 정보제공

추천된 여행지에 대한 상세하고 매력적인 정보를 제공하기 위해 LLM을 활용하여 주요 명소, 액티비티, 음식 등에 대한 설명을 생성한다. 이 과정은 여행지 관련 데이터 수집, 데이터 전처리, LLM 기반 콘텐츠 생성, 그리고 품질 검증의 네 단계로 구성된다. 먼저, 공신력 있는 여행 정보 사이트와 리뷰 플랫폼 등에서 관련 데이터를 수집한다. 수집된 데이터는 정제 및 구조화 과정을 거친 후, LLM에 입력되어 매력적이고 정보가 풍부한 여행지 설명문으로 생성된다. 생성된 콘텐츠는 정확성과 적절성 검증을 거쳐 필요시 수정된다. 또한, 고품질 이미지 및 동영상을 통해 여행지 정보를 시각화하여 사용자의 기대감을 높이고 여행 계획 수립을 효과적으로 지원한다.

## III. 결 론

본 연구는 LLM과 RAG 시스템을 활용하여 실시간 개인 맞춤형 여행지 추천 웹 서비스를 설계하였다. 이 서비스는 사용자와의 실시간 상호작용을 통해 사용자의 선호도를 정밀하게 파악하고, 최적의 여행지와 관련된 정보를 제공함으로써 사용자 경험을 크게 향상시킬 것으로 기대된다. 특히 대규모 언어 모델과 RAG 기반 검색 엔진을 도입하여 시스템 성능을 극대화하고, 사용자 만족도와 여행 계획 수립의 편의성을 크게 증진시킬 수 있을 것으로 전망된다.

그러나 본 연구는 몇 가지 한계를 지니고 있다. 첫째, 실제 사용자 데이터를 충분히 확보하지 못해 시스템 성능 평가가 제한적이었다는 점이다. 둘째, 다양한 언어 및 문화적 배경을 가진 사용자들에 대한 고려가 부족하여 글로벌 사용자층에 대한 적용 가능성에 한계가 있었다. 또한, LLM과 RAG 시스템의 윤리적 사용과 개인정보 보호 문제 또한 추가적인 연구가 필요하다.

향후 연구에서는 이러한 한계를 극복하기 위해 서비스 프로토타입을 구현하고 실제 사용자를 대상

으로 한 평가를 통해 시스템의 효과성과 사용성을 검증할 예정이다. 또한, 다중 언어 지원 및 문화적 차이를 반영한 맞춤형 정보 제공 기능을 개발하여 글로벌 사용자층에 적합한 서비스를 제공할 계획이다. 아울러, 다른 도메인으로 기술 확장을 모색하여 시스템 범용성을 높일 계획이다.

## 참 고 문 헌

- [1] Smith J., Jones A., "Collaborative filtering and content-based recommendation systems: A review", Journal of Information Technology, 2020.
- [2] Brown T., Mann B., Ryder N., et al., "Language models are few-shot learners", Advances in Neural Information Processing Systems, 2020.
- [3] Devlin J., Chang M.-W., Lee K., Toutanova K., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", Proceedings of NAACL-HLT, 2019.

# 지방 거점 도시의 집값 예측을 통한 재테크 시스템

이상기\*, 이세훈\*, 신민혁\*, 김민호\*, 강병권\*\*

## A Financial Technology System Based on the Forecast of House Prices in Local Base Cities

Sang-Gi Lee\*, Se-Hun Lee\*, Min-Hyuk Shin\*, Min-Ho Kim\*, and Byeong-Gwon Kang\*\*

### 요 약

본 연구는 충청남도 천안시 동남구를 대상으로 주택 가격 예측과 재테크 시스템 구축을 목적으로 한다. 2014년부터 2023년까지의 주택 거래 정보와 경제적 요인을 활용해 아파트 매매 예측 모델을 개발하였고, 건축 연도를 기준으로 신축, 준신축, 구축으로 구분하여 예측 성능을 최적화하였다. 다양한 머신러닝 모델 중 CatBoost 모델이 가장 높은 성능을 보였으며, 이를 통해 사회초년생과 신혼부부가 재정적 의사 결정을 지원하는 맞춤형 재테크 시스템을 구상하였다.

### Abstract

This study aims to predict housing prices and develop a financial strategy system focused on Dongnam District, Cheonan City. Using transaction data and economic factors from 2014 to 2023, a housing price prediction model was developed. Apartments were classified into newly built, recently built, and old based on construction year for model optimization. Among various machine learning models, CatBoost showed the highest accuracy, supporting a personalized financial planning system that assists young adults and newlyweds in making informed financial decisions.

### Key words

Real Estate, XGBoost, Regression, CatBoost, Random Forest

## 1. 서 론

최근 몇 년 동안 지방 거점 도시의 주택 시장은 주택 가격 상승과 경제적 불안정성으로 인해 급격한 변화를 겪고 있으며, 이는 사회초년생과 신혼부부와 같은 주거 준비 계층에게 큰 재정적 부담을

주고 있다[1]. 지방 거점 도시는 지역적 경제 요인과 주거 환경이 복합적으로 작용하여 단순한 예측 모델을 적용하기 어려운 특성을 지닌다[2]. 이로 인해 지역별 주택 시장 예측과 맞춤형 재정 계획의 필요성이 더욱 강조된다. 그러나 기존 연구들은 주택 가격 예측에만 초점을 맞추어 개개인의 재정 상

\* 순천향대학교 정보통신공학과 학부과정, lsg7455@naver.com

\*\* 순천향대학교 정보통신공학과 교수, bgkang@sch.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 연구 결과로 수행되었음(2021-0-01399)

황에 맞춘 구체적인 구매 의사 결정을 돕는 기능이 미흡하다. 따라서 본 연구는 지방 거점 도시의 주택 가격 예측을 통해 사회초년생과 신혼부부를 위한 맞춤형 재테크 시스템을 구축하고자 한다. 이를 위해 주택 가격에 영향을 미치는 경제적·환경적 변수를 분석하고, 예측 정확도가 높은 모델을 선정하며, 사용자별 위험 감수 성향과 투자 기간을 고려한 맞춤형 재테크 전략을 도출한다. 이를 통해 지방 거점 도시 주택 시장의 예측 가능성을 높이고, 사용자에게 적합한 재테크 방안을 제시하여 주거와 자산 형성의 부담을 완화할 수 있을 것으로 기대한다.

## II. 데이터셋 구축 및 전처리

### 2.1 데이터셋

본 연구는 충청남도 천안시 동남구의 아파트 실거래가 데이터를 바탕으로 주택 가격 예측을 목표로 한다. 데이터는 주택 거래 정보와 생활 편의성, 경제적 요소를 포함하며, 2014년부터 2023년까지의 자료로 구성된다. 아파트 실거래가는 국토교통부에서 제공하는 데이터를 사용하였고, 주택 금융 환경 반영을 위해 한국은행의 기준금리 데이터를 추가하였다. 교통 및 생활 편의성 평가를 위해 국토교통부의 버스 정류장, 서울교통공사의 지하철 정보, 공공 데이터포털의 공원 및 학교 위치 정보를 수집하고, 지오코드 API를 활용해 각 아파트 단지의 위도와 경도를 산출하였다.

### 2.2 데이터 전처리

본 연구의 데이터 전처리는 아파트 가격 예측 모델의 성능 향상을 위해 연식 계산, 날짜 및 수치형 변환, 면적 구분을 포함하여 진행되었으며, 데이터의 일관성과 모델 학습 최적화를 중점으로 하였다.

먼저, 아파트 건축 연도를 활용하여 연식을 계산하고 이를 '신축(5년 이하)', '준신축(6~10년)', '구축(10년 이상)'으로 구분하여 '건물 구분' 변수로 추가하였다.

다음으로, 전용면적과 거래금액을 수치형 데이터

로 변환하였고, 전용면적은 평형대별 수요를 반영하여 59㎡ 미만을 '59타입', 60~85㎡ 미만을 '84타입', 그 이상을 '기타'로 구분하였다. 이러한 면적 구분을 통해 면적별 선호도를 예측 모델이 학습할 수 있도록 준비하였다.

마지막으로, 계약 연도와 일자를 결합하여 거래 계약일을 날짜형 데이터로 생성하여 시계열 분석의 일관성을 확보하였다.

### 2.3 데이터 시각화

시군구별 평균 거래금액 분석 결과, 지역 간 거래 금액 차이가 뚜렷하게 나타났다. 상위 3개 지역의 평균 거래금액은 표와 같으며, 지역적 특성이 아파트 가격 형성에 중요한 영향을 미침을 확인할 수 있었다.

표 1. 시군구별 평균 거래금액 상위 3개 지역  
Table 1. Top 3 Districts by Average Transaction Price

순위	행정구역	평균 거래금액 (만원)
1	청당동	31,234.76
2	용곡동	27,256.84
3	신부동	22,708.94

상관관계 히트맵 분석을 통해 전용면적, 건축년도, 가장 가까운 역과의 거리가 아파트 가격 형성에 중요한 영향을 미치는 주요 변수임을 확인하였다.

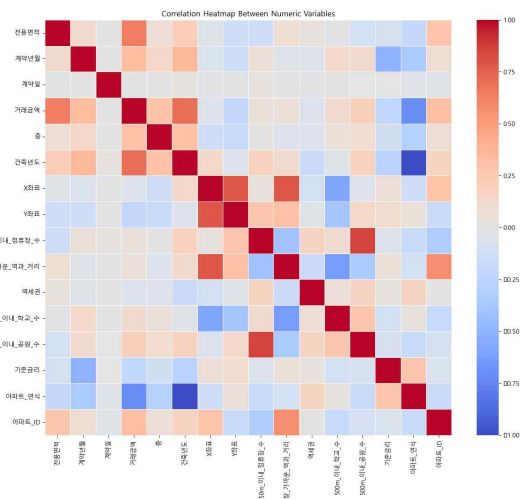


그림 1. 주요 변수 간 상관관계 히트맵  
Fig. 1. Correlation Heatmap of Key Variables

## 2.4 변수 중요도 분석

랜덤 포레스트 모델을 통한 변수 중요도 분석 결과, 전용면적, 건축년도, 연식이 주택 가격 형성에 가장 큰 영향을 미치는 변수로 나타났다. 전용면적은 아파트 크기와 가격 간의 관계를 반영하며, 건축년도와 연식은 신축 선호도를 보여준다. 그 외 계약월, 공원 접근성, 역세권 여부 등의 변수는 거래 금액에 영향을 미치지만 상대적으로 낮은 중요도를 보였다.

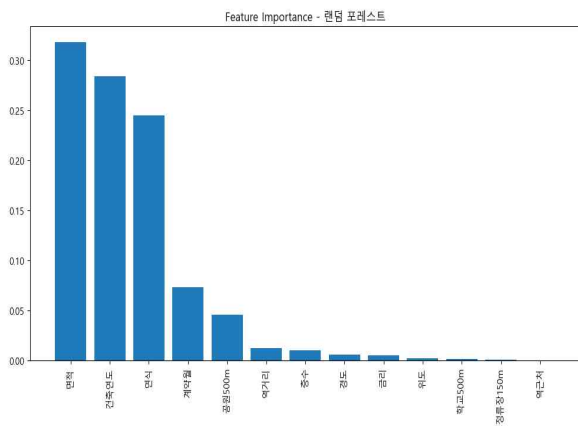


그림 2. 변수 중요도 (랜덤 포레스트)  
Fig. 2. Feature Importance (Random Forest)

### III. 결론 및 향후과제

다양한 머신러닝 모델을 활용하여 주택 가격 예측 성능을 비교한 결과, CatBoost 모델이 가장 우수한 성능을 보였다. 각각의 모델 성능을 RMSE(평균 제곱근 오차)로 비교한 결과, CatBoost와 XGBoost 모델이 상대적으로 낮은 RMSE 값을 기록하여 예측 정확도가 높음을 확인하였다. 아래 표는 주요 모델의 성능을 요약한 것이다.

표 2. 모델별 성능 비교 (RMSE)  
Table 2. Performance Comparison by Model (RMSE)

모델	RMSE
CatBoost	1277.53
XGBoost	1301.44
Random Forest	1365.23
Linear Regression	4091.23

이와 같은 예측 모델을 기반으로 사용자 맞춤형 재테크 시스템을 설계하였다. 이 시스템은 주택 가격 예측, 투자 전략 추천, 주요 변수 영향 분석 기능을 포함하며, 사용자가 UI를 통해 예측된 주택 가격과 관련 변수의 시각적 분석을 손쉽게 확인할 수 있도록 한다. 또한, 투자 목표와 위험 성향에 따른 맞춤형 전략을 제안하며, 거래 시점과 가격 변동에 맞춘 알림 설정 기능을 제공하여 사용자 중심의 자산 관리와 재정 의사결정을 지원한다. 이를 통해 지방 거점 도시 주택 시장에 대한 신뢰성 높은 자산 형성 전략을 제시할 수 있다.

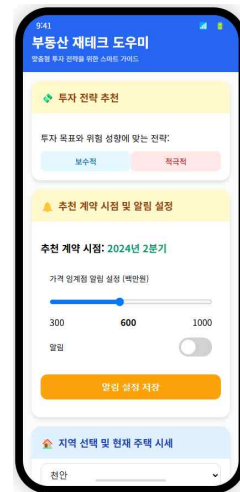


그림 3. 재테크 시스템 사용자 인터페이스(UI)  
Fig. 3. User Interface (UI) of Financial Technology System

### 참 고 문 헌

- [1] 최령, 이준민. (2014). 사회초년생의 주거실태 및 요구에 관한 연구. 한국주거학회 2014 추계학술 발표대회 논문집, 26(2), 97.
- [2] 박정아, 황효정, 송재백. (2024). 거시경제 및 인구특성의 변화가 수도권 및 비수도권 주택 매매시장에 미치는 영향에 관한 연구. 대한부동산학회지, 42(2), 5-23.
- [3] 김학현, 유환규, 오하영. (2023). 딥러닝과 머신러닝을 이용한 아파트 실거래가 예측. 정보처리학회 논문지 (KTSDE), 12(2), 59-76.

# 앙상블 기법을 활용한 향수 추천 알고리즘에 관한 연구

박원규\*, 윤지원\*, 김민호\*, 강병권\*\*

## A Study on Perfume Recommendation Algorithm Using Ensemble Techniques

Won-gyu Park\*, Jiwon Yoon\*, Minho Kim\*, and Byeong-Gwon Kang\*\*

### 요약

본 논문에서는 사용자 맞춤형 향수 추천 시스템을 위해 랜덤 포레스트와 코사인 유사도를 결합한 앙상블 기법을 제안한다. 여러 입력 요소와 향수 데이터를 분석하여 최적의 향수를 추천하며, 랜덤 포레스트로 평점을 예측하고 코사인 유사도로 사용자 입력과 향수 간 유사성을 계산해 정확도를 높였다. 최종적으로 웹 기반 인터페이스를 구축하여 사용자에게 최적의 향수를 추천하는 큐레이팅 서비스를 제공한다.

### Abstract

This paper proposes an ensemble method combining Random Forest and Cosine Similarity to develop a personalized perfume recommendation system. By analyzing multiple input factors and perfume data, it recommends the optimal fragrance. Random Forest predicts perfume ratings, while Cosine Similarity calculates the similarity between user inputs and perfume features to improve accuracy. Finally, a web-based interface is implemented to provide a curating service that recommends the ideal perfume for each user.

### Key words

Perfume, Curation, YOLOv8, mood, Random forest regression, Cosine similarity, Ensemble technique, Web

## 1. 서론

향기가 감정에 미치는 영향은 많은 연구를 통해 입증 되어왔다. 실제로 향은 다양한 분야에서 여러 목적으로 사용된다. 향수는 사용자의 기분을 전환시키거나 자신의 개성을 나타내는 한 가지의 도구가

될 수 있다. 또한 최근에는 개인의 취향에 맞춘 큐레이팅 서비스가 활발하게 발전되고 있다. 본 연구는 큐레이팅 서비스에 개성을 표현할 수 있는 향수를 접목시켜 향수 구매에 어려움을 겪는 사람들에게 도움을 주고자 프로젝트를 진행하게 되었다. 사용자에게 대한 정보 중 기분에 초점을 맞춰 향수를 추천

\* 순천향대학교 정보통신공학과 학부과정 pwg0408@sch.ac.kr

\*\* 순천향대학교 정보통신공학과 교수 bgkang@sch.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2024년도 SW중심대학사업의 연구 결과로 수행되었음 (2021-0-01399)

하는 알고리즘을 개발하였다. 사전 연구로 객체인식과 apriori 연관분석을 활용한 향수 추천 알고리즘을 구현하였는데 [1], 기분에 기반한 향수 추천에 대한 타당성이 부족했다. 이를 해결하기 위해 본 연구에서는 랜덤 포레스트와 코사인 유사도를 결합한 앙상블 기법을 적용하여 향수 추천 알고리즘에 정확도를 높이고자 한다. 또한 사용자 기분과 향수의 특성을 연결하기 위해 사전 설문조사를 통해 기분에 어울리는 향을 매핑하였고 이를 통해 모델의 타당성을 높였다. 사전 연구보다 더 많은 데이터를 사용하여 전처리 과정을 거친 1002개의 향수 데이터를 활용하여 사용자에게 추천하고자 한다.

## II. 본 론

본 연구에서 구현한 향수 추천 프로그램은 사용자로부터 정보를 입력받고 표정 인식을 통해 기분 정보를 수집한다. 랜덤 포레스트와 코사인 유사도를 활용하여 사용자에게 어울리는 향수의 점수를 계산하고 높은 점수를 받은 향수를 추천하는 흐름을 갖는다.

향수 추천을 위해 향수의 노트 정보를 라벨 인코딩(Label Encoder)을 사용하여 범주형 데이터를 수치형 데이터로 변환하여 랜덤 포레스트와 코사인 유사도에 적합한 형태로 전처리한다.

기분을 판단하기 위해 객체인식 모델인 YOLOv8을 사용하여 사용자의 표정을 인식하여 표정을 분석하였으며 640개의 표정 데이터를 학습시켰다.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

학습된 모델의 정확도는 계산식(1)에 따라 61.2%의 표정 인식률을 기록하였다. 이를 통해 행복, 분노, 평온, 슬픔, 놀람의 다섯 가지 표정을 기반으로 기분 정보를 추출한다.

향수 추천을 위한 모델로 랜덤 포레스트 모델을 사용하였다. 랜덤 포레스트란 여러 개의 결정 트리를 학습하여 각 트리의 예측을 종합해 더 정확하고 안정적인 결과를 도출하는 앙상블 학습 알고리즘으로 이를 통해 추천의 정확도를 높이고자 하였다[2].

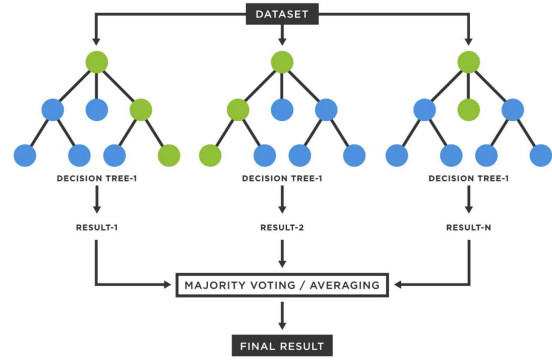


그림 1. 랜덤 포레스트  
Fig. 1. Random Forest

랜덤 포레스트를 활용한 이유는 많은 features를 처리하는 데 강점을 갖고 있기 때문이다. 랜덤 포레스트를 각 향수의 특성을 분석하여 향수의 평점을 예측하는 회귀 모델로 사용하였다. 향수의 전반적인 향, 성별, 베이스 노트, 향의 지속성 등 다양한 특성을 사용하여 랜덤 포레스트 모델을 학습하였고 향수 데이터의 item\_rating을 타깃 변수로 설정하여 각 향수가 사용자에게 받을 평점을 예측하는 데 사용한다.

앙상블 기법으로 코사인 유사도를 결합하였다. 코사인 유사도는 사용자로부터 입력된 성별, 선호하는 향, 원하는 향의 지속성 정도, 카메라를 통해 인식된 기분 정보를 향수의 특징 데이터와 유사도를 비교하여 사용자 선호도를 반영한다. 코사인 유사도는 두 벡터 간의 유사성을 측정하는 방법으로 두 벡터의 방향을 고려하여 유사도를 계산하며(2), 고차원 데이터에서도 높은 성능을 보인다[3].

$$\sim ilarity = \cos(\Theta) = \frac{A \cdot B}{\|A\| \|B\|} \quad (2)$$

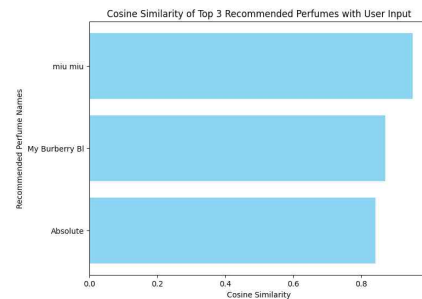


그림 2. 최종 추천 3가지 향수의 코사인 유사도 결과  
Fig. 2. Results of cosine similarity of 3 perfumes final recommendations

그림 2에서 확인할 수 있듯이 사용자의 입력과 각 향수의 특징 간 코사인 유사도가 높은 값을 보인다. 이를 통해 추천된 향수들이 사용자의 선호도와 높은 연관성을 가진다는 것을 의미한다.

본 연구의 핵심은 랜덤 포레스트와 코사인 유사도를 앙상블 하여 최종 추천 점수를 산출하는 방법이다. 랜덤 포레스트로 예측한 향수 점수에 30% 가중치를, 코사인 유사도에 10을 곱하고 70%의 가중치를 부여하여 최종 점수를 계산하였다. 코사인 유사도에 높은 가중치를 부여한 이유는 향수의 특징이 사용자의 선호도 및 정보와 얼마나 연관성이 있는지 더 중점적으로 고려하기 위함이다. 또한 랜덤 포레스트의 타깃 값은 0 ~ 5점 사이의 범위를 가지고 이 데이터에서 코사인 유사도의 범위는 0 ~ 1이다. 이러한 값의 차이를 조정하고 두 모듈의 가중치를 효과적으로 반영하기 위해 정규화를 적용하여 최종 점수를 계산하였다. 따라서 각 향수의 최종 점수는 0점에서 최대 8.5점의 범위를 갖는다.

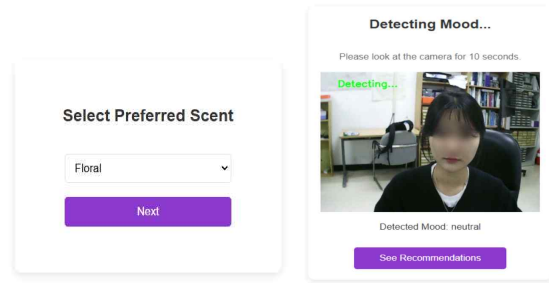


그림 4. 사용자 입력 웹 화면  
Fig. 4. User Input Web Screen

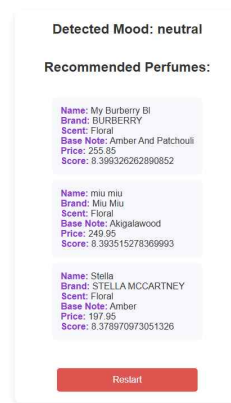


그림 5 향수 추천 결과 웹 화면  
Fig. 5. Perfume recommendation result web screen

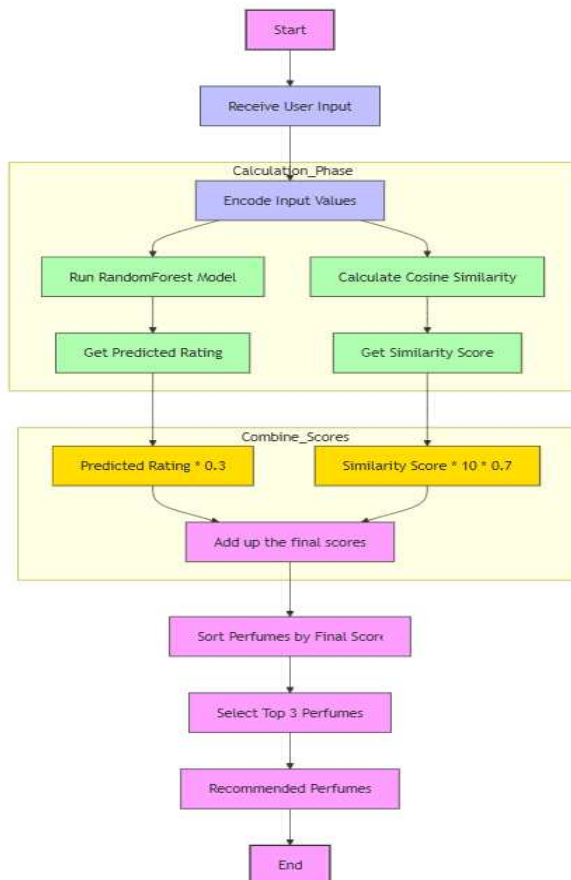


그림 3. 전체 플로우  
Fig. 3. Full Algorithm Flow

### III. 결 론

본 논문에서는 사전 연구보다 향수 추천 시스템의 타당성과 정확성을 높이기 위해 앙상블 기법을 제안하였다. 제안된 방법은 랜덤 포레스트와 코사인 유사도를 결합하여 더욱 고차원적인 데이터에서도 분석 능력을 높였다. 최종적으로 웹 인터페이스를 구성하여 사용자가 보다 쉽게 향수 추천 프로그램을 이용할 수 있게 하였다.

### 참 고 문 헌

- [1] 박원규, et al. "객체 인식을 활용한 향수 추천 알고리즘에 관한 연구." 한국통신학회 학술대회 논문집 (2024): 1199-1200.
- [2] Biau, Gérard, and Erwan Scornet. "A random forest guided tour." Test 25 (2016): 197-227.
- [3] 김성철, et al. "코사인 유사도 기법을 이용한 top-k 관련쌍 검색 방법 조사." 한국정보처리학회 학술대회논문집 24.1 (2017): 808-809.



# 시각장애인 보행 안전을 위한 Few-shot Learning과 Grounding DINO 통합 기반 YOLO 최적화 프레임워크

정소미\*, 모아사랑\*\*, 정은미\*\*\*

## YOLO optimization framework based on integrating Few-shot Learning and Grounding DINO for blind pedestrian safety

So-Mi Jeong\*, A-Sa-Rang Mo\*, and Eun-Mi Jung\*\*\*

### 요 약

본 연구는 시각장애인의 보행 안전을 위한 Few-shot Learning과 Grounding DINO 통합한 YOLO 최적화 프레임워크를 제안한다. 제안된 프레임워크는 제한된 데이터셋에서도 효과적인 객체 탐지 성능을 달성하며, 기존 YOLO 모델의 대규모 레이블링 데이터셋 요구 문제를 해결한다. 실험 결과, 제안된 방법은 기존 YOLO 모델과 유사한 성능을 유지하면서 필요한 레이블 데이터를 크게 줄였고, 다양한 환경 조건에서도 안정적인 성능을 보였다. 본 연구는 시각장애인 보행 안전 향상뿐만 아니라 소규모 데이터셋 환경에서의 객체 탐지 기술 발전에도 기여할 것이다.

### Abstract

This study presents an optimized YOLO framework that integrates Few-shot Learning and Grounding DINO to enhance the walking safety of visually impaired individuals. The proposed framework effectively addresses the challenges associated with limited dataset environments, which are often a requirement for conventional YOLO models that rely on large-scale labeled datasets. Experimental results indicate stable performance across diverse environmental conditions, as well as improved real-time processing capabilities in mobile settings. This research is anticipated to contribute significantly to the enhancement of walking safety for visually impaired individuals, while also advancing object detection methodologies in environments characterized by small datasets across various fields.

### Key words

Few-shot Learning, Grounding DINO, YOLO optimization, visually impaired walking safety, semi-supervised learning, object detection, limited dataset, real-time processing, environmental adaptability, computer vision

---

\* 국립안동대학교 컴퓨터공학과, someee0706@naver.com

\*\* 국립안동대학교 소프트웨어융합학과, software000329@gmail.com

\*\*\* 국립안동대학교 SW융합교육원, emjung@anu.ac.kr(교신저자)

※“본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)

## I. 서 론

시각장애인의 안전한 보행권 보장은 현대 사회의 핵심적인 도전 과제로 부각되고 있다[1]. 최근 통계에 따르면, 국내 시각장애인의 보행 중 교통사고 발생률이 지속적으로 증가하고 있어, 기존의 음향 신호기 중심 보조 시스템만으로는 안전한 보행 환경을 보장하기 어렵다는 점이 드러나고 있다[2].

컴퓨터 비전 기술의 비약적 발전은 이러한 문제에 대한 새로운 해결 가능성을 제시하고 있다. 특히 YOLO(You Only Look Once) 계열의 객체 탐지 모델은 실시간 처리 능력과 높은 정확도를 바탕으로 보행 안전 시스템 구현의 기술적 토대를 마련하였다[2]. YOLO는 단일 단계 접근 방식을 사용하여 전체 이미지를 한 번에 처리함으로써 실시간 객체 탐지를 가능하게 하지만, 이러한 모델들은 대규모의 레이블링된 학습 데이터를 필요로 한다는 근본적인 제약이 존재한다[3]. 특히 시각장애인의 보행 안전을 위해서는 보행자 시점에서의 객체 탐지가 필수적이거나, 현재 공개된 데이터셋의 대부분은 차량 시점에 초점이 맞추어져 있어 실제 적용에 한계가 있다[4].

보행자 시점의 데이터셋 구축에는 다양한 환경 변수들이 고려되어야 한다. 기상 조건과 조도 변화는 물론, 시각장애인의 실제 보행 높이와 시야각에 따른 특수성이 반영되어야 하며, 이는 상당한 시간과 비용을 수반한다[4]. 이러한 맥락에서 제한된 데이터셋으로도 효과적인 학습이 가능한 Few-shot Learning과, 레이블링 비용을 절감할 수 있는 Grounding DINO의 활용이 주목받고 있다.

따라서 본 연구에서는 Few-shot Learning과 Grounding DINO의 상호보완적 특성을 활용한 새로운 준지도학습 프레임워크를 제안하고, 이를 YOLO 모델과 통합함으로써 데이터 수집의 한계를 극복하고 시각장애인의 보행 안전을 위한 실용적이고 고성능의 객체 탐지 시스템을 개발하고자 한다. 제안된 방법론은 소규모 데이터셋 환경에서도 효과적인 객체 탐지가 가능하며, 실시간 처리가 가능한 경량화된 모델 구조를 통해 실제 보행 환경에서의 적용 가능성을 최적화하였다.

## II. 관련 연구

### 2.1 YOLO와 객체 탐지 모델

객체 탐지 분야에서 YOLO(You Only Look Once)는 단일 신경망을 통한 실시간 객체 탐지로 주목받고 있다. YOLOv8은 이전 버전에 비해 정확도와 속도 면에서 개선되었으며, 앵커프리 방식의 도입으로 객체 탐지의 효율성을 크게 향상시켰다. 그러나 최적의 성능 발휘를 위해 클래스당 상당한 수의 레이블링된 이미지가 요구되는 점은 시각장애인 보행 안전과 같은 특수 목적 응용 분야에서 중요한 제약 요인으로 작용한다[5].

### 2.2 Few-shot Learning

Few-shot Learning은 제한된 학습 데이터 환경에서도 효과적인 모델 학습을 가능하게 하는 접근 방식이다. 특히 표준 미세 조정(fine-tuning) 방법은 극히 제한된 데이터 환경(예: 5-way 1-shot 작업)에서도 다른 적응 알고리즘들보다 우수한 성능을 보였다. 그러나 미세 조정 방법은 높은 시간 복잡도를 가진다는 한계가 있어, 실시간 추론 과정에서의 계산 비용 문제는 여전히 해결해야 할 과제로 남아있다[6].

### 2.3 Grounding DINO

Grounding DINO는 텍스트 프롬프트만으로도 객체의 위치를 특정할 수 있는 Vision-Language 모델이다. 이 모델은 카테고리 이름이나 참조 표현과 같은 자연어 입력을 바탕으로 임의의 객체를 탐지할 수 있는 능력을 보여준다. 특히 레이블링되지 않은 데이터에서도 높은 정확도의 객체 탐지가 가능하며, 다양한 도메인에서 우수한 일반화 능력을 입증하였다[7].

## III. Few-shot Learning과 Vision-Language 모델의 통합적 접근을 통한 YOLO 최적화 프레임워크

본 연구에서는 시각장애인의 안전한 보행을 지원하기 위한 객체 탐지 시스템 구현에 있어, Few-shot Learning과 Vision-Language 모델을 통합한 준지도학

습 기반 YOLO 최적화 프레임워크를 제안한다. 그림1은 제안하는 프레임워크의 전체 구조와 데이터 흐름을 보여준다. 그림 1은 보행자 시점의 데이터가 Few-shot Learning과 Grounding DINO를 통해 처리되는 과정, 그리고 이 결과가 YOLO 모델의 학습 최적화에 어떻게 통합되는지를 도식화하고 있다. 특히, 특징 추출과 자동 레이블링 과정이 병렬적으로 이루어지며, 이 결과가 YOLO 모델의 성능 향상에 기여하는 흐름을 명확히 보여준다.

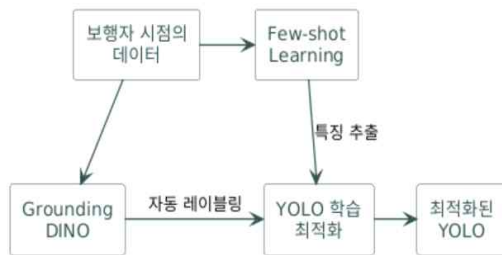


그림 1. 준지도학습 기반 YOLO 최적화 프레임워크  
Fig. 1. A semi-supervised YOLO optimization framework

### 3.1 Few-shot Learning을 통한 보행 환경 특징 추출

보행자 시점의 제한된 데이터 환경에서 효과적인 특징 학습을 위해 Few-shot Learning 기반의 접근 방식을 도입한다. 본 연구에서는 시각장애인의 보행 특성을 고려하여 횡단보도 영역, 신호등 상태, 그리고 장애물 요소를 주요 객체 클래스로 정의하였다. 각 클래스별 특징 추출은 프로토타입 네트워크를 기반으로 수행되며, 이는 공간적 구조, 시각적 속성, 위치 관계 등을 포괄적으로 고려한다.

특히 횡단보도 영역의 경우, 보행 방향성과 횡단 거리를 고려한 기하학적 특징 추출에 중점을 두었으며, 신호등 상태 인식을 위해서는 색상 변화와 시인성에 관한 특징을 중점적으로 추출한다. 장애물 요소의 경우, 보행자와의 상대적 위치와 이동성을 기준으로 한 특징 추출을 수행한다.

### 3.2 Grounding DINO 기반 자동 레이블링

시각장애인 보행 환경에 특화된 자동 레이블링을 위해 Grounding DINO의 Vision-Language 기반 접근 방식을 활용한다. 그림 2는 제안하는 자동 레이블링 프로세스를 보여준다.

그림 2는 입력 이미지가 Grounding DINO 모델을 통해 처리되는 과정과, 생성된 레이블이 다단계 검증을 거쳐 최종 레이블로 확정되는 흐름을 보여준다. 자동 레이블링 과정에서는 보행 환경의 특수성을 반영한 텍스트 프롬프트를 설계하여 활용한다. 예를 들어, 횡단보도 영역 검출을 위해서는 "보행자 진행 방향의 횡단보도 시작점과 종료점"과 같은 구체적인 프롬프트를 사용하며, 신호등 상태 인식을 위해서는 "보행자 신호등의 현재 점등 상태"와 같은 상황 특정적 프롬프트를 활용한다.



그림 2. 자동 레이블링 프로세스  
Fig. 2. Automatic Labeling Process

생성된 레이블의 신뢰성 확보를 위해 다단계 검증 프로세스를 도입하였다. 이 프로세스는 세 가지 주요 단계로 구성된다. 우선, 객체 검출 결과의 신뢰도 점수를 평가하여 임계값 이상의 결과만을 선별하는 필터링 과정을 수행한다. 다음으로, 객체 간의 공간적 관계성을 검증하여 물리적으로 타당한 레이블 조합만을 유지하는 공간 일관성 검사를 실시한다. 마지막으로, 연속된 프레임에서의 시간적 일관성을 확인하여 레이블의 안정성을 보장하는 시간 연속성 분석을 수행한다. 이러한 다층적 접근을 통해 자동 생성된 레이블의 정확성과 신뢰성을 크게 향상시킬 수 있다.

### 3.3 YOLO 모델 통합 최적화

Few-shot Learning과 Grounding DINO의 결과를 통합하여 YOLO 모델을 최적화한다. 본 연구는 시각장애인 보행 안전을 위해 정확도와 실시간성의 균형에 중점을 두었다. 그림 3은 이 통합 최적화 과정을 도식화한다.

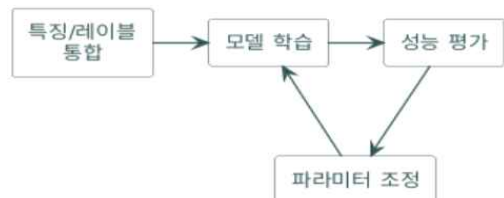


그림 3. YOLO 통합 최적화 프로세스  
Fig. 3. YOLO Integration Optimization Process

그림 3은 Few-shot Learning을 통해 추출된 특징과 Grounding DINO의 레이블링 결과가 YOLO 모델의 학습 과정에 어떻게 통합되는지를 나타낸다. 모델 최적화 과정에서는 신뢰도 기반의 가중치 손실 함수를 도입하여 자동 생성된 레이블의 품질을 학습에 반영한다. 높은 신뢰도를 가진 레이블에는 더 큰 가중치를 부여하고, 상대적으로 낮은 신뢰도의 레이블은 보조적인 학습 신호로 활용한다. 또한, Few-shot Learning을 통해 획득한 특징 정보를 정규화 방향으로 활용하여, 모델이 도메인 특화된 특징을 유지하면서도 일반화 성능을 확보할 수 있도록 하였다.

특히 시각장애인의 보행 환경을 고려하여 다양한 조도 조건에서의 강건성 확보를 위한 최적화 전략을 도입하였다. 배치 정규화 레이어의 통계량을 보행 환경의 특성에 맞게 조정하였는데, 구체적으로는 주간과 야간의 밝기 분포 차이를 고려하여 배치 정규화의 이동 평균과 분산을 동적으로 조정하는 방식을 적용하였다. 이를 통해 날씨나 시간대의 변화에도 안정적인 성능을 유지할 수 있도록 하였다.

#### IV. 실험

##### 4.1 데이터셋

AI Hub에서 제공하는 인도 보행 영상 데이터셋을 기반으로 실험용 데이터셋을 구성하였다. AI Hub의 데이터셋은 다양한 도시 환경에서 수집된 고해상도 보행자 시점 영상으로, 객체 인식을 위한 풍부한 어노테이션을 포함하고 있다. 본 연구에서는 전체 데이터셋 중 36,808장의 이미지를 학습에 사용하였으며, 5,000장의 이미지를 검증에 활용하였다. 특히 시각장애인의 보행 특성을 고려하여 다양한 환경 조건(주간/야간, 맑음/우천)과 보행 상황이 포함되도록 데이터를 구성하였다.

##### 4.2 실험 환경

제안된 프레임워크의 성능 검증을 위한 실험 환경은 표1과 같이 구성하였다.

표 1. 실험 시스템 환경 구성  
Table 1. System Environment Configuration

구분	값
CPU	Intel Xeon 24 Cores / 2.1Ghz
GPU	NVIDIA A100
운영체제	Ubuntu 20.04
프로그래밍 언어	Python

##### 4.3 성능 평가

제안된 방법의 성능을 기존 YOLO 모델과 비교한 실험 결과는 Table 2와 같다. 실험 결과를 통해, 제안된 방법이 기존 YOLO 모델과 비교하여 유사한 수준의 성능(mAP 0.80)을 유지하면서도 필요한 레이블 데이터를 클래스당 평균 150장으로 크게 줄일 수 있음을 보여준다. 또한 42 FPS의 처리 속도와 3.5GB의 메모리 사용량을 달성하여 실시간 처리가 가능한 수준의 성능을 확보하였다. 특히 메모리 사용량의 감소는 실제 모바일 환경에서의 활용 가능성을 높였다는 점에서 의미가 있다.

표 2. Few-shot Learning과 Vision-Language 모델 기반 YOLO 프레임워크의 성능 평가

Table 2. Performance Evaluation of Few-shot-DINO YOLO Framework

지표	YOLO	Few-shot-DINO YOLO
mAP	0.82	0.80
FPS	45	42
메모리(GB)	4.2	3.5
학습 샘플 수*	3,000	150

\*클래스당 필요한 레이블링된 이미지 수

##### 4.4 환경 조건별 성능 분석

시각장애인의 실제 보행 환경은 시간과 날씨에 따라 크게 변화하므로, 다양한 환경 조건에서의 시스템 안정성 검증은 필수적이다. 이를 위해 주/야간 및 날씨 조건별 성능 평가를 수행하였다. 주간 맑은 날씨에서 mAP 0.85를 달성하였으며, 우천 시에도 0.82의 준수한 성능을 유지하였다. 야간의 경우 맑은 날씨에서 0.80, 우천 시 0.78의 mAP를 보여, 열악한 환경 조건에서도 안정적인 객체 탐지가 가능함을 확인하였다. 특히 기존 YOLO 모델과 비교하여 환경 변화에 따른 성능 저하가 평균 20% 감소하였다. 예를 들어, 야간 우천 조건에서 기존 YOLO 모델의 mAP가 0.65로 떨어진 반면, 제안된 모델은 0.78을

유지하여 환경 적응성이 크게 향상되었음을 입증하였다.

## V. 결론

본 연구에서는 시각장애인의 보행 안전을 위한 Few-shot Learning과 Grounding DINO 통합 기반 YOLO 최적화 프레임워크를 제안하였다. Few-shot Learning과 Grounding DINO를 통합적으로 활용한 본 접근 방식은 제한된 데이터셋 환경에서도 효과적인 객체 탐지 성능을 달성하였다. 특히 기존 YOLO 모델이 요구하는 대규모 레이블링된 데이터셋의 한계를 극복하면서도, 실제 보행 환경에서의 적용 가능성을 높였다.

본 연구는 시각장애인을 비롯한 교통 약자들의 안전한 이동을 지원하는 실용적 시스템 구현을 위한 기술적 기반을 제시한다. 또한, 제안된 방법론은 대규모 데이터셋 구축이 어려운 다양한 분야에서 효과적인 객체 탐지 모델 개발에 새로운 패러다임을 제시할 수 있을 것으로 기대된다. 이는 의료 영상 분석, 희귀 생물종 탐지, 산업 현장의 불량품 검출 등 소규모 데이터셋으로 인한 제약이 있는 영역에서 특히 유용할 것이다.

향후 연구에서는 모바일 환경에서의 실시간 처리 최적화를 위한 모델 경량화 기법 개발과, 극한 기상 조건에서의 강건성 향상에 주력할 필요가 있다. 또한, 다중 센서 데이터를 통합적으로 활용하는 멀티모달 접근 방식은 보행 안전성을 더욱 향상시킬 수 있는 잠재력을 지니고 있다. 이러한 연구의 지속적 발전은 궁극적으로 데이터 희소성 문제를 가진 다양한 분야에서의 객체 탐지 기술 향상에 기여하며, 더 나아가 안전하고 포용적인 사회 구현에 실질적으로 이바지할 수 있을 것으로 전망된다.

## 참 고 문 헌

[1] 보건복지부, "2023년 등록장애인 264만 2,000명, 전체 인구 대비 5.1%," 보건복지부 보도자료, 2024. [Online]. Available: [https://www.mohw.go.kr/board.es?act=view&bid=0027&list\\_no=1481120&mid=a](https://www.mohw.go.kr/board.es?act=view&bid=0027&list_no=1481120&mid=a)

10503000000

- [2] International Transport Forum, "Korea: Road Safety Country Profile 2023," 2023. [Online]. Available: <https://www.itf-oecd.org/sites/default/files/korea-road-safety.pdf>
- [3] Ikomia, "Grounding DINO: Revolutionizing AI with Zero-Shot Detection," Ikomia. [Online]. Available: <https://www.ikomia.ai/blog/grounding-dino-zero-shot-detection-explained>
- [4] Viso.ai, "Grounded-SAM Explained: A New Image Segmentation Paradigm?" Viso.ai. [Online]. Available: <https://viso.ai/deep-learning/grounded-sam/>
- [5] Ultralytics, "Is YOLOv8 an anchor-free detector? #3424," GitHub, 2023. [Online]. Available: <https://github.com/ultralytics/ultralytics/issues/3424>
- [6] J. Snell, K. Swersky, and R. Zemel, "Prototypical Networks for Few-shot Learning," in Advances in Neural Information Processing Systems, 2017. [Online]. Available: [https://www.cs.toronto.edu/~zemel/documents/prototypical\\_networks\\_nips\\_2017.pdf](https://www.cs.toronto.edu/~zemel/documents/prototypical_networks_nips_2017.pdf)
- [7] Ikomia, "Grounding DINO: Revolutionizing AI with Zero-Shot Detection," Ikomia Blog, 2023. [Online]. Available: <https://www.ikomia.ai/blog/grounding-dino-zero-shot-detection-explained>



# 장애인 특별교통수단의 대기시간에 영향을 미치는 주요 요인 분석과 최적화 전략

김수인\*†, 최민화\*†, 정서윤\*, 최하은\*, 윤웅창\*\*

## Analysis of key factors affecting waiting time of special transportation for the disabled and optimization strategies

Su-in Kim\*†, Min-hwa Choi\*†, Seo-yun Jung\*, Ha-eun Choi\* and Woong-chang Yoon\*\*

### 요 약

부산시 장애인 특별교통수단인 '두리발' 콜택시는 교통약자의 이동권을 보장하는 핵심 서비스지만 긴 대기 시간 문제로 많은 이용자의 불편을 초래하고 있다. 이에 본 연구에서는 2020년부터 2022년까지 수집한 두리발 운행 데이터를 바탕으로 대기시간을 체계적으로 분석하고, 배차 최적화 방안을 제시한다. 시간대, 월, 휴일 유무, 출발지, 목적지, 주행 거리와 대기시간과의 연관 관계를 분석하였다. 이를 통해 특정 시간대와 공휴일 유무에 따라 대기시간이 현저히 증가하는 문제점을 확인하였다. 이를 해결하기 위한 전략을 제안함으로써 장애인을 위한 교통수단 서비스의 운영 효율성과 이용자 만족도를 향상할 수 있을 것으로 기대된다.

### Abstract

The "Duribal" call taxi service, designed specifically for individuals with disabilities in Busan, plays a pivotal role in ensuring the mobility of vulnerable populations. Nevertheless, lengthy waiting periods have resulted in considerable inconvenience for a considerable number of users. This study presents a systematic analysis of waiting times based on data collected from 2020 to 2022, and proposes methods for optimizing vehicle dispatch. Through an exhaustive analysis of the fluctuations in waiting times contingent on variables such as time of day, month, holiday status, departure points, destinations, and distance, we discerned pivotal periods and presence of holidays where waiting times exhibited a marked increase. By proposing a range of strategies to address these issues, it is anticipated that improvements in operational efficiency and user satisfaction for transportation services catering to individuals with disabilities will be achieved.

### Key words

Duribal, Transportation services, Disabled, Waiting time analysis

---

\* 경상국립대학교 컴퓨터공학부, btj1011@gnu.ac.kr, 220161105@gnu.ac.kr, jsy00319@gmail.com, haeunannie@gmail.com

\*\* 경상국립대학교 컴퓨터공학과/바이오의료빅데이터학과, neobits@gnu.ac.kr

† These authors contributed equally to this work.

## 1. 서 론

부산시 장애인 특별교통수단인 ‘두리발’ 콜택시는 교통약자, 특히 장애인의 이동권 보장을 위한 핵심 서비스로, 사회 복지 측면에서 중요도가 높은 교통수단이다 [1].

이러한 장애인 전용 교통수단은 장애인의 이동권을 보장하기 위한 필수적인 서비스임에도 불구하고 긴 대기시간 문제로 탑승이 원활하지 않아 많은 이용자의 불편을 야기하고 있다 [2].

장애인 특별교통수단의 서비스 최적화와 대기시간 분석에 관한 연구는 교통약자의 이동권 보장을 위한 중요한 연구 분야로 자리 잡고 있다 [3]. 기존 연구에서는 서울시 장애인 콜택시의 이용 특성과 대기시간에 미치는 영향을 분석하며, 빅데이터와 공간 통계 모델을 활용하여 장애인의 이동성에 영향을 미치는 요인을 파악한 사례가 있다 [4]. 또한, 서울시 장애인 콜택시의 대기시간 특성을 다양한 요인에 따라 K-평균 알고리즘을 이용해 분석한 연구도 있다 [5]. 이러한 연구들은 장애인 교통 서비스 개선을 위한 기초자료를 제공하며, 서비스의 효율성을 높이는 데 기여하고 있다.

이에 따라 본 연구는 두리발 서비스의 대기시간 문제를 다차원적으로 분석하고, 데이터에 기반한 차량 배치 및 운영의 최적화 방안을 제안하고자 한다.

## II. 연구 방법

### 2.1 데이터 수집 및 전처리

본 연구에서 사용된 데이터 세트는 2020년 1월부터 2022년 7월까지의 ‘두리발’ 이용 정보를 담은 데이터로, 총 656,984개의 개별 운행 기록과 관련 정보로 이루어져 있다 [6]. 각각의 데이터는 접수일시, 호출지 위·경도, 목적지 위·경도, 배차일시, 승차일시, 요금 등 총 15가지의 다양한 상세 이용 정보를 포함하고 있다.

데이터 전처리 과정에서는 메모, 상세 위치 등과 같은 불필요한 열을 삭제하고 분석의 편리함을 위해서 배차일시, 승차일시 등의 기존 데이터를 기반으로 정보를 얻을 수 있는 배차 대기시간, 승차 대기시간

등과 같은 열을 추가하였다. 또한, 이상치를 제거하고 고유향 추출 후 동일 지역 표기를 통일하였다 (표 1).

표 1. 주요 요인별 분포

Table 1. Distribution by major factors

	mean	med	min	max	std
배차대기 시간 (분)	16.043	9.787	0.002	67.554	16.521
승차대기 시간 (분)	17.601	17.057	0.001	37.007	6.809
총 대기 시간 (분)	33.645	29.406	0.008	84.021	17.742
승차 시간 (분)	21.318	19.350	0.016	51.883	10.487
이동 직선 주행 거리 (km)	9998.594	10005.190	15.519	20015.086	4734.205

### 2.2 탐색적 데이터 분석과 통계적 유의성

데이터 세트에서 대기시간과 가장 연관성이 높을 것이라 예상되는 시간과 장소와 관련된 요인을 선정하여 EDA(Exploratory Data Analysis)를 진행하였다. 분석에는 월, 하루 중 시간대, 공휴일 유무 (평일/공휴일/주말), 출발지, 목적지, 이동 거리를 선정하여 해당 요인과 대기시간에 따른 상관관계를 분석하였다. 각 요인과 대기시간과의 상관관계는 피어슨 상관관계를 통하여 확인하였으며, 통계적 유의성은 ANOVA를 통해서 확인하였다. 통계적 유의수준은  $p\text{-value} < 0.05$ 를 채택하였다.

### 2.3 대기시간 감소를 위한 전략 제시

대기시간 감소를 위한 전략을 수립하기 위해서 대기시간에 영향을 미치는 시간별 및 장소별 주요 요인을 분석하여 상관관계가 큰 요인 중에서 핵심적인 요인을 선정하였다. 이를 바탕으로 대기시간이 짧은 범주의 운행 대수 일부를 대기시간이 긴 범주로 배차함으로써 전체적인 대기시간 감소를 기대할 수 있도록 하였다. 이를 위해서 대기시간이 긴 범주와 짧은 범주와의 비율의 차이를 구하고 이를 최소화하는 방식으로 대안을 제시하여 결과적으로 대기시간의 최솟값과 최댓값이 평균에 가까워지도록 하였다.



### III. 연구 결과

EDA 분석을 위한 요인 탐색을 시행하여 가장 먼저 월별 대기시간의 차이에 대한 분석을 수행하였다. 그 결과 2월과 3월의 대기시간은 30분 미만을 보여, 다른 월에 비해 상대적으로 짧았다. 반면 6월과 7월의 대기시간은 39-42분으로, 이들 월에서는 대기시간이 증가한 것으로 확인되었다 (그림 1).

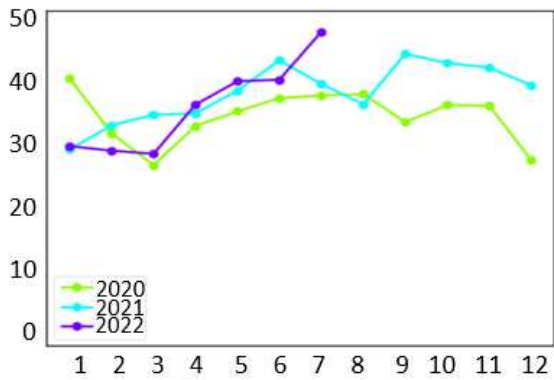


그림 1. 연도에 따른 월별 두리발 탑승 대기시간  
Fig. 1. Monthly waiting times for Duribal by year

시간대와 대기시간과의 관계를 분석한 결과, 오전 5시의 대기시간이 약 72분으로 가장 길었고, 다음은 오전 6시와 오후 9시가 52분으로 나타났다. 대기시간이 약 23분으로 가장 짧았던 오후 8시와 대기시간이 가장 길었던 오전 5시는 49분 정도의 편차가 발생하는 것을 확인하였다 (그림 2).

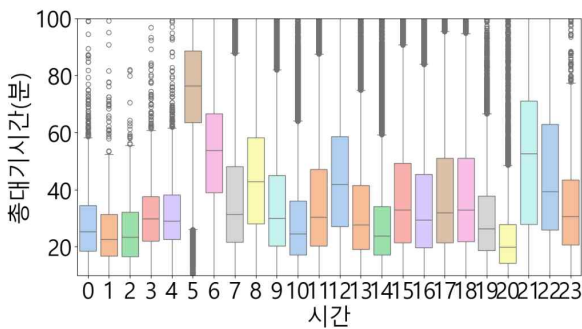


그림 2. 시간대별 두리발 탑승 대기시간  
Fig. 2. Waiting time for Duribal rides by hour

공휴일 유무와 대기시간과의 관계에서는 평일의 두리발 탑승 대기시간은 주말과 공휴일의 대기시간에 비해 1.5배 이상 긴 것으로 나타났다 (그림 3).

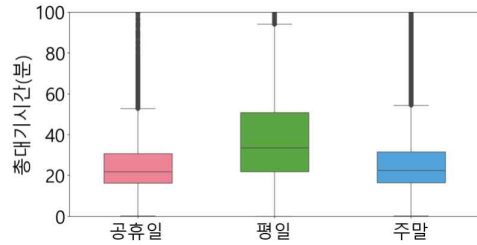


그림 3. 공휴일 유무에 따른 두리발 탑승 대기시간  
Fig. 3. Waiting time for Duribal rides on the presence of holidays

출발지와 대기시간과의 관계를 분석한 결과, 사하구의 대기시간이 약 38분으로 가장 길었고, 북구의 대기시간은 약 30분으로 가장 짧았다. 이에 따라 출발지별 가장 긴 대기시간과 가장 짧은 대기시간 사이의 차이는 약 8분을 나타내었다 (그림 4).

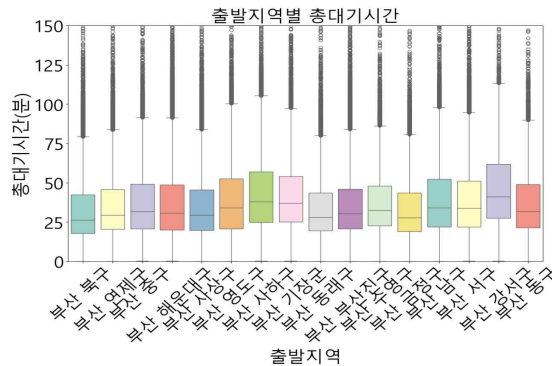


그림 4. 출발지에 따른 두리발 대기시간  
Fig. 4. Waiting time for Duribal rides by departure point

목적지와 대기시간과의 관계를 분석한 결과, 서구에서 대기시간이 약 38분으로 가장 길었고, 북구의 대기시간이 약 31분으로 가장 짧았다. 목적지별 대기시간의 편차는 약 7분으로 나타났다 (그림 5).

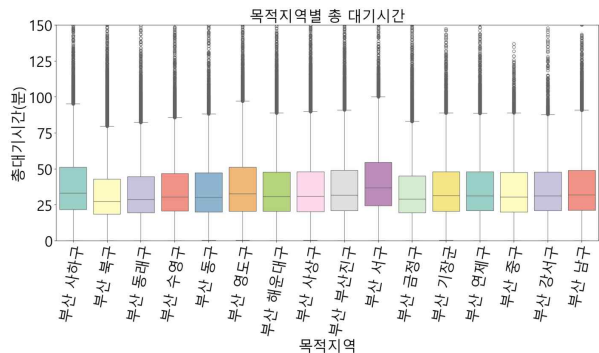


그림 5. 목적지에 따른 두리발 탑승 대기시간  
Fig. 5. Waiting time for Duribal rides by destination

주행 거리와 대기시간과의 관계를 분석한 결과 20km+인 구간에서 대기시간이 약 33분으로 가장 길게 나타났고, 15-18km 구간에서 대기시간이 약 145분으로 가장 짧게 나타났다. 주행 거리별 대기시간의 차이는 약 9분으로, 이는 이동 주행 거리와 대기시간 간 큰 관련이 없음을 의미한다 (그림 6).

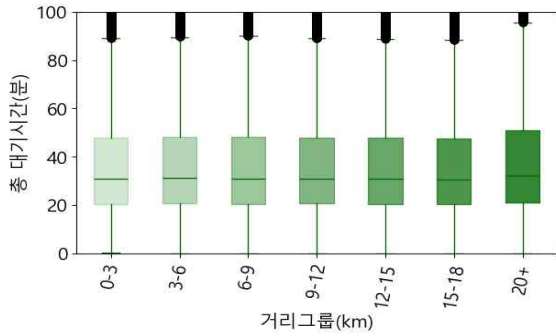


그림 6. 주행 거리별 두리발 탑승 대기시간  
Fig. 6. Waiting time for Duribal rides by travel distance

### 3.1 통계적 유의성 검증과 상관관계 분석

위의 다양한 요인과 대기시간과의 상관관계 분석을 통해서 각 요인 내에서 범주 간의 차이를 확인하였으며, 이러한 차이에 대해서 통계적으로 유의성 검증 수행하였다. 그 결과, 모든 요인에서 p-value가 0.0001보다 낮은 값을 보여서, 대기시간과의 관계가 대한 통계적으로 유의함을 확인하였다 (표 2).

표 2. 요인별 대기시간 차이에 대한 ANOVA 분석  
Table 2. ANOVA analysis of waiting time differences by factor

대기시간 차이	p-value	F-statistic
시간대별	< 0.0001	3545.1
월별	< 0.0001	1522.7
공휴일 유무별	< 0.0001	3991.1
출발지별	< 0.0001	816.6
목적지별	< 0.0001	135.2
주행 거리별	< 0.0001	26.7

각 요인과 대기시간 간의 상관관계는 피어슨 상관계수 분석을 통해서 확인하였다. 그 결과 공휴일 유무와 대기시간 간의 상관계수는 0.988로 강한 큰 것을 확인하였으며, 다음으로는 시간대와 대기시간 간의 상관계수가 -0.551인 것을 확인하였다 (표 3).

표 3. 대기시간과 관련 요인 간의 상관계수  
Table 3. Correlation coefficient between waiting time and related factors

대기시간과의 상관 요인	상관계수
공휴일 유무별	0.988
시간대별 운행 대수	-0.551
목적지	0.212
출발지	-0.196
월별 이용 건수	0.067
주행 거리	0.019

### 3.2 최적화 결과 및 전략 제안

각 요인과 대기시간 간의 상관관계 분석 결과, 공휴일 유무와 시간대의 운행 대수가 가장 큰 관계를 보이는 것을 확인하였다 (표 3). 때문에, 2가지 요인에 대한 운행 대수 조정을 통하여 두리발 운행의 효율성을 높이기 위한 최적화 전략을 다음과 같이 제시할 수 있다 (표 4).

표 4. 최적화 전략 적용 전후 두리발 평균 운행 대수  
Table 4. Average number of Duribal before and after the optimization strategy is applied

시간	적용 전		적용 후	
	평일	공휴일/주말	평일	공휴일/주말
5시	2	2	5	2
7시	53	24	58	19
8시	18	16	21	13
9시	38	36	45	29
13시	47	35	50	28
21시	3	3	4	2
22시-익일 5시	2	2	2	2
Total	163	118	185	95

첫째, 대기시간이 비교적 짧은 시간대에 운영되는 차량을 대기시간이 긴 시간대로 조정하는 방안을 수립하였다. 오전 5시와 오후 8시의 대기시간을 조정하는 것이 가장 큰 효과를 기대할 수 있다. 대기시간은 운행 대수에 반비례하므로, 현재 오전 5시 (2대, 72분)와 오후 8시 (41대, 23분)의 대기시간 비율을 바탕으로 조정할 운행 대수의 비율을 구하고, 이를 바탕으로 오후 8시에 운행되는 차량의 약 6%를 오전 5시에 배

치함으로써 대기시간을 효과적으로 단축할 수 있을 것이다.

$$\frac{n_{\text{오전5시}} + xn_{\text{오후8시}}}{(1-x)n_{\text{오후8시}}} = \frac{72}{23} \quad (1)$$

둘째, 공휴일과 주말 운행 차량 일부를 평일에 배치함으로써 한정된 운행 대수 내에서 평일의 긴 대기시간을 효과적으로 감소시킬 수 있다. 대기시간은 운행 대수에 반비례하고, 평일 대기시간이 주말보다 1.5 배 길다는 점을 고려하여 운행 대수를 조정하는 것이 가장 효율적이다.

$$\frac{n_{\text{평일}} + xn_{\text{공휴일/주말}}}{(1-x)n_{\text{공휴일/주말}}} = 1.5 \quad (2)$$

때문에, 주말 운행 대수의 약 19%를 평일로 전환하면 평일 대기시간을 효율적으로 줄일 수 있을 것이다.

이러한 조정을 통해 대기시간이 짧은 범주의 배차를 대기시간에 긴 범주로 이동하여 고객의 대기시간을 효과적으로 감소시킬 수 있을 것이다 (표 4).

#### IV. 결 론

본 연구에서는 부산시의 장애인 특별교통수단인 '두리발'의 긴 배차 대기시간 문제를 개선하기 위해 대기시간을 체계적으로 분석하고 배차 최적화 방안을 제시하였다.

EDA 분석을 통해 시간대, 월, 공휴일 유무, 출발지, 목적지, 주행 거리와 대기시간 간의 상관관계가 통계적으로 유의미한 차이가 있음을 확인하였으며, 특히 평일, 새벽, 자정 등 특정 시간대의 대기시간이 매우 긴 경향성을 확인하였다.

이에 따라, 대기시간이 짧은 공휴일과 주말의 운행 차량을 평일로 전환하고, 대기시간이 짧은 시간대의 차량 일부를 대기시간이 긴 시간대로 조정하는 방안을 제안하였다. 이러한 전략을 통해 대기시간의 편차를 감소시켜 서비스 품질을 향상과 함께 고객 만족도를 또한 높일 수 있을 것으로 기대된다.

이번 연구를 통해서 두리발의 운행에 시간 및 장소에 관련된 요인들이 대기시간에 미치는 영향을 확인할 수 있었으며, 본 연구에서 다루지 못한 다양한 요인에 대한 추가 분석을 통해서 향상된 최적화 방안을 향후 추가적인 연구를 통해 제시할 수 있을 것이다.

결론적으로 본 연구를 통해 장애인들의 운송 수단에 관련된 문제를 해결함으로써 이동 편리성을 증진하고, 다른 교통수단의 최적화 연구에도 기여할 수 있을 것으로 기대한다.

#### 참 고 문 헌

- [1] 부산광역시, “제4차 부산광역시 교통약자 이동 편의 증진계획(2022-2026)”, 2022.09, <https://www.aurum.re.kr/Legal/LegalSub.aspx?pcode=B10>.
- [2] 정현영 and 박기준, "순서형 로짓 모델을 이용한 두리발 이용자의 일반택시로의 수단전환에 관한 연구", 대한교통학회지, Vol. 31, No. 5, pp. 79-88, 2013.
- [3] 박태윤 and 박경아, "교통약자 이동지원 서비스의 사회적 가치 추정 연구: 장애인콜택시의 통행시간 질적 수준 개선을 중심으로", 교통연구, Vol. 30, No. 4, pp. 1-13, 2023.
- [4] 이지민 and 이수기, "서울시 장애인 콜택시 이용특성 및 대기시간 영향 요인 분석 : 장애인 콜택시 빅데이터와 Community Detection을 활용하여", 국토계획, Vol. 58, No. 1, pp. 75-90, 2023.
- [5] 김홍균 and 김도경, "K-평균 알고리즘 기반의 클러스터링을 통한 장애인 콜택시 대기시간 특성분석", 교통기술과정책, Vol. 19, No. 3, pp. 52-56, 2022.
- [6] 서창갑 and 정대현, "사회적가치 창출 관점의 장애인 콜택시 운영 최적화를 위한 빅데이터 분석", 인터넷전자상거래연구, Vol. 22, No. 6, pp. 145-154, 2022.

# 지역 활성화를 위한 지역 챗봇 시스템

정현섭\*, 이세민\*\*, 손민섭\*\*, 정용성\*\*\*, 김진우\*\*\*, 김희선\*\*\*

## Local Chatbot System for Local Revitalization

Hyeonseop Jeong\*, Semin Lee\*\*, Minseop Son\*\*, Yongsung Jung\*\*\*, Jinwoo Kim\*\*\*,  
and Heesum Kim\*\*\*

### 요약

오늘날 지방의 지역들은 인구 유출 및 고령화로 인해 경제와 문화가 쇠퇴하고 있다. 이러한 문제를 해결하기 위해서는 지역의 특색을 살린 자원과 매력을 외부에 알리고, 이를 기반으로 지역 경제와 관광 산업을 활성화하는 것이 중요하다. 본 시스템은 지역 특산물, 관광 명소 그리고 지역에 연관된 해당 위인을 통해 정보를 사용자에게 손쉽게 제공하여, 지역의 매력을 널리 알린다. 더 나아가 지역 경제 활성화와 관광 촉진에 기여할 수 있는 솔루션을 제안하고자 한다.

### Abstract

Today, local areas are experiencing economic and cultural decline due to population outflow and aging. In order to solve these problems, it is important to promote the unique resources and charm of the region to the outside world and revitalize the local economy and tourism industry based on this. This system promotes the charm of the region by easily providing users with information about local specialties, tourist attractions, and famous people associated with the region. Furthermore, we would like to propose solutions that can contribute to revitalizing the local economy and promoting tourism.

### Key words

local revitalization, population decline, regional chatbot, social issue solution, local resource promotion

## 1. 서론

최근 한국의 지방 지역은 인구 감소와 고령화, 경제적 쇠퇴로 인해 심각한 사회적 문제를 겪고 있다[1]. 이러한 인구 유출과 경제적 어려움은 지방

소멸이라는 위기를 불러일으키며, 이는 단순히 지역의 문제를 넘어 국가 전체의 경제와 문화적 다양성에도 영향을 미치고 있다. 지방의 인구가 줄어들면서 점차 경제 활동이 감소하게 되고, 이로 인해 지역 상권이 침체되고 일자리 부족 현상이 심화되고

\* 국립안동대학교 멀티미디어공학과, jhs789654123@gmail.com

\*\* 국립안동대학교 AI융합학과, dltpls7498@gmail.com, spcle2412@gmail.com

\*\*\* 국립안동대학교 소프트웨어융합학과, pororojys@naver.com, rlawlsdn2797@gmail.com, hskim@anu.ac.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01113)

있다.

이러한 문제를 해결하기 위해서는 지역의 매력과 자원을 외부에 널리 알리고, 지역 경제와 관광 산업을 활성화하는 노력이 필요하다. 지역마다 고유한 특산물, 역사적 인물, 그리고 관광 명소 등이 존재하며, 이를 효과적으로 홍보함으로써 지방을 다시 매력적인 장소로 만들 수 있다. 그러나 많은 지방들이 끊임없는 노력에도 불구하고 다양한 어려움을 직면하고 있는 상황이다[2].

본 논문에서는 이러한 문제를 해결하기 위해 '지역 활성화를 위한 지역 챗봇 시스템'을 제안한다. 이 시스템은 지방의 특산물, 유명 인물, 관광 명소 등 다양한 정보를 사용자에게 제공하여 지역의 매력을 널리 알리고, 더 나아가 지역 경제와 관광 산업의 활성화를 일으킨다. 특히, 인공지능 기반의 챗봇을 통해 사용자에게 맞춤형 정보를 제공함으로써 사용자가 손쉽게 지역 정보를 얻을 수 있도록 돕는다. 이러한 정보는 지역 사회에 대한 이해를 높이고, 지역 방문을 유도함으로써 지방 경제에 긍정적인 영향을 미칠 수 있을 것이다[3].

## II. 본 론

본 연구에서는 지역 활성화를 위한 챗봇 시스템을 설계하고 구현하는 과정에서 다양한 기술적 요소를 사용한다. 특히 OpenAI의 GPT 기반 언어 모델과 RAG (Retrieval-Augmented Generation) 개념을 활용하여 챗봇이 사용자의 질문을 정확하고 관련성 높은 답변을 제공할 수 있도록 한다.

RAG는 대형 언어 모델(LLM)과 정보 검색 기능을 결합한 기술로, 챗봇이 사용자 질문에 대해 실시간으로 관련된 정보를 검색하고 이를 바탕으로 응답을 생성할 수 있게 한다. 예를 들어, 사용자가 특정 지역의 역사적 배경이나 관광 명소에 대해 선택하고 질문하면, 챗봇은 미리 업로드된 관련 문서들에서 해당 정보를 검색하고, 이를 기반으로 답변을 제공한다[그림1]. 이러한 방식은 단순히 언어 모델의 학습된 데이터를 기반으로 답변을 생성하는 것보다 더 정확하고 신뢰할 수 있는 정보를 제공하는 데 도움이 된다.

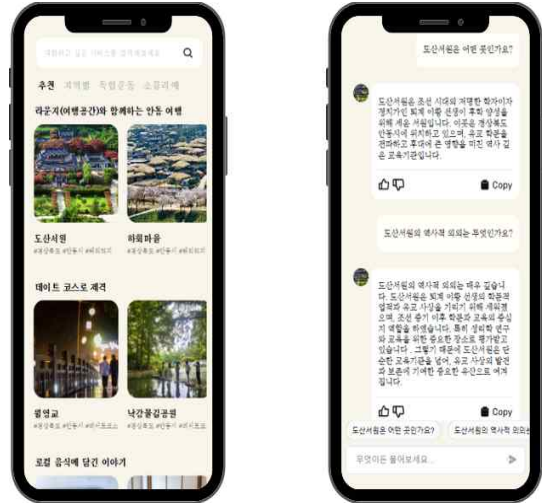


그림 1. 관리자 페이지 - 어시스턴트 등록  
Fig. 1. Administrator Page - Assistant Registration

또한, 본 연구에서는 챗봇 시스템의 관리 기능을 강화하기 위해 관리자 페이지를 구축한다[그림2]. 이를 통해 새로운 지역이나 주제에 대한 챗봇을 손쉽게 추가하고 관리할 수 있도록 하였으며, 각 챗봇은 고유한 어시스턴트 ID와 문서 ID를 활용하여 특정 주제나 지역에 대한 정보 검색 및 답변을 제공할 수 있도록 설계한다. 어시스턴트 ID와 문서 ID를 활용한 구조는 챗봇이 다양한 주제를 효과적으로 다룰 수 있도록 하는 핵심 요소로, 각 주제에 맞는 정보를 정확하게 제공하는 데 큰 역할을 한다.

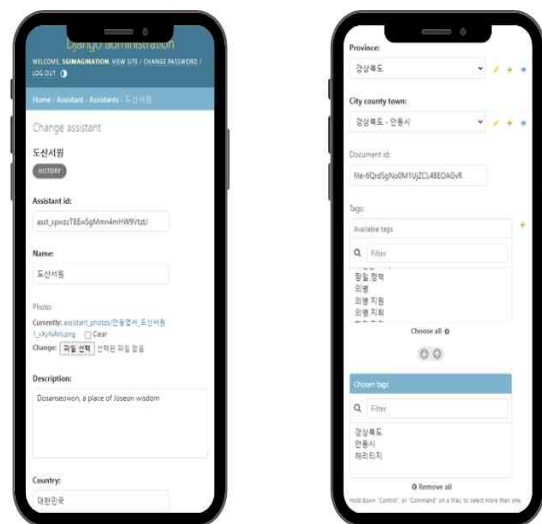


그림 2. 관리자 페이지 - 어시스턴트 등록  
Fig. 2. Administrator Page - Assistant Registration

챗봇 시스템은 Django 프레임워크와 OpenAI API를 사용하여 개발되었으며, 사용자가 웹 페이지에서 직접 챗봇과 대화할 수 있는 UI를 제공한다. 사용자 인터페이스는 질문 입력창과 답변 영역으로 구성되어 있으며, 사용자는 원하는 질문을 입력하고 그에 대한 답변을 실시간으로 받을 수 있다. 또한, 챗봇의 답변은 자연스러운 대화 형식으로 제공되며, 사용자는 이전 대화의 맥락과 상관없이 독립적인 질문을 반복해서 할 수 있도록 설계되었다. 이를 통해 사용자는 더욱 편리하고 유연한 대화 경험을 누릴 수 있다.

### III. 결 론

본 논문에서는 지역 활성화를 위한 챗봇 시스템의 개발과 적용 가능성에 대해 연구한다. 인공지능 기반의 챗봇은 사용자가 지역의 특산물, 역사적 인물, 관광 명소 등 다양한 정보를 손쉽게 얻을 수 있도록 돕는 중요한 도구로 활용될 수 있다. 이를 통해 지역의 매력을 널리 알리고, 관광객 유치와 지역 경제 활성화에 기여할 수 있을 것이다.

본 연구에서 개발한 챗봇 시스템은 사용자 맞춤형 정보를 제공하기 위해 RAG 개념을 도입하여 LLM과 정보 검색을 융합한다. 이를 통해 챗봇은 단순한 정보 제공을 넘어, 사용자에게 신뢰할 수 있는 맞춤형 답변을 제공할 수 있게 된다. 또한, 관리자 페이지를 통해 새로운 주제의 챗봇을 손쉽게 추가하고 관리할 수 있는 기능을 구현하여, 시스템의 확장성과 유지보수성을 높였다.

앞으로의 연구에서는 챗봇 시스템의 성능을 더욱 개선하고, 다양한 지역과 주제에 대한 챗봇을 포함으로써 더 많은 사용자에게 유익한 정보를 제공할 수 있도록 할 것이다. 또한, 사용자 피드백을 기반으로 챗봇의 대화 품질을 지속적으로 개선하여 사용자 경험을 향상시키는 것을 목표로 하고 있다.

### 참 고 문 헌

[1] Park Eun Ji, "An In-Depth Analysis of the Latest Artificial Intelligence Music Generators", The

Korean Journal of Arts Studies, No. 44, pp. 5-30, 2024.

[2] Young-Gil Kwon (2020). Searching for the Policy Alternatives for the Activation of the Local Culture Industry in Korea. Journal of the Korea Entertainment Industry Association, 14(7), 175-191, 10.21184/jkeia.2020.10.14.7.175

[3] Dongsu Lim and Soongu Lee. (2022). A study on the impact of mobile chatbot quality in the tourism industry on social presence, trust in suppliers, and intention to continue using: Focusing on personal innovation and anthropomorphism control variables. Journal of Business Management, 29(2), 207-235.

# 세그멘테이션을 활용한 초상화 드로잉 로봇 구현

최윤호\*, 장민우\*, 박주환\*\*, 옥승호\*\*\*

## Implementation of a Portrait Drawing Robot Using Segmentation

Yun-Ho Choi\*, Min-Woo Jang\*, Joo-Hwan Park\*\* and Seung-Ho Ok\*\*\*

### 요 약

최근 AI 및 매니플레이터를 활용한 드로잉 연구가 활발히 진행되고 있다. 초상화 드로잉에 주로 사용되는 그레이스케일 방식은 미세한 노이즈로 인해 얼굴 특징을 표현하는 드로잉 경로가 과도하게 생성되어 불필요한 동작이 발생하는 문제점이 있다. 본 논문에서는 세그멘테이션을 적용하여 불필요한 노이즈를 제거한 초상화 드로잉 방식을 제안한다. 학습된 세그멘테이션 모델을 통해 불필요한 선을 제거하고 윤곽선이 선명한 이미지로 변환한 후 이로부터 윤곽선을 추출하여 매니플레이터로 초상화를 그린다. 실험 결과 기존 그레이스케일 방식 대비 노이즈 드로잉 선분이 약 83.48% 감소하였으며, 매니플레이터가 얼굴 특징을 정확하게 표현함을 확인하였다.

### Abstract

Recent research on drawing techniques using AI and manipulators has been actively conducted. The grayscale method, commonly used for portrait drawing, has a drawback in that excessive drawing paths are generated due to fine noise, leading to unnecessary movements. In this paper, we propose a portrait drawing method that reduces unnecessary noise by applying segmentation. By using a trained segmentation model, unnecessary lines are removed, and the image is converted to a sharper contour outline. The manipulator then uses this refined outline to draw the portrait. Experimental results show that noise in drawing line segments was reduced by approximately 83.48% compared to the conventional grayscale method, and it was confirmed that the manipulator accurately represented facial features.

### Key words

portrait drawing, manipulator, segmentation, edge detection

---

\* 동의대학교 로봇자동화공학과, kjn8026@office.deu.ac.kr, 20192116@office.deu.ac.kr

\*\* 동의대학교 인공지능학과, 20246146@office.deu.ac.kr

\*\*\* 동의대학교 로봇자동화공학과, osh@deu.ac.kr(교신저자)

※ 이 논문은 교육부의 지방대학활성화사업 사업비를 지원받아 작성되었음

※ 본 논문(저서)는 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임

## 1. 서 론

최근 AI 및 매니플레이터를 활용한 드로잉 연구가 활발히 진행되고 있다[1-3]. 일반적인 초상화 드로잉 로봇은 입력된 얼굴 이미지를 그레이스케일로 변환하고 이진화를 통해 노이즈를 제거한 후 윤곽선을 추출하여 초상화를 그리는 방식을 사용한다 [4]. 그러나 이러한 방식은 머리카락과 눈썹 등에서 발생하는 노이즈로 인해 얼굴 특징을 표현하는 드로잉 경로가 과도하게 생성되어 불필요한 로봇의 모션이 발생하는 문제점이 있다. 이에 본 논문에서는 세그멘테이션을 활용하여 불필요한 로봇 드로잉 경로를 최소화한 얼굴 이미지를 생성하고 이를 엣지 검출한 뒤 초상화를 그리는 드로잉 방식을 제안한다.

## II. 본 론

그림 1은 본 논문에서 제시하는 세그멘테이션 초상화 드로잉 흐름도이다. 원본 이미지에서 얼굴 부분만 크롭(crop)한 후 배경을 제거한다. 이후 세그멘테이션을 사용하여 얼굴 윤곽, 입술 및 눈썹의 윤곽선을 선명하게 한 후 머리카락 영역을 검은색으로 변환하여 불필요한 드로잉 경로를 제거하였다. 세그멘테이션 모델을 사용하기 위해 공개 데이터 셋인 CelebAMask-HQ을 사용하여 얼굴 구성요소에 대한 세그멘테이션 데이터를 학습시켰다. 세그멘테이션을 거친 이미지에서 엣지 검출을 한 후 매니플레이터가 드로잉 작업을 수행한다.

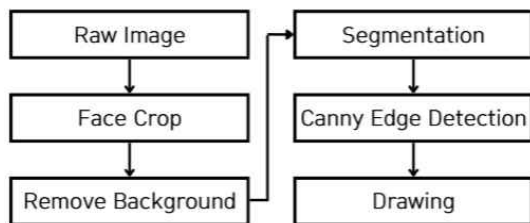


그림 1. 초상화 드로잉 흐름도  
Fig. 1. Portrait drawing flowchart

그림 2는 기존 방식과 제안 방식의 윤곽선 검출 결과를 비교한 것이다. 그림 2(a)는 그레이스케일

변환 후 임계값 조절을 통해 윤곽선을 검출한 결과이며 그림 2(b)는 제안하는 세그멘테이션 기반 윤곽선 검출 결과이다. 두 결과를 비교 분석한 결과, 세그멘테이션을 적용한 경우 드로잉 선분의 개수가 811개에서 134개로 약 83.48% 감소하였다. 이를 통해 제안하는 방식이 불필요한 드로잉 경로 제거와 윤곽선 추출에서 개선 효과가 있음을 확인하였다.

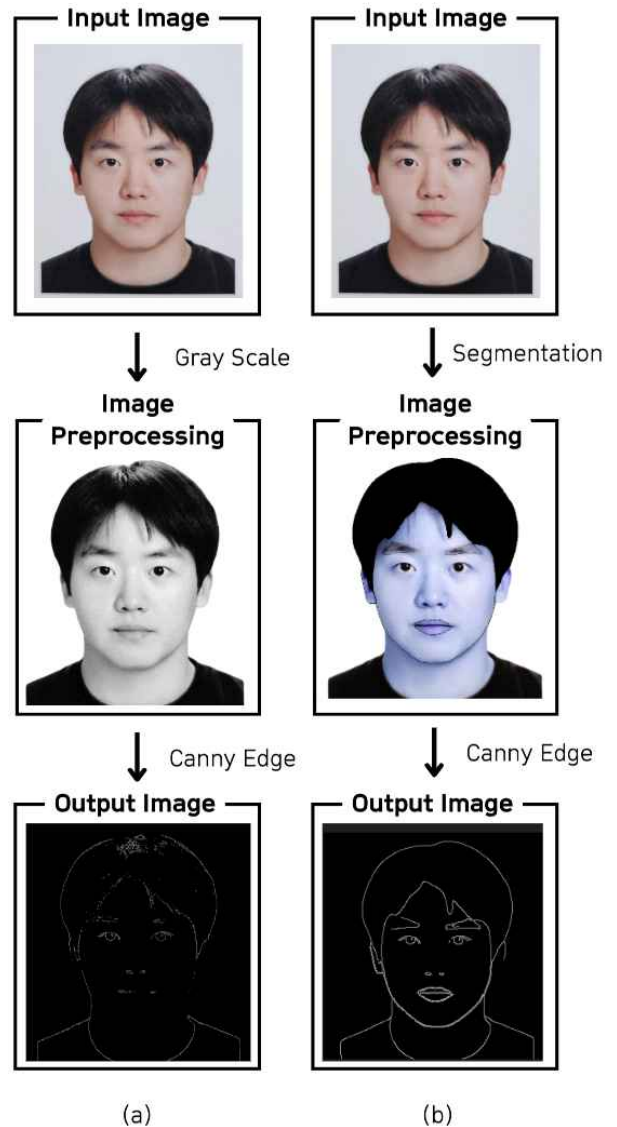


그림 2. 이미지 전처리 결과 비교 (a) 그레이스케일 적용 (b) 세그멘테이션 적용

Fig. 2. Comparison of image preprocessing results: (a) application of grayscale (b) application of segmentation

본 논문에서는 그림 3과 같이 ZEUS사의 6축 매니플레이터인 ZRA-503P를 사용하였으며



엔드이펙터에 붓펜을 설치하여 A4용지 위에 드로잉하였다. 그림 4는 세그멘테이션을 적용한 최종 드로잉 결과이다.

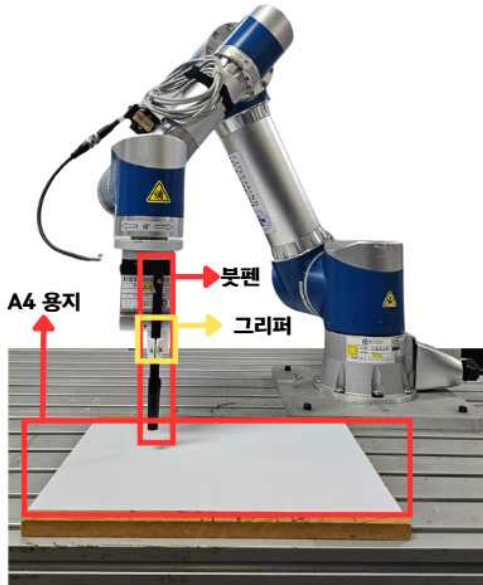


그림 3. 드로잉 환경  
Fig. 3. Drawing environment

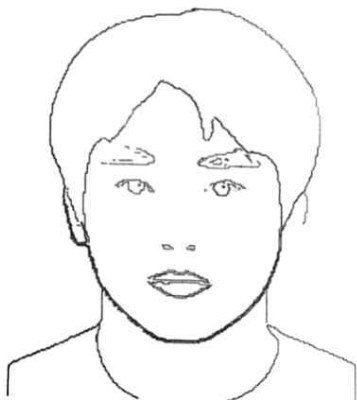


그림 4. 드로잉 결과물  
Fig. 4. Drawing output

### III. 결 론

본 논문에서는 세그멘테이션을 적용하여 얼굴 특징 드로잉 경로에 발생된 불필요한 노이즈 선분을 제거한 후 초상화를 그리는 드로잉 로봇을 제안하였다. 제안하는 방법은 사용자가 학습시킨 세그멘테이션 모델을 통해 드로잉 선분이 적은 이미지로 변환한다. 변환된 이미지에서 윤곽선 추출 후 매니퓰레이터를 사용하여 초상화를 그린다. 실험 결과 그

레이스케일 방식에 비해 드로잉 선분의 개수가 약 83.48% 감소하였으며, 매니퓰레이터가 입력된 얼굴 이미지의 특징을 살려 그려내는 것을 확인하였다.

### 참 고 문 헌

- [1] P. Schaldenbrand, J. McCann and J. Oh, "FRIDA: A Collaborative Robot Painter with a Differentiable, Real2Sim2Real Planning Environment", IEEE International Conference on Robotics and Automation, London, United Kingdom, pp. 11712-11718, 2023.
- [2] Y. Park, S. Jeon and T. Lee, "Robot Learning to Paint from Demonstrations", IEEE/RSJ International Conference on Intelligent Robots and Systems, Kyoto, Japan, pp. 3053-3060, 2022.
- [3] P. Schaldenbrand, G. Parmar, J. Y. Zhu, J. McCann and J. Oh, "CoFRIDA: Self-Supervised Fine-Tuning for Human-Robot Co-Painting", IEEE International Conference on Robotics and Automation, Yokohama, Japan, pp. 2296-2302, 2024.
- [4] J. K. Kim, S. P. Lee, H. L. Jung, and H. K. Cho, "Development of a Dual-Arm Drawing Robot using Line Segment Approximation of Image Edges", The Journal of Korea Robotics Society, vol. 9, no. 3. The Korea Robotics Society, pp. 140-146, 2014.

# ATP 테니스 데이터 기반의 랭킹 시스템 비교 분석

김상희\*, 이주찬\*, 김진수\*\*, 한다성\*\*\*

## Comparative Analysis for Ranking Systems Based on ATP Tennis Data

Sanghee Kim\*, Juchan Lee\*, Jin S. Kim\*\*, and Daseong Han\*\*\*

### 요 약

랭킹 시스템은 테니스와 같은 경쟁 스포츠에서 중요한 역할을 한다. 그러나 현재의 랭킹 시스템은 선수들의 능력을 보다 객관적으로 평가하는 데 한계가 있다. 본 연구는 다양한 랭킹 시스템을 비교 분석하고 랭킹에 영향을 미치는 요소들을 파악하여 선수들의 능력을 객관적으로 평가할 수 있는 랭킹 시스템을 제안한다. 제안된 랭킹 시스템은 효과적인 랭킹 계산을 위한 데이터의 적절한 기간을 결정하고, 테니스 커뮤니티에 객관적인 기술 평가를 제공하는 데 기여할 것으로 기대된다.

### Abstract

The ranking system plays a critical role in competitive sports such as tennis. However, current ranking systems have limitations in more objectively assessing players' abilities. This study presents a ranking system that can objectively evaluate players' abilities by comparing and analyzing various ranking systems and identifying factors that influence rankings. The proposed ranking system is expected to contribute to determining an appropriate period for data used in effective ranking calculations, ultimately providing the tennis community with an objective technical evaluation.

### Key words

Ranking system, Comparative analysis, ATP tennis data, Evaluation metric

## 1. 서 론

랭킹 시스템은 스포츠와 게임뿐만 아니라 추천 시스템, 검색 엔진 알고리즘 등 다양한 분야에서 중요한 역할을 한다. 이러한 시스템은 공정하고 신뢰할 수 있는 비교를 위한 중요한 지표를 제공한다. 특히 프로 테니스와 같은 경쟁 스포츠에서는 선수

들의 능력과 성과를 평가하는 데 중요한 도구로 사용된다. ATP(Association of Tennis Professionals) 랭킹은 남자 프로 테니스 분야에서 국제적으로 인정받는 랭킹 시스템이다. 그러나 ATP 랭킹 시스템은 주로 선수들의 대회 참여를 장려하기 위해 설계되었으며, 특정 대회의 결과에 크게 의존하고 있어 선수들의 전반적인 능력과 꾸준한 성과를 정확

\* 한동대학교 ICT창업학부, {heexi, dlwncks}@handong.ac.kr

\*\* 풋프린트스퀘어(주), jin.kim@footprintsquare.com

\*\*\* 한동대학교 ICT창업학부, dshan@handong.edu(교신저자)

히 반영하지 못하는 문제가 있다. 예를 들어, ATP 마스터스 1000에서 우승한 선수는 52주 동안 1000 점을 유지하게 된다. 그러나 다음 해에 동일한 대회에서 우승하지 못하면 53주 후에 해당 점수를 잃게 된다. 이는 1~2주라는 짧은 기간 동안 랭킹이 크게 변동할 수 있음을 의미한다. 또한, 주요 대회에서 상위 랭킹의 선수들과 맞붙는 하위 랭킹의 선수들은 랭킹을 올리기 어려운 문제에 직면한다. 이는 대회 초기 라운드에서 하위 랭킹 선수들이 상위 랭킹 선수들과 경기를 펼칠 수밖에 없기 때문에 승리하기 어렵고, 랭킹을 개선할 기회가 제한되기 때문이다. 이러한 문제들은 ATP 랭킹 시스템이 선수들의 실제 성과와 시즌 전반에 걸친 일관성을 충분히 반영하지 못한다는 것을 시사한다.

본 연구는 ATP 데이터를 활용하여 널리 사용되는 랭킹 시스템(즉, Elo 모델[1], Bradley-Terry-Luce 모델[2], Eigenvector Centrality 모델[3])을 비교 분석하여 이러한 문제를 다루고자 한다. 또한, 랭킹 알고리즘에 입력할 데이터의 최적 기간을 결정함으로써, 보다 정확하고 공정한 랭킹 시스템을 제안한다.

본 연구의 목표는 랭킹 알고리즘에 입력할 적절한 경기 기간을 결정하고, 다양한 평가 지표를 기반으로 더 효과적인 랭킹 시스템을 개발하는 것이다. 이 연구를 통해 테니스 선수들의 능력 평가 지표가 개선될 수 있으며, 최적의 입력 변수를 사용하여 더 정교한 스포츠 분석과 데이터 활용이 가능해질 것이다. 이러한 연구 결과는 테니스 커뮤니티에 실질적인 혜택을 제공하고 스포츠 과학 분야 전반에 걸쳐 중요한 기여를 할 것으로 기대된다.

## II. 선행 연구

본 절은 테니스 경기 결과를 기반으로 랭킹 시스템을 개발하고 비교한 이전 연구를 검토하기 위해 두 가지 주요 랭킹 시스템 알고리즘인 통계 알고리즘과 중심성 알고리즘을 분석한다.

### 2-1. 통계 알고리즘

통계 알고리즘은 테니스 경기 결과를 통계적으로 분석하여 랭킹을 산출한다. 대표적인 알고리즘으로

Elo와 BTL(Bradley-Terry-Luce)이 있다.

Elo 모델은 선수의 상대적 기술 수준을 계산하여 랭킹을 매긴다[1]. 경기가 시작되기 전에 각 선수의 승리 확률을 계산하고, 경기 결과에 따라 Elo 점수를 업데이트한다. 예상치 못한 결과는 더 큰 점수 변화를 초래한다. Elo 랭킹은 선수들의 상대적 기술을 양적으로 측정하여 성과를 더 정확하게 반영한다.

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}} \quad (1)$$

$$R'_A = R_A + K(S_A - E_A)$$

$E_A$ 는 선수 B가 선수 A를 이길 확률을 나타내며,  $R'_A$ 는 선수 A의 업데이트된 점수를,  $R_A$ 와  $R_B$ 는 각각 선수 A와 B의 현재 점수를,  $S_A$ 는 선수의 실제 점수를, K는 변동성을 조정하는 상수를 나타낸다.

Bradley-Terry-Luce(BTL) 모델은 선수 i가 선수 j를 이길 확률을 추정한다[2]. 초기 확률 매개변수를 설정한 후 결과 데이터를 기반으로 확률을 반복적으로 업데이트하여 최종 추정값에 수렴한다. BTL 모델은 계산 과정이 복잡하고 시간이 오래 걸리지만 경기 결과에 기반한 보다 정밀한 랭킹을 제공한다.

$$P(i > j) = \frac{p_i}{p_i + p_j} \quad (2)$$

$$p_i = \frac{W_i}{\sum_{j \neq i} \frac{W_{ij} + W_{ji}}{p_i + p_j}}$$

$P(i > j)$ 는 선수 i가 선수 j를 이길 확률을 나타내며,  $p_i$ 와  $p_j$ 는 각각 선수 i와 j의 실력 추정치를,  $W_i$ 는 선수 i의 총 승리 횟수를,  $W_{ij}$ 는 선수 i가 선수 j를 이긴 총 횟수를 나타낸다.

### 2-2. 중심성 알고리즘

네트워크 분석은 노드와 엣지로 구성된 네트워크 구조를 분석하는 방법론이다. 중심성 분석은 네트워크 내 노드의 중심성을 계산하며, 다양한 중심성 지표가 랭킹을 산출하는 데 사용된다.

Eigenvector Centrality는 한 노드의 중심성이 그와

연결된 다른 노드들의 중심성에 의존하여 계산되며, 선수들 간의 상호작용을 통해 실력 차이를 반영한다[3].

$$e_u = \frac{1}{\lambda} \sum_{v \in V} A[u,v]e_v \quad (3)$$
$$\lambda c = Ac$$

$e$ 는 노드의 고유벡터 중심성 값이며, 각 노드의 중요도를 나타낸다.  $V$ 는 노드의 집합을,  $\lambda$ 는 고유값을,  $c$ 는 고유벡터를,  $A$ 는 네트워크의 인접 행렬을 나타낸다.

### III. 연구 방법

본 연구는 데이터 수집 및 전처리, 하이퍼파라미터 탐색 및 튜닝, 평가지표로 구성된 시스템을 제안한다. 전처리된 경기 결과 데이터를 각 랭킹 모델에 입력하여 랭킹을 도출한다. 도출된 랭킹을 기반으로 각 모델의 평가 지표를 계산하여 어떤 모델이 가장 높은 성능을 보이는지 분석한다.

#### 3.1. 데이터

본 연구에서는 남자 프로 테니스(ATP) 경기 데이터를 사용한다. 경기 데이터는 1968년부터 2023년까지의 정보를 포함하는 Jeff Sackmann의 GitHub 저장소에서 수집하였다. 데이터는 경기 날짜, 대회 유형, 코트 표면, 선수 정보, 경기 결과, 세부 성과 통계 등이 포함된다.

#### 3.2. 매개변수 탐색

매개변수 탐색은 기존 모델을 수정 및 변형하여 추가 매개변수를 탐색하는 것을 목표로 한다. 본 연구에서는 경기 데이터를 입력 데이터로 사용한다. 입력 데이터 기간을 달리하여 입력 데이터 세트의 크기를 조정하고, 이에 따라 모델을 학습시킨다. 이 방법은 입력 기간에 따른 성능 차이를 확인하고, 가장 정확한 랭킹을 도출할 수 있는 입력 기간을 확인하는 것을 목표로 한다. 모델은 입력 기간 매개변수를 1년, 3년, 5년으로 설정하여 학습한다.

#### 3.3. 평가 지표

본 연구에서는 랭킹 시스템을 비교하기 위해 4가지 평가 지표를 사용한다. 첫번째, 예측 정확도(Predictive Accuracy) 지표는 상위 랭킹 선수들이 하위 랭킹 선수들을 이길 가능성이 더 높다는 가정에 기반하여 랭킹이 얼마나 정확하게 선수들의 능력을 반영하는지 수치화한다[4]. 예측 정확도는 상위 랭킹 선수가 이긴 경기 수를 총 경기 수로 나누어 계산된다. 두번째, 공정성(Calibration)은 예측된 승리 확률이 실제 결과와 얼마나 일치하는지를 나타내는 지표이다. 이 지표는 더 높은 랭킹을 가진 선수가 이긴 경기 수에서 더 높은 랭킹을 가진 선수가 이길 확률의 합을 나누어 계산한다. 1에 가까울수록 모델이 편향이 적고 공정하게 평가되었음을 의미한다. 세번째, 모델 구별력(Model Discrimination)은 랭킹을 얼마나 정확히 구별하였는지를 평가한다. 더 높은 랭킹의 선수가 이긴 경기들의 예측된 승리 확률의 평균에서 더 높은 랭킹의 선수가 이긴 경기들의 예측된 승리 확률의 평균을 제하여 계산한다. 이 지표는 값이 높을수록 모델의 구별력이 높다는 것을 의미한다. 네번째, 오차(Brier Score)는 예측 정확도를 측정한다. 예측된 승리 확률에서 실제 결과를 뺀 값을 제곱한다. 제곱된 오차들을 모두 더한 후 경기 횟수로 나누어 계산한다. 이 지표는 실제 결과값과의 오차를 나타내며 0에 가까울수록 오차가 적음을 의미한다[5].

또한, 각 랭킹 모델이 특정 기간의 입력 데이터에만 의존하지 않도록 슬라이딩 윈도우 기법을 사용한다. 슬라이딩 윈도우 기법은 고정 크기의 입력 데이터로 모델을 반복적으로 학습시키면서 데이터 기간을 조정하여 평균 성능을 도출하는 방식이다.

### IV. 실험 결과

본 절에서 실험 결과의 간략한 표기를 위해 기본 랭킹 모델, 슬라이딩 윈도우 기법 적용 모델에 대해 각각 Basic, Sliding Window 라는 용어를 사용한다.

랭킹 점수를 계산하는 과정에서, 주로 승리 횟수가 적은 선수들 사이에서 동일한 점수를 가지는 경우가 발생했다. 동일한 랭킹의 부여를 피하기 위해, 동일한 점수를 가진 선수들은 경기 수와 나이를 기준으로 추가적으로 차별화했다. 동일한 점수를 가진

경우에는 승리한 경기 횟수가 더 적은 선수, 나이가 더 어린 선수, 총 경기 횟수가 더 적은 선수를 기준으로 높은 랭킹을 부여했다.

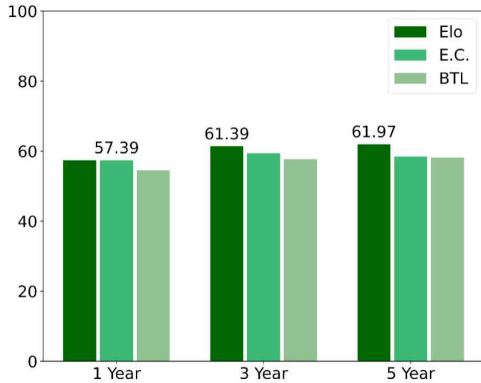


그림 1. 슬라이딩 윈도우 모델들의 연도별 예측 정확도  
Fig. 1. Yearly Predictive Accuracy of Sliding Window Models

그림 1은 입력 데이터를 1년, 3년, 5년으로 변경하며 학습시킨 모델들의 예측 정확도를 나타낸다. 입력 기간에 따른 모델간의 차이를 확인하고, 보다 정확한 랭킹을 도출하기 위한 입력 기간을 검증하였다. 통계 기반의 Elo, BTL 모델은 입력 기간이 늘어남에 따라 높은 성능을 나타냈지만, 높은 연산 시간을 필요로 했다. 반면에 Eigenvector Centrality 모델은 입력 기간이 5년으로 길어지자 성능 저하를 보였다.

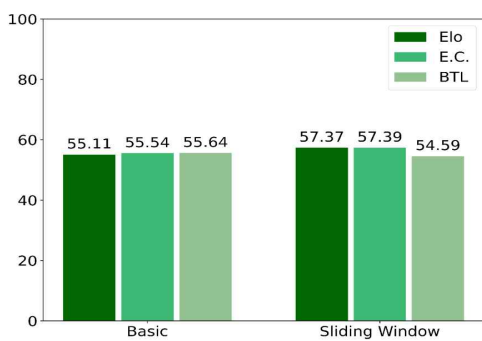


그림 2. 모델들의 예측 정확도  
Fig. 2. Predictive Accuracy of Models

이후 모델들은 입력 데이터를 1년으로 고정하였다. 그림 2는 모델들의 예측 정확도를 나타낸다. 슬라이딩 윈도우 방법을 적용한 모델들은 기본 모델에 비해 성능이 전체적으로 향상하였다.

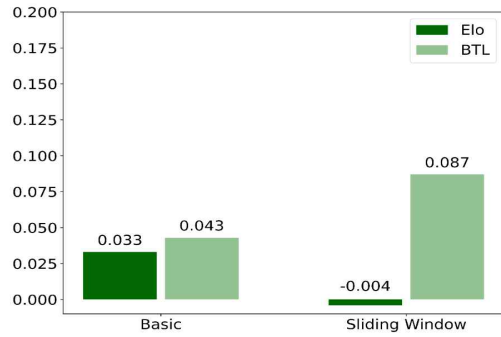


그림 3. 모델들의 구별력  
Fig. 3. Model Discrimination of Models

그림 3은 모델들의 구별력을 나타내며 BTL에 슬라이딩 윈도우 방법을 적용한 모델이 가장 높은 성능을 나타냈다.

그림 4는 모델들의 공정성을 나타낸다. 기본 BTL 모델이 가장 좋은 성능을 보였다.

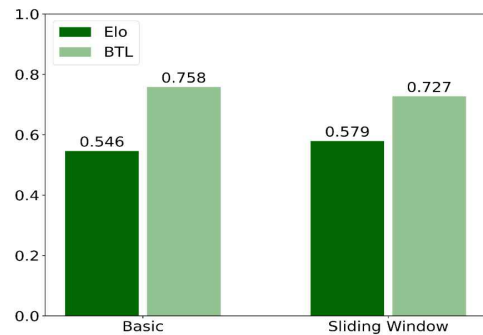


그림 4. 모델들의 공정성  
Fig. 4. Calibration of Models

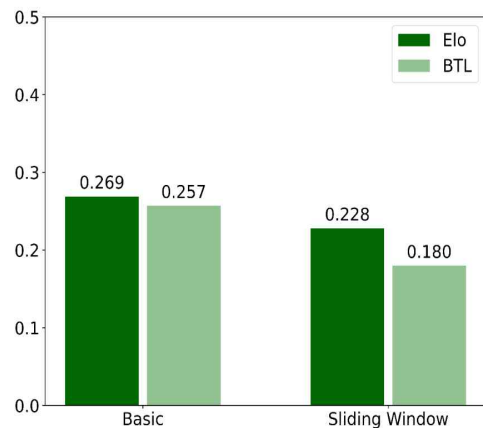


그림 5. 모델들의 오차  
Fig. 5. Brier Score of Models

그림 5는 모델들의 오차를 나타낸다. BTL에 슬라이딩 윈도우 방법을 적용한 모델이 0.18로 가장 낮은 오차를 보여줬다. Elo도 슬라이딩 윈도우 방법만을 적용한 모델이 가장 낮은 오차를 보여줬다. 이에, 슬라이딩 윈도우 방법과 같이 랭킹을 매일 업데이트 하는 것이 실제 경기 결과와 승리 예측 확률의 오차를 줄이는 방법임을 시사한다.

## V. 결 론

본 연구는 테니스 선수들의 전반적인 실력을 객관적으로 측정할 수 있는 랭킹 시스템을 구축하는 것을 목표로 한다. 이를 위해 ATP 테니스 경기 데이터를 수집하고, 다양한 랭킹 시스템을 비교 분석하여 다양한 평가지표를 계산한다. 또한 입력 변수의 적절한 기간을 식별하여 랭킹 모델을 수정한다. 실험을 통해 성능을 상대적으로 개선할 수 있는 모델과 매개변수를 확인했다. 통계 모델은 긴 입력 기간을 사용한 학습이 랭킹 계산에서 더 나은 결과를 도출했다. 계산 시간, 예측 정확도를 고려했을 때 Elo 모델이 높은 성능을 보여준다. 반면에, BTL 모델은 구별력, 공정성, 오차에서 높은 성능을 보였지만 높은 연산 시간을 보인다. 슬라이딩 윈도우 기법은 평가 방법으로 적합했으며, 동일한 점수를 가진 경우 경기 수와 나이를 고려함으로써 시스템의 정확도를 높일 수 있다. 모델에 슬라이딩 윈도우 방법을 적용하면 예측 정확도를 높일 수 있지만 모델의 공정성은 저하될 가능성이 있다.

본 연구는 테니스의 랭킹 시스템을 발전시키는데 기여할 것으로 기대된다. 데이터 분석을 통해 다양한 변수의 영향을 체계적으로 평가함으로써, 향후 랭킹 시스템의 신뢰성과 공정성을 높이는 데 기여할 수 있다. 또한, 제안된 랭킹 시스템은 다양한 스포츠 분야에 적용될 수 있는 가능성을 지니고 있어, 더 넓은 범위의 스포츠 분석 및 성과 평가에 기여할 수 있는 기반을 마련한다.

## 참 고 문 헌

[1] G. Angelini, V. Candila, and L. De Angelis,

"Weighted Elo rating for Tennis Match Predictions," *European Journal of Operational Research*, vol. 297, no. 1, pp. 120-132, February 2022.

[2] I. McHale and A. Morton, "A Bradley-Terry Type Model for forecasting tennis match results," *International Journal of Forecasting*, vol. 27, no. 2, pp. 619-630, 2011.

[3] A. Arcagni, V. Candila, and R. Grassi, "A new model for predicting the winner in tennis based on the eigenvector centrality," *Annals of Operations Research*, vol. 325, no. 1, pp. 615-632, March 2022.

[4] A. Aronson, "TennisRank: A New Ranking of Tennis Players Based on PageRank," 2015.

[5] W. Leighton, C. Chunping, D. Lerato, and G. Hannah, "How well do Elo-based ratings predict professional tennis matches?," *Journal of Quantitative Analysis in Sports*, December 2020.

# 환자의 생리적 지표를 활용한 간경화 진단 모델 구축 및 응용

정은서, 노유민, 임도예, 기희수\*, 광근창\*\*, 염찬욱\*\*\*

## Establishment and Application of a Cirrhosis Diagnosis Model Using Patient Physiological Indicators

Eun-Seo Jung, Yu-Min Roh, Do-Ye Lim, Hui-Su Ki\*, Keun-Chang Kawk\*\*, Chan-Uk Yeom\*\*\*

### 요 약

본 연구는 간경화 진단을 위한 모델을 개발하고, 그 성능을 개선하는 방법을 제시한다. 특히, Grid-ANFIS와 Subtractive-ANFIS 모델을 결합하여, 간경화 진단의 예측 정확성을 높이는 것을 목표로 한다. Kaggle의 "Cirrhosis Prediction Dataset"을 사용하여 실험을 진행하였으며, Grid-ANFIS는 4개의 주요 변수를 사용해 효율적인 예측 성능을 나타냈고, Subtractive-ANFIS는 15개의 변수를 바탕으로 복잡한 패턴을 학습하여 더욱 정밀한 진단을 수행하였다. 본 연구는 이러한 접근을 통해 간경화 진단의 실질적 도구로서 활용 가능성을 확인하며, 향후 연구에서는 가중 앙상블 기법이나 추가적인 임상 데이터, 시계열 분석 기법을 도입하여 성능을 고도화할 가능성을 논의한다.

### Abstract

This study presents a model for cirrhosis diagnosis and suggests methods to improve its performance. Specifically, it aims to enhance diagnostic accuracy by combining the Grid-ANFIS and Subtractive-ANFIS models. Experiments were conducted using Kaggle's "Cirrhosis Prediction Dataset," where Grid-ANFIS demonstrated efficient predictive performance with four key variables, while Subtractive-ANFIS utilized 15 variables to learn complex patterns for more precise diagnosis. This study confirms the feasibility of this approach as a practical tool for cirrhosis diagnosis, and discusses the possibility of further enhancement in performance by incorporating weighted ensemble techniques, additional clinical data, and time-series analysis in future research.

### Key words

Cirrhosis Diagnosis, Grid-ANFIS, Subtractive-ANFIS, Machine Learning, Predictive Modeling

### 1. 서 론

간경화는 만성 간 질환으로, 조기 진단이 환자 예후와 의료 자원 효율성에 중요하다. 대부분 많은 연구에서는 의료 영상 기반 예측 모델에 집중했으나, 생리적 지표 분석을 추가하면 진단과 예측의 정확도가 높아질 것이다.[1]

기존 연구에서 다양한 머신러닝 및 퍼지 시스템

기반 예측 모델(Grid-ANFIS, Subtractive-ANFIS, FCM-ANFIS)을 비교한 결과, Grid-ANFIS이 비교적 우수한 성능을 보이는 것으로 확인되었다[2]. 그러나 이는 매개변수의 수가 적은 경우에 해당하며, 매개변수 수가 증가할 경우 Subtractive-ANFIS는 성능이 향상되는 경향을 보여[3], 간경화 진단에 있어 새로운 가능성을 제시한다. 이에 본 연구는

\* 조선대학교 의생명과학과 학부생, jes0607@chosun.kr, dbalsl025@chosun.kr, doye8208@chosun.kr, ks281537@chosun.kr,

\*\* 조선대학교 전자공학부 교수, kwak@chosun.ac.kr, \*\*\*조선대학교 AI융합대학사업단 연구교수, walt18@chosun.ac.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 지원을 받아 수행된 AI 헬스케어 융합대학 사업 연구임.

Grid-ANFIS와 Subtractive-ANFIS의 강점을 결합하고, 이를 퍼지 클러스터링, 뉴로 퍼지 시스템, 랜덤 포레스트와 통합함으로써 간경화 진단의 정확성을 높이고 의료 현장에서 활용 가능한 정밀 진단 모델을 구축하고자 한다.

본 연구는 환자의 생리적 지표 데이터를 활용하여 다양한 머신러닝 기법으로 간경화 진단의 정확도와 신뢰성을 향상시키고, 이를 통해 기존 연구의 한계를 보완하며 실질적인 진단 모델을 제시하는 것을 목표로 한다.

## II. 관련 연구

### 2-1 ANFIS

ANFIS(Adaptive Neuro-Fuzzy Inference System)는 퍼지 추론과 신경망을 결합하여 모호한 정보를 처리하고 학습 능력을 통해 추론 성능을 최적화하는 시스템이다. 이 시스템은 인간의 사고 방식을 모방하여 다양한 분야에서 예측 및 제어 문제를 해결하는 데 사용된다. ANFIS는 퍼지 규칙을 생성하고 학습 과정을 통해 규칙의 매개변수를 최적화하여 예측 성능을 높인다.

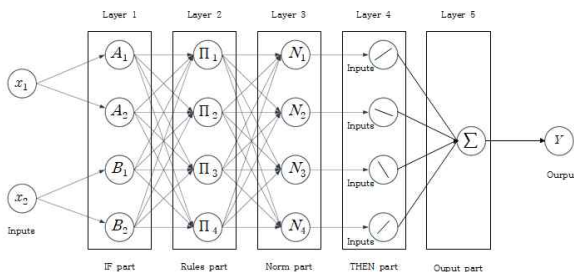


그림 1. ANFIS 구조  
Fig. 1. ANFIS Structure

### 2-2 Grid-ANFIS

Grid-ANFIS는 입력 공간을 격자형 구조로 나누어 중첩이 없는 특정 영역을 생성한다. 이 방법은 입력 변수의 수가 적을 때 유리하며, 입력 공간의 각 구역마다 퍼지 규칙을 적용하여 예측을 수행한다. 예를 들어, 입력 변수가 10개이고 각 변수가 두 개의 멤버십 함수를 가진다면, 2의 10제곱인 1024개의 구역으로 나뉘어지며, 각 구역마다 규칙이 생성되어 매우 복잡한 구조를 형성하게 된다. 따라서 Grid-ANFIS는 소수의 입력 변수로 구성된 저차원

공간에 적합하다.

### 2-3 Subtractive-ANFIS

Subtractive-ANFIS는 Grid-ANFIS의 규칙 수가 증가하는 단점을 보완하기 위해 입력 공간을 클러스터링 알고리즘을 사용해 구역을 나누는 방법이다. 이 방법에서는 클러스터의 반경을 사용자가 설정하며, 반경에 따라 클러스터의 크기와 개수가 결정된다. 작은 반경은 더 많은 클러스터와 규칙을 생성하고, 큰 반경은 클러스터의 수를 줄여 규칙의 수 또한 감소시킨다. 이러한 방식으로 특정 규칙의 수를 줄여 예측 성능을 향상시킬 수 있다.

### 2-4 가중 앙상블 기법 적용

본 연구에서는 개별 모델의 예측 성능을 반영한 가중 앙상블(Weighted Ensemble) 기법을 통해 Grid-ANFIS와 Subtractive-ANFIS의 예측 결과를 결합하여 예측 정확도를 높이는 것을 제안한다. 각 모델의 평균 제곱 오차(MSE)에 반비례하는 가중치를 설정함으로써, 성능이 높은 모델이 최종 예측값에 더 큰 영향을 미치도록 설계하였다. 가중 앙상블 기법은 개별 모델의 예측값에 가중치를 부여하고 이들을 합산하여 최종 예측값을 산출하는 방식이다. 이때 가중치는 각 모델의 예측 성능에 기반하여 설정되며, MSE가 낮을수록 더 높은 가중치를 부여한다. 두 모델의 개별 예측값  $\hat{y}_{Grid}$ 와  $\hat{y}_{Subtractive}$ 에 대한 최종 예측값  $\hat{y}$ 는 다음과 같이 정의된다.

$$\hat{y} = w_1 \cdot \hat{y}_{Grid} + w_2 \cdot \hat{y}_{Subtractive} \quad (1)$$

여기서,  $\hat{y}$ 는 최종 예측값,  $\hat{y}_{Grid}$ 와  $\hat{y}_{Subtractive}$ 는 각각 Grid-ANFIS와 Subtractive-ANFIS의 예측값,  $w_1$ 과  $w_2$ 은 두 모델의 예측 기여도를 나타내는 가중치이며, 두 가중치의 합은 1로 정의된다. 각 모델의 가중치는 MSE의 역수에 비례하도록 설정하여, 성능이 높은 모델의 예측이 최종 예측에 더 큰 기여를 하도록 한다.

이 기법의 구체적인 성능 평가와 실제 적용 가능성은 향후 연구에서 실험적으로 검증될 예정이다.



이를 통해 두 모델의 상보적 특성을 최대한 활용함으로써 더 높은 예측 정확도를 달성할 수 있을 것으로 보인다.

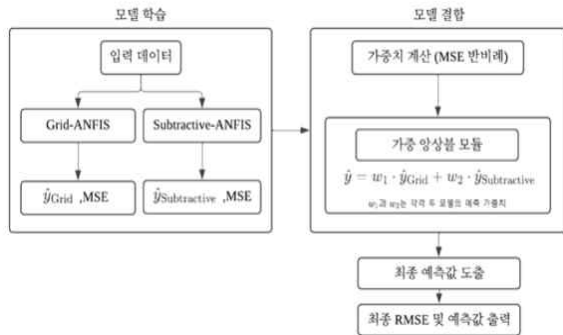


그림 2. 가중 앙상블 기법 구조  
Fig. 2. Weighted Ensemble Technique Structure

### III. 실험 및 결과분석

#### 3-1 데이터베이스

데이터 셋은 기존과 같이 Kaggle에서 제공된 "Cirrhosis Prediction Dataset"을 활용하였다[3]. 해당 데이터셋은 418명의 간경화 환자들로부터 나온 20개의 변수로 구성되어 있다. 데이터의 형식은 CSV이며, 각 샘플은 환자의 간경화 상태와 관련된 정보를 담고 있다. 입력 변수로 성별, 나이, 복수(Ascites) 여부, 비장 비대(Hepatomegaly) 여부, 빌리루빈 수치(Bilirubin), 알부민 수치(Albumin), 혈소판 수(Platelets), 프로트롬빈(Prothrombin) 등이 있고 출력 변수로는 간경화 병기(Stage)를 사용하였다.

이 데이터셋은 무작위 위약 대조 실험의 결과이므로, 본 연구에서는 치료 효과보다는 위약 그룹의 데이터를 중심으로 분석을 수행하였다. 기존 분석에서 생존일수(N\_Days) 변수는 예측 성능과의 관련성이 낮다고 판단되어 이번 분석에서 제외했다.

데이터 전처리 과정에서는 범주형 변수를 분석에 적합한 형식으로 변환하기 위해 원핫 인코딩(One-Hot Encoding)을 사용하였고, 결측값이 있는 경우에는 해당 열의 평균값으로 대체했다. 이러한 전처리 과정을 통해 데이터의 품질을 높이고, 머신러닝 알고리즘의 학습 성능을 높였다.

#### 3-2 실험 방법 및 결과 분석

Grid-ANFIS 모델은 앞선 연구에서 본 것처럼 적

은 수의 변수를 사용할 때 우수한 성능을 보이는 특징이 있었다. 이에 따라 Random Forest 알고리즘을 이용하여 상위 변수 4개를 선정하였고, 생존일수를 제외한 결과로 혈소판(Platelets), 프로트롬빈(Prothrombin), 알부민(Albumin), 그리고 추가적으로 중요한 변수로 나이(Age)가 도출되어 이 4개의 변수를 사용하여 Grid-ANFIS 모델을 구축하였다[4].

반면, Subtractive-ANFIS 모델은 서브트랙티브 클러스터링(Subtractive Clustering)을 통해 데이터를 클러스터링하고, 이를 바탕으로 보다 복잡한 패턴을 반영하는 모델이다. 그래서 Subtractive-ANFIS 모델을 구축할 때는 기존과 달리 총 15개의 매개변수를 모두 사용하여 모델을 생성하였으며, 이를 통해 더욱 정교한 예측을 수행하였다.

정규화와 데이터 증강을 통해 모델의 훈련을 최적화하였다. 학습 데이터와 테스트 데이터는 60:40 비율로 나누어졌으며, FIS 모델이 효과적으로 훈련되도록 설정하였다.

Grid-ANFIS 모델은 입력 데이터에서 4개의 주요 변수를 사용하였다. MF 수를 2~6으로 지정하고 mfType과 outputMfType을 'gaussmf', 'constant'로 설정하여 가장 예측 성능이 좋은 모델을 선정하였다. 또한, 3000번의 epoch를 사용하여 성능을 최대한 끌어올렸다. 훈련 과정에서 에포크가 증가함에 따라 오차가 점진적으로 줄어들었으며, 최소 RMSE는 0.1904, MSE는 0.0363으로 기록되었다. 이는 모델이 적은 변수로도 간경화 진단에서 일정 수준의 예측 성능을 확보할 수 있음을 보여준다. 특히, epoch를 늘릴수록 성능이 향상되었다. 이는 충분한 학습이 이루어질수록 모델이 더 정밀한 패턴을 학습한다는 것을 시사한다.

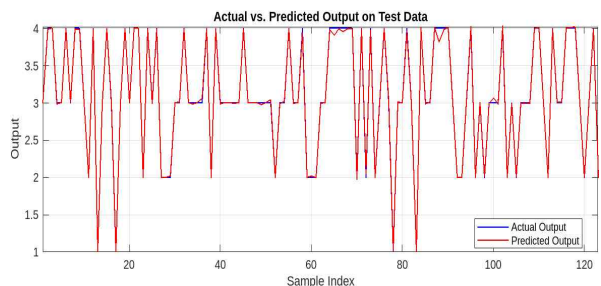


그림 3. grid-ANFIS 실제 및 예측 출력 비교  
Fig. 3. grid-ANFIS real and predicted outputs comparison

Subtractive-ANFIS 모델은 전체 변수를 사용하여 Subtractive clustering 방식을 적용하였다. Subtractive-ANFIS 모델의 성능 최적화를 위해 입력 (inputData) 및 출력 변수(outputData)에 표준화를 적용하였다. 이 과정은 각 변수의 평균을 0으로, 표준편차를 1로 맞추어 변수의 분포 차이가 모델 학습에 미치는 영향을 줄이고, 모델의 수렴 속도를 높이는 데 유리하다. 특히 소량의 데이터 및 변수 간 분포 차이가 클 때 표준화는 민감도를 감소시켜 모델의 예측 성능과 안정성을 개선하는 데 효과적이다. 클러스터 반경은 0.2~0.7 사이에 가장 예측능이 좋은 0.2로 설정 하였다. 이 모델은 상대적으로 적은 epoch(500회)에서 빠르게 수렴하는 경향을 보였으며, 최소 RMSE는 0.1530, MSE는 0.0234로 나타났다. 이는 복잡한 변수를 모두 사용했을 때 모델의 예측 성능이 크게 향상됨을 의미하며, 변수 간의 상호작용을 효과적으로 학습했다는 점을 보여준다.

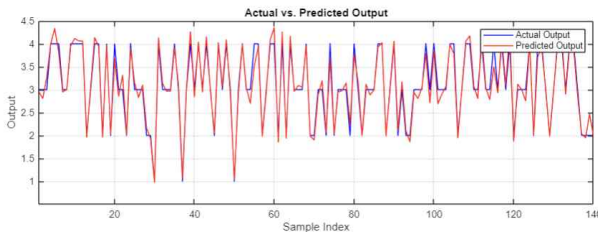


그림 4. Subtractive-ANFIS 실제 및 예측 출력 비교  
Fig. 4. Subtractive-ANFIS real and predicted outputs comparison

표 1. Grid-ANFIS와 Subtractive-ANFIS의 성능 비교  
Table 1. Performance Comparison between Grid-ANFIS and Subtractive-ANFIS

Model	Hyper Parameters	Epoch	Testing RMSE
Grid-ANFIS	MFs : 4	3000	0.1904
Subtractive-ANFIS	Rdus : 0.2	500	0.1530

Grid-ANFIS이 적은 수의 변수로도 효율적인 학습을 보여주었고, Subtractive-ANFIS는 더 많은 변수와 복잡한 패턴을 효과적으로 반영한다.

#### IV. 결론

본 연구에서는 Grid-ANFIS와 Subtractive-ANFIS

모델을 통해 간경화 진단 예측 성능을 향상시키는 방법을 제안하였다. Grid-ANFIS 모델은 적은 수의 변수로도 일정 수준의 예측 성능을 보였고, Subtractive-ANFIS 모델은 더 복잡한 변수를 포함하여 정교한 패턴을 학습하였다. 이는 두 모델이 각기 다른 방식으로 데이터를 처리함으로써 상호 보완적인 역할을 수행할 수 있을 것이다.

향후 연구에서는 간경화 진단 모델의 성능을 더욱 개선하기 위해 몇 가지 방안을 고려할 필요가 있다. 가중 앙상블 기법을 사용하여 두 모델을 결합하여 모델을 고도화 하거나 추가적인 임상 데이터를 통합하면 모델의 일반화 성능을 높일 수 있다. 특히 유전 정보나 생활 습관 같은 외부 요인을 포함함으로써 더 다양한 환자군에 적용할 수 있을 것이다. 또한, 시계열 분석 기법을 도입하면 환자의 상태 변화를 추적하고 장기적인 예측을 가능하게 만들어, 진단의 정확도를 높일 수 있을 것이다. 이를 통해 본 연구에서 제안한 모델은 더욱 실용적이고 발전된 진단 도구로 확장될 수 있을 것이다.

#### 참고 문헌

- [1] Stefan Lucian Popa, et al, "Diagnosis of Liver Fibrosis Using Artificial Intelligence: A Systematic Review", *Medicina*, vol.59, no.5, pp. 1-3, May 2023.
- [2] E. Jung, Y. Roh, D. Lim, H. Ki, K. Kawk, and C. Yeom, "Performance Comparison of Cirrhosis Prediction Model Using Fuzzy Clustering", *Proc. KIPS Conf.*, pp. 1-5, 2024.
- [3] C. Yeom, K. Kwak, Performance Comparison of ANFIS Models by Input Space Partitioning Methods. *Symmetry* 2018, 10(12), 700.
- [4] Kaggle, "Cirrhosis Prediction Dataset", Kaggle, 2024.
- [5] L. Breiman, "Random Forests", *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.

# 빈 컨테이너 부족 문제 해결을 위한 LiDAR 및 GPS 기반 실시간 위치 관리 시스템

주민선\*, 온병원\*\*, 정동원\*\*

## LiDAR and GPS-Based Location Management System for Solving the Empty Container Shortage

Min Sun Ju\*, Byung-Won On\*\*, and Dongwon Jeong\*\*

### 요 약

홍해 사태의 장기화와 미국의 관세 인상 예고로 중국이 수출 물량을 급증 시켜 항만의 혼잡과 빈 컨테이너의 부족 현상이 심화되고 있다. 더불어, 글로벌 화물 수요가 증가하여 선박 물동량이 한계에 이르며 빈 컨테이너 순환 속도가 느려지고 운임이 급등하고 있다. 본 논문에서는 이러한 빈 컨테이너 부족 현상을 해결하기 위해 LiDAR 센서와 GPS를 활용하여 빈 컨테이너의 위치를 추적하고 관리하는 시스템을 제안한다. LiDAR 센서를 통해 빈 컨테이너의 유무를 확인하고 GPS를 통해 빈 컨테이너의 위치를 추적한다. 이러한 방식을 통해 빈 컨테이너 부족 문제와 물류 혼잡을 완화하고, 전체 물류비용과 시간을 절감할 수 있다.

### Abstract

The prolonged Red Sea crisis and the anticipated U.S. tariff hikes have led to a surge in Chinese export volumes, exacerbating port congestion and the shortage of empty containers. Furthermore, the increase in global cargo demand has pushed shipping volumes to their limits, slowing down the circulation of empty containers and causing freight rates to surge. This paper proposes a system that uses LiDAR sensors and GPS to track and manage the location of empty containers to address this shortage. The LiDAR sensors detect whether containers are empty, while GPS tracks their location. This approach is expected to alleviate the empty container shortage and ease logistics congestion, ultimately reducing overall logistics costs and time.

### Key words

empty container shortage, LiDAR sensors, GPS

## 1. 서 론

컨테이너의 사전적 정의는 물건을 담을 수 있는  
견고한 용기이다. 국제운송과 관련해서는 반복적으

\* 국립군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과, {jooms0815}@kunsan.ac.kr

\* 국립군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과, {bwon, djeong}@kunsan.ac.kr(교신저자)

※ 본 연구는 2024학년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065).

로 사용이 가능하도록 규격화된 직육면체형의 용기로, 화물의 단위운송을 실현시켜주는 수송용기로 정의된다. 즉, 컨테이너는 화물을 능률적이고 경제적으로 수송하기 위해 사용하는 상자형 용기를 이르는 말이다. 컨테이너는 정의에서 알 수 있듯이 항만에서 매우 중요한 역할을 한다.

하지만 홍해 사태가 장기화되면서 선박들이 희망봉을 우회하게 되어 항해 거리가 길어지고 운항 정시성이 떨어지면서 항만에 컨테이너 적체가 발생해 빈 컨테이너의 순환이 어려워졌다. 또한, 미국의 관세 인상 예고로 중국이 수출 물량을 급증시키며 항만 혼잡과 빈 컨테이너 부족 현상이 더 심화되었다. 이로 인해 일부 수출 허브에서는 빈 컨테이너의 확보가 어려워졌고, 미주 항로의 선사들이 수익성을 위해 선복 할당을 줄이면서 수출 기업들이 선복 확보에 어려움을 겪고 있다. 더불어, 글로벌 화물 수요가 증가하여 선박 물동량이 한계에 다다르면서 빈 컨테이너 순환 속도가 느려지고 운임이 급등하고 있다.

본 논문에서는 클라우드 기반 시스템을 사용하여 IoT 센서와 GPS를 통해 각 항만과 물류센터의 빈 컨테이너 위치와 상태를 실시간으로 추적하고, 클라우드 서버에 저장된 데이터를 AI 기반으로 분석하여 각 지역의 빈 컨테이너 밀집도를 알 수 있다. 이를 통해 빈 컨테이너 공급 불균형을 해결하고 클라우드 플랫폼을 통해 관계자들이 실시간으로 정보를 공유할 수 있다. 이를 통해 빈 컨테이너 부족 문제를 해결하고, 전체 물류비용과 시간을 절감을 기대한다.

본 논문의 구성은 다음과 같다. II장에서는 관련 연구에 대해서 소개하고, III장에서는 결론을 기술한다.

## II. 관련 연구

### 1.1 LiDAR 구성 요소

본 논문에서는 LiDAR 센서를 사용하여 컨테이너가 비어있는지에 대한 여부와 만약 비어있다면 얼마나 여유 공간이 있는지 측정할 것을 제안한다.

LiDAR 센서는 레이저 송신기와 펄스 에너지, 수신기의 민감도를 이상적으로 조율함으로써 반사율이 낮으면서 먼 거리에 떨어져 있는 대상물도 확실하게 측정할 수 있다. 이를 활용해 컨테이너 내부의 상태를 확인할 수 있으며, 특히 다층 스캐너와 함께 사용될 경우 x, y, z 방향 위치 정보까지 얻을 수 있어 여러 각도에서 데이터를 수집하여 공간의 빈 비율을 정확히 파악할 수 있다. MRS1000은 SICK에서 개발한 3D LiDAR 센서로, 네 개의 스캔 레이어를 통해 275°의 넓은 시야각을 제공한다. 이 센서는 HDDM+ 기술과 멀티 에코 기능을 적용하여 비, 먼지, 안개 등 다양한 환경 조건에도 높은 신뢰성과 내후성을 보장한다[1].

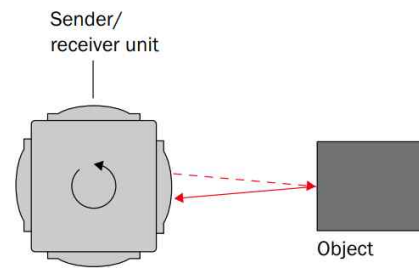


그림 1. MRS1000의 구조  
Fig. 1. MRS1000 Structure

### 2.2 GPS 기반 컨테이너 관리, 위치추적 시스템

GPS란 인공위성을 활용하여 지구상의 위치를 정확하게 파악하는 시스템이다. GPS는 지구 궤도를 도는 24개의 위성으로 구성되며, 각 위성은 고유의 신호와 궤도 정보를 전송한다. 지상의 GPS 수신기는 이러한 신호를 수신하여 위성과의 거리를 계산하고, 이를 통해 정확한 위치를 산출한다. 이러한 GPS 기능을 통해 빈 컨테이너의 위치를 추적한다. 그림 2는 시스템의 흐름도이다. 위치와 이동 상태 데이터를 GPS를 통해서 수집하고 GPS 데이터를 RS-232 통신 포트를 통해 NMEA 형식으로 전송한다. 여기서 NMEA는 GPS 장치 간 데이터 교환에 사용되는 표준 프로토콜이다. 수신된 데이터를 GGA와 VTG로 구분한다. GGA는 GPS로부터 현재 위치 정보를 나타내고, VTG는 GPS의 속도 및 이동 방향 상태를 나타낸다[2].

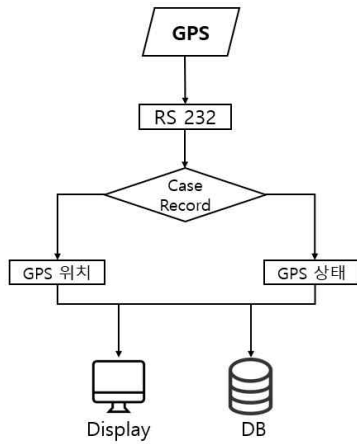


그림 2. 시스템 흐름도  
Fig. 2. System Flowchart

또한 GPS를 사용한 모니터링 시스템을 활용하여 컨테이너를 관리할 수 있다. 컨테이너에서 자신의 수신기로 위성으로부터 위치정보를 받게 되고, 컨테이너에서 수신된 위치 정보는 측정된 위치 데이터와 관계된 부가 데이터로써 NMEA(LLQ) 메시지 형식으로 저장되어 통째실로 전송하여 사용자에게 제공되며 전송된 데이터를 이용하여 모니터링 프로그램에서 시각화 시켜준다. 그림 3은 컨테이너 모니터링 시스템의 구조이다.

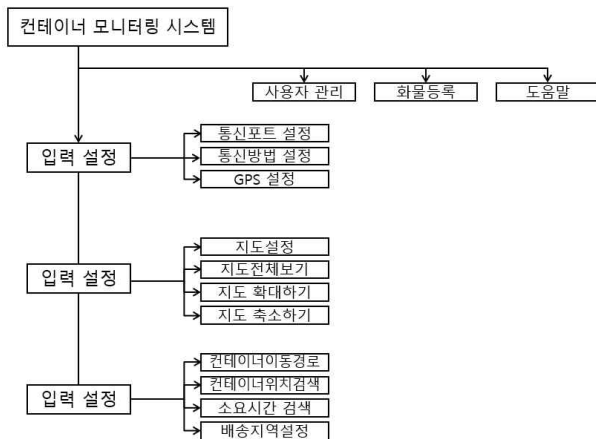


그림 3. 컨테이너 모니터링 시스템 구조  
Fig. 3. Container Monitoring System Architecture

### III. 결 론

LiDAR 센서를 활용함으로써 컨테이너가 비어 있는지의 여부를 확인할 수 있으며, 다층 스캐너를 통

해 컨테이너 내부의 공간을 정밀하게 측정하여 얼마만큼 컨테이너가 비어있는지에 대한 데이터를 제공한다. 이러한 상태 정보는 컨테이너의 효율적인 재배치와 활용도를 높이는 데 기여할 수 있다.

또한 GPS를 통해 빈 컨테이너의 실시간 위치를 추적함으로써 물류 흐름을 원활하게 관리할 수 있다. 이 시스템은 RS-232 통신과 NMEA 프로토콜을 이용하여 위치 데이터를 신속하게 전송하고, 이를 통해 관계자들이 각 컨테이너의 위치 및 이동 상태를 실시간으로 파악할 수 있도록 한다.

본 논문에서는 LiDAR 센서와 GPS를 결합하여 컨테이너의 상태와 위치를 효과적으로 모니터링할 수 있는 시스템을 제안하여 빈 컨테이너의 위치를 파악하여 빈 컨테이너 부족 현상을 해결하기를 기대한다.

### 참 고 문 헌

- [1] Harald Webe, Head of Product Unit Ranging LiDAR sensors. (2018). LiDAR 센서의 작동 방식과 다양한 이형. SICK AG in Waldkirch / Deutschland
- [2] 최병길, 진세일, 홍상기. (2004). GPS를 이용한 컨테이너 모니터링 시스템에 관한 연구. 한국측량학회지, 22(4), 401-409.

# ERP 시스템 전환 시 GPT 기반 데이터 자동 매핑 방법 제안

김영민\*, 권세종\*, 김능회\*\*, 정동원\*\*

## Proposed automatic data mapping method based on GPT when switching ERP systems

Youngmin Kim\*, Sejong Kwon\*, Neunghoi Kim\*\*, and Dongwon Jeong\*\*

### 요 약

ERP 시스템은 실행에 상당한 비용과 전문성이 요구된다. 따라서 개인사업자 와 소상공인은 ERP 시스템 제공 업체에 비용을 지불하고 시스템을 이용한다. 다양한 목적과 사용자의 요구사항을 충족하기 위해 여러 시스템을 사용해 보는 것은 중요한 과정이다. 하지만 시스템마다 데이터 속성이 달라 적용이 어렵다는 한계점이 있다. 따라서 본 논문은 이러한 문제를 해결하기 위해 ERP 시스템 전환 시 GPT 기반 데이터 자동 매핑 방법을 제안한다. 제안 방법은 GPT-4o를 프롬프트 엔지니어링 기법을 활용하여 규칙을 반복적으로 학습시켜 일관된 결과를 유도하도록 만든다. 해당 모델을 사용하여 기존 시스템의 데이터를 다른 환경에서 사용 가능한 새로운 데이터로 자동 변환시킨다. 제안 방법을 통하여 비용 감소, 정확성 향상과 같은 효과를 기대할 수 있다.

### Abstract

ERP systems require significant cost and expertise to run. Therefore, self-employed individuals and small business owners pay ERP system providers and use the systems. It is an important process to try out various systems to meet various purposes and user requirements. However, there is a limitation that it is difficult to apply because the data properties of each system are different. Therefore, this paper proposes a GPT-based data automatic mapping method when converting ERP systems to solve this problem. The proposed method repeatedly learns rules using the prompt engineering technique of GPT-4o to induce consistent results. Using the model, the data of the existing system is automatically converted into new data that can be used in a different environment. Through the proposed method, we can expect effects such as cost reduction and accuracy improvement.

### Key words

ERP, GPT, GPT-4o, automatic data mapping, switching ERP system

---

\* 국립군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과, {kimmik7119, }@kunsan.ac.kr

\*\* 국립군산대학교 컴퓨터소프트웨어특성화대학 소프트웨어학과, {nunghoi, djeong}@kunsan.ac.kr(교신저자)

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065)

## 1. 서 론

ERP(Enterprise Resource Planning, 이하 ERP)는 기업 전사 프로세스 기반으로 설계되어 표준화된 업무 기준을 엄격하게 제시하는 기업 솔루션으로서 생산, 구매, 물류, 회계 등 기업 전반의 경영활동 프로세스를 통합 연계 관리해주는 전사적 통합시스템이다[1].

현재 국내에는 여러 ERP 시스템이 존재한다. 이중 사용 목적과 사용자의 시스템 사용 패턴에 따라 적합한 시스템이 존재하기 마련이다. 적합한 시스템을 선택하기 위해 여러 시스템을 경험해 보는 것 또한 필요한 단계이다.

기존 ERP 시스템은 타 시스템과 데이터 이전이 복잡하다는 한계점이 존재한다. 여러 시스템을 경험하기 위해서는 시스템 간 안전한 데이터 이전이 보장되는 기능 지원이 필요하다.

따라서 본 논문은 Chat-GPT를 통해 서로 다른 시스템에서 비롯되는 형태의 데이터를 원하는 시스템에 맞도록 변환시켜주는 데이터 자동 매핑 방법에 대해 제안한다.

## II. 필요 배경 및 관련 기술

### 2.1 시스템 필요 배경

중소기업청과 지방상공회의소 등을 통해 국내 중소기업들도 정부기관의 지원을 받거나 자체적으로 시스템을 도입하고 있다. 그러나 ERP 시스템의 장점을 인식하고도 도입하는 기업들 중 ERP시스템 실행을 포기하는 기업도 발생하고 있는데, 그 이유는 ERP 시스템 실행에 상당한 비용과 전문성이 요구되기 때문이다[2].

이러한 이유로 개인사업자 또는 소상공인의 경우는 ERP 시스템 제공 업체에 비용을 지불하고 시스템을 사용하는 경우가 많다.

개인사업자 또는 소상공인은 업종이 넓은 만큼 사용 목적이 다양하다. 따라서 더욱 선택의 폭이 넓어야 한다. 하지만 다른 두 시스템 간 데이터 이전이 원활하지 못하다면 사용자는 시스템 교체를 망설이거나 포기할 것이다.

현재 운영하는 시스템은 CSV, XLS, JSON 등의

형식을 통하여 데이터 Export 한다. 같은 시스템에 해당 데이터를 Import 한다면 문제가 없겠지만 다른 시스템에 Import를 시도할 경우 Column의 속성이 달라 해당 파일을 사용하지 못하는 경우가 발생할 수 있다.

### 2.2 관련 기술

GPT 기반 데이터 자동 매핑 시스템을 개발하기 위해 본 논문에서 제안하는 핵심 기술은 Chat-GPT이다. 최신 모델인 GPT-4o를 활용하여 데이터를 변환하여 과정을 자동화 및 간결화할 수 있다.

GPT-4o를 사용할 경우, 변환 규칙을 학습시켜야 한다. 학습 없이 데이터를 업로드할 경우 정확한 처리가 이루어지지 않을 수 있다. 이로 인해 날짜 형식, 단위 등과 같은 변환 규칙을 반복적으로 학습시키는 것이 중요하다. 이를 위해 프롬프트 엔지니어링 기법을 활용하여 모델이 제시하는 일관된 결과를 유도하도록 지침을 제공해야 한다.

또한, 데이터 변환 시 발생할 수 있는 변환 오류나 예외를 자동으로 GPT가 감지하여 처리할 수 있도록 해야 한다. 필드 값이 예상 범위를 벗어날 경우, 관리자에게 경고를 보내도록 설정하는 것과 같이 GPT가 감지함으로써 데이터의 정확성 및 무결성을 유지할 수 있다.

## III. 데이터 자동 매핑 방법

먼저, 해당 시스템이 원활하게 작동하려면 현재 서비스하고 있는 ERP 시스템들의 Column 속성을 파악해야 한다. 이 속성들은 ERP 시스템 제공 업체로부터 제공받아야 하며, 받은 속성들을 Chat-GPT에 사전 학습을 시켜줘야 한다.

둘째, 기존 시스템에서 Export된 데이터가 필요하다. 해당 데이터의 속성을 파악하고, 저장해둔다.

셋째, Chat-GPT를 통해 변환하고자 하는 회사에 입력하고 CSV 파일을 업로드한다. 이 과정을 통해 A 회사에서 Export 된 데이터는 B 회사에 바로 Import 가능한 형식으로 자동 변환되게 된다.

이러한 과정을 통해, 기존 A 회사에서 사용하던 데이터를 간단한 과정을 통해서 B 회사에 사용할 수 있는 환경이 된다.

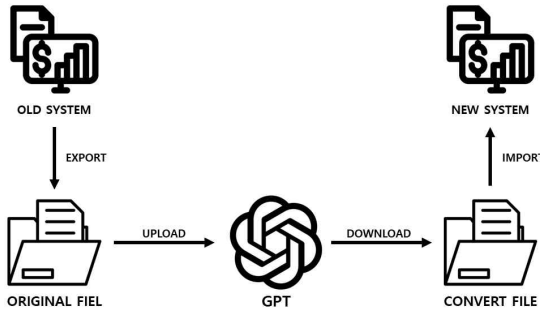


그림 1. 데이터 자동 매핑 절차  
Fig. 1. Data automatic mapping procedure

#### IV. 기대 효과 및 한계점

##### 4.1 기대 효과

본 논문에서 제안하는 방법을 통하여 기대할 수 있는 효과는 다음과 같다.

첫째, ERP 전환 비용이 절감된다. 종종 수작업으로 진행되던 매핑 작업을 학습된 GPT 모델을 통해 자동화하여 인력과 시간 절약하고, 비용을 절감할 수 있다.

둘째, 데이터의 일관성과 정확성을 향상시킬 수 있다. 수작업으로 인해 발생할 수 있는 오류를 학습된 GPT를 통하여 데이터 손실을 줄이고, 신뢰성을 높일 수 있다. 또한, 규칙 기반의 검증 및 오류를 자동으로 감지하여 관리 및 보고에 용이해지며, 효율적으로 유지관리도 가능하다.

마지막으로, 확장성과 호환성이 높다. 다양한 ERP 시스템 확장이나 전환 시 추가적인 매핑 작업 없이 데이터 이전이 가능하다. 따라서 장기적으로 유연성이 강화되고, 변화에 대응하는 속도를 효과적으로 개선할 수 있다.

##### 4.2 한계점 및 보완점

본 논문에서 제안하는 데이터 전환 방법에는 몇 가지 한계점과 보완점이 존재한다.

첫째, GPT-4o 사용 시 보안의 위험이 존재한다. 본 방법론과 밀접하게 존재하는 보안 위험으로는 데이터 유출 및 탈취이며, 대화를 통해 학습을 개선하기 위해 대화 내용을 저장, 해킹을 통해 인공지능의 데이터를 탈취를 예로 들 수 있다.[3] 이로 인해, 개인 정보나 민감한 정보는 값을 난독화하는 과정

을 통해 보안성을 강화할 필요가 있다.

둘째, GPT-4o를 사용하여 데이터를 변환하는 방법의 특성상 변환 과정에 개입하기 힘들다. 이로 인해 사전에 방지하는 것이 중요하다. 따라서 변환 규칙을 사전에 철저하게 설정하고, 업로드 전 난독화하는 등 보안 조치를 해야 한다.

셋째, API 사용 시 보안 연결(HTTPS)를 통해 네트워크 전송 중 데이터가 보호되도록 하는 것도 하나의 방법이 될 수 있다.

마지막으로, 구현 논문이 아닌 제안 논문으로 해당 방법에 대해 검증되지 않았고, 실제 적용 시 발생할 수 있는 문제점에 대해 파악하지 못한다는 한계가 존재한다. 따라서 후속 연구를 통하여 해당 방법에 대한 검증이 필요하다.

#### V. 결 론

본 논문에서는 GPT-4o 기반의 데이터 자동 매핑 방법을 통해 ERP 시스템 간 전환을 자동화하고 간소화하는 방법을 제안한다. 이를 통해 ERP 전환 시 비용과 시간을 절약하고, 데이터의 일관성과 정확성을 향상시켜 효율적인 유지관리가 가능할 것으로 기대된다. 하지만, 보안적인 위험도 존재하므로 데이터 난독화 및 보안 연결(HTTPS) 등의 조치가 필요하다. 시스템 특성상 중간에 개입이 불가능하므로, 변환 규칙 설정과 전처리가 중요하다. 향후 구현과 검증을 통한 추가적인 연구가 필요하다는 과제가 남아있다.

#### 참 고 문 헌

- [1] 김성민, "ERP 시스템 기능 확장성에 관한 구현 사례연구", 2021.
- [2] 송신근 and 전진영, "개인적 특성이 ERP시스템의 지식이전에 미치는 영향에 관한 연구", 국제회계연구, Vol. 11, pp. 195-210, 2004.
- [3] 박종현, 우지은, and 우성희, "Chat GPT 활용과 보안 이슈", 한국정보통신학회 종합학술대회 논문집, 경기, pp. 518-520, 2023-10-26.



# 에코 챔버 현상과 추천 알고리즘이 사용자의 정보 소비 습관에 미치는 영향과 완화 방안

김용민\*, 강명진\*, 김능회\*\*

## Effects of echo chamber phenomena and recommendation algorithms on users' information consumption habits and suggestions for mitigating them

Yongmin Kim\*, Myung-jin Kang\*, and Neung-hoe Kim\*\*

### 요 약

본 연구는 디지털 플랫폼의 추천 알고리즘이 에코 챔버 현상을 형성하고 강화하는 방식과 이를 완화하기 위한 방안을 분석한다. 기존 추천 알고리즘은 사용자의 선호 콘텐츠를 반복 추천하여 특정 관점에 고착되게 할 수 있으며, 정보 다양성을 제한할 위험이 있다. 이를 해결하기 위해 본 연구는 중립적 콘텐츠 주기적 노출, 다양한 시각 선택이 가능한 사용자 인터페이스 개선, 그리고 이 효과를 검증하기 위한 추적 실험을 제안한다. 이러한 접근을 통해 정보 편향성을 줄이고, 균형 잡힌 정보 환경을 조성하는 데 기여하고자 한다.

### Abstract

This study analyzes the way the digital platform's recommendation algorithm forms and reinforces the eco-chamber phenomenon and the measures to mitigate it. Existing recommendation algorithms can repeatedly recommend users' preferred content to stick to a specific point of view, and there is a risk of limiting information diversity. To solve this problem, this study proposes neutral content periodic exposure, improvement of a user interface that allows various visual choices, and tracking experiments to verify this effect. Through this approach, we intend to reduce information bias and contribute to creating a balanced information environment.

### Key words

Recommendation Algorithms, Echo Chamber Phenomenon, Information Bias, Mitigation Measures

## 1. 서 론

---

\* 국립군산대학교 컴퓨터소프트웨어학부 소프트웨어학전공, gr1gr2kym@gmail.com,  
소프트웨어학전공, mljc0123@gmail.com,

\*\* 국립군산대학교 컴퓨터소프트웨어학부 소프트웨어학전공(교신저자), nunghoi@kunsan.ac.kr

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065)

현대 사회에서 SNS와 디지털 플랫폼의 발전은 정보 소비 방식을 크게 변화시켰다. 시장조사업체 DMC미디어의 ‘2024 소셜 미디어 동향보고서’에 따르면 2024년 1월 기준으로 우리나라의 소셜미디어 활성도는 93.4%로 전 세계 평균 71%보다 크게 상회하는 것을 볼 수 있다[1]. 우리나라 사람들은 보통 콘텐츠 소비를 목적으로 SNS 및 디지털플랫폼을 사용한다. 특히, 이러한 플랫폼에서 사용되는 추천 알고리즘은 사용자의 관심사에 맞춘 콘텐츠를 자동으로 추천하여 사용자 경험을 향상시켰다. 하지만, 이러한 경험 향상은 정보 편향과 에코 챔버(Echo Chamber) 현상을 강화할 위험이 있다. 여기서 에코 챔버 현상은 사용자가 동일한 시각의 정보에 반복적으로 노출되면서 특정 관점에 고착되는 문제로, 정보의 다양성을 제한한다[2]. 예를 들어, 넷플릭스의 추천 알고리즘은 사용자가 선호하는 콘텐츠와 유사한 장르를 반복적으로 추천하여 사용자가 다양한 시각을 접할 기회를 줄인다.

본 연구는 추천 알고리즘이 에코 챔버 형성에 미치는 영향을 분석하고, 정보 소비 편향성을 줄이기 위한 대안을 모색하는 것을 목적으로 한다.

## II. 에코 챔버 완화 방안 제안

추천 알고리즘은 사용자가 과거에 선호했던 콘텐츠 유형을 학습하여, 이와 유사한 정보를 반복적으로 제공함으로써 사용자를 특정 시각에 고착시키는 경향이 있다. 이를 통해, 사용자가 선호하는 콘텐츠와의 상호작용이 많아질수록 알고리즘은 이를 강화하여 더욱 유사한 콘텐츠를 추천하게 된다. 이런 피드백 루프는 다양한 시각을 접할 기회를 줄이고, 특정 관점에 대한 고착을 가속화한다.

이러한 에코 챔버 현상을 완화하기 위해, 본 연구에서는 추천 알고리즘과 사용자 인터페이스(UI) 개선 방안을 제안한다. 구체적으로는, 자극적 콘텐츠와 반대되는 시각의 중립적 콘텐츠를 주기적으로 노출하는 알고리즘 설계, 다양한 시각의 콘텐츠를 선택적으로 볼 수 있는 UI 개선, 그리고 이러한 방안을 실제로 적용하여 에코 챔버 효과를 평가할 수 있는 추적 실험을 제시한다.

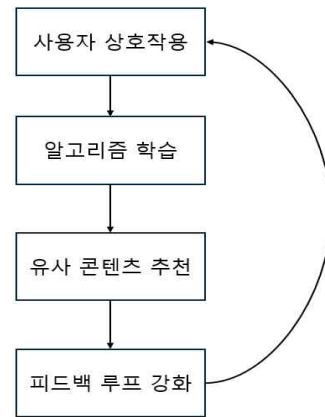


그림 1. 추천 알고리즘의 에코 챔버 현상 가속화 과정  
Fig. 1. Acceleration process of echo chamber phenomenon In recommendation algorithm

표 1. 추천 알고리즘 특징  
Table 1. Recommendation Algorithm Features

	기존 알고리즘 특징	개선 알고리즘 특징
데이터 활용	사용자 행동 데이터에 의존	다양한 콘텐츠 노출
콘텐츠 추천	선호 콘텐츠 반복 추천	중립적 콘텐츠 주기적 추천
피드백 방식	피드백 루프 강화	다양성 보장 알고리즘 적용

표 1과 같이 추천 알고리즘은 사용자의 선호 콘텐츠를 반복적으로 추천하여 특정 관점에 고착될 가능성이 높다. 이는 사용자가 편향된 정보만을 소비하게 만들어 다양한 시각을 접하는 기회를 제한할 수 있다. 이에 대한 개선 방안으로, 자극적 콘텐츠와 반대되는 시각의 중립적 콘텐츠를 주기적으로 노출하는 알고리즘을 설계하여 사용자가 새로운 관점을 접할 수 있도록 한다[3].

사용자 인터페이스 개선 방안으로는 ‘다양한 시각 보기’ 옵션을 제공하여 사용자가 자신과 다른 시각의 콘텐츠를 선택적으로 볼 수 있는 UI를 설계한다. 이를 통해 에코 챔버 효과를 줄이고 정보의 균형을 촉진하는 환경을 조성할 수 있다. 이러한 UI 개선은 사용자가 무의식적으로 편향된 정보만을 소비하는 것을 방지하고, 자발적으로 다양한 정보를 탐색할 수 있는 선택지를 제공한다.



그림 2. 제안 방법의 순서도  
Fig. 2. Flowchart of the proposed solution

추적 실험에서는 대안 알고리즘과 개선된 UI를 통해 사용자들에게 다양한 시각의 콘텐츠가 노출되는 환경을 조성하고, 특정 사용자 그룹을 대상으로 에코 챔버 현상의 완화 여부를 실험하여 그 효과를 측정한다. 이 실험을 통해, 추천 알고리즘이 에코 챔버 현상을 어떻게 강화하는지에 대한 실질적 데이터를 수집하고, 이를 완화할 수 있도록 구체적으로 제안할 수 있다.

본 연구는 학술적 기여로서 디지털 기술이 정보 소비와 에코 챔버 형성에 미치는 메커니즘을 분석함으로써, 정보 편향성과 사회적 갈등의 원인을 심도 있게 이해하는 데 도움을 줄 것이다. 또한, 실무적 기여로서 에코 챔버를 완화하기 위한 대안적 추천 알고리즘과 사용자 인터페이스 설계를 제안함으로써, 디지털 플랫폼에서 정보의 다양성을 보장하고 사용자에게 균형 잡힌 정보 환경을 제공할 수 있는 방안을 제시한다.

### III. 결 론

본 연구는 추천 알고리즘이 에코 챔버 형성에 미치는 영향을 분석하고, 이를 완화하기 위한 대안으로 중립적 콘텐츠 주기적 추천 알고리즘과 다양한 시각을 제공하는 사용자 인터페이스 및 추적 실험 방안을 제안했다.

이러한 접근은 정보 소비와 에코 챔버 형성 메커니즘을 이해하고, 정보 다양성 촉진 방안을 제시하는 데 중요한 역할을 한다. 또한, 에코 챔버 완화를 통해 디지털 플랫폼이 균형 잡힌 정보 환경을 제공하여 사회적 갈등을 줄이고 건강한 정보 소비 문화를 조성할 수 있기를 기대한다.

### 참 고 문 헌

- [1] DMC 리포트, 2023. 『2024 소비자 행태 분석: 디지털 미디어와 콘텐츠 소비 트렌드』, DMC 미디어.
- [2] Propelland. "How AI Shapes Digital Echo Chambers by Reinforcing User Preferences and Limiting Exposure to Diverse Viewpoints." 2023.
- [3] 조진형, 김규정, 2022, 『소셜미디어에서 에코챔버에 의한 필터버블 현상 개선 방안 연구』, 숭실대학교

# 컨테이너 관리 시스템의 안정성 향상을 위한 자원 해제 자동화 및 오류 복구 기법

허유정\*, 김창인\*\*, 최창범\*\*

## Automation of Resource Release and Error Recovery Method for Improving Stability of Container Management System

Youjeng Heo\*, Changin Kim\*\*, and Changbeom Choi\*\*

### 요 약

본 연구는 컨테이너 관리 시스템의 안정성을 높이기 위한 자원 해제 자동화와 오류 복구 기법을 제안한다. 컨테이너 종료 시 PID 파일, 네트워크 인터페이스(veth), 마운트된 파일 시스템을 자동으로 해제하여 시스템 자원이 불필요하게 점유되지 않도록 하였으며 자원 해제 실패 시 재시도 메커니즘을 통해 연속성을 보장한다. 이러한 기법을 통해 시스템의 자원 효율성과 안정성을 향상시킬 수 있었으며 향후 경량화된 컨테이너 환경에서 오케스트레이션 기능 통합도 가능할 것이라 기대한다.

### Abstract

In this study, we propose resource release automation and error recovery techniques to increase the stability of container management systems. Automatically disconnects PID files, network interfaces, and mounted file systems at the end of the container to prevent unnecessary occupancy of system resources, and ensures continuity through a retry mechanism in case of resource release failure. We expect that these techniques will improve the resource efficiency and stability of the system, and that orchestration capabilities will be integrated in a lighter container environment in the future

### Key words

Container Management, Resource Release Automation, System Stability, Lightweight Virtualization

## 1. 서 론

컨테이너 기술은 가상화보다 경량화된 환경을 제공하여 애플리케이션의 배포와 자원 효율성 개선에 이바지한다. 컨테이너는 가상머신과 달리 애플리케이션 계층에서 가상화되어 호스트 OS 커널을 공유하며 리소스 사용이 적다는 장점이 있다[1]. 이러한 컨테이너를 효과적으로 관리하고 안정성을 보장하기 위해서는 인스턴스 자원을 적절히 해제하고, 종료 단계에서 발생할 수 있는 오류를 자동

\*국립한밭대학교, 20211939@edu.hanbat.ac.kr,

\*\*국립한밭대학교, 20217140@edu.hanbat.ac.kr, \*\*국립한밭대학교, cbchoi@hanbat.ac.kr

으로 처리해 주는 시스템이 필수적이다[2]. 특히, 컨테이너 종료 시 자원이 해제되지 않거나 네트워크 네임스페이스가 남아 있는 경우, 시스템 성능 저하 및 안정성 문제가 발생할 수 있다.

본 논문에서는 자원 해제의 자동화 및 오류 복구를 통해 컨테이너 관리 시스템의 안정성을 향상하는 방법을 제안한다. 컨테이너 런타임 구조에서 생애주기 관리의 핵심인 생성, 실행, 중지, 삭제 기능을 개발했으며, 이를 통해 컨테이너 종료 시 발생할 수 있는 시스템 성능 저하 문제를 개선하고자 한다[3].

## II. 컨테이너 런타임 생애주기

본 논문에서 제안한 컨테이너 관리 시스템 아키텍처는 자원 해제의 효율성을 높이는 방향으로 설계되었다. 본 시스템은 컨테이너의 생성, 실행, 중지, 삭제의 생애주기 관리를 위한 런타임 모듈과 자원 해제를 자동화하는 모듈을 포함하며 네트워크 인터페이스와 프로세스 ID(PID)과 같은 주요 자원을 효율적으로 관리하는 기능을 제공한다.

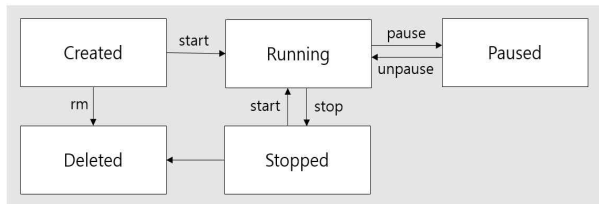


그림 1. Container Lifecycle

그림 1은 컨테이너의 생성, 실행, 중지, 삭제의 내용을 포함한 생애주기를 표현한 것이다. 컨테이너 생성 단계에서는 Linux 네임스페이스와 Cgroups를 활용하여 컨테이너별로 독립된 네트워크와 프로세스 환경을 제공한다. 이를 통해 각 컨테이너는 격리된 네트워크 인터페이스와 프로세스 공간을 확보하게 되며 자원 충돌 없이 안정적으로 동작한다. 실행 단계에서는 생성된 네트워크 인터페이스와 호스트 간의 원활한 통신을 보장하며 컨테이너의 리소스 활용도를 높인다

## III. 자원 해제 메커니즘

컨테이너의 종료 단계에서는 PID 파일, 네트워크 인터페이스(veth), 마운트된 파일 시스템의 해제를 자동화한다. 컨테이너 종료와 동시에 PID 파일이 삭제되며 불필요한 프로세스 참조가 방지된다. 또한 네트워크 인터페이스(veth)도 자동으로 해제되어 네트워크 자원이 누적되는 것을 방지한다. 이러한 자원 해제 메커니즘은 컨테이너 종료 시 호출되어 시스템의 자원 효율성을 높이고 하였다.

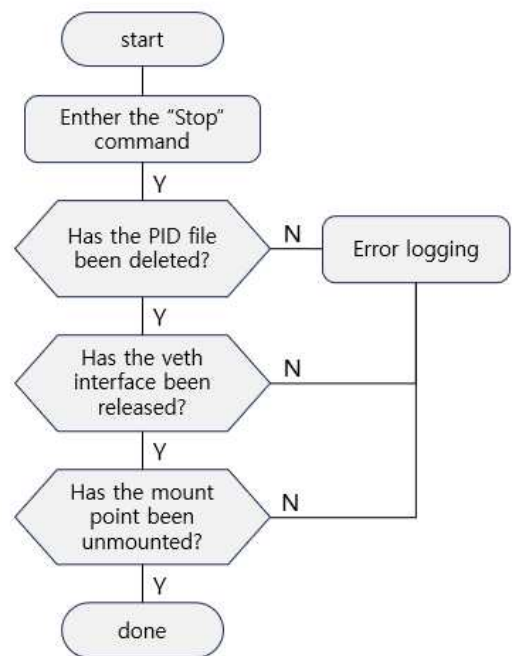


그림 2. Container Resource Release Flow

또한 본 논문에서는 자원 해제에 실패하거나 프로세스가 비정상적으로 종료되는 경우를 고려하여 오류 복구 방안을 설명한다. 만약 네트워크 인터페이스(veth)나 마운트 포인트 해제 및 삭제에 실패할 경우 시스템에서 오류 메시지를 기록하고 재시도를 수행하여 자원 해제를 보장한다. 이를 통해 시스템 관리자가 오류를 확인하고 조치를 할 수 있다. 이는 남아 있는 네트워크 인터페이스와 PID 파일이 시스템 성능에 미치는 부정적인 영향을 줄일 수 있다. 이러한 오류 관리와 자동화된 복구 메커니즘은 시스템의 연속성과 안정성을 보장하고, 자원 관리의 신뢰성을 높이는 데 이바지한다.

#### IV. 사례 연구

#### 참 고 문 헌



그림 3. "Carte start" command



그림 4. "Carte stop" command

그림 3는 "Carte start" 명령어의 수행 결과를 보여주며 그림 4은 "Carte stop" 명령어를 통해 수행된 결과를 보여준다. "Carte create", "Carte start", "Carte stop", "Carte remove"의 명령어를 통해 컨테이너 생애주기를 구현하였으며 특히 "Carte stop" 명령어를 실행하면, 자원 해제 과정을 거쳐 컨테이너가 종료되는 모습을 확인할 수 있다. 이러한 자원 해제 프로세스를 통해, 시스템 자원이 불필요하게 점유되지 않고 시스템의 효율성과 안정성이 유지됨을 확인할 수 있었다.

#### V. 결 론

본 연구에서 제안한 자원 해제 자동화 및 오류 복구 기법은 자원 관리의 효율성을 향상하고, 컨테이너 관리 시스템의 안정성을 높이는 데 이바지할 것이다.

향후 연구로는 본 연구에서 구현한 자원 해제 및 복구 기법을 확장하여, 경량화된 컨테이너 환경에서 오케스트레이션 기능을 통합할 수 있을 것이라 기대한다.

#### 사 사 문 구

"본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음"(2022-0-01068)

- [1] Kim Hye Su, Shim Jun Yong, & Lee Won Sik (2023-02-08). Considerations for Docker-Based Server Virtualization to Design Control System Laboratory. Proceedings of Symposium of the Korean Institute of communications and Information Sciences, Gangwon.
- [2] Sang-Jae Kim, Hyeon-Jun Jang, Hyun-Wook Jin, & Chung-il Yang (2023-12-20). Container-based Management System for Edge Devices with Restricted Network Connection. Korean Institute of Information Scientists and Engineers, Busan.
- [3] JungBin Kim, HyeonJin Yu, JaeHyuck Shin, & SeoYoung Noh (2022-12-20). Performance Analysis of Container Runtime in Kubernetes based Storage Environments. Korean Institute of Information Scientists and Engineers, Jeju.

# 계측기 지시침 검출을 위한 딥러닝 모델 성능 비교

신진홍\*, 전체욱\*\*, 옥승호\*\*\*

## Comparison of Deep Learning Model Performance for Meter Pointer Detection

Jin-Hong Shin\*, Chae-Wook Jeon\*\*, and Seung-Ho Ok\*\*\*

### 요 약

산업 현장에서 사용되는 아날로그 계측기는 주기적인 점검 및 유지 보수가 필수적이다. 이에 따라 이미지 처리와 딥러닝 알고리즘을 활용하여 지시계의 지시침을 자동으로 검출하는 솔루션이 연구되고 있다. 본 논문에서는 지시침 검출에 적합한 모델을 선정하기 위해 네 가지 딥러닝 모델을 비교 분석하였다. 실제 환경에서 수집한 지시계기 이미지 500장을 활용하여 모델에 학습 후 지시침을 검출하였다. 실제 지시침과의 각도 차이를 계산하여 오차가 4° 이하인 이미지를 선별하고 모델들의 정확한 이미지 비율을 비교하였다. 성능 평가 결과, YOLOv8이 92%의 정확도로 가장 우수한 성능을 보였다.

### Abstract

Analog gauges in industrial environments require regular inspection and maintenance. In response, research has been conducted to develop solutions using image processing and deep learning algorithms for automatic detection of gauge needles. This paper presents a comparative analysis of four deep learning models to identify the most suitable one for needle detection. Using a dataset of 500 images of industrial gauges collected from real environments, the models were trained to detect gauge needles. The angle differences between the actual and detected needle positions were calculated, with images showing an error margin of 4° or less selected for analysis. In the performance evaluation, YOLOv8 demonstrated the highest accuracy, achieving a success rate of 92%.

### Key words

indicator instrument, image processing, deep learning, detection, accuracy

## 1. 서 론

아날로그 계측기는 다양한 산업 현장에서 활용되고 있으며 주기적으로 점검 및 유지 보수가 필요하

다. 그러나 수동 점검 방식은 많은 인력과 시간이 소요되는 비효율적인 문제점을 안고 있다[1]. 특히 계측기의 측정 오차가 누적될 시 고도의 정확성을 요구하는 산업 분야에서 심각한 문제를 초래할 수

\* 동의대학교 로봇자동화공학과, 0421skyjh@office.deu.ac.kr

\*\* 동의대학교 인공지능학과, 20246145@office.deu.ac.kr, 20246146@office.deu.ac.kr

\*\*\* 동의대학교 로봇자동화공학과, osh@deu.ac.kr(교신저자)

※ 이 논문은 교육부의 지방대학활성화사업 사업비를 지원받아 작성되었음

※ 본 논문(저서)는 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임

있다[2]. 이러한 문제점을 해결하기 위해 이미지 처리 및 딥러닝 알고리즘을 결합하여 지시계기를 자동으로 판독하고 측정하는 연구가 진행되고 있다 [3][4]. 여러 연구 중 S. Chavan 의 연구[5]에서 색상 변화, 형태학적 처리 그리고 이진화 등의 이미지 전처리를 적용하였고 S. Li 의 연구[6]에서 딥러닝 기반의 알고리즘을 활용하여 기존보다 지시침 검출 성능을 높이는 성과를 보였다.

본 논문은 실제 환경에서 수집한 지시계기 이미지를 활용하여 네 가지 딥러닝 모델의 성능을 비교 분석한다. 실제 값과 예측값 간의 오차가 작은 이미지 비율을 통해 모델별 정확도를 비교하여 지시계기 지시침 검출에 적합한 모델을 선정한다.

## II. 본 론

그림 1은 지시계기 이미지 수집 환경을 나타내며 휴엔텍의 HU205(렌즈 : HULE24310) 카메라 모듈로 30cm 거리에서 지시계기 이미지 500장을 수집하였다. 지시침에 대한 라벨링을 진행한 후 표 1과 같이 데이터 셋을 구성하였다. 이미지의 정확성을 높이기 위해 형태학적 침식 연산을 적용하여 이미지 내의 객체를 축소하고, 작은 객체나 잡음을 제거하여 경계선을 매끄럽게 변형시킨다[7].

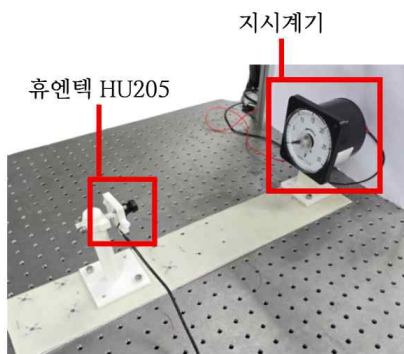


그림 1. 지시계기 이미지 수집 환경  
Fig. 1. Image collection environment

표 1. 지시침 학습 데이터 구성

Table 1. Gauge needle detection training data configuration

train	test	validation
400	50	50

## III. 실험 결과

이미지 검출에 많이 사용되는 네 종류의 딥러닝 모델을 선정하였으며 각각의 특징은 다음과 같다. YOLOv8[8]은 하나의 CNN 패스를 통해 이미지를 처리하며 이전보다 네트워크 아키텍처와 학습 전략을 개선하였다. EfficientDet[9]는 EfficientNet 백본을 사용하고 BiFPN로 객체 탐지 성능을 높인 모델이다. Faster R-CNN[10]은 RPN(Region Proposal Network)을 통해 잠재적인 객체 영역을 파악 후 CNN을 사용하여 정확한 객체 분류와 바운딩 박스를 예측한다. Mask R-CNN[11]은 Faster R-CNN에 인스턴스 마스크 예측 기능을 추가하여 객체 탐지와 동시에 객체의 픽셀 단위 분할을 수행할 수 있는 모델이다.

모델 비교를 위한 검증 방법은 실제 지시침과 예측한 지시침을 직선으로 시각화한 후 두 직선 사이의 각도를 벡터로 계산한다. 오차가 4° 이하일 때 정확하게 지시계기에서 지시침을 검출한 이미지로 간주하고 50장의 이미지 중 지시침 검출 비율이 높은 모델을 선정한다. 그림 2는 실제 지시침(빨간색)과 예측한 지시침(파란색)의 이미지로 시각화한 모습이다.

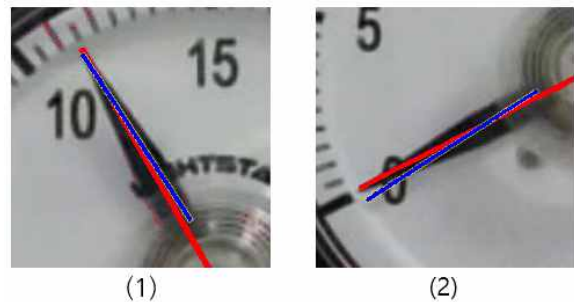


그림 2. 지시침 검출 결과 (1) 오차가 4° 이하인 이미지, (2) 오차가 4° 초과인 이미지

Fig. 2. Detection results for gauge needle (1) An image with an error of 4° or less, (2) An image with an error exceeding 4°

표 2는 모델별 정확하게 예측한 이미지 수를 백분율로 나타낸 결과이며 이를 통해 YOLOv8 모델이 4개의 모델 중 92%로 가장 높은 정확도를 보여



주고 있다.

표 2. 지시침 예측 정확도  
Table 2. Pointer prediction accuracy

Model	YOLO v8	Efficient Det	Faster R-CNN	Mask R-CNN
Accuracy	92%	74%	78%	88%

#### IV. 결 론

본 논문에서는 지시계의 정확한 지시침 검출을 위해 네 가지 딥러닝 모델을 학습시키고, 실제 지시침과 예측한 지시침 간의 각도 차이를 비교하여 지시침 검출에 적합한 최적의 모델을 선정하고자 하였다. 지시침 검출 결과 오차가 4° 이내인 이미지 비율을 비교한 결과 YOLOv8은 92%, EfficientDet은 74%, Faster R-CNN은 78%, Mask R-CNN은 88%로 YOLOv8 모델이 92%로 가장 우수한 성능을 보였다. 이를 통해 모델별로 지시침의 검출 여부와 성공률을 확인하여 위의 모델 중 가장 우수한 모델을 선정할 수 있었다. 향후 연구에서는 지시계의 관측 각도, 이미지 선명도, 왜곡 등의 다양한 환경 변수를 고려한 학습을 통해 YOLOv8 모델을 활용한 지시침 검출 정확도를 향상할 계획이다.

#### 참 고 문 헌

[1] J. Huang et al., "An Automatic Analog Instrument Reading System Using Computer Vision and Inspection Robot", in Proc IEEE Trans. Instrum. Meas., vol. 69.9, pp. 6322-6335, 2020.

[2] J. Liu et al., "Novel method of automatic recognition for analog measuring instruments", in Proc. 2015 6th Int. Conf. Manufacturing Science and Engineering, Atlantis Press, pp. 67-75, 2015.

[3] J. Leon-Alcazar et al., "Learning to read analog gauges from synthetic data", in Proc. IEEE/CVF Winter Conf. Applications of Computer Vision, pp. 8616-8625, 2024.

[4] J. Wang et al., "Automatic reading system for

analog instruments based on computer vision and inspection robot for power plant", in Proc. 2018 10th Int. Conf. Modelling, Identification and Control, IEEE, pp. 1-6, 2018.

[5] S. Chavan et al., "High precision analog gauge reader using optical flow and computer vision", in Proc. 2022 IEEE Int. Conf. Electro Information Technology, IEEE, pp. 171-175, 2022.

[6] S. Li et al., "A visual identification method of analog instrument panel based on Faster R-CNN", in Proc. 2022 34th Chinese Control and Decision Conference, IEEE, pp. 6118-6123, 2022.

[7] K.A.M. Said et al., "Analysis of image processing using morphological erosion and dilation", Journal of Physics: Conference Series, vol. 2071, no. 1, pp. 012033, 2021.

[8] M. Safaldin et al., "An improved YOLOv8 to detect moving objects", IEEE Access, vol. 12, pp. 59782-59806, 2024.

[9] M Tan et al., "EfficientDet: Scalable and Efficient Object Detection", in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit, pp. 10781-10790, 2020.

[10] S Ren et al., "Faster R-CNN: Towards real-time object detection with region proposal networks" in Proc. IEEE Trans. Pattern Anal. Machine Intelligence, vol. 39.6, pp. 1137-1149, 2016.

[11] K. He et al., "Mask R-CNN," in Proc. IEEE Int. Conf. Computer Vision, pp. 2961-2969, 2017.

# 전이 학습을 활용한 병원 전화 상담 전사 성능 향상을 위한 Whisper 모델 파인튜닝

윤종인\*, 윤웅창\*\*

## Fine-tuning Whisper for enhanced transcription of hospital telephone consultations using transfer learning

Jong-In Yun\* and Woongchang Yoon\*\*

### 요 약

환자의 초기 진단과 처방에 있어서 전화 상담은 중요한 역할을 하며 이를 관리하고 기록하는 것은 다양한 문제점으로 인하여 쉽지 않은 일이다. 본 연구에서는 Whisper 모델을 기반으로 하여 병원 전화 상담 내용을 자동으로 전사하고 관리하는 시스템을 개발하고자 하였다. 기존의 사전 학습된 Whisper 모델을 실제 의료진과 환자의 음성 데이터 세트를 이용한 전이학습을 통해 파인튜닝 모델을 얻었다. 병원 전화 상담을 위해 최적화된 파인튜닝 모델은 기존 모델과 비교하여 감소된 character error rate를 보였으며 이것은 통계적으로 유의함을 확인하였다. 제안된 모델을 통해 병원 전화 상담 음성을 신뢰성 높은 텍스트로 변환함으로써 의료 커뮤니케이션 기록 관리의 정확성과 효율성을 향상시킬 수 있다.

### Abstract

Telephone consultations play an important role in the initial diagnosis and prescription for patients, yet managing and documenting these consultations is challenging due to various issues. This study aimed to develop a system based on the Whisper model to automatically transcribe hospital telephone consultations. We obtained a fine-tuned model by applying transfer learning to the pre-trained Whisper model using actual voice datasets from healthcare professionals and patients. The fine-tuned model, optimized for hospital telephone consultations, demonstrated a reduced character error rate compared to the baseline model, with statistically significant improvements. The proposed model enables the conversion of hospital telephone consultations into highly reliable text, thereby enhancing the accuracy and efficiency of managing medical communication records.

### Key words

Whisper model, Automatic transcription, Healthcare documentation, Transfer learning, Fine-tuning

---

\* 경상국립대학교 컴퓨터과학부, godxxy@gnu.ac.kr

\*\* 경상국립대학교 컴퓨터공학과/바이오의료빅데이터학과, neobits@gnu.ac.kr

## 1. 서 론

전화 상담은 일차 의료에서 환자와의 중요한 소통 수단으로, 진료 예약, 검사 결과 확인, 간단한 증상에 대한 조언 등에 활용된다 [1]. 그러나, 전화 상담 내용의 문서화는 매우 미흡한 실정으로 연구에 따르면 상담 내용의 절반 정도만 기록되고 있다 [1]. 이는 정보 누락, 불명확성으로 이어질 수 있으며, 법적 분쟁의 소지가 있다 [1]. 또한, 의료진이 녹음된 내용을 청취하고 기록하는 것은 많은 시간과 노력이 소모되는 비효율적인 작업이다 [2].

OpenAI의 Whisper는 680,000시간의 다국어 음성 데이터로 학습된 최신 음성인식 모델로써 다양한 도메인에서 뛰어난 성능을 보이며, 영어뿐만 아니라 98개 언어를 지원하는 장점이 있다 [3]. Whisper는 대규모 음성 데이터로 사전 학습된 강력한 음성인식 모델이지만, 의료 분야 특유의 전문용어와 언어적 특성으로 인해 모델을 그대로 적용하기에는 한계가 있다 [3]. 때문에, 실제 도메인에 최적화된 모델로 성능을 향상시키기 위해서는 해당 도메인의 데이터를 이용하여 전이 학습 (Transfer learning)을 통한 파인튜닝 모델을 얻는 것이 권장된다 [3].

이에 본 연구에서는 Whisper 모델을 병원 전화 상담 데이터를 이용하여 전이 학습 방법을 통해 전사 성능을 향상시킨 파인튜닝 모델 기반의 병원 전화 상담 녹취 자동 관리 시스템을 제안한다.

## II. 실험 방법

### 2.1 병원 전화 상담 녹취 자동 관리 시스템의 개요

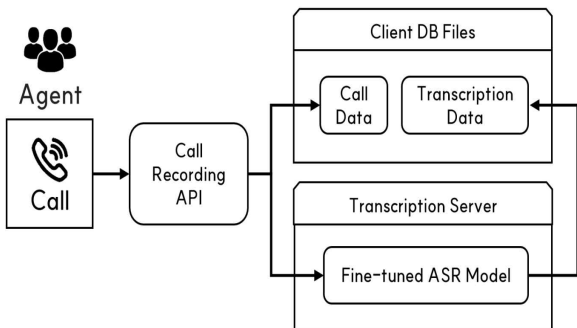


그림 1. 전화 상담 녹취 자동 관리 시스템 구조도  
Fig. 1. Diagram of the automatic transcription management system for telephone consultation recordings

제안하는 시스템은 다음과 같은 과정으로 구성된다. 사용자가 병원에 전화를 걸면 통신사에서 제공하는 전화 기록 API를 이용해 서버에 통화 내역을 저장한다. 저장된 음성기록은 전사 모델을 통해서 음성 데이터를 텍스트로 변환한다. 모델에 의해 전사된 내용은 파이프라인에 의해 구조화되어 진료 기록 데이터베이스에 자동으로 저장된다 (그림 1).

### 2.2 데이터 세트 특성과 전처리

본 연구에서는 의료진과 환자의 다양한 대화가 포함된 총 1,300시간의 비대면 진료 음성 데이터 세트를 AI-HUB에서 수집하여 사용하였다 [4]. 해당 데이터 세트는 의료진 (의사와 간호사)의 발화와 환자의 발화로 구분되어 있으며, 각 그룹별 특성을 반영하기 위해서 그룹별 데이터를 구분하여 전사하고 분석하였다 (표 1).

표 1. 병원 전화 상담 음성 데이터 세트  
Table 1. hospital telephone consultations voice data set

내 용	데이터 규모
의사 음성	300시간 (약300,000문장)
간호사 음성	400시간 (약 400,000문장)
환자 음성	600시간 (약 450,000문장)
의료진 문장 말뭉치	33,000 단어
환자 문장 말뭉치	150,000 단어

모델의 효율적인 학습을 위해서 전처리를 수행하여 모든 음성 파일을 16kHz 샘플링 레이트 및 16비트 깊이로 변환하였다. 또한, 물음표, 온점과 같은 특수문자는 평가 신뢰도에 영향을 미칠 수 있으므로 제거하였다.

### 2.3 Whisper 모델의 전이 학습

일반적으로 Whisper 모델의 전이 학습은 저자원 언어나 특정 도메인에서도 3000-4000 스텝 정도면 충분한 것으로 알려져 있으며, 본 연구에서 전이 학습은 총 2,800 스텝을 수행하였다 [5].

학습률은 사전학습 모델의 학습된 가중치를 크게 변화시키지 않으면서도 새로운 도메인에 적응할 수

있는 적절한 값으로 알려져 있는  $1e-5$ 로 설정 하였다 [6]. 또한, Transformer 기반 모델 전이 학습 시 일반적으로 사용되는 방법인 초기 500 스텝 동안 학습률을 점진적으로 높이는 워밍업을 적용하여 학습 초기에 급격한 가중치 변화를 방지하고 안정적인 학습이 이뤄지도록 하였다 [7].

전이 학습 시 배치의 크기는 GPU 메모리 제약을 최소화한 값인 16과 8을 학습과 평가단계에서 각각 사용하였다 [8]. 그래디언트 누적은 1로 설정하였으며, GPU 메모리의 활용성을 최대화할 수 있는 16비트 부동 소수점 연산 (FP16)을 사용하였다 [5].

평가 주기는 일반적으로 권장되는 전체 스텝의 1/4인 700 스텝마다 평가를 수행하였고, 25 스텝마다 로깅을 하여 진행 상황을 모니터링 하였다 [6]. 생성된 텍스트의 최대 길이는 225 토큰으로 제한하였다 [6].

전이 학습 종료 후 CER이 가장 낮은 모델을 최종적으로 선택하였으며, 이를 통해 과적합을 방지하고 일반화 성능을 높이고자 하였다 [6].

#### 2.4 기본 모델과 파인튜닝 모델의 성능 평가

기본 모델과 전이 학습을 통해서 생성한 파인튜닝 모델의 성능 평가를 위한 성능 지표로는 WER (Word Error Rate)과 CER (Character Error Rate)을 사용하였다. WER은 인식된 문장과 정답 문장 간의 단어 삭제, 삽입, 대체와 같은 단어 단위 오류율을 평가하며, 이는 문장 단위의 인식 성능을 파악하는데 유용하다.

하지만, 한국어는 교착어 (침가어)로서 조사나 어미 사용의 다발성과 형태소 단위의 다변화로 인하여 단어 경계가 모호하다. 때문에, 한국어 음성 인식 성능 측정에서는 단어의 오류를 측정하는 WER보다 문자 단위의 오류인 CER이 성능을 보다 정확하게 반영한다 [9].

WER과 CER 모두 낮은 값을 가질 수록 전사 성능이 높음을 나타내며 아래와 같이 정의된다.

$$WER = \frac{D+I+S}{N}, CER = \frac{D_C+I_C+S_C}{N_C}$$

WER에서  $D$ 는 삭제된 단어 수,  $I$ 는 삽입된 단어 수,  $S$ 는 대체된 단어 수,  $N$ 은 전체 단어 수를 의미하고 CER에서  $D_C$ 는 삭제된 문자 수,  $I_C$ 는 삽입된 문자 수,  $S_C$ 는 대체된 문자 수,  $N_C$ 는 전체 문자 수를 의미한다.

#### 2.5 기본 모델과 파인튜닝 모델의 성능 차이 검증

Whisper 기본 모델과 파인튜닝 모델의 성능은 오류율 (WER, CER)을 기준으로 세 발화 그룹인 간호사, 의사, 환자로 나누어 분석하였다.

전이 학습 전후 모델 간의 성능 차이는 통계적 유의성을 통해서 확인하였다. 성능 평가를 위해 구한 CER과 WER 값의 정규성 검정은 Q-Q Plot과 Shapiro-Wilk 검정을 이용하였다. 정규성 검정 결과 정규성을 만족하지 않아서 Wilcoxon Signed-Rank 검정을 통해서 두 그룹의 성능 차이를 통계적으로 검증하였다. 통계적 유의수준은  $p\text{-value} < 0.05$ 를 채택하였다.

### III. 실험 결과

#### 3.1 전이 학습 전후 모델 간의 발화자별 성능 비교

전이 학습 이전의 기본 모델과 음성 데이터를 통해 전이 학습을 진행한 후의 파인튜닝 모델 간의 성능은 CER과 WER을 통해서 비교하였다.

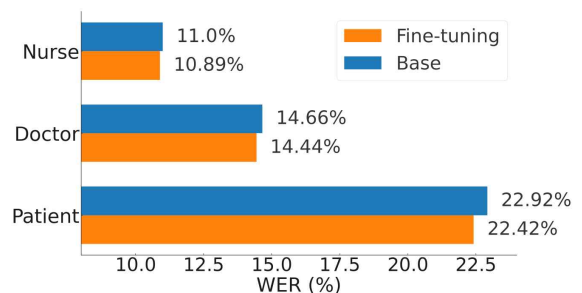


그림 2. 전이 학습 전후 모델 간의 WER 비교  
Fig. 2. The comparison of WER between models before and after transfer learning

간호사 그룹에서 기본 모델의 평균 WER은 11.0%였으나, 파인튜닝 모델은 10.89%로 감소한 것을 확인하였으며, 의사 그룹의 평균 WER은 기본 모델의 14.66%에서 파인튜닝 모델의 14.44%로 감소한 값을 보였다. 환자 그룹에서의 WER 또한 위의 두 그룹에서 보여준 결과와 유사하게, 전이 학습을 통해서 22.92%에서 22.42%로 감소한 것을 확인하였다 (그림 2).

기본 모델과 파인튜닝 모델 간의 CER을 비교한 결과, WER 보다 모델 간의 성능 차이가 더 크게 나타남을 확인하였다. 간호사 그룹에서 기본 모델의 평균 CER은 2.50%였으나, 파인튜닝 모델은 2.44%로 감소하였으며, 의사 그룹의 평균 CER은 기본 모델의 3.35%에서 파인튜닝 모델의 3.20%로 감소하였다. 마지막으로 환자 그룹에서의 CER은 전이 학습 전 모델의 5.32%에서 학습 이후 모델의 4.98%로 감소한 것을 확인하였다 (그림 3).

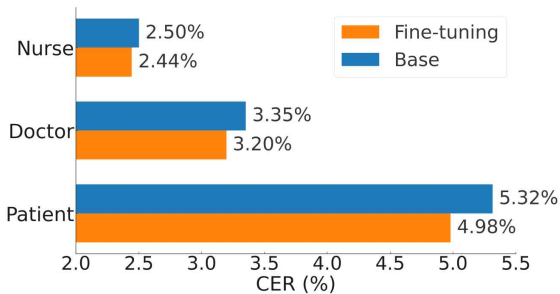


그림 3. 전이 학습 전후 모델 간의 CER 비교

Fig. 3. The comparison of CER between models before and after transfer learning

환자 그룹에서의 전이 학습을 통한 WER과 CER의 차이는 간호사와 의사 그룹보다 더 큰 경향성을 보였다. 이는 환자 그룹은 의료진보다 비구조화된 대화 방식의 사용, 비표준적 발음, 억양의 다양성 때문에 보이는 현상으로 유추할 수 있다 [10].

### 3.2 통계분석을 통한 모델 간의 성능 차이 검증

전이 학습 이전과 이후의 모델 간의 성능 차이를 통계적으로 분석하기 위해서, 먼저 정규성 검정을 수행하였다. 이를 위해 Q-Q Plot과 Shapiro-Wilk 검

정을 사용하였다.

Q-Q Plot을 통해서 파인튜닝 모델의 CER 값이 대부분 0 근처에 집중되어 분포가 비대칭적임을 확인하였으며 결과적으로 CER 값의 분포가 비정규성을 나타냄을 확인하였다 (그림 4).



그림 4. 파인튜닝 모델의 CER에 대한 Q-Q Plot  
Fig. 4. Q-Q Plot for CER on fine-tuning model

다른 정규성 검정인 Shapiro-Wilk 검정 결과에서도 두 모델 모두 p-값이 0.05 이하로 나와 정규성을 만족하지 않는 것으로 나타났다 (표 2).

표 2. 각 모델의 CER에 대한 Shapiro-Wilk 검정 결과  
Table 2. Shapiro-Wilk test results for CER of each model

Model	W	p-value
Base	0.6203	< 0.0001
Fine tuning	0.6675	< 0.0001

비모수적 검정의 대표적인 Wilcoxon Signed-Rank 검정 결과, 전체 데이터 세트에서 WER과 CER의 경우 모두 매우 작은 p-value를 보여서 통계적으로 유의함을 보였다. 각 그룹별 비교에서 간호사 그룹을 제외하고 의사와 환자 그룹 모두 WER과 CER의 경우 통계적으로 유의함을 확인하였다 (표 2).

다른 비모수적 비교 방법인 부호 검정 (Sign Test)를 통해서도 간호사 그룹을 제외한 의사와 환자 및 모든 데이터 세트의 경우에도 WER과 CER 값의 차이가 통계적으로 유의함을 만족하였다 (표 3).

표 3. Wilcoxon signed-rank 검정을 통한 모델 간 성능 차이 비교

Table 3. Comparison of performance between each model using the Wilcoxon Signed-Rank test

Method	Group	p-value
WER	Nurse	0.1573
	Doctor	0.0050
	Patient	0.0273
	Overall	0.0002
CER	Nurse	0.1797
	Doctor	0.0004
	Patient	0.0180
	Overall	< 0.0001

표 4. 부호 검정을 통한 모델 간 성능 차이 비교

Table 4. Comparison of performance between each model using the Sign test

Method	Group	P-Value
WER	Nurse	0.5
	Doctor	0.0020
	Patient	0.0313
	Overall	< 0.0001
CER	Nurse	0.5
	Doctor	< 0.0001
	Patient	0.0157
	Overall	< 0.0001

위의 결과들을 종합하여 볼 때, 기본 모델에서 전이 학습을 통해서 얻은 튜닝모델의 성능 개선은 통계적으로 유의한 차이가 있음을 확인하였다.

### 3.3 학습 과정에서의 손실 변화 분석

모델의 학습 과정에서의 성능을 평가하고 과소적합 (Under-fitting)과 과대적합 (Over-fitting)의 발생을 확인하기 위해서 학습 손실 (Training Loss)과 테스트 손실 (Test Loss)의 변화를 분석하였다. 초기 700 스텝까지 학습 및 테스트 손실이 급격하게 감소하다가 이후에는 변화가 미미한 것을 확인하였다. 이를 통하여 과소적합이 일어나지 않고 충분히 모델이 학습되었음을 확인하였다 (그림 5).

이후 학습이 진행됨에 따라 학습 손실과 테스트 손실의 차이가 점차적으로 증가하는 것을 보였으며, 과대적합의 발생을 피하기 위해서 2800 스텝까지만 학습을 진행하였다 (그림 5).

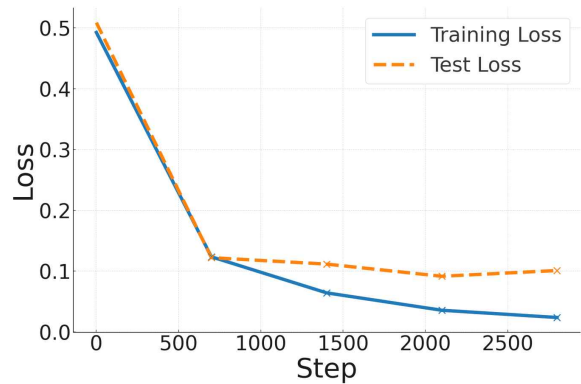


그림 5. Step에 따른 Training Loss 및 Test Loss 변화  
Fig. 5. Training Loss and Test Loss Change with Step

## IV. 결 론

본 연구에서는 Whisper 모델의 전이 학습을 통해서 얻은 파인튜닝 모델이 한국어 음성인식에서 전사 성능을 개선한 것을 확인하였다. 특히, 한국어와 같은 교착어에서는 CER이 WER보다 좋은 성능을 보이는 평가지표라는 것을 확인하였다. 또한, 의료진 그룹보다는 환자 그룹의 데이터 세트에서 전사 정확도를 크게 향상시킨 것을 보였다.

향후 연구에서는 더 늘어난 대용량 음성 데이터의 확보와 함께 다양한 환경에서 만들어진 데이터를 사용하여 정확성과 다양성이 향상된 모델로 발전시킬 수 있을 것이다.

우리의 연구를 통해서 음성 인식 기술의 활용성을 높이고, 의료 현장에서의 실질적인 활용 가능성을 더욱 강화할 수 있는 기반을 마련할 것이다. 더 나아가 우리의 연구가 의료진의 기록 업무 부담을 경감하고 데이터 활용도를 제고함으로써 의료 커뮤니케이션 및 기록 관리 프로세스 혁신에 기여할 것을 기대한다.

## 참 고 문 헌

- [1] Car, Josip, and Aziz Sheikh. "Telephone consultations." *Bmj* 326.7396 (2003): 966-969.
- [2] Khan, Muhammad Naseer Babar. "Telephone consultations in primary care, how to improve their safety, effectiveness and quality." *BMJ Open Quality* 2.1 (2013): u202013-w1227.

- [3] Radford, Alec, et al. "Robust speech recognition via large-scale weak supervision." International conference on machine learning. PMLR, 2023.
- [4] <https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=208>
- [5] Liu, Yunpeng, Xukui Yang, and Dan Qu. "Exploration of Whisper fine-tuning strategies for low-resource ASR." EURASIP Journal on Audio, Speech, and Music Processing 2024.1 (2024): 29.
- [6] Chang, Jungwon, and Hosung Nam. "Exploring the feasibility of fine-tuning large-scale speech recognition models for domain-specific applications: A case study on Whisper model and KsponSpeech dataset." Phonetics and Speech Sciences 15.3 (2023): 83-88.
- [7] Ma, Hao, et al. "Extending Whisper with prompt tuning to target-speaker ASR." ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2024.
- [8] Harrathi, Faycel, David van Leeuwen, and Djoerd Hiemstra. "Fine-tune Whisper for Speaker Recognition" (2024).
- [9] James, Jesin, and Deepa P. Gopinath. "Advocating Character Error Rate for Multilingual ASR Evaluation." arXiv preprint arXiv:2410.07400 (2024).
- [10] Nigmatulina, Iuliia, Tannon Kew, and Tanja Samardzic. "ASR for Non-standardised Languages with Dialectal Variation: the case of Swiss German." Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects. 2020.

# 오픈소스 기반 SW개발자 역량 진단 방법 연구

이예영\*, 김장원†

## A Study on Competency Assessment Methods for Open-Source Software Developers

Yeyeong Lee\* and Jangwon Gim†

### 요 약

디지털 기술의 확산과 IT 산업의 성장으로 SW 개발자의 역량을 체계적이고 정확하게 평가할 필요성이 커지고 있다. 그러나 기존 평가 방식은 주로 정량적 지표나 주관적 판단에 치우쳐 있어, 개발자의 다양한 역량을 균형 있게 반영하지 못하는 한계가 있다. 본 논문에서는 GitHub의 정량적 데이터와 설문 기반의 정성적 데이터를 통합한 O-STAR 모델을 통해 SW 개발자의 기술력, 협업 능력, 문제 해결 능력을 다각적으로 평가하는 아키텍처를 제안한다. 이 모델은 지시적 분석과 SW 전공 역량 분석을 통해 개발자에게 맞춤형 성장 방향을 제시하여, 실질적인 역량 강화를 지원하고자 한다.

### Abstract

As digital technology becomes widespread and interest in the IT industry grows, the need for systematic and accurate evaluation of software developers' competencies has increased. However, traditional evaluation methods often rely on quantitative metrics or subjective judgments, which limits the comprehensive assessment of developers' diverse skills. This study proposes the O-STAR model, an architecture that integrates quantitative data from GitHub with survey-based qualitative data to evaluate SW developers' skills, collaboration abilities, and problem-solving competencies from multiple perspectives. The model employs prescriptive analytics and software-specific competency analysis to provide personalized development recommendations, supporting the practical enhancement of developer capabilities.

### Key words

O-STAR Model, Quantitative and Qualitative Integration, prescriptive analysis

## 1. 서 론

디지털 전환의 가속화와 함께 소프트웨어 개발은 다양한 산업에서 핵심적인 역할을 차지하고 있으며

\* 국립군산대학교 소프트웨어학과 (lyy21@kunsan.ac.kr)

† 국립군산대학교 소프트웨어학과 (교신저자) (jwgim@kunsan.ac.kr)

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065).



[1], 이에 따라 SW 개발자의 역량을 체계적으로 평가하고 강화할 필요성이 높아지고 있다. SW 개발자는 고품질 소프트웨어를 신속하고 안정적으로 개발하기 위한 기술적 능력뿐만 아니라, 협업, 문제 해결, 적응력 등 다각적인 역량이 요구된다[2]. 그러나 기존의 개발자 역량 평가는 주로 커밋 횟수나 코드 라인 수와 같은 단순 정량적 지표에 의존하거나 평가자의 주관적 판단에 치우쳐 있어, 개발자의 실제 역량을 종합적이고 객관적으로 파악하는 데 한계가 있다[3, 4]. 이러한 한계를 보완하기 위해 개발자의 역량을 정확하고 종합적으로 평가할 수 있는 방법론이 필요하다[5, 6]. 본 논문에서는 오픈소스 데이터를 기반으로 SW 개발자 역량을 진단하는 시스템 아키텍처를 제안한다. 오픈 SW를 통해 획득가능한 SW 개발 활동 및 설문 등을 통한 정량적, 정성적 데이터를 통합하여 개발자의 기술력, 협업 능력, 문제 해결 능력을 객관적으로 진단할 수 있는 O-STAR 모델을 제안한다. 이때 5W1H 방식의 지시적 분석(Prescriptive Analytics)[7]을 통해 개발자가 목표를 달성하기 위한 행동 지침을 SW 개발자에게 제공한다. 또한, 8대 SW 전공 역량 분석을 통해 개발자의 구체적 역량 수준을 진단하고 유형을 분류하여 맞춤형 성장 방향을 제시함으로써, 실질적인 역량 강화를 지원하고자 한다.

## II. 제안 모델

### 2.1 O-STAR 모델 개요

본 논문에서는 SW 개발자 역량 진단을 위해 정량적·정성적 평가를 통합한 O-STAR(Open-Source Talent Assessment and Recommendation) 모델을 제안한다. 그림 1은 제안된 모델의 전체 개요를 나타낸다. O-STAR 모델은 GitHub을 비롯하여 오픈된 SW 허브 및 리포지터리에 저장된 기초 통계를 활용하여, 개발자의 기술력, 협업 능력, 문제 해결 능력 등을 객관적으로 평가할 수 있다. 이를 위해 O-STAR 모델에서는 개발자 특성을 S(Skill)는 Speed와 Careful, T(Technology)는 Depth와 Breadth, A(Adaptability)는 Agile과 Focus, R(Responsibility)는

Reliable과 Initiative로 각각 대조되는 특성으로 구분한 총 16가지 유형으로 구성한다.

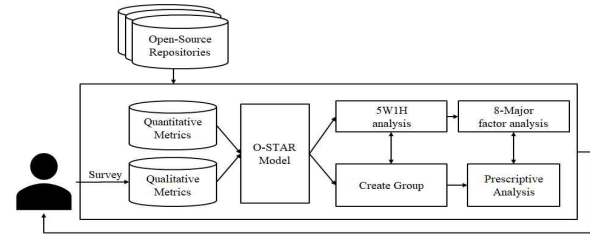


그림 1. O-STAR 모델 개요도

Fig. 1. Overview of the proposed O-STAR model

이렇게 도출된 SW 개발자 유형과 SW 개발자의 정성 평가 결과를 기반으로 5W1H 분석과 8대 SW 전공 역량 분석을 통해 심층 평가를 수행한다. 최종적으로, 유사한 역량을 갖춘 개발자 그룹으로 군집화를 통해 개별 개발자가 희망하는 개발자 그룹으로 성장하기 위한 조건, 방법, SW 개발 활동 등을 제시하는 지시적 분석 결과를 제공한다.

### 2.2 모델 아키텍처의 구성 요소

지시적 분석 모듈은 개발자 역량을 5W1H 기반 분석 방법론[7]을 적용하여 평가한다. 'Who'는 개발자의 롤 모델 그룹이 누구인지 도출하고, 'What'은 개발자의 SW 개발 관련 장단기 목표 설정 정보를 도출, 'When'은 시간 흐름에 따른 SW 개발 성과 추적을, 'Where'는 개발자의 작업 환경과 협업 관계 분석을, 'Why'는 개발자의 내적 동기와 실제 행동 패턴 비교를, 'How'는 목표 달성을 위한 현재 작업 방식과 성장 전략 수립을 다룬다. 8대 SW 전공 역량 분석 모듈은 O-STAR 모델과 지시적 분석 결과를 통합하여 소프트웨어 전공을 학습하는데 필요한 여덟 가지 핵심 요소에 대한 개발자의 역량 수준을 진단한다. 또한, 그룹 생성 모듈은 유사한 개발자 유형을 군집화하고 다양한 개발자 그룹을 추천한다. 마지막으로, 지시적 분석 모듈은 진단 결과를 기반으로 개발자별 맞춤형 성장 방향을 제시한다. 부족한 역량을 보완하기 위한 구체적인 활동을 추천하고, 유사 그룹 내 우수 사례를 기반으로 개선점을 도출한다.

### III. 실험

#### 3.1 실험 데이터

본 실험은 O-STAR 모델의 검증을 위해 약 1000 명의 SW 개발자에 대한 디미 데이터를 생성하여 사용하였다. O-STAR 모델에서 정의한 총 16가지 STAR 유형별로 정상 데이터 900개와 이상치 데이터 100개로 구분하였다.

#### 3.2 실험 결과

O-STAR 모델의 16가지 유형 분류 결과는 그림 2 와 같다. 각 점들은 개별 개발자 데이터를 나타내며, 색상은 STAR 유형을 구분한다. 실험 결과, 각 유형이 명확한 경계를 형성하는 것을 확인하였다. 특히 대조되는 특성 유형은 반대 영역에 분포하는 것을 보였으며, 이상치 데이터 군집 경계에 위치하는 것을 확인하였다.

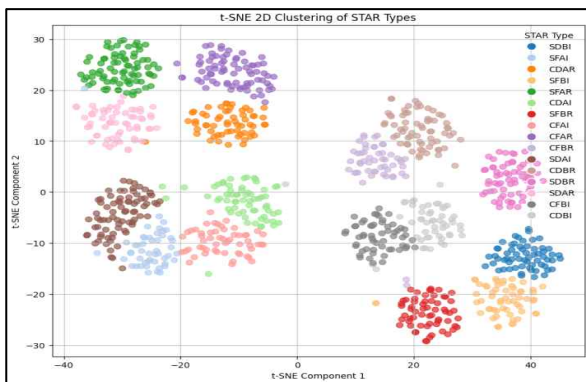


그림 2. O-STAR 모델 기반 개발자 유형 군집화 결과  
Fig. 2. Clustering Results of Developer Types Based on O-STAR Model

### IV. 결론

본 논문에서는 오픈소스 기반 SW 개발자 역량 진단을 위한 O-STAR 모델 아키텍처를 제안하였다. 제안된 아키텍처는 정량적 GitHub 데이터와 정성적 설문 데이터를 통합하여 개발자 역량을 다면적으로 진단한다. O-STAR 모델과 지시적 분석을 통한 체계적 평가 방법은 기존 기초적인 정량평가 방법의

한계를 보완할 수 있다. 샘플 데이터를 활용한 실험 결과, O-STAR 모델은 SW 개발자의 다양한 역량 요소를 효과적으로 구분할 가능성을 보였으며, 제안된 아키텍처의 구현 방향성을 제시하였다.

향후 연구에서는 제안한 아키텍처를 기반으로 실제 시스템을 구현하고, GitHub 데이터를 수집하여 모델을 검증할 필요가 있다. 또한, O-STAR 모델의 유형 분류 알고리즘을 정교화하고 지시적 분석 결과를 효과적으로 제공하는 사용자 인터페이스 개발이 요구된다.

### 참고 문헌

- [1] Popoola, O. A., Adama, H. E., Okeke, C. D., & Akinoso, A. E. (2024). Conceptualizing agile development in digital transformations: Theoretical foundations and practical applications. *Engineering Science & Technology Journal*, 5(4), 1524-1541.
- [2] Liang, J. T., Zimmermann, T., & Ford, D. (2022, November). Understanding skills for OSS communities on GitHub. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (pp. 170-182).
- [3] Bergstra, J., Delen, G., & van Vlijmen, B. (2011). Outsourcing competence
- [4] Borges, H., Valente, M. T., Hora, A., & Coelho, J. (2015). On the Popularity of GitHub Applications: A Preliminary Note. *arXiv preprint arXiv:1507.00604*.
- [5] 한상곤, 류사오, 우균. (2022). 교육 수준 및 목적을 기반으로 한 온라인 저지 시스템의 평가 방법 및 분류. *정보과학회 컴퓨팅의 실제 논문지*, 28(10), 487-492, 10.5626/KTCP.2022.28.10.487
- [6] 김수연, 조희권, 이승아, 박지선. (2014). 정성적 평가의 평가준거와 전공역량간의 관계 연구. *입학전형연구*, 3(0), 35-64.
- [7] 김장원, 정한민, 정도현, 송사광, 황명권. (2015). 지시적 분석 기반 역량 강화 시스템. *정보과학회 컴퓨팅의 실제 논문지*, 21(1), 46-51.

# 스마트 글래스 및 음성 인식을 사용한 모바일 로봇의 원격 제어 시스템 구현

김예찬\*, 권도현\*, 강승욱\*, 전채욱\*\*, 옥승호\*\*\*

## Implementation of a Remote Control System for Mobile Robots Using Smart Glasses and Voice Recognition

Ye-Chan Kim\*, Do-Hyun Kwon\*, Seung-Uk Kang\*, Chae-Uk Jeun\*\*, Seung-Ho Ok\*\*\*

### 요 약

스마트 글래스는 다양한 콘텐츠와의 접목을 통해 산업 현장에서 작업자의 안전성과 업무 효율성을 향상시키기 위한 웨어러블 디바이스로 발전되고 있다. 본 논문에서는 스마트 글래스로 원격 로봇의 실시간 상황을 모니터링하며 음성 명령으로 모바일 로봇을 제어하는 시스템을 구현한다. Flask 서버를 이용해 실시간 웹 스트리밍을 구현하여 원격에서 실시간으로 상황을 확인할 수 있으며, GPT-3.5-Turbo 기반 음성 챗봇을 통해 명령값을 전송함으로써 모바일 로봇을 원격 제어한다.

### Abstract

Smart glasses are evolving into wearable devices designed to enhance worker safety and operational efficiency in industrial settings by integrating with various content. This paper implements a system that monitors the real-time status of remote robots using smart glasses and controls a mobile robot through voice commands. By utilizing a Flask server for real-time web streaming, users can check the situation remotely. The system remotely controls the mobile robot by sending command values through a voice chatbot based on GPT-3.5-Turbo.

### Key words

smart glass, voice assistant, robot control, flask, GPT

## 1. 서 론

스마트 글래스는 실시간 정보 제공, 원격 협업, 모니터링 등의 기능을 통해 작업 효율성을 크게 향

상시킬 수 있어 제조업, 물류, 의료, 교육 등 다양한 분야에서 응용되고 있다. 특히 산업 현장에서 실시간 모니터링을 통해 작업자의 안전성과 업무 효율성을 높일 수 있어 원격 제어 및 모니터링 시

\* 동의대학교 로봇자동화공학과, 20222289@office.deu.ac.kr, 20202253@office.deu.ac.kr, 20202316@office.deu.ac.kr

\*\* 동의대학교 인공지능학과, 20246145@office.deu.ac.kr

\*\*\* 동의대학교 로봇자동화공학과, osh@deu.ac.kr(교신 저자)

※ 이 논문은 교육부의 지방대학활성화사업 사업비를 지원받아 작성되었음

※ 본 논문(저서)는 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구임

시스템에 대한 수요가 증가하고 있다[1].

본 논문은 GPT3.5-Turbo를 기반으로 음성 챗봇을 구현하여 이를 모바일 로봇에 적용하였고, Flask 웹 스트리밍을 구축하여 스마트 글래스를 통해 모바일 로봇을 원격으로 명령 제어하며, 모니터링하는 시스템을 구현하였다.

## II. 본 론

그림 1은 원격 모니터링과 음성 명령을 통한 로봇제어 시스템 전체 구조를 나타내며, 그림 2는 모바일 로봇과 스마트 글래스의 구성 요소를 나타낸다.

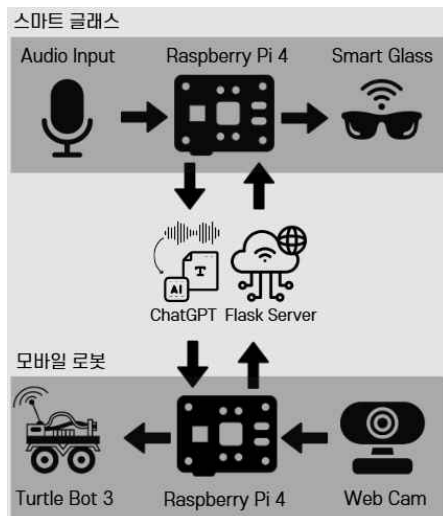


그림 1. 원격 모니터링과 음성 명령을 통한 로봇제어 시스템

Fig. 1. Operation Structure of Command System through Remote Monitoring and Voice

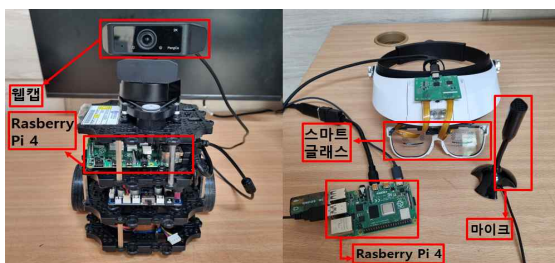


그림 2. 모바일 로봇과 스마트 글래스의 구성 요소  
Fig. 2. Components of a mobile robot and smart glasses.

표 1은 모바일 로봇을 제어하기 위한 로봇의 구성 요소이다. 표 1에서 마이크로 입력된 음성을 명령 값으로 변환하기 위해 GPT-3.5 Turbo를 사용하였다. GPT-3.5 Turbo는 GPT-3의 개선된 버전으로, 더욱 효율적인 처리와 더 나은 대화 생성 능력을 제공한다[2][3].

표 2는 모바일 로봇의 구성 요소를 나타낸다. 모바일 로봇에는 임베디드 보드와 웹캠이 연결되어있다. Flask를 활용해 웹소켓(WebSocket) 통신 방식의 실시간 스트리밍 서버를 구축하여 원격 모니터링 시스템을 구현하였다. 사용자가 마이크로 모바일 로봇에 음성 명령을 내리면, 모바일 로봇은 액션을 수행한다. 마이크로 입력된 음성 명령은 STT(Speech to Text)를 거쳐 GPT-3.5 Turbo에서 명령 값에 맞는 행동으로 ‘전진’, ‘후진’, ‘좌회전’, ‘우회전’, ‘정지’를 생성하고, 행동에 대한 명령 값을 모바일 로봇으로 전달한다. 명령 값을 받은 모바일 로봇은 linear x, angular z 값을 조정하며 작동한다.

표 1. 스마트 글래스의 구성 요소

Table 1. Components of smart glasses

	구성
Smart Glass	FrontiAr lens
Embedded Board	Raspberry Pi 4 Model B
음성 입력	Mic

표 2. 모바일 로봇의 구성 요소

Table 2. Components of mobile robots

	구성
Mobile Robot	Turtlebot 3
Embedded Board	Raspberry Pi 4 Model B
Camera	Web Cam

그림 3은 Flask 서버를 활용한 실시간 웹 스트리밍 구동 환경을 나타낸다. 모니터로 출력되는 스트리밍 영상은 모바일 로봇의 웹캠의 영상으로써 스마트 글래스를 통해 Flask 서버에 접속하여 실시간 작업 영상을 확인할 수 있다.



그림 3. Flask 서버를 활용한 실시간 웹 스트리밍 구동 환경  
Fig. 3. Real-time web streaming environment using a server

그림 4는 스마트 글래스로 로봇의 영상을 원격 모니터링 하고, 마이크로 음성 명령을 통해 모바일 로봇을 제어하는 예시를 나타낸다. 이를 통해 작업자는 현장 상황을 스마트 글래스로 모니터링 하며 음성 명령으로 로봇을 원격 제어할 수 있다.



그림 4. 원격 모니터링과 음성 명령을 통한 모바일 로봇제어  
Fig. 4. Remote Monitoring and Voice Command-Based Control of a Mobile Robot

### III. 결 론

본 논문은 스마트 글래스를 통한 실시간 모니터링 시스템을 제안하였다. 스마트 글래스를 통해 로

봇의 웹캠에서 송출되는 영상을 모니터링 하고, OpenAI의 GPT-3.5-Turbo 모델을 활용한 음성 챗봇을 통해 로봇을 원격 제어 가능하다. 향후 웹 스트리밍 시 객체 탐지 모델을 추가하여 작업자가 작업할 때 안전을 위협하는 요소들을 파악하여 사용자에게 알림을 줄 수 있는 연구를 진행할 예정이다.

### 참 고 문 헌

- [1] K. Xiang, "A Observation on the Design Trend and Factors of HMD -Focusing on case analysis-", Journal of the Korean Society of Contents, DBpia, January 2023.
- [2] J. Shin, S. Park, H. Kim, and J. Lee, Search-based Generation Techniques for Enhancing LLM Responses: A Comparative Study of GPT-3.5 and GPT-4 in Zero-shot and RAG, in Proceedings of the 2023 Autumn Conference of the Korean Institute of Information and Communication Sciences (KICS), 2023.
- [3] S. H. Jang, C. U. Jeun, and S. H. Ok, ROS based Mobile Robot Control using LLM and Voice Interface, in Proceedings of the 2024 Summer Conference of the Korean Institute of Information Technology (KIT), pp. 532-534, May, 2024.

# KoBERT 모델을 활용한 텍스트 기반 리뷰 감정 분류

정창현, 설성중, 이재혁, 임지후\*, 염찬욱\*\*, 곽근창\*\*\*

## Text-Based Review Emotion Classification Using the KoBERT Model

Chang-Hyeon Jeong, Sung-Joong Seol, Jae-Hyuk Lee, Ji-Hoo Lim\*, Chan-Uk Yeom\*\*,  
and Keun-Chang Kwak\*\*\*

### 요약

본 논문은 KoBERT 모델을 활용하여 한국어 텍스트에서 기쁨, 분노, 슬픔, 불안, 당황, 상처의 6가지 감정을 분류하는 모델을 제안한다. 첫 번째 단계는 KoBERT 모델을 말뭉치 데이터셋으로 파인튜닝 하는 과정이다. 두 번째 단계는 하이퍼파라미터 조정을 통해 모델 성능을 향상시키는 과정이며, 마지막 단계는 웹 크롤링으로 확보한 실제 리뷰 데이터를 사용하여 성능을 평가 및 검증하는 과정이다. 이를 위해 말뭉치 데이터셋을 활용하여 KoBERT 모델을 파인튜닝하고, 실제 리뷰 데이터를 통해 성능을 검증하였다. KoBERT 기반 감정 분류 모델은 하이퍼파라미터 최적화를 통해 Test Accuracy 0.67, F1-Score 0.69의 성능을 기록하며, 기존 모델인 LSTM, GRU 보다 성능이 향상되었다. 또한, 텍스트 데이터를 정확히 반영하여 감정 분류의 신뢰성을 높였다. 실제 영화 관람평을 입력값으로 테스트한 결과, 감정 분류 성능이 우수하여 실생활 응용 가능성도 확인되었다.

### Abstract

This paper proposes a model for classifying six emotions—joy, anger, sadness, anxiety, embarrassment, and hurt—in Korean text using the KoBERT model. The first step involves fine-tuning the KoBERT model with a corpus dataset. The second step focuses on enhancing the model's performance by adjusting hyperparameters, and the final step involves evaluating and validating performance using actual review data collected via web crawling. For this purpose, the KoBERT model was fine-tuned with a corpus dataset, and its performance was validated with real review data. The KoBERT-based emotion classification model recorded a Test Accuracy of 0.67 and an F1-Score of 0.69 through hyperparameter optimization, outperforming previous models such as LSTM and GRU. Furthermore, it demonstrated high reliability in emotion classification by accurately reflecting textual data. Testing with actual movie reviews as input confirmed the model's excellent emotion classification performance and practical applicability.

### Key words

sentiment analysis, KoBERT, text classification, emotion classification, Korean text, product reviews, web crawling

\*조선대학교 컴퓨터통계학부, jch2209@chosun.kr, tjftjdwnd@chosun.kr, jehyug0212@chosun.kr, huhu4282@chosun.ac.kr,  
\*\*조선대학교 AI융합대학사업단 연구교수, walt18@chosun.ac.kr, \*\*\*조선대학교 전자공학부 교수, kwak@chosun.ac.kr

※ 이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 지원을 받아 수행된 AI 헬스케어 융합대학 사업 연구임.

## I. 서 론

최근 온라인 쇼핑과 소셜 미디어의 급격한 성장에 따라 소비자 리뷰 데이터의 양이 폭발적으로 증가하고 있다. 이러한 리뷰는 소비자들이 제품을 평가하고 의사 결정을 내리는 데 중요한 역할을 하지만, 방대한 리뷰 데이터를 소비자가 직접 읽고 분석하기에는 시간적, 인지적 한계가 존재한다. 리뷰의 감정 상태는 작성자의 순간적인 기분이나 개인적 경험에 따라 왜곡될 수 있어, 리뷰의 실질적인 가치를 정확히 평가하는 데 어려움이 있다. 이에 따라 리뷰의 감정을 자동으로 분석하여 기쁨, 분노, 슬픔, 불안, 당황, 상처의 6가지 감정을 분류하는 감정 분석 기술의 필요성이 대두되고 있다.

기존의 감정 분석 모델들은 단순한 키워드 기반 분석에 의존하거나 문맥을 충분히 반영하지 못해 감정 분류의 정확도가 떨어지는 한계가 있었다. 특히, 한국어 텍스트의 경우 복잡한 문법 구조와 다의어 표현으로 인해 감정 분석이 더욱 어려워, 기존 모델로는 만족스러운 성능을 얻기 어려웠다. 이러한 문제점을 해결하기 위해서는 한국어에 특화된 모델이 필요하다[1].

본 연구는 이러한 한계점을 극복하고자, KoBERT(Korean Bidirectional Encoder Representations from Transformers) 모델을 활용하여 한국어 텍스트에서 정확히 분류할 수 있는 감정 분류 모델을 제안한다. 특히 본 연구에서는 다양한 한국어 말뭉치 데이터를 학습시켜, 감정 분류의 정확도를 높이고 실질적인 활용 가능성을 확인하는 것을 목표로 한다. 이를 통해 판매자도 리뷰 데이터를 보다 효율적으로 활용할 수 있을 것으로 기대된다.

## II. 기존 모델의 한계점 및 KoBERT 모델의 필요성

감정 분석을 위한 기존 모델인 LSTM(Long Short-Term Memory)과 GRU(Gated Recurrent Unit)는 순환 신경망(RNN) 기반의 모델로, 시간에 따라 변화하는 데이터의 장기 의존성을 학습하도록 설계

되었다. LSTM은 셀 상태와 게이트 구조를 통해 장기적인 정보를 효과적으로 유지하며, GRU는 LSTM의 구조를 단순화해 연산 효율을 높이면서도 유사한 성능을 제공한다[2]. 그러나 이러한 모델들은 감정 표현을 분류하는 데 있어 여전히 제한적인 성능을 보인다. 기존 감정 분류 시스템은 단순한 키워드 매칭이나 통계적 방법론에 의존하여 감정을 분류했으며, 이러한 방식은 문맥을 고려하지 못해 정확한 분석이 어려웠다. 예를 들어, 한국어의 경우 단어가 문맥에 따라 다른 감정을 표현할 수 있는데, 기존의 키워드 기반 모델들은 이러한 문맥적 변화를 반영하지 못해 감정 분류의 성능이 제한되었다.

이러한 한계를 극복하기 위해 딥러닝 기반 모델들이 도입되었으며, 특히 BERT 기반의 자연어 처리 모델들이 탁월한 성능을 발휘하고 있다. BERT(Bidirectional Encoder Representations from Transformers)는 문맥을 양방향으로 이해할 수 있는 모델로, 텍스트의 의미와 맥락을 효과적으로 파악하여 감정 분석의 정확도를 크게 향상시킨다[3]. 특히 KoBERT는 한국어 텍스트에 최적화된 BERT 모델로, 한국어의 복잡한 문법 구조와 표현 방식에 맞춰 설계되었기 때문에 한국어 감정 분석에서 더 높은 성능을 발휘할 수 있다[4].

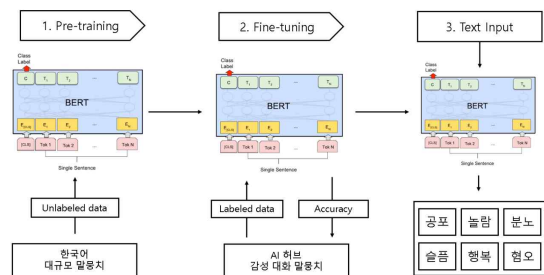


그림 1. KoBERT 모델 학습 및 감정 분류 과정  
Fig. 1. KoBERT Model Training and Emotion Classification Process

본 연구에서는 KoBERT 모델을 사용하여 텍스트 데이터의 감정을 6가지 감정(기쁨, 분노, 슬픔, 불안, 당황, 상처)으로 분류하는 감정 분류 모델을 제안한다. 이 모델은 특히 다양한 한국어 말뭉치 데이터를 기반으로 학습되어, 실제 리뷰 데이터를 통해 검증되었다. KoBERT 모델은 기존 모델들과 비교하여 문

맥을 더 깊이 이해하고 처리할 수 있어, 리뷰 작성자의 의도를 보다 정확하게 반영할 수 있다. 이를 통해 감정 분류 결과의 신뢰성을 높이고, 다양한 응용 분야에서 활용 가능성을 기대할 수 있다.

### III. 모델 학습 및 검증

#### 1. 데이터셋 설명

본 연구에서는 감정 분류를 위한 KoBERT 모델의 학습을 위해 AI 허브에서 제공하는 감성 대화 말뭉치 데이터셋을 사용하였다. 이 데이터셋은 한국어 감성 대화 데이터를 기반으로 구축되었으며, 기쁨(joy), 분노(anger), 슬픔(sadness), 불안(anxiety), 당황(embarrassment), 상처(hurt)의 6가지 감정으로 라벨링 되어있다. 데이터셋은 다양한 세대와 상황을 반영한 대화 텍스트로 구성되어 있어, 한국어 텍스트의 감정 분류에 적합하다.

데이터 규모: 본 연구에서 사용된 데이터셋은 총 51,630개의 감성 대화 문장으로 구성되어 있으며, 각 문장은 앞서 언급한 6가지 감정으로 레이블링 되었다. 학습 및 테스트 데이터 분할: 전체 데이터셋의 80%인 41,304개는 학습용 데이터로 사용되었으며, 나머지 20%인 10,326개는 모델 성능 평가를 위한 테스트 데이터로 사용되었다. 감정 분류의 편향을 방지하기 위해 각 감정별 데이터는 고르게 분포되도록 샘플링되었다. 이 데이터셋은 다양한 감정과 상황을 포함하고 있어 감정 분류 모델의 학습에 이상적이며, 한국어 문장의 미묘한 감정 변화를 학습시키기에 적합한 특성을 포함하고 있다.

#### 2. 모델 학습 과정

본 연구에서는 KoBERT 모델을 기반으로 감정 분류 작업을 수행하였다. KoBERT는 한국어 자연어 처리에 최적화된 BERT 기반 모델로, 문맥을 양방향으로 분석하여 텍스트의 의미를 보다 깊이 이해할 수 있다. 감정 분류 작업에서는 각 감정의 복잡한 표현을 잘 반영할 수 있는 KoBERT의 특성을 최대한 활용하였다.

하이퍼파라미터 설정:최적의 학습 성능을 위해 하이퍼파라미터를 조정하였다[5]. 초기에는 KoBERT 모델의 기본 설정을 따랐으나, 성능 향상을 위해 여러 번의 실험을 진행했다. 이를 통해 최종적으로 사용된 하이퍼파라미터는 아래와 같다.

표 1. 하이퍼파라미터 수치 조정  
Table 1. Hyperparameter Tuning

하이퍼파라미터 수치		
변수	변경 전	변경 후
max_len	128	64
batch_size	10	64
num_epoch	5	10
learning_rate	5e-5	3e-6
dr_rate	-	0.2
log_interval	200	200

학습률(learning\_rate)은 5e-5에서 3e-6으로 조정되었으며, num\_epoch를 10으로 늘려 모델이 충분히 학습할 수 있도록 하였다. 또한, batch\_size는 32로 설정하여 학습 효율성을 높였고, dropout\_rate를 0.2로 적용하여 과적합을 방지하였다.

위 표는 하이퍼파라미터 최적화 전후의 변화를 나타낸다. 먼저, 입력 시퀀스 처리 효율성을 높이기 위해 max\_len을 128에서 64로 줄여 처리 시간을 단축하였다. 또한, batch\_size는 10에서 64로 증가하여 학습의 안정성을 높이고 수렴 속도를 빠르게 하였다. learning\_rate는 1e-1에서 5e-6으로 줄여 학습 과정의 수렴을 세밀하게 조정하였으며, dropout\_rate는 0.1에서 0.5로 설정하여 과적합을 방지하였다. 이러한 최적화 과정을 통해 Test Accuracy는 0.67, F1-Score는 0.69의 값을 얻을 수 있었다.

#### 3. 최종 모델 비교

최적의 감정 분류 모델을 결정하기 위해 LSTM, GRU, KoBERT 모델 간의 Test Accuracy와 F1-Score 비교 분석을 수행하였다. 그 결과, KoBERT 모델이 Test Accuracy 0.67과 F1-Score 0.69로 가장 높은 성능을 기록하며, LSTM (Test Accuracy: 0.57, F1-Score: 0.58), GRU(Test Accuracy: 0.54, F1-Score:



0.55) 보다 향상된 성능을 보였다. 이러한 결과를 바탕으로 KoBERT 모델을 감정 분류에 적합한 모델로 선정하였다. 여기서 정확도(Accuracy)는 전체 예측 중 올바르게 분류된 샘플의 비율을 의미하며, 수식 (1)과 같이 정의된다.

여기서 TP는 True Positive, TN은 True Negative, FP는 False Positive, FN은 False Negative를 나타낸다. 반면, F1-Score는 모델의 정밀도(Precision)와 재현율(Recall)을 모두 고려하여 조화 평균을 구한 값으로, 특히 불균형한 데이터셋에서 모델 성능을 평가할 때 유용하다. F1-Score는 수식 (2)로 표현된다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2)$$

이때, 정밀도(Precision)는  $\frac{TP}{TP + FP}$ , 재현율(Recall)은  $\frac{TP}{TP + FN}$ 로 정의된다. F1-Score는 높은 정밀도와 재현율을 동시에 달성한 모델을 선호하게 하며, 본 연구에서는 KoBERT 모델이 가장 높은 F1-Score를 기록하여 이러한 평가 지표를 종합적으로 고려했을 때 감정 분류에 적합함을 확인하였다.

표 2. 모델 간 정확도 비교

Table 2. Comparison of Accuracy Between Models

Model	Test_Accuracy	F1-Score
LSTM	0.57	0.58
GRU	0.54	0.55
KoBERT	0.67	0.69

#### 4. 웹 크롤링을 통한 KoBERT 모델 성능 검증

본 연구에서는 KoBERT 기반 감정 분류 모델의 성능을 평가하기 위해 웹 크롤링 기술을 사용하여 실제 영화 리뷰 데이터를 수집하였다. 웹 크롤링은 자동화된 프로그램을 통해 웹사이트에서 데이터를 수집하는 방법으로, 본 연구에서는 특정 영화 리뷰

사이트에서 리뷰 텍스트를 수집하는 데 사용되었다 [6]. 수집된 리뷰 데이터는 XLSX 파일로 저장되었고, 이를 KoBERT 모델에 입력하여 각 리뷰를 기쁨(joy), 분노(anger), 슬픔(sadness), 불안(anxiety), 당황(embarrassment), 상처(hurt)의 6가지 감정으로 분류하였다. 그 결과는 다시 XLSX 파일로 출력되었으며, 이를 통해 실제 데이터 환경에서 모델의 감정 분류 성능을 평가하였다. 이 실험을 통해 KoBERT 모델이 실제 리뷰 데이터를 처리할 때도 높은 정확도로 감정을 분류할 수 있음을 확인하였으며, 다양한 실생활 응용 가능성을 검증할 수 있었다.

#### IV. 결론

본 연구는 KoBERT 모델을 활용하여 한국어 텍스트에서 기쁨, 분노, 슬픔, 불안, 당황, 상처의 6가지 감정을 분류하는 감정 분류 모델을 제안하였다. 기존 감정 분류 모델들이 문맥을 충분히 반영하지 못해 감정 분류의 정확도가 떨어지는 문제를 극복하고자 하였으며, 이를 위해 AI 허브 감성 대화 데이터셋을 학습 데이터로 사용하였다. 실험 결과, 모델의 Test Accuracy는 0.67, F1-Score는 0.69로 기록되었으며, 기존 연구 대비 성능이 향상되었음을 확인하였다.

이 연구는 특히 한국어 텍스트에 특화된 감정 분류 모델을 제시함으로써, 실질적인 감정 분류에서 우수한 성능을 확인할 수 있다. 이를 통해 판매자들이 소비자 리뷰 데이터를 보다 효율적으로 분석하고 활용할 수 있는 방안을 제시하였으며, 감정 분석을 통해 소비자의 감정 상태를 명확히 파악하여 제품 평가와 의사결정을 돕는 것에 도움을 줄 수 있다.

하지만 본 연구에는 몇 가지 한계점이 존재한다. 사용된 데이터셋은 감성 대화를 기반으로 하고 있어, 보다 복잡한 문맥을 포함하는 텍스트나 다양한 상황에서의 감정 표현에 대해 추가적인 검증이 필요하다. 또한, 데이터셋의 범위가 특정 대화 유형에 제한적이기 때문에, 더 다양한 데이터셋을 통해 모델의 일반화 능력을 확장할 필요가 있다.

향후 연구에서는 본 연구의 결과를 바탕으로 실

시간 감정 분류 시스템을 다양한 온라인 플랫폼에서 적용할 수 있도록 할 것이다. 또한, 감정 분류 결과를 사용자 맞춤형 피드백 시스템에 통합하여 사용자 경험을 향상시키고, 고객의 감정 상태에 따라 맞춤형 제품 추천이나 고객 지원 서비스에 활용할 수 있는 실질적인 방안도 연구할 수 있다. 이를 통해 감정 분류 모델이 더 많은 응용 분야에서 중요한 역할을 할 것으로 기대된다.

24(5), 381-388, 10.5762/KAIS.2023.24.5.381

## 참 고 문 헌

- [1] 이재아 (2023). A Study on the Korean Multi-class Emotion Analysis Based on KoBERT in Text Messages of Radio Listeners [Master's Thesis, 서울과학기술대학교].
- [2] Jee-Hyun Kim, Ye-Rim Kim, & Hae-Won Byun (2023). SNS Context-based Emoji Recommendation Using Hierarchical KoBERT. Journal of Digital Contents Society, 24(6), 1361-1371, 10.9728/dcs.2023.24.6.1361
- [3] Cheoneum Park, & Changki Lee(2019). Sentimental Analysis of Korean Movie Review using Variational Inference and RNN based on BERT. KIISE Transactions on Computing Practices,25(11), 552-558, 10.5626/KTCP.2019.25.11.552
- [4] Eunmi Kim, Seung-Jin Nam, Taeyi Kim, & Taeho Hong(2024). The Prediction of Review Helpfulness by Integrating Large Language Models and Deep Learning based on KoBERT and KoGPT2. Journal of Intelligence and Information Systems, 30(2), 195-209, 10.13088/jiis.2024.30.2.195
- [5] 우종현,&김형중(2023-05-31).Bayesian Optimization-based optimal hyperparameter selection for FDM process defect detection. 대한산업공학회 춘계공동학술대회 논문집, 여수.
- [6] Hanin Lee, & Jinho Cha (2023). A Study on the Recognition of Korea Armed Forces Nursing Academy and Military Nursing Officers through Web Crawling and Text Mining. Journal of the Korea Academia-Industrial cooperation Society,

# 맛집 추천을 위한 음식 분류체계 기반 온톨로지 구축 연구

한혜지\*, 김장원†

## A Study on Developing an Ontology-Based Food Classification System

Hyeji Han\* and Jangwon Gim†

### 요 약

맛집 검색 결과를 제공하는 서비스와 기존에 공개된 음식 분류체계는 최근 다양성과 트렌드에 관심이 높은 사용자들의 세부적인 요구를 충분히 반영하지 못하는 한계가 있다. 따라서, 본 논문에서는 음식점의 특징과 다양한 음식 스타일을 고려하여 기존 분류체계를 개선한 온톨로지 기반 음식 분류체계를 제안한다. 개선된 분류체계를 활용함으로써 사용자는 자신에게 맞는 맛집 검색 결과를 얻을 수 있음을 보였다. 향후 제안한 분류 체계는 사용자 맞춤형 추천 시스템 개발에 다양한 사용자의 요구를 반영한 검색 결과 및 추천 정보를 제공할 수 있을 것으로 기대한다.

### Abstract

Existing restaurant search services and publicly available food classification systems have limitations in meeting the specific needs of users who are increasingly interested in diversity and current trends. This paper proposes an ontology-based food classification system that improves upon existing frameworks by incorporating various food styles and restaurant characteristics. The enhanced classification system allows users to obtain personalized restaurant search results that better align with their preferences. The proposed system is expected to support the development of personalized recommendation systems, providing search results and recommendations that reflect diverse user requirements.

### Key words

Food Classification System, Ontology, recommendation system

### 1. 서 론

현대 사회에서는 사용자 맞춤형 추천에 대한 수

요가 증가하고 있다[1]. 음식점에 대한 정보를 더 세밀하게 분류하고 사용자 요구에 맞춘 추천 시스템이 필요해지고 있다[2]. 그렇지만 기존의 맛집 추

\* 국립군산대학교 소프트웨어학과 (hanhyeji0126@kunsan.ac.kr)

† 군산대학교 컴퓨터소프트웨어학부 교신저자 (jwgim@kunsan.ac.kr)

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065)

천 시스템은 주로 단순한 위치 정보나 제한된 카테고리 기반을 기반으로 하여 사용자 개개인의 세부적인 요구를 충분히 반영하지 못하는 한계가 있다[3]. 또한, 기존의 음식 분류체계가 정교하지 않아 사용자 맞춤형 음식 추천 시 상세한 추천을 제공하는 데 어려움이 있다. 이러한 한계를 극복하기 위해, 본 논문에서는 음식점의 특징과 음식 스타일 등을 참고하여 기존 분류체계를 개선한 새로운 음식 분류체계를 설계하고, 이를 활용한 사용자 맞춤형 추천 시스템을 위한 온톨로지 구축 연구를 제안한다.

## II. 관련 연구

### 2.1 온톨로지

온톨로지란 특정 도메인 내 개념과 그들 간의 관계를 공식적으로 표현한 것으로, 구조화된 지식을 나타낸다[4]. 또한, 정보의 의미적 관계를 기반으로 한 추천을 가능하게 해준다. 기존 맛집 추천 시스템이 단순히 위치나 카테고리 정보에 의존하는 것과 달리 온톨로지를 활용하면 음식점의 편의시설, 리뷰 등의 요소를 정교하게 연결하여 사용자 요구를 세분화하여 반영할 수 있다.

### 2.2 맛집 추천 시스템

맛집 추천 시스템과 관련된 다양한 연구가 진행되었다. [5]는 머신러닝 기법을 활용하고, [6]은 인스타그램 데이터를, [7]은 소셜 네트워크 정보를 활용하여 추천 시스템을 제안하였다. 이러한 연구들은 음식 분류체계의 세분화나 온톨로지 활용에 한계가 있어 사용자들의 구체적인 음식 선호도를 충분히 반영하지 못하는 문제점이 있다. 따라서 음식의 다양한 특징을 구조화하여 추천에 활용하는 방법에 관한 연구가 필요하다. 본 논문에서는 이러한 선행 연구를 바탕으로 음식 분류체계와 온톨로지를 활용한 새로운 맛집 추천 시스템을 제안하고자 한다.

## III. 음식 분류체계

통계청의 통계분류포털에서는 음식점을 분류할 수 있는 분류체계를 공개하고 있다[8]. 그러나 기존 분류체계는 한식, 중식, 일식 등 대분류와 육류 요리 전문점, 면 요리 전문점 등으로 세분화하지만, 음식의 다양한 특징을 충분히 반영하지 못해 사용자들의 구체적인 요구를 만족시키기 어렵다. 예를 들어, 다양한 스타일의 면 요리 전문점을 찾고자 할 때 한식 면 요리 전문점만 추천되는 문제가 있다. 이처럼 세부적인 요구사항이 분류체계에 반영되지 않아 사용자는 원하는 정보를 얻기 위해 추가적인 노력이 필요하게 된다.

본 논문에서는 기존 음식 분류체계를 개선한 새로운 음식 분류체계를 제안한다. 그림 1은 제안된 분류체계를 보여준다. 개선된 분류체계는 사용자 요구를 반영하여 면 요리 전문점, 육류 요리 전문점 등을 한식, 일식, 서양식 등으로 더욱 세분화하였다(회색 음영). 이를 통해 사용자는 원하는 음식 유형이나 특성에 맞는 음식점을 더 쉽게 찾아 사용자 맞춤형 추천 시스템의 효과성을 높일 수 있다.

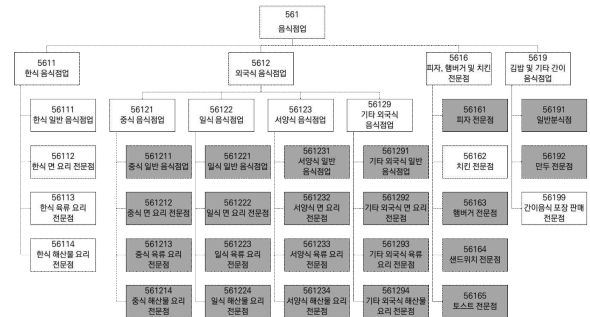


그림 1. 제안하는 음식 분류체계  
Fig. 1. Proposed Food Classification System

## IV. 실험

### 4.1 음식 분류체계 RDF 구축

그림 2는 제안하는 음식 분류체계의 일부를 온톨로지 형태로 보인다. 각 분류는 skos:broader와 skos:narrower 속성을 통해 상위 및 하위 개념의 관계를 표현한다.

```

@prefix kofid: <http://knowledgemap.kr/ontology/food/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

# 음식점 및 음식점
kofid:FoodAndBeverageServiceActivities a rdfs:Class, skos:Concept ;
skos:prefLabel "음식점 및 음식점" ;
skos:prefLabel "Food and beverage service activities" ;
skos:broader kofid:RestaurantAndMobileFoodServiceActivities .

# 음식점
kofid:Restaurant a rdfs:Class, skos:Concept ;
skos:prefLabel "음식점" ;
skos:prefLabel "Restaurant and mobile food service activities" ;
skos:broader kofid:FoodAndBeverageServiceActivities ;
skos:narrower kofid:KoreanFoodRestaurant, kofid:ForeignFoodRestaurant, kofid:FastFoodServiceActivity, kofid:DriedSoybeanEddisAndJigilFoodServiceActivities .

# 한식 음식점
kofid:KoreanFoodRestaurant a rdfs:Class, skos:Concept ;
skos:prefLabel "한식 음식점" ;
skos:prefLabel "Korean food restaurant" ;
skos:broader kofid:RestaurantAndMobileFoodServiceActivities ;
skos:narrower kofid:KoreanGeneralFoodRestaurant, kofid:KoreanNoodleFoodRestaurant, kofid:KoreanMeatFoodRestaurant, kofid:KoreanSeafoodRestaurant .

# 한식 일반 음식점
kofid:KoreanGeneralFoodRestaurant a rdfs:Class, skos:Concept ;
skos:prefLabel "한식 일반 음식점" ;
skos:prefLabel "Korean General Food Restaurant" ;
skos:broader kofid:KoreanFoodRestaurant .
    
```

그림 2. 제안하는 음식 분류체계 온톨로지 (일부)  
Fig. 2. Proposed Food Classification System Ontology (partial)

#### 4.2 음식 분류체계 기반 질의 예시

제안하는 분류체계에 대한 사용자 맞춤형 검색 결과 확인을 위해 인스턴스를 RDF 형태로 구축 및 저장하였다. 그림 3는 트리플 저장소의 RDF를 활용한 SPARQL 쿼리문 결과를 보인다. 개선된 분류체계를 기반으로 '면 요리 전문점'에 대한 검색에서 한식 면요리 전문점뿐만 아니라 서양식, 일식 등 다양한 검색 결과를 얻을 수 있었다.

```

1 PREFIX schema: <http://schema.org/>
2 PREFIX ex: <http://example.org/restaurants/>
3 PREFIX kofid: <http://knowledgemap.kr/ontology/food/>
4
5 SELECT ?restaurantName ?address ?category ?participantCount ?additionalInfo
6 WHERE {
7   ?restaurant a schema:Restaurant ;
8     schema:name ?restaurantName ;
9     schema:address ?address ;
10    ex:participantCount ?participantCount ;
11    ex:additionalInformation ?additionalInfo ;
12    schema:foodEstablishmentCategory ?category .
13
14 FILTER(?category IN (kofid:한식면요리전문점, kofid:서양식면요리전문점, kofid:일식면요리전문점,
15                    kofid:중식면요리전문점, kofid:기타외국식면요리전문점))
16 }
17 ORDER BY DESC(xsd:integer(?participantCount))
    
```

restaurantName	address	category
영미탈국수	전북 군산시 근영길 26	http://knowledgemap.kr/ontology/food/한식면요리전문점
황산중화요리 군산	전북 군산시 대학로 209 황산중화요리	http://knowledgemap.kr/ontology/food/중식면요리전문점
요쿠미치 군산	전북 군산시 구영3길 49 요쿠미치	http://knowledgemap.kr/ontology/food/일식면요리전문점
리틀유러	전북 군산시 수송남로 40-1 리틀 YouLuv	http://knowledgemap.kr/ontology/food/서양식면요리전문점
영미탈국수 직영점	전북 군산시 수송남로 20	http://knowledgemap.kr/ontology/food/한식면요리전문점
파라디소 뷔페두부	전북 군산시 한밭로 76-11 파라디소 뷔페두부	http://knowledgemap.kr/ontology/food/서양식면요리전문점
모든	전북 군산시 미충남로 111-5 1층	http://knowledgemap.kr/ontology/food/서양식면요리전문점
라노	전북 군산시 나충안길 35	http://knowledgemap.kr/ontology/food/서양식면요리전문점
전영원	전북 군산시 구영1로 70-1	http://knowledgemap.kr/ontology/food/중식면요리전문점
오스트레한	전북 군산시 옥산면 남내안길 14	http://knowledgemap.kr/ontology/food/서양식면요리전문점
리조리	전북 군산시 대야면 서안자길 6	http://knowledgemap.kr/ontology/food/서양식면요리전문점
대복명	전북 군산시 신영1길 3 대복명	http://knowledgemap.kr/ontology/food/중식면요리전문점

그림 3. '면 요리 전문점'에 대한 SPARQL 쿼리문과 결과  
Fig. 3. SPARQL queries and results on 'Noodle specialty'

#### V. 결론 및 향후 연구

본 논문은 기존 맛집 추천 시스템과 음식 분류체계가 사용자들의 세부적인 요구를 충분히 반영하지 못하는 한계를 해결하기 위해, 음식점의 특징과 음식

스타일을 고려한 새로운 음식 분류체계를 설계하고 이를 온톨로지로 구축하였다. SPARQL 쿼리를 활용하여 사용자가 원하는 음식 유형이나 특성에 맞는 음식점 검색이 가능한 것을 보였다. 향후에는 제안한 분류체계를 기반으로 사용자 맞춤형 추천 시스템을 개발할 것이다.

#### 참고 문헌

[1] Chen, W., Huang, P., Xu, J., Guo, X., Guo, C., Sun, F., ... & Zhao, B. (2019, July). POG: personalized outfit generation for fashion recommendation at Alibaba iFashion. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 2662-2670).

[2] Blanco-Mallo, E., Remeseiro, B., Bolón-Canedo, V., & Alonso-Betanzos, A. (2020, August). On the effectiveness of convolutional autoencoders on image-based personalized recommender systems. In Proceedings (Vol. 54, No. 1, p. 11). MDPI.

[3] Dev, H., & Hamooni, H. (2020, October). Profiling us restaurants from billions of payment card transactions. In 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA) (pp. 449-458). IEEE.

[4] Funk, M., Hosemann, S., Jung, J. C., & Lutz, C. (2023). Towards ontology construction with language models. arXiv preprint arXiv:2309.09898.

[5] 최환석, 팽전, & 이우섭. (2020). 머신러닝 기반 음식점 추천시스템 설계 및 구현. 한국디지털콘텐츠학회 논문지, 21(2), 259-268.

[6] 정한조, 송은수, 최현승, & 박원정. (2020). 인스타그램 포스트 데이터를 이용한 협업 필터링 기반 맛집 추천 시스템. 한국컴퓨터정보학회 학술발표논문집, 28(2), 279-280.

[7] LIANHUA, T. (2013). 소셜 정보 기반의 맛집 추천시스템. 이화여자대학교 대학원 석사학위논문, pp24-43.

[8] 통계청 한국표준산업분류, (2024). <https://kostat.go.kr/>

# CCTV 영상 분석을 통한 소방 방제용 다중 군중 움직임 감지 및 위험도 평가 시스템 연구

정상진\*, 조영현\*, 김승현\*, 한동진\*\*

## Fire Prevention System for Multi-Crowd Movement Detection and Risk Assessment Using CCTV Video Analysis

Sangjin Jeong\*, Young-Hyun Cho\*, Seung-Hyun Kim\*, and Domgjin Han\*\*

### 요 약

본 연구는 CCTV 영상 분석을 통해 다중 군중의 움직임을 실시간으로 감지하고 위험도를 평가하는 소방 방제용 시스템을 제안한다. 대규모 군중이 모이는 장소에서 발생할 수 있는 화재나 구조물 붕괴와 같은 비상 상황에 대비하여, 군중의 밀집도와 이동 패턴을 분석하여 위험 요소를 조기에 감지하고 경고를 발령하는 기능을 구현하기에 필요한 요소를 파악하였다. 연구 결과는 군중 밀도, 이동 속도 및 방향 변화를 통해 비상 상황의 위험도를 정확히 평가할 수 있는 방안을 모색한다, 제안된 시스템이 재난 발생 시 초기 대응에 효과적으로 기여할 수 있도록 한다.

### Abstract

This study proposes a fire prevention system that utilizes CCTV video analysis to detect multi-crowd movements in real-time and evaluate risk levels. The system is designed to respond to emergencies, such as fires and structural collapses in densely populated areas, by analyzing crowd density and movement patterns to identify potential hazards early and issue warnings. The findings demonstrate that the system can accurately assess emergency risks based on crowd density, movement speed, and direction changes, confirming the system's effectiveness in facilitating rapid initial responses during disasters.

### Key words

Crowd Movement Detection, Risk Assessment System, Fire Prevention Technology

### 1. 서 론

대규모 군중이 모이는 행사나 장소에서 발생할 수 있는 사고는 예측이 어렵고, 화재나 구조물 붕괴와 같은 비상 상황에서는 즉각적인 대피가 어려워

대형 인명 사고로 이어질 위험이 크다. 이러한 상황에서 군중의 안전을 보장하기 위한 소방 방제 시스템의 중요성이 커지고 있으며, 특히 군중이 밀집된 상태에서는 더욱 신속하고 정확한 위험 감지와 대처가 필요하다.

\* 경일대학교 소방방재학부 \*\* Email(교신저자 한동진 han@kiu.ac.kr), ...

※경일대학교 2024 대학혁신지원사업(UI)

현재 대부분의 군중 밀집 상황에서는 CCTV를 통한 관제와 물리적 경비 인력에 의존하고 있으나, 실시간으로 군중의 움직임 분석을 통한 비정상적인 상황을 감지하고 위험성을 평가하는 기술적 접근은 부족한 실정이다. 특히 재난 발생 시 군중의 움직임과 비정상적 행동을 조기에 파악할 수 있는 시스템의 필요성이 대두되고 있다.

본 연구는 이러한 문제를 해결하기 위해, CCTV 영상 데이터를 기반으로 군중의 움직임과 밀집 상태를 실시간으로 분석하고, 비상 상황에서 위험도를 평가할 수 있는 소방 방제 시스템의 설계 방안을 모색하고자 한다. 이를 통해 군중 밀집 상태에서의 화재나 압사 등의 위험을 사전에 감지하고 경고를 발령할 수 있는 알고리즘을 제안함으로써, 대규모 재난 발생 시 효과적인 초기 대응이 가능하도록 하는 것을 목표로 한다.

## II. 군중 밀도, 이동 속도 및 위험도 평가 모델

### 2.1 연구 배경 및 관련 기술

대규모 군중이 모이는 장소에서 화재, 압사 등의 사고가 발생할 경우 즉각적인 대응이 필수적이며, 이러한 상황에서는 실시간으로 군중의 밀집도와 이동 경로를 분석하여 위험도를 평가하는 시스템이 필요하다. 현재 CCTV와 컴퓨터 비전 기술이 발전함에 따라, 군중의 움직임을 분석하는 다양한 연구가 진행되고 있다. 특히, 컴퓨터 비전 기술은 CCTV 영상을 통해 사람들의 위치, 이동 속도, 방향 등을 파악하고, 이를 통해 군중의 행동 패턴과 밀집 상태를 실시간으로 모니터링할 수 있는 기반을 제공한다. 이러한 기술적 발전을 바탕으로 군중의 밀집도와 이동 패턴을 분석하여 위험 요소를 사전에 감지하는 소방 방제 시스템을 설계하는 것이 본 연구의 주요 목표이다.

### 2.2 연구 방법론 (Methodology)

#### 2.2.1 데이터 수집 및 전처리

본 연구에서는 대규모 군중이 모이는 장소에서 CCTV 영상 데이터를 수집하고, 이를 분석하기 위

한 전처리 과정을 수행한다. CCTV 영상 데이터는 다양한 각도와 거리에서 촬영되므로, 영상의 해상도와 조명 조건에 따라 데이터가 달라질 수 있다. 따라서, 수집된 데이터는 군중의 이동 방향, 밀집도 등을 정확히 분석할 수 있도록 조정된다.

#### 2.2.2 군중 움직임 분석 기법

군중 밀도 분석: 군중이 밀집되는 구역을 감지하는 알고리즘을 사용하여, 특정 구역에 모이는 사람들의 수를 실시간으로 파악한다. 이를 통해 밀집도가 높은 지역을 구분하고, 경고 수준에 따라 적절한 대응을 할 수 있도록 한다.

$$v_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} / \Delta t \quad (1)$$

속도 및 방향 분석: 군중의 이동 경로와 속도 변화를 분석하여 비정상적 움직임을 감지하는 알고리즘을 적용한다. 예를 들어, 특정 구역에서 군중이 갑자기 한 방향으로 빠르게 이동하거나 밀집된 지역에서 혼란스러운 움직임이 발생할 경우, 비상 상황의 징후로 인식할 수 있다.

위험 요소 감지: 화재나 연기와 같은 외부 위험 요소를 감지할 수 있는 컴퓨터 비전 알고리즘 (Optical flow) 등을 사용하여, 단순한 군중 움직임뿐만 아니라 주변 환경의 변화도 동시에 모니터링할 수 있을 것이다. 이러한 외부 요소들은 군중의 대피를 유도하거나 위험 구역을 즉시 알릴 수 있는 경고 체계를 구축하는 데 사용된다.

#### 2.2.3 위험도 평가 모델

간단한 위험도 평가 모델은 군중 밀집도, 이동 속도, 방향 변화, 외부 위험 요소와 같은 요인을 종합적으로 분석하여 위험도를 계산하는 방식을 사용해 볼 수 있을 것이다. 위험도 지표는 수치화하여 일정 임계치에 도달하면 경고가 발령되며, 위험도가 높은 구역을 시각화하여 소방 방제 인력에게 실시간으로 제공한다. 이를 통해 특정 구역에서 위험도가 급격히 증가할 경우, 인명 피해를 줄이기 위한 즉각적인 대응이 가능하도록 한다. 초기 실험을 위해 다음과 같은 식을 사용할 수 있을 것이다.

$$R = w_1 \frac{D}{D_{threshold}} + w_2 \frac{\bar{v}}{v_{threshold}} + w_3 \frac{\Delta\theta}{\theta_{threshold}} \quad (2)$$

$D, \bar{v}, \Delta\theta$ 는 각각 군중 밀집도, 평균속도, 방향 변화량을 나타낸다. 군중 밀집도는 단순히 사람의 수를 면적으로 나눈 군중 밀도를 사용해 볼 수 있을 것이다.

### 2.2.4 실시간 경고 시스템

비상 상황이 감지되면 경고 알람이 즉시 발령되며, 군중의 이동 경로를 안전하게 유도하기 위한 안내 메시지나 경로 표시 시스템과 연동된다. 경고 시스템은 위험도에 따라 다단계로 설정되며, 일반 경고에서부터 긴급 대피 안내까지 상황에 맞는 대응이 이루어지도록 설계한다. 이러한 실시간 경고 시스템은 소방 방재 관제 센터와도 연결되어, 관제 요원이 전체 상황을 신속하게 파악하고 현장 인력에게 필요한 정보를 즉각 전달할 수 있도록 한다.

## 2.3 시스템의 기대 효과 및 한계

### 2.3.1 기대 효과

제안된 시스템은 대규모 군중이 모이는 장소에서의 사고를 사전에 감지하고 예방할 수 있는 중요한 기능을 제공한다. 군중 밀집 상태에서의 위험 요소를 조기에 파악함으로써, 대형 인명 사고 발생을 방지하고, 소방 방재 인력이 신속하게 대응할 수 있도록 돕는다. 특히, 화재 발생 시 초기 단계에서 위험을 감지하고 경고를 발령함으로써 인명 피해를 줄일 수 있으며, 시스템이 자동으로 경고를 발령하여 군중을 안전한 경로로 유도하는 기능을 통해 혼란을 최소화할 수 있다.

### 2.3.2 연구의 한계

본 연구는 실제 실험이 아닌, 시스템 설계 및 방법론을 모색하는 단계에 있는 연구로서, 실제 적용 시 발생할 수 있는 기술적 문제나 환경적 요인(조

명, CCTV의 해상도, 기기 간의 연동 문제 등)에 대한 한계가 존재할 수 있다. 또한, 군중의 움직임을 예측하는 데 있어 개인의 심리적 반응이 반영되지 않으며, 이는 대규모 대피 상황에서 예측과 실제 행동 간의 차이를 유발할 수 있다.

## 2.4 향후 연구 방향

향후 연구에서는 본 시스템의 실험적 검증을 통해 군중의 이동 패턴과 위험 감지의 정확성을 높이고, 다양한 환경에서도 적용할 수 있도록 시스템의 최적화를 도모할 예정이다. 추가로, 군중의 심리적 반응을 반영할 수 있는 알고리즘을 개발하여 대피 상황에서의 군중 행동을 더욱 정확히 예측하고, 다양한 재난 상황에 대응할 수 있는 포괄적 소방 방재 시스템으로 발전시킬 예정이다.

## III. 결 론

본 연구는 대규모 군중이 모이는 장소에서 발생할 수 있는 비상 상황에 대비하여, CCTV 영상 데이터를 활용한 군중 움직임 분석 및 위험도 평가 시스템을 설계하는 방안을 제안하였다. 기존의 군중 관리 시스템은 실시간 군중 밀집도와 이동 패턴을 감지하고 이에 따라 즉각적인 대응을 유도하는 데 한계가 있었으나, 본 연구에서는 군중의 밀집도, 이동 속도, 방향 변화, 외부 위협 요소를 종합적으로 고려하여 위험도를 평가할 수 있는 시스템을 모색하였다.

이 시스템은 대규모 재난 상황에서 발생할 수 있는 혼란을 최소화하고, 소방 방재 인력이 효율적으로 대응할 수 있도록 실시간 경고 및 대피 유도 기능을 포함하고 있다. 연구 결과는 실시간 위험도 평가와 경고 발령이 가능하여 인명 피해를 최소화할 수 있는 시스템의 잠재적 효용성을 보여준다.

현재 본 시스템을 실험적으로 검증하고 성능을 최적화하는 실험을 진행하고 있다. 각 요소에 대한 실험을 통해 최적화된 알고리즘을 만들 것이다. 또한 다양한 환경에서도 적용할 수 있도록 개선할 필요가 있다. 또한, 군중의 심리적 반응을 반영한 추가적인 연구를 통해 더욱 정확하고 포괄적인 위험



감지 및 대응 시스템으로 발전시켜 나가야 할 것이다. 이를 통해 본 연구에서 제안한 시스템이 군중 밀집 장소에서의 안전성을 제고하고, 스마트 도시의 방재 시스템 구축에도 기여할 수 있을 것으로 기대된다. .

## 참 고 문 헌

- [1] V. R. T. Das, S. Velastin, and A. Vicencio-Silva, "Crowd analysis using computer vision techniques: A survey," *ACM Computing Surveys*, 2017.
- [2] A. Alahi, V. Ramanathan, and L. Fei-Fei, "Real-time crowd monitoring using CCTV and machine learning for public safety," *IEEE Trans. Pattern Anal. Machine Intell.*, 2016.
- [3] R. Mehran, A. Oyama, and M. Shah, "Risk assessment in large crowds using real-time video analysis," *IEEE Trans. Image Process.*, 2013.
- [4] M. Solmaz, B. E. Moore, and M. Shah, "Automatic detection of abnormal crowd behavior using computer vision," *Computer Vision and Image Understanding*, 2012.
- [5] W. Ge and R. T. Collins, "A survey of crowd management using video analytics," *Computer Vision and Image Understanding*, 2017.
- [6] T. Muhammad and A. H. K. Javed, "Fire detection in video sequences using convolutional neural networks," *Sensors*, 2018.
- [7] S. Helbing, A. Johansson, and H. Z. Al-Abideen, "The role of crowd density in crowd disaster management," *Safety Science*, 2014.



# 이더리움 거래 네트워크에서 사기 거래의 분리된 표현 학습을 위한 쿼드로플렛 손실 기반 그래프 컨볼루션 오토인코더

남민우, 이현주, 부석준\*

## Quadruplet Loss-based Graph Convolutional Autoencoder for Learning Disentangled Representation of Scam Transactions in Ethereum Transaction Network

Min-Woo Nam, Hyeon-Ju Lee, and Seok-Jun Buu\*

### 요약

암호화폐 사용 증가와 함께 사기 활동도 증가하여, 신뢰성 확보를 위한 정확한 사기 탐지가 중요하지만, 거래 패턴의 유사성과 데이터 불균형 문제로 기존 방법으로는 정확한 사기 탐지가 어렵다. 이를 해결하기 위해, 본 논문에서는 사기와 정상 거래 간의 미묘한 패턴을 효율적으로 포착하여 분리하기 위한 쿼드로플렛 손실과 그래프 컨볼루션 오토인코더 모델을 결합한 쿼드로플렛 손실 기반 그래프 컨볼루션 오토인코더(QGCAE) 모델을 제안하여 사기 거래의 분리된 표현을 학습한다. 우리의 모델은 기존 모델과의 성능 비교 실험을 통해 F1-score에서 기존의 최고 성능 대비 0.152%p의 향상을 이루어 내고, 분리된 표현 학습을 확인함으로써 쿼드로플렛 손실이 암호화폐의 사기와 정상간의 미묘한 패턴 인식 및 분리에 유효함을 입증한다.

### Abstract

As cryptocurrency usage rises, so does fraudulent activity, making accurate scam detection crucial for establishing trust. However, due to similar transaction patterns and data imbalance, existing methods struggle with precise scam detection. To address these challenges, this paper introduces the Quadruplet Loss-based Graph Convolutional Autoencoder (QGCAE) model, which combines quadruplet loss and a graph convolutional autoencoder to effectively capture and distinguish subtle patterns between scam and benign transactions. Our model achieves a 0.152 percentage point improvement in F1-score over previous best-performing models, demonstrating the efficacy of quadruplet loss in recognizing and separating subtle patterns between scam and benign transactions in cryptocurrency.

### Key words

Disentangled Representation Learning, Quadruplet Loss, Graph Convolutional Autoencoder, Scam Detection

---

\* 경상국립대학교 컴퓨터공학과(Dept. of Computer Engineering, Gyeongsang National University).  
{2019010878, ouo\_ofo, sjjuu\*}@gnu.ac.kr

## I. 서론

최근 암호화폐의 사용은 블록체인 기술의 발전, 탈중앙화된 플랫폼, 저렴한 수수료, 그리고 스마트 계약의 효율성에 힘입어 전세계적으로 급증하고 있다. 하지만 이와 함께 암호화폐 사기도 증가하고 있으며, 이는 암호화폐의 신뢰성을 위협하기에, 암호화폐 거래에서의 사기를 효과적으로 탐지할 수 있는 모델의 개발은 필수적이다. 하지만 암호화폐 거래에서 사기와 정상 거래는 유사한 패턴을 보이기에 기존의 모델들은 이러한 차이를 잘 식별하고 정확하게 탐지하는데 한계가 존재한다.

본 연구에서는 사기 거래와 정상 거래의 미묘한 차이를 명확히 구분하고 분리하기 위한 퀴드로플렛 손실과 그래프 컨볼루션 오토인코더를 결합한 퀴드로플렛 손실 기반 그래프 컨볼루션 오토인코더(QGCAE) 방법을 제안한다. 제안된 접근법은 정상 거래와 사기 거래 간의 미묘한 차이를 효과적으로 식별하여, 기존 방법들의 최대 F1-score 대비 0.152%p의 성능 향상을 이루어 내어 미묘한 패턴 인식 및 분리의 유효성을 입증한다.

표 1. 암호화폐 거래내역에서 사기탐지 연구

Table 1. Related works for scam detection on ryptocurrency Transaction network

Approach	Method	Dataset Source (Scam/Benign/Ratio)	Performance
Machine Learning	XGBoost with 10-fold CV [1]	Labeled by Etherscamdb (2,179/2,502/46.54%)	F1: 0.96 AUC: 0.99
	Augmentation (SMOTE) + LightGBM [2]	Ethereum Classic Blockchain (482/1,697/22.12%)	F1: 0.97 Precision: 0.94
Graph Traversing	Embedding (Node2Vec) + classifier [3]	Labeled by Etherscamdb&Etherscan (1,866/1,175/61.36%)	F1: 0.95 Precision: 0.96 Recall: 0.94
	Embedding (DeepWalk) + Transformer [4]	Labeled by Etherscan (1,165/2,972,324/0.03%)	F1: 0.93 Precision: 0.90 Recall: 0.96
Graph Neural Network	Embedding (GCN Autoencoder) + classifier [5]	Labeled by Etherscan (1,157/2,972,225/ 0.03%)	F1: 0.23 Precision: 0.72 Recall: 0.14 AUC: 0.57
	GRU-GAT based on GRU + GAT [6]	Elliptic dataset (4,545/42,019/9.76%)	F1: 0.97 Precision: 0.95 Recall: 0.99
	SIEGE based on GraphSAGE [7]	Labeled by Etherscan&Blacklists (6,588/~19,764/~25.00%)	F1: 0.72 Precision: 0.64 Recall: 0.82 AUC: 0.83

## II. 관련 연구

표 1은 암호화폐 거래내역에서 사기를 탐지하기 위한 기존 연구들을 종합적으로 정리한다. 기존 연구들은 전통적인 기계학습 기법, 노드 임베딩 생성을 위한 그래프 탐색, 그래프 구조를 활용한 그래프 신경망 등 다양한 접근 방법을 탐구해온 것을 확인 가능하다. 최근에는 그래프 구조를 활용한 그래프 신경망 기법이 특히 주목받고 있으며, 이는 복잡한 네트워크 구조에서 데이터를 효과적으로 처리할 수 있는 가능성을 제공한다.

본 연구에서는 암호화폐 거래내역의 불균형을 고려한 그래프 컨볼루션 오토인코더(GCAE)를 사용하며, 사기 거래와 정상 거래 간의 미묘한 차이를 효율적으로 포착하기 위해 퀴드로플렛 손실을 결합한다. 이러한 접근법을 통해 암호화폐 거래에서 사기 거래와 정상 거래 간의 미묘한 패턴을 보다 효율적으로 포착하고, 사기 거래를 효과적으로 탐지할 수 있는 방법론을 제시한다.

### III. 제안하는 방법

#### 3.1 제안한 방법의 개요

그림 1처럼 쿼드로플렛 손실 기반 그래프 컨볼루션 서널 오토인코더로 거래 패턴을 분리한 재구축 그래프와 잠재공간을 생성한다. 이후, 주성분 분석을 적용하여 추가적인 특성을 도출하며, 재구축 그래프, 잠재공간 및 원본 그래프와 합쳐 새로운 그래프 데이터를 생성하여 그래프 샘플링 기법을 활용하여 복잡한 그래프 데이터 학습에 효율적인 GraphSAGE 모델을 사용하여 최종 탐지한다.

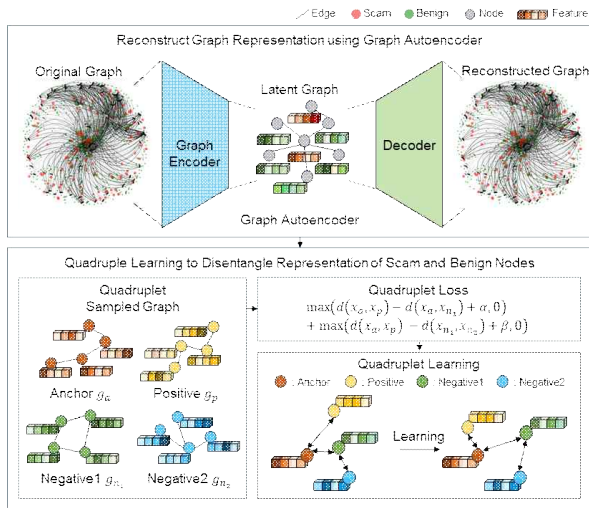


그림 1. 쿼드로플렛 손실 기반 그래프 재구축  
Fig. 1. Quadruplet Loss-based graph reconstruction

#### 3.2 표현 분리 학습을 위한 쿼드로플렛 손실

본 연구에서는 쿼드로플렛 손실을 통해 사기와 정상간의 미묘한 차이를 포착하여 효과적으로 분리한다. 쿼드로플렛 손실은 수식(1)와 같이 정의한다.

$$L_{quadruplet} = \max(d(x_a, x_p) - d(x_a, x_{n_1}) + \alpha, 0) + \max(d(x_a, x_p) - d(x_{n_1}, x_{n_2}) + \beta, 0) \quad (1)$$

$x_a, x_p$ 는 사기 노드 클래스의 anchor와 positive 샘플이며,  $x_{n_1}, x_{n_2}$ 는 정상 노드 클래스의 negative 샘플들이다.  $d(x_1, x_2)$ 는 두 샘플  $x_1, x_2$  사이의 거

리 함수이며,  $\alpha$ 는 anchor와  $n_1$ 간의 최소 마진,  $\beta$ 는 두 샘플  $n_1, n_2$  간의 최소 마진이다.

첫 번째 항은 anchor 샘플  $x_a$ 와 positive 샘플  $x_p$  사이의 거리가  $x_a$ 와 첫 번째 negative 샘플  $x_{n_1}$  사이의 거리보다 최소  $\alpha$  만큼 더 작아지도록 하여 사기 노드 간의 거리가 정상 노드와의 거리보다 가까워지게 한다.

두 번째 항은 두 negative 샘플  $x_1, x_2$  사이의 거리가 anchor와 positive 샘플 간의 거리보다 최소  $\beta$  만큼 더 크게 하여 사기 노드들이 더 밀집되게 모이도록 한다. 이는 사기 노드들이 더 밀집되게 모이도록 함으로써, 쿼드로플렛 손실을 통한 재구축 그래프에서 사기 노드들이 명확하게 구분될 수 있도록 한다.

#### Algorithm 1. Quadruplet Sampling with Hard Negative Mining

The algorithm samples training quadruplets consisting of two positive Scam samples and two negative Benign samples, with half the negatives selected via hard negative mining.

**Input:** (Array of class label), (Number of quadruplets to be sampled)

**Output:** (Sampled quadruplets)

1. Initialize:
  - Create mapping of each to its corresponding indices.
2. Sample Quadruplets:
  - For each iteration
    - a. Select a label for the positive(Scam) pair and a different label for the negative(Benign) pair.
    - b. Sample two indices from and two indices from .
    - c. Append the quadruplet to the list .
3. Hard Negative Replacement:
  - For the first quadruplets:
    - a. Replace the negative pair with two indices from a hard negative label , chosen based on proximity to the positive pair .
4. Return:
  - Return the list of quadruplets.

#### 3.3 그래프 컨볼루션 오토인코더 기반

그래프 재구축 및 쿼드로플렛 인덱스 샘플링

우리의 방법은 암호화폐 거래내역에서 사기 거래가 정상 거래에 비해 적은 빈도 불균형 문제를 고려한다. 이를 위해 그래프 컨볼루션 오토인코더를 활용하여 사기와 정상 거래를 효과적으로 분리하는 그래프를 재구축 하게 된다. 그래프 컨볼루션 오토

토인코더의 입력과 출력은 하나의 큰 암호화폐 거래내역 그래프이다. 하지만 쿼드로플렛 손실의 계산을 위해서는 사기 노드의 anchor, positive 샘플과 정상 노드의 negative 1,2, 샘플들이 필요하다. 이를 위해 재구축된 그래프의 인덱스 샘플링이 필요하며, 이는 알고리즘 1에서 정리한다. 샘플링 과정 중 인덱스의 절반은 hard negative mining 기법을 통해 결정되며, 이는 모델이 혼동하기 쉬운 anchor, positive의 표현과 가까운 거리에 있는 부정 샘플을 추출하여 어려운 negative 샘플들에 대한 추가적인 손실을 계산하여 사기 클래스와 정상 클래스 간의 표현을 더 확고히 분리시켜 주게 된다.

이렇게 샘플링된 인덱스들을 사용하여 쿼드로플렛 손실을 계산하게 된다. 이후, 이 손실을 기반으로 그래프 컨볼루션 오토인코더를 학습하게 되며, 최종적으로 사기 및 정상 거래 간의 미묘한 패턴차이를 식별하여 명확한 구분이 가능하도록 하는 재구축 그래프를 생성하게 된다.

#### IV. 실험

##### 4.1 이더리움 데이터셋 특성 확인

표 2. 이더리움 거래내역 데이터셋 특성

Table 2. Ethereum transaction network dataset features

	Num Nodes	Average in-degree	Avg out-degree	Ratio(total nodes)	Ration(total transaction)
Scam	2,972,324	4.55	4.55	99.97%	99.73%
Benign	1,165	32.19	20.43	0.03%	0.27%

표 3. 노드 크기 별 비교군 모델과의 비교 실험

Table 3. Comparative Experiment with the Comparison Model Based on Node Size

Approach	Node Size	5,000			10,000			20,000			30,000		
		Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Machine Learning	XGBoost	0.860	0.827	0.843	0.832	0.705	0.763	0.823	0.629	0.713	0.803	0.550	0.653
	LightGBM	0.849	0.801	0.824	0.861	0.714	0.781	0.817	0.656	0.728	0.808	0.590	0.682
Graph Traversing	DeepWalk	0.550	0.852	0.668	0.537	0.626	0.578	0.631	0.846	0.723	0.710	0.882	0.786
	Node2Vec	0.507	0.872	0.641	0.525	0.858	0.651	0.637	0.606	0.621	0.690	0.646	0.667
Graph Neural Network	GCN	0.873	0.856	0.864	0.769	0.768	0.768	0.799	0.782	0.790	0.810	0.832	0.821
	GAT	0.839	0.856	0.847	0.780	0.804	0.792	0.604	0.836	0.701	0.588	0.844	0.693
	Ours(w/o Q-Loss)	0.964	<b>0.964</b>	0.964	0.917	0.910	0.913	0.943	0.930	0.936	0.934	<b>0.934</b>	0.934
	<b>Ours (QGCAE)</b>	<b>0.969</b>	0.962	<b>0.965</b>	<b>0.934</b>	<b>0.912</b>	<b>0.923</b>	<b>0.947</b>	<b>0.938</b>	<b>0.942</b>	<b>0.960</b>	0.926	<b>0.942</b>

본 연구에서 사용한 암호화폐는 비트코인에 이어 많은 거래량을 가지는 이더리움 거래내역 데이터셋이다. 데이터셋은 1,165개의 사기 노드, 2,972,324개의 정상 노드로 구성되어 있으며, 총 13,551,303개의 엣지로 구성되어 있다. 데이터셋의 추가적인 정보는 표 2에서 설명한다.

##### 4.2 기존 모델들과 노드 크기별 비교 실험

전체 데이터셋은 총 2,972,324개의 방대한 노드가 있기 때문에, 사기 노드를 기준으로 하여 너비 우선 탐색(BFS)으로 그래프 크기를 설정한다. 그래프 크기는 그래프의 노드 개수로 설정되며, 비교 실험을 위해 5,000, 10,000, 20,000, 30,000 개의 노드를 사용한다. 표 3에서는 각 노드의 개수에 따라 기존 모델들의 성능을 확인한다. 우리의 모델은 Precision, Recall, F1-score 등에서 기존의 모델들의 성능을 능가한다. 특히, 쿼드로플렛 손실을 사용하지 않은 모델과의 성능 비교를 통해, F1-score, Precision의 큰 성능 향상을 확인 가능하며 이는 쿼드로플렛 손실이 사기와 정상간의 미묘한 패턴을 효과적으로 구별하는데 기여함을 시사한다.

### 4.3 쿼드로플렛 손실 적용 여부에 따른 모델별 탐지 오류 케이스 분석

그림 2와 표 4는 쿼드로플렛 손실을 적용하지 않은 비교군 모델(표 3의 w/o Q-Loss 모델)과 우리의 모델인 쿼드로플렛 손실 기반 그래프 컨볼루션서널 오토인코더 모델이 오탐지한 경우를 보여주며, Case 1은 비교군 모델만 오탐지한 경우이며, Case 2는 쿼드로플렛 손실 적용 여부와 상관없이 비교군과 우리 모델 전부 탐지에 실패한 경우이다.

Case 1에서는, 암호화폐 거래 중 연결된 노드와 거래량이 많은 복잡한 거래에 대해서 쿼드로플렛 손실이 적용된 모델(QGCAE)은 사기 및 정상 거래의 차이를 잘 포착하여 탐지한 반면, 쿼드로플렛 손실이 적용되지 않은 모델(비교군)은 복잡한 거래에 대한 정확한 탐지가 힘든 것을 확인한다. Case 2는 연결된 노드와 거래량이 적은 간단한 거래에 대해서 우리 모델과 비교군 모델 모두 정확한 탐지가 어려움을 확인한다.

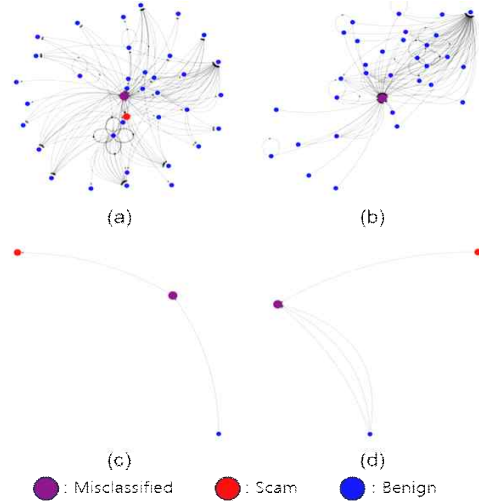


그림 2. 모델의 쿼드로플렛 손실 여부별 탐지 오류 케이스 분석. (a), (b): 쿼드로플렛 손실 없는 비교군 모델의 탐지 오류, (c), (d): 비교군과 우리의 모델 모두 탐지 오류  
Fig. 2. Case Analysis of Detection Errors Based on the Presence of Quadruplet Loss, (a), (b): Detection Errors of Comparison Model Without Quadruplet Loss, (c), (d): Detection Errors from Both the Comparison Model and Our Model

### 4.4 t-SNE 활용 분리 확인

그림 3은 t-SNE 기법을 사용하여 암호화폐 거래 내역의 원본 그래프와 재구성된 그래프의 피쳐 공간을 시각화한 결과이다. 원본 그래프 (a)에서는 사기와 정상 피쳐 간의 명확한 구분이 어렵다. 그러나 재구성 그래프 (b)에서는 사기와 정상 피쳐 간의 경계와 군집이 나타난다. 이는 쿼드로플렛 손실이 사기와 정상의 미묘한 패턴을 인식하여 구분하고 있다는 것을 뒷받침한다.

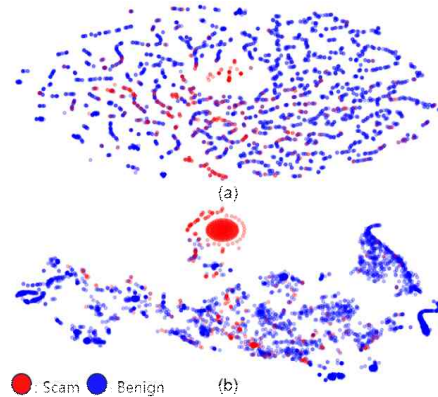


그림 3. t-SNE 활용한 피쳐 공간 시각화, (a) 원본 그래프 (b) 재구성된 그래프  
Fig. 3. Feature space visualization using t-SNE, (a) original graph (b) reconstructed graph

표 4. 쿼드로플렛 손실 여부에 따른 탐지 오류 케이스 상세 분석

Table 4. Detailed Analysis of Detection Errors Based on the Presence of Quadruplet Loss

Case	Account	Nodes	Edges	Transaction (ETH)	ID
Case 1. Detection Errors Comparison Model (without Quadruplet Loss)	0x3a878eddb991ebcbc7c8052055b2e5ed5d0d1ba4	In: 8 Out: 35	In: 8 Out: 60	In: 2.70 Out: 3.09	(a)
	0x8bfe5ebb128ee82f4ba80f56bb32409cc87bc6fb	In: 30 Out: 1	In: 47 Out: 1	In: 211.63 Out: 0	(b)
Case 2. Detection Errors Comparison Model and Our Model	0x80ff1325226351c65c15b6eaa40c46122203a28f	In: 1 Out: 1	In: 1 Out: 1	In: 11.76 Out: 11.76	(c)
	0x5700a44aaf77c76fad4f05d780258ef44c1790e1	In: 2 Out: 0	In: 4 Out: 0	In: 1.89 Out: 0	(d)

#### 4.5 결론 및 향후 연구

본 연구는 퀴드로플렛 손실 기반 그래프 컨볼루션 서널 오토인코더(QGCAE) 모델을 제안하여 이더리움 거래 내역에서 사기 거래와 정상 거래의 미묘한 패턴을 효과적으로 식별하여 구분할 수 있음을 보여준다. 제안된 모델은 기존 방법들보다 우수한 성능을 보였으며, 암호화폐 거래 내역에서 퀴드로플렛 손실의 활용 가능성을 입증한다.

다만, 본 연구에서는 한계점도 발견된다. 표 4의 Case 2에서 연결된 노드와 거래 수가 적은 거래 내역에서 탐지 정확도가 낮은 현상이 나타난다. 이러한 한계를 해결하기 위해, 향후 연구에서는 선택적으로 가중치를 반영하기 위한 어텐션 메커니즘과 연결성을 증강하기 위한 GANs, SMOTE 같은 데이터 증강 기법을 통해 연결성이 부족한 거래내역에서도 높은 탐지 성능을 보일 수 있는 방법을 검토할 예정이다.

#### 감사의 글

본 과제(결과물)는 교육부와 한국연구재단의 지원으로 지원을 받아 수행된 3단계 산학연협력 선도 대학 육성사업(LINC 3.0)의 연구결과입니다.

#### 참 고 문 헌

- [1] S. Farrugia, J. Ellul and G. Azzopardi, "Detection of Illicit Accounts Over the Ethereum Blockchain," *Expert Systems with Applications*, vol. 150, p. 113318, 2020.
- [2] R. M. Aziz, M. F. Baluch, S. Patel and A. H. Ganie, "LGBM: A Machine Learning Approach for Ethereum Fraud Detection," *Int. Journal of Information Technology*, vol. 14, pp. 3321-3331, 2022.
- [3] R. Tan, Q. Tan, P. Zhang and Z. Li, "Graph Neural Network for Ethereum Fraud Detection." *IEEE Int. Conf. on Big Knowledge*, pp. 78-85, 2021.

- [4] S. H. Choi and S. J. Buu, "Learning to Traverse Cryptocurrency Transaction Graphs Based on Transformer Network for Phishing Scam Detection," *Electronics*, vol. 13, p. 1298, 2024.
- [5] L. Chen, J. Peng, Y. Liu, J. Li, F. Xie and Z. Zheng, "Phishing Scams Detection in Ethereum Transaction Network," *ACM Trans. on Internet Technology*, vol. 21, pp. 1-16, 2020.
- [6] Y. Zheng, "GRU-GAT Model for Blockchain Bitcoin Abnormal Transaction Detection," *IEEE Conf. on Telecommunications, Optics and Computer Science*, pp. 666-674, 2022.
- [7] S. Li, R. Wang, H. Wu, S. Zhong and F. Xu, "SIEGE: Self-Supervised Incremental Deep Graph Learning for Ethereum Phishing Scam Detection," *Proc. of the ACM Int. Conf. on Multimedia*, pp.8881-8890, 202



# Mamba를 활용한 하향식 자세 추정 방법 연구

원지예\*, 김인수\*, 문준범\*, 정순기\*\*

## Top-down Pose Estimation method using Mamba Architecture

Jiye Won\*, InSu Kim\*, Junbeom Moon\*, and SoonKi Jung\*\*

### 요 약

본 논문은 자세 추정 연구를 위해 Mamba를 기반으로 한 MambaVision 아키텍처를 활용한 방법을 제안하였다. 제안한 방법은 하향식 자세 추정 방법으로써 사람을 검출하고 검출된 대상에 대해서 주요 부위를 히트맵 형식으로 추정할 수 있다. 제안한 모델을 MS COCO 데이터셋에 대해 학습하고 테스트 데이터에 대해 추정한 결과 제안하는 모델이 검출된 사람에 대해서 각 주요 부위에 대한 히트맵을 잘 추정하는 것으로 확인하였다.

### Abstract

In this paper, we proposed a method utilizing the MambaVision architecture based on Mamba for posture estimation research. The proposed method is a top-down pose estimation method that can detect a person and estimate the main parts of the detected object in the form of a heatmap. By training the proposed model on the MS COCO dataset and estimating it on test data, we found that the proposed model can estimate the heatmap of each major part of the detected person well.

### Key words

pose estimation, artificial neural networks, deep learning

## 1. 서 론

자세 추정 연구는 이미지나 영상에서 탐색한 사람 객체에 대해 인체의 여러 부위를 추정하고 그 상관 관계를 구하는 연구이다. 이 기술은 헬스케어, 증강현실, 스포츠 등 사람 객체가 나타나는 다양한 분야에서 활용할 수 있어 그 활용 방안이 무궁무진

하다. 딥러닝 기반의 자세 추정 기술은 합성곱 신경망을 기반으로 한 High-Resolution Net(HRNet)[1], Open Pose[2], Deep Pose[3]과 같은 모델에서부터 시작하여 최근에는 Transformer 아키텍처를 기반으로 하는 ViTPose[4], PCT[5] 등의 모델에 이르기까지 다양한 모델 구조와 기법들이 연구되었다. 자세를 추정하는 방법으로는 크게 골격 기반 모델, 모양 기

\* 경북대학교 IT 대학 컴퓨터학부, oneone8773@gmail.com, sky5166@knu.ac.kr, jpm04135@knu.ac.kr

\*\* 경북대학교 IT 대학 컴퓨터학부, skjung@knu.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 정보통신·방송 기술개발 사업의 일환으로 하였음  
IRS-2024-00336663, 다중 영상 기반 마약사범 추적 및 검거를 위한 AI 기술 개발

반 모델, 볼륨 기반 모델 세가지로 나누어진다. 이들 중 2D 이미지에서는 일반적으로 골격 기반 모델이 많이 사용된다. 또한 학습 방법으로는 하향식과 상향식 두 방법으로 학습이 이루어진다.

현재 주류로 평가받는 Transformer 기반의 다양한 모델들은 정확도면에서 많은 이점이 있으나 Transformer 아키텍처가 가지는 고질적인 문제점인 수 많은 파라미터로 인한 메모리 사용량과 긴 처리 시간으로 인해 이를 대체하기 위한 다양한 접근법들이 연구되고 있다. 이러한 접근법들 중 상태 공간 모델(SSM)을 이용한 방법은 Transformer와 비교하여 상대적으로 적은 메모리 사용량과 긴 시퀀스 데이터에 대해 더 나은 성능을 보인다는 점에서 각광받고 있다. 상태 공간 모델의 일종인 Mamba[6]는 현재 모델 파라미터가 입력 크기에 대응하여 변화할 수 있다는 장점을 통해 Transformer를 대체할 기반 모델 구조로 주목받고 있다.

본 논문에서는 Mamba 기반의 MambaVision[7]을 이용한 자세 추정 방법을 제안한다. 제안하는 모델은 MambaVision 기반의 Encoder를 통해 특징을 추출하고 제안하는 디코더를 사용하여 하향식 학습 방법을 통해 자세를 추정하였다.

## II. 관련 연구

자세 추정은 입력 영상에서 객체의 자세를 추정하는 연구이다. 자세에서 주요 부위를 추정하는 방법에는 골격에 맞게 추정하는 골격 기반 모델, 전체 모양을 추정하는 모양 기반 모델 그리고 주로 3D 자세 추정에서 활용되는 볼륨 기반 모델이 있다. 또한, 영상 내의 여러 사람에 대한 자세를 추정하는 방법에 따라 하향식과 상향식 방법으로 나누어진다. 하향식은 먼저 사람 객체를 검출하고 검출된 객체의 주요 부위를 추정하는 방법이며, 상향식은 영상에서 주요 부위들을 먼저 추정하고 이를 연결해서 자세를 추정한다. 본 논문에서는 하향식 및 골격 기반 모델로 자세를 추정하는 방법을 제안한다.

Mamba 아키텍처는 상태 공간 모델(SSM)을 기반으로 한 구조로, Transformer에 비해 대규모의 컨텍스트, 긴 시퀀스에 대해서도 강건하게 동작하며 유사한 성능을 보여주면서도 상대적으로 추론 속도가 빠르다는 장점을 가지고 있다. 본 논문에서는 자세 추정을 위해서 Mamba를 적용하여 자세 추정에서의 해당 구조의 유용성을 확인해보고 이를 통해 향후 후속 연구의 기반을 다지기 위한 연구를 수행한다.

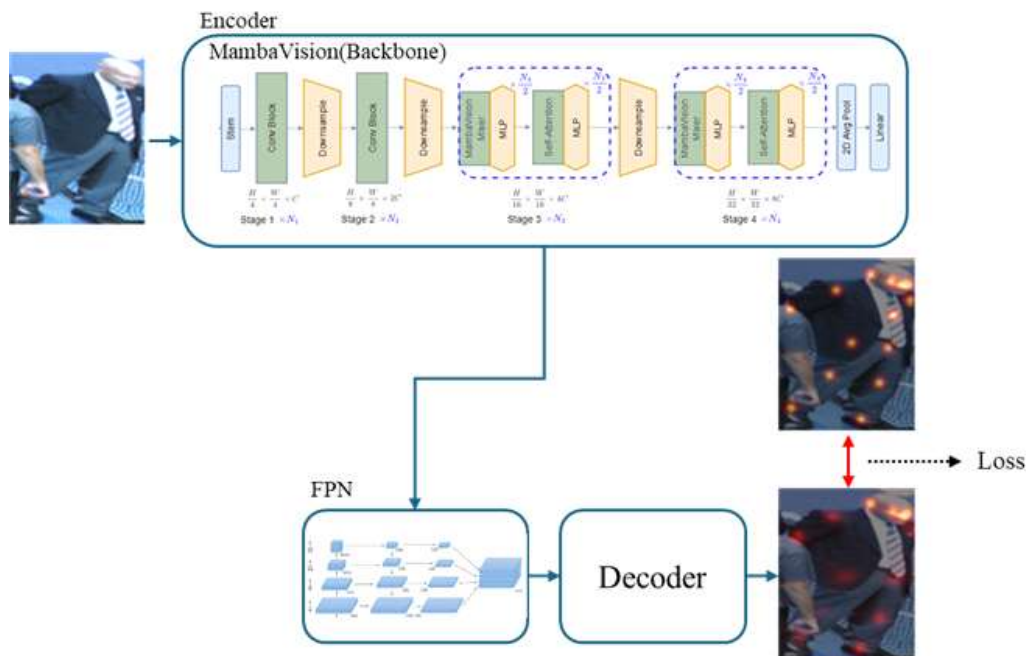


그림 1. 제안하는 방법의 전체 구성도

### III. 제안하는 방법

본 논문에서는 자세 추정을 위한 MambaVision 기반의 모델 구조를 제안한다. 그림 1에서 입력 이미지는 별도의 사람 객체 검출 모듈을 통해 각각의 사람 이미지로 분리된다. 이후 MambaVision 기반의 인코더를 통해 사람 이미지로부터 특징맵을 추출한다. 마지막으로 디코더에서는 추출된 특징맵을 이용하여 신체 주요 부위 추정을 수행한다.

사람 객체 검출 모듈은 입력된 이미지에서 사람을 검출하고 전체 이미지에서 검출된 사람에 해당하는 영역만 따로 잘라내서 인코더로 전달하는 역할을 수행한다. 인코더에서는 사람의 주요 부위(주요 관절, 얼굴 요소 등)를 추정하기 위한 특징맵을 추출한다. MambaVision은 4단계에 걸쳐 서로 다른 크기의 특징맵을 추출하는데 이를 디코더에서 활용하기 위해서 Feature Pyramid Network(FPN) 모듈을 추가하였다. 이후 특징맵을 결합하고 역합성곱 신경망을 통해 원본의 이미지로 복원한 후 최종 1x1 합성곱 신경망으로 신체의 주요 부위를 히트맵 방식으로 추정하였다.

자세 추정 모델의 학습에는 Adaptive Wing Loss(AWing Loss) 손실 함수를 사용하였다.

$$AWing(p, \hat{p}) = \begin{cases} w \ln(1 + \frac{p - \hat{p}}{c} |^{\alpha - p}) & \text{if } |p - \hat{p}| < \theta, \\ A|p - \hat{p}| - C & \text{otherwise,} \end{cases} \quad (1)$$

디코더에서 추론된 결과 히트맵과 ground-truth 히트맵간의 손실 함수를 통해 학습이 된다.

$$L_{AWing}(P, \hat{P}) = \frac{1}{K} \sum_{k=1}^K \sum_{j=1}^{H'} \sum_{i=1}^{W'} AWing(p_k^{i,j}, \hat{p}_k^{i,j}). \quad (2)$$

### IV. 실험 결과

제안된 모델은 MS COCO 데이터셋을 이용한 학습을 진행하였다. MS COCO 자세 추정 데이터셋은 이미지 내의 사람에 대한 경계 상자의 정보와 함께 17개의 주요 부위에 대한 정보를 가지고 있다. 본 논문에서 제안하는 모델은 위 데이터셋을 사용하여 64의 배치 사이즈 및 50 epoch로 학습되었다.

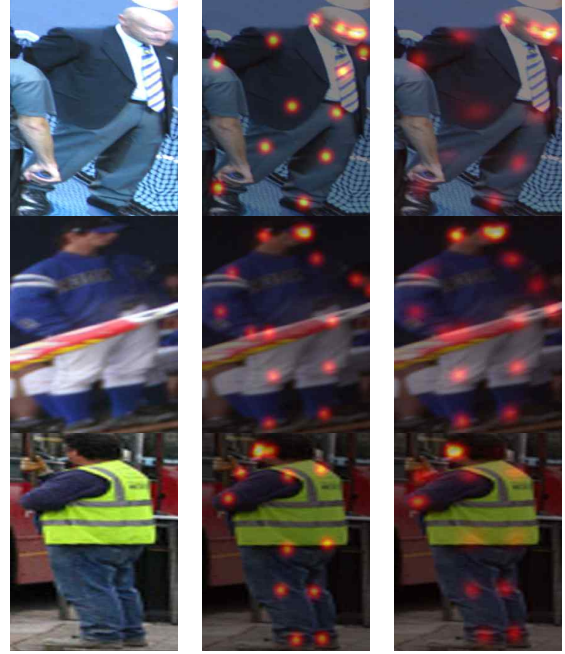


그림 2. 실험 결과

그림 2에서 본 논문에서 제안하는 방법을 통해 추정된 히트맵의 결과를 확인할 수 있다. 가장 왼쪽 열은 원본 입력 이미지이며 두 번째 열은 ground-truth 이미지이다. 세 번째 열이 제안한 방법으로 예측된 히트맵 이미지이다. 그림에서 확인할 수 있듯이 비교적 신체 특징이 뚜렷한 얼굴에 대한 히트맵이 확연히 밝게 나타나고, 그 외의 주요 부위들에도 잘 나타나는 것을 확인할 수 있다. 또한 일부 보이지 않거나 겹쳐있는 부위에 대해서도 자세 추정이 가능한 것을 확인할 수 있다. 이러한 실험 결과를 통해 본 논문에서 제안한 Mamba 아키텍처 기반의 모델을 이용한 자세 추정이 가능하다는 것을 확인할 수 있다.

### V. 결론

본 논문에서는 하향식 방법의 자세 추정을 위해 Mamba 아키텍처 기반의 모델을 제안하였다. 제안된 모델은 자세 추정을 위해 MambaVision의 4단계 특징맵을 FPN을 통해 결합하고 역합성곱 신경망으로 주요 부위 히트맵을 생성하여 자세 추정을 수행한다. MS COCO 데이터셋을 이용한 실험 결과, 제안된 모델이 주요 부위에 대한 추정을 수행할 수 있음을 입증 하였다. 향후 자세 추정을 위한 효과적

인 방법을 위해 Mamba 아키텍처를 활용한 다양한 연구를 통해 자세 추정에 대한 성능을 향상 시키고자 한다.

### 참 고 문 헌

- [1] Wang, Jingdong, et al. "Deep high-resolution representation learning for visual recognition." IEEE transactions on pattern analysis and machine intelligence 43.10 (2020): 3349-3364.
- [2] Martinez, Ginés Hidalgo. Openpose: Whole-body pose estimation. Diss. Carnegie Mellon University, 2019.
- [3] Toshev, Alexander, and Christian Szegedy. "Deeppose: Human pose estimation via deep neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.
- [4] Xu, Yufei, et al. "Vitpose: Simple vision transformer baselines for human pose estimation." Advances in Neural Information Processing Systems 35 (2022): 38571-38584.
- [5] Guo, Meng-Hao, et al. "Pct: Point cloud transformer", Computational Visual Media 7 (2021): 187-199.
- [6] Gu, Albert, and Tri Dao. "Mamba: Linear-time sequence modeling with selective state spaces." arXiv preprint arXiv:2312.00752 (2023).
- [7] Hatamizadeh, Ali, and Jan Kautz. "Mambavision: A hybrid mamba-transformer vision backbone." arXiv preprint arXiv:2407.08083 (2024).

# 생성형 AI를 활용한 개인 맞춤형 문제 생성 시스템 개발

허동빈, 유원근, 하현수, 김현근, 박효미, 송정화, 조호연, 허성연, 최한용\*\*

## Development of a Personalized Problem Generation System Using Generative AI

Dongbin Heo, Wongeun Yu, Hyunsoo Ha, Hyungeun Kim, Hyomi Park, Jeonghwa Song, Houyeon Cho, Seongyeon Heo, and Hanyong Choi\*\*

### 요약

본 연구는 OCR 기술과 GPT-4 기반 생성형 AI를 결합하여 맞춤형 문제집과 해설을 자동으로 생성하는 혁신적인 시스템을 설계하고 개발하였다. 기존의 문제 생성 서비스들은 미리 구축된 데이터베이스에 제약을 받으며 문제를 생성하였다. 본 시스템은 사용자가 제공하는 이미지나 PDF의 문제 유형을 실시간으로 인식하여 유사한 새로운 문제를 생성한다. 따라서 사용자의 특정 요구사항과 학습 목표에 완벽히 부합하는 무한히 다양한 문제와 해설을 제공함으로써, 개인화 교육의 새로운 패러다임을 제시한다.

### Abstract

This study designed and developed an innovative system that automatically generates personalized problem sets and explanations by combining OCR technology with GPT-4-based generative AI. Existing problem generation services were limited to creating problems from pre-built databases. This system recognizes problem types in real-time from user-provided images or PDFs and generates similar new problems. As a result, it presents a new paradigm for personalized education by providing an infinite variety of problems and explanations that perfectly align with users' specific requirements and learning objectives.

### Key words

OCR, Generative AI, Customized Problem Generation, Personalized Learning, Educational Technology

### I. 서론

현재 시중에 나와 있는 문제집이나 학습 자료는

대부분 표준화된 내용을 담고 있어, 개별 학습자의 수준과 요구를 충족시키기에는 한계가 있다. 특히 대학생들을 위한 맞춤형 학습 도구는 상대적으로

---

신한대학교 소프트웨어융합학과, mailsky9687@gmail.com, ywg6525@icloud.com, {huns0905, hyungeun0313, phyomi3, sjhw7200, cesunny54, hsy3160}@naver.com, hychoi@shinhan.ac.kr

※ "본 연구는 2023년도 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업지원을 받아 수행되었음" (2023-0-00089)

부족한 실정이다[1,2]. 따라서 AI 기술을 활용한 맞춤형 문제집 생성 서비스는 대학생들의 학습 효율성을 높일 수 있는 중요한 도구가 될 수 있다. 기존의 문제 생성 서비스들은 주로 미리 구축된 문제 데이터베이스를 활용하여 인공지능 기술로 문제를 선별하거나 변형하는 방식을 채택해 왔다[3,4]. 이러한 접근 방식은 대규모 문제 풀을 제공할 수 있다는 장점이 있으나, 사용자의 특정 요구사항이나 학습 목표에 완벽히 부합하는 문제를 생성하는 데에는 한계가 있었다[5-8]. 본 연구에서는 이러한 한계를 극복하고자 OCR(Optical Character Recognition) 기술과 최신 생성형 AI를 결합한 혁신적인 맞춤형 문제 생성 서비스를 제안한다. 본 시스템의 가장 큰 특징은 사용자가 제공하는 이미지나 PDF 형태의 문제 유형을 OCR 기술로 정확히 인식하고, 이를 바탕으로 GPT-4 기반의 생성형 AI가 유사한 유형의 새로운 문제를 실시간으로 생성한다는 점이다. 이를 통해 사용자는 자신이 원하는 특정 유형의 문제를 즉각적으로 얻을 수 있으며, 기존 데이터베이스의 한계에 구애받지 않고 다양한 유형의 문제를 생성할 수 있게 된다.

## II. 시스템 구성

### 2.1. 시스템 블록 다이어그램

PDF 또는 이미지 파일로 데이터를 입력받으면 OCR 기술을 통해 텍스트를 추출한다. 그림 1과 같이 추출된 텍스트를 생성형 AI에 프롬프트에 맞는 형식으로 변환하기 위해 텍스트를 처리하는 과정을 거친 후 생성형 AI를 통해 비슷한 유형의 문제집과 답안지를 생성한다.

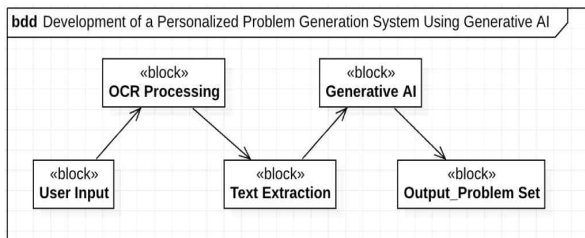


Fig. 1. System Block Diagram

### 2.2. 시퀀스 다이어그램

그림 2와 같이 사용자는 출제희망 자료 정보를 PDF나 이미지 파일로 메인 시스템에 업로드한다. 메인 시스템은 OCR 기술을 통해 해당 파일에서 텍스트를 추출하는 과정을 거친 후 생성형 AI의 API를 호출하여 유사 유형의 문제를 생성하여 사용자에게 문제집을 출력한다.

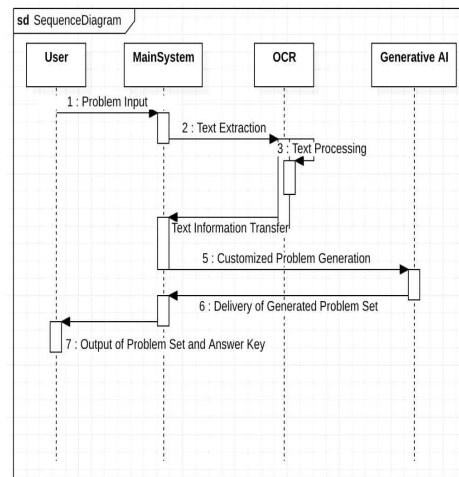


Fig. 2. System Sequence

## III. 시스템 구현

### 3.1 OCR 기술 구현

본 연구에서 개발한 맞춤형 문제 생성 시스템은 웹 플랫폼과 안드로이드 앱 두 가지 버전으로 구현되었으며, 각각 다른 OCR 기술을 적용하여 최적의 성능을 도출하였다. 웹 버전에서는 Google이 개발한 오픈소스 OCR 엔진인 Tesseract를 활용하였다. Tesseract는 다양한 언어와 폰트에 대해 우수한 인식률을 보이며 특정 요구사항에 맞게 최적화가 가능하다. Tesseract의 처리 과정은 다음과 같다. 이미지 전처리 → 텍스트 추출 → 데이터 변환 → 출력모바일 환경에 최적화된 성능을 위해 안드로이드 앱에서는 Google의 ML Kit을 사용하였다. ML Kit은 모바일에 최적화되어 있어 온디바이스 처리로 빠른 속도와 오프라인 사용이 가능하며 고급 머신러닝 모델을 쉽게 사용할 수 있는 API를 제공한다. 무엇

보다 카메라 입력의 실시간 처리가 가능하다는 장점이 있다. 이러한 OCR 기술들을 효과적으로 활용함으로써, 본 시스템은 다양한 형태의 문제 이미지로부터 정확하고 신속하게 텍스트를 추출하여 맞춤형 문제 생성의 기반을 마련하였다.

### 3.2 생성형 AI 활용

GPT-4o-mini 모델을 활용하여 고품질의 맞춤형 문제를 생성하는 시스템을 구현하였다. 이 과정에서 프롬프트 엔지니어링 기법을 적용하여 OCR을 통해 인식된 문제 유형을 정확히 파악하고, 이를 바탕으로 유사한 새로운 문제를 생성하는 데 성공하였다. GPT-4o-mini는 OpenAI에서 개발한 GPT-4의 경량화 버전으로 높은 언어 이해도와 텍스트 생성, 효율적인 리소스 활용으로 해당 기술을 채택하였으며 프롬프트 엔지니어링을 통해 원본과 유사하지만 독창성을 유지할 수 있도록 하였다. 또한 문제의 핵심 요소를 분석하여 분석된 문제 유형에 맞는 문제를 지속적으로 생성할 수 있도록 유도하였다.

이러한 방식을 통해 높은 유사성의 다양성을 확보한 문제를 신속하게 생성할 수 있는 문제집 생성 서비스를 개발하였다. 본 시스템의 구현을 통해 OCR 기술과 생성형 AI의 결합이 맞춤형 교육 자료 생성에 매우 효과적임을 입증하였으며, 향후 개인화 학습 플랫폼 개발에 중요한 기반을 마련하였다.

## IV. 실험 및 평가

### 4.1 실험

GPT-4o-mini 모델을 활용하여 문제 생성 모듈의 실성능을 비교하기 위해, 간단한 문제지를 만들어 실험을 진행하였다. 해당 실험은 총 3개의 분야(생명공학(Biotechnology), 경영학(Business administration), 컴퓨터 과학(Computer science))에 대한 자료로 진행 하였고, 문제 생성 시 15가지 문제를 출제하도록 설계 하였다. 각 분야마다 문제 생성과 문제 재생성 2번, 총 3번을 진행하여 45가지 문제 생성으로 측정을 진행 하였다.

### 4.2 평가지표

평가지표로는 총 3가지의 평가 지표를 사용하였다. 각각 정확도(Accuracy),중복률(Duplication rate), 명확성(Answerability)이다.

중복률은 문제 재생성 시 동일한 문제가 출제되었는지, 정확도는 문제에 대한 정답이 일치한지, 명확성은 정답이 답인 이유가 명확한지를 평가하는 항목이다.

### 4.3 실험 결과

본 실험은 경영학, 컴퓨터 과학, 생명공학 총 3개의 분야에 대해 각 45문제의 실험 결과를 그림 3과 같은 형태로 정리하였다.

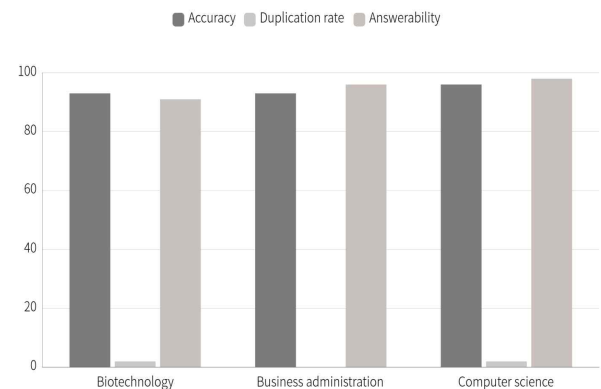


그림 3 . 분야별 정확도,중복률,명확성 비교  
Fig. 3. Comparison of accuracy, duplication rate, and clarity by field

생명공학의 경우, 45문제 중 3문제의 틀린 답이 발생하여 약 93%의 정확도를 보였으며, 1문제의 중복이 발생하여 약 2%의 중복률을 보였다. 또한, 명확성은 4문제의 잘못된 해설이 발생하여 약 91%를 나타냈다.

경영학의 경우, 정확도, 중복률, 명확성의 순서로 각각 약 93%, 0%, 96%의 결과가 나타났다.

컴퓨터 과학에 대해서는 약 96%, 2%, 98%로 나타났다. 각 분야마다 정확도, 중복률, 명확성에서 유의미한 차이를 보이지 않아, 더 많은 분야에서 사용 될 수 있을 것으로 나타난다.

## V. 결 론

본 논문에서는 OCR 기술과 생성형 AI를 결합한 혁신적인 맞춤형 문제 생성 시스템을 개발하였다. 이 시스템은 사용자가 제공하는 문제 유형을 정확히 인식하고, 이를 바탕으로 유사하면서도 새로운 문제를 자동으로 생성하는 기능을 구현하였다. 기존 커스텀 데이터베이스와 LLM을 기반으로 문제를 생성하는 방식에서 벗어나 OCR기술을 접목하여 사용자가 간편하게 문제 유형을 입력할 수 있다. 빠른 문제 재생성 등 사용자 편의성을 중심으로 생성형 AI와 교육 분야를 접목하여 새로운 패러다임을 제시하였다. 향후 연구 방향으로 다양한 난이도 지원과 개인별 최적화된 학습 경로를 제시하여 개인화된 학습 서비스를 통해 학생들의 학습 교화를 더욱 향상시킬 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1] 허동석, 김기태, 송형우, and 서봉원, '프롬프트 개발을 통한 수능 국어 맞춤형 문제 생성 시스템 제안,' in 한국HCI학회 학술대회, 2024, pp. 183-189.
- [2] 김은정. "웹 기반 기사시험 학습 시스템에서의 문제 출제 알고리즘." 한국산업정보학회 논문지 제 9.3 pp.11-18. 2004.
- [3] OpenAI. GPT-4 Technical Report. arXiv:2303.08774. November 10, 2023.
- [4] . Novikova, O. Dušek, A. Cercas Curry, and V. Rieser, 'Why We Need New Evaluation Metrics for NLG,' in Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark: Association for Computational Linguistics, 2017, pp. 2241-2252. doi: 10.18653/v1/D17-1238.
- [5] 박지원, 이재욱, 최제범, 홍서운, 홍성준, and 정은성, '대규모 언어 모델을 이용한 대학수학능력 평가 영어 문제 생성 모델 구현 및 평가,' in 대한전자공학회 하계학술대회, 2024, pp.2580-2584.
- [6] Clavié, Benjamin, et al. "Large Language Models

- in the Workplace: A Case Study on 프롬프트 엔지니어링 for Job Type Classification." International Conference on Applications of Natural Language to Information Systems. Cham: Springer Nature Switzerland, pp 3-17. 2023.
- [7] Khodeir, N. A., Elazhary, H., Wanas, N. Generating story problems via controlled parameters in a web-based intelligent tutoring system. The International Journal of Information and Learning Technology, 35(3), pp. 199-216. 2018.
  - [8] G. Kurdi, J. Leo, B. Parsia, U. Sattler, and S. Al-Emari, 'A Systematic Review of Automatic Question Generation for Educational Purposes,' Int J Artif Intell Educ, vol. 30, no. 1, pp. 121-204, Mar. 2020, doi: 10.1007/s40593-019-00186-y.



# Prophet 모델을 이용한 서울시 일일 오존 농도 예측

이의진\*, 고흥석\*, 이민수\*, 김재현\*\*

## Prediction of Daily Ozone Concentration in Seoul Using the Prophet Model

Lee Eui Jin\*, Ko Hyeong Seok\*, Lee Min Soo\*, and Kim Jae Hyun\*\*

### 요 약

대기 중 오존 농도가 높아짐에 따라 서울과 같은 대도시에서는 오존 농도 예측이 환경 관리와 시민 건강 보호를 위해 필수적인 과제가 되고 있다. 특히 자치구별 오존 농도는 계절적 변동과 추세를 반영한 예측이 필요하다. 본 연구에서는 서울시 25개 자치구의 오존 농도 데이터를 활용하여 시계열 예측 모델인 Prophet 모델을 적용해 예측력을 높인 모델을 구축하고자 한다. 이를 통해 지역별 특성에 맞는 대기 관리 정책 수립과 시민 건강 증진에 기여할 것으로 기대한다.

### Abstract

As urban ozone levels continue to rise, predicting ozone concentration has become essential for environmental management and protecting public health in large cities like Seoul. In particular, it is crucial to account for seasonal fluctuations and trends when forecasting ozone levels by district. This study aims to build an enhanced predictive model by applying the Prophet model time-series model to ozone concentration data from 25 districts in Seoul. The results are expected to support the development of localized air quality management policies tailored to each area's unique characteristics and contribute to the promotion public health.

### Key words

Prophet model, Predicting ozone concentration time-series data, predictive model

## 1. 서 론

대기 중 오존(O<sub>3</sub>)은 성층권에 있을 때 자외선을 차단하여 생태계를 보호하는 역할을 하지만, 대류권에서는 유해한 오염 물질로 작용한다. 특히 도시 지역에서는 차량 배출가스와 산업 활동으로 인해 오

존 농도가 높아질 가능성이 크며, 이는 호흡기 질환과 같은 건강 문제를 유발할 수 있다. 심할 경우 폐 기능 저하를 가져오고, 천식 및 만성 폐쇄성 폐 질환 등 호흡기 질환이 있는 사람에서 심한 폐 기능 저하, 기도과민성 증가, 기도 염증 유발 등의 문제를 일으킨다. 따라서 서울시와 같은 대도시에서는

\* 서경대학교, lee9191lee@naver.com, kkhs4988@gmail.com, minsujang22@naver.com

\*\* 서경대학교 교수, statsr@skuniv.ac.kr(교신저자)

대기 오염 관리가 중요한 과제로 화두가 되고 있으며, 오존 농도는 그 주요 지표 중 하나이다.

이러한 점에서 서울의 자치구별 오존 농도를 효과적으로 예측하는 것은 환경 관리 및 시민 건강 보호 측면에서 중요한 의미를 가지며, 예측 결과는 오존 농도를 줄이기 위한 정책 수립과 실행에 중요한 자료로 활용될 수 있다.

본 연구에서는 Prophet 모델을 이용해 서울시 25개 자치구의 오존 농도를 예측하고자 한다. Prophet 모델은 시계열 데이터에 강점을 지닌 예측 모델로, 계절성과 추세성을 동시에 고려할 수 있다. 이를 통해 보다 정확하고 세밀한 오존 농도 예측하여 각 지역의 특성에 맞는 대기 관리 정책을 마련하는 데 기여하고자 한다.

## II. 본문

### 2.1 선행연구

Prophet 모델은 간편한 모델링과 연산 효율성으로 인해 시계열 데이터를 활용한 미래 예측 분야에서 많이 사용되고 있으며, 특히 국외뿐만 아니라 국내에서도 다양한 예측 연구에 채택되고 있다.

Prophet 모델을 이용한 일일 지하철 이용객 인원 예측 연구(고형석, 이동규, 이슬기, 이준석, 하정원, 김재현; 2023)는 서울 열린 데이터 광장에서 제공하는 2015년부터 2023년 4월까지의 수도권 일별 지하철 이용 데이터에 기반해 이루어졌다. 이 연구는 Prophet 모델의 성능을 평가하기 위해 LSTM 모델과 비교 분석을 수행했으며, MAPE 기준으로 Prophet 모델이 LSTM 모델(1.76)에 비해 향상된 성능(1.56)을 보였다. Prophet 모델에 주간, 월간, 공휴일 등의 요소를 반영함으로써 예측 정확도가 더욱 향상되었다. 이를 통해 본 연구는 Prophet 모델이 시계열 데이터 예측의 정확성을 높이고 도시 환경 관리에 기여할 가능성을 제시하고 있다.

K-Means 클러스터링과 Prophet 모델을 이용한 서울특별시 미세먼지 농도 예측(김재현, 한기울; 2023)에서는 서울특별시를 주거지역, 상업지역, 혼합지역, 생산공장 밀집 지역 등 4개의 클러스터로 구분하여 각 클러스터별 특성에 따른 미세먼지 농도 예측을

수행했다. 이 연구는 예측 모델의 성능을 평가하기 위해 2015년부터 2021년까지의 데이터를 학습에 사용하고, 2022년 데이터를 검증용으로 사용했다. Prophet 모델은 RMSPE의 값이 0.1을 나타냈으며, 이는 예측 정확도 향상에 기여한 것으로 나타났다. 본 연구에서는 위 연구에서 제시된 클러스터 분류 방식을 활용하여, 각 클러스터를 대표할 수 있는 자치구 한 곳씩을 선정, 이들 지역의 오존 농도 변화를 분석함으로써, 서울시 전체의 대기질 변화 추이를 종합적으로 이해하고자 한다.

### 2.2 데이터 수집 및 전처리

분석에 사용된 오존 농도 데이터는 Air Korea에서 제공하는 2015년 1월부터 2024년 4월까지의 1시간 간격으로 측정된 서울시의 25개 구에 대한 데이터를 활용하였다. 데이터 전처리 과정에서 일별 평균 오존 수치를 구하여 일일 데이터로 변환, 결측치는 선형 보간법을 통해 처리하여 모델의 예측 정확도를 향상하였다.

본 연구는 서울시의 25개 자치구 중에서 선행 연구를 통해 각 클러스터를 대표할 수 있는 종로구, 은평구, 구로구, 송파구를 대상으로, 모델의 예측력과 예측 결과의 정확성을 검증하였다.

### 2.3 Prophet 모델

Prophet 모델은 시계열 데이터를 입력 값으로 받아들이며 미래를 예측하도록 훈련시키는, Facebook에서 공개한 시계열 예측 라이브러리이다. 계절 효과가 강하거나 과거 데이터에 여러 계절성이 존재하는 시계열 데이터 예측에서 좋은 성능을 보이는 것으로 알려졌다. Prophet 모델에서는 직관적인 하이퍼파라미터를 사용하여 모형을 쉽게 조정할 수 있고, 계절성 조절과 추세 변화 값을 통해 더욱 정확한 예측 결과를 제공한다. 또한 시계열의 시간 종속적인 특성을 고려하는 기존의 시계열 분석 방법론들과는 달리 곡선 적합(curve-fitting)으로 모형을 적합시킨다. 곡선 적합을 사용하여 복잡한 패턴을 더 잘 파악하고 변동성이 큰 데이터에 대해 정확한 분석 결과를 제공할 수 있다.

Prophet 모델의 기본 구조는 추세(trend), 계절성(seasonality), 이벤트 효과(holiday)의 합으로 다음과 같이 구성되어 있다.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (1)$$

수식(1)에서  $g(t)$ 는 비주기적인 변동인 trend를 나타내는 추세 함수,  $s(t)$ 는 주간 또는 연간 계절성과 같은 주기적 변화를 반영하는 함수,  $h(t)$ 는 불규칙적으로 발생하는 휴일 효과 함수이고  $\epsilon_t$ 는 모델에 표현되지 않는 모든 특이점을 의미한다. 이러한 효과들을 적극 반영하여 예측 성능을 개선하였다.

## 2.4 연구 결과

[그림 1]은 Prophet 모델을 사용하여 시계 순서대로 종로구, 은평구, 송파구, 구로구의 오존 농도 데이터를 분석한 결과를 나타낸다. 자치구의 각 그래프는 Prophet 모델 하이퍼파라미터인 전체 추세, 주간 변동성, 연간 변동성을 나타낸다.

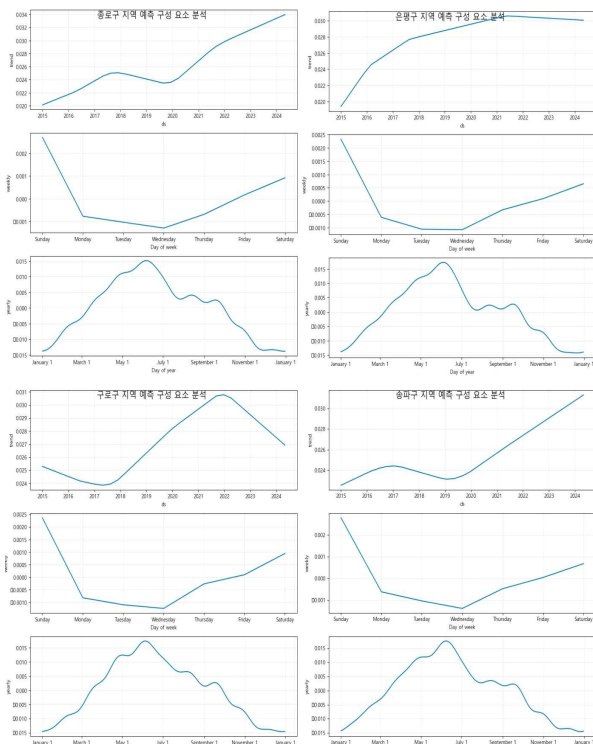


그림 1. Prophet 모델의 구성요소 그래프  
Fig. 1. Prophet Model Component Graphs

먼저 추세 그래프에서 종로구와 송파구는 2020년 이후 오존 농도가 꾸준히 상승하는 경향을 보였고, 구로구는 2017년 이후로 상승하고, 은평구는 계속 상승하는 모습을 보였다. 이는 각 자치구별로 오존 농도에 영향을 미치는 다양한 요인들을 이해하고 분석하는데 중요한 지표로 활용할 수 있다. 주간 변동성 그래프는 공통적으로 수요일에 오존 농도가 가장 낮았으며, 주말로 갈수록 오존 농도가 증가하는 것으로 나타났으며, 연간 변동성 그래프에서도 공통적으로 5월~6월에 최고점에 도달하는 것을 확인할 수 있었다.

성능 개선을 위해 Prophet 모델 파라미터 중 seasonality\_scale 값을 따로 설정하여 분석을 진행한 결과 scale 값은 기본값인 10이 성능이 좋은 것으로 나타나 기본값으로 설정하여 분석을 진행하였다.

[표 1]은 이러한 효과를 반영하여 예측 모델을 생성하고 검증해 본 결과이다. 각 4개의 자치구의 RMSE 지표와 MAPE 지표를 나타낸 결과 은평구의 RMSE는 0.00666, MAPE는 16.4975로 상대적으로 우수한 예측 성능을 보였다.

표 1. 지역에 따른 성능지표

Table 1. Performance Metrics by Region

지역	RMSE	MAPE
종로구	0.00734	19.9976
은평구	0.00666	16.4975
구로구	0.00730	19.9048
송파구	0.00758	24.3807

[그림 2]는 Prophet 모델을 관측값과 예측값을 그래프로 나타낸 것이다.

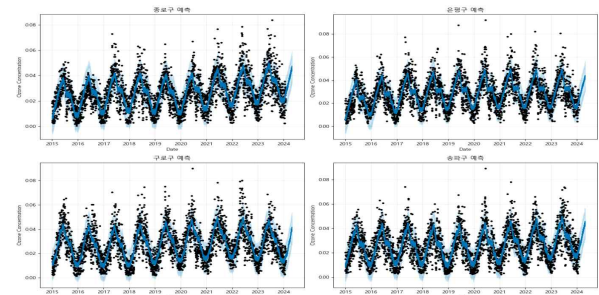


그림 2. 오존 농도 예측 결과  
(종로구, 은평구, 송파구, 구로구)

Fig. 2. Ozone Concentration Prediction Results  
(Jongno-gu, Eunpyeong-gu, Songpa-gu, Guro-gu)

### III. 결 론

본 연구는 2015년 1월부터 2024년 4월까지의 서울시 시간별 오존 농도 측정 데이터를 바탕으로 일일 오존 농도를 예측하였다. Prophet 모델을 이용하여 계절성, 휴일 효과 등의 특성을 통해 학습시키고 2024년 데이터로 검증을 진행하였는데 [표 1]에서 나온 것 처럼 평균적으로 RMSE는 0.008 미만, MAPE는 20 미만을 나타내고 있다. 본 연구의 결과를 바탕으로 실시간 대응 체계를 강화하고, 오존 농도가 높을 것으로 예상되는 기간에 대해 사전에 대처가 가능할 것으로 보인다. 또한, 지역별 오존 농도 패턴 파악으로 정책 수립에 도움을 줄 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] 에어 코리아, <https://www.airkorea.or.kr/>
- [2] 김재현, 한기율, “클러스터링과 Prophet 모형을 이용한 서울특별시 미세먼지 농도 예측”, 융복합지식학회논문지 제 11권 3, pp. 73-82, Sep 2023.
- [3] 고희석, 이동규, 이슬기, 이준석, 하정원, “Prophet 모델을 이용한 일일 지하철 이용자 인원 예측”, 한국정보기술학회 2023년도 하계종합 학술대회 및 대학생논문경진대회, pp. 427-431, Jun. 2023
- [4] 김준석, 강재환, 김성희, 윤주상, “Prophet 을 사용한 일변량의 시계열 예측”, 한국정보통신학회 2020년도 춘계종합학술대회 논문집 제24권 제1호, pp 329-331, Jul. 2023

# EasyOCR 성능 향상을 위한 최적의 Optimizer와 Scheduler 연구

김소현\*, 박인혁\*, 김민석\*, 정재진\*\*

## A Study on the Optimizer for EasyOCR Performance Improvement

SoHyun Kim \*, InHyuk Park\*, MinSeok Kim\* and JaeJin Jeong\*\*

### 요약

현재 다양한 OCR(광학문자인식)이 존재한다. 기존의 OCR은 특정 글자에서는 인식의 정확성이 떨어지는 경향이 있다. 이를 해결하기 위해, 본 논문은 EasyOCR을 기반으로 하여 VGG와 ResNet 각각의 모델에서 Adam과 Adadelta를 비교하였고, Adam에서의 Scheduler의 사용 전후의 정확도를 비교하며 Scheduler가 성능향상에 필수적임을 제안한다. 평가 지표로는 정확도, 신뢰도, CER을 사용하였으며, kaggle에서 노이즈가 섞인 데이터셋을 각 모델에 학습시켜 성능을 비교하였다. 실험결과 ResNet모델을 Adam Optimizer와 Scheduler를 사용하여 학습하였을 때 90%의 정확도가 개선되는 것을 확인할 수 있었다. 향후 다양한 Optimizer와 Scheduler를 조합한다면 더 좋은 성능을 낼 수 있을 것이라 기대한다.

### Abstract

Currently, various OCR (optical character recognition) exist. Existing OCR tend to be less accurate in certain characters. To solve this problem, this paper compares Adam and Adelta in each model of VGG and ResNet based on EasyOCR, compares the accuracy before and after the use of Scheduler in Adam, and suggests that Scheduler is essential for performance improvement. Accuracy, reliability, and CER were used as evaluation indicators, and performance was compared by learning a dataset mixed with noise in kaggle to each model. As a result of the experiment, it was confirmed that the accuracy improved by 90% when the ResNet model was learned using Adam Optimizer and Scheduler. It is expected that better performance will be achieved if various optimizers and Scheduler are combined in the future.

### Keywords

EasyOCR, English/Digit Recognition, Optimizer, Scheduler, Model

### I. 서론

현대 사회에서 다양한 필기체와 복잡한 인쇄체로

구성된 문서들이 여전히 수작업으로 처리되며 많은 시간과 비용이 소요된다. 이러한 문제를 해결하기 위해 OCR(광학 문자 인식) 기술을 활용한 문서 자동

\* 국립금오공과대학교 전자및전파공학과 ekfvodgl02@naver.com, cheol0519@naver.com, kimm3274@naver.com

\*\* 국립금오공과대학교 전자및전파공학과 jeong@kumoh.ac.kr

인식 및 처리 시스템의 필요성이 커지고 있다.

대표적인 OCR(Optical Character Recognition) 으로는 Google의 Tesseract OCR과 JaidedAI의 EasyOCR[7]이 있으며, 각각 CPU와 GPU 기반으로 작동한다. Tesseract는 LSTM(Long Short-Term Memory) 구조로 설계되어 특정 언어에서 높은 인식률을 보이며, 구조가 단순해 사용이 용이하다. 반면, EasyOCR은 필기체와 왜곡된 문자 및 숫자 인식에서 우수한 성능을 보인다.[1]

학습한 모델의 성능 향상을 위해 Learning rate를 동적으로 조정하는 Scheduler를 사용하였다. Learning rate는 Optimizer가 loss의 최소값에 도달하는 단계의 크기를 제어한다. 기존의 Optimizer만 사용 시 Learning rate의 크고 작음에 따라 가중치가 변화하게 되고, 이는 가중치가 overflow(과적합)하거나 학습을 정체시켜 모델의 성능과 학습 시간에 영향을 미친다. 위와 같은 이유로, 적절한 Learning rate를 찾는 것이 중요한데, Scheduler를 사용하게 되면 학습 과정에서 학습률을 동적으로 조정하게 함으로서 해당 문제를 해결하고 성능 향상에 도움을 줄 수 있다.[11]

본 논문에선 CRNN(Convolutional Recurrent Neural Network) 구조를 기반으로 하는 EasyOCR을 VGG와 ResNet 모델에 대해 학습시키고 다양한 Optimizer와 Scheduler를 적용 여부에 따라 모델들의 성능을 높이고자 한다. 경사 하강법의 학습이 오래 걸리는 단점을 개선한 Adam과 Adadelata를 사용하였다.[11] Adam은 보폭과 방향 모두 조절하여 최적해는 찾는 방식으로, Adadelata는 보폭만을 조절하여 학습 경로의 가중치를 조정하는 방식으로 학습률을 조정한다. 실험결과, Adam이 더 좋은 성능을 보였으며, 이를 Scheduler를 통하여 성능을 높일 수 있었다.[8][9]

## II. EasyOCR 학습 모델 별 성능 비교

### 2.1 학습 데이터 선정 및 전처리

이 논문에서 CRNN(CNN + RNN) 기반의 EasyOCR의 학습을 위해 구글 캐글 (Kaggle)에서 제

공하는 노이즈가 포함된 데이터셋을 사용하였다. 표 1에 언급한 바와 같이 각 데이터셋 4가지에서 각각 1000개씩 랜덤 추출하여 9:1 비율로 Training 데이터셋과 Test 데이터셋으로 분류하여 학습을 진행하였다. [2-5] Training 데이터셋과 Test 데이터셋을 라벨링하여 확장자를 문자로부터 분리하고, lmbd 파일 형식으로 변환하여 학습을 진행하였다.

표 1. 학습 데이터 세트

Table 1. learning dataset

데이터개수	training	test
1000	900	100
1000	900	100

### 2.2 Optimizer 비교

본 논문에서는 설명의 용이성을 위해 None-VGG-BiLSTM-CTC를 VGG로 TPS-ResNet-BiLSTM-CTC를 ResNet으로 약칭하겠다.

EasyOCR의 학습 결과는 표 2에 보인 바와 같이 VGG모델에서는 Adam에서 79%로 Adadelata에 비해 7% 높게 나타났으며, ResNet모델에서는 Adadelata가 8.25% 높게 나타났다. 그러한 결과로 VGG모델에서는 Adam이, ResNet모델에서는 Adadelata가 더 좋은 성능을 낸다는 것을 확인하였다.

ResNet에서 Adadelata보다 Adam이 더 성능이 좋을 것이라 예상했지만, 그 반대의 결과가 나오게 되었는데 이는 Adadelata인 큰 초기치 값이 ResNet 모델의 residual connect 영향으로 큰 초기치 값이 장점으로 작용하여 생긴 문제로 판단된다.

표 2. 모델 최대 정확도

Table 2. model best accuracy

model-optimizer	MODEL BEST ACCURACY	
	None-VGG-BiLSTM-CTC	TPS-ResNet-BiLSTM-CTC
adadelata	72	87.5
adam	79	79.25

### 2.3 Scheduler 적용

Adam의 성능을 더 올리고자, Scheduler를 사용하

였다. Learning rate(lr)를 2500, 4000에서 gamma=0.5로 설정하여 lr에 0.5가 곱해져, 지정 구간에서 lr을 감소시켜 학습이 진행되도록 하였다.

표 3에서의 결과와 같이 Adam모델에 Scheduler를 적용하였을 때 VGG에서는 5% 높은 79%의 성능을 보였고 ResNet에서는 10.75% 높은 90%의 성능을 보인다. Scheduler를 사용하기 전에는 VGG에서만 좋은 성능을 보였으나 Scheduler를 통하여 러닝레이트를 감소시켜 학습한 결과 ResNet모델에서도 Adam이 좋은 성능을 낸다는 것을 확인하였다.

표 3. 스케줄러 적용 후 최고 정확도

Table 3. Best Accuracy After Scheduler Application

model-optimizer	MODEL BEST ACCURACY	
	None-VGG-BiLSTM-CTC	TPS-ResNet-BiLSTM-CTC
adam	79	79.25
adam+scheduler	84	90

학습된 모델을 Test 데이터에 적용하여 표 3과 같이 모델의 신뢰도를 비교한 결과 Scheduler를 사용한 모델이 VGG에서는 0.82, ResNet에서는 0.80으로 Adam만 사용했을 때보다 높은 신뢰도를 얻을 수 있었다.

표 4. 신뢰도

Table 4. confidence

model-optimizer	MODEL CONFIDENCE	
	None-VGG-BiLSTM-CTC	TPS-ResNet-BiLSTM-CTC
adam	0.78	0.72
adam+scheduler	0.82	0.80

CER(Character Error Rate)을 통해 모델의 오차율 비교를 통하여 각 모델의 성능을 비교하였다.[6]

$$CER = \frac{S + D + I}{N}$$

CER은 문자열들 사이에 오차율을 계산하는 지표로써 S = 잘못 대체된 음절 수, D = 잘못 삭제된 음절 수, I = 잘못 추가된 음절 수, N = 정답 음절 수를 뜻한다. 이는 0에 가까울수록 더 좋은 성능을

가짐을 의미한다. 각 모델에 따른 CER은 표 4와 같으며, 스케줄러를 사용했을 때 VGG는 0.154, ResNet은 0.143으로 더 낮은 CER을 가지는 것으로 보아 스케줄러가 성능 향상에 도움이 된다는 것을 확인할 수 있었다.

표 5. 오차율

Table 5. CER

MODEL	CER
None-VGG	0.183
None-VGG + Scheduler	0.179
TPS-ResNet	0.154
TPS-ResNet + Scheduler	0.143

## 2.4 연구 결과

기존 OCR은 특정 글자에서의 오류로 인해 실무에서 사용하기에는 어려움이 있어, 성능을 향상시키고자 연구를 진행하였다. 본 논문에서는 ResNet + Adam + Scheduler 조합을 통해 학습 성능을 최적화하는 방안을 제안한다. ResNet은 복잡한 데이터셋에서도 효율적 학습을 지원하며, Adam은 데이터 및 모델 파라미터가 많이 필요한 문제에 적합하다. 또한 추가적으로 Scheduler를 사용해서 lr을 적절하게 감소시키는 방법으로 성능을 Improve하여 그림 1에 보이는 바와 같이 VGG 모델에 스케줄러를 적용한 결과인 84%보다 높은 90%의 정확성을 얻을 수 있었다.

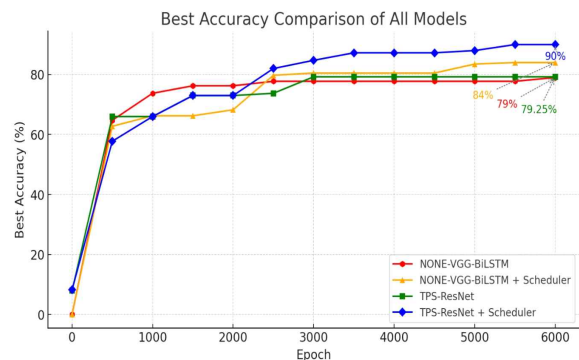


그림 1. 모델 학습후 정확도

Fig 1. Accuracy after model training

### III. 결 론

본 논문에선 VGG와 ResNet 모델에서 Adam과 Adadelta라는 Optimizer를 사용하여 결과를 비교하고, Scheduler의 유무에 따라 성능의 개선 여부를 보여줌으로써 Optimizer 선택 및 Scheduler의 필요성을 보여주었다. 향후 추가 연구에서, 다양한 Optimizer와 최적의 Scheduler를 조합한다면 더 좋은 성능을 낼 수 있을 것이라 기대한다.

19(2), 41-48.

[11]L. Wen, L. Gao, X. Li and B. Zeng, "Convolutional Neural Network With Automatic Learning Rate Scheduler for Fault Classification," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-12, 2021,

### 참 고 문 헌

- [1] OCR Engine Comparison Tesseract vs. EasyOCR(2020).<https://medium.com/mllearning-ai/tesseract-vs-keras-ocr-vseasyocr-ec8500b9455b>
- [2] Kaggle, CAPTCHA Images, (2018). <https://www.kaggle.com/datasets/fournierp/captcha-version-2-images>
- [3] Kaggle, 66k Captchas Dataset, (2023). <https://www.kaggle.com/datasets/jassoncarvalho/comprasnet-captchas>
- [4] Kaggle, CAPTCHA Dataset, (2020). <https://www.kaggle.com/datasets/parsasam/captcha-dataset/>
- [5] Kaggle, Large Captcha Dataset, (2021). <https://www.kaggle.com/datasets/akashguna/large-captcha-dataset>
- [6] <https://github.com/hyeonsangjeon/computing-Korean-STT-error-rates>
- [7] EasyOCR (2024) Repository [Source Code]. <https://github.com/JaidedAI/EasyOCR>
- [8] Santos-Bustos Daniel Fernando, Nguyen Binh Minh, Espitia Helbert Eduardo. "Towards automated eye cancer classification via VGG and ResNet networks using transfer learning" Engineering science and technology, an international journal
- [9] 주기훈, 박치현, 임현승. (2020). 기계학습 옵티마이저 성능 평가. 전기전자학회논문지, 24(3), 766-776.
- [10] Kim Chan, & Yun Young-Sun (2023). Implementation of an Automated Grading System for Formalized Answer Sheet Using EasyOCR. Journal of Software Assessment and Valuation,



# 산업 현장의 근골격계 질환을 실시간으로 예방할 수 있는 인공지능 기반 시스템

이강산\*, 여인규\*, 권현우\*, 장광익\*, 정해성\*, 박주혁\*, 김한서\*, 채훈희\*, 이육현\*\*

## AI-Based System for Real-Time Prevention of Musculoskeletal Disorders in Industrial Workplaces

Kangsan Lee\*, Ingyu Yeo\*, Hyunwoo Kwon\*, Kwangik Jang\*, Haeseong Jeong\*, Juhyeok Park\*, Hanseo Kim\*, Hunhee Chae\*, and Ukhyun Lee\*\*

### 요 약

산업 현장에서 근골격계 질환은 반복적이거나 비정상적인 작업 자세로 인해 빈번하게 발생하며, 이는 생산성 저하와 장기적인 건강 악화를 초래한다. 본 연구는 현장직 근로자의 근골격계 위해 요인을 예방하기 위한 AI 프로그램 개발에 중점을 두고 있다. YOLOv8-pose 모델을 사용하여 근로자의 작업 자세를 실시간으로 추적하고, EfficientNet-B7 모델을 통해 관절 움직임을 분석하였다. 이러한 분석 결과는 근골격계 위험 요인 평가 기준인 OWAS 및 REBA 평가를 통해 근로자의 근골격계 부담 점수로 환산된다. CCTV에서 실시간 저장된 영상 데이터를 웹 서버에서 받아서 AI서버에 전달하고, AI서버에서 분석된 결과를 웹 페이지와 모바일 알림으로 관계자에게 제공한다. 본 개발의 결과는 현장직 근로자의 작업 환경 개선과 근골격계 질환 예방에 기여할 수 있다.

### Abstract

Musculoskeletal disorders in the industrial workplace are frequently caused by repetitive or abnormal work postures, which leads to reduced productivity and long-term health deterioration. This research focuses on the development of an AI programme to prevent musculoskeletal risk factors in field workers. The YOLOv8-pose model was used to track workers' working postures in real time, and the EfficientNet-B7 model was used to analyse joint movements. The results of these analyses are converted into a worker's musculoskeletal burden score through OWAS and REBA assessments, which are criteria for evaluating musculoskeletal risk factors. The video data stored in real time from CCTV is received by the web server and delivered to the AI server, and the results analysed by the AI server are provided to the relevant parties through web pages and mobile notifications. The results of this development can contribute to improving the working environment of field workers and preventing musculoskeletal diseases.

### Key words

Musculoskeletal disorders, AI application, YOLOv8-pose model, EfficientNet-B7 model, REBA, OWAS

---

\* 신한대학교 소프트웨어융합학과 학부생, kslee0527@naver.com, ci0430@naver.com, harryer@naver.com, rhkdemforhs2@naver.com, canpoo3@naver.com, wngur3033@naver.com, srszgg@gmail.com, gnsgml1218@gmail.com

\*\* 신한대학교 소프트웨어융합학과 교수 (교신저자) uhlee@shinhan.ac.kr

※ 감사의 글 : "본 연구는 2023년도 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음" (2023-0-00089)

## I. 서 론

산업 현장에서 근골격계질환은 반복적이거나 비정상적인 작업 자세로 인해 빈번하게 발생하며, 이러한 인체공학적 고위험 작업은 생산성 저하와 장기적인 건강 악화를 초래한다. 산업계 노동자들의 질환 증상은 팔꿈치, 어깨, 손 또는 손목, 허리 부분 등에서 높은 것으로 보고되었다[1]. 근골격계 질환을 효과적으로 예방하기 위해서는 작업의 근골격계 부담을 정확히 평가하고 그 결과에 근거하여 작업을 개선하는 것이 필요하다. 이를 해결하기 위해, 국내 대부분의 사업장은 근로자의 자세를 분석하는 대표적인 인간공학적 정밀 도구인 OWAS (Ovako Working posture Analysing System), RULA (Rapid Upper Limb Assessment), REBA (Rapid Entire Body Assessment)를 사용하여 근골격계 부담을 평가하고 있다[2].

그러나 근골격계 질환 유해 요인 조사는 근골격계 질환이 많이 발생하고 있는 제조업에서 실시율이 매우 저조하다[2]. 또한, 근골격계 질환 유해 요인 조사를 실시 하여도 근로자의 작업 자세를 인간공학 전문가와 인간공학 비전문가 간의 평가 결과 차이를 비교 분석하였을 때 정확도와 신뢰도 측면에 한계가 있는 것으로 알려졌다[3]. 따라서 위에서 언급한 인간공학 정밀 평가 도구의 정확도와 신뢰도의 향상 및 평가의 편차를 줄이기 위한 연구가 많이 진행되었다[3][4].

이러한 연구를 배경으로 본 연구에서는 근로자의 자세를 실시간으로 추적하고 관절 움직임을 분석하여 근골격계 부담을 인간공학 정밀 평가 도구 중 OWAS와 REBA로 AI 기술을 통해 평가하는 프로그램을 개발하고자 한다. 그리고 근로자의 자세 분석의 정확성을 높이고 실시간 피드백을 통해 근로자가 안전한 자세를 유지할 수 있도록 하는 시스템을 개발하고자 한다.

## II. 관련 연구

기존에는 인간공학적 정밀 도구인 OWAS, REBA, RULA을 사용하여 육안으로 근로자의 동작 분석에

집중해 왔으나, 평가를 판단하는 정확도와 신뢰도 측면에 한계가 있다[3]. 그리고 실시간으로 작업 자세를 평가하고 근골격계 부담을 정량화하는 연구는 아직 부족하다. 특히 AI 기술의 발달로 실시간 데이터 처리와 분석이 가능해지면서 작업 환경에서의 적용 가능성이 높아졌다[4]. 이러한 필요성과 환경에 의해 본 연구는 AI 기술을 근골격계 부담 평가에 적용하여 더 정확하고 신속한 분석 및 평가의 정량화를 목표로 하였다. 즉, 근골격계 부담을 실시간으로 분석하고 위험을 평가하여 예방 조치를 제안하는 AI 기술을 적용한 SW 개발을 목표로 한다.

## III. 서비스의 구성 및 설계

### 3.1 AI 모델 선정 과정

AIHub에서 제공하는 제조업 노동자 관련 데이터는 다양한 작업 자세 분석에 유용한 모델들을 포함하고 있다[5]. 기존 AIHub에서는 작업 자세 추적을 위해 YOLOv5, HRNet 모델을, 관절 움직임 분석을 위해 EfficientNet 모델을 제안하였다[5][6]. 본 개발에서는 실시간 작업 자세 추적을 위해 YOLOv5이 아닌 YOLOv8버전의 사전학습모델인 YOLOv8-pose 모델을 선택하였으며 관절 움직임 분석을 위해서는 EfficientNet의 사전학습모델인 EfficientNet-B7 모델을 선정하였다. 이러한 모델 선정은 성능과 효율성을 모두 고려하여 비교 분석한 후 선정하였다.

YOLOv8-pose 모델은 실시간으로 근로자의 자세를 정확하게 추적할 수 있는 객체 탐지 기반 모델로, 신속한 반응이 요구되는 작업 환경에 적합하다[7][8].

EfficientNet-B7 모델은 구글 AI 연구팀이 개발한 모델로 복잡한 패턴의 관절 움직임을 분석하는 데 탁월한 성능을 제공한다. 복잡 스케일링 전략을 통해 적은 파라미터로도 높은 성능을 발휘하며, 대규모 비디오 데이터 분석에 적합하다[9].

### 3.2 서비스 구성

그림1은 근골격계 위해 요인 평가 시스템의 서비

스 구성이다. 크게 웹 서버, 파일 시스템, AI, DB로 구성되어 있다.

해당 서비스는 HTTP 프로토콜을 사용하여 사용자와 통신하며, Flask 웹 서버를 통해 AI 기반으로 근로자의 작업 자세를 실시간으로 수집 및 분석하고, 웹 인터페이스를 통해 인간공학적 정밀도구인 OWAS, REBA 점수를 제공하는 웹서비스 S/W로 설계되었다. 이 시스템은 YOLOv8-pose 모델과 EfficientNet-B7 모델을 사용하여 근로자의 자세를 추적하며, 결과 데이터를 평가 도구인 OWAS, REBA의 점수로 환산하여 근로자의 근골격계 부담을 평가 및 분석한다. 또한, MySQL 데이터베이스를 사용하여 관련된 데이터를 저장하고 관리한다.

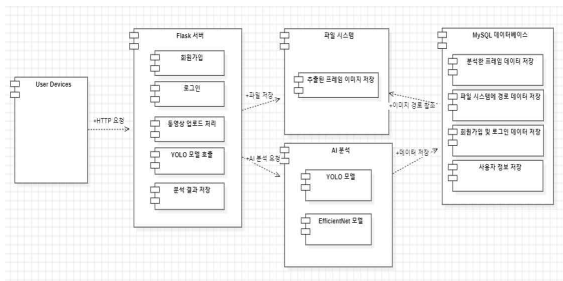


그림 1. 어플리케이션 디플로이먼트 다이어그램  
Fig. 1. Application deployment diagram

### 3.3 시스템 구성 요소

그림 2는 어플리케이션의 데이터 흐름을 나타낸다. 크게 데이터 가공, AI 추론 및 분석, 데이터 활용으로 이루어져 있다.

서버는 Flask 기반의 웹 서버로, 실시간 영상을 지정된 디렉터리 경로에 저장한다. 디렉터리에 있는 영상을 AI가 프레임별로 자르고 프레임별 이미지 안에 있는 근로자의 관절을 추론한다. 추론 후 OWAS, REBA 평가기법을 사용해 근골격계 질환 유해 요인에 대한 점수를 매긴다.

데이터베이스 관리시스템은 MySQL을 사용하였다. 근로자 정보와 작업 데이터 결과를 저장 및 관리하고 데이터베이스를 통해 사용자 인증 및 세션을 관리한다.

파일 저장은 로컬 서버에서 실시간으로 저장된 영상과 AI 모델이 추론한 이미지를 파일 시스템을

이용해 로컬 파일로 저장한다. 그리고 이미지에 대한 정보와 OWAS, REBA에 대한 정보를 데이터베이스에 저장한다.

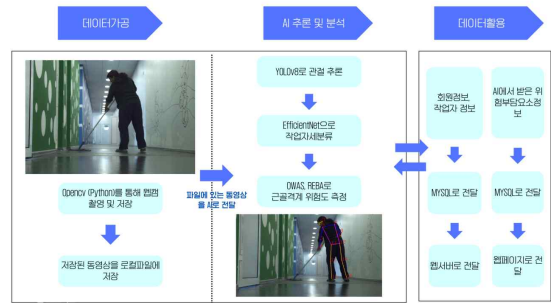


그림 2. 어플리케이션 데이터 흐름도  
Fig. 2. Application Data Flowchart

### 3.4 설계

웹 어플리케이션은 그림 3과 같이 구성되어 있다. 근로자의 상태를 모니터링하고 분석 결과를 확인할 수 있는 다양한 화면으로 구성된다. 주요 화면은 로그인 및 회원가입 화면, 메인 화면, 근골격계 위험 감지 화면, 관리자 화면으로 이루어져 있다.



그림 3. 어플리케이션 화면 구성도  
Fig. 3. Application Screen Configuration Chart

로그인 및 회원가입 화면에서 사용자는 웹서비스에 접속하여 데이터베이스에 저장된 사용자 정보를 통해 로그인 및 회원가입을 할 수 있다. 사용자는 로그인을 통해 자신의 작업 자세 분석 데이터를 조회할 수 있으며, 시스템에 저장된 과거 작업 데이터를 확인할 수 있다.

메인 화면에서는 작업 자세가 실시간으로 표시된다. YOLOv8-pose 모델과 EfficientNet-B7을 사용하여 수집된 자세 데이터를 바탕으로 근로자의 OWAS 및 REBA 점수가 실시간으로 계산되며, 위험한 자

세가 감지될 경우 근골격계 위험 감지 화면으로 분석 결과가 전달되고, 경고 메시지가 제공된다. 사용자는 현재 작업 자세 상태를 확인할 수 있다.

근골격계 위험 감지 화면에서 그림 4와 같이 실시간 작업 자세 분석 결과가 표시된다. 관리자는 자신이 수행한 작업 데이터를 MySQL 데이터베이스 및 로컬 디렉터리에서 불러와 조회할 수 있다. 각 작업의 OWAS 및 REBA 점수는 시간대별로 기록되며, 이를 통해 근로자의 자세 변화 및 근골격계 부담 요인으로 인한 점수 변동을 확인할 수 있다. 부담 분석 결과는 화면에 제공된다.



그림 4. AI 근골격계 분석 결과 사진  
Fig. 4. AI Musculoskeletal Analysis Photo

관리자 화면에서는 관리자가 MySQL 데이터베이스의 모든 근로자의 정보를 관리 할 수 있다. 그리고 작업 환경을 개선하는 데 필요한 데이터를 확인할 수 있다. 또한, 관리자 화면에서는 작업에 따라 OWAS와 REBA 처리 코드의 하중을 조정하는 등 설정을 조정하여 각 작업 환경에 맞춤형 피드백을 제공할 수 있다.

#### IV. 결론

본 연구는 AI 모델을 활용하여 근로자의 작업 자세를 실시간으로 분석하고, 이를 근골격계 부담 평가 도구를 활용하여 점수를 환산하고 위험 자세가 나오는 시점을 보여주는 프로그램을 개발하였다. YOLOv8-pose 모델과 EfficientNet-B7 모델을 통해 자세와 관절 움직임을 인식하고, OWAS 및 REBA 평가 기법을 통해 근골격계 부하값을 산출하였다. 이 시스템은 AI를 적용하여 유해 요인 조사의 객관

성을 향상시킬 수 있고, 소요시간을 절감할 수 있다. 그러므로 작업 현장에서 근골격계 질환을 예방하는 데 크게 기여할 수 있다.

#### V. 향후 연구

향후 연구에서는 더 많은 작업 환경에서 다양한 데이터 수집과 AIHub의 데이터를 기반으로 선정한 모델을 분석 후 튜닝하고 더 많은 데이터 학습을 하여 모델의 성능을 더욱 높일 예정이다. 즉, 95% 이상의 신뢰도를 가지는 판독 결과를 생성할 수 있도록 개발할 것이다. 또한, 실시간 피드백 기능을 강화하여 근로자가 작업 중 실시간으로 자세를 수정할 수 있도록 하는 알람 시스템과 모바일 알림 시스템을 개발할 계획이다. 나아가, 미세한 손목 동작 작업 환경 등 더욱 다양한 작업 현장에서 적용할 수 있는 AI 솔루션 시스템을 제공하기 위한 연구를 계속 진행할 예정이다.

#### 참고 문헌

- [1] M. Zare, J. Bodin, E. Cercier, R. Brunet and Y. Roquelaure, "Evaluation of ergonomic approach and musculoskeletal disorders in two different organizations in a truck assembly plant", *International Journal of Industrial Ergonomics*, Vol. 50, pp. 34-42, 2015.
- [2] J. H. Park, I. S. Lee, D. H. Kee, H. S. Jung and J. K. Park, "Survey on performance of the risk assessment of musculoskeletal disorders." *Journal of the Korean Society of Safety*, Vol. 26 No. 1, pp. 49-57, 2011.
- [3] W. H. Cheon and K. H. Jung, "Analysis of Accuracy and Reliability for OWAS, RULA, and REBA to Assess Risk Factors of Work-related Musculoskeletal Disorders", *J. Korea Saf. Manag. Sci*, Vol. 22 No. 2, pp. 31-38, June 2020.
- [4] J. O. Seo and S. H. Lee, "Automated postural ergonomic risk assessment using vision-based posture classification", *Automation in Construction*,

Vol. 128, pp. 103725, 2021.

- [5] <https://www.aihub.or.kr/aihubdata/data>. [Accessed: May. 23, 2024]
- [6] Y. J. Kang, T. K. Roh, K. H. Kim and S. C. Jeong, "Class Classification and Validation of a Musculoskeletal Risk Factor Dataset for Manufacturing Workers", *The Korea Journal of BigData*, Vol. 8, No. 1, pp. 45-59, 2023.
- [7] D. Maji, S. Nagori, M. Mathew and D. Poddar, "YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss.", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2637-2646, 14 April 2022.
- [8] H. D. Kim, S. H. Kee, M. K. Sohn, and J. K. Kim, "A Study on the face and body pose estimation based on YOLOv8", *Proceedings of Symposium of the Korean Institute of communications and Information Sciences*, pp. 640-641, November 2023.
- [9] Mingxing Tan and Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", *Computer Vision and Pattern Recognition*, 11 Sep 2020.

# GPS 조건을 활용한 PointNet 기반 CVAE 모델의 포인트 클라우드 학습

최서인\*, 조준영\*, 임성진\*, 오창석\*, 온병원\*\*, 정동원\*\*

## PointNet-based CVAE Model for Point Cloud Learning Using GPS Conditions

Seoin Choi\*, Joonyoung Jo\*, Seongjin Lim\*, ChangSeok Oh\*, ByungWon On\*\*,  
and Dongwon Jeong\*\*

### 요 약

본 연구는 3D 공간 데이터의 효과적인 처리를 위해, 포인트 클라우드 데이터의 학습과 재구성을 위한 PointNet 기반 조건부 변분 오토인코더(CVAE) 모델을 제안한다. 이 모델은 포인트 클라우드 데이터의 3차원 구조를 학습하고, GPS 데이터를 조건부 입력으로 활용하여 시공간적 맥락을 강화함으로써 재구성 성능을 향상시킨다. 연구에서는 KITTI 데이터셋의 도시 데이터를 사용해 k-means 클러스터링과 정규화를 포함한 데이터 전처리 과정을 거쳤다. 실험 결과, 제안된 CVAE 모델은 기존 VAE보다 높은 정확도의 포인트 클라우드 재구성을 달성하였으며, Chamfer Loss를 통해 원본 데이터와 재구성된 데이터 간의 형상 유사성을 보장했다.

### Abstract

This study proposes a PointNet-based Conditional Variational Autoencoder (CVAE) model for learning and reconstructing point cloud data to enable effective processing of 3D spatial data. The model learns the 3D structure of point cloud data and improves reconstruction performance by incorporating spatiotemporal context using GPS data as conditional input. The research involved using the city data from the KITTI dataset and included data preprocessing steps such as k-means clustering and normalization. Experimental results demonstrated that the proposed CVAE model achieved higher reconstruction accuracy compared to the conventional VAE, ensuring shape similarity between the original and reconstructed data through the use of Chamfer Loss.

### Key words

PointNet, Conditional Variational Autoencoder(CVAE), GPS, K-means Clustering

---

\* 국립군산대학교 소프트웨어학과, {seoinc46, joon0zo1022, 2001595, ocs2852}@kunsan.ac.kr

\*\* 국립군산대학교 소프트웨어학과, {bwon, djeong}@kunsan.ac.kr(교신저자)

※ 본 논문은 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065).

## I. 서론

최근 몇 년간 자율 주행 차량, 로봇 공학, 정밀 농업 등 다양한 공간 데이터 기반 응용 분야의 발전은 3D 공간 데이터를 정확하고 효율적으로 처리하는 것을 필수적인 요구사항으로 만들고 있다. 이러한 응용 분야는 실시간 환경을 인식하고, 정확하게 위치를 추적하며, 고도화된 데이터 분석을 통해 향상된 의사결정을 지원하도록 하고 있다. 특히 포인트 클라우드(point cloud) 데이터는 물체의 3차원 구조를 인식하고 복잡한 지형 정보를 수집하는 데 중요한 역할을 하고 있다[1]. 그러나 포인트 클라우드 데이터는 고차원의 복잡성을 지니고 있어, 이를 효과적으로 분석하고 해석하는 데에는 여러 제약이 따르고 있다[2].

GPS 데이터는 기기의 움직임과 위치 정보를 제공하며, 공간 인식의 정확성을 높이는 중요한 보조 요소로 작용하고 있다[3]. 다중 센서 데이터를 효과적으로 결합하고 활용하기 위해서는 고도화된 데이터 통합 기법이 필요하다. 최근 연구에서는 딥러닝 기반 모델, 특히 오토인코더(Autoencoder) 구조[4]가 복합 데이터를 처리하는 데 있어 주목받고 있다. 그 중에서도 조건부 변분 오토인코더(Conditional Variational AutoEncoder, CVAE)는 특정 조건을 고려하여 데이터를 보다 정확하게 재구성할 수 있어, 고차원 데이터의 특성을 학습하고 유용한 표현을 추출하는 데 뛰어난 성능을 보이고 있다.

본 연구에서는 포인트 클라우드 데이터를 입력으로 하고, GPS 데이터를 조건부 입력으로 활용하는 CVAE 모델을 제안하고자 한다. 이를 통해 포인트 클라우드의 시공간적 맥락을 강화하고, 데이터 간의 상관관계를 보다 명확히 학습하여 정교한 분석과 다양한 응용 가능성을 탐구하고자 한다.

## II. 관련연구

### 2.1 Point Cloud Data

포인트 클라우드(point cloud) 데이터[1]는 물체의 3차원 구조와 복잡한 지형 정보를 인식하고 수집하는 데 핵심적인 역할을 하고 있다. 포인트 클라우드는

주로 라이다(LiDAR) 센서나 3D 스캐너와 같은 센서를 통해 얻어지며, 각 점은 공간상의 좌표와 함께 때로는 색상 및 반사 강도와 같은 추가 정보를 포함할 수 있다. 이러한 데이터는 자율 주행 차량, 로봇 공학, 정밀 농업, 도시 계획, 증강 현실 등의 다양한 분야에서 활용된다. 그러나 포인트 클라우드는 고차원의 복잡성과 방대한 데이터 양으로 인해, 이를 효율적으로 처리하고 분석하는 데 여러 도전 과제를 수반한다. 특히, 노이즈 제거, 데이터의 희소성, 불규칙한 구조 등은 포인트 클라우드 분석의 주요 과제로 남아 있다.

### 2.2 변분 오토인코더(VAE)

변분 오토인코더(Variational AutoEncoder, VAE)[4]는 오토인코더 구조를 기반으로 하면서 확률적 모델링을 도입하여 데이터의 잠재 공간(latent space)을 학습하는 데 사용된다. VAE는 입력 데이터를 인코더를 통해 잠재 벡터로 변환하고, 디코더를 통해 이를 다시 재구성한다. 이 과정에서 VAE는 입력 데이터의 분포를 학습하며, 잠재 벡터 공간에서 데이터의 잠재 구조를 포착한다. VAE의 특징적인 점은 확률적 성질을 통해 잠재 공간의 연속성을 보장하고, 새로운 데이터 생성을 가능하게 한다는 것이다. 이 모델은 이미지 생성, 이상 탐지, 데이터 보강 등 다양한 분야에서 활용되고 있으며, 고차원의 데이터 처리에 있어 강력한 성능을 발휘한다.

### 2.3 조건부 변분 오토인코더(CVAE)

조건부 변분 오토인코더(Conditional Variational AutoEncoder, CVAE)[5]는 VAE의 확장된 형태로, 입력 데이터와 함께 조건부 정보를 사용하여 데이터의 재구성을 수행한다. CVAE는 인코더와 디코더 단계에서 조건부 정보를 입력하여, 조건에 따라 다르게 학습된 잠재 벡터를 생성할 수 있다. 이를 통해 특정 조건을 고려한 데이터의 재구성 및 생성을 가능하게 하며, 데이터의 의미적 표현을 더 잘 학습할 수 있다. CVAE는 다중 모달 데이터의 통합, 조건부 샘플링, 지도 학습 등 다양한 응용 분야에서 활용되고 있으며, 특히 고차원의 복합 데이터를 다루는 데 있어 탁월한 성능을 보인다.

### III. 제안방안

#### 3.1 데이터 전처리

연구를 진행하기 위해 KITTI 데이터셋[6]의 도시 데이터를 활용하였다. 전체 데이터를 다운로드한 후, 원본 포인트 클라우드와 GPS 데이터를 확보하고, 효율적인 데이터 처리를 위해 k-means 클러스터링을 수행하였다. 클러스터 개수 k는 50으로 설정하였으며, 이를 통해 8,307개의 포인트 클라우드 및 GPS 데이터 포인트를 생성하였다. 클러스터링된 데이터는 공간적 특성과 분포를 구조적으로 파악하는데 기여하였다. 이후, 데이터의 모델 학습 안정성과 효율성을 높이기 위해 정규화 과정을 거쳤다. 정규화는 데이터의 범위를 일정한 스케일로 조정하여 학습 시 과도한 편차와 불균형을 방지하였다. 이러한 전처리 과정을 통해 포인트 클라우드와 GPS 데이터는 일관된 학습 가능한 형태로 변환되었으며, CVAE 모델 학습을 위한 준비가 완료되었다.

#### 3.2 제안 모델

본 연구에서는 포인트 클라우드 데이터를 효율적으로 처리하기 위해 PointNet 기반의 변분 오토인코더(VAE)에 GPS 데이터를 조건부 입력으로 추가한 조건부 변분 오토인코더(Conditional Variational

AutoEncoder, CVAE) 모델을 제안한다. 이 모델은 포인트 클라우드의 3차원 구조적 정보를 학습하는 동시에, GPS 데이터를 활용하여 공간적 맥락을 강화한다.

그림 1을 살펴보면, 인코더(Encoder) 부분에서는 PointNet 구조를 기반으로 포인트 클라우드 데이터를 처리하며, GPS 데이터는 조건부 입력  $c$ 로 사용된다. 포인트 클라우드 입력  $x$ 는 GPS 데이터와 결합(concatenate)되어 인코더로 들어가며, PointNet은 포인트 클라우드의 주요 특징을 추출하고 이를 고차원 잠재 벡터  $z$ 로 변환한다. PointNet의 구조적 특징인 좌표 변환 네트워크와 다중층 퍼셉트론(MLP) 구조는 데이터의 불변성을 보장하며, GPS 조건은 인코더의 학습 과정에서 결합되어 시공간적 상관관계를 학습한다.

디코더(Decoder) 부분에서는 인코더에서 생성된 잠재 벡터  $z$ 와 조건부 GPS 입력  $c$ 를 다시 결합하여 입력으로 사용된다. 디코더는 입력된 벡터를 원래의 포인트 클라우드 형태로 복원하기 위해 다층 퍼셉트론(MLP)과 업샘플링 레이어를 사용하여 데이터를 재구성한다. 이 과정에서 GPS 데이터는 재구성 과정에 반영되어 모델의 공간적 정확성을 높인다.

제안된 CVAE 모델의 학습은 수식 1에서 보이는 것처럼 KL 손실(Kullback-Leibler Divergence)과 Chamfer Loss를 결합한 손실 함수를 최적화하는 방식으로 이루어진다.

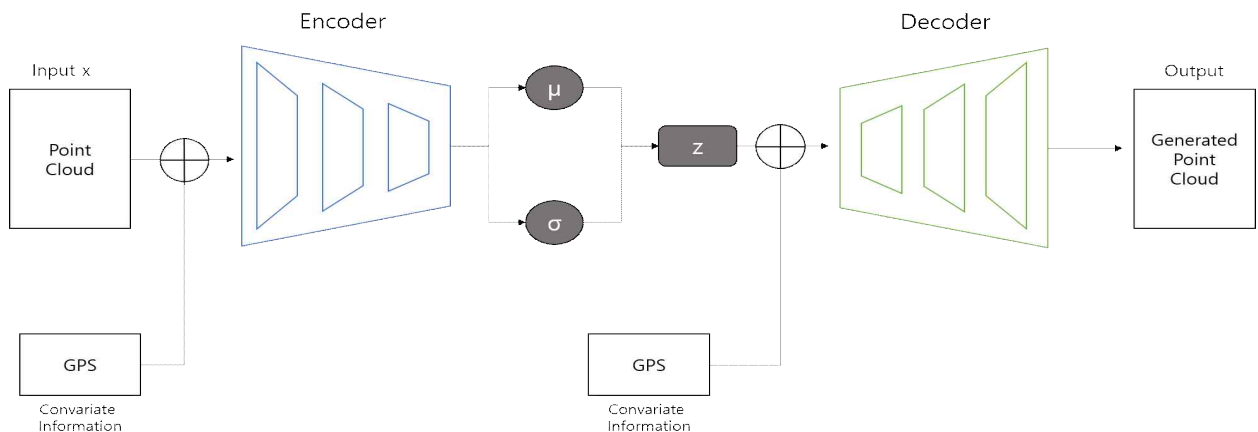


그림 1. 제안 모델 아키텍처  
Fig. 1. Proposed Model Architecture



$$L = E_{q(z|x,c)}[ChamferLoss(x,\hat{x})] - D_{KL}(q(z|x,c)||p(z|c)) \quad (1)$$

수식 2는 KL 손실을 나타내며, 인코더에서 학습된 잠재 분포  $q(z|x,c)$ 와 가우시안 prior 분포  $p(z|c)$  간의 차이를 최소화하여 잠재 공간의 연속성과 일반화 능력을 보장한다.

$$D_{KL}(q(z|x,c)||p(z|c)) = \frac{1}{2} * \sum_{i=1}^d (\sigma_i^2 + \mu_i^2 - \log(\sigma_i^2) - 1) \quad (2)$$

수식 3은 Chamfer Loss를 나타내며, 재구성된 포인트 클라우드  $\hat{x}$ 와 원본 포인트 클라우드  $x$  간의 거리를 측정하여, 두 데이터 간의 유사성을 극대화한다. Chamfer Loss는 각 점에서 가장 가까운 점까지의 거리를 평균하여 계산 된다.

$$ChamferLoss(x,\hat{x}) = \sum_{x \in X} \min_{\hat{x} \in \hat{X}} \|x - \hat{x}\|^2 + \sum_{\hat{x} \in \hat{X}} \min_{x \in X} \|\hat{x} - x\|^2 \quad (3)$$

이 Chamfer Loss는 포인트 간의 비대칭성을 고려하여 원본 데이터와 재구성된 데이터 간의 형상 유사성을 보장하며, 포인트 클라우드의 정확한 재구성을 가능하게 한다. 이러한 손실 함수 설계를 통해 제안된 PointNet 기반 CVAE 모델은 포인트 클라우드의 구조적 특징을 효과적으로 학습하며, GPS 데이터를 조건부로 사용하여 포인트 클라우드의 시공간적 맥락을 강화할 수 있다.

#### IV. 실험 및 결과

그림 2는 데이터 전처리 과정에서 수행된 클러스터링 결과를 보여준다. 그림 2의 (a)는 데이터 정규화(normalization) 이전의 k-means 클러스터링 결과로, 각 클러스터가 색상으로 구분되어 나타난다. 그림 2의 (b)는 데이터 정규화 후의 클러스터링 결과를 나타내며, 데이터의 분포가 보다 균일해지고 클러스터 간의 경계가 명확해졌음을 확인할 수 있다. 이러한 결과는 정규화를 통해 모델 학습 시 데이터의 스케



그림 2. 정규화 전후의 K-means 클러스터링 시각화  
Fig. 2. Visualization of K-means Clustering Before and After Normalization

일 문제를 해결하고 안정성을 높일 수 있음을 시사한다. 이러한 시각적 비교는 제안된 PointNet 기반 CVAE 모델의 학습 과정에서 데이터 전처리와 조건부 GPS 입력의 중요성을 강조하며, 모델이 고차원 포인트 클라우드 데이터를 효과적으로 처리하고 재구성하는 데 있어 유의미한 성능을 발휘함을 보여준다.

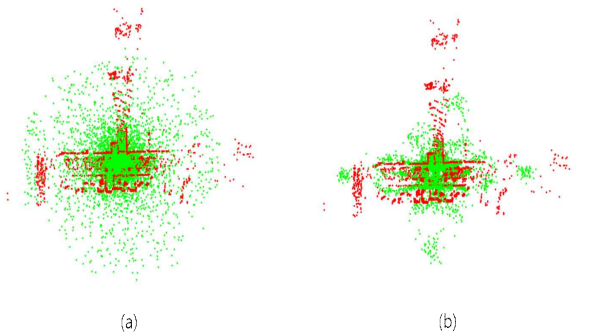


그림 3. CVAE 모델 학습 전후의 포인트 클라우드 출력  
Fig. 3. Point Cloud Output Before and After CVAE Model Training

그림 3은 학습 전후의 모델 출력 결과를 비교하여 보여준다. 빨간색 점은 실제 포인트 클라우드 데이터를 나타내며, 초록색 점은 모델의 출력 결과를 나타낸다. 그림 3의 (a)는 학습 이전의 모델 출력으로, 실제 포인트 클라우드와의 차이가 크게 나타난다. 그림 3의 (b)에서는 학습이 완료된 후의 결과를 보여주며, 모델이 실제 데이터와 더욱 유사하게 재구성되었음을 확인할 수 있다. 제안된 CVAE 모델은 GPS 데이터를 조건으로 사용하여 포인트 클라우드의 시공간적 맥락을 반영함으로써 재구성 성능이 향상되었음을 시각적으로 확인할 수 있다.

## V. 결론

본 연구에서는 PointNet 기반의 조건부 변이 오토 인코더(CVAE) 모델을 제안하여 포인트 클라우드 데이터를 효율적으로 학습하고 재구성하는 방법을 탐구하였다. GPS 데이터를 조건부 입력으로 사용함으로써 포인트 클라우드의 시공간적 맥락을 더욱 정교하게 반영할 수 있었다. 실험 결과, 제안된 모델은 기존의 단순한 VAE 모델에 비해 더 정확하게 포인트 클라우드를 재구성하였으며, 학습 전과 후의 시각적 비교를 통해 성능 향상을 확인할 수 있었다. 특히, Chamfer Loss를 사용하여 원본 데이터와 재구성된 데이터 간의 유사성을 극대화함으로써, 재구성의 정확도를 높였다.

또한, 데이터 전처리 과정에서 클러스터링 및 정규화가 모델의 학습 효율성과 안정성에 미치는 영향을 분석하였다. 정규화된 데이터는 클러스터링 결과에서 보다 균일한 분포를 보였으며, 이는 학습 과정에서 스케일 불균형을 줄여 모델 성능을 향상시켰다.

## 참고 문헌

- [1] Rusu, R. B., & Cousins, S. (2011). "3D is here: Point Cloud Library (PCL)." IEEE International Conference on Robotics and Automation (ICRA), pp. 1-4.
- [2] Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., & Bennamoun, M. (2020). "Deep learning for 3D point clouds: A survey." IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(12), 4338-4364.
- [3] Groves, P. D. (2013). Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems. Artech House.
- [4] Kingma, D. P., & Welling, M. (2014). "Auto-Encoding Variational Bayes." Proceedings of the International Conference on Learning Representations (ICLR).
- [5] Sohn, K., Lee, H., & Yan, X. (2015). "Learning

Structured Output Representation Using Deep Conditional Generative Models." Advances in Neural Information Processing Systems (NIPS), pp. 3483-3491.

- [6] Geiger, A., Lenz, P., & Urtasun, R. (2012). "Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3354-3361.

# 다양한 날씨 조건을 반영한 이미지 증강을 통한 자율주행을 위한 강화학습 데이터 생성

조준영\*, 최서인\*, 온병원\*\*† 정동원\*\*†

## Generating autonomous driving reinforcement learning data through image augmentation reflecting various weather condition

Junyoung Jo\*, Seoin Choi\*, Byung-Won On\*\*†, and Dongwon Jeong\*\*†

### 요 약

최근 험지 환경에서의 자율주행 연구는 다양한 변인 요소(토지 표면, 날씨 등)를 고려해야 하지만, 시뮬레이션에서 이러한 모든 조건을 구현하는 것은 현실적으로 어려운 과제이다. 본 연구에서는 단일 환경에서도 FUNIT을 활용하여 다양한 날씨 조건을 반영한 증강 이미지를 생성하고, 이를 강화학습에 사용할 수 있는 데이터로 활용하는 방안을 제안한다. 단일 환경에서 촬영된 이미지를 FUNIT으로 변환하여 비, 눈, 안개 등 다양한 날씨 조건을 반영한 이미지를 생성하고, 이를 강화학습 에이전트의 학습 데이터로 활용한다. 이를 통해 시뮬레이션 환경에서 제한된 조건의 데이터를 다양하게 확장하여, 에이전트의 일반화 성능을 높일 수 있는 가능성을 제시한다.

### Abstract

Recent research on autonomous driving in roughland environments requires consideration of various variable factors (land surface, weather, etc.), but implementing all these conditions in simulation is practically a difficult task. In this study, we propose to use FUNIT even in a single environment to generate augmented images reflecting various weather conditions and to use them as data that can be used for reinforcement learning. Images taken in a single environment are converted into FUNIT to generate images reflecting various weather conditions such as rain, snow, and fog, which are used as training data for reinforcement learning agents. Through this, we present the possibility of enhancing the generalization performance of agents by expanding the data of limited conditions in a simulation environment in various ways.

### Key words

autonomous driving, off-road driving, data augmentation, simulation

---

\* 국립군산대학교 소프트웨어학부 학부생 {joon0zo1022, seoinc46}@kunsan.ac.kr

\*\* 국립군산대학교 소프트웨어학부 교수, {bwon, djeong}@kunsan.ac.kr(교신저자)

† 공동교신저자

※ 본 논문은 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음(2023-0-00065).

## I. 서론

자율주행 기술은 빠르게 발전하고 있으며, 특히 험지 환경에서의 자율주행은 도로 조건이 예측 불가능하고 환경 변화가 심하기 때문에 더욱 주목받고 있다[1]. 험지 환경에서 자율주행 차량이 안정적으로 주행하기 위해서는 다양한 조건에서의 학습 데이터가 필수적이다. 하지만 현실적인 시뮬레이션 환경에서 토지 표면, 날씨, 장애물 등 모든 변인 요소를 고려하여 환경을 구축하는 것은 기술적·자원적 한계로 인해 매우 어려운 과제이다. 따라서 이러한 다양한 환경 요소를 반영한 데이터를 확보하지 못하면 자율주행 에이전트의 일반화 성능이 저하될 수 있으며, 실제 환경에서의 적응력이 떨어질 가능성이 있다.

기존의 이미지 증강 기법들은 주로 기하학적 변형(회전, 크기 조절, 색상 조정 등)에 집중되어 있어, 현실적인 날씨 조건이나 지형 특성을 충분히 반영하지 못하는 한계가 있다[2]. 특히 비, 눈, 안개와 같은 날씨 변인은 자율주행 성능에 큰 영향을 미치지만, 이를 시뮬레이션에서 모두 재현하는 것은 매우 어렵다. 이러한 문제를 해결하기 위해, 본 연구에서는 단일 환경 이미지에서 다양한 날씨 조건을 생성할 수 있는 FUNIT(Few-shot Unsupervised Image-to-Image Translation)[3]을 활용하여 자율주행을 위한 데이터 다양성을 확보하고자 한다.

FUNIT은 적은 양의 타겟 스타일 데이터를 바탕으로 도메인 전환을 수행할 수 있는 모델로, 단일 환경 이미지에서 비, 눈, 안개 등의 다양한 날씨 조건을 반영한 증강 이미지를 생성할 수 있다. 본 연구는 FUNIT을 통해 생성된 증강 이미지를 자율주행 강화학습에서 사용할 수 있는 학습 데이터로 활용하는 가능성을 탐구한다. 이러한 접근은 시뮬레이션에서 구현하기 어려운 날씨 변인을 반영한 데이터를 제공함으로써, 자율주행 연구의 데이터 부족 문제를 해결하고 에이전트의 일반화 성능을 향상시킬 수 있는 방안을 제시한다.

본 논문에서는 FUNIT을 활용하여 단일 환경 이미지에서 다양한 날씨 조건을 반영한 증강 이미지를 생성하는 방법을 제안하고, 이를 통해 자율주행

학습에서의 데이터 다양성을 확보하는 가능성을 논의한다. 이를 바탕으로 험지 자율주행 연구에서 데이터 증강의 가능성과 중요성을 제시하고자 한다.

## II. 제안 방안

### 2.1 데이터셋 구축

본 연구에서는 BDD100K 데이터셋[4]을 사용하여 다양한 날씨 조건에서의 증강 이미지를 생성하고자 한다. BDD100K 데이터셋은 날씨와 시간대를 고려한 다양한 자율주행 환경 데이터를 제공하며, 총 6개의 날씨 속성과 3개의 시간대 속성이 포함된다. 각각의 이미지는 이 두 가지 속성을 모두 가지지만, 날씨와 시간대를 동시에 고려하여 데이터를 구축할 경우 클래스 간 불균형이 심화되고, 특정 조합의 이미지 수가 적어져 모델 학습에 제한이 생긴다.

이를 해결하기 위해, 본 연구에서는 날씨 속성만을 기준으로 데이터셋을 분류하여 학습에 사용하였다. 날씨에 따라 분류된 데이터셋은 각 클래스 간의 데이터 불균형 문제를 최소화하고, 다양한 날씨 조건에서의 증강 이미지 생성을 가능하게 한다. 이를 통해 생성된 데이터는 자율주행 강화학습을 위한 학습 데이터로 사용될 수 있으며, 시뮬레이션 환경에서 구현하기 어려운 다양한 날씨 변인을 반영한다.

표 1. BDD100K 데이터셋 분포  
Table 1. BDD100K Dataset Distribution

대분류 \ 소분류	소분류	개수	총합
날씨	맑음	37,342	61,728
	비	5,058	
	눈	5,547	
	흐림	8,770	
	부분적 흐림	4,881	
	안개	130	
시간대	낮	29,699	61,728
	새벽/황혼	4,500	
	밤	27,529	

## 2.2 FUNIT 모델을 이용한 이미지 증강

본 논문은 FUNIT(Few-shot Unsupervised Image-to-Image Translation) 모델을 사용하여 단일 환경 이미지에서 다양한 날씨 조건의 증강 이미지를 생성하는 방안을 제안한다. FUNIT은 적은 양의 타겟 스타일 이미지를 바탕으로 특정 도메인에서 다른 도메인으로 이미지를 전환할 수 있는 모델이다. 자율주행을 위한 데이터 다양성을 확보하기 위해, FUNIT을 통해 단일 환경 이미지를 비, 눈, 안개 등 다양한 날씨 조건으로 변환한다.

FUNIT의 학습 과정에서, 날씨 속성별로 분류된 BDD100K 데이터셋을 사용하여 각 날씨 조건의 특징을 학습하도록 모델을 훈련시켰다. 이 모델은 타겟 스타일 이미지의 날씨 조건을 소량의 데이터로부터 학습하여, 입력 이미지에 반영할 수 있다. 예를 들어, 맑은 날씨의 원본 이미지에 비 조건의 스타일을 적용하면, 흐려진 배경과 물방울이 있는 이미지로 변환된다. 이러한 방식으로, 다양한 날씨 조건을 반영한 증강 이미지가 생성된다.

FUNIT을 통해 생성된 증강 이미지는 실제 주행 환경에서 경험할 수 있는 다양한 날씨 조건을 반영하며, 강화학습 에이전트의 학습 데이터로 사용될 수 있다. 이를 통해 단일 환경에서도 다양한 조건을 고려한 학습 데이터를 확보하여, 에이전트의 일반화 성능을 높이고 실제 환경에서 더 잘 적응할 수 있도록 한다.

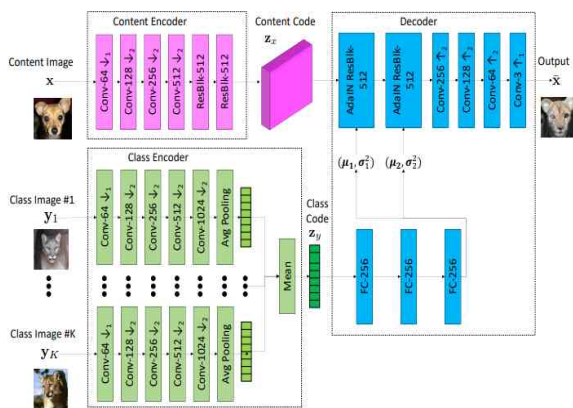


그림 1. FUNIT 아키텍처  
Fig. 1. FUNIT Architecture

## III. 실험

### 3.1 실험 환경

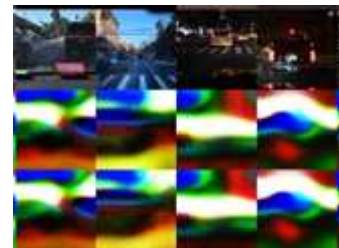
본 논문에서 제안하는 증강방안을 검증하기 위한 실험 환경은 표 2와 같다.

표 2. 실험 환경  
Table 2. Experimental environments

구분	사양
OS	Windows 11
CPU	13 <sup>th</sup> Gen Intel(R) Core(TM) i9-13900K
Memory	32GB
GPU	NVIDIA GeForce RTX A6000
Python	3.8.19
Pytorch	2.2.2

### 3.2 실험 결과

FUNIT 모델은 학습이 진행됨에 따라 점진적으로 날씨 조건을 반영하는 능력을 습득하며, 최종 학습 단계에서 다양한 조건을 비교적 높은 품질로 구현할 수 있게 된다. 초기 단계에서는 이미지 왜곡이 심하고 날씨 조건의 표현이 불분명하지만, 학습이 진행되면서 변환된 이미지의 질감과 날씨 특성이 점차 개선된다. 최종 학습 단계에서는 비, 눈, 안개 등의 조건을 이미지에 반영하였다.



(a) 2,500 에포크 결과



(b) 10,000 에포크 결과



(c) 20,000 에포크 경과

그림 2. FUNIT 모델의 에포크별 생성 이미지 변화  
Fig. 2. Generated Image Variations by Epochs in FUNIT Model

#### IV. 결론 및 향후 연구

본 연구에서는 험지 환경에서의 자율주행을 위한 다양한 날씨 조건의 학습 데이터를 생성하기 위해 FUNIT을 활용한 이미지 증강 방법을 제안하였다. FUNIT 모델을 통해 단일 환경 이미지에 비, 눈, 안개와 같은 다양한 날씨 조건을 반영하여 데이터 다양성을 확보하고, 강화학습 에이전트의 학습 데이터로 활용할 가능성을 탐색하였다.

실험 결과, FUNIT 모델은 학습이 진행됨에 따라 점진적으로 날씨 조건을 반영하는 능력을 습득하며, 최종 학습 단계에서는 다양한 날씨 조건을 비교적 높은 품질로 구현할 수 있게 되었다. 초기 단계에서는 왜곡이 심하고 날씨 조건의 표현이 명확하지 않았으나, 학습이 진행됨에 따라 이미지의 질감과 날씨 특성이 점차적으로 개선되었다. 최종적으로, 비, 눈, 안개 등 다양한 조건을 사실적으로 반영하여 시각적 품질과 조건 반영의 정확성이 크게 향상되었다. 이러한 증강 이미지는 실제 시뮬레이션 환경에서 구현하기 어려운 다양한 날씨 조건을 시뮬레이션 환경에서 간접적으로 학습할 수 있게 하여, 강화학습 에이전트의 성능 향상에 기여할 수 있을 것으로 기대된다.

향후 연구에서는 다음과 같은 방향으로 연구를 진행할 수 있다. 첫째, FUNIT 모델을 더욱 발전시켜 안개와 같은 특정 조건에서의 표현력이 향상되도록 모델을 조정하거나 추가적인 데이터 증강 방법을 모색할 필요가 있다. 둘째, 생성된 증강 데이터를 실제 강화학습 환경에 적용하여, 다양한 날씨

조건이 에이전트의 학습 성능에 미치는 영향을 평가하고, 에이전트의 실제 환경 적응 능력을 검증할 필요가 있다. 셋째, 날씨 조건 외에도 다양한 지형, 장애물, 조도 변화 등을 반영하여, 보다 복잡한 조건에서 자율주행 학습 데이터를 증강하는 방안을 제안할 수 있다.

#### 참고 문헌

- [1] Liu, O., Yuan, S., & Li, Z. (2020). A survey on sensor technologies for unmanned ground vehicles. In 2020 3rd International Conference on Unmanned Systems (ICUS) (pp. 638-645). IEEE.
- [2] Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621. Tracts in Advanced Robotics, vol. 42, 2008.
- [3] Liu, M.-Y., Huang, X., Mallya, A., Karras, T., Aila, T., Lehtinen, J., & Kautz, J. (2019). Few-shot unsupervised image-to-image translation. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 10551-10560.
- [3] Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F. & Darrell, T. (2020). Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2636-2645).

# YOLO 모델을 활용한 컨테이너 표면 손상 및 녹 탐지 기반 위험도 평가 시스템

박정일\*, 온병원\*\*, 정동원\*\*, 박재민\*\*

## A YOLO-Based System for Risk Assessment of Container Surface Damage and Corrosion Detection

Jeongil Park\*, Byungwon On\*, and Dongwon Jeong\*

### 요 약

본 연구는 YOLO 모델을 활용하여 컨테이너 표면의 파손과 녹을 탐지하고, 위험도에 따라 유지보수 우선순위를 결정하는 시스템을 제안한다. 이를 통해 유지보수를 자동화하고 비용 효율성과 화물 안전성을 높인다. 손상 부위의 크기, 위치, 중첩 여부를 바탕으로 고위험 구역을 우선 유지보수하며, 중위험 구역에는 예방적 유지보수를 권장한다. 위험도는 색상으로 시각화하여 신속한 인식과 대응을 지원한다. 향후 연구에서는 손상 확산 예측 모델과 환경 조건에 따른 녹 면적 가중치 조정 방법을 제안한다.

### Abstract

This study proposes a YOLO-based system for detecting container surface damage and corrosion, prioritizing maintenance by risk levels. The system automates maintenance, enhancing cost efficiency and cargo safety by assessing damage size, location, and overlap. High-risk zones receive prioritized maintenance, and medium-risk zones are recommended for preventive actions. Risk levels are visualized with color coding to support rapid recognition and response. Future work will include predictive modeling of damage progression and dynamic adjustment of corrosion weighting based on environmental factors.

### Key words

YOLO, Risk Assessment, Container, Corrosion, Detection

---

\* 국립군산대학교 소프트웨어학과, 1901596@kunsan.ac.kr

\*\* 국립군산대학교 소프트웨어학과, {bwon, djeong}@kunsan.ac.kr(교신저자)

※ 본 논문은 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065).

## I. 서 론

국제 물류와 해운 산업에서 컨테이너는 대량의 화물을 안전하게 보호하고 운송하기 위한 필수적인 장비로 사용된다. 그러나 장기적인 해상 운송과 보관 과정에서 컨테이너는 염분과 습기에 지속적으로 노출되어 부식이 발생하거나 외부 충격으로 인한 물리적 손상을 입기 쉽다. 이러한 손상은 컨테이너의 구조적 강도를 저하시키며, 심각한 경우 내부 화물의 안전에도 위협을 줄 수 있다. 해양 환경의 특성상 염분과 습도는 금속 부식을 가속화시키며, 시간이 지남에 따라 손상 부위가 급속히 확산되어 심각한 위협을 초래할 수 있다.

현재 컨테이너 점검은 대부분 육안 검사를 통해 이루어지며, 인력이 직접 손상 상태를 평가하고 수리를 수행하는 방식이 일반적이다. 그러나 이러한 점검 방식은 대규모 물류 환경에서 매우 비효율적이며, 시간과 비용이 과도하게 소요되는 문제가 있다. 특히, 점검 과정에서 작업자의 숙련도와 경험에 따라 손상이 제대로 발견되지 않거나, 평가 결과가 일관되지 않을 가능성도 존재한다. 이러한 한계는 컨테이너 유지보수의 신뢰성을 낮추고, 결과적으로 운송 안전성에도 부정적인 영향을 미칠 수 있다.

이에 따라, 최근에는 딥러닝 기반의 컴퓨터 비전 기술을 활용한 컨테이너 자동 점검 시스템의 필요성이 대두되고 있다. 특히, 객체 탐지 모델 중 하나인 YOLO(You Only Look Once) 모델은 실시간 탐지 성능이 뛰어나며, 컨테이너와 같은 대규모 구조물의 손상을 효과적으로 식별할 수 있는 잠재력을 가지고 있다. YOLO 모델을 활용한 자동화된 손상 탐지 시스템은 대규모 컨테이너 운송과 보관이 이루어지는 산업에서 효율적인 유지보수 관리의 기반이 될 수 있다.

본 연구에서는 YOLO 모델을 활용하여 컨테이너 표면의 파손 및 녹 발생 여부를 자동으로 탐지하고, 이를 바탕으로 손상의 위험도를 평가하는 시스템을 제안한다. 제안된 시스템은 손상 부위의 크기와 위치를 분석하여 위험도를 산출하며, 녹이 발생한 경우 추가적인 가중치를 부여하여 위험도를 높인다. 이러한 자동화된 위험도 평가 시스템을 통해 컨테

이너 유지보수의 효율성을 높이고, 예방적 유지보수를 통해 화물의 안전성을 확보함으로써 유지보수 비용 절감과 사고 예방에 기여할 수 있을 것으로 기대한다.

## II. 관련 논문

컨테이너 및 선박의 주요 대상을 식별하기 위한 연구로[1],YOLO 기반 경량화와 트리플 주의 메커니즘을 적용하여 실시간 탐지 성능을 높이고, 해상 환경에서의 중요한 대상을 효율적으로 식별하는 방법을 제안하였다[1]. 이 연구는 경량화된 YOLO 모델과 주의 메커니즘을 결합하여 모델의 정확성과 속도를 개선하였으며, 컨테이너 관리에서도 유사한 접근이 가능함을 시사한다.

또한, 풍력 터빈의 다중 결함 탐지를 위해 YOLOv8 모델을 활용한 다중 스케일 결함 탐지 알고리즘 YOLOv8-WTDD를 제안하였다[2]. 이 연구는 결함의 크기와 위치에 따라 다중 스케일 탐지를 수행하며, 풍력 터빈의 상태를 실시간으로 모니터링하는 데 유용하다. 본 연구와 유사하게, 이 접근법은 다양한 크기의 결함을 효과적으로 탐지하고 분석하여, 손상된 부분의 우선순위를 결정하는 데 참고할 수 있다.

한편, 해양 환경에서 물고기 탐지 및 추적을 위한 연구로 개선된 프레임 차이 기법과 YOLO-CTS 모델을 결합한 탐지 및 추적 시스템을 제안하였다 [3]. 이 시스템은 물고기 위치와 이동을 실시간으로 추적하며, 해양 환경에서의 객체 탐지에 최적화된 모델로 설계되었다. 이러한 연구는 컨테이너의 손상 부위와 상태 변화를 실시간으로 모니터링하고 추적하는 데 유용한 방법론을 제공하며, 지속적인 유지보수 및 관리에 응용될 가능성을 보여준다.

## III. 제안 방안

본 연구에서 제안하는 시스템은 크게 (1) 손상 탐지, (2) 데이터 전처리 및 통합, (3) 위험도 산출, (4) 유지보수 우선순위 도출의 네 가지 주요 단계로 구성된다. 각 단계에서의 방법론은 다음과 같다.



### 3.1 손상 탐지

본 연구에서는 YOLO 모델을 활용하여 컨테이너 표면의 파손과 녹 발생 여부를 자동으로 탐지하는 방안을 제안한다. YOLO 모델은 객체 탐지 모델로서, 손상 영역을 바운딩 박스 형태로 추출하여 각 손상의 위치와 크기를 파악할 수 있다. 이를 통해 손상 부위의 면적을 계산하여 파손의 규모를 정량화할 수 있다.

본 연구에서는 '파손'과 '녹'을 두 가지 클래스로 정의하여 YOLO 모델을 학습시켰다. 모델은 고해상도 이미지를 입력으로 받아 손상 부위를 정밀하게 탐지할 수 있으며, 실시간 처리가 가능하도록 YOLO의 빠른 처리 속도를 활용하였다. 이렇게 추출된 바운딩 박스 정보를 통해 손상 부위의 위치와 면적을 확보함으로써, 후속 위험도 평가에 필요한 기초 데이터를 제공한다.



그림 1. YOLO를 통한 객체 탐지  
Fig. 1. Object detection with YOLO

표 1. YOLO 실험 결과

Table 1. YOLO Experiment Results

class	Box(p)	R	mAP(50)	mAP(50-95)
All	0.742	0.574	0.595	0.524
Broken	0.911	0.981	0.987	0.936
Rust	0.573	0.168	0.203	0.111

### 3.2 데이터 전처리 및 통합

표 2. 실험 데이터 세트

Table 2. Experimental Data Sets

데이터 개수	녹슨 컨테이너 데이터셋	파손 컨테이너 데이터셋	총 데이터 셋
All	6,223	5,181	11,404
Train	5,913	4,539	10,452
Valid	154	426	580
Test	156	216	372

YOLO 모델로 탐지된 손상 부위의 바운딩 박스 좌표는 YOLO 형식으로 제공되며, 이를 통해 손상의 크기와 위치 정보를 추출할 수 있다. 추가적으로 파손과 녹이 중첩된 영역을 분석하기 위해 Shapely 라이브러리를 사용하여 바운딩 박스를 다각형 형태로 변환하고, 중첩 여부를 계산한다. 이 과정에서는 다음과 같은 지수를 산출하여 통합 데이터셋을 구축한다.

파손 면적 지수 (Damage Index, DI): 탐지된 파손 면적을 기반으로 산출한 지수로, 손상의 심각도를 나타낸다.

녹 면적 지수 (Rust Area Index, RAI): 녹이 발생한 영역의 면적을 기반으로 산출한 지수로, 추가적인 위험 요소를 반영한다.

중첩 면적 지수 (Critical Rust-Damage Index, CRI): 파손과 녹이 중첩된 영역을 산출하여, 중첩된 손상이 일반적인 손상보다 심각한 위험을 초래할 가능성이 크다는 점을 반영한 지수이다.

이러한 지수들은 후속 위험도 산출 단계에서 통합 데이터셋의 일관된 형태로 저장되어, 체계적인 분석을 가능하게 한다.

### 3.3 위험도 산출

위험도 산출 단계에서는 손상 부위의 크기, 위치, 녹 발생 여부를 종합적으로 고려하여 위험도를 평가한다. 구체적으로, 각 지수를 바탕으로 다음과 같은 위험도 산출식을 적용한다:

$$Risk\ Score = \alpha \times RAI + \beta \times DI + \gamma \times CRI$$

여기서  $\alpha$ ,  $\beta$ ,  $\gamma$ 는 각 지수에 할당된 가중치로, 실험을 통해 최적화된 값으로 설정한다. 이를 통해 산

출된 위험도는 유지보수 우선순위 도출의 기초 데이터로 사용된다.

#### IV. 실험 결과 및 분석

##### 4.1 고위험 구역 우선 수리 제안

YOLO 모델을 통해 컨테이너 표면의 파손과 녹 발생 여부를 탐지하고, 이를 바탕으로 위험도 지표를 산출한 결과, High 등급에 해당하는 손상 부위가 다수 확인되었다. High 리스크 등급은 리스크 점수가 일정 임계값 이상인 영역으로, 구조적 강도 저하와 내부 화물 손상의 원인이 될 수 있는 심각한 손상을 의미한다. 따라서 이러한 고위험 구역은 유지보수에서 우선적으로 수리해야 할 부위로 판단되며, 신속한 조치가 필요하다.

##### 4.2 예방적 유지보수 추천

탐지 결과에서 Medium 등급 이상의 위험도를 가진 손상 부위에 대해서는 예방적 유지보수를 권장한다. Medium 등급은 아직 심각한 위험을 초래하지는 않지만, 지속적으로 노출될 경우, 손상이 확산될 가능성이 있는 영역이다. 따라서 이러한 영역은 추가적인 손상을 방지하기 위해 사전적인 유지보수 조치가 필요하다. 예방적 유지보수는 컨테이너의 수명을 연장하고, 추후 발생할 수 있는 대규모 수리 비용을 절감하는 효과를 기대할 수 있다.

#### V. 결론 및 향후 연구

본 연구에서는 YOLO 모델을 활용하여 컨테이너 표면의 녹과 파손을 자동으로 탐지하고, 이를 기반으로 위험도 지표를 산출하여 유지보수 우선순위를 도출하는 방법을 제안하였다. 제안된 방법은 유지보수 인력을 대체하거나 보조하는 방식으로, 컨테이너 관리의 효율성을 크게 향상시킬 수 있다. 또한, 위험도 분석을 통해 고위험 구역을 우선적으로 처리함으로써, 유지보수 비용 절감과 작업의 신속성을 높이는 데 유용하게 활용될 수 있다. 이러한 접근은

대규모 물류 환경에서 더욱 효과적일 것으로 기대되며, 운송 화물의 안전성을 보장하는 데 기여할 수 있다.

향후 연구에서는 시간 경과에 따른 손상 예측 모델을 추가 개발하여, 손상의 진행 속도를 파악하고 이를 예방할 수 있는 구체적인 유지보수 계획을 수립하는 데 집중할 예정이다. 이를 통해 컨테이너의 장기적인 관리와 유지보수에 필요한 자원을 효율적으로 배분할 수 있을 것이다.

또한, 본 연구에서는 녹 면적 지수를 단순히 위험도에 추가 가중치로 반영하였으나, 향후 연구에서는 녹의 확산 가능성과 파손의 중첩 정도를 고려하여 가중치를 더욱 정교하게 조정하는 방안을 탐구할 필요가 있다. 특히, 녹이 발생한 부위가 파손과 중첩될 경우, 해당 영역의 부식이 파손을 더욱 악화시킬 가능성이 크므로, 중첩되는 면적의 위험도를 별도로 평가하는 지표를 개발할 수 있다.

이를 위해 녹의 면적뿐만 아니라 녹이 발생한 위치와 녹의 심각도를 기반으로 한 가중치 설정이 필요하다. 예를 들어, 특정 환경 조건(예: 높은 습도나 염분이 많은 환경)에서 녹이 더 빠르게 확산될 수 있는 경우, 이러한 조건을 반영한 동적 가중치를 적용할 수 있는 예측 모델을 구축할 수 있다. 이러한 추가적인 가중치 설정 연구는 단순히 현재 상태를 평가하는 것을 넘어서, 향후 손상 위험성을 예측하여 예방적 유지보수 계획을 세우는 데 중요한 역할을 할 것이다.

따라서 향후 연구에서는 녹 면적 지수에 대한 가중치를 다양한 환경 변수와 손상 확산 경향을 고려하여 더욱 세분화하고, 이를 기반으로 보다 정밀한 유지보수 전략을 수립할 수 있는 방안을 제시할 예정이다.

#### 참 고 문 헌

- [1] P. Wang, S. Yang, G. Chen, W. Wang, Z. Huang, et al., "A Ship's Maritime Critical Target Identification Method Based on Lightweight and Triple Attention Mechanisms," *Journal of Marine Science and Engineering*, vol. 12, pp. 1839, 2024.
- [2] X. Xu, J. Hu, J. Yang, Y. Ran, and Z. Tan, "A

Fish Detection and Tracking Method based on Improved Inter-Frame Difference and YOLO-CTS," IEEE Transactions on Instrumentation and Measurement, 2024.

- [3] X. Yu, P. Yan, S. Zheng, Q. Du, and D. Wang, "YOLOv8-WTDD: Multi-Scale Defect Detection Algorithm for Wind Turbines," The Journal of Supercomputing, 2024.



# 유해조수 인식 데이터 및 이미지 전송을 위한 연구

백대현\*, 하영서\*, 조대제\*\*, 이영학\*\*\*,

## Study on Transmission of Hazardous Animal Recognition Data and Images

Daehyeon Baek\*, Yeongseo Ha\*, Daejae Jo\*\*, and Yeunghak Lee\*\*\*

### 요 약

유해조수는 멧돼지, 고라니, 까마귀, 까치 등이 있고, 유해조수에 의한 피해는 매년 100억 원 이상 꾸준히 발생하고 있다. 전체 피해액의 13%는 조류에 의한 피해로 큰 부분을 차지하고 있다. 유해조수를 퇴치 및 사용자의 조치가 이루어지기 위해서는 정확한 감지 정보를 제공하는 것이 중요하다. 본 논문은 유해조수 감지 정보를 컴퓨터 간의 전송을 위해서 윈도우에서 제공하는 OpenSSH 서버 기능을 이용하여 SFTP 환경을 구축하고 유해조수 인식 정보와 이미지를 전송하는 방법에 관해서 설명한다.

### Abstract

Hazardous birds and animals include wild boars, water deer, crows, and magpies, and damage caused by harmful birds and animals is steadily occurring more than 10 billion won every year. 13% of the total damage is caused by birds, accounting for a large portion. In order to deal with harmful birds and take user action, it is important to provide accurate detection information. This thesis describes how to establish an SFTP environment and transmit bird recognition information and images using the OpenSSH server function, which provides harmful bird detection information in a window for transmission between computers.

### Key words

SFTP(SSH File Transfer Protocol), YOLO(You Only Look Once), OpenSSH, MySQL

## I. 서 론

유해조수로 인한 농작물 피해는 매년 끊이지 않고 발생하고 있다. 유해조수에는 멧돼지, 고라니, 까마귀, 까치 등 다양한 종류가 있는데, 2018년부터 2022년까지 5년간 유해조수가 발생시킨 농작물 피

해는 약 540억 원에 달하고, 전체 피해액의 약 13%는 조류로 인한 피해인 것으로 확인되었다.[1]

데이터 송수신을 위한 방법에는 FTP(File Transfer Protocol), SFTP(SSH File Transfer Protocol), HTTP(Hyper Text Transfer Protocol), SMTP(Simple Network Management Protocol) 등이 있다. 본 논문에서

\* 국립안동대학교 컴퓨터공학과, eogus702@naver.com, ysha941118@gmail.com

\*\* 국립안동대학교 소프트웨어융합학과, djcho@anu.ac.kr,

\*\*\* 국립안동대학교 SW융합교육원, yhyi@anu.ac.kr(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)

서는 SFTP를 활용한 컴퓨터 간의 통신으로 유해조 수를 인식한 이미지를 전송하는 것을 목표로 한다.

본 논문의 2장에서는 데이터를 전송하기 위해 사용한 기술인 SFTP에 대해서 설명하고, 3장에서는 데이터 전송을 위한 SFTP 환경 구축 방법을 설명하고, 4장에서는 데이터 전송을 구현한 모습을 설명하고, 5장에서는 결론을 맺는다.

## II. 관련 기술

FTP는 인터넷을 통한 파일 송신과 수신만을 위해서 고안된 서비스 프로토콜로, 빠른 파일 전송을 위하여 데이터가 모두 다운로드 된 후 확인할 수 있는 특징을 갖는다[2]. FTP 서비스를 제공하는 서버와 접속하는 사용자 사이에 데이터 전송을 제어하기 위한 신호를 주고받기 위한 연결과 실제 데이터 전송을 위한 연결이 생성되는데, 이것을 이용하여 서버와 클라이언트 사이의 파일을 송수신한다.

그리고, FTP는 프로토콜 구조가 매우 단순해서 공격자가 전송되는 데이터를 가로챌 때 데이터 정보가 그대로 노출된다는 단점을 가지고 있어서 보안이 매우 취약하다[3].

본 논문에서 사용한 기술인 SFTP는 SSH(Secure Shell) 프로토콜의 연결계층에서 제공하는 기능을 바탕으로 파일 접근 및 전송 등의 기능을 수행하고, 파일을 전송할 때 암호화해서 전송하는 방식을 말한다. 데이터가 암호화해서 전송되므로 보안 측면에서 FTP보다 뛰어나다는 특징을 가지고 있다[4].

## III. SFTP 환경 구축

### 3.1 개발 환경

개발환경은 윈도우(Windows) 운영체제에서 설계 되었으며, 윈도우에서 자체적으로 제공하는 OpenSSH 서버 기능을 이용하여 SFTP 통신을 구현했다. OpenSSH는 SSH 프로토콜을 이용하여 구현한 오픈소스이다. SSH는 도청이나 하이재킹 등의 공격을 막기 위해 모든 트래픽을 암호화시키며 클라이언트와 서버의 통신에서 안전한 접속을 할 수 있는 통신 채널을 제공한다[5].

### 3.2 SFTP 설정

윈도우에서 제공하는 OpenSSH 서버를 설치한 후, SFTP 설정을 진행한다. SFTP는 SSH를 매개체로 이용하기 때문에 SSH 폴더로 이동하여 설정을 진행한다. SSH 폴더는 C 드라이브의 ProgramData 폴더로 이동하면 찾을 수 있는데, 기본 설정이 숨김으로 되어있어서 숨김파일 보기를 선택하여 접근한다.

SSH 폴더 안의 sshd\_config 파일을 열어서 설정 변수를 확인한다.

```
65 #PermitUserEnvironment no
66 #ClientAliveInterval 0
67 #ClientAliveCountMax 3
68 #UseDNS no
69 #PidFile /var/run/sshd.pid
70 #MaxStartups 10:30:100
71 #PermitTunnel no
72 ChrootDirectory "C:\Users\USER\Desktop\webpage\public\image"
73 #VersionAddendum none
```

그림 1. SFTP 폴더 경로 지정  
Fig. 1. Set SFTP folder path

그림 1을 보면 UseDNS, PidFile, MaxStartups, PermitTunnel, ChrootDirectory 등 여러 가지 설정 변수들이 있는 것을 확인할 수 있는데, 이 중 ChrootDirectory 설정의 주석을 해제하여 변경해 준다. 해당 설정을 변경하는 이유는 SFTP를 이용하여 수신받은 이미지를 저장할 디렉토리를 지정해 주는 것으로, 다른 컴퓨터에서 전송한 이미지들은 그림1에서 설정한 경로의 폴더로 모이게 된다.

### 3.3 SFTP 사용자 설정

SFTP를 사용하기 위해 사용자를 구성한다. 사용자를 구성하기 위해 윈도우의 컴퓨터 관리 항목으로 이동하여 로컬 사용자 및 그룹을 선택하여 새 사용자를 추가할 수 있다. 새 사용자의 이름은 임의로 설정하고 암호를 설정한다. 그 후 생성된 사용자의 속성에서 소속 그룹을 제거하면 사용자 설정이 완료된다. 명령 프롬프트에 “사용자 이름@컴퓨터아이피”와 사용자를 생성할 때 설정한 비밀번호를 입력하여 SFTP에 접속한다.

해당 내용은 그림2에 나타나 있고, 그림 3에서는 lpwd 명령어를 이용하여 데이터를 전송받기 위해 설정한 디렉토리를 확인하는 모습을 나타내었다.

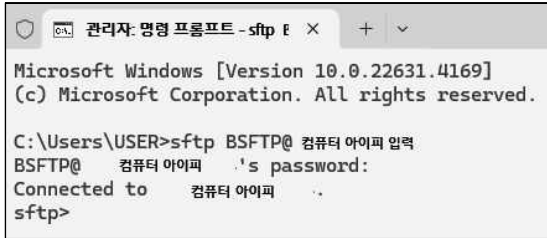


그림 2. SFTP 접속  
Fig. 2. SFTP Connection



그림 3. SFTP 디렉토리 확인  
Fig. 3. Verifying the SFTP Directory

#### IV. 데이터 전송 구현

유해조수 인식을 위해 YOLO(You Only Look Once) 모델을 사용하였다. 그 중 YOLOv8을 사용하였으며, 해당 모델은 이미지를 그리드로 나누고, 각 그리드에서 객체의 존재 여부, 바운딩 박스를 예측하고, 이미지 감지, 분류의 기능을 지원한다[6]. 그림 4에서는 YOLO 모델을 이용하여 탐지한 유해조수의 이미지를 앞에서 지정한 경로로 SFTP를 이용하여 전송하는 부분에 대한 코드를 나타내고 있다.

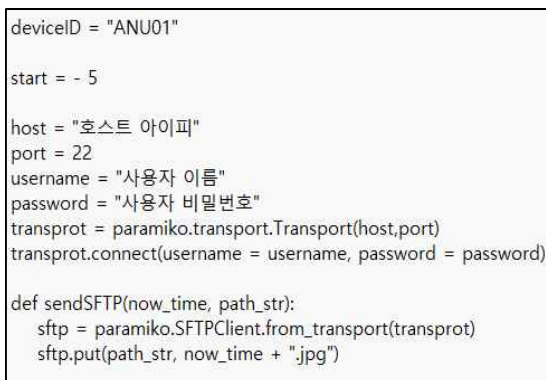


그림 4. SFTP 이미지 전송 코드  
Fig. 4. SFTP Image Transfer Code

그림 4를 보면 데이터를 수신할 호스트 컴퓨터의 아이피와 SFTP의 기본 포트 번호인 22와 앞에서 생성한 사용자의 이름과 비밀번호를 입력하여 연결한다. SFTP를 이용하여 전송할 이미지의 이름은 유해조수가 인식된 시간으로 지정한다.

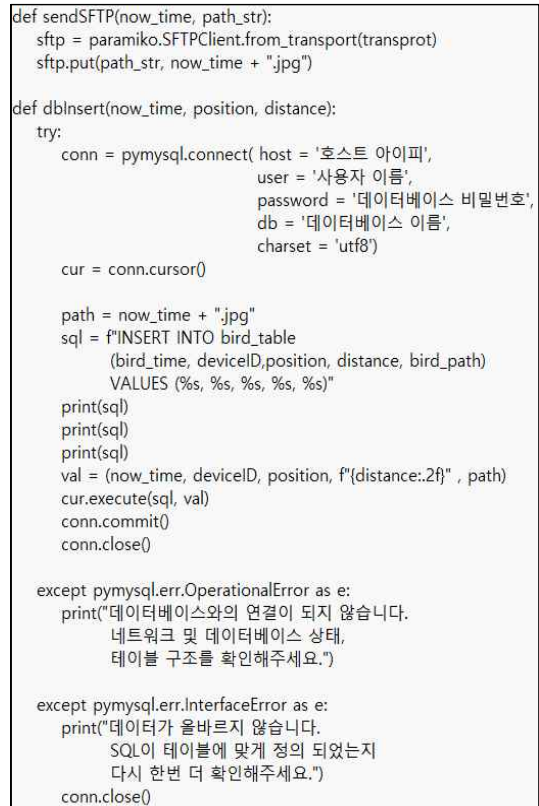


그림 5. 데이터베이스로 데이터 전송 코드  
Fig. 5. Data transfer code to database

그림 5를 보면 인식한 정보를 데이터베이스로 전송하는 코드를 나타낸다. 유해조수를 인식한 정보를 담기 위해 선택한 데이터베이스 관리 시스템(Database Management System)은 단순 처리에 적합하며, 다양한 스토리지 엔진을 병합해 데이터베이스를 구성할 수 있고, 대용량 데이터 처리가 가능한 MySQL을 사용했다[7]. 데이터베이스에 담은 데이터는 인식한 시간, 인식한 장비 이름, 유해조수의 위치, 유해조수와 카메라 사이의 거리, 인식한 이미지 이름으로 구성하였다.

그림 6을 보면 SFTP를 이용하여 호스트 컴퓨터로 전송한 이미지가 잘 전송되었는지 확인할 수 있다.

```
sftp> ls -al
-rw-***** 1 - - 457676 Sep 24 20:40 20240924_204048.jpg
-rw-***** 1 - - 468132 Sep 24 20:40 20240924_204053.jpg
-rw-***** 1 - - 431885 Sep 24 20:41 20240924_204059.jpg
-rw-***** 1 - - 427761 Sep 24 20:41 20240924_204104.jpg
-rw-***** 1 - - 461392 Sep 24 20:41 20240924_204109.jpg
-rw-***** 1 - - 467403 Sep 24 20:41 20240924_204114.jpg
-rw-***** 1 - - 460548 Sep 24 20:41 20240924_204120.jpg
-rw-***** 1 - - 456702 Sep 24 20:41 20240924_204125.jpg
-rw-***** 1 - - 428033 Sep 24 20:41 20240924_204130.jpg
-rw-***** 1 - - 439424 Sep 24 20:41 20240924_204135.jpg
-rw-***** 1 - - 428172 Sep 24 20:41 20240924_204141.jpg
-rw-***** 1 - - 457676 Oct 1 09:38 20241001_093845.jpg
-rw-***** 1 - - 462421 Oct 1 09:38 20241001_093851.jpg
-rw-***** 1 - - 468132 Oct 1 09:38 20241001_093856.jpg
-rw-***** 1 - - 435742 Oct 1 09:39 20241001_093902.jpg
```

그림 6. 전송한 이미지 확인  
Fig. 6. Verifying for Sent Images

## V. 결 론

본 논문에서는 매년 100억 원 이상의 농작물 피해를 발생시키는 유해조수에 의한 피해를 예방하기 위해 YOLO 모델을 이용하여 유해조수를 인식한 데이터 및 이미지를 SFTP를 활용하여 전송하는 방법에 대해 제안하였다.

여러 개의 장치에서 인식한 유해조수의 데이터를 하나의 장치로 전송시켜서 데이터베이스에 저장하고, 분류된 이미지 또한 효율적으로 관리할 수 있다. 이렇게 수집한 데이터를 이용해 유해조수가 자주 나타나는 시각과 위치 등을 분석한 데이터를 활용하여 출몰 빈도수가 잦은 구역에 유해조수를 쫓아내는 장치를 추가로 설치하고, 출몰 빈도수가 적은 구역엔 경계를 완화하는 등의 방법으로 유해조수에 의한 피해를 더 효율적으로 예방하는 효과를 얻을 수 있을 것으로 기대한다.

## Acknowledgements

“본 연구는 2024년 과학기술정보통신부 및 정보통신기획 평가원의 SW중심대학사업의 연구결과로 수행되었음”

(2019-0-01113)

## 참 고 문 헌

[1] “5년간 야생동물 농작물피해 542억원 ...”멧돼지 피해 가장 커, 연합뉴스, 2023년 9월 21일 수정, 2024년 10월 9일 접속, <https://www.yna.co.kr/>

view/AKR20230920120800003

[2] 김광수, ”오브젝트 스토리지를 이용한 소산 백업 성능 연구“ 숭실대학교 대학원, pp.7-8, 2021.

[3] 안재원, 최범진, 옥성진, 강정하, 김재영, 김은기 ”보안 기능을 지원하는 파일 전송 프로토콜의 설계 및 구현“ 한밭대학교 정보통신학과, pp. 30 86-3087, 2014.

[4] 김경륜, 김명섭, 김형중 ”EM 클러스터링을 이용한 SSH 트래픽 식별“ 한국통신학회논문지 12-1 2 vol.37B No.12, pp. 1161-1163, 2012.

[5] 정지인, 이재혁, 이경률 ”정적 분석 도구 성능 비교 및 분석: 오픈소스의 보안 취약점을 기반으로“ The Journal of Korean Institute of Communications and Information Sciences Vol.47 No.04, p p. 681. 2022.

[6] 조영래, 김홍준, 박병훈, 신수연, 이치훈 ”YOLO v8을 활용한 디지털 문서의 핵심 객체 추출 및 분류 시스템 설계“ ASK 2024 학술발표대회 논문집 31권 1호, pp. 597, 2024.

[7] 김은기, ”MongoDB와 MySQL 대용량 데이터 처리 비교를 통한 Nosql 활용 방안 연구.“ 국내석사학위논문 숭실대학교, pp. 24-26, 2016.



# 수소연료전지 기반 드론의 효율적 장시간 비행을 위한 최적화 연구

김정아\*, 이승학\*, 박성호\*, 천종훈\*\*, 홍성호\*\*\*

## A Study on the Optimization of Hydrogen Fuel Cell-Based Drones for Efficient Long Flight

Kim Jeong A\*, Rhee Seung Hak\*, Park Sung Ho\*, Chun Jong Hun\*\*, and Hong Seong Ho\*\*\*

### 요약

본 논문은 수소연료전지 중에서도 프로톤 교환막 연료전지(PEMFC)를 중심으로, 드론의 효율적인 장시간 비행을 위한 최적화 설계를 다룬다. PEMFC는 높은 에너지 밀도와 신속한 반응 특성을 통해 드론의 비행 시간을 크게 늘릴 수 있다. 본 연구는 연료전지 스택의 구조적 설계, 수소 저장 기술, 그리고 추진 시스템과 에너지 관리 방안의 통합적인 최적화를 목표로 한다. 이를 통해 드론의 비행 성능을 극대화하고, PEMFC 기반 드론의 실용성을 확인하는 실험 데이터를 제공한다.

### Abstract

This research focuses on optimizing long-duration flights of drones powered by Polymer Electrolyte Membrane Fuel Cells (PEMFC). With increasing demand for efficient aerial solutions in logistics, surveillance, and disaster response, PEMFC technology offers significant advantages due to its high power density and rapid response. The study explores critical aspects such as the design of PEMFC systems, various hydrogen storage methods, and the optimization of propulsion systems to enhance the drone's operational efficiency. By addressing the balance between weight, power management, and energy storage, this research aims to contribute to the development of sustainable hydrogen fuel cell technology in the drone industry, paving the way for advanced applications and improved flight performance.

### Key words

Hydrogen fuel cell, Hydrogen storage, UAV, Long-endurance drone, Energy efficiency

---

\* 전남도립대학교, otxf0@naver.com, stenly69@daum.net, shpark@dorip.ac.kr,

\*\* 전남도립대학교, jhchun@dorip.ac.kr,

\*\*\* 호그린에어, hogreen@hogreen.com

※ 본과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 첨단분야 혁신융합대학사업의 연구결과입니다.

## 1. 서론

수소연료전지를 적용한 드론은 기존의 배터리 기반 드론보다 장시간 비행이 가능하며, 특히 물류, 감시, 재난 구조 등에서의 활용도가 높다. 이 중 **\*\*PEMFC(Polymer Electrolyte Membrane Fuel Cell)\*\***는 높은 출력 밀도와 빠른 반응 속도로 주목받고 있다. 본 연구는 PEMFC를 기반으로 한 드론의 장시간 비행을 위한 최적화 방안을 심도 있게 탐구하며, 수소 저장 방식, 추진 시스템, 에너지 관리 시스템 등의 종합적 설계를 통해 효율성을 극대화하는 방법을 다룬다.

## II. 수소연료전지 기술

### 2.1 PEMFC의 원리 및 구조

PEMFC는 수소를 연료로 사용해 전기화학 반응을 통해 전력을 생성하는 방식이다. 수소 연료는 아노드에서 전자를 방출하며, 이 전자는 외부 회로를 통해 전력을 공급한 뒤, 캐소드에서 산소와 결합해 물을 생성한다. 이때 프로톤은 고분자 전해질막을 통해 아노드에서 캐소드로 이동하며, 이를 통해 지속적인 전기 발생이 가능하다.

PEMFC의 주요 구성 요소는 다음과 같다:

- 아노드(Anode): 수소가 산화하여 프로톤과 전자로 분리되는 반응이 일어나는 곳. 백금 기반 촉매가 사용되어 높은 반응성을 보장한다.
- 캐소드(Cathode): 산소가 수소에서 전달된 프로톤과 결합해 물을 생성하는 반응이 일어나는 곳. 여기에서도 백금 촉매가 사용되지만, 최근 비용 절감과 효율 향상을 위한 대체 촉매 연구가 활발히 진행 중이다.
- 고분자 전해질막(PEM): 프로톤만을 선택적으로 통과시키는 막으로, PEM의 성능이 연료전지의 전체 성능에 중요한 영향을 미친다. Nafion과 같은 소재가 일반적으로 사용된다.

PEMFC의 주요 장점은 다음과 같다:

- 높은 에너지 밀도: 리튬이온 배터리에 비해 같

은 무게 대비 더 많은 에너지를 저장할 수 있다.

- 빠른 반응 속도: PEMFC는 낮은 작동 온도 (60~80°C)에서 작동하므로 시동 시간이 짧고, 환경 변화에 대한 적응력이 뛰어나다.
- 친환경성: 수소와 산소의 결합으로 물만 생성되므로, 탄소 배출이 없다는 점에서 매우 환경 친화적이다.

그러나 PEMFC는 백금과 같은 고가의 촉매를 사용하기 때문에 초기 비용이 높다. 이를 해결하기 위해 촉매의 대체 물질 개발 및 촉매의 양을 최소화하는 연구가 진행 중이다.

### 2.2 PEMFC의 성능을 좌우하는 요소

PEMFC의 성능은 크게 전극 소재, 전해질막의 수분 관리, 반응물 전달 방식에 의해 결정된다. 특히 전해질막의 수분 유지는 프로톤의 이동을 촉진해 전기 전도도를 높이는 중요한 역할을 한다. 반면, 수분이 과도하게 공급되면 막이 젖어 효율이 감소하므로, 적정 수분 수준을 유지하는 것이 핵심이다.

연료전지의 효율성을 높이기 위해서는 수소의 반응성을 극대화해야 하며, 이를 위해 연료와 산소의 균형 있는 공급이 필수적이다. 연료전지 스택 내에서의 균일한 연료 공급을 위해서 다공성 소재와 정밀한 흐름 제어 시스템을 사용하는 방법이 효과적이다.

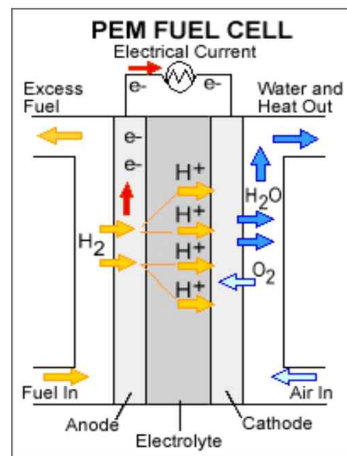


그림 1. 연료전지 스택 구성도

### 2.3 PEMFC 스택 설계

PEMFC 스택은 여러 개의 연료전지를 직렬 또는 병렬로 연결하여 원하는 출력 전압과 전력을 얻는 구조다. 각 연료전지는 얇은 판으로 분리되며, 이를 통해 경량화가 가능하다. 드론에 적용되는 PEMFC 스택은 다음과 같은 설계 목표를 달성해야 한다:

- 경량화: 비행체의 무게를 줄이기 위해 얇은 전극과 전해질막을 사용하며, 이를 통해 출력 밀도를 높이는 것이 중요하다.
- 냉각 시스템의 최소화: PEMFC는 낮은 온도에서 작동하기 때문에 추가적인 냉각 시스템이 불필요하거나 최소화할 수 있어 전체 시스템의 복잡성을 줄인다.

고분자 전해질막과 전극 간의 접촉 저항을 최소화하는 것이 성능을 높이는 데 중요한 역할을 한다. 이를 위해 고전도성의 얇은 전극 재료와 균일한 압력을 제공하는 프레임 설계가 요구된다.

## III. 수소 저장 방식

### 3.1 압축 수소 저장

가장 널리 사용되는 수소 저장 방식은 압축 수소 저장이다. 이는 복합재 소재로 제작된 고압 탱크에 수소를 350~700 bar 압력으로 압축해 저장하는 방식이다. 이 방식은 고에너지 밀도를 제공하지만, 압축 과정에서 에너지 손실이 발생할 수 있고, 안전 문제가 뒤따른다.

장점: 기술적 성숙도 및 경제성, 높은 에너지 밀도 제공.

단점: 고압 탱크의 복잡성 및 무게로 인해 드론의 비행 시간에 영향을 줄 수 있으며, 사고 발생 시 큰 위험성을 내포.

### 3.2 액체 수소 저장

액체 수소는 영하 253°C에서 저장되며, 압축 수소보다 더 높은 에너지 밀도를 제공한다. 이는 특히 장시간 비행이 필요한 드론 시스템에서 유리하지만, 저장과 관리가 까다롭고, 액화 과정에서 많은 에너

지가 소모된다.

장점: 부피 대비 높은 에너지 밀도.

단점: 매우 낮은 온도를 유지해야 하며, 탱크 설계가 복잡하고 무겁다.

항목	고압기체수소	액화수소
저장 밀도	33.3 kg/m <sup>3</sup>	71 kg/m <sup>3</sup>
운용 온도	상온	-253°C

### 3.3 금속 수소화물 저장

금속 수소화물은 금속이 수소를 흡수하여 저장하는 방식으로, 수소를 저압에서 안전하게 저장할 수 있다는 장점이 있다. 이는 고압 시스템보다 안전성이 높지만, 무게가 많이 나가고 저장 용량이 제한적이다.

장점: 안전성, 저압 수소 저장 가능.

단점: 금속의 무게로 인해 드론의 비행 성능에 영향을 미칠 수 있다.

### 3.4 나트륨 боро하이드라이드 기반 저장

나트륨 боро하이드라이드(NaBH<sub>4</sub>)를 이용한 저장 방식은 화학적으로 수소를 생성해내는 방식으로, 상대적으로 안전하고 경량화된 저장 시스템을 가능하게 한다. 나트륨 боро하이드라이드는 물과 반응해 수소를 발생시키므로, 복잡한 저장 탱크가 필요하지 않다.

장점: 안전성, 경량화.

단점: 수소 생성 후의 부산물 처리 문제 및 추가적인 화학 반응 장치의 필요성.

## IV. 추진 시스템 최적화

### 4.1 모터 및 프로펠러 설계

PEMFC 드론의 추진 시스템은 매우 중요한 역할을 한다. 드론의 에너지를 효율적으로 사용하는 것은 추진 시스템의 설계에 따라 크게 좌우된다. 일반적으로 \*\*브러시리스 모터(BLDC)\*\*가 드론의 추진 시스템에 사용되며, 이 모터는 전력 소비가 적고,

효율이 높다.

프로펠러 최적화: 드론의 비행 효율을 극대화하기 위해서는 프로펠러의 크기와 각도가 중요한 요소다. 프로펠러가 크면 추력이 증가하지만 전력 소모도 커진다. 반면, 작은 프로펠러는 효율이 떨어질 수 있으므로 기체의 중량과 비행 목적에 맞는 최적의 크기와 각도를 찾아야 한다.

#### 4.2 전력 관리 시스템

PEMFC는 일정한 전력을 생성하므로, 비행 중 요구되는 전력 변동에 대해 효율적인 전력 관리 시스템이 필수적이다. 이를 통해 추진 시스템에 공급되는 전력을 실시간으로 제어하고, 불필요한 에너지 낭비를 최소화한다. 특히, 배터리와 연료전지의 하이브리드 시스템을 구축해 전력 수요에 따라 유연하게 대응하는 것이 중요하다.

배터리 보조 시스템: 드론이 비행 중 특정 순간에 높은 출력을 요구하는 경우, 배터리를 보조 전력원으로 활용할 수 있다. 이러한 시스템은 연료전지가 일정 속도 이하로 작동할 때 더 높은 효율을 발휘하며, 전체 비행 시간을 늘릴 수 있다.

모니터링 및 제어 알고리즘: 연료전지와 배터리의 효율적인 관리를 위해 실시간 데이터 모니터링 및 예측 알고리즘을 사용할 수 있다. 이를 통해 비행 중 에너지 소비를 최적화하고, 연료전지의 수명과 성능을 극대화할 수 있다.

### V. 결 론

본 연구에서는 PEMFC를 기반으로 한 드론의 장시간 비행을 위한 최적화 방안을 심도 있게 분석하였다. PEMFC의 특성과 성능을 향상시키기 위한 다양한 기술적 접근이 요구되며, 이와 더불어 수소 저장 방식과 추진 시스템의 최적화가 필수적이다. 본 연구는 드론 산업에서의 지속 가능한 에너지 솔루션으로서 수소연료전지의 가능성을 제시하며, 향후 기술 발전과 상용화를 위한 기초 자료를 제공할 것이다.

### 참 고 문 헌

- [1] K. Jung et al., "PEM Fuel Cell System for Drone Applications: Optimization and Testing," *International Journal of Hydrogen Energy*, vol. 45, no. 10, pp. 512-523, 2022.
- [2] P. O'Hayre et al., "Fuel Cell Fundamentals," 4th Edition, John Wiley & Sons, 2016.
- [3] S. Park et al., "Advanced Hydrogen Storage Technologies for UAVs," *Energy Storage Materials*, vol. 32, pp. 125-135, 2021.
- [4] H. Kim et al., "Energy Management Strategies for Hydrogen Fuel Cell UAVs," *Journal of Aerospace Engineering*, vol. 28, no. 4, pp. 204-213, 2020.

# YOLOv5와 스테레오 카메라를 활용한 실시간 유해조류 탐지 및 거리 측정 시스템

최민정\*, 정윤걸\*\*, 이상훈\*\*\*, 이영학\*\*\*\*

## Real-Time Harmful Bird Detection And Distance Measurement System Using YOLOv5 and Stereo Camera

Min-Jeong Choi\*, Yun-Geol Jung\*\*, Sang-Hun Lee\*\*\*, Yeung\_Hak Lee\*\*\*\*

### 요 약

본 연구는 과수원과 같은 농가에서 유해조류로 인한 피해를 방지하기 위한 실시간 새 탐지 및 거리 측정 시스템을 제안한다. 임베디드 플랫폼인 Jetson Nano에서 YOLOv5 모델을 활용하고, Intel RealSense Depth Camera(D435)를 적용하여 스테레오 비전과 Focal length를 통해 유해조류의 위치 추적을 수행한다. 2D 카메라의 초점 거리 기반 거리 측정 방법과 스테레오 비전 방식을 비교하여 3m 이내 범위에서 정확한 거리 측정을 위한 최적화된 방안을 제시하였다. 실험 결과, 허용 가능한 오차율 내에서 최소한의 오차율을 보여주어, 과수원 환경에서 실시간 유해조류 방제의 효율성을 높일 수 있음을 확인하였다.

### Abstract

This study presents a real-time bird detection and distance measurement system aimed at mitigating harm caused by pest birds in agricultural environments. Utilizing the YOLOv5 model on an embedded platform (Jetson Nano) and leveraging the Intel RealSense Depth Camera (D435), this system classifies and tracks the spatial position of birds using stereo vision. By comparing the depth estimation accuracy between a 2D focal length-based method and the stereo vision approach, we propose an optimized solution for precise distance measurement within a range of 3 meters. Experimental results demonstrate minimal error rates within acceptable limits, enhancing the efficacy of real-time pest bird control in orchard settings.

### Key words

Streo Vision, Focal Length, YOLO,

### 1. 서 론

유해조류 방제는 환경 보전과 인류의 안전을 위

한 중요한 주제로, 최근 인공지능(AI) 기술의 발전에 따라 유해조류 탐지 및 대응 기술이 더욱 주목받고 있다. 매년 유해조류로 인한 피해가 지속되고

\* 국립안동대학교 컴퓨터공학과, chlalswj4@naver.com, \*\*국립안동대학교 컴퓨터공학과, ykjung0917@kakao.com, \*\*\* 구미전자정보기술원, legotec@geri.re.kr, \*\*\*\*국립안동대학교 SW융합교육원, yhyi@anu.ac.kr(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)

있으나, 이에 대한 해결 방안은 제한적이며, 유해조류가 지닌 높은 적응력과 퇴치 저항성으로 인해 효과적인 대책 마련이 어렵다. 이러한 문제를 해결하기 위해 AI 기반의 탐지 및 거리 측정 시스템을 활용한 유해조류 실시간 탐지 연구가 필요하다.

본 연구는 YOLOv5 모델을 사용하여 유해조류의 분류 및 거리 측정 시스템을 개발하고, 이를 임베디드 플랫폼인 Jetson Nano에서 실행할 수 있도록 최적화하고자 한다. 특히 Intel RealSense Depth Camera (D435)를 활용해 스테레오 비전을 통해 유해조류와의 거리를 보다 정확히 계산하며, 이를 통해 유해조류의 위치와 움직임을 실시간으로 파악할 수 있도록 한다. 2D 카메라에 기반한 focal length를 통해 거리를 계산하는 알고리즘과 스테레오 비전을 통해 사물 간의 거리를 계산하는 알고리즘의 정확도 차이를 비교한다. 이러한 시스템을 통해 유해조류 방제 작업의 정확성과 효율성을 높이고, 환경에 미치는 영향을 최소화하면서도 방제 효과를 극대화할 수 있는 솔루션을 제시하고자 한다.

## II. 임베디드 플랫폼 환경

본 연구에서 사용된 임베디드 모델은 Jetson Nano로 4-core Cortex-A78 64bit 모델을 기반으로 연구를 진행하였다. 자세한 사양은 아래의 Table 1과 같다.

Table 1. 임베디드 플랫폼 성능표

Machine	Jetson Nano(P3450)
CPU model	4-core ARM Cortex-A57
GPU	128-core Maxwell GPU
Memory	4G LPDDR4
Networking	Gigabit Ethernet
Storage	64G microSD
Power	5V 20W DC

## III. 구현 과정

### 3.1 YOLO 모델

2D 딥러닝 YOLO 알고리즘을 활용하여 YOLOv5의 Nano version을 활용한 YOLO모델을 개발하여

진행하였다. 학습 데이터로 활용한 데이터는 휴대폰, 카메라 그리고 동영상을 프레임 단위로 추출하여 annotation한 총 4,000장의 ‘새(Bird)’ 이미지 데이터를 활용하였다. Train, Test, 그리고 Value를 각각 8:1:1의 비율로 분류하여서 학습을 진행하였다.

또한 임베디드 플랫폼이라는 환경에서 가장 적합한 모델을 선정하고자 하였다. 그림 1은 YOLOv5 버전의 모델별 특징을 나타낸다.

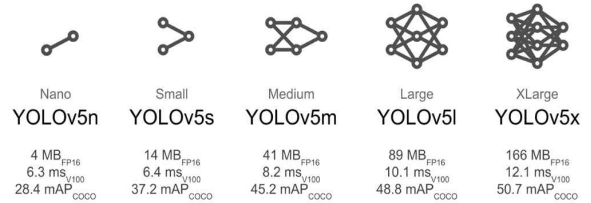


그림 1. YOLOv5 모델별 특징

임베디드 플랫폼이라는 한정된 자원 안에서는 경량화 모델이 가장 적합하다고 판단하여 YOLOv5(nano) 모델로 실험을 진행하였다. 이는 동일 프레임 기준 적은 용량을 차지하고, 속도가 빠르다는 특징이 있기 때문에, 본 연구의 환경인 임베디드 플랫폼에서 가장 적합한 모델이라 판단하였다.

### 3.2 스테레오 비전 실험 환경

본 연구에서는 Intel RealSense depth camera의 스테레오 비전과, 2D camera의 focal length를 통한 객체와의 측정 거리 정확성을 도출하고자 하였다. 이를 위해 그림 2와 같은 환경을 구성하였으며, 각 미터 별로 지점을 설정하여 새 이미지를 두고, 이에 따른 측정된 거리의 정확성을 오차율을 사용하여 구하였다. Intel RealSense depth camera의 최대 측정 거리는 약 10m로, 본 연구에서 활용한 D435의 경우에는 약 3m 범위까지 측정이 가능하다. 그렇기에 본 연구에서는 측정 범위를 3m 이내로 조정하여서 진행하였다.

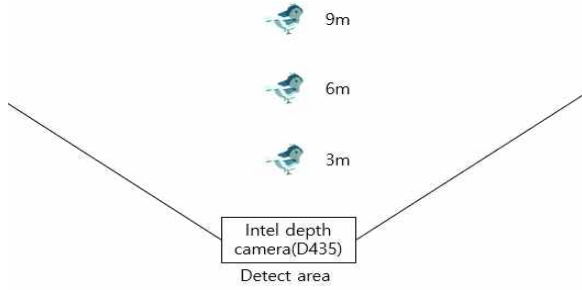


그림 2. 실험 환경

먼저 YOLOv5 모델을 활용해 새를 검출하면, 검출된 새의 바운딩 박스 좌표가 반환된다. 바운딩 박스는 (xmin, ymin, xmax, ymax) 형태로 주어지며, 각각 좌상단과 우하단 모서리의 좌표를 나타낸다. 이후 객체의 중심 좌표를 계산하고, 이를 위해 식(1)과 식(2)를 사용한다.[2]

$$A_{senter} = \frac{A_{min} + A_{max}}{2} \quad (1)$$

$$B_{senter} = \frac{B_{min} + B_{max}}{2} \quad (2)$$

Intel RealSense Depth Camera는 스테레오 비전 기술을 활용하여 각 픽셀의 깊이 값을 제공한다. 스테레오 비전 기술이란, 두 개의 카메라가 서로 다른 각도에서 동일한 장면을 촬영하고, 이로부터 생기는 차이를 바탕으로 각 픽셀의 거리를 계산하는 방식이다. Intel RealSense Depth Camera는 두 개의 적외선(IR) 카메라로 구성되며, 일정한 간격(Base line)으로 배치된 두 카메라로부터 깊이 정보를 얻는다.

$$Z = \frac{f \times B}{d} \quad (3)$$

- $Z$  : 깊이 값(카메라와 픽셀 간의 거리)
- $f$  : 카메라의 초점 거리(Focal length)
- $B$  : 두 카메라 사이의 거리(Base line)
- $d$  : 시차(Disparity)

이후 식(3)을 통해 시차로부터 계산된 깊이 값을 실제 거리로 반환하여, 각 픽셀에 대해 미터 단위의 깊이 값을 구하였다.

### 3.3 Focal length 실험 환경

Focal length는 렌즈나 카메라의 중심과 이미지가 맺히는 지점 사이의 거리로, 그림 3과 같이 두 변을 공유하는 직각삼각형은 닮음 상태를 가진다. 즉, 각 삼각형의 비율이 동일하다는 것이다. 이를 통해 실제 크기, 초점 거리, 픽셀 높이를 알면 실제 거리를 구하는게 가능하다.

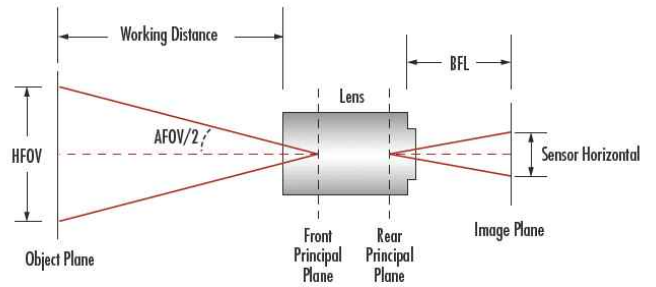


그림 3. Focal length

그림 3과 같이 두 변을 공유하는 두 직각삼각형은 맞꼭지각을 형성하여 닮음 상태를 취하고 있음을 알 수 있다. 이는 각 삼각형의 길이 비율이 동일함을 의미하며, Real Height(실제 크기), Focal length(초점 거리), 그리고 Pixel Height(픽셀 높이)를 알면 Real Distance(실제 거리)를 구할 수 있다는 것을 말한다. 이를 의미하는 것이 식(3)과 식(4)이다.

$$\frac{Real\ Height}{Real\ Distance} = \frac{Pixel\ Height}{Focal\ length} \quad (4)$$

$$Real\ Distance = \frac{Focal\ length \times Real\ Distance}{Pixel\ Height} \quad (5)$$

- $Real\ Height$  : 실제 크기
- $Focal\ Length$  : 카메라의 초점 거리
- $Pixel\ Height$  : 픽셀 높이
- $Real\ Distance$  : 실제 거리

## IV. 실험

본 논문에서는 그림 2와 같이 1m, 2m, 3m 간격으로 새를 검출하고, 이에 따른 거리의 정확도를 Intel RealSense Depth camera와 2D 카메라 별로 측정

하여 비교하였다.

Table 2. Intel RealSense Depth camera - Depth 결과

Id	1m	2m	3m
1	1.1	2	3
2	1	2	3
3	1	2.1	3.1
4	1	2	3.1
5	1	2	3
Average	1.02	2.02	3.04
Result	-0.02	-0.02	-0.04

Table 2. 2D camera - Focal length 결과

Id	1m	2m	3m
1	1	2	3
2	1.1	2.1	3.1
3	1	2	3
4	1	2	3
5	1	2	3
Average	1.02	2.02	3.02
Result	-0.02	-0.02	-0.02

## VI. 결론

본 연구에서는 Jetson Nano라는 임베디드 플랫폼에서 YOLOv5(nano) 모델을 활용하여 새 검출 모델을 개발하고, 이를 Intel RealSense Depth camera와 2D 카메라를 활용하여 깊이 기반의 거리 측정 방식과 focal length 기반의 거리 측정 방식 간의 차이를 비교 분석한 것이다. 실험을 통해 제안한 각 알고리즘의 성능을 측정된 결과, 약 0.4m 이내의 오차를 보였다. 두 거리 측정 값 모두 극미한 오차를 보였으며, 이는 허용 가능한 오차 범위 내에서 보다 정확히 객체 사이의 거리 측정이 가능함을 의미한다. 이러한 연구 결과는 Intel RealSense Depth Camera(D435)가 3m 이내에서 가장 정확한 측정이 가능하다는 특성을 가지기 때문이라고 유추되는데, 빛이나 감도 등의 다양한 환경 조건을 고려하지 않았다는 점도 큰 요인으로 작용했을 것이라고 생각된다.

본 연구는 과수원과 같은 농가에서 유해조류 방조를 위한 대안으로서 진행되었다. 유해조류 방조에는 실시간 객체 탐지와 거리 측정이 필요한데, 이를

위한 다양한 알고리즘 중 Depth(깊이)를 활용한 알고리즘과, Focal length를 활용한 알고리즘의 성능을 비교하여 유해조류 방조 시스템 설계에 사용할 수 있는 거리 측정 알고리즘 간의 차이점을 나타내었다.

향후 연구로는 다양한 환경적 요인을 고려하고, 보다 더 넓은 범위를 측정하여 종합적인 성능과 정확도를 파악하고자 한다. 이를 통해 실제 과수원이라는 배경적 요건을 고려하고, 연구 결과의 오차를 줄여나가도록 알고리즘을 발전시키고자 한다.

## 참 고 문 헌

- [1] 송장훈, 이한수, 양창열, 이두표, “배 과수원의 조류피해 실태와 봄철 집단 포살이 까치 밀도에 미치는 영향”, 한국조류학회지, Vol.8, No.2, pp. 117-126, 2001.
- [2] 장시웅, 정동훈, “레이더 센서를 이용한 거리 측정 시스템의 설계 및 구현”, 한국정보통신학회 논문지, Vol.22, No.7, pp. 1009-1014, 2018.



# YOLOv8을 활용한 유해 조류 감지 및 추적 연구

김민지\*, 이영학\*\*

## A Study on Detecting and Tracking Harmful Birds Using YOLOv8

Min-Ji Kim\* and Yeung-hak Lee\*\*

### 요약

이 연구는 YOLOv8 모델을 활용하여 유해 조류를 감지하고 추적하는 방법에 대한 연구이다. 유해 조류는 농작물과 환경에 피해를 주는 주요 요인 중 하나로, 효과적인 감지와 추적이 필요하다. 본 연구에서는 객체 추적 기술 중 하나인 Deep SORT를 활용하여, 유해 조류를 정확하게 감지하고 추적하는 방법을 제안함으로써 유해 조류 감지 및 추적의 정확도와 효율성을 향상시킨다.

### Abstract

This study explores a method for detecting and tracking harmful birds using the YOLOv8 model. Harmful birds are a major factor causing damage to crops and the environment, making effective detection and tracking necessary. In this study, we propose a method that leverages Deep SORT, an object tracking technique, to accurately detect and track harmful birds, thereby improving the accuracy and efficiency of harmful bird detection and tracking.

### Key words

Harmful Bird Detection, YOLOv8, Object Tracking, Computer Vision, Deep Learning, Crop Protection

## I. 서론

유해 조류는 농작물과 환경에 심각한 피해를 유발하는 주요 요인 중 하나로, 이에 대한 효과적인 감지와 추적은 매우 중요하다. 유해 조류가 미치는 피해는 농작물의 생산성을 저해할 뿐만 아니라, 생태계의 균형에도 악영향을 미친다. 특히, 농업 지역에서 유해 조류를 감지하고 관리하는 것은 경제적 손실을 줄이고 생태적 보호를 위해 필수적이다.

기존의 유해 조류 감지 방법들은 주로 단일 객체 감지에 집중하였으며, 추적과 관련된 연구는 제한적이었다. 이러한 방법들은 유해 조류의 이동 경로를 실시간으로 파악하는 데 한계가 있으며, 감지의 정확도와 추적의 효율성을 저해하는 문제점을 가지고 있다. 최근 인공지능 기술의 발전과 함께, 객체 감지 및 추적에 대한 다양한 연구가 진행되었고, 특히 YOLO(You Only Look Once)와 같은 실시간 객체 감지 모델이 주목받고 있다. 그러나 기존의 YOLO

\* 국립안동대학교 컴퓨터공학과, ddaa637878@gmail.com,

\*\* 국립안동대학교 SW융합교육원, yhyi@anu.ac.kr(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)

모델 기반 감지 기법들은 추적 기능과의 결합에 있어서 성능의 한계가 존재하였다.

본 연구에서는 최신 YOLOv8 모델과 객체 추적 기술 중 하나인 [1]Deep SORT를 결합하여 유해 조류를 감지하고 추적하는 방법을 제안한다. YOLOv8 모델은 높은 정확도와 빠른 속도를 제공하며, Deep SORT는 감지된 객체를 지속적으로 추적하는 데 탁월한 성능을 보여준다. 이를 통해, 유해 조류의 실시간 감지 및 추적의 정확도를 향상시키고, 농업 및 환경 보호에 기여하고자 한다.

## II. 연구방법

### 2.1 데이터셋

본 연구에서는 유해 조류의 감지를 위해 자체 구축한 데이터셋을 사용하였다. 데이터셋은 총 2000개의 이미지로 구성되어 있으며, 각 이미지에는 유해 조류의 위치를 나타내는 라벨이 포함되어 있다. 데이터 전처리 과정에서는 이미지 리사이징과 데이터 증강 기법을 적용하여 모델의 일반화 성능을 향상시켰다.

### 2.2 YOLOv8 모델 구성 및 학습

YOLOv8 모델은 YOLO 시리즈의 최신 버전으로, 본 연구에서는 높은 성능과 효율성을 고려하여 YOLOv8n 모델을 사용하였다. 모델 학습을 위해 배치 크기 16, 학습률 0.001, 에포크 수 100회의 설정으로 진행하였으며, 학습 데이터와 검증 데이터를 8:2의 비율로 분할하여 모델을 훈련하였다.[2]

### 2.3 객체 추적 알고리즘: Deep SORT[1]

유해 조류의 이동 경로를 실시간으로 추적하기 위해 객체 추적 알고리즘인 Deep SORT를 적용하였다. Deep SORT는 감지된 객체의 특성을 기반으로 연속된 프레임에서 동일 객체를 식별하며, 이를 통해 조류의 이동 경로를 추적한다. 이 알고리즘을 통해 실시간 감지 및 추적의 효율성을 높일 수 있었다.

## III. 결과 및 분석

본 연구에서는 YOLOv8 모델을 사용하여 유해 조류를 감지하고, Deep SORT를 활용하여 유해 조류의 실시간 추적을 수행하였다. 이를 통해 감지된 조류의 수, 위치 및 이동 경로를 파악할 수 있었다. 그림 1과 그림 2는 Deep SORT 적용 전과 감지 후의 결과 이미지 비교를 나타낸다.



그림 1. 기본 실험 후 유해 조류 감지 결과  
Fig. 1. Detection Result of Harmful Birds After Baseline Experiment



그림 2. Deep SORT 적용 후 유해 조류 감지 결과  
Fig. 2. Detection Result of Harmful Birds After Applying Deep SORT

### 3.1 유해 조류 감지 결과

그림 2에서 볼 수 있듯이, YOLOv8 모델은 조류의 정확한 위치를 감지하고, 각 객체에 대해 신뢰도(confidence score)를 부여하였다. 감지된 조류의 경계 상자에는 “bird” 라벨과 함께 각 조류의 confidence score가 표시되었다. 이 결과는 모델이 설정된 confidence threshold(신뢰도 임계값)를 기준으로 조류를 감지하고 있음을 의미한다.

### 3.2 추적 결과 분석

YOLOv8 모델과 Deep SORT를 결합함으로써, 감지된 조류의 위치와 이동 경로를 실시간으로 추적할 수 있었다. 이는 특히 농작물 보호를 위해 조류의 비행 패턴을 분석하고, 조류의 출현 빈도가 높은 구역을 파악하는 데에 유용하다. 본 연구에서는 실험을 통해 YOLOv8 모델이 다양한 환경에서 유해 조류를 정확하게 감지하고, 추적 기법을 통해 이동 경로를 지속적으로 모니터링할 수 있음을 확인하였다.

### 참 고 문 헌

- [1] “동영상 분석을 위한 객체 추적 기술“, <https://blog-ko.superb-ai.com/object-tracking-technology-for-video-analysis/>, (Accessed on: Octpber 30, 2024).
- [2] "Yolov5\_DeepSort\_Pytorch", Available: [https://github.com/gongliyu/Yolov5\\_DeepSort\\_Pytorch](https://github.com/gongliyu/Yolov5_DeepSort_Pytorch), (Accessed on: Octpber 30, 2024).

# 무대 조명을 위한 스포트라이트 드론 제작

김태현\*, 김정아\*, 천종훈\*, 박성호\*, 이승학\*\*, 양승호\*\*\*

## Manufacture of Spotlight Drone for Stage Lighting

Kim Tae Hyun\*, Kim Jeong A\*, Chun Jong Hun\*, Park Sung Ho\*, Rhee Seung Hak\*\*,  
and Yang Seung-Ho\*\*\*

### 요약

최근 드론 시장의 확대로 다양한 산업군에 드론 응용 기술이 접목되고 있으며, 역동적인 연출을 위한 다각도 조명의 활용기술 개발도 지속적으로 개발이 요구되고 있다. 본 논문에서는 드론 호버링 시스템 및 AI 비행 알고리즘을 활용하여 기존의 고정된 조명 문제와 공간적 제약을 극복한 엔터테인먼트용 스포트라이트 드론을 제작하고자 한다.

### Abstract

Recently, expansion of the drone market, drone application technology is being applied to various industries, and the development of technology utilizing multifaceted lighting for dynamic production is also required. In this paper, we intend to produce a spot light drone for entertainment that overcomes the existing fixed lighting problems and spatial constraints by using the drone hovering system and AI flight algorithm.

### Key words

drone, hovering system, AI flight algorithm, spot light, entertainment

## I. 서론

최근 드론 시장의 확대로 다양한 산업군에 드론 응용 기술이 접목되고 있으며, 평창 올림픽 및 다양한 드론 엔터테인먼트에서 조명을 설치한 AI 알고리즘 기반 드론[1]이 활용되고 있다. 연극, 콘서트 등 무대를 더욱 화려하게 꾸미는 기술은 계속해서 발전해 왔으며, 역동적인 연출을 위한 다각도 조명

의 활용기술 개발도 지속적으로 개발이 요구되고 있다.

본 논문에서는 드론 호버링 시스템[2] 및 AI 비행 알고리즘을 활용하여 기존의 고정된 조명 문제와 공간적 제약을 극복한 엔터테인먼트용 스포트라이트 드론을 제작하고자 한다.

## II. 스포트라이트 드론 설계 및 제작

\* 전남도립대학교, thkim4852@naver.com, otxf0@naver.com, jhchun@dorip.ac.kr, shpark@dorip.ac.kr,

\*\* 전남도립대학교, stenly69@daum.net,

\*\*\* (주)유니컴퍼니, help1001@unicompany.co.kr

\* 본과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 첨단분야 혁신융합대학사업의 연구결과입니다.

설계한 드론용 스팟라이트 장치는 고휘도 LED 등의 특수 조명 모듈, 조명의 밝기, 색상, 각도를 실시간으로 제어하기 위한 제어 알고리즘, 조명 모듈의 발열 문제를 해결하기 위한 냉각 시스템으로 구성하였다.

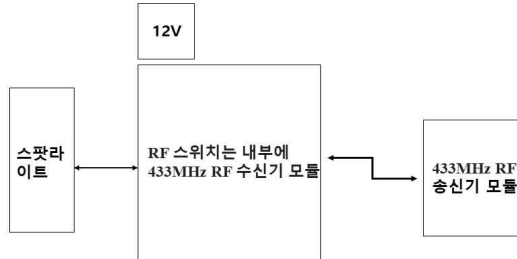


그림 1. 드론용 스팟라이트 장치 구성  
Fig. 1. Drones Spotlight Device Configuration

다음으로 설계한 스팟라이트 장치를 장착한 드론을 설계하기 위해 다음 같은 기술을 적용하였다.

첫째, 호버링 시스템의 불안정성 문제를 해결하기 위해 GPS, IMU, LiDAR 등을 활용한 다중 센서 융합 기술을 적용[2]하여 다양한 센서 데이터를 통합하고, 드론의 안정적인 비행을 위해 PID 및 SLAM 등을 활용한 고급 제어 알고리즘을 사용해 호버링 안정성을 확보하였다.

둘째, 배터리 소모 문제로 인한 실시간 비행 및 조명 연출 제한 문제를 해결하기 위해 에지 컴퓨팅 기술을 적용[3]하여 데이터 처리 효율성을 높이고 [1], 에너지 효율적인 부품[4]사용하여 배터리 소모를 줄였으며, 드론과 조명 시스템에 전력을 효율적으로 공급하는 전력 분배 시스템의 설계, 배터리 용량을 늘리거나 교체 가능한 배터리 시스템을 도입하였다.

셋째, 비상 상황에서의 드론 사고 위험을 방지하기 위해 비상 착륙 시스템과 충돌 회피 기술을 추가하여 안전 프로토콜 제작하였다.

넷째, 사용자 인터페이스의 복잡성을 줄이기 위해 관련 UI/UX 연구 결과분석[5] 사용자 테스트를 기반으로 직관적이고 사용하기 쉬운 인터페이스를 설계하였으며, 드론의 상태, 위치, 조명 상태 등을 실시간으로 모니터링할 수 있는 기능을 구현하였다.

다섯째, 실시간 제어를 위해 RF신호를 이용하여 드론과 지상 제어 장치 간의 실시간 데이터 송수신을 위한 통신 프로토콜을 포함한 원격제어장치를 설계하였다.

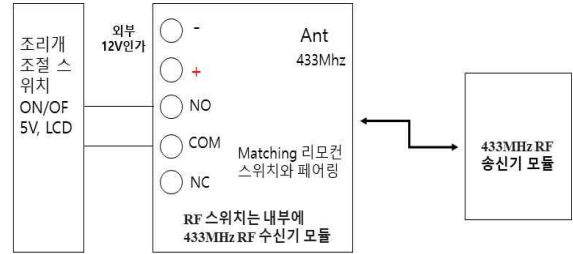


그림 2. 제어 화면 설계  
Fig. 2. Control Screen Design



그림 3. RF신호를 이용한 실시간 원격제어장치  
Fig. 3. Real-time remote controller using RF signal

### III. 결 론

본 논문에서는 드론 호버링 시스템 및 AI 비행 알고리즘을 활용하여 엔터테인먼트용 스팟라이트 드론을 제작하였다. 제작한 드론은 다양한 조명 효과 제공, 실시간 제어기능, 에너지 절감 기능 등을 통해 라이브 행사, 영화 촬영, 도시 장식, 공사 현장 조명, 긴급 구조 작업 등에 활용할 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] 박민수, 최정훈, “AI 비행 알고리즘의 실시간 처리 성능과 에너지 효율성”, IEEE Transactions on Aerospace and Electronic Systems, 2021
- [2] 김철수, 이영희, “드론 호버링 시스템의 안정성 확보와 다중 센서 융합 기술 적용에 관한 연구”, 한국항공우주학회 논문집, 2022
- [3] 최현우, 김소영, “에지 컴퓨팅 기술의 적용과 실시간 데이터 처리”, 한국정보과학회, 2023
- [4] 장민호, 이수정, “지속 가능한 발전을 위한 친환경 소재 사용 및 에너지 효율성 개선 방안”, Journal of Sustainable Development, 2021
- [5] 정유진, 김지훈, “UI/UX 디자인과 사용자 경험 최적화”, 디자인학연구, 2020

# 컴퓨터 비전을 활용한 디스플레이 외관 결함 검출 시스템

손혜원\*, 권동욱\*, 서나라\*, 임진수\*, 이원일\*\*

## Display Surface Defect Detection System Using Computer Vision

Hye Won Son\*, Dong Uk Gwon\*, Na Ra Seo\*, Jin Su Lim\* and Won Il Lee\*\*

### 요 약

설비의 오류, 작업자의 부주의, 제조 환경 관리 미흡 등은 스크래치, 얼룩 또는 패널 내 이물과 같은 디스플레이의 외관 결함을 유발할 수 있다. 그러므로 품질 보장과 고객 만족을 위해 결함 검출 시스템이 필요하다. 초기의 디스플레이 외관 결함 검출 방법들은 육안 검사나 비싼 검출 센서를 사용하여 인적, 비용적 문제가 있었다. 이 논문에서는 간단한 H/W 장비를 활용하여 디스플레이 외관 결함을 검출하고 결함을 구분하는 시스템을 제안하여 결함 검출의 효율을 향상시킨다.

### Abstract

Errors, physical damage, etc. during the display manufacturing process can cause defects. Therefore, a defect detection system is required for quality assurance and customer satisfaction. Early display appearance defect detection methods had human and cost problems using visual inspection or expensive detection sensors. In this paper, we improve the efficiency of defect detection by proposing a system that detects display appearance defects and distinguishes defects using simple H/W equipment.

### Key words

display, defect detection, surface defect, automation, computer vision

## I. 서 론

현대 사람들은 TV, 태블릿 PC, 스마트폰 등 디스플레이가 탑재된 전자제품을 많이 사용하고, 사용 추세 또한 증가하고 있다[1-2]. 따라서 디스플레이는 많은 사용자에게 긍정적인 경험을 주기 위한 중

요한 부품이다.

하지만 제조 과정에서 생긴 디스플레이의 외관 결함으로 사용자에게 부정적 경험을 줄 수 있으며, 외관 결함이 반사나 반응성 문제를 초래할 수 있다. 이러한 문제들은 디스플레이의 성능과 제품의 품질에 대한 심각한 문제로 여겨질 수 있다.

\* 국립금오공과대학교 전자공학부(반도체시스템전공), hyewon.135@daum.net, ehddnr0116@naver.com, snls369527@naver.com, jinsu2011@naver.com

\*\* 국립금오공과대학교 전자공학부(반도체시스템전공), wilee@kumoh.ac.kr (교신저자)

※ 이 연구는 정부(교육부-산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0022161, 반도체전공트랙사업)

종래 결함 검사 방법은 작업자가 맨눈으로 검사하거나 비싼 결함 탐지 센서를 이용하여 결함을 검출했다. 하지만 작업자별 검출 기준이 다르거나, 작업자의 육체적 정신적 상태가 검사 결과에 영향을 미치는 인적 문제와 비용 문제가 있다[3-4]. 이러한 문제를 해결하고 결함을 검출하고자 디스플레이 외관 결함 검출 시스템을 구성하려고 한다.

결함 검출의 수행 시간과 성능은 생산 효율 및 제품 품질과 직결되기 때문에 개별 공정 환경에 맞는 적합한 검출 알고리즘이 필요하며 [5], 이를 위한 다양한 결함 검출 알고리즘이 제안되었다[6-8]. [6]에서는 이미지 전처리 후 결정 트리 휴리스틱 알고리즘을 적용하여 결함을 분류하고, [7]에서는 템플릿을 사용하여 이미지를 정렬하며, [8]에서는 합성곱 신경망 기반 객체 검출 알고리즘 중 하나인 YOLOv4를 활용하여 결함 영역을 검출하는 기법을 제안하였다.

본 논문에서는 컴퓨터 비전을 활용한 제품의 외관 결함 검출 방식으로 검사 대상의 길이, 구성요소, 순서, 라벨 유무 등의 다양한 매개변수를 분석하여 결함을 식별한다. 이를 위해 자동화된 광학 검사 시스템은 에지 검출 기술을 이용해 제품의 윤곽선을 추출하고, 윤곽선의 길이 정보를 기반으로 결함 유무를 판단한다. 또한, 부품의 위치를 정확히 파악하여 결함이 발생할 수 있는 위치도 함께 분석하도록 구현한다.

## II. 시스템 하드웨어

본 시스템에서는 젯슨 나노, YUROBOT 사의 LED 조명, ArduCAM 사의 18MP AR1820HS 카메라, LCD Display, 컨베이어 벨트를 사용한다. 또한, 실제 결함 있는 디스플레이를 실험에 사용하기 한계가 있으므로 아크릴판을 사용한다.

젯슨 나노는 리눅스 환경을 지원하며 높은 GPU 성능으로 영상 처리 라이브러리를 활용하기 좋은 장점이 있으므로 선정하였다. YUROBOT 사의 LED 조명을 선택한 이유는 하나의 조명 모듈이 8개의 다이오드로 구성되어 있어 공간 활용을 최적화할 수 있으며 GPIO로 제어에 용이성이 있기 때문이다.

디스플레이의 작은 미세 결함까지 관측하기 위해 높은 해상도의 18MP 카메라를 선정하였으며 AR1820HS 카메라는 수동으로 초점 조절이 가능하므로 시인성을 확보하기에 수월하여 선정하였다. 또한, HMI 프로그램으로 시스템 동작을 제어하기 위해 10.1인치 크기의 Display를 사용한다.

## III. 시스템 소프트웨어

본 논문에서는 결함 검출 알고리즘과 HMI를 구현하였다. 결함 검출 알고리즘으로는 디스플레이 영역 검출, 이미지 전처리, 결함 검출 및 구분으로 나뉜다. 우선 본 시스템에서 정의한 결함은 표 1로 정리하였다.

표 1. 결함 정의  
Table 1. Definition of Defects

결함	정의
얼룩	디스플레이에 남은 크고 작은 자국
스크래치	1cm 이상이며 가로, 세로로 굽힌 자국
이물	지름 3mm 미만 이물질

먼저 카메라로 얻은 이미지에서 디스플레이의 외곽선을 얻기 위해 회색조 이미지로 변환한다. 그리고 적절한 이진화를 통해 디스플레이 영역만 강조한 뒤 Canny 에지 검출 알고리즘과 HoughTransform을 거쳐 외곽선의 정보를 얻는다. 마지막으로 회색조 이미지에서 해당 정보를 활용하여 원근 변환을 수행하고, 디스플레이 영역만 세로로 변환한 이미지를 획득한다. 그림 1에 각 과정의 결과 이미지를 나타내었다.

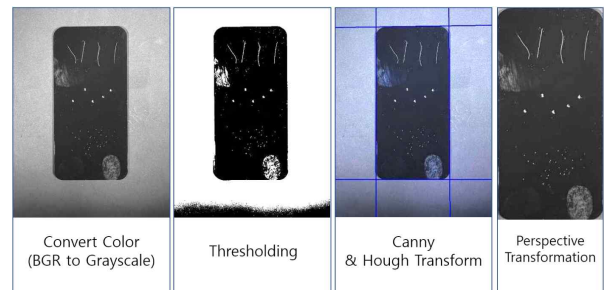


그림 1. 디스플레이 영역 검출 과정  
Fig 1. Display Area Detection Process

이미지 전처리 과정으로는 모폴로지 닫기 연산을 통해 지문이 강조된 이미지를 얻어 지문 결함의 영역 중 끊긴 영역이 먼저 결함으로 검출되는 것을 방지하고 온전히 지문 결함으로 검출되도록 한다. 또한 그 과정에서 결함이 아닌 노이즈를 제거하기 위해 양방향 필터링을 수행하며, 외곽선을 훼손하지 않도록 한다.

최종적으로 결함 검출 및 구분을 하기 위해 이미지에 connectedComponentsWithStats 메서드를 적용하여 이미지의 구성요소를 추출하고 동시에 x 좌표, y 좌표, 넓이, 높이 정보를 얻는다. 그리고 해당 정보를 활용하여 크기와 중형비를 구해 결함을 구분한다. 또한, 이진화 이미지에 검출된 결함을 색상으로 구분해 바운딩 박스로 표시한다. 디스플레이 영역을 검출한 회색조 원본 이미지와 결함 검출된 이진화 이미지를 그림 2에 나타내었다.

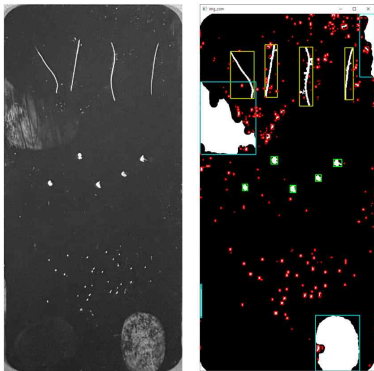


그림 2. 회색조 이미지와 결함 검출 이미지  
Fig 2. A gray tone image and defect detection image

HMI는 그림 3과 같이 QtDesigner 프로그램으로 디자인하고 PyQt 라이브러리로 설계하였다. 왼쪽 위부터 검사 시작 버튼, 카메라, 검사 모드, 로그 파일 저장 여부, LED를 설정하는 라디오 버튼과 수동 모드 시 벨트 조작을 위한 버튼, 디스플레이 모델 선택 Combo Box가 있다. 또한, 카메라 화면과 결함 검출 이미지를 볼 수 있는 부분, 그리고 아래의 TextView에 시스템 진행 로그가 출력되고 오른쪽 TextView에 한 디스플레이당 검사 결과가 출력된다.

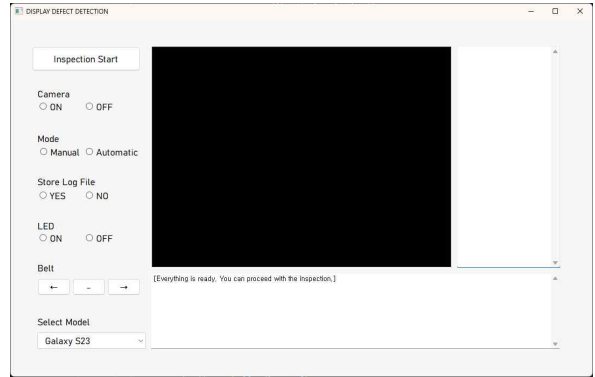


그림 3. HMI 디자인

Fig 3. HMI Design

#### IV. 시스템 동작

카메라를 켜고 검사 모드와 로그 파일 저장 여부를 설정한 다음 Inspection Start 버튼을 눌러 검사를 시작한다. 검사 모드는 수동과 자동으로 벨트 동작과 LED 설정을 결정한다. 수동으로 설정하여 벨트와 LED를 작업자가 직접 움직인 후 검사를 시작하면 해당 카메라 화면을 저장하여 전처리 과정과 결함을 검출하고 자동으로 설정하면 LED를 켜 상태에서 벨트가 움직이며 디스플레이가 완전히 카메라에 보일 때 화면을 저장하여 검사를 진행한다.

한 디스플레이의 검사를 마친 뒤 로그 파일 저장이 설정되어 있다면 시스템 진행 로그와 결과 로그가 하나의 파일로 저장된다.

#### V. 결론

본 논문에서는 디스플레이의 외관 결함을 검출하기 위한 자동 검사 시스템을 제안하였다. 아크릴판에 임의로 제작한 결함을 검출하고 3가지로 분류하는 데 성공하였다. 결론적으로 제안된 방법은 기존의 육안으로 하는 검사 방법보다 더 일관성 있으며, 비싼 결함 탐지 센서를 사용하는 방법보다 비용적으로 우수한 검출 결과를 보여준다.

제안한 방법을 실제 제조 현장에 적용함으로써 실시간 자동 검사를 통해 이상 원인을 찾고 제거하여 공정의 안정성을 높일 수 있을 것으로 기대된다. 또한 제안한 시스템은 기존의 관측된 유형의 화질 결함 검출에는 쉽지만, 새로운 유형의 결함 검출에



는 적용되기가 어려운 단점이 있어 딥러닝을 활용하여 검출하기 힘든 새로운 유형의 화질 결함이나 디스플레이의 외관 결함까지 자동 검출할 수 있는 시스템을 구현하는 연구가 추가로 필요하다.

of Intelligent Systems, vol. 32, no. 4, pp. 275-285, 2022.

## 참 고 문 헌

- [1] Gallup Korea, "25th National Assembly Election Public Opinion Poll", Gallup Korea, <https://www.gallup.co.kr/gallupdb/reportContent.asp?seqNo=1215>, October 27, 2023.
- [2] Gallup Korea, "Public Opinion on Social Issues", Gallup Korea,, <https://www.gallup.co.kr/gallupdb/reportContent.asp?seqNo=1497>, October 27, 2023.
- [3] G. H. Jo, Y. W. Heo, H. W. Cho, Y. J. Song, "Implementation of OLED Display Defect Detection System using CNN," Journal of KIIT., Vol. 20, No. 6, pp. 1-8, 2022 (in Korean).
- [4] C. Oh, Y. K. Ryu, S. J. Shin, H. J. Yoo, B. U. Jun, "An Algorithm Development for Detecting Blister Defects of Display Glasses," Journal of the Korean Society for Precision Engineering, Vol. 24, No. 4, pp. 15-20, 2007 (in Korean).
- [5] J. H. Kim, T. Y. Lee, Y. H. Ko, "Self-Reference PCSR-G Method for Detecting Defect of Flat Panel Display," Journal of Korea Multimedia Society, Vol. 18, No. 3, pp. 312-322, 2015 (in Korean).
- [6] A. Cubero-Fernandez, F.J. Rodriguez-Lozano, R. Villatoro, et al., "Efficient pavement crack detection and classification," J. Image Video Proc., vol. 39, 2017.
- [7] C. Jian, J. Gao, Y. Ao, "Automatic surface defect detection for mobile phone screen glass based on machine vision," Appl. Soft Comput., vol. 52, pp. 348-358, 2017.
- [8] S.-H. Lee, S.-H. Kang, Y.-S. Shin, O.-K. Choi, S. Kim, and J.-M. Kang, "YOLO-Based Detection of Metal Surface Defects," Journal of Korean Institute

# 운송조건 자율판단 시스템이 탑재된 고중량(100kg) 물류 운송 드론 개발

박홍준\*, 김정아\*, 이승학\*, 박성호\*, 천종훈\*\*, 최흥기\*\*\*

## Development of a heavy-duty (100kg) logistics transport drone equipped with an autonomous transport condition judgment system

Park Hong Jun\*, Kim Jeong A\*, Rhee Seung Hak\*, Park Sung Ho\*, Chun Jong Hun\*\*, and Choi Heung ki\*\*\*

### 요 약

본 논문에서는 물류 운송을 위한 100kg급 드론의 운용을 위해 FEA를 이용하여 BLDC 모터를 설계 및 개발하였다. 안정적인 비행을 위해 8개의 모터를 가진 옥토크터를 기준으로 설계하였으며 페이로드와 ESC 및 배터리, 카본 프레임 등 주요 부품의 무게까지 고려하여 중량해석 모듈을 구성했다. BLDC 모터의 출력 효율을 향상 시키기 위해 Halbach magnet array구조의 자기적 특성을 이용하여 기존 BLDC 모터와 비교 시, 동등한 출력을 유지하면서도 전류 밀도와 손실을 저감하고, 높은 효율을 낼 수 있음을 확인하였다.

### Abstract

In this paper, a BLDC motor was designed and developed using FEA for the operation of a 100 kg drone for logistics transportation. For stable flight, it was designed based on an octocopter with eight motors, and a weight analysis module was constructed taking into account the weight of major components such as payload, ESC, battery, and carbon frame. In order to improve the output efficiency of the BLDC motor, it was confirmed that by using the magnetic characteristics of the Halbach magnet array structure, it was possible to reduce current density and loss and achieve high efficiency while maintaining equivalent output when compared to existing BLDC motors.

### Key words

Drone, Brushed DC Motor, Artificial Intelligence(AI), Control Algorithm, Air Transport

---

\* 전남도립대학교, ghdwns221@naver.com, otxf0@naver.com, stenly69@daum.net, shpark@dorip.ac.kr,

\*\* 전남도립대학교, jhchun@dorip.ac.kr

\*\*\* 유니컴퍼니, pamo23@naver.com

※ 본과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 첨단분야 혁신융합대학사업의 연구결과입니다.

## 1. 서 론

드론의 주요 구성품은 크게 항법장치, 구동모터, 구동드라이버, 배터리 총 4가지로 나눌 수 있으며, 이 중 구동모터, 구동드라이버와 함께 프로펠러를 통틀어 구동모듈이라 하며, 구동모듈은 드론의 운송 능력을 결정하는 핵심 부품으로서 고출력밀도 개발이 요구되고 있다.

전 세계적으로 온라인 쇼핑의 급증으로 인해 물류 시스템의 효율성이 매우 중요한 이슈이며 고중량 화물을 처리할 수 있는 드론은 빠르고 효율적인 배송 수단으로 특히 대형 물품이나 다량의 제품을 배송하기 위하여 자세제어가 필수이며, 이를 위해 고효율의 모터가 장착되어야 한다.

특히, 산간 지역, 도서 지역, 또는 재난 현장과 같은 접근이 어려운 지역에서 고중량 물류 드론의 안정성 있는 제어와 그 역할을 원활히 수행하기 위하여 고효율의 모터가 필요하며, 100kg급 드론의 운용을 위해서는 최소 중량을 100kg급으로 산정하여야 하기 때문에 해당 중량을 공중에 띄우고 안정된 비행을 위해서는 적절한 추력의 효율을 낼 수 있는 모터가 필요하다.

드론에 적용할 구동모터는 전체 시스템의 효율 개선과 운전시간 증가를 위해 무게를 낮추면서 동일한 출력을 갖는 성능이 요구된다. [1].

구동 모터의 일반적인 요구조건은 아래와 같다.

- 한번 충전으로 주행 가능한 거리를 확대하기 위해 저속에서 고속까지 전 운전 영역에서 효율이 높아야 한다.
- 소형,경량으로 출력밀도가 높아야한다.
- 넓은 속도범위에서 정출력특성을 가질 것.
- 고속운전이 가능할 것.
- 진동에 강한 기계적 구조로 수명이 길고 저속시 고토크 특성을 가질 것

이러한 전기모터는 개발초기에 비교적 간단한 직류 전동기(Blushed DC Motor)가 사용되었으나,최근에는 DC 모터, 브러쉬리스 DC모터, AC모터로 크게 3가지로 구분된다[2].

최근 DC 모터의 단점을 보완한 Brushless DC모터는 내구성이 우수하고 마찰열이 없어 효율이 높아서 점차 사용이 확대되고 있다.

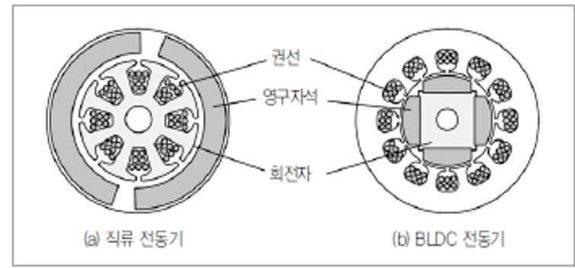


그림 1. DC/BLDC 전동기의 구조

표 1. 모터의 종류 및 특징

종류	공급전기	특징
DC Motor	DC	브러쉬로 인해 내구성이 낮음 제어가 간편 전류로 토크를 제어
BLDC Motor		브러쉬가 없어 내구성이 높음 순간출력이 낮음
영구자석 동기모터	AC	브러쉬가 없어 내구성이 우수 토크의 직접제어 순간출력이 높으나 인버터 가격이 고가

## II. BLDC 모터를 적용하여 100kg급 화물 운송이 가능한 드론 모터 개발

이 논문에서는 물류 운송을 위한 100kg급 드론의 운용을 위한 BLDC 모터의 개발을 진행하였다.

### 1) BLDC 모터 설계 및 시뮬레이션

BLDC 모터의 설계 진행을 위해 FEA를 이용하였으며, FEA에서 도출된 특성 결과는 그 해의 오차가 크지 않다고 하여 설계의 근거자료로 이용되고 있으며, 제안한 BLDC 모터의 설계를 진행하였다.

안정적인 비행을 위해 옥토크터를 기준으로 설계를 진행할 시, 페이로드 및 모터와 ESC, 배터리, 카본 프레임 등 주요 부품 무게까지 고려하여 설계를 진행해야 한다[3].

주요 부품 등 이륙중량을 특정 할 시, 그 중량을 기준으로 최대 추력(Thrust)은 중량의 1/8 kg 이상, 바람 등 다이내믹한 기상환경을 고려했을 시 1/4 kg 이상 최대추력을 내는 모터가 일정 개수 이상 필요하다. 이상적인 모터의 개수는 8개(옥토크터) 이상이다[3].

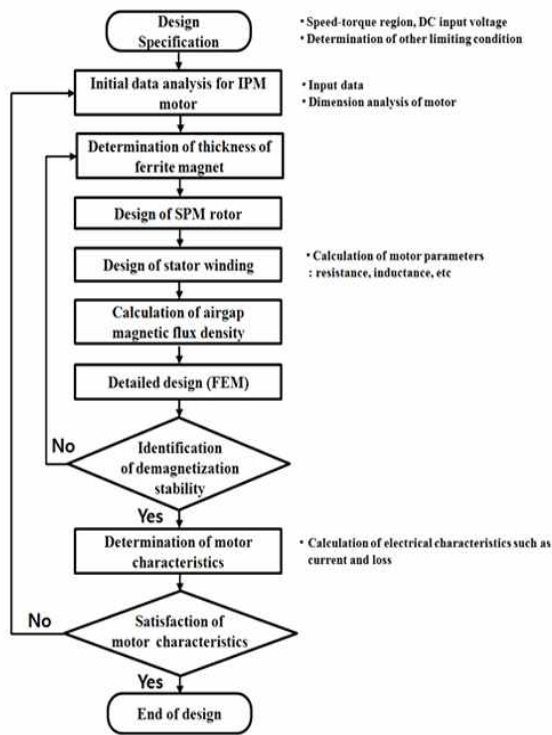


그림 2. BLDC모터의 설계를 위한 플로우차트

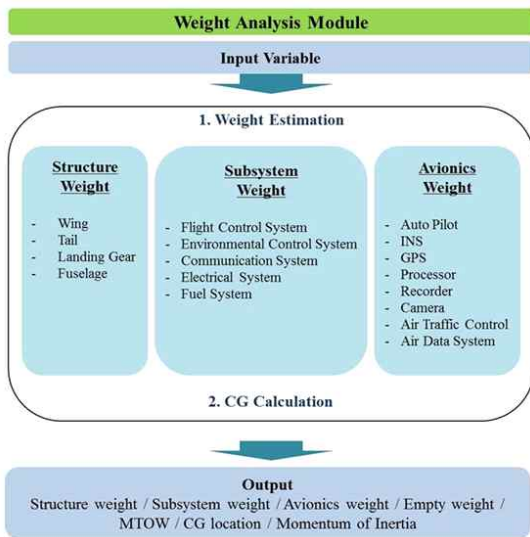


그림 3. 중량해석 모듈 구성[5]

BLDC 모터의 출력 특성을 향상시키기 위해 Halbach magnet 구조를 적용한 개선 모델 설계를 진행하여, 극 효율에 따라 출력특성을 비교분석 한 후 최적의 극효율을 선정한다. 그 후, Halbach magnet array 구조의 자기적 특징을 이용하여 회전자 외측의 철심 두께를 감소시키며, 고정자 치 폭도 저감한 개선 모델을 도출한다[4].

## 2) BLDC 모터 제작

이 논문에서는 물류 운송을 위한 100kg급 드론의 운용을 위한 BLDC 모터의 개발을 진행하였다.

영구자석의 배열과 길이를 조절하여 토크와 출력은 감소하지만 전류 밀도와 손실 및 효율 면에서 개선되는 현상을 보이는 Halbach magnet array 구조의 모터를 제작 할 수 있다[4].

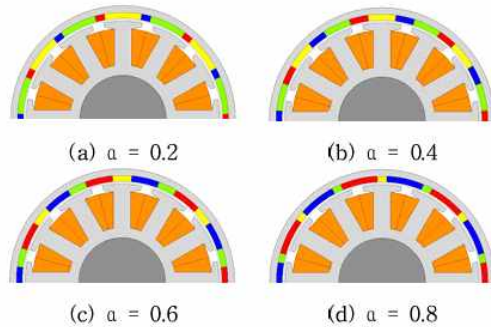


그림 4. Halbach magnet array 구조 극효율

변경시킨 a값을 기준으로 한 해석한 결과를 바탕으로 동등한 수준의 출력을 유지하면서 전류 밀도와 손실을 저감하고, 높은 효율을 낼 수 있는 BLDC 모터를 설계, 개발할 수 있다[4].

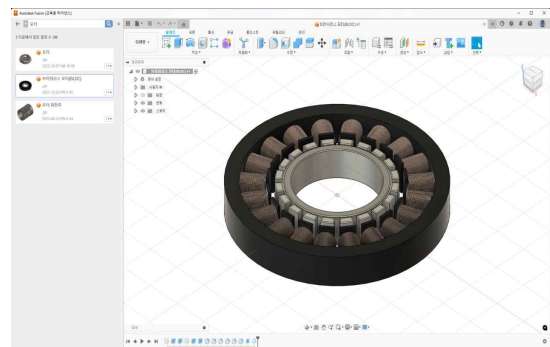


그림 5. 개선한 설계기반 Modeling 및 모터 제작

### III. 결 론

본 논문에서는 물류 운송을 위한 100kg급 드론의 운용을 위한 BLDC 모터를 개발을 진행하여, 고중량 드론을 통해 대용량 화물의 배송이 가능해짐에 따라 배송시간 절감 및 접근이 어려운 도서산간 지방에 배송을 원활히 할 수 있으며, 재난 현장이나 응급 상황시에도 고중량 드론을 활용한 시기적절한 구호활동을 진행할 수 있다.

### 참 고 문 헌

- [1] Soohwang Chae , Sooyoung Cho , Seong-Hyeon Go , Sang-Hwan Ham, Ju Lee, “Power density improvement design of drone BLDC motor considering noise and vibration” pp.1072-1073, 2019
- [2] 친환경 전기자동차 시장의 실태와 전망 ,BIRResearchGroup, BIR.
- [3] 김현우, 건축물 화재 인명구조 활동용 드론 개발 및 운용방안 연구. pp.45-76, 2022
- [4] Ye-Ji Park\*, Young-Geun Min\*, Sang-Hwan Ham\*\*, Sol Kim\*\*\*, Ho-Joon Lee\*\*\*\*, Ju Lee\*A study on Improvement of Power Characteristics of Permanent Magnet BLDC for Drone. pp.869-870, 2019
- [5] Possibility Based Design and Optimization for Electric-Powered Unmanned Aerial Vehicles. pp.16-40, 2013

# LLM을 활용한 항공사 고객서비스 에이전트 시스템

박희중\*, 박지환\*, 이민수\*, 강종연\*\*, 김현우\*\*, 고석주\*\*\*

## LLM-Based Agent Systems for Enhanced Airline Customer Support

Heejung Park\*, Jihwan Park\*, Minsu Lee\*, Jongyeon Kang\*\*, Hyunwoo Kim\*\*, and Seokjoo Koh\*\*\*

### 요약

코로나19 팬데믹이 끝난 이후 항공 서비스 수요가 급증했다. 본 논문에서는 대규모 언어 모델(LLM)을 기반으로 한 AI Agent를 개발해 항공사 고객 서비스 팀의 효율성과 고객 편의성을 증대시키고자 한다. LLM과 트랜잭션을 결합한 고객 서비스 에이전트는 예약, 일정 관리, 이메일 발송 등 다양한 편의 기능을 통합 지원한다. 또 RAG, ReAct를 통해 사실에 기반하여 논리적이고 합리적인 서비스를 제공할 수 있다. 이로써 고객 편의성이 증대되고, 항공사는 업무 효율성과 비용 절감 효과를 기대할 수 있다.

### Abstract

The demand for airline services has skyrocketed since the end of the COVID-19 pandemic. In this paper, we develop an AI Agent based on a large-scale language model (LLM) to improve the efficiency of airline customer service teams and customer convenience. The customer service agent, which combines LLM and transactions, provides integrated support for various convenient functions such as reservations, schedule management, and email sending. In addition, it can provide logical and reasonable services based on facts through RAG and ReAct. This increases customer convenience, and airlines can expect work efficiency and cost reduction effects.

### Key words

large language model, prompt engineering, google gemini, ai agent, rag, react

### I. 서론

심은 현재 폭발적으로 증가하고 있으며, 그에 따라 항공 서비스에 대한 고객의 수요 역시 높아지고 있다. 그러나 항공 산업은 규정과 약관이 자주 변경  
코로나19 팬데믹으로 위축되었던 여행에 대한 관

\* 경북대학교 컴퓨터학부, heejung000321@gmail.com, pcbmlh73@gmail.com, minsoo030232@gmail.com

\*\* 경북대학교 컴퓨터학부, whddus0789@gmail.com, lag4420@gmail.com

\*\*\* 경북대학교 컴퓨터학부 교수, sjkoh@knu.ac.kr(교신 저자)

※ 이 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 통해 수행되었음(2021-0-01082)

되고 복잡하여 신뢰할 수 있는 최신 정보를 제공하는 것이 중요한 과제로 대두되고 있다. 여러 항공사의 규정은 항공사 홈페이지나 대행사 웹사이트에서 확인할 수 있지만, 고객은 항공사별로 일일이 확인해야 하거나, 대행사 사이트의 정보가 최신인지 확인하기 어려운 불편함이 있다. 예를 들어, 대행사 웹사이트에서 조회한 ‘A 항공사의 위탁 수하물 규정’ 이 수년 전 정보라면, 이는 최신 정책과 다를 수 있어 고객에게 혼란을 줄 뿐만 아니라 항공사의 서비스 신뢰도에 부정적인 영향을 미칠 가능성이 크다. 최근 이러한 문제를 해결하기 위해 기업에서는 LLM을 단순한 Q&A 용으로 활용하는 것을 넘어 다양한 업무를 자동으로 수행하는 에이전트 개념의 애플리케이션에 대한 관심이 높아지고 있다.

이에 본 연구는 대규모 언어 모델(LLM)을 활용하여 항공사 관련 최신 정보를 반영하고 신뢰할 수 있는 사실 기반 응답을 제공하는 고객 서비스 에이전트를 설계하는 것을 목표로 한다. 또한 이 에이전트 시스템은 고객의 질문에 정확한 정보 제공과 함께, 예약, 일정 등록, 이메일 발송과 같은 트랜잭션까지 처리하는 일원화된 서비스를 지원한다.

이를 위해 LLM과 트랜잭션 기능을 결합해 개인별 맞춤형 응답을 구현하며, 고객 데이터를 바탕으로 한 개인화된 서비스 시나리오를 제공할 수 있다. 특히 ReAct 아키텍처(Reasoning and Acting) 기반의 시스템을 통해, 고객의 의도를 추론하고 적절히 응답할 수 있는 고도화된 기능을 수행한다.

또한, 본 시스템은 Google Gemini를 기반으로 개발되었으며, 프롬프트 엔지니어링을 통해 질의의 의도와 맥락을 정확하게 파악하여 자연스러운 다국어 지원이 가능하다. 이를 통해 고객은 항공사로부터 시간과 장소에 구애받지 않고, 정확하고 신뢰할 수 있는 정보를 손쉽게 제공받아 이용 경험을 개선할 수 있다.

따라서 본 연구에서 제안하는 에이전트 시스템은 고객 서비스의 신뢰도와 만족도를 높일 뿐만 아니라, 항공사의 업무 효율성 및 비용 절감에도 기여할 수 있는 솔루션이 될 것이다.

## II. 관련 연구

### 2.1 프롬프트 엔지니어링

프롬프트 엔지니어링은 생성형 AI 시스템에서 원하는 결과를 얻기 위해 최적화된 요청을 구성하는 작업이다. 프롬프트 유형은 직접 프롬프트와 예시 프롬프트로 구분할 수 있다. 제로샷이라고도 불리는 직접 프롬프팅은 모델에 예시를 제공하지 않고 명령어만을 통해 답변을 생성하도록 하는 방식이다.

예시 프롬프팅은 모델이 작업의 패턴이나 형식을 이해할 수 있도록 예시를 제공하는 방식으로, 예시의 개수에 따라 원샷과 퓨샷 및 멀티샷 방식으로 나눌 수 있다. 원샷 프롬프팅은 모델의 작업 수행에 필요한 하나의 예시를 제공하는 방식이고, 퓨샷 및 멀티샷 프롬프팅은 모델에게 여러 예시를 제공하여 보다 더 복잡한 태스크에 적합하게 응답을 생성하도록 한다[1].

### 2.2 RAG(Retrieval-Augmented Generation)

RAG는 LLM 성능 강화를 위해 개발된 Architecture이다. 이는 단순히 학습된 데이터만을 기반으로 답변하던 기존 LLM이 외부 데이터를 기반으로 답변하도록 한다.

RAG의 목표는 LLM이 최신 정보나 특정 분야의 지식에 접근하여 보다 정확한 응답을 생성하도록 하는 것이다. 이를 위해 RAG는 크게 두 단계로 동작한다: (1) 검색(Retrieval) (2) 생성(Generation)[2].

검색 단계에서는 RAG Model이 질문에 답변하기 위해 필요한 추가 정보를 외부 데이터베이스에서 찾는다. 두 번째 생성 단계에서는 이전 단계에서 찾은 정보를 기반으로 사용자 질문에 대한 최종 응답을 생성한다.

이처럼 RAG는 훈련 데이터에 포함되지 않은 최신 정보나 특정 분야의 전문 지식을 필요로 할 때 효과적이다. 규모가 큰 Gemini나 ChatGPT 특성상 최신 데이터를 매번 학습하기엔 제약이 있는데, 이를 효과적으로 보완할 수 있는 방법이다. 따라서 최신 제품 매뉴얼 챗봇, 의학 정보 검색 등 사실에 기반한 정확한 답변을 해야하는 시스템에 주로 적

용된다.

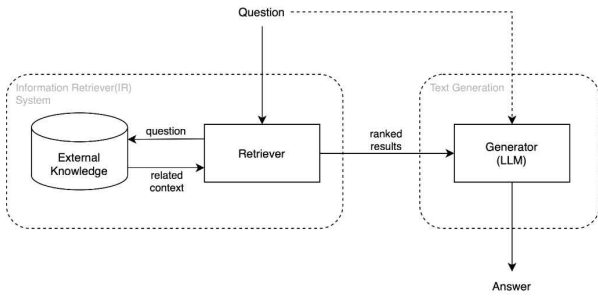


그림 1. RAG Architecture

### 2.3 ReAct

ReAct는 AI 에이전트가 복잡한 작업을 (1) 논리적 추론(Reasoning)과 (2) 행동(Act)을 통해 수행하도록 하는 방법론이다[3]. 기존 LLM은 단순히 주어진 질문만을 기반으로 대답하도록 설계되었기 때문에 복잡한 논리적 추론이 필요한 문제를 해결하지 못할 수 있다.

ReAct는 스스로 (1) 추론 - (2) 행동을 반복하며 위 단점을 보완한다. (1) 추론 단계에선 주어진 정보를 바탕으로 다음에 해야 할 행동을 판단한다. 이 단계에서 수행할 행동 후보들을 평가하고 선택한다. (2) 행동 단계에선 이전에 내린 추론을 바탕으로 특정 행동을 실행한다. 행동 단계 후 그 결과를 바탕으로 다시 추론이 이어진다.

ReAct는 논리적 추론만 사용하는 “CoT(Chain-of-Thought)” 방식에 비해 환각 효과(Hallucination)를 절반 이상 감소시킨다[4].

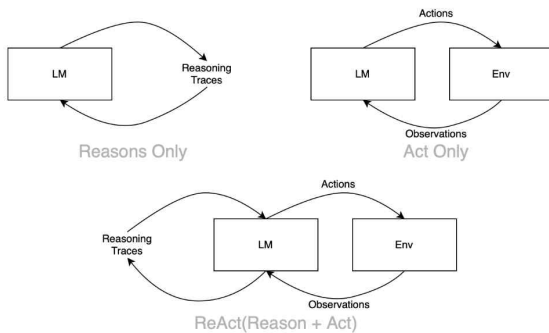


그림 2. ReAct Prompting

## III. LLM을 활용한 항공사 고객센터 에이전트 시스템

### 3.1 시스템 설계

본 항공사 에이전트 시스템은 항공사 고객의 다양한 요청을 상황에 맞게 효과적으로 처리하고 정보를 제공할 수 있도록 구성되어 있다. 시스템의 실행 과정은 그림 3과 같으며 프롬프트에 사용자 요청이 들어왔다고 가정해 보자. Function Declaration으로 선언된 다양한 툴 목록에서 해당 요청에 가장 적합한 작업을 수행할 수 있는 툴을 선택한다. 이 툴들은 사용자의 요청에 맞춰 RAG 아키텍처 기반의 정보 검색 기능 또는 예약 신청과 같은 트랜잭션 수행 기능을 선택적으로 실행한다.

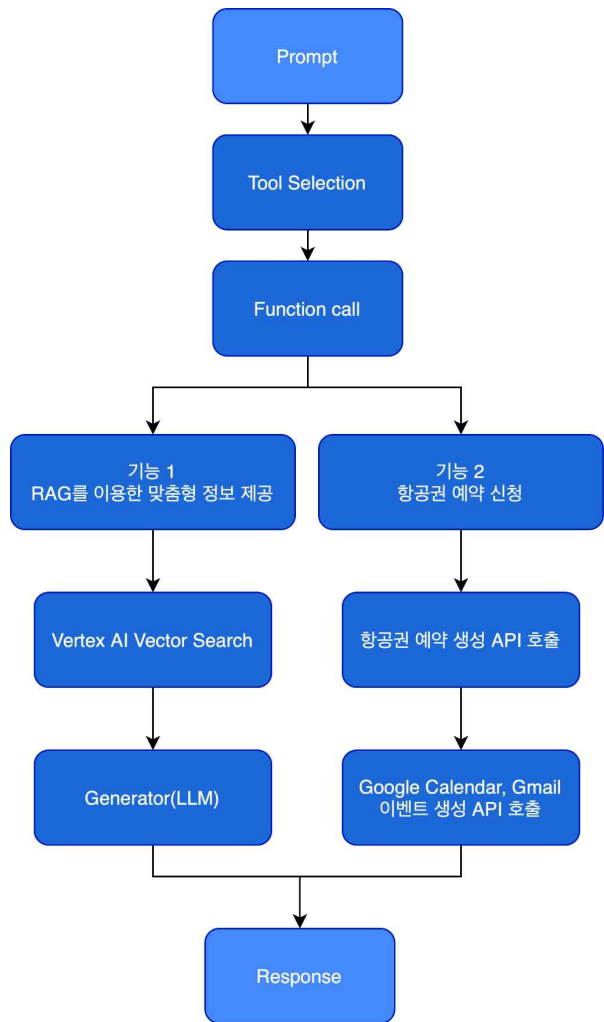


그림 3. 시스템 다이어그램



기능 1은 사용자가 “기내에 인라인스케이트를 소지하고 탑승할 수 있어?”처럼 항공사의 약관 및 정책과 관련된 질문을 했을 때 외부 데이터베이스를 토대로 정확한 답변을 제공할 수 있다. 따라서 고객이 직접 약관을 찾아봐야 하는 수고를 크게 덜어 준다. 특히 구체적인 세부 약관에 대해서는 직접 찾아보기 어려움이 있는 경우에도 손쉽게 정보를 얻을 수 있다는 이점이 있다.

기능 2는 사용자가 “N월 N일 인천에서 도쿄로 가는 항공편을 예약해줘” 라는 요청을 했을 때 예약 생성 API호출로 예약 생성을 할 수 있으며 예약 생성 성공 시 구글 캘린더 및 Gmail 이벤트 생성까지 진행하면서 사용자 편의성을 올릴 수 있다.

### 3.2 RAG 아키텍처를 이용한 맞춤형 정보 제공

항공사의 규정 및 정책에 관한 질문에 대한 응답을 제공하기 위해서는 최신 정보가 담겨있는 데이터가 필요하다. 본 연구에서는 ‘대한항공’ 에서 제공하는 공식 문서를 정적 데이터로 활용하였다. 그림 4와 같이 여객 운송 약관과 운임표, 기내 반입 금지 물품 규정 등의 문서를 Data Store에 저장하여 이 데이터를 기반으로 답변을 생성하는 RAG 아키텍처를 이용하였다.

Name	Size	Type	Created	Storage class
대한항공_CIS_운임표.pdf	245.9 KB	application/pdf	Oct 30, 2024, 6:42:23 PM	Standard
대한항공_국내선_운임표.pdf	62.1 KB	application/pdf	Oct 12, 2024, 6:10:44 PM	Standard
대한항공_국제선_운임표.pdf	388.1 KB	application/pdf	Oct 12, 2024, 6:10:45 PM	Standard
대한항공_국제선_운임표_보조문.pdf	633.3 KB	application/pdf	Oct 12, 2024, 6:10:46 PM	Standard
대한항공_기내반입물품규정.pdf	99.5 KB	application/pdf	Oct 12, 2024, 6:10:47 PM	Standard
대한항공_내항주_운임표.pdf	243.3 KB	application/pdf	Oct 30, 2024, 6:42:22 PM	Standard
대한항공_동남아_운임표.pdf	270.9 KB	application/pdf	Oct 30, 2024, 6:42:22 PM	Standard
대한항공_러시아_운임표.pdf	178.9 KB	application/pdf	Oct 12, 2024, 6:10:46 PM	Standard
대한항공_미주_운임표.pdf	260.6 KB	application/pdf	Oct 30, 2024, 6:42:22 PM	Standard
대한항공_유럽_운임표.pdf	265.9 KB	application/pdf	Oct 30, 2024, 6:42:23 PM	Standard
대한항공_일본_운임표.pdf	267.1 KB	application/pdf	Oct 30, 2024, 6:42:23 PM	Standard
대한항공_중국_운임표.pdf	275.8 KB	application/pdf	Oct 30, 2024, 6:42:23 PM	Standard

그림 4. Data Store에 저장한 정적 데이터

RAG 아키텍처는 정보 검색기(Retriever)와 생성기(Generator)를 결합하여 작동하는데, 본 연구에서는 Vertex AI Vector Search를 정보 검색기로, Gemini를 생성기로 채택하였다. 사용자가 질문을 입력하면, Vector Search를 통해 관련 문서를 Data Store에서 찾아낸다. 이후 Gemini는 검색된 문서의

내용을 바탕으로 응답을 생성한다.

그림 5와 같이, 사용자가 “기내에 인라인스케이트를 소지하고 탑승할 수 있어?”라는 질문을 입력했을 때, RAG는 기내 반입 금지 물품에 관한 공식 문서를 검색한다. 해당 문서에는 스포츠 용품에 대한 명확한 규정이 포함되어 있으며, 이 문서에 따르면 인라인스케이트는 반입이 가능한 물품 목록에 포함되어 있다. 따라서, 챗봇은 “인라인스케이트는 기내에 소지하고 탑승이 가능합니다.”라는 명확한 답변을 제공할 수 있다.

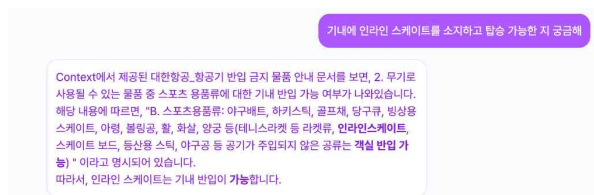


그림 5. RAG 적용 예시

### 3.3 ReAct 기반의 질문 강화

AI Agent가 RAG를 통해 사실에 기반한 정확한 답변을 할 수 있게 됐다면, ReAct를 통해 보다 합리적인 추론 능력을 갖출 수 있게 된다.

예를 들어 인천에서 도쿄로 가는 항공편에 대한 정보를 알고 있는 AI Agent에게, 인천에서 나리타 공항으로 가는 항공편에 대한 요청을 보냈다고 가정해 보자. 나리타 공항은 도쿄 주변에 있어서 도쿄 여행을 가는 사람들이 주로 이용하는 공항이다. 이때 만약 단순 LLM에 RAG 구조만 적용되어 있다면, AI Agent는 ‘나리타 공항’에 대한 정보가 없어 그림 6과 같이 적절한 응답을 찾지 못하게 된다.

ReAct를 사용해 질문을 강화하면 LLM은 그에 맞는 추론-행동(즉, Reasoning-Act)을 하게 된다. 논리적인 추론과 그에 따른 행동, 그리고 그 행동으로 관찰한 결과를 바탕으로 하는 추론이 연쇄적으로 일어나 알맞은 응답을 도출할 수 있게 된다.

아래는 실제로 ReAct가 적용되기 전후를 비교한 것이다. ReAct가 적용되기 전 그림 6에선 나리타 공항이 도쿄 근처인지 모르고 이상한 답변을 한다. 그림 7에 ReAct가 적용된 후엔 나리타 공항이 도쿄 근처임을 성공적으로 추론해 제대로 된 답변을 주

고 있다.

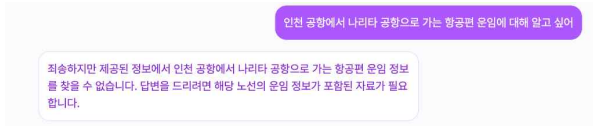


그림 6. ReAct 적용 전



그림 7. ReAct 적용 후

#### IV. 결론

본 논문에서는 Google Gemini Pro 1.5를 기반으로 항공사 고객 서비스 AI 에이전트를 개발하였다. 이 에이전트는 고객이 직접 찾기 어려운 항공사의 세부 규정부터 항공편 예약, 조회 등 다양한 편의 기능을 제공한다.

항공사 고객 서비스 AI 에이전트는 사실에 근거한 답변을 제공하며, 고객의 요청에 논리적으로 대응할 수 있어야 한다. 그러나 LLM 단독으로 답변을 제공할 경우 환각 현상(hallucination)과 같은 문제가 발생할 가능성이 있다.

따라서 본 연구는 RAG(Retrieval-Augmented Generation)와 ReAct 프롬프트 엔지니어링 기법을 적용하여 LLM의 한계를 보완하고자 했다. RAG는 외부 데이터베이스에서 데이터를 참조해 LLM이 사실 기반의 응답을 생성할 수 있도록 지원한다. ReAct는 추론과 행동의 반복을 통해 LLM이 보다 합리적이고 논리적인 답변을 생성하도록 한다.

또한, 사용자 편의를 극대화하기 위해 Google Calendar나 Gmail과 같은 서비스와 연동하여 항공 일정 확인, 항공편 예약 등을 하나의 서비스 내에서 손쉽게 처리할 수 있도록 구현하였다.

본 논문에서 제안한 시스템은 항공사 고객의 편의를 증대하고 고객 응대의 효율성을 높인다. 추가

적인 보완과 검증을 통해 현장에 도입될 경우, 고객에게는 편리하고 효율적인 서비스를 제공하고, 항공사에는 운영 효율성과 비용 절감 효과를 기대할 수 있다.

#### 참고 문헌

- [1] "Prompt Engineering for Generative AI", Google for Developers, last modified 8 Aug. 2023, <https://developers.google.com/machine-learning/resources/prompt-eng>, accessed 30 Oct. 2024
- [2] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, Douwe Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks"
- [3] Yao S, Zhao J, Yu D, Du N, Shafran I, Narasimhan K R, Cao Y, "ReAct: synergizing reasoning and acting in language models." In: Proceedings of the 11th International Conference on Learning Representations. 2023
- [4] J. Wei, X. Z. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. Le, and D. Zhou, "Chain-of-thought prompting elicits reasoning in large language models," arXiv preprint arXiv: 2201.11903, 2022.

# 인접한 셀에 의한 트랜지스터의 특성 변화를 최소화하는 표준 셀 배치 최적화

김다연\*, 김근식\*, 한창호\*\*

## Placement Optimization of Standard Cells to Minimize Transistor Characteristic Variations due to Neighboring Cells

Dayeon Kim\*, and Geunsik Kim\*, Changho Han\*\*

### 요 약

트랜지스터는 인접한 트랜지스터의 디퓨전 크기에 따라 특성(threshold voltage, drain current)이 달라지는데 이를 NDE(Neighbor Diffusion Effect)라고 한다. 초미세 공정에서 트랜지스터 간의 거리가 점점 줄어들기 때문에 NDE에 의한 트랜지스터 특성 변화가 더욱 심해진다. 디지털회로는 다양한 디퓨전 크기를 갖는 트랜지스터를 사용하여 표준 셀을 만들고 설계 자동화 도구를 이용하여 자동으로 셀을 배치하기 때문에 다수의 NDE가 발생하게 되고 이는 실리콘으로 제작된 회로와 설계단계에서 예측한 회로 간의 성능 차이를 유발한다. 이를 해결하기 위해 같은 디퓨전 크기를 갖는 셀들이 좌우에 인접하도록 재배치하는 알고리즘을 제안하였다. 최적화 영역을 활용한 다중 행 재배치 알고리즘은 코어 영역을 여러 개의 영역으로 나누고 그 안에서 셀을 재배치함으로써 최소한의 이동 거리로 성능 변화를 방지할 수 있는 셀 배치 방법을 구현하였다.

### Abstract

The characteristics of a transistor (threshold voltage, drain current) vary based on the diffusion size of adjacent transistors, a phenomenon known as Neighbor Diffusion Effect (NDE). As the distance between transistors decreases in deep sub-micron process, the impact of NDE on transistor characteristics become more pronounced. For the digital circuit design, standard cells are crated using transistors with various diffusion sizes, and design automation tools are used to automatically place these cells, resulting in lots of transistors having NDE. This effect causes a discrepancy between the performance of the fabricated silicon circuits and the design-stage predictions. To address this issue, we propose an algorithm that relocates cells so that those with the same diffusion size are adjacent to each other. This multi-row relocation algorithm, which leverages optimization regions, divides the core area into multiple sections and rearranges cell within each section thereby implementing a placement strategy that minimizes performance variation with minimal displacement.

Key words : neighbor diffusion effect (NDE), multi-row (MR) placement optimization

---

\* 국립금오공과대학교, 반도체시스템전공, {ekdus2376, yyx0008}@naver.com

\*\* 국립금오공과대학교, 전자공학부, 반도체시스템전공, changho.han@kumoh.ac.kr

※ 이 연구는 정부(교육부-산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0022161, 반도체전공트랙사업)

## I. 서론

최근 반도체 기술의 미세화가 발전됨에 따라 디바이스 크기는 점점 작아지고, 셀 간 배치 공간이 줄어드는 현상이 발생하고 있다. 이에 따라 Fin의 개수가 다른 셀들이 인접할 때 서로 다른 diffusion height 차이(Diffusion step)로 인해 NDE(Neighbor Diffusion Effect)가 발생하여 트랜지스터의 임계전압(threshold voltage)과 드레인 전류(drain current)를 변화시키는데, 이로 인해 트랜지스터의 성능이 달라진다. 이는 결국 집적회로의 신뢰성을 떨어뜨려 품질에 부정적인 영향을 미친다. 기존에는 NDE의 영향을 줄이기 위해 셀 배치 시 서로 다른 diffusion height를 가진 셀들을 띄워 필러 셀을 삽입하였다. 그림 1을 통해 NDE와 필러 셀 삽입에 대해 이해할 수 있다.

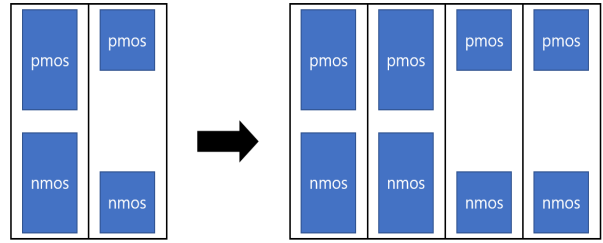


그림 1. NDE와 필러 셀 삽입에 대한 예시  
Fig. 1. Examples of NDE and filler cell insertion

그러나 최신 기술에서는 디바이스가 줄어들어 따라 여유 공간 확보가 어려워지고 있어 단순히 인접한 셀들을 띄워서 Filler cell을 넣는 것은 사실상 한계에 다다랐다. 따라서 NDE를 최소화하기 위해서는 셀들을 재배치하는 방식이 필요하다[1].

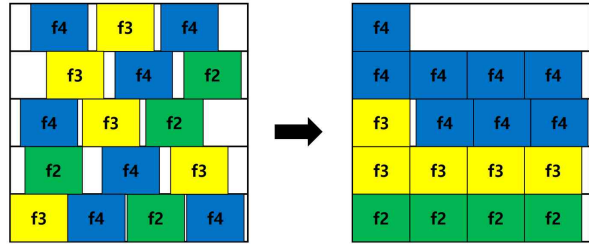


그림 2. 전체 행 재배치 결과  
Fig. 2. Results of replacement in the entire row

본 논문에서는 모든 셀을 재배치하는 placement 최적화 연구 과정을 설명하고, 최소한의 움직임으로 셀을 재배치하는 방법을 기술하였다. 이 배치 방법은 단일 행, 전체 행 재배치를 기반으로 연구되었다. 마지막으로 초기 상태의 배치와 전체 행 재배치, 단일 행 재배치, 최적화 영역에서의 다중 행 재배치를 활용한 배치를 비교함으로써 최적화 영역에서의 다중 행 재배치의 필요성에 대해 분석한다.

며, 전체 셀 중에서 20%는 f2, 30%는 f3, 50%는 f4이다. core area의 모든 행에 배치된 셀을 추출한다. f2, f3, f4끼리 저장한다. NDE는 서로 다른 diffusion height를 가질 때 나타나는 현상이므로 fin의 개수가 같은 셀끼리는 인접하게 배치하여도 성능에 무관하다는 점을 이용하여 재배치를 진행한다.

## II. 전체 행 재배치

추출한 셀 중에서 f2를 core area 가장 아래 행부터 인접시키며 배치한다. 한 행을 다 채우면 바로 위 행으로 이동하여 셀 배치를 계속한다. f2의 배치가 완료되었다면 필러 셀이 삽입될 수 있도록 필러 셀의 너비만큼 공간을 띄운 다음 f3의 배치를 시작한다. f4 역시 f3과 필러 셀의 너비만큼 간격을 둔 다음 배치가 된다. 전체 행 재배치의 결과는 그림 2와 같이 나타난다.

이 섹션에서는 전체 행 재배치 알고리즘과 한계점에 대해 설명한다.

입력: 셀의 너비, 셀의 fin 개수

출력: 최적화된 전체 행에 대한 셀의 재배치

### A. 알고리즘

각 셀들은 fin의 개수를 구분하기 용이하도록 fin의 개수가 2개이면 f2, 3개이면 f3, 4개이면 f4라고 지칭한다. 각 셀의 fin 개수는 임의로 정해진 것이

### B. 한계점

모든 셀을 추출한 다음 재배치를 하는 것이기에 NDE를 확실하게 최소화할 수 있지만 모든 행에 대해서 이동이 이루어지기 때문에 wire의 길이가 초기배치 대비 평균 116% 증가한다. wire 길이의 증가는 디바이스의 속도 저하를 야기한다.

## III. 단일 행 재배치

이 섹션에서는 단일 행 재배치 알고리즘과 한계점에 대해 설명한다.

입력: 셀의 너비, 셀의 fin 개수

출력: 최적화된 단일 행에 대한 셀의 재배치

#### A. 알고리즘

전체 행 재배치와 유사하게 셀을 추출하여 fin의 개수가 같은 셀끼리 인접하여 재배치한다. 다만, 단일 행 재배치는 모든 행에서 셀을 추출하여 모든 행에 대해 재배치하는 것이 아닌, 한 행 안에서만 셀의 정보를 추출하고 fin의 개수에 따라 셀을 구분한 다음, 셀이 기존에 위치해 있었던 행 내에서만 셀을 배치한다. 한 행 내에 f2, f3, f4 순서대로 배치하고 fin 개수가 다른 셀 사이에는 필러 셀 너비만큼의 간격이 띄어져야 한다. 단일 행 재배치의 결과는 그림 3과 같이 나타난다.

#### B. 한계점

전체 행 재배치보다 wire 길이가 평균 29% 감소하였다. 그러나 여전히 x축 방향으로의 이동 거리가 크기에, 기존 배치보다 wire 길이가 53% 증가한 모습을 보여 여전히 디바이스 성능에 문제가 생길 가능성이 있다.

### IV. 최적화 영역에서의 다중 행 재배치

#### 1) 최적화 영역에서의 다중 행 최적화

이 섹션에서는 최적화 영역에서의 다중 행 최적화의 알고리즘에 대해 설명한다.

입력: 셀의 기존 좌표, 셀의 너비, 셀의 fin 개수

출력: 최적화 영역에서의 다중행 재배치

전체 행 재배치에서는 chip 내 모든 셀을 core area의 전체 행에서 배치하였다. 다만, 배치 가능 영역이 넓어지면서 셀이 초기배치에서 이동하는 거리가 증가하였고, 증가한 거리에 비례하여 wire 길이 역시 과도하게 커진 것을 확인하였다. 이에 착안하여 셀을 재배치할 수 있는 영역을 한 행으로 제한한 단일 행 재배치를 진행하였으나, 이 방법 역시 wire 길이를 효과적으로 줄이지 못하였다.

위 방법들을 기반으로 NDE 측면에서도, wire 길이 측면에서도 효과적인 방법을 고안하여 최적화 영역이라는 개념을 도입하였다. 최적화 영역은 코어

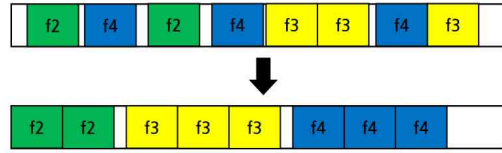


그림 3. 단일 행 재배치의 결과  
Fig. 3. Results of replacement in single-row

영역을 일정한 크기의 바운더리로 나누는 것으로, 크기는 사용자가 선택할 수 있다.

#### A. 알고리즘

본 논문의 경우 코어 영역을 x축 기준으로 3열로 나누는 다음, y축에 대해서 두 행씩 나누어 최적화 영역을 정해주었다. 단일 행의 개수가 홀수라면 두 행씩 나누는 다음 가장 마지막에 있는 세 행을 같은 최적화 영역으로 설정해주었다. 그 후 기존의 MR 최적화와 같이 추출한 셀을 fin의 개수가 같은 셀들끼리 구분한 다음 최적화 영역 가장 아래부터 f2, f3, f4 순서대로 셀을 배치해주었다.

최적화 영역을 사용하면서 배치 영역을 축소함에 따라 셀을 배치할 공간이 부족해지고, 그 결과 한번에 배치되지 못하는 셀들이 발생하였다. 이를 해결하기 위해 기존 최적화 영역 내의 빈 공간을 탐색하는 알고리즘과 다른 최적화 영역 내의 빈 공간을 탐색하는 알고리즘을 추가하였다.

셀이 초기배치에서 크게 벗어나는 것을 방지하기 위해 먼저 기존에 셀이 있던 최적화 영역 내에서 빈 공간을 탐색하는 것을 우선순위로 한다. 셀의 초기위치에 따라 탐색의 순서가 달라진다.

만약 기존의 최적화 영역에서 셀이 배치될 공간이 없는 경우에는 셀이 기존의 최적화 영역에서 벗어나 배치될 공간이 있는 다른 최적화 영역으로 이동할 수 있도록 한다. 먼저 그림 4에 나와 있는 것처럼 셀이 초기 배치된 행에 있는 다른 최적화 영역에서 공간을 탐색한다. 셀이 배치될 공간을 찾지 못했다면 그림 5에 나타난 대로 탐색 영역을 셀이 초기 배치된 행의 바로 위아래 행으로 옮겨준다. 이때에는 셀이 초기에 배치된 위치에 따라 탐색하는 영역의 우선순위를 지정해줌으로써 최대한 이동 거리와 wire 길이가 증가하지 않고 배치될 수 있도록 한다. 셀이 배치될 수 있는 공간을 찾을 때까지 탐색 행을 위아래로 계속 늘려가며 탐색을 진행한

다.

B. 한계점

최적화 영역 내의 셀들을 추출한 다음 배치할 때 각 셀의 초기배치를 고려하지 않고 추출된 순서대로 배치를 진행하기 때문에 wire 길이 및 이동 거리가 완전히 줄어들지 않는 문제가 발생한다.

2). 초기배치를 고려한 최적화 영역에서의 다중 행 최적화

기존의 최적화 영역에서의 다중 행 재배포보다 더 감소한 wire 길이와 이동 거리를 얻기 위해서는 셀들이 초기배치에서 최대한 적게 움직여야 한다. 이 방법에서는 셀이 초기배치에서 갖는 좌표와 최적화 영역 내에서 배치될 좌표를 통해 셀이 움직이게 될 거리를 계산하여 셀들이 가장 적게 움직이는 재배포를 찾아낸다.

<알고리즘>

초기배치에서 갖는 좌표와 배치될 좌표의 거리를 비교하고, 가장 짧은 거리 차이를 갖는 셀에 우선순위를 부여하여 먼저 배치될 수 있도록 한다. 예를 들어 셀이 초기배치에서 갖는 좌표를  $(x_1, x_2)$  라고 하고, 배치되어야 할 좌표를  $(x_2, y_2)$ 라고 할 때 두 좌표 간의 거리 차이는

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

으로 구한다.

이 최적화 방법 역시 한번에 배치되지 못하는 셀들이 나타난다. 기존의 최적화 영역에서 사용하는 빈 공간 탐색 알고리즘이 셀의 초기배치를 고려하는 방법이기때문에 이 섹션의 재배포 방법에도 기존의 빈 공간 탐색 알고리즘을 동일하게 사용한다.

V. 실험

단일 행, 전체 행, 최적화 영역에서의 다중 행, 위치를 고려한 최적화 영역에서의 다중 행 재배포를 각 5개의 Utilization에서 실험을 수행하였다. 사용된 디자인은 GCD이며, ICC2를 활용하여 데이터를 추출하였다.

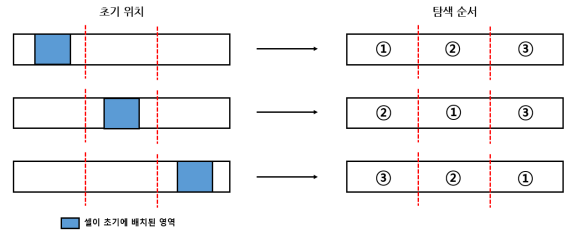


그림 4. 초기 배치된 행에 있는 다른 최적화 영역 탐색 순서

Fig. 4. Exploration order of other optimization areas in the initially placed row

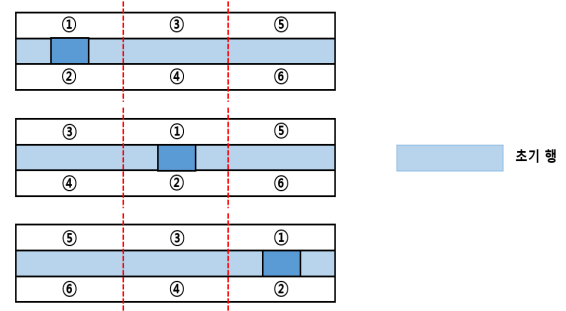


그림 5. 다른 행에 있는 최적화 영역 탐색 순서

Fig. 5. Exploration order of optimization areas in other rows

그림 6은 각각의 Utilization에서 실험한 결과이다. 각 배치 방법을 사용한 5가지 Utilization 실험 모두 NDE를 최소화시킬 수 있다.

특히 위치를 고려한 최적화 영역에서의 배치 방법은 다른 배치 방법들보다 초기배치의 Wire 길이를 최소한으로 증가시키면서 인접한 셀을 모두 제거하고, 또한 셀 이동 거리가 전체 행 재배포 대비 평균 76% 감소한 것을 알 수 있다. 따라서 해당 배치 방법은 Trade off 문제를 가장 효과적으로 보완하는 방법이다.

VI. 결론 및 향후 연구

본 논문에서는 wire 길이 및 이동 거리를 최소한으로 증가시키면서 NDE 발생을 최소화하는 재배포 방법을 제시하였고, 두 가지 개선점을 도출했다. 1) fin의 개수가 다른 셀이 인접한 경우를 제거하여 NDE를 100% 해결하였다. 2) 초기배치 대비 wire 길이 증가율이 평균 22% 정도 감소하였다.

Utilization = 63%	인접한 셀[ea]	Wire 길이[um]
초기배치	53	1795.108
전체 행 재배치	0	3661.97(+104%)
단일 행 재배치	0	2607.429(+45.3%)
최적화 영역에서의 다중 행 재배치	0	2178.65(+21.37%)
초기배치를 고려한 최적화 영역에서의 다중 행 재배치	0	2116.312(+17.89%)

Utilization = 68%	인접한 셀[ea]	Wire 길이[um]
초기배치	64	1734.013
전체 행 재배치	0	3605.953(+107.95%)
단일 행 재배치	0	2584.785(+49.06%)
최적화 영역에서의 다중 행 재배치	0	2121.535(+22.35%)
초기배치를 고려한 최적화 영역에서의 다중 행 재배치	0	2100.451(+21.13%)

Utilization = 75%	인접한 셀[ea]	Wire 길이[um]
초기배치	70	1673.893
전체 행 재배치	0	3640.194(+117.47%)
단일 행 재배치	0	2497.372(+49.2%)
최적화 영역에서의 다중 행 재배치	0	2146.456(+28.23%)
초기배치를 고려한 최적화 영역에서의 다중 행 재배치	0	2058.195(+22.96%)

Utilization = 80%	인접한 셀[ea]	Wire 길이[um]
초기배치	92	1643.668
전체 행 재배치	0	3628.168(+120.74%)
단일 행 재배치	0	2614.289(+59.05%)
최적화 영역에서의 다중 행 재배치	0	2062.755(+25.5%)
초기배치를 고려한 최적화 영역에서의 다중 행 재배치	0	1992.612(+21.23%)

Utilization = 87%	인접한 셀[ea]	Wire 길이[um]
초기배치	105	1586.355
전체 행 재배치	0	3626.938(+128.63%)
단일 행 재배치	0	2549.575(+60.72%)
최적화 영역에서의 다중 행 재배치	0	2182.776(+37.6%)
초기배치를 고려한 최적화 영역에서의 다중 행 재배치	0	2006.349(+26.48%)

그림 6. 5개의 utilization에서 수행한 실험 결과  
 Fig. 6. Experimental results conducted on 5 utilizations

본 논문에서 제시한 새로운 다중 행 배치 방법을 바탕으로 AES, JPEG, VGA 등 셀의 개수가 기존 디자인(GCD)보다 많은 디자인에서 구현할 계획이다.

IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS, VOL. 38, NO. 9, SEPTEMBER 2019.

### 참 고 문 헌

[1] Changho Han, Andrew B. Kahng, Fellow, IEEE, Lutong Wang, Student Member, IEEE, and Bangqi Xu, Student Member, IEEE, "Enhanced Optimal Multi-Row Detailed Placement for Neighbor Diffusion Effect Mitigation in Sub-10 nm VLSI",

# 유해 조류 검출을 위한 YOLOv5와 YOLOv8 성능 비교 연구

양수빈\*, 김민지 \*\*, 이영학\*\*\*

## \*\*A Comparative Study on the Performance of YOLOv5 and YOLOv8 for Harmful Bird Detection\*\*

Subin Yang\*, Minji Kim and Young-hak Lee\*\*

### 요 약

농업 생산성은 유해 조류로 인한 피해로 큰 영향을 받을 수 있으며, 이러한 문제를 해결하기 위해 빠르고 정확한 탐지 시스템이 필요하다. 본 연구는 실시간 객체 검출 분야에서 널리 쓰이는 YOLOv5 모델과 성능이 향상된 YOLOv8 모델을 비교하였다. 두 모델은 같은 데이터셋과 학습 환경에서 훈련되었고, 테스트 영상에서의 탐지 결과를 바탕으로 성능을 평가하였다.

### Abstract

In the agricultural field, damage caused by harmful birds has a significant impact on productivity, and the need for a fast and accurate detection system to solve this problem is increasing. In this study, we evaluated the performance of harmful bird detection by comparing the YOLOv5 model, which is widely used for real-time object detection, with the latest YOLOv8 model. The two models were trained under the same dataset and learning conditions, and their performance was analyzed focusing on detection results in test images.

### Key words

agricultural field, harmful birds, detection system, YOLOv5 model, YOLOv8 model, real-time object detection

## I. 서 론

농업에서 유해 조류로 인한 피해는 심각한 문제로, 이를 해결하기 위해 빠르고 정확한 탐지 기술이 필요하다. [1] 실시간 객체 탐지에 널리 쓰이는 YOLO(You Only Look Once) 모델은 높은 정확도와 속도로 주목받고 있으며, 최신 버전인 YOLOv8은 특히 성능이 개선된 것으로 알려져 있다. 본 연구는

YOLOv5와 YOLOv8 모델을 이용해 유해 조류 탐지 성능을 비교하고, 그 실제적 적용 가능성을 분석하는 데 중점을 두었다.

## II. 연구 방법

### 2.1 데이터셋 구성

본 연구는 다양한 유해 조류 이미지와 그 위치

\* 국립안동대학교 컴퓨터공학과, 양수빈, didtnqls9341@naver.com,

\*\* 국립안동대학교 SW융합교육원, 이영학, yhyi@andong.ac.kr(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)



정보를 포함한 데이터셋을 활용하였다. 데이터셋은 학습, 검증, 테스트 용도로 나뉘어 각 모델이 동일한 조건에서 학습될 수 있도록 구성되었다. 이러한 방식은 두 모델의 성능을 공정하게 비교하고, 검출 결과의 차이를 명확히 파악하는 데 도움이 되었다.

## 2.2 모델 학습

YOLOv5와 YOLOv8 모델은 동일한 설정 하에 학습되었다. Google Colab 환경에서 --batch 16 및 --epochs 100의 설정을 사용하여 훈련이 이루어졌으며, 학습 중 기록된 지표는 성능 평가의 기준으로 활용되지 않고 검출 결과 중심의 비교에 초점을 맞추었다. 이러한 설정은 두 모델의 학습 안정성을 유지하면서 실제 검출 성능을 명확히 평가하기 위함이다.

## III. 성능 비교

### 3.1 검출 결과 비교

YOLOv5와 YOLOv8 모델의 성능을 테스트 영상에서 비교한 결과, YOLOv8이 유해 조류 탐지에서 더 뛰어난 정확도를 나타냈다. YOLOv5는 대부분의 조류를 잘 탐지했으나, 작은 크기나 멀리 있는 조류를 놓치거나 낮은 신뢰도로 탐지하는 경우가 있었다. 이는 YOLOv5가 작은 객체를 인식하는 데 있어 제약이 있을 수 있음을 시사한다.

반면, YOLOv8은 작은 객체나 먼 거리에 있는 대상도 높은 신뢰도로 탐지하며, 더욱 정확한 경계 상자를 제공하는 특징을 보였다. 특히, 작은 조류나 먼 거리의 조류를 탐지할 때에도 YOLOv8은 안정적인 성능을 유지했으며, YOLOv5보다 일관된 결과를 보였다. 예시 그림에서는 YOLOv8이 다수의 조류를 높은 신뢰도로 탐지한 반면, YOLOv5는 일부 탐지에 실패하거나 신뢰도가 낮은 결과를 보였다. 이러한 결과는 YOLOv8이 작은 객체나 먼 거리의 탐지에서 더 나은 성능을 발휘한다는 것을 보여준다



그림1. Yolov5의 검출 결과

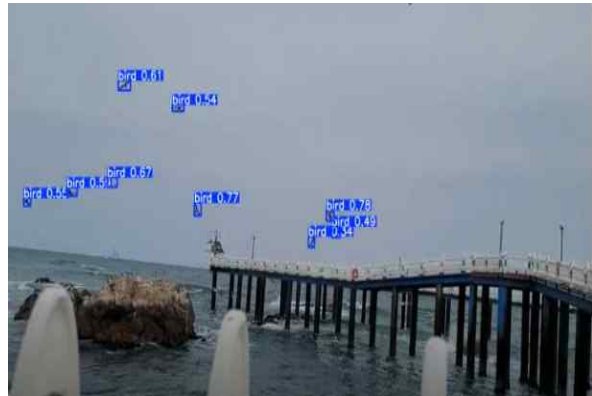


그림2. Yolov8의 검출 결과

## IV. 결론

본 연구를 통해, 유해 조류 탐지에서 YOLOv8 모델이 YOLOv5보다 뛰어난 성능을 보임이 드러났다. YOLOv8은 특히 조류의 크기가 작거나 멀리 떨어진 경우에도 높은 신뢰도로 탐지하며, 위치 정보를 정확하게 제공하는 데 강점을 보였다. 이러한 발견은 YOLOv8이 실시간 조류 탐지 시스템에 효과적으로 사용될 수 있음을 시사한다.

향후 연구에서는 다양한 거리에서의 촬영, 조명 변화, 그리고 여러 환경적 요인을 고려해 두 모델의 성능을 보다 면밀히 비교할 예정이다. 또한, 대규모 데이터셋을 활용해 YOLOv8의 일반화 성능을 심층적으로 평가하고, 실시간 객체 탐지 분야에서의 응용 가능성을 더욱 확대할 계획이다.

참 고 문 헌

- [1] Y. J. Kim, S. M. Hong, H. S. Kim, C. H. Kang, and S. Y. Kim, "Analysis of Detection Performance According to Distance and Size Based on YOLOv8," *Proceedings of the Korean Society of Mechanical Engineers Conference*, pp. 824-825, November 2023.

# AI 모델 상호운용을 위한 메타데이터 구축 방법 연구

고동환\*, 김장원†

## A Study on Metadata Construction Methods for AI Model Interoperability

Donghwan Ko\*, Jangwon Gim†

### 요 약

최근 AI 모델 공유 플랫폼에 등록된 모델 수가 급격히 증가하고 있다. AI 모델 문서는 운영과 관리에 필요한 정보를 제공하지만, 개발사마다 문서화된 키워드와 구조가 달라 상호운용성과 일관성에 어려움이 존재한다. 본 논문은 AI 모델 간 상호운용성을 지원하는 메타데이터 모델링 프로세스를 제안하며, 이를 통해 메타데이터의 표준화와 통합 체계를 구축할 수 있음을 기대한다. 향후에는 제안된 메타데이터 모델을 기반으로 지식그래프를 구축하여 AI 모델 상호운용성을 강화하는 연구를 진행할 예정이다.

### Abstract

The number of AI models registered on model-sharing platforms has been rapidly increasing. While AI model documentation provides essential information for operation and management, inconsistencies in terminology and structure across different developers hinder interoperability and standardization. This paper proposes a metadata modeling process to support interoperability between AI models, aiming to establish a standardized and integrated framework for metadata. Future research will focus on building a knowledge graph based on the proposed metadata model to further enhance AI model interoperability.

### Key words

AI model interoperability, metadata modeling, Ontology

## 1. 서 론

최근 인공지능(AI) 모델 공유 플랫폼에 등록된 모델의 수가 급증하고 있으며 허깅페이스(Hugging Face)는 2024년 9월말 기준 100만 개 이상의 AI 모

델이 등록되고 있다[1]. 또한, PyTorch 허브[2], TensorFlow 허브[3], ONNX 허브[4] 등의 플랫폼에서도 AI 모델의 저장, 공유, 배포를 지원하고 있다. 이러한 플랫폼들은 AI 연구자와 개발자들이 모델의 재사용성을 높이고 공유가 가능한 정보와 라이브러리

\* 국립군산대학교 소프트웨어학과 1901586@kunsan.ac.kr

† 국립군산대학교 소프트웨어학과 (교신저자) jwgim@kunsan.ac.kr

※ 본 연구는 2024년도 과학기술정보통신부 및 정보통신기획평가원의 “SW중심대학사업” 지원을 받아 수행되었음 (2023-0-00065)

등을 제공하여 AI 생태계의 성장을 촉진하고 있다. 그렇지만 수많은 AI 모델의 공개로 인하여 모델들 간의 상호운용성 또한 중요하다. 대표적 허브인 허깅페이스에서는 각 모델별로 AI 모델의 운영 및 관리에 필요한 주요 정보를 제공하는 역할을 한다. 그렇지만 다수의 AI 모델 허브들과 비교하였을 때 제공되는 메타데이터의 정보량이 상이하여 모델의 상호운용을 위한 정보를 생산하고 관리함에 한계가 있다. 따라서 본 논문에서는 AI 모델 간 상호운용성을 지원하기 위해 표준화된 메타데이터 모델링 방법을 제안한다.

## II. 관련 연구

### 2.1 인공지능(AI) 모델 온톨로지

AI 모델과 관련된 다양한 온톨로지들이 개발되었으며, 이들은 특정 도메인에 맞춰 메타데이터 항목을 정의하고 있다. AIO(Artificial Intelligence Ontology)는 주로 딥러닝 모델의 레이어 구성, 뉴럴 네트워크 아키텍처 등 기술적 사항을 정의하는 온톨로지로서, 모델의 구조와 동작 방식을 설명한다[5]. 이를 통해 개발자와 연구자가 모델을 이해하고 최적의 활용 방안을 모색하는 데 기여할 수 있다. AIRO(Artificial Intelligence Risk Ontology)는 EU AI Act 및 ISO/IEC의 가이드라인에 따라 개발된 온톨로지로서, 모델과 관련된 위험 요소를 정의하는 데 중점을 둔다[6]. AIRO는 모델 설계 및 사용 단계에서 발생할 수 있는 윤리적, 법적, 사회적 위험 요소를 규명하여, 모델의 안전성과 신뢰성을 보장하는 데 도움을 준다. MLS(Machine Learning Schema)는 머신러닝 알고리즘의 구조와 특성을 정의하는 온톨로지로서, 학습 과정에서 사용하는 알고리즘 유형과 데이터 처리 방식을 명시한다[7]. MLS는 머신러닝 모델의 학습 방식과 특성에 대한 표준화된 정보를 제공하여 다양한 알고리즘 간 비교 및 분석을 용이하게 한다. RAINS(Realising Accountable Intelligent Systems)는 모델의 생애주기 전반을 관리하는 메타데이터를 제공하는 온톨로지이다[8]. RAINS는 모델의 설계 초기부터 배포 및 유지보수 단계까지 모든 주기적 정

보를 제공하여, 모델의 지속적 관리와 효율적 운영을 지원한다. 이러한 온톨로지들이 각각 특정 도메인(예: 기술적 세부사항, 윤리적·사회적 위험, 머신러닝 알고리즘 특성, 생애주기 관리)을 목표로 하고 있으므로 개별적인 사용은 가능하지만 통합된 프레임워크로 적용하여 AI 모델의 전반적인 메타데이터를 구성하는 시도는 아직 초기 단계에 머물러 있다.

## III. AI 모델 메타데이터 구축 방법

### 3.1 AI 모델 키워드 추출 및 메타데이터 어휘 선정

AI 모델 공유 플랫폼에서 제공하는 문서의 구조를 분석하여 키워드를 추출하고 등장 빈도 정렬 및 표기, 의미상 중복 제거 등 정제 작업을 하였다. AI 모델에 대한 기존 온톨로지(AIO, AIRO, MLS, RAINS 등)와 비교하여 초기 메타데이터 항목을 선정하였다. [표 1]은 카테고리별 메타데이터 주요 항목을 보인다.

정제된 키워드들을 초기 메타데이터 항목과 유사도를 측정하여 메타데이터 항목에 대해 유사한 키워드를 매핑하여 매핑 테이블을 생성하였다. 이러한 매핑 테이블은 신규 AI 모델 설명 페이지의 키워드에 대한 통합을 지원할 수 있다.

LLM은 AI 분야에 대한 이해력을 바탕으로 온톨로지 어휘에 대한 확장, 유지보수, 업데이트를 지원 및 간소화할 수 있다. LLM을 활용한 온톨로지 개발은 새로운 AI 개념에 대한 통합을 도울 수 있다.

초기 메타데이터 항목에 대해 LLM을 활용하여 메타데이터 항목에 대한 구조화 및 필수, 공통, 커스텀 메타데이터를 분류하였다. AI 모델 운영에 필요한 메타데이터를 필수 메타데이터, AI 모델 간 상호운용에 필요한 메타데이터를 공통 메타데이터, 이를 제외한 나머지 메타데이터를 커스텀 메타데이터로 정의하였다. 분류된 메타데이터 항목에 대한 정확도를 수작업으로 검증하고 메타데이터 항목 간 연계를 목적으로 메타데이터 항목 간 관계에 대한 정보를 정의하였다.

표 1. 카테고리별 메타데이터 주요 항목  
Table 1. Key Metadata Items by category

Category	Metadata Items
Basic Information	Identification, Version, Pipeline Tag, ModelArchitecture, ModelTrainingTime, ModelUpdateFrequency
Model Characteristics	Input Type/Output Type, Supported Languages, Library Name
Performance and Evaluation	Performance Metrics(ClassificationMetric, ObjectDetectionMetric, SequenceGenerationMetric), Benchmark Dataset, Inference Speed, ExternalValidation
Dataset	Identification, Dataset Description, Dataset License, Version, Dataset URL, DataDetails(DataSplitMethod, FineTuningDataset, ValidationDataset)
Derived Model Management	BaseModel, AdaptedModel, FinetunedModel, QuantizedModel, MergedModel, Derivative Method, Modification Details
Usage and Restrictions	License Type, Intended Use, Known Limitations, Ethical Considerations
Management Information	Last Updated, Maintainer Contact, Related Documentation, Community Feedback

[그림 1]은 제안 모델에 대한 개요도를 보인다.

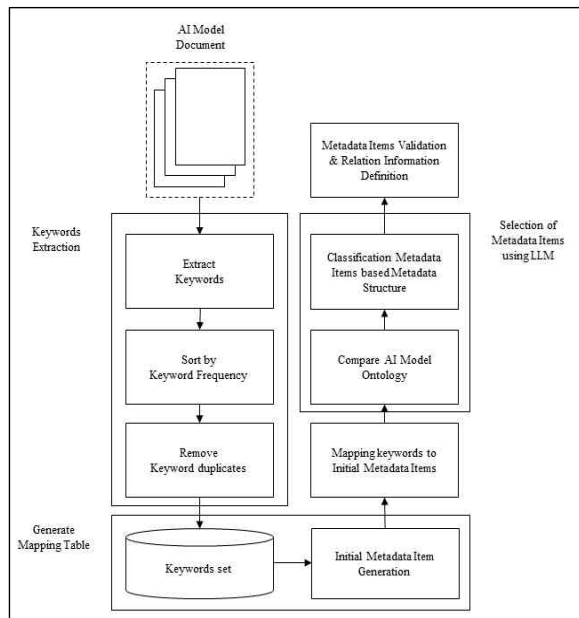


그림 1. 제안 모델 개요도

Fig 1. Proposed Model Overview

능을 제공하여 메타데이터를 표현하는 데 유용하다. 웹 기반 기술 문서나 표준 문서 작성에 사용되며, 웹 기반의 문서를 생성하고 표준화된 형식으로 제공할 수 있도록 돕는다. 따라서, ReSpec을 활용하여 AI 모델 메타데이터 항목에 대한 웹페이지를 구축하였다. [그림 2]은 작성된 메타데이터 항목 사이트의 예시로, 각 항목의 정의 및 설명(예: Label, Domain, Range)을 포함한다.

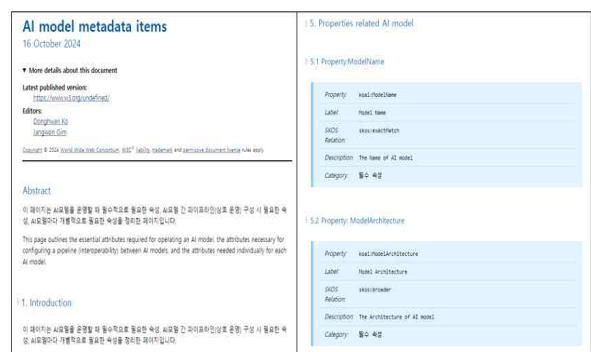


그림 2. 메타데이터 항목 사이트(일부)

Fig 2. Metadata Item Site (Partial)

### 3.2 AI 모델 메타데이터 명세

ReSpec은 W3C(Web Consortium)에서 개발한 도구로, 웹 표준 및 기술 명세 문서 작성을 지원하는 기

### IV. 결론 및 향후 연구

본 논문은 AI 모델 간 상호운용을 위한 메타데이

터 구축 방법을 제안하였다. 제안된 방법을 활용하여 AI 모델 간 상호운용성, AI 모델의 운영 및 관리 효율성 향상이 기대된다. 향후 연구는 메타데이터 모델을 기반으로 지식그래프를 구축하여 AI 모델 간 상호운용에 대한 연구 및 신규 AI 모델에 대한 메타데이터가 자동으로 확장되는 시스템에 대한 연구를 진행할 계획이다.

## 참 고 문 헌

- [1] <https://huggingface.co/>[Accessed: 2024-11-04]
- [2] <https://pytorch.org/hub/>[Accessed: 2024-11-04]
- [3] <https://www.kaggle.com/models?tfhub-redirect=true>[Accessed: 2024-11-04]
- [4] <https://onnx.ai/models/>[Accessed: 2024-11-04]
- [5] Joachimiak, Marcin P., et al. "The Artificial Intelligence Ontology: LLM-assisted construction of AI concept hierarchies." arXiv preprint arXiv:2404.03044 (2024).
- [6] Golpayegani, Delaram, Harshvardhan J. Pandit, and Dave Lewis. "Airo: An ontology for representing ai risks based on the proposed eu ai act and iso risk management standards." Towards a Knowledge-Aware AI. IOS Press, 2022. 51-65.
- [7] Publio, Gustavo Correa, et al. "ML-Schema: An interchangeable format for description of machine learning experiments." Semantic Web 0.0 (2020): 1-11.
- [8] Naja, Iman, et al. "A semantic framework to support AI system accountability and audit." The Semantic Web: 18th International Conference, ESWC 2021, Virtual Event, June 6-10, 2021, Proceedings 18. Springer International Publishing, 2021.

# 효율적인 헤어 이미지 생성을 위한 StyleCLIP 모델의 프롬프트 최적화 기법에 관한 연구

김현지\*, 정유진\*, 유상훈\*, 최영찬\*, 미르잘롤로브미라보스\*, 김재수\*

## A Study on Prompt Optimization Techniques for Efficient Hair Image Generation Using the StyleCLIP Model

Hyeon Ji Kim\*, Yujin Jung\*, Sang Hoon Ryu\*, Young Chan Choi\*, Mirjalolov Mirabbos\*,  
and Jae Soo Kim\*

### 요 약

본 연구는 인공지능 기반의 헤어 이미지 생성 성능을 개선하기 위해 StyleCLIP 모델에 다양한 프롬프트 엔지니어링 기법을 적용하여 그 효과를 분석하였다. 생성된 이미지의 품질을 평가하기 위해 CLIP 모델을 사용해 텍스트-이미지 일관성을 측정하였으며, latent optimization을 통해 최적의 결과를 얻기 위한 추가적인 성능 개선을 진행하였다. 실험 결과, 프롬프트 엔지니어링을 적용한 모든 경우에서 CLIP Loss가 감소하여 이미지의 일관성이 향상됨을 확인하였고, 특히 Chain of Thoughts 기법이 StyleCLIP 모델에 가장 효과적임을 입증하였다. 이러한 결과는 생성형 AI 모델의 성능 향상에 프롬프트 엔지니어링이 중요한 기여를 할 수 있음을 보여주었다.

### Abstract

This study analyzed the effect by applying various prompt engineering techniques to the StyleCLIP model to improve the artificial intelligence-based hair image generation performance. Text-image consistency was measured using the CLIP model to evaluate the quality of the generated image, and additional performance improvement was performed to obtain optimal results through late optimization. As a result of the experiment, it was confirmed that the CLIP loss was reduced in all cases where prompt engineering was applied, improving the consistency of the image, and in particular, the Chain of Thought technique proved to be the most effective for the StyleCLIP model. These results showed that prompt engineering can make an important contribution to the performance improvement of the Generative AI model.

### Key words

Stable Diffusion, StyleGAN, Hair Image Generation, CLIP, Generative AI, Prompt Engineering

---

\* 경북대학교 IT대학 컴퓨터학부,  
{guswlrkgus25, ah12gk3, ryusam012, cyc1215, 2020110372}@knu.ac.kr, kjs@knu.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(2021-0-01082)

## I. 서 론

국내에서 이미지는 지식재산권의 일부분으로 인정되며, 원 저작권자의 승인 없이 이미지를 사용하는 것은 저작권 침해에 해당한다. 그러나 소상공인이 운영하는 미용실, 쇼핑몰 등에서는 저작권을 고려하지 않은 이미지가 빈번하게 사용되고 있다. 이는 소상공인이 카탈로그나 안내 페이지와 같은 홍보물에 사용할 사진을 위해 큰 비용을 지출하기 어려운 현실에 기인한다.

최근에는 DALL-E, Stable Diffusion, Adobe Firefly와 같은 이미지 생성형 인공지능 모델들이 등장하면서 주목받고 있다. 이러한 모델들은 디자이너와 모델 없이도 텍스트나 카테고리 선택과 같은 간단한 입력만으로 다양한 이미지를 생성할 수 있어 비용과 시간을 획기적으로 절약할 수 있다. 또한 저작권으로부터 자유롭다는 점에서 매력적이다.

한편, 이미지 생성형 인공지능에 맞는 프롬프트 엔지니어링 연구는 적은 실정이다. 이에 본 연구는 인공지능에 텍스트 데이터를 입력하여 사용자의 의도에 맞는 이미지를 생성할 수 있는 프롬프트 엔지니어링 기법을 제안하고자 한다. 이를 위해 distribution quality metric에서 최적의 성능을 보이는 StyleGAN 모델에 여러 프롬프트 엔지니어링 기법을 적용하여 이미지를 생성하고, 생성된 이미지의 적합성을 Contrastive Language-Image Pretraining(이하 CLIP) 모델의 Loss 비교를 통해 평가하여 제안 기법의 유효성을 입증하였다.

## II. 본 론

### 2.1 프롬프트 구성

프롬프트 엔지니어링은 생성형 AI 모델의 출력을 조정하여 사용자 요구에 맞는 결과물을 생성하는데 중요한 역할을 한다[1, 2]. 본 연구에서는 다음과 같은 프롬프트 엔지니어링 기법을 적용하여 원문 문장을 다양한 방식으로 변형하였다.

- **Weighted Terms**

특정 단어나 개념에 가중치를 부여하여 모델이

이를 강조하도록 유도하는 방법이다. 예를 들어, 이미지 생성 시 ‘빨간색:2’와 같이 가중치를 주어 특정 색상에 집중하게 할 수 있다.

- **Quality Boosters**

“high quality,” “ultra-realistic” 등과 같은 표현을 추가해 모델이 더 고품질의 결과물을 생성하도록 유도하는 방법이다. 세부 묘사나 사실성을 강화하는데 유용하다.

- **Chain of Thoughts**

모델이 논리적 사고를 할 수 있도록 단계별 지시를 주는 기법이다. 예를 들어, “첫째로 ~, 그 다음에 ~”와 같은 방식으로 문제 해결에 논리적 접근을 유도할 수 있다.

- **Fix Deformed Generations**

인물의 비율이 비정상적이거나 문법적 오류가 있는 출력을 수정하기 위해 프롬프트에 “anatomically correct” 등의 문구를 추가하는 방법이다. 이를 통해 비정상적인 출력을 줄이고 자연스러운 결과를 생성할 수 있다.

위와 같은 다양한 프롬프트들을 헤어 이미지 생성에 필요한 "Long Blonde Curly Casual hairstyle for Oval faces"라는 문장에 적용하였고, 다음과 같은 문장의 변화값들을 얻어냈다.

표 1. 프롬프트 유형에 따른 적용 문장 비교  
Table 1. Comparison of Applied sentences by various prompt types

Prompt	Applied Sentence
Original (no Apply)	Long Blonde Curly Casual hairstyle for Oval faces
(a) Weighted Terms	Long::10 Blonde::1.5 Curly::5 Casual hairstyle::2 suited for Oval faces::1.5
(b) Quality Boosters	high-quality luxurious Long detailed Blonde elaborate Curly Casual Hairstyle for Oval Face
(c) Chain of Thoughts	Imagine a person with an oval face shape. The hairstyle should be long and casual, featuring blonde, curly hair. The curls add a natural, relaxed look to the overall style



(d) Fix Deformed Generations	Create a long blonde curly hairstyle in a casual style for oval faces. Avoid overly formal or structured styling. Exclude straight or short hairstyles.
------------------------------	---

## 2.2 평가 지표

### 1) CLIP 지표 활용

CLIP 점수는 이미지와 텍스트 설명 간의 일치도를 측정한다[3, 4]. CLIP은 이미지와 텍스트를 임베딩 공간으로 변환한 후, 코사인 유사도를 계산하여 텍스트-이미지 일관성을 평가한다. 특히 생성된 이미지에서 머리카락과 같은 세부 사항이 프롬프트와 얼마나 일치하는지를 측정하는 데 유용하다[4].

### 2) latent optimization

source latent code  $w_s \in W+$  와 텍스트 프롬프트가 주어지면 다음 최적화 문제를 해결하는 방식이다[4].

$$\arg \min_{w \in W+} D_{CLIP}(G(w), t) + \lambda_{L2} \|w - w_s\|_2 + \lambda_{ID} \mathcal{L}_{ID}(w), \quad (1)$$

이때, G는 사전 훈련된 StyleGAN2 생성기이고  $D_{CLIP}$ 는 두 인수의 CLIP 임베딩 사이의 코사인 거리다. 입력 이미지와의 유사성은 잠재 공간의  $L_2$  거리와 identity loss에 의해 제어된다.

$$L_{ID}(w) = 1 - \langle R(G(w_s)), R(G(w)) \rangle \quad (2)$$

여기서 R은 얼굴 인식을 위해 미리 훈련된 ArcFace 네트워크이고, 인수 간의 코사인 유사도를 계산한다. G와 CLIP 이미지 인코더를 통해 수식(1)의 목표 기울기를 역전파함으로써 기울기 하강을 통해 최적화 문제를 해결한다.

## 2.3 StyleCLIP 모델 기반 실험 진행

본 실험에서는 StyleCLIP 모델에 FFHQ(Flickr Faces High Quality) 데이터셋을 사용해, 원본 문장 (Original)과 7개의 프롬프트 변형 문장을 각각 입력하여 프롬프트 당 150회 최적화 반복을 수행하고 손실 값을 도출하였다. 이를 10회 반복해 총 1500회

의 실험을 통해 각 프롬프트 유형별 손실 값을 비교하였다.

## 2.4 실험 결과 분석

본 실험에서는 StyleCLIP 모델을 사용하여 프롬프트 엔지니어링을 통한 이미지 생성 성능을 평가하였다.

실험 결과(표 2), 모든 경우에서 프롬프트 엔지니어링을 적용한 결과가 적용하지 않은 결과보다 loss가 작으므로 프롬프트 엔지니어링을 활용하는 것이 오류 이미지를 생성할 확률이 낮음을 확인하였다.

표 2. 프롬프트 엔지니어링 적용 전 Loss 값 통계  
Table 2. Statistics of Loss values before applying prompt engineering

	Loss 평균	Loss 증감값	표준 편차
(a)	0.7015	0.691	0.016
(b)	0.6591	0.66	0.0039
(c)	0.7015	0.691	0.0196
(d)	0.6591	0.66	0.0039

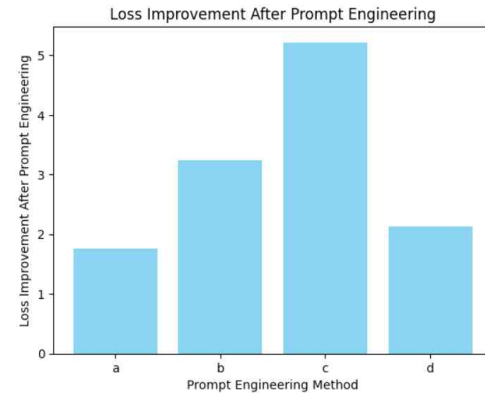


그림 1. 프롬프트 엔지니어링 적용 전후 CLIP Loss 비교((a) Weighted Terms, (b) Quality Boosters, (c) Chain of Thoughts, (d) Fix Deformed Generations)

Figure 1. Comparison of CLIP Loss before and after application of prompt engineering

표 3. 프롬프트 엔지니어링 적용 후 Loss 값 통계  
Table 3. Statistics of Loss values after applying prompt engineering

	Loss 평균	Loss 증감값	표준 편차
(a)	0.689	0.689	0.0029
(b)	0.638	0.639	0.0026
(c)	0.665	0.665	0.0036
(d)	0.645	0.644	0.0017

이 실험을 통해 이미지 생성형 AI에 특화된 효율적인 프롬프트가 존재하고 Chain of Thoughts 방식이 styleCLIP 모델에 적합함을 입증하였다.

아래의 그래프는 프롬프트 엔지니어링을 적용하지 않은 원본 문장과 프롬프트 엔지니어링을 적용한 7개의 문장을 입력하여 실험한 CLIP Loss 값을 나타낸다. 원본 문장에 비해 프롬프트 엔지니어링을 적용했을 때 Loss가 전반적으로 더 낮은 것을 확인할 수 있다.

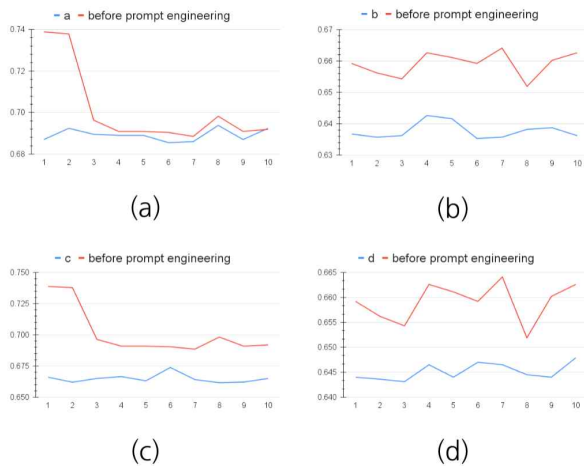


그림 2. 프롬프트 적용 전후의 CLIP Loss 비교  
Figure 2. Comparison of CLIP Loss before and after application of prompt engineering

### III. 결 론

본 연구에서는 StyleCLIP 모델에 다양한 프롬프트 엔지니어링 기법을 적용하여 헤어 이미지 생성 성능을 개선하고자 하였다. 실험 결과, 프롬프트 엔지니어링을 적용한 경우가 그렇지 않은 경우보다 CLIP Loss가 전반적으로 낮아, 텍스트-이미지 일관성이 더욱 높아짐을 확인할 수 있었다. 이는 프롬프트 엔지니어링이 이미지 생성형 AI 모델의 성능 향상에 중요한 역할을 한다는 것을 시사한다.

본 연구는 AI 기반 이미지 생성 작업에서 프롬프트 엔지니어링의 중요성을 강조하며, 이를 통해 이미지의 품질과 일관성을 높일 수 있음을 입증하였다. 나아가 프롬프트 엔지니어링의 세부 기법들이 서로 다른 성능 향상을 제공할 수 있다는 점에서, 생성 작업의 목적에 따라 최적의 프롬프트 설계를 위한 연구가 지속적으로 필요함을 시사한다.

### 참 고 문 헌

- [1] Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed Chi, Quoc Le, and Denny Zhou, published in NeurIPS 2022.
- [2] Prompt Design and Engineering: Introduction and Advanced Methods, Xavier Amatriain, gives an in-depth overview of core and advanced prompt engineering techniques, Xavier Amatriain, 2024
- [3] Towards Counterfactual Image Manipulation via CLIP" Yingchen Yu, Fangneng Zhan, Rongliang Wu, Jiahui Zhang, Shijian Lu, Miaomiao Cui, Xuansong Xie, Xian-Sheng Hua, Chunyan Miao, 2022
- [4] StyleCLIP: Text-Driven Manipulation of StyleGAN Imagery, Or Patashnik, Zongze Wu, Eli Shechtman, Daniel Cohen-Or, Dani Lischinski, Hebrew University of Jerusalem, Tel-Aviv University, Adobe Research, 2022, NIPS

# 인공지능 기술을 활용한 에너지 사용량 예측 웹 어플리케이션 설계 및 구현

배민석\*, 명세민\*, 이승민\*, 유주오\*, 권순용\*\*, 고석주\*\*\*

## Design and Implementation of a Web Application for Energy Consumption Measurement Using Artificial Intelligence

Minseok Bae\*, Semin Myeong\*, Seungmin Lee\*, Juo Yu\*, Soonyong Kwon\*\*,  
and Sukjoo Koh\*\*\*

### 요 약

본 연구는 기업의 에너지 소비 패턴을 분석하고 최적화하기 위한 FEMS(Factory Energy Management System) 솔루션 개발을 목적으로 한다. 기존의 에너지 관리 시스템은 에너지 소비 데이터를 모니터링하고 분석하는데 그쳤으나, 본 연구는 인공지능 기술을 도입해 에너지 소비 패턴을 예측하고, 이를 바탕으로 사용자가 실시간으로 예측 데이터를 확인할 수 있도록 한다. 본 시스템은 웹 서버 환경에서 안정적이고 빠른 데이터 처리를 지원하며, 사용자는 웹 인터페이스를 통해 시각적으로 예측된 에너지 사용량을 확인할 수 있다.

### Abstract

This study aims to develop a Factory Energy Management System (FEMS) solution to analyze and optimize energy consumption patterns within companies. While traditional energy management systems have primarily focused on monitoring and analyzing energy consumption data, this research integrates machine learning algorithms to predict energy usage patterns. Based on these predictions, users can access real-time forecast data, enabling more proactive management. The system supports stable and fast data processing in a web server environment, allowing users to visually confirm the predicted energy usage through a web interface.

### Key words

Factory Energy Management System(FEMS), Artificial intelligence

---

\* 경북대학교 컴퓨터학부, {otuera1123, msemin009, milouw56, juoyu78}@gmail.com

\*\* 넥사디에스, sykwon@nexa-ds.com

\*\*\* 경북대학교 컴퓨터학부 교수(교신저자), sjkoh@knu.ac.kr

※ 이 연구는 과학기술정보통신부 및 정보통신산업진흥원의 “지역산업 SW인재양성 기반조성 사업” 지원으로 수행되었음.

※ 이 연구는 과학기술정보통신부의 재원으로 한국지능정보사회진흥원의 지원을 받아 구축된 "전력신사업을 위한 전라남도 지역 전력소비패턴 데이터"를 활용하여 수행된 연구입니다. 본 연구에 활용된 데이터는 AI 허브(aihub.or.kr)에서 다운로드 받으실 수 있습니다.

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(2021-0-01082)

## I. 서론

데이터 기반 의사결정은 현대 산업에서 필수적인 요소로 자리 잡고 있으며, 특히 에너지 관리 분야에서는 데이터를 활용하여 에너지 소비를 최적화하는 것이 필요하다. 제조업에서 에너지 소비를 효율적으로 관리하는 것은 비용 절감뿐만 아니라 환경 보호 측면에서도 매우 중요하다. 이를 위해 공장 에너지 관리 시스템(FEMS, Factory Energy Management System)은 에너지 사용 데이터를 모니터링하고 분석하여 최적화된 에너지 사용 방안을 제시하는 역할을 한다. 그러나 기존 시스템은 단순히 과거 에너지 사용량에 대한 모니터링과 데이터 분석에 그쳐, 에너지 사용 패턴을 예측하거나 실시간으로 결과를 사용자에게 제공하는 데는 한계가 있다.

본 연구는 인공지능을 활용하여 에너지 소비 패턴을 예측하고, 이를 웹 서비스를 통해 사용자에게 실시간으로 제공하는 시스템을 제안한다. 인공지능 기술은 대규모 데이터를 분석하고 복잡한 패턴을 예측할 수 있는 강력한 도구로, 에너지 사용량의 변동성을 예측하는 데 적합하다. 이를 통해 사용자들은 에너지 소비에 대한 예측 데이터를 기반으로 운영 전략을 수립하고, 불필요한 에너지 낭비를 줄일 수 있다.

본 논문에서는 수집된 에너지 소비 데이터를 통해 실시간으로 예측된 결과를 웹 서비스 형태로 제공하는 시스템을 설계하고 구현한다. 이 시스템은 웹 서버를 기반으로 구축되며, Node.js와 Express를 활용하여 안정적이고 효율적인 데이터 처리를 보장한다. 사용자는 웹 인터페이스를 통해 예측된 데이터를 시각적으로 확인하고, 그에 따라 효율적인 에너지 관리 결정을 내릴 수 있다.

본 연구의 목적은 사용자에게 예측된 에너지 소비 정보를 실시간으로 제공함으로써, 에너지 관리 효율성을 극대화하는 것이다. 이를 통해 제조업체는 예측된 데이터를 바탕으로 더 나은 의사결정을 내릴 수 있으며, 장기적으로는 비용 절감과 환경 보호에 기여할 수 있다. 본 연구는 웹 서비스와 인공지능을 결합한 새로운 FEMS 솔루션의 가능성을 제시하며, 실시간 데이터 기반의 예측 서비스를 통해 제

조업의 에너지 관리 방식을 개선하고자 한다.

## II. 시스템 설계 및 구현

본 논문에서는 효율적인 개발을 위해 React와 Node.js, Express를 활용해 인공지능을 기반으로 한 웹 애플리케이션을 설계한다. 이 애플리케이션은 React를 사용하여 프론트엔드를 구현하고, Axios와 React-Query를 통해 백엔드와의 API 통신을 관리하며, 예측된 에너지 소비 데이터를 사용자에게 실시간으로 제공한다.

백엔드 서버는 2개로 분리되어 다른 역할을 수행한다. 첫 번째 백엔드 서버는 Node.js와 Express로 구축되었으며, 사용자가 요청한 데이터를 처리하고 기존의 에너지 소비 데이터를 관리하는 역할을 한다. 주로 사용자에게 에너지 소비 현황과 통계를 제공하는데 집중한다. 두 번째 백엔드 서버는 AI모델과 연결된 서버로 인공지능 모델이 예측한 결과를 웹 서버와 통신을 통해 주고받는다. 이러한 방식으로 서버를 목적에 따라 두 가지로 구분하여 분업화함으로써, 시스템의 성능을 최적화하고, 유지 관리를 용이하게 할 수 있다.

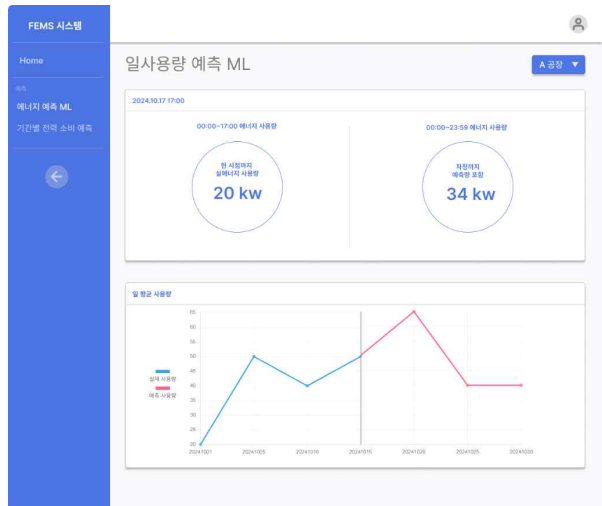


그림 1. 일사용량 예측 페이지

Fig. 1. User Interface for Daily Consumption Prediction

API는 RESTful로 설계되어 있으며, 과거 에너지 사용량 데이터와 미래 에너지 예측 결과는 API 엔드포인트를 통해 프론트엔드에 전달된다. 사용자는

예측된 에너지 소비량을 웹 인터페이스를 통해 확인할 수 있으며, React-Query를 사용하여 API에서 받아온 데이터를 처리하고, 효율적으로 상태 관리를 수행한다.

예측된 에너지 데이터는 데이터 시각화 라이브러리를 통해 표나 그래프 등으로 표시되며, 이를 통해 사용자는 직관적으로 과거의 에너지 소비량과 미래 사용량을 파악할 수 있다. 또한, 인증 및 권한 관리 기능을 통해 로그인한 사용자만 자신의 예측 데이터를 확인할 수 있도록 설정되었다.

### III. 인공지능 모델 선정 및 검증

최근에는 인공지능을 통해 에너지 소모량과 같은 시계열 데이터를 예측하고자 하는 많은 연구가 이루어지고 있다[1,2,3]. 이러한 연구들은 대부분 기존의 벤치마크 데이터 셋[4,5,6]을 통한 검증 방식을 사용한다. 하지만 이러한 벤치마크 데이터 셋은 제조업 등 기업체의 에너지 사용량에 국한된 데이터가 아니며, 환경 또한 한국을 배경으로 하지 않는다. 따라서, 국내 FEMS 시스템에 기존 연구 결과를 맹신하고 성능이 좋은 모델을 바로 적용하기에는 무리가 있을 수 있다. 이러한 맥락에서, 우리는 기존의 모델들을 한국 전라남도에 위치한 제조업의 전력 사용량 데이터를 사용하여 재검증하고, 이를 토대로 모델을 선택한다. 검증의 대상이 되는 모델은 DLinear, NLinear[1], Autoformer[2], TimeXer[3] 4가지이며, 제조업의 전력 사용량 데이터의 경우 AI hub[7]에서 제공하는 “전력신사업을 위한 전라남도 지역 전력소비패턴 데이터” 중 제조업에 속하는 5개 기업의 데이터를 사용하였다. 표 1은 실험에 사용된 데이터의 특성을 나타낸다.

표 1. 실험에 사용된 데이터의 특성  
Table 1. Characteristics of data used in experiments

기업	지역	시간 간격	기간
제조업 A	순천시	1시간	2020.09.01. ~ 2022.08.31.
제조업 B	나주시		
제조업 C	목포시		
제조업 D	여수시		
제조업 E	광양시		

### IV. 실험

4가지 기존 모델을 평가하기 위해, 본 논문에서는 기존 연구를 따라[1] Mean Squared Error (MSE)와 Mean Absolute Error (MAE) 두 지표를 사용한다. 모델이 예측한 전력 값을  $\hat{y}$ , 실제 전력 값을  $y$ , 평가용 데이터의 수를  $n$ 이라고 할 때, MSE와 MAE를 수식으로 나타내면 다음과 같다:

$$MSE = \frac{\sum^n (y - \hat{y})^2}{n} \quad (1)$$

$$MAE = \frac{\sum^n |y - \hat{y}|}{n} \quad (2)$$

표 2는 각 모델에 대한 각 데이터의 MSE, MAE 값을 나타내며, 실험 결과 TimeXer의 성능이 가장 높게 나타나 해당 모델을 인공지능 모델로 선정하였다.

표 2. 실험에 사용된 데이터의 특성  
Table 2. Characteristics of data used in experiments

Model	NLinear[1]		DLinear[1]		Autoformer[2]		TimeXer[3]	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
제조업 A	0.142	0.259	0.144	0.262	0.229	0.349	0.166	0.262
제조업 B	1.022	0.604	0.940	0.557	0.761	0.554	0.437	0.317
제조업 C	0.391	0.466	0.358	0.440	0.386	0.465	0.324	0.393
제조업 D	0.317	0.388	0.317	0.388	0.411	0.458	0.285	0.366
제조업 E	1.003	0.662	1.076	0.682	0.581	0.477	0.667	0.523
평균	0.575	0.476	0.567	0.466	0.474	0.461	0.376	0.372

### V. 결론

본 논문에서는 인공지능을 활용하여 제조업에서 에너지 소비 패턴을 예측하고, 이를 실시간으로 제공하는 웹 애플리케이션을 설계하고 구현하였다. 제안된 시스템은 데이터 수집과 이를 통한 예측, 그리

고 예측 결과를 사용자에게 제공하는 웹 서비스를 통합한 형태이다. 또한, 본 논문은 에너지와 같은 시계열 데이터를 예측하는 모델의 유효성을 한국 지역 제조업의 실제 전력 데이터를 통해 검증하였다. 적절한 인공지능 모델을 사용하여 FEMS 서비스를 구축한다면 사용자는 미래 에너지 소비량 예측을 통해 주도적으로 이에 대응할 수 있을 것이며, 에너지의 현명한 소비를 통해 비용 절감이나 환경 보호와 같은 가치를 실현할 수 있을 것이다.

forecasting”, In The Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021  
[7] ai hub, <https://www.aihub.or.kr/>

### 참 고 문 헌

- [1] A. Zeng, M. Chen, L. Zhang, and Q. Xu, “Are transformers effective for time series forecasting?”, arXiv preprint arXiv:2205.13504, 2022.
- [2] H. Wu, J. Xu, J. Wang, and M. Long, “Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting,” Advances in Neural Information Processing Systems, vol. 34, pp. 22 419–22 430, 2021.
- [3] Y. Wang, H. Wu, J. Dong, Y. Liu, Y. Qiu, H. Zhang, J. Wang, and M. Long, “TimeXer: Empowering Transformers for Time Series Forecasting with Exogenous Variables”, arXiv preprint arXiv:2402.19072
- [4] R. A. Angryk, P. C. Martens, B. Aydin, D. J. Kempton, S. S. Mahajan, S. Basodi, A. Ahmadzadeh, X. Cai, S. F. Boubrahimi, S. M. Hamdi, M. A. Schuh, and M. K. Georgoulis, “Multivariate time series dataset for space weather data analytics”, Scientific Data, 2020.
- [5] G. Lai, W. Chang, Y. Yang, and H. Liu, “Modeling long- and short-term temporal patterns with deep neuralnetworks”, international acm sigir conference on research “and” development in information retrieval, 2017
- [6] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, “Informer: Beyond efficient transformer for long sequence time-series

# 언론 기사 작성 자동화를 위한 AI 기반 제로샷 및 퓨샷 학습 기법 적용 연구

곽시현\*, 백슬기\*, 전희연\*, 신연순\*\*

## A Study on Applying AI-Based Zero-Shot and Few-Shot Learning Techniques for Automation of News Article Writing

Siheon Kwak\*, Seulgi Baek\*, Heeyeon Jeon\*, Youn-Soon Shin\*\*

### 요 약

본 논문은 언론사의 특정 스타일과 구조를 반영한 AI 기반 기사 생성 방안을 제안한다. 제로샷 및 퓨샷 학습을 통해 AI 모델이 주어진 카테고리 및 형식에 맞춰 기사를 생성하며, 신입 기자가 기사 작성 능력을 빠르게 향상시킬 수 있도록 돕는다. GPT-4와 Chain of Thought(CoT) 프롬프팅 기법을 활용해 GPT 4.0 모델의 효과를 평가하였다. 이 연구는 언론사의 자동화된 기사 작성에 기여할 수 있을 것으로 기대된다.

### Abstract

This paper proposes an AI-based article generation approach that reflects the specific style and structure of news agencies. Through zero-shot and few-shot learning, the AI model generates articles tailored to given categories and formats, assisting junior reporters in rapidly improving their writing skills. The effectiveness of the AI model was evaluated using GPT-4 and Chain of Thought (CoT) prompting techniques. This research is expected to contribute to the automation of article writing in the media industry.

### Key words

Large Language Model (LLM), Prompt Engineering, Chain-of-Thought Prompting(CoT), In-Context Learning, Few-Shot Learning, Zero-Shot Learning

## 1. 서 론

현대 언론 환경은 신속하고 광범위한 정보 제공을 필수적으로 요구한다[1]. 숙련된 기자들은 정보

수집 및 가공, 기사 작성 및 검토를 신속하게 수행할 수 있지만, 신입 기자들은 기사 작성의 기본을 익히는 데 어려움을 겪는 경우가 많다. 이에 따라 많은 언론사들이 기사 작성과 편집 과정을 자동화

\* 동국대학교 컴퓨터공학전공, {kwakshi1025, seulgi.baek, hee0713y} @ dgu.ac.kr,

\*\* 동국대학교 AI소프트웨어융합학부(교신저자), ysshin@dongguk.edu

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음 (2023-0-00049)

하기 위해 인공지능(AI) 기술을 도입하고 있다.

기존의 기사 생성 도구는 일반적인 텍스트 생성에 중점을 두고 있으나, 언론사의 스타일이나 구체적인 기사 구성 방식을 충족시키지 못하는 한계가 있다[2]. 다양한 형식의 기사 작성 요구를 만족시키지 못해 카테고리별 특성에 맞는 기사 생성 기술 개발의 필요성이 대두되었다. 이를 위해 본 논문은 Zero-shot learning 및 Few-shot learning 방법을 활용하여, AI 모델이 주어진 카테고리과 형식을 정확하게 반영하도록 기사를 작성하게 한다.

본 논문의 목적은 GPT-4과 Chain of Thought(CoT) 프롬프팅 기법을 활용하여 Zero-shot learning 및 Few-shot learning에서 신입 기자를 지원할 수 있는 AI 기반 기사 생성 기술의 효과를 평가하는 것이다[3]. 이를 달성하기 위해 본 연구는 특정 내용과 형식을 반영한 기사 생성 방법론을 제안한다.

## II. 시스템 설계

### 2.1 데이터 수집

본 논문에서는 헬스경향 웹사이트에서 '건강정보', '뷰티', '뉴스 정책' 카테고리로 분류된 기사 데이터를 크롤링하여 수집하였다. 수집된 데이터는 각각 건강정보에서 10,400건, 뷰티에서 4,900건, 뉴스 정책에서 4,500건으로, 총 19,800건의 기사를 분석 대상으로 하였다.

### 2.2 데이터 분류 체계 설명

본 연구에서 수집된 기사 데이터는 헬스경향의 기존 카테고리 시스템을 기반으로 하여 '상위-중위-하위-소분류'의 계층적 구조로 추가 분석을 통해 분류되었다. 이를 확장하고 세분화하기 위해 기사의 제목만을 크롤링하여 GPT-4를 이용한 클러스터링을 진행한 결과, 유사한 주제를 가진 기사들이 그룹으로 묶였다. 이러한 결과를 토대로 주요 주제를 선정하고, 전체 기사에 대한 레이블링을 수행하였다.

이러한 분석을 통해 각 카테고리에 특화된 내용

별로 요구되는 정보 요소를 식별하고, 이를 바탕으로 기사의 '내용 형식'을 정립하였다. 정립된 내용 형식은 취재자료나 보도자료를 재구성하는 기준으로 활용되었으며, 이 과정에서 생성된 재구성된 자료를 '포함내용정리본'으로 명명하였다. 표 1은 이러한 기사 내용에 따른 분류 체계를 나타낸다.

표 1. 내용 분석  
Table 1. Content Analysis

상위	중위	하위	소분류
건강 정보	건강 일반	연구 결과	-
		질병 정보	-
		정보 전달	-
	먹거리 건강	음식 및 식재료	-
		정보 전달	-
	한방	연구 결과	-
		질병 정보	-
		음식 및 식재료	-
		한의원 홍보	-
		정보 전달	-
	헬스 신간	-	-
뷰티	성형	-	-
	피부 미용	-	-
	화장품	-	-
뉴스	정책	국회의원	개최
		단체 소식	협력협약
			수상소식
	홍보		

### 2.3 내용 구성 및 형식 분석

본 연구에서는 헬스경향의 요구사항을 반영하여 다양한 문체, 구성 방식, 스타일을 채택하였다. 문체는 공식적이고 간결한 문체와 친근하고 부드러운 문체로 구분되며, 구성 방식에는 피라미드형, 역피라미드형, 다이아몬드형, 혼합형이 포함된다. 또한, 서술 방식인 스타일은 서술형, Q&A형, 핵심 포인트형으로 나뉜다.

이러한 형식적 접근은 기사 제목과 보도자료를 바탕으로, 상위분류부터 소분류에 이르는 내용 카테고리과 형식을 선택하는 과정에서 중요하게 활용된다. 선택된 내용에 맞게 GPT-4은 보도자료를 재구성하고, 필요한 내용이 누락된 경우 빈칸으로 처리



하여 재구성 자료를 생성한다. 이 과정에서 기사의 문체, 구성, 스타일을 포함하여 철저한 분석이 이루어진다. 이러한 분석 결과를 통해 최종적으로 기사 형식에 따른 분류 체계가 명확히 정리되며, 이는 표 2에서 상세히 나타난다.

표 2. 형식 분석  
Table 2. Format Analysis

대분류	소분류
문체	딱딱한 문체의 기사
	부드러운 문체의 기사
구성	역피라미드형 구성
	피라미드형 구성
	다이아몬드형 구성
	혼합형 구성
스타일	서술형
	Q&A형
	핵심 포인트형

#### 2.4 기사 생성 도구의 매커니즘

본 연구는 헬스경향 기자가 작성한 것처럼 기사를 생성할 수 있는 도구를 구현하기 위해 GPT-4 모델을 기반으로 한 기사 생성 시스템을 설계하였다. 이 시스템은 기자들의 기사 생성 도구의 매커니즘을 모사하여 다음과 같은 단계를 거쳐 기사를 작성한다.

첫 번째 단계는 내용 선별로, GPT-4 모델에 취재 자료나 보도자료를 입력하여 기사에 반영할 중요한 내용을 선별한 포함내용정리본을 생성하는 과정이다. 이때 각 기사 유형에 맞는 내용 형식을 기반으로 프롬프트를 제공하여, 모델이 핵심 정보만을 추출할 수 있도록 한다.

두 번째 단계는 기사 초안 작성 단계로, 생성된 포함내용정리본을 바탕으로 기사의 형식(문체, 구성, 스타일)에 맞춰 초안을 작성한다. GPT-4 모델은 제공된 형식에 따라 문장과 구조를 구성하여 기사의 초기 초안을 생성한다.

세 번째 단계는 최종 기사 생성 단계로, 작성된 기사 초안에 헬스경향이 제시한 표기 규칙을 반영하고 최종 기사를 완성한다.

### III. 시스템 실험

#### 3.1 실험 설계

본 실험의 목적은 설계된 기사 생성 도구의 매커니즘 각 단계가 기사 작성 과정에 어떻게 기여하는지 평가하는 것이다. 시스템 설계에서 정의된 단계별 매커니즘을 기반으로, 다양한 학습 기법과 입력 조건을 설정하여 성능을 비교 분석한다. 이를 위해 대조군, Zero-shot CoT, Few-shot learning의 세 가지 실험군을 설정하였다.

대조군은 GPT-4에 예시 데이터를 제공하지 않고, 주어진 재구성 자료만을 입력하여 기사를 작성하도록 설정한다. 이를 통해 GPT 4.0이 독립적으로 형식을 추론하고 기사를 작성하는 능력을 평가한다. 그림 1은 이러한 대조군의 실험 과정을 시각화하여 나타낸 것이다.

Zero-shot CoT 실험군은 GPT-4에 기사 제목과 보도자료를 제공하여, 이 자료만을 사용해 필요한 내용을 재구성하고 기사를 작성하도록 유도하는 방식이다. 이때, 필요한 정보가 부족한 경우 해당 부분을 빈칸으로 두어 정보 누락을 방지하며, 특정 기사 형식의 구성 방식을 모델이 스스로 분석하여 적용하도록 한다. 이 실험군은 형식적 지침 없이도 모델이 문서 구조를 형성하는 능력을 평가하는 데 중점을 둔다. 그림 2은 이러한 zero shot CoT의 실험 과정을 시각화하여 나타낸 것이다.

Few-shot learning 실험군은 GPT-4에 예시 데이터 쌍을 제공하여, 모델이 학습한 형식과 내용을 기사의 작성에 보다 정확하게 반영할 수 있도록 한다. 이를 통해 GPT-4가 학습한 특정 형식에 맞춰 기사를 생성하는 능력을 평가하며, 이 실험군에서 생성된 기사의 질적 수준을 분석한다. 그림 3은 이러한 few shot learning의 실험 과정을 시각화하여 나타낸 것이다.

#### 3.2 평가 방법

각 실험군에서 생성된 기사는 헬스경향 기자가 작성한 원본 기사와 비교하여 평가된다. 본 연구에

서는 BERTScore를 사용하여 생성된 기사와 원본 기사 간의 유사도를 측정하며, BERTScore는 문장 간 의미적 유사성을 평가하기 위해 사전 학습된 BERT 모델의 임베딩을 활용하는 척도이다[4]. 이를 통해 GPT-4가 생성한 기사가 원본 기사의 문맥적 정확성과 의미 일관성을 얼마나 잘 반영하는지를 평가한다.

실험 설계에서 정의된 대조군, Zero-shot CoT, Few-shot learning 각 실험군에서의 결과를 비교하여, GPT-4의 특정 형식에 맞춘 기사 작성 능력을 평가하고, 각 학습 기법이 성능에 미치는 영향을 분석한다.

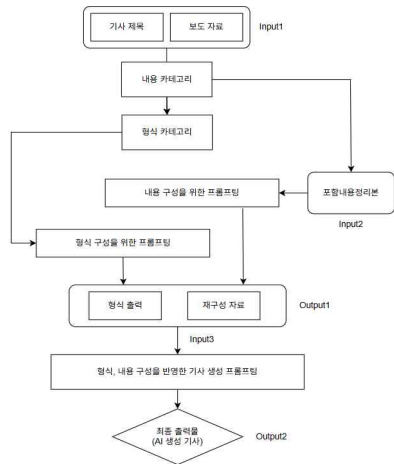


그림 1. 대조군  
Fig. 1. control group

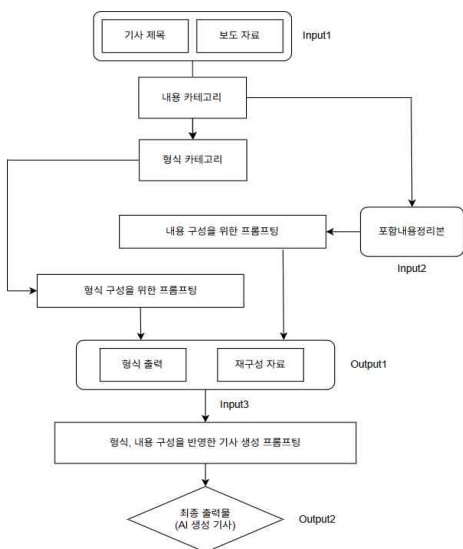


그림 2. 제로 샷 학습  
Fig. 2. zero shot CoT

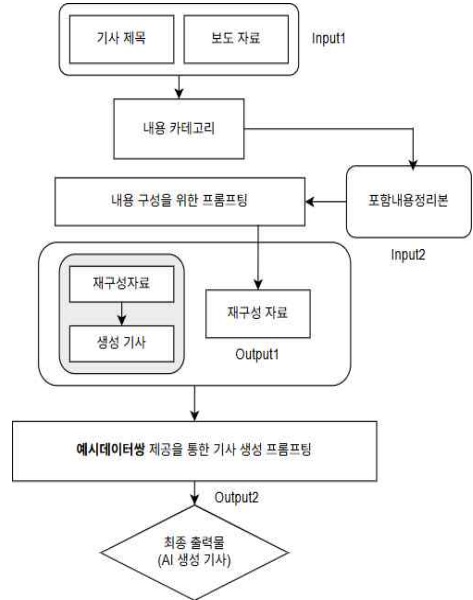


그림 3. 퓨 샷 학습  
Fig. 3. few shot learning

### 3.3 실험 결과

본 연구는 다양한 기사 형식과 문체에서 zero-shot과 few-shot 학습 기법의 성능을 BERTScore로 비교하였다.

딱딱한 문체에서는 대조군이 0.9082, zero-shot이 0.9049, few-shot이 0.9051로 유사한 성능을 보였으나, 부드러운 문체에서는 대조군이 0.8959, zero-shot이 0.8947, few-shot이 0.8922로 few-shot의 효과가 다소 낮았다.

구성 측면에서 역피라미드형은 zero-shot이 0.9030으로 가장 높았고, 피라미드형과 다이아몬드형 구성에서는 few-shot이 각각 0.9070과 0.9230으로 최고 유사도를 기록했다.

스타일에서는 Q&A형에서 zero-shot이 0.9068, 핵심 포인트형에서 few-shot이 0.9116으로 가장 높은 점수를 보였다.

전체 평균 BERTScore는 대조군이 0.9049, zero-shot이 0.9058, few-shot이 0.9052로, zero-shot이 대조군보다 약간 우수한 성능을 보였으며, few-shot도 대조군과 유사한 수준을 유지하였다. 이는 표 3에서 상세히 나타난다. 이러한 결과는 Few-shot 기법이 전반적으로 가장 높은 성능을 보였음을 보여주며 제안한 방법의 우수성을 입증한다.

표 3. 결과 분석  
Table 3. Results Analysis

대분류	소분류	대조군	zero shot	few shot
문체	딱딱한 문체의 기사	0.9082	0.9049	0.9051
	부드러운 문체의 기사	0.8959	0.8947	0.8922
구성	역피라미드형 구성	0.8999	0.9030	0.9011
	피라미드형 구성	0.9001	0.9056	0.9070
	다이아몬드형 구성	0.9087	0.9126	0.9230
	혼합형 구성	0.8999	0.8981	0.9009
스타일	서술형	0.9205	0.9146	0.9105
	Q&A형	0.9017	0.9058	0.8955
	핵심 포인트형	0.9087	0.9127	0.9116
average BERTScore		0.9049	0.9058	0.9052

#### IV. 결론

본 연구에서는 BERTScore를 활용하여 AI가 생성한 기사와 원본 기사 간의 의미적 유사성을 평가하였다. 평균 F1 점수가 0.9 이상으로 나타났으며, 이는 AI 생성 콘텐츠가 원본의 의미적 내용을 충실히 반영하고 있음을 확인시켜 준다. Few-shot 학습 방식과 Zero-shot 학습 방식을 비교한 결과, Few-shot 방식이 평균 F1 점수에서 0.9058로 Zero-shot의 0.9052에 비해 소폭 높은 결과를 보였다. 이는 Few-shot 학습이 제공하는 추가적인 컨텍스트가 AI에게 더 정확한 문체, 구성, 스타일을 반영하도록 도와주어 의미적 유사성을 높이는 데 기여한 것으로 해석된다[5]. 이러한 결과는 Few-shot 방식이 AI 기반 기사 생성에 있어 더욱 유의미한 성능 향상을 제공함을 시사한다.

대조군 포함 결과를 통해 볼 때, Few-shot 방식이 일관된 의미적 유사성을 유지하면서도 미세하게 높은 성능을 제공하는 것이 확인되었다. 그러나 BERTScore는 문체나 구조적 일관성을 평가하는 데 한계가 있음을 발견하였으며, 이는 텍스트 임베딩을

바탕으로 한 평가가 형식적 요소를 완전히 반영하지 못하기 때문이다. 이에 형식 평가를 보완할 수 있는 추가적인 평가 지표의 도입이 필요하다. 이러한 추가적인 평가는 단순한 내용 유사성을 넘어 다양한 텍스트 특성을 포괄적으로 분석하는 데 기여할 것으로 기대된다. 향후 연구에서는 수동 검토 및 형식을 학습한 분류 모델을 포함시켜 평가의 정확성과 종합성을 더욱 강화할 계획이다.

본 연구는 AI 기반 기사 생성 시스템의 효과를 입증하며, 특히 Few-shot 학습 방식이 제공하는 성능의 유의미한 향상을 강조한다[6]. 이는 실제 언론 환경에서의 적용 가능성을 높이며, AI 생성 텍스트의 의미적 충실성 및 문체와 형식적 요구사항을 평가하기 위한 정교한 방법론의 필요성을 제시한다.

#### 참고 문헌

- [1] 박상언, 강주영, "ChatGPT 및 거대언어모델의 추론 능력 향상을 위한 프롬프트 엔지니어링 방법론 및 연구 현황 분석," 지능정보연구, vol. 29, no. 4, 12월 2023.
- [2] A. Webson and E. Pavlick, "Do prompt-based models really understand the meaning of their prompts?," NAACL, June 2022.
- [3] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large language models are zero-shot reasoners," arXiv:2205.11916, May 2022.
- [4] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "What makes good in-context examples for GPT-3?," arXiv:2101.06804, Jan. 2021.
- [5] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, et al., "Language models are few-shot learners," NeurIPS, vol. 33, Dec. 2020.
- [6] T. Gao, A. Fisch, and D. Chen, "Making pre-trained language models better few-shot learners," ACL, Aug. 2021.



# 스태이블 디퓨전과 로라를 사용한 유니티 에셋 생성 도구

최성우\*, 흥범수\*, 윤준식\*, 김차종\*\*

## Unity Asset Generation tool with Stable Diffusion and LoRA

Seongwoo Choi\*, Beomsu Hong\*, Junsik Yoon\*, and Chajong Kim\*\*

### 요약

인디 게임 개발자가 늘어나는 가운데 유니티 프로그램을 사용하는 사람들이 많아지고 있다. 유니티를 이용한 게임 개발에 난관을 겪는 이유 중 하나가 자신의 게임 테마와 맞는 에셋을 구해야 하는 것과 텍스처 또는 배경을 직접 만들어야 한다는 어려움이 있다. 이 논문에서는 그러한 문제점을 해결하기 위해 LoRA를 학습해 스태이블 디퓨전 모델에 적용하여 더욱 정교한 이미지를 생성한다. 생성한 이미지에서 Florence-2를 이용하여 객체를 인식하고 이미지를 잘라와 배경을 제거한 후 바로 유니티에서 사용할 수 있도록 하는 도구를 개발하였다. 또한, 생성한 이미지를 직접 유니티와 게임에 적용함으로써 그 실효성을 보였다.

### Abstract

With the increasing number of indie game developers, more people are using the Unity program. However, one of the challenges in game development using Unity is finding assets that match the game's theme and the difficulty of creating textures or backgrounds manually. This paper addresses these issues by training LoRA and applying it to the Stable Diffusion model to generate more refined images. The generated images are processed using Florence-2 to recognize objects, extract them, and remove the backgrounds, allowing for immediate use in Unity. Furthermore, the effectiveness of the tool was demonstrated by directly applying the generated images in Unity and games.

### Key words

Stable Diffusion, LoRA, Florence-2, Unity

## 1. 서론

인디게임 산업은 지난 몇 년 동안 놀랄 만한 성장을 이루어 내며 전 세계적으로 주목을 받고 있다. 인디 개발자들이 창의적이고 독특한 아이디어로 다양한 게임을 창조하면서 시장은 다양성과 혁신의 새로운 파도에 휩싸이고 있다. 이러한 급격한 성장

은 인디게임 개발자에게 더 높은 수준의 경쟁과 함께 작업 속도의 중요성을 강조하고 있다[1].

거기에 인디게임 스튜디오들은 더 빠른 게임 출시를 통해 성공을 거두고 있다. 개발자 작업 시간의 감소는 이러한 트렌드의 핵심이 되며, 신속한 게임 개발은 경쟁에서 앞섰을 때 더 큰 경제적 이익을 가져온다. 최근에는 작업 속도와 생산성 향상을 위

\* 국립한밭대학교 컴퓨터공학과, wsdewsde27@gmail.com, qjatn3801@naver.com, frog0328@naver.com

\*\* 국립한밭대학교, 컴퓨터공학과 교수(교신저자), cjkim@hanbat.ac.kr

해 이미 만들어진 예셋의 적극적인 활용이 주목받고 있다[2].

이러한 이유로 게임 개발 과정에서 예셋은 중요한 역할을 하며 캐릭터, 배경, 사운드 효과 등 다양한 요소를 구성한다. 하지만 자신이 찾는 예셋이 존재하지 않을 수 있으며, 직접 예셋을 제작하기엔 시간과 노력이 많이 필요하다.

하지만 생성형 AI 기술의 발전으로 DALL-E, Midjourney등을 이용하여 예셋을 제작하는 시도를 해볼 수 있었다. 그런데 기존 생성형 AI 도구들은 다음과 같은 문제점을 가지고 있다.

낮은 제어력: 사용자가 원하는 예셋을 정확하게 제작하기 어렵다.

낮은 호환성: 다른 AI 도구에서 생성된 이미지를 바로 유니티에 사용할 수 없다.

번거로운 작업 과정: 여러 단계를 거쳐야 원하는 예셋을 제작할 수 있다.

유료서비스: 몇몇 생성형 AI 도구들(DALL-E, NovelAI, Midjourney)은 요금 청구 모델을 도입하고 있어, 많은 양의 예셋을 필요로 하는 대규모 프로젝트에서는 부담이 될 수 있다.

이에 위 문제점들을 해결하기 위해 사용자가 텍스트 입력과 한 번의 클릭과 같은 간단한 입력만으로 수월하게 원하는 예셋을 제작할 수 있도록 하는 Stable Diffusion 1.5를 기반으로 한 유니티 예셋 생성 도구를 개발하였다.

## II. 본 론

### 2.1 학습 데이터

예셋 중 텍스처 이미지를 생성하기 위해 필요한 LoRA를 학습하기 위해 저작권이 없는 학습 데이터를 준비하였다.

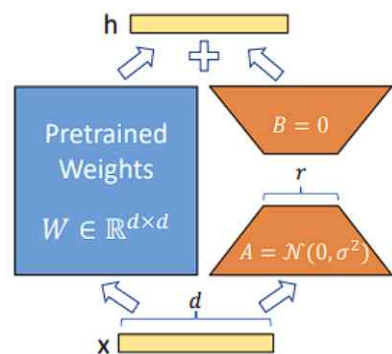
표 1에 나오는 것처럼 실제 게임을 만들 때 필요로 할 수 있는 텍스처들의 이미지를 512x512 크기로 준비하여 각 이미지에 txt파일로 캡션을 진행했다. 데이터개수가 모자라는 이미지들은 반복수를 더 높게 설정하여 학습이 고루 진행되도록 하였다.

표 1. 학습 데이터  
Table 1. Training Data

이미지 종류	데이터개수	반복수
asphalt	50	10
grass	57	10
gravel	58	10
leather	50	10
marble	48	10
metal	60	10
paper	68	10
sand	50	10
snow	50	10
soil	50	10
tile	50	10
wall_brick	50	10
wall_concrete	50	10
wall_stone	50	10
wood_texture	50	10
real_rock_texture	25	15
real_wood_texture	25	15

### 2.2 LoRA모델 학습

모델이 클수록 파인튜닝을 할 때 많은 파라미터를 학습시켜야하기 때문에 시간이 오래 걸리고 비효율적이다. 그리하여 저차원의 intrinsic rank를 이용해 파인튜닝하는 방법론인 LoRA를 사용하였다. (그림1)



LoRA only train A and B (reparametrization)

그림 1. LoRA의 작동방식[3]  
Fig. 1. How LoRA works[3]

LoRA를 사용하게 되면 전체 파라미터가 아닌 일부분의 파라미터만 업데이트 하는 것이기 때문에 시간과 비용 면에서 훨씬 효율적이다.(그림2)

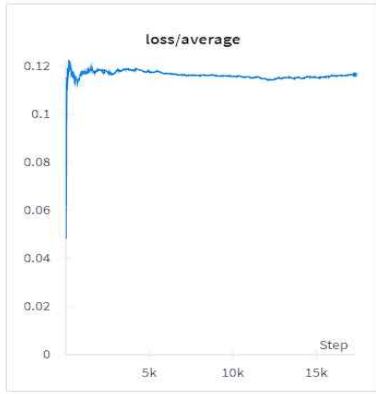


그림 2. LoRA학습 및 이미지 생성 결과  
Fig. 2. Training LoRA and Image Generation Results

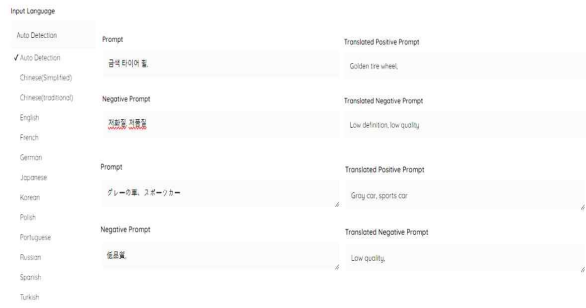


그림 3. 자동 번역 예시  
Fig. 3. Auto translate sample

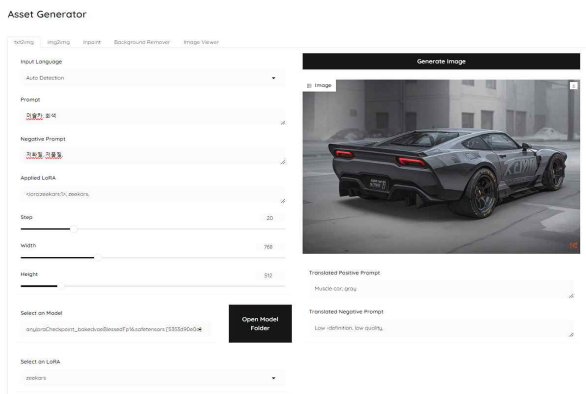


그림 4. txt2img 예시  
Fig. 4. txt2img sample

### 2.3 이미지 생성 인터페이스 및 개발내용

인터페이스는 Gradio를 이용하여 개발하였고 txt2img, img2img, inpaint, Background Remover, Image Viewer 탭을 가지고 있다.

모든 프롬프트는 googletrans를 이용하여 자동 번역을 구현하였기 때문에 사용자가 입력한 언어를 자동으로 감지하고 영어로 번역하여 이미지 생성에 사용하게 된다.(그림3)

txt2img에서 프롬프트 입력, 로라, 모델 선택 및 파라미터 조절을 하여 Stable Diffusion을 이용하여 이미지를 생성하고(그림4) img2img에서는 txt2img에서 이미지가 입력값으로 추가되어 기존 이미지를 기반으로 이미지를 생성할 수 있다.(그림5) inpaint에서는 img2img에서 추가적으로 마스킹된 이미지가 입력값으로 추가되며 직접 마우스 포인터를 통해 마스킹이 가능하고, 마스킹된 부분만 img2img를 진행하여 이미지를 수정하는 것이 가능하다.(그림6)

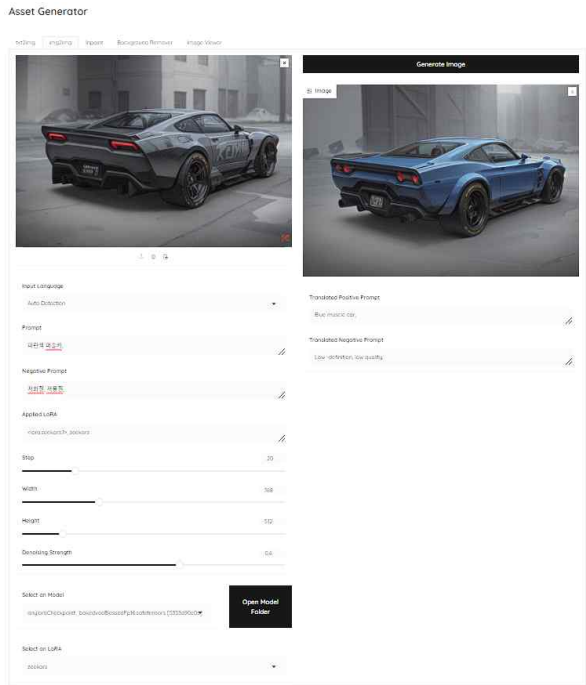


그림 5. img2img 예시  
Fig. 5. img2img sample

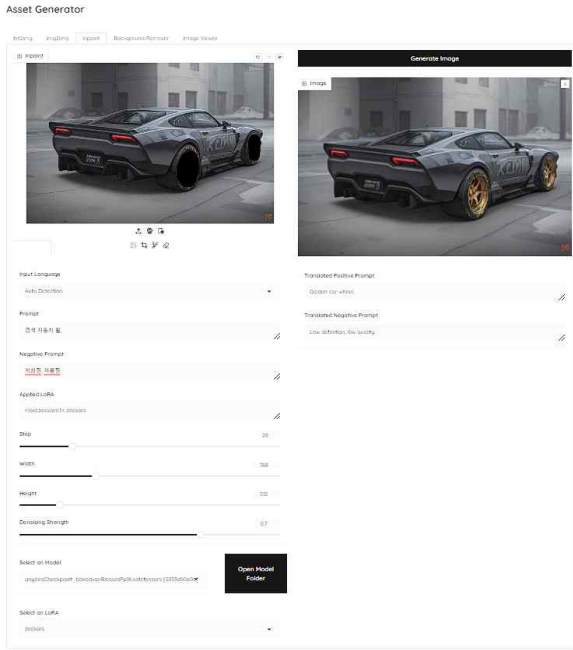


그림 6. inpaint 예시  
Fig. 6. inpaint sample

Background Remover 탭에서는 이미지를 삽입하고 실행을 누르게 되면 Florence-2를 이용하여 객체를 인식을 진행하고[4], 생성된 bounding box에 맞춰 이미지를 잘라와 배경을 제거하는 과정을 한 번에 거치게 된다. 즉 바로 에셋으로 사용이 가능하도록 가공해주는 것이다.(그림7)

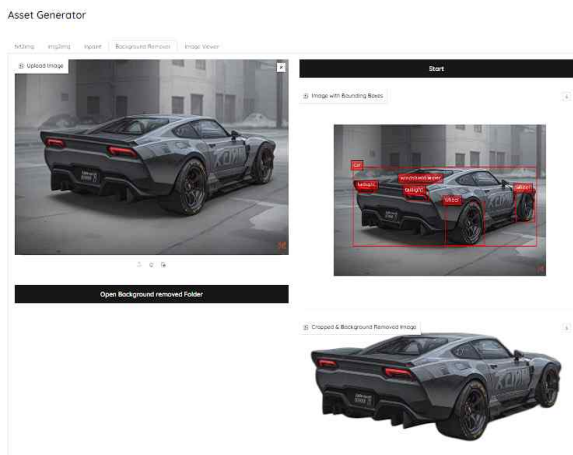


그림 7. 객체 인식 및 배경 제거  
Fig. 7. Object Detection and Background Removal

Image Viewer 탭에서는 지금까지 생성한 이미지들을 생성 날짜별로 선택하여 확인 가능하고, 파일 관리자에서 이미지 폴더를 열 수도 있다.(그림8)

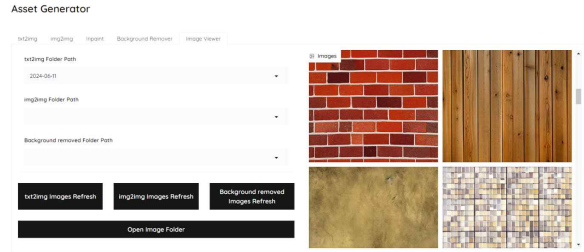


그림 8. 이미지 뷰어 예시  
Fig. 8. Image Viewer sample

## 2.4 이미지 생성 예시

이미지 생성에는 Stable Diffusion 1.5 모델을 사용하게 된다.(그림9) 현재 SDXL 등 더 성능이 높은 모델도 존재하지만, 이 모델은 더 높은 컴퓨팅 성능을 요구한다. 따라서 더 많은 사용자가 접근할 수 있도록 Stable Diffusion 1.5에 LoRA를 적용하는 방법을 선택하였다. Stable Diffusion과 LoRA를 활용한 이미지 생성 예시로는 배경, 텍스처, 스프라이트, 게임 내 아이콘, 효과, 프로필 이미지 등이 있다.(그림 10)

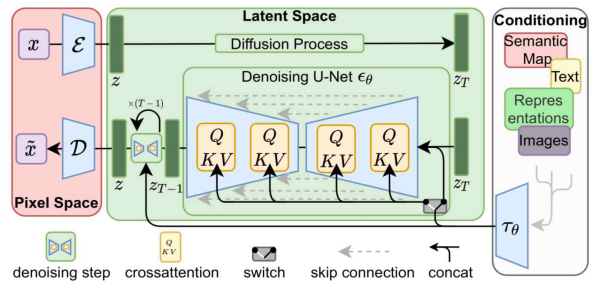


그림 9. 스테이블 디퓨전 구조[5]  
Fig. 9. Stable Diffusion Architecture[5]



그림 10. 제작한 도구로 생성한 이미지들  
Fig. 10. Images generated with the tools



## 2.5 적용 예시

생성한 이미지들을 직접 Unity와 게임에 적용시킨 예시이다. Unity에서도 직접 서비스 중인 Unity Muse와 비교해보았을 때에도 뒤쳐지지 않는 모습을 보여준다.(그림11)

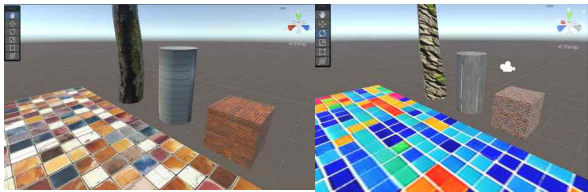


그림 11. 제작한 도구와 Unity Muse로 생성한 이미지 비교

Fig. 11. Compare images generated by the tool and Unity Muse

플레이어와 적기의 이미지를 직접 생성하여 실제 제작한 게임에 적용시켜 보았을 때에도 이질감 없이 잘 적용되는 모습을 확인할 수 있었다.(그림12)

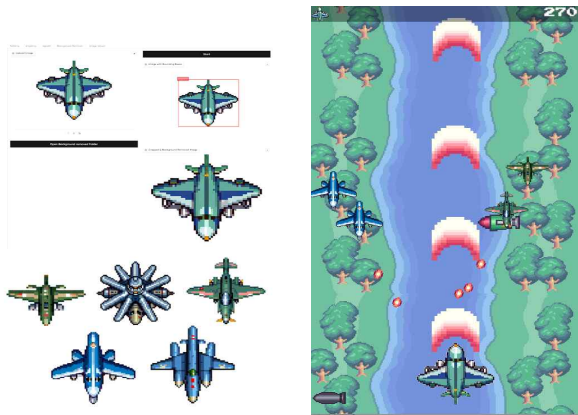


그림 12. 게임에 생성한 에셋을 적용한 모습

Fig. 12. Application of the created asset in the game

## III. 결 론

본 논문에서는 Stable Diffusion과 LoRA를 활용하여 유니티 에셋 생성 도구를 제안하였다. 제안된 도구는 인디 게임 개발자와 초보 개발자들이 손쉽게 자신만의 게임에 맞는 에셋을 생성할 수 있도록 돕는 시스템이다. 이를 통해 게임 개발에 소요되는 시간과 비용을 절감할 수 있으며, Florence-2를 이용한

객체 인식 기능을 추가하여 생성된 이미지의 배경을 자동으로 제거함으로써 유니티에서 바로 사용할 수 있도록 개발하였다. 본 도구는 Unity Muse와 비교했을 때에도 품질 면에서 뒤떨어지지 않는 성능을 보였으며, 직접 개발한 게임에서 생성된 에셋을 적용했을 때에도 자연스럽게 통합되는 결과를 확인하였다. 향후 연구에서는 더 다양한 LoRA를 학습하여 보다 많은 사용자들에게 도움이 될 수 있는 도구로 발전시켜 나갈 계획이다.

## 참 고 문 헌

- [1] 비즈한국, “대기업도 눈여겨보는 시장 '인디게임'의 시대 올까”, <https://www.bizhankook.com/bk/article/26742>
- [2] "유니티 2023 게임 업계 보고서", <https://unitysquare.co.kr/growwith/resource/form?id=335>
- [3] HU, Edward J., et al. Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685, 2021.
- [4] XIAO, Bin, et al. Florence-2: Advancing a unified representation for a variety of vision tasks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. p. 4818-4829.
- [5] CAMPOS, Ronald, et al. High-Resolution Image Synthesis with Latent Diffusion Models (Stable Diffusion). 2022.

# 8bit ADPCM : 잔차 양자화를 통한 RAW PCM 데이터의 새로운 압축 방식

정성원\*, 김민영\*

## 8bit ADPCM : A Novel Compression Method for Raw PCM data through Residual Quantization

Seongwon Jeong\* and Minyoung Kim\*

### 요 약

본 논문은 기존의 ADPCM 압축 방식을 개선해 새로운 8bit ADPCM 오디오 압축 방식을 제안한다. 다양한 오디오 관련 압축 방식이 존재하지만, 기존의 여러 압축 방식들은 헤더 데이터가 필수적이고, 사람의 가청 주파수를 벗어나는 범위에서 많은 손실을 가진다는 한계가 있었다. 반면, 본 논문에서 제안하는 8bit ADPCM은 RAW 데이터만을 가지고 압축 및 해제가 가능하며, 적은 Loss와 짧은 실행 시간이라는 장점을 가진다. 또한 해당 방식은 헤더 데이터를 포함하지 않기 때문에 다양한 AI 부문에서도 크게 활용될 수 있을 것으로 예측한다.

### Abstract

This paper proposes an improved 8-bit ADPCM audio compression method that builds upon traditional ADPCM compression techniques. While various audio compression methods exist, many conventional approaches require header data and suffer from significant loss outside the range of human-audible frequencies. In contrast, the proposed 8-bit ADPCM method enables compression and decompression using only raw data, offering the benefits of reduced loss and shorter execution time. Furthermore, as it does not include header data, this approach is anticipated to be highly applicable across a variety of AI domains.

### Key words

audio, quantization, music, compression, machine-learning

## I. 서 론

음악 데이터는 매우 큰 크기를 가지고 있기 때문에, 이를 효율적으로 압축시키는 여러 방법이 연구되어 왔다. 기존의 여러 압축 방식은 사람의 가청 주파수를 벗어난 범위의 주파수 대역에서 큰 손실

을 가져가는 방식으로 효율적인 오디오 압축을 진행한다. 하지만, 다양한 변환에 관련된 정보를 담기 위한 헤더 데이터가 필수적으로 필요하다는 한계를 가지고 있다. 따라서, 대부분의 오디오 관련 연구는 압축되지 않은 raw PCM 데이터를 이용해 수행하는데, WAV 파일의 큰 크기는 연구에 큰 장벽이 된다.

\* These authors contributed equally to this work.

\* 성균관대학교, sw5717312@naver.com, \*성균관대학교, kmny0128@naver.com

이에 본 저자들은 새로운 오디오 압축 방식, 8bit ADPCM을 제안한다. 기존의 4bit ADPCM은 굉장히 작은 용량으로 오디오를 압축할 수 있지만, 그만큼 큰 손실을 가진다는 단점이 있었다. 따라서, 이를 8bit로 확장해 손실을 최소화하도록 하는 방법을 연구하였다. 뿐만 아니라, raw 데이터를 이용해 압축 및 해제를 진행하기에, 다양한 연구에서도 해당 압축 방식을 활용할 수 있을 것으로 예측한다.

본 논문의 흐름은 다음과 같다.

II. 관련 연구 : 기존의 4bit ADPCM 방법을 알아본다. 또한, 일반적인 WAV 파일의 구조와 기존의 오디오 압축 방식인 MP3, Vorbis 등의 구조에 대해 탐구한다. 기존의 압축 방식이 어떠한 한계를 가지고 있는지를 살펴봄으로써 8bit ADPCM이 어떤 차별점을 가지는지를 더 잘 이해할 수 있도록 한다.

III. 8bit ADPCM 구조 : ADPCM이 데이터를 양자화하는 과정을 상세히 설명한다.

IV. 실험 및 결과 : 8bit ADPCM과 다른 압축 방식들의 손실률, 압축 및 해제에 걸리는 시간, Mel-spectrogram Loss를 비교, 분석한다.

V. 결론: 이 연구의 성과와 앞으로의 발전 방향에 대해 서술한다.

## II. 관련 연구

### 1. 4bit ADPCM

4 bit ADPCM 방식은 ~32K ~ 32K 범위의 PCM 데이터를 -7 ~ 7 범위의 데이터로 압축해 표현한다. 데이터 압축은 다음의 데이터를 예측하고 그 예측값과 실제 값의 차이를 현재 스텝 크기에 맞춰 양자화하는 방식으로 진행된다. 위의 그림 1은 실제

ADPCM 변환 과정의 한 예시를 보여준다.

구체적인 양자화 과정은 아래와 같으며, 1 ~ 4 과정을 반복해 진행한다.

1. 예측한 샘플값과 실제 PCM 데이터 사이의 차이를 계산 (이하 차이값)
2. 현재의 스텝 크기를 이용해 차이값을 양자화해 ADPCM 배열에 저장
3. 양자화된 차이값을 이용해 예측 샘플 업데이트
4. 양자화된 차이값을 이용해 인덱스 테이블을 참조하여 스텝 크기 값 업데이트를 진행

### 2. WAV 파일의 구조와 PCM 데이터

WAV 파일은 마이크로소프트와 IBM이 개발한 오디오 파일 포맷으로, 주로 무압축 상태의 PCM 데이터를 저장하는 데 사용된다. WAV 파일은 크게 파일 헤더와 실제 오디오 데이터로 구성된다. 파일 헤더는 파일의 기본 정보를 담고 있으며, 주파수, 채널 수 등 오디오 데이터의 특성을 정의한다. 실제 오디오 데이터는 PCM 데이터 형식으로 저장된다.

PCM 데이터는 아날로그 오디오 신호를 디지털화한 데이터로, 연속적인 아날로그 신호를 일정한 시간 간격으로 샘플링하여 기록한다. PCM 방식은 주파수와 진폭의 정보를 유지하는 비압축 방식으로, 높은 음질을 제공하지만, 큰 파일 크기를 요구한다.

### 3. MP3

MP3 코딩 방식은 변형 이산 코사인 변환(MDCT)을 사용하여 오디오 신호를 주파수 도메인으로 변환한다. MDCT를 적용하기 이전, 신호를 작은 블록으로 나누고 이를 이용해 주파수 정보를 분석한다.

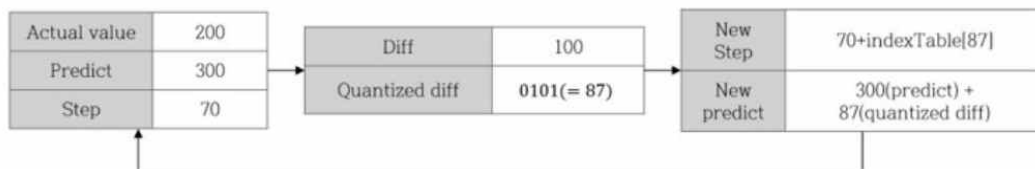


그림 1. ADPCM convert 과정의 한 예시

Fig. 1. An Example of the ADPCM Conversion Process

그리고 청취 심리학적 모델을 적용해 인간의 귀로 덜 인지되는 소리를 제거, 축소하여 데이터 크기를 줄인다. 주파수 대역은 중요도에 따라 서브밴드 필터링으로 나뉘어 처리되며, 낮은 대역은 더 정밀하게, 높은 대역은 덜 정밀하게 압축시킨다. 양자화를 통해 각 주파수 성분을 비트로 표현하고, 또한, 자주 나타나는 값을 더 적은 비트로 표현하는 허프만 코딩을 사용한다. MP3는 비트레이트 제어기능을 통해 고정 또는 가변 비트레이트 방식으로 데이터를 효율적으로 압축하며, 이를 통해 비교적 작은 파일 크기에서도 좋은 음질을 유지한다.

#### 4. Vorbis

Vorbis 코딩 방식은 여러 압축 방식을 조합하여 고효율의 손실 압축을 구현한다. MP3와 비슷하게 수정 이산 코사인 변환(MDCT)을 사용하여 오디오 신호를 시간 도메인에서 주파수 도메인으로 변환하며 인간이 인지할 수 없는 영역의 소리를 제거하거나 축소시킨다. 하지만, Vorbis는 서브밴드 필터링을 진행하지 않고 전체 대역폭에 MDCT를 적용한다. 이후 벡터 양자화를 통해 주파수 데이터를 코드북 엔트리로 인코딩하고, 중요한 주파수 성분만 남겨 데이터를 최소화시킨다. 그리고 허프만 코딩을 적용시킨다. 마지막으로, 잔여 코딩(Residue Coding)을 통해 예측이 어려운 신호를 별도로 인코딩한다. 이러한 과정을 통해 Vorbis는 고품질의 오디오를 상대적으로 작은 파일 크기로 압축한다.

#### 5. AAC

AAC는 MP3 압축 방식을 개선시킨 압축 방식이다. AAC는 더 유연한 블록 크기를 사용하여, 신호에 따라 작은 블록(120개의 샘플) 또는 큰 블록(1024개의 샘플)을 선택적으로 적용하여 복잡한 신호는 더 정밀하게, 단순한 신호는 더 효율적으로 압축한다. 또한, AAC는 서브밴드 필터링 없이 바로 MDCT를 적용해 전 대역을 처리하며, PNS(Perceptual Noise Substitution) 기법을 도입하여 특정 잡음을 데이터로 저장하지 않고 복원할 때 대체하는 방식을 사용한다. 또한 TNS(Temporal Noise Shaping)를 통해 시간적 변동이 큰 신호에서 잡음을 더욱 효과적으로 제어할 수 있으며, 가변 비트레이

트(VBR)를 기본으로 제공하여 MP3보다 더 나은 음질-파일 크기 비율을 제공한다.

### III. 8bit ADPCM 구조

본 저자들은 4bit ADPCM에서 활용한 인덱스 테이블과 스텝 크기 테이블의 경향성을 그대로 반영하되, 4bit에서 8bit로의 비트 수 확장으로 인한 변화만을 반영하여 두 테이블을 구성하였고, 테스트를 진행해 보았다. 하지만 테스트 결과가 기대에 미치지 못하여 보다 좋은 성능을 낼 수 있는 방법을 고민하던 중, 스텝 크기를 고정하고, 두 테이블은 사용하지 않는 방법을 고안하게 되었다.

본 저자들은 여러 스텝 크기를 실험해보며 최적의 스텝 크기를 찾았고, 4K의 스텝 크기를 가진 8bit ADPCM 방법을 제안한다.

이렇게 스텝 크기를 고정하면서, 인덱스와 스텝 크기를 업데이트 하는 과정이 생략되고, 차이값을 계산하고 양자화하여 저장하는 과정만으로 압축 과정이 단순해진다.

이 방식으로 int16 타입의 데이터인 PCM 데이터를 int8 형태로 표현할 수 있다. 이는 부동 소수점 16bit 영역에서 정확히 표현될 수 있는 장점을 가지고 있다.

### IV. 실험 및 결과

본 저자들은 다양한 압축 방식과 8bit ADPCM 압축 방식의 L1 Loss 차이와 압축 및 해제에 걸리는 시간, 그리고 Mel-spectrogram Loss를 계산해 비교하였다.

8bit ADPCM 방법 내에서도 다양한 값으로 설정한 고정 스텝 크기 버전과 가변 스텝 크기 버전의 성능을 비교하였다.

DEAP experiment에 사용된 40개의 음악 데이터를 대상으로 실험을 진행한다. 아래는 실험 결과를 나타낸 표이다. 값들을 보면, MP3, AAC, OGG 등 기존의 압축 방식들은 PCM 데이터를 압축 및 해제했을 때 데이터 하나당 약 4000이 넘는 로스가 발생하고 있으나, 본 논문에서 제시한 8bit ADPCM 방식은 이보다 약 100배 이상 감소한 34 정도의 값을

보이고 있다. 하지만 기존에 제시되었던 4bit ADPCM은 훨씬 큰 손실률을 가지는 모습을 보여주었다. 즉, 단순히 ADPCM을 활용하는 것이상으로 8bit로 범위를 확장한 것이 유의미하게 작동함을 확인할 수 있다.

표 1. 압축 방식 별 압축 및 해제 시 L1 loss, 압축 해제 시간

Table 1. L1 Loss and Compression-Decompression Time for Each Compression Method

Codec		L1 Loss	duration	Mel Loss
ADPCM With Fixed stepsize	1k	581.0762	1.8303	1.5417
	2k	176.4196	1.8679	0.8068
	<b>4k</b>	<b>34.3018</b>	<b>1.5788</b>	<b>0.6187</b>
	8k	32.5912	1.3134	1.0205
	16k	63.9604	1.1151	2.0145
ADPCM_8bit_table		124.7908	3.2071	1.0325
ADPCM_4bit		934.3615	0.7791	1.3635
mp3		4295.3301	1.0562	5.4391
aac		5873.7469	3.9146	3.7397
libvorbis		4354.7215	0.9603	5.3234

또한, 가변 스텝 크기에서보다 고정 스텝 크기에서 더 Loss가 적은 모습을 확인할 수 있다. 특히, 4K의 스텝 크기에서는 L1 Loss뿐만 아니라 Mel-spectrogram Loss에서도 타 압축 방식보다 월등히 좋은 값을 보이는 것을 확인할 수 있다.

본 저자들은 가변 스텝 크기에서 더 결과가 좋지 않은 이유를 기존의 4bit 방식을 8bit로 그대로 확장했을 때, 오디오의 변화 흐름을 step size 표와 index size 표가 제대로 표현하지 못 하고 있을 것으로 예측한다. 또한, 사용할 수 있는 수가 2배 늘어났기에, 고정 스텝 크기에서도 변화값을 다소 정확하게 양자화하는 것으로 보여진다.

다음으로 Duration 면에서 비교했을 때, 기존 압축 방식들은 압축 및 해체에 훨씬 더 많은 시간을 소요하여 약 4~5초 정도의 시간이 필요했던 반면에 8bit ADPCM은 그보다 약 3배 이상 빨라진 1.3초 정도의 시간 만에 압축 및 해제를 완료하였다. 4bit ADPCM의 경우는 loss는 좋지 않았지만, 압축 방식이 8bit 버전보다 단순하기에 압축 및 해제의 속도만큼은 더 빠른 모습을 보여주고 있다.

하지만, 8bit ADPCM 방식은 기존의 데이터를 2배 압축하는 반면, 기존의 압축 방식은 그보다 더 좋은 압축률을 제공하고 있다. 따라서 해당 방식은 좋은 압축률보다는 데이터의 정확성과 빠른 응답 시간을 제공하는 것으로 볼 수 있다. 그렇기에 다양한 데이터 가공 혹은 머신러닝 학습 분야에서 활용될 것을 기대한다.

## V. 결론

본 저자들은 과거 사용되었던 ADPCM 압축 방식을 8bit 크기로 확장한다면 현재 오디오 영역에서도 적용될 수 있음을 증명하였다. 이 방식은 압축 및 해제 시간을 상당히 감소시켰으며, 특히 손실 압축 방식에는 불가피한 loss를 매우 낮출 수 있었다.

추후 연구에서는 해당 방식을 통해 압축한 ADPCM 데이터를 활용하여 오디오 데이터를 활용하는 인공지능 모델의 학습을 부동 소수점 32bit 영역이 아닌 부동 소수점 16bit 영역에서 수행하는 것을 연구해 볼 수 있을 것으로 기대된다. 이를 통해 리소스 사용의 감소 및 학습 속도 증진 효과를 얻을 수 있을 것으로 예상된다.

또한 해당 방식은 기존 압축 방식과 달리 헤더를 필요로 하지 않기 때문에 음악 생성형 AI 모델에서 활발히 활용될 수 있을 것을 기대한다.

## 참고 문헌

- [1] Karlheinz Brandenburg, "MP3 and AAC explained", AES 17th International Conference on High Quality Audio Coding, 1999
- [2] Paulius Micikevicius et al, "Mixed Precision Training ", 2018
- [3] Xiph.Org Foundation, "Vorbis," GitHub repository, 1994-2020. <https://github.com/xiph/vorbis>
- [4] R. Richey, \*Adaptive Differential Pulse Code Modulation using PICmicro™ Microcontrollers\*, Microchip Technology Inc., 1997.

# 광대역 주파수 스펙트럼과 영상처리를 이용한 AI 기반 날씨 예측 시스템 설계

정보석\*, 이세현\*\*, 홍익표\*\*\*

## Designing an AI-powered weather prediction system using broadband frequency spectrum and image processing

Bo-seok Jeong\*, Se-Hyeon Lee\*\*, and Ic-Pyo Hong\*\*\*

### 요 약

본 연구에서는 1MHz~6GHz의 SDR 기기인 Hack RF와 Raspberry Pi를 활용하여 주파수 스펙트럼 및 영상 처리를 통한 인공지능 기반 정밀 날씨 예측 시스템을 구현하였다. 정밀한 날씨 예측을 위하여 온도, 습도, 기압 데이터는 Raspberry Pi의 BME 280 센서를 사용하여 수집하고, Hack RF를 통해 100MHz~300MHz, 600MHz~800MHz, 2.3GHz~2.5GHz 총 3개의 주파수 대역에서의 PSD(Power Spectrum Density) 변화와 날씨 영상의 RGB 비율 변화를 감지하였다. 수집된 데이터 텍스트 파일로 전처리 후 인공지능 모델에 입력되어 날씨 예측을 할 수 있으며, 이를 기존 기상 예보와 비교하여 본 연구에서 제안한 새로운 접근법의 정확성을 검증하였다.

### Abstract

In this paper, an artificial intelligence-based precise weather prediction system was implemented through frequency spectrum and image processing by using Hack RF and Raspberry Pi, which are SDR devices of 1 MHz to 6 GHz. For more precise weather prediction, temperature, humidity, and atmospheric pressure data are collected using Raspberry Pi BME 280 sensor, and through Hack RF. we detected PSD (Power Spectrum Density) changes and RGB ratio changes in weather images in a total of three frequency bands: 100 MHz tcco 300 MHz, 600 MHz to 800 MHz, and 2.3 GHz to 2.5 GHz. After preprocessing with the collected data text file, it can be entered into an artificial intelligence model to make weather prediction and compared with the existing weather forecast to verify the accuracy of the new approach proposed in this paper.

### Keywords

AI, Radio Frequency, Hack RF, Raspberry Pi, Image Processing

---

\* 국립공주대학교 스마트정보기술공학과, information\_stone@nate.com,

\*\* 국립공주대학교 스마트정보기술공학과, project259@naver.com,

\*\*\* 국립공주대학교 스마트정보기술공학과, iphong@kongju.ac.kr(교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음 (2024-0-00073)

## I. 서론

예상보다 더욱 빠르게 다가오는 지구 온난화로 인해 최근 기후변화가 매우 심각해지고 있다. 국립기상과학원의 기상기술 정책정보 동향 보고서에 따르면 The Guardian의 IPCC 전문가 백여 명을 대상으로 설문을 진행한 결과, 응답자의 3/4 이상이 2100년까지 지구 평균기온이 2.5°C 이상 상승할 것으로 예측하였으며[1], 실제로 폭염으로 인해 최근 온열 질환자의 수가 전년 대비 80.2% 증가하였다[2]. 또한, 2023년 여름 기록적인 장마로 전국적으로 인명 및 재산 피해가 발생했으며 도로, 철도 등 교통 인프라에도 많은 영향을 끼치는 등[3] 다양한 기상이변 현상이 계속 발생하고 있다. 따라서 급격한 기후변화에 대한 정확한 일기예보의 중요성이 높아지고 있으나, 현재 국내 기상청 예측 정확도는 최근 5년 사이 83%에서 69%로 하락하였다[2]. 이러한 문제를 해결하기 위해 야외의 날씨 영상의 채도와 명도 값을 이용하여 온도를 예측하고 분류할 수 있는 모델을 연구하거나[4], CCTV 영상을 이용해 명암도와 RGB 분석을 통해 맑음, 흐림, 비, 안개를 구분하는 알고리즘을 사용하였다[5]. LSTM 모델을 이용해 날씨를 예측하거나 [6~8]. 딥러닝 기술을 이용해 인공지능 기반의 다양한 연구가 진행되어왔다 [9~12].

본 논문에서는 정밀한 날씨 예측을 위하여 BME 280 센서의 온도, 습도, 기압 뿐만 아니라 영상 센서를 이용하여 실제 날씨 영상과 Hack RF 이용하여 100MHz~300MHz, 600MHz~800MHz, 2.3GHz~2.5GHz 주파수 대역에서의 스펙트럼 변화를 사용하여 초단기 예보 날씨 예측이 가능한 시스템을 설계하였다.

## II. 설계 및 시뮬레이션

### 2.1 시스템 구성도

본 논문에서는 그림 1에 나타난 바와 같이 Hack RF를 이용하여 100MHz~300MHz, 600MHz~800MHz, 2.3GHz~2.5GHz인 주파수 대역에서 측정된 PSD (Power Spectrum Density) 값과 실제 날씨 영상의

RGB 비율값, 그리고 센서를 통해 측정된 온도, 습도, 기압의 변화값 총 3가지 방식을 측정된 날씨 정보를 활용하여 AI 학습을 통해 초단기 예보 날씨 예측이 가능한 시스템을 설계하였다.

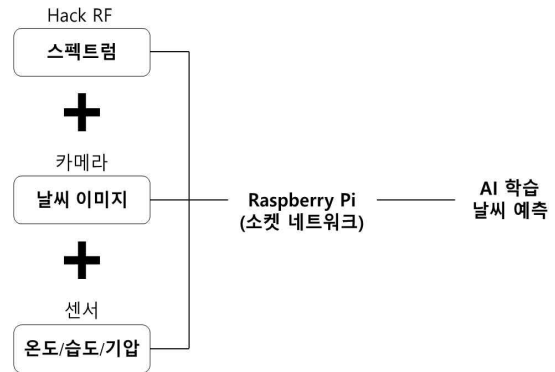


그림 1. 시스템 구성도

Fig. 1. System Configuration Chart

### 2.2 스펙트럼 변화

앞에서 설명한 바와 같이 제안한 시스템은 스펙트럼 변화와 날씨 영상 그리고 온도, 습도, 기압 변화 등 총 3가지의 변화값을 사용하여 날씨 예측이 가능하며, 2.1장은 스펙트럼 변화를 이용한 날씨 구분 방법을 나타내었다.

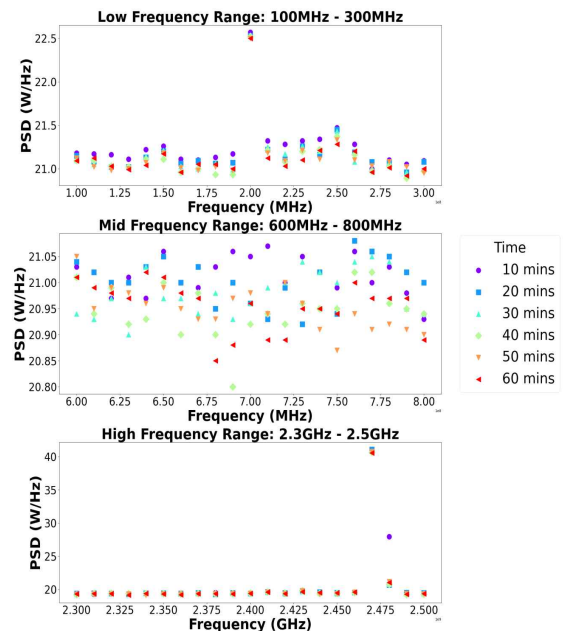


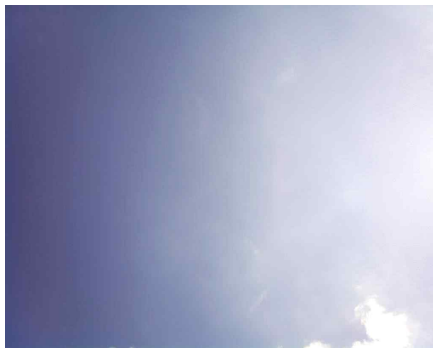
그림 2. 시간에 따른 스펙트럼 PSD 변화

Fig. 2. Spectral PSD changes over time

측정 주파수 대역은 Raspberry Pi와 Hack RF의 성능 및 야외 측정 환경을 고려하여 낮은 주파수 대역 (100MHz~300MHz), 중간 주파수 대역 (600MHz~800MHz), 높은 주파수 대역(2.3GHz~2.5GHz) 총 3가지 주파수 대역으로 나누어 설정하였다.

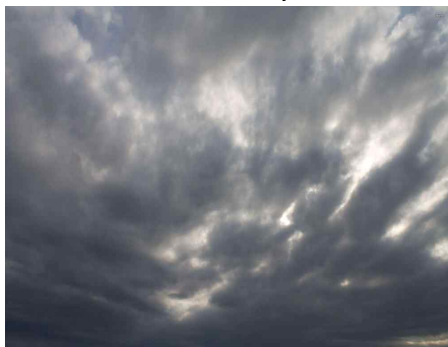
### 2.3 날씨 영상 변화

2.3장에서는 날씨 영상 변화값을 이용한 날씨 구분 방법에 대하여 설명한다. 그림 3은 맑은 날씨와 흐린 날씨 영상의 적색, 녹색, 청색의 분포도를 나타낸 것으로, 이때의 날씨 영상은 120°의 넓은 화각의 카메라인 OV5647 모듈을 이용하여 실제 2024년 8월의 하늘을 촬영한 것이다.



**Red Ratio: 26.10%**  
**Green Ratio: 32.55%**  
**Blue Ratio: 41.35%**

(a) 맑은 날  
 (a) clear day



**Red Ratio: 32.48%**  
**Green Ratio: 33.00%**  
**Blue Ratio: 34.52%**

(b) 흐린 날  
 (b) cloudy day  
 그림 3. 날씨 영상

Fig. 3. weather images

그림 3을 통해 알 수 있듯이, 맑은 날씨의 경우 영상의 RGB 분포가 적색 26.10%, 녹색 32.55%, 청색 41.35%로 청색의 분포가 높고 적색 분포가 낮았으며 흐린 날씨에는 적색 32.48%, 녹색 33.00%, 청색 34.52%로 RGB 분포가 전반적으로 비슷하게 분포되어 있다는 것을 알 수 있다. 8월 셋째 주부터 9월 첫째 주 동안 측정한 결과에 따라 RGB 분포 비율을 활용하여 화창한 날과 구름이 많은 날을 구분할 수 있으며, 이를 AI 학습에 적용하면 보다 정확한 날씨 예측이 가능하다.

### 2.4 인공지능 설계

초단기 날씨예측 인공지능 시스템을 설계하기 위해 가장 먼저 기상청의 종관기상관측 4개년의 일 자료 중에서 온도, 습도, 기압과 당일 날씨 아이콘을 토대로 날씨 정보를 표 1과 같이 0:맑음, 1:구름조금, 2:구름많음, 3:흐림, 4:비, 5:눈 순으로 숫자로 변환하여 학습 데이터로 활용하였다. 주파수 스펙트럼과 날씨 영상에 관한 데이터는 기상청에서 제공하지 않음으로 2.1장과 2.3장에 언급한 바와 같이 직접 맑은 날과 흐린 날에 여러번 측정 실험을 진행하였으며, 해당 데이터를 학습 데이터로 활용하였다. 날씨예측 인공지능은 파이썬의 RandomForest Classifier 기계학습 알고리즘을 사용하였다. 랜덤 포레스트로 여러 개의 결정트리 분류기가 배경을 기반으로 데이터를 샘플링하고 학습을 한 후 예측하여 나타낸다.

표 1. 기상청 데이터

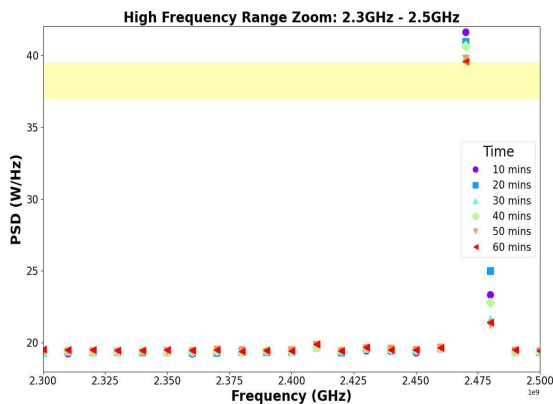
Table 1. Meteorological Administration data

일시	기온 (°C)	상대 습도 (%)	현지 기압 (hPa)	날씨	날씨 (숫자)
2023-01-01	-2.3	65	1020.6	구름조금	1
2023-01-02	-4.8	54.8	1021.8	맑음	0
2023-01-03	-6.3	66.5	1020.4	구름많음	2
2023-01-04	-3.9	69.1	1020.4	구름조금	1

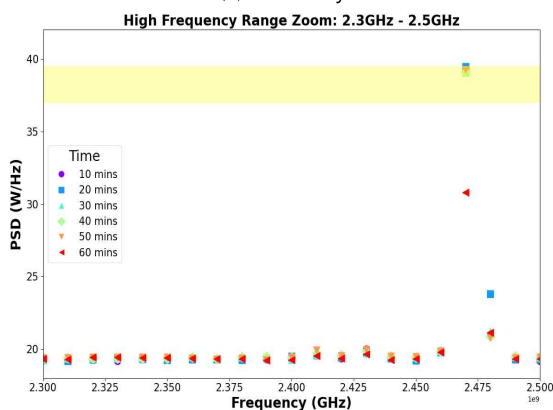


### III. 제안한 시스템 구성도 및 측정 결과

그림 4는 Hack RF를 활용하여 시간 변화에 따른 주파수 스펙트럼의 PSD 변화값을 그래프로 나타낸 것이다. 이때 측정 주파수 대역은 Raspberry Pi와 Hack RF의 성능 및 야외 측정 환경을 고려하여 낮은 주파수 대역 (100MHz~300MHz), 중간 주파수 대역 (600MHz~800MHz), 높은 주파수 대역(2.3GHz~2.5GHz) 총 3가지 주파수 대역으로 나누어 설정하였다. 측정 결과, 시간이 변화함에 따라 낮은 주파수 대역과 중간 주파수 대역에서는 PSD 변화값이 비교적 일정하며 큰 변화가 없지만 높은 주파수 대역에서는 맑은 날과 흐린 날에 따라 PSD 값이 달라지는 것을 확인할 수 있다.



(a) 맑은 날  
(a) clear day



(b) 흐린 날  
(a) clear day

그림 4. 14시 높은(2.3GHz~2.5GHz) 대역의 PSD 변화  
Fig. 4. PSD change in high (2.3GHz~2.5GHz) band at 14:00

특히 PSD 값이 37W/Hz~39W/Hz 사이에 나타날 경우에는 날씨가 흐리며, 소량의 비가 내리는 것을 확인할 수 있지만, PSD 값이 40W/Hz 이상일 경우에는 날씨가 맑은 경향이 나타났다. 실제 측정 시간 때의 기상청 예보에서는 비 예측은 없이 흐린 날씨만 예측하였지만, 실제 기상청 중관기상관측 결과에 따르면 흐린 날씨에 0.4mm 강수량의 비가 내렸다.

### IV. 결 론

본 연구에서 제안한 인공지능 기반 날씨예측 모델은 주파수 스펙트럼과 야외 영상, 그리고 기상 데이터(온도, 습도, 기압)를 활용하여 초단기 예보의 정확성을 보다 높일 수 있다. 특히 주파수 2.3GHz~2.5GHz 대역에서 PSD값의 변화를 통해 흐린 날씨에서의 소량의 비소식을 보다 정확하게 예측할 수 있음을 확인할 수 있다. 향후 연구에서는 소나기와 같은 특이기상과 같은 다양한 날씨에 대한 주파수 스펙트럼 데이터와 다른 학습 데이터를 수집하여 제작한 인공지능 모델의 성능을 향상시킬 예정이다.

### 참 고 문 헌

- [1] 2024-Meteorological Technology Policy Information Trends, NIMS, July 2024
- [2] Reporter Jeong Gu-hee, "It rained 'twice as much as forecast', Why the ups and downs of the 'rainy season'?", SBS News, July 2024
- [3] Korea Meteorological Administration and government jointly. "2023 Abnormal Climate Report" Korea Disease Control and Prevention Agency, April 2024
- [4] W. J. Choi and S. J. Shin, "A Study on Temperature Prediction Model Using Weather Image", Journal of Next-generation Convergence Technology Association, Vol. 4, No. 4, pp. 400-405, August 2020.
- [5] L. S. Han, J. S. Ha, G. Y. Jin, J. B. Seo, and S. C. Kim, "A Study on the weather prediction using LSTM systems", Proceedings of Symposium of the

- Korean Institute of communications and Information Sciences, Vol. 2020, No. 11, pp. 15-16, November 2020
- Transactions of the Royal Society A, vol. 379, no 2194, February 2021.
- [6] B. Y. Bak, S. Namkoong, and H. S. Lee, "Development of the Weather Information Detection Algorithm using CCTV Images", Proceedings of the KOR-KST Conference, Vol. 52, No 53, pp. 525-530, June 2006.
- [6] Baste, Pravin, and D. Dighe, "Low cost weather monitoring station using raspberry pi", International Research Journal of Engineering and Technology (IRJET) 4.5, vol. 4, no. 5, pp. 3184-3189, May 2017.
- [7] Charlton-Perez, Andrew J., et al. "Do ai models produce better weather forecasts than physics-based models? A quantitative evaluation case study of storm ciarán", npj Climate and Atmospheric Science, vol. 7, no. 93, April 2024.
- [8] B. Kosovic, S. E. Haupt, D. Adriaansen, S. Alessandrini, G. Wiener, L. D. Monache, Y. Liu, S. Linden, T. Jensen, W. Cheng, M. Politovich and P. Prestopnik, "A comprehensive wind power forecasting system integrating artificial intelligence and numerical weather prediction." *Energies*, vol. 13, no. 6, March 2020.
- [9] R. Xiaoli, L. Xiaoyong, R. Kaijun, S. Junqiang, X. Zichen, D. Kefeng and W. Xiang, "Deep learning-based weather prediction: a survey." *Big Data Research*, vol. 2, no. 3, 2021.
- [10] W. Fang, Q. Xue, L. shen and V. S. Sheng "Survey on the application of deep learning in extreme weather prediction." *Atmosphere*, vol. 12, no. 6, April 2021.
- [11] Hess, Philipp, and Niklas Boers. "Deep learning for improving numerical weather prediction of heavy rainfall." *Journal of Advances in Modeling Earth Systems*, vol. 14, no. 3, March 2022.
- [12] P. Grönquist, C. Yao, T. B. Nun, N. Dryden and P. Dueben, "Deep learning for post-processing ensemble weather forecasts." *Philosophical*

# 겨울철 결빙 구간 교통사고 방지를 위한 자동화 시스템 구현

김영재\*, 최영우\*, 서방원\*\*

## Implementation of Automated System to Prevent Traffic Accidents on Icy Road Sections in Winter

YoungJae Kim\*, Youngwoo Choi\*, and Bangwon Seo\*\*

### 요 약

겨울철 도로 결빙 구간에서 발생하는 교통사고는 매년 심각한 인명 및 재산 피해를 초래한다. 기존의 결빙 방지 장치인 염수 살포 장치와 열선 시스템은 노즐 동결, 차량 부식, 높은 설치 및 유지 비용 등의 문제를 지닌다. 본 연구는 ATmega328 기반 자작 프로세서를 사용하여 저비용 자동화 결빙 방지 시스템을 구현했으며, 실시간 데이터 분석과 제어를 통해 노즐 동결 방지, 차량 유무 확인, 결빙 상황에 맞춰 염수 살포를 자동화하여 도로 안정성을 향상시킨다.

### Abstract

Road accidents due to icy road sections in winter cause significant annual losses of life and property. Traditional frost prevention systems, such as brine spray and heating wires, face issues like nozzle freezing, vehicle corrosion, and high installation and maintenance costs. This study implements a cost-effective automated frost prevention system using an ATmega328-based custom processor, which prevents nozzle freezing, detects vehicle presence, and automates brine spraying according to icy conditions through real-time data analysis and control, ultimately enhancing road safety.

### Key words

Frost prevention, automated system, ATmega328, brine spray

## 1. 서 론

그림 1은 전체 교통사고 치사율과 결빙 교통사고 치사율을 보여준다. 겨울철 도로 결빙 구간에서 발생하는 교통사고는 매년 수 많은 인명 및 재산 피해를 초래한다 [1]. 특히 결빙이 자주 발생하는 터널 입구 및 출구, 교량, 급경사 구간에서는 사고 위험이 더욱 높다. 이를 예방하기 위해 널리 사용되는 시스템으로는 염수 살포 장치와 열선 장치가 있지

만, 이들 시스템은 염수 살포 노즐의 동결로 인한 오작동, 차량 부식, 높은 설치 및 유지 비용 등의 문제를 안고 있다. 본 연구에서는 이러한 문제를 해결하기 위해 ATmega328 기반의 저비용 자동화 결빙 방지 시스템을 설계하고 구현하였다. 제안하는 시스템은 실시간 환경 데이터를 기반으로 결빙 방지를 위한 여러 장치들을 자동으로 제어하며, 필터링 기법을 통해 데이터 안정성을 높였다.

\* 국립공주대학교, \*\* 국립공주대학교 (교신저자: seobw@kongju.ac.kr)

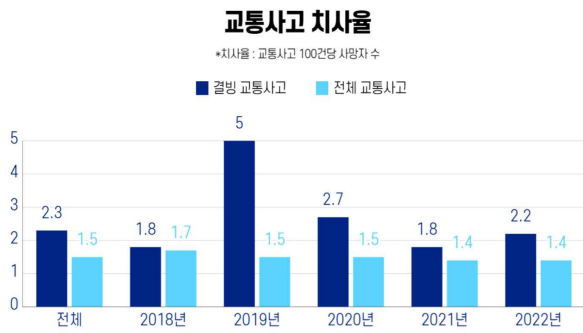


그림 1. 교통사고 치사율  
Fig 1. Fatality rate of traffic accidents

## II. 기존 결빙 방지 시스템의 한계

기존 결빙 방지 시스템은 주로 염수 살포 장치와 열선 장치로 구분된다. 염수 살포 장치는 차량의 존재 여부와 관계없이 결빙 조건이 형성되면 염수를 살포하는 방식으로 동작하는데, 이로 인해 차량이 부식 문제가 발생할 수 있으며, 겨울철 극한 상황에서 염수 살포 노즐이 동결될 경우 제 기능을 발휘하지 못할 수 있다. 또한 열선 시스템은 1m당 약 100만 원의 설치 비용과 매년 약 360만 원의 관리 비용이 소요되어 실용성이 제한적이다. 이러한 문제들을 해결하기 위해 새로운 저비용 자동화 결빙 방지 시스템이 필요하다.

## III. 제안하는 결빙 방지 자동화 시스템 설계

### 3.1 ATmega328 기반 자작 프로세서

ATmega328을 기반으로 한 자작 프로세서는 다양한 센서 및 장치들에서 수집된 데이터를 실시간으로 처리하고 자동으로 결빙 방지 시스템을 제어한다 [2]. 이를 통해 LED 경고, 염수 살포 및 발열 장치 등의 작동을 자동 제어하여 효율성을 극대화했고, 해당 프로세서는 비용 절감과 소형화를 위해 불필요한 부품을 최소화하였다.

### 3.2 HC-SR04 초음파 거리 센서와 EMA 필터

HC-SR04 초음파 센서는 차량의 존재 여부를 감지하여 결빙 구간 내에 차량이 없을 때만 염수 살

포가 이루어지도록 한다. 그러나 초음파 센서는 노이즈가 많아 데이터의 신뢰성이 떨어지기 때문에 지수 이동 평균(EMA) 필터를 적용하였다 [3], [4]. EMA 필터의 수식은 다음과 같다.

$$y[n] = \alpha \cdot x[n] + (1 - \alpha) \cdot y[n-1] \quad (1)$$

여기서  $\alpha$ 는 필터의 평활화 계수로, 본 연구에서 0.5로 설정하였다.  $\alpha$  값을 크게 설정한 이유는 차량이 빠르게 이동하는 상황에서 실시간 응답이 중요했기 때문이며, 이를 통해 차량 감지 데이터가 실시간으로 반영되고 신호가 부드럽게 처리되어 시스템의 실시간 성능이 강화되었다.

### 3.3 DS18B20 온도 센서와 버터워스 필터

도로 노면의 온도 변화는 급격하게 발생할 수 있으므로 DS18B20 온도 센서에서 측정된 데이터를 처리할 때 2차 버터워스 저역통과 필터를 적용하였다. 버터워스 필터는 데이터의 노이즈를 줄이면서도 부드러운 신호를 제공하여 시스템의 안정성을 향상시킨다. 2차 버터워스 저역 통과 필터의 이산 시스템 동작을 나타내는 전달 함수는 다음과 같다.

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{a_0 + a_1 z^{-1} + a_2 z^{-2}} \quad (2)$$

시간 영역에서의 필터 적용은 다음과 같은 수식인 차분 방정식으로 표현된다.

$$y[n] = b_0 \cdot x[n] + b_1 \cdot x[n-1] + b_2 \cdot x[n-2] - a_1 \cdot y[n-1] - a_2 \cdot y[n-2] \quad (3)$$

이때 필터 계수는 차단 주파수  $f_c = 0.1\text{Hz}$ 와 샘플링 주파수  $f_s = 1\text{Hz}$ 에 따라 설정되었으며, 차단 주파수를 낮게 설정 함으로써 도로 노면의 온도 변화가 천천히 반영되면서도 노이즈를 효과적으로 제거하여 실제 도로 온도 변화에 민감하게 반응하도록 설계되었다.

### 3.4 염수 살포 및 동결 방지 시스템

도로 노면 온도가 3도 이하로 내려가고 수분 임계값인 600 이하에 도달했을 때만 염수 살포가 자동으로 이루어진다. 염수 살포 노즐에는 캡톤 필름 히터가 적용되어 염수 노즐의 동결을 방지하고 안정적인 작동을 보장한다. 또한 초음파 센서를 통해 결빙 구간 내 차량의 존재 여부를 실시간으로 감지하고 차량이 없을 경우에만 염수 살포가 이루어지도록 설계하여 염수 낭비와 차량의 부식을 방지하였다.

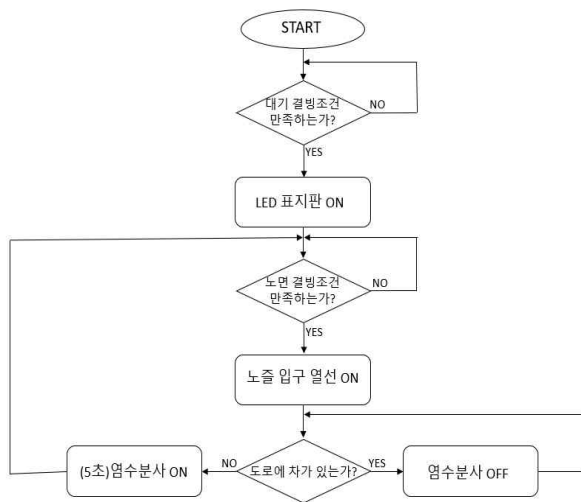


그림 2. 결빙 방지 시스템 순서도  
Fig 2. Frost Prevention System Flowchart

### IV. 시스템 구현 및 성능 평가

제안 시스템은 결빙 구간에서 실시간 환경 조건에 따라 LED 경고, 발열 장치 제어, 차량 감지, 염수 살포 등의 기능을 자동으로 수행한다.



그림 3. 결빙 방지 시스템 구현  
Fig. Implementation of Frost Prevention System

EMA 필터를 적용한 HC-SR04 초음파 센서는 차량 감지 정확도를 높였으며, DS18B20 온도 센서에 적용한 버터워스 필터는 노이즈를 줄여 데이터 안정성을 크게 높였다. 실험 결과, 제안된 시스템은 기존 시스템의 문제를 해결하며, 실시간 결빙 방지 기능을 효율적으로 수행하는 것을 확인할 수 있었다.

### V. 활용 방안 및 추가 연구 방향

본 시스템은 겨울철 결빙 방지뿐만 아니라 여름철 폭염 대응에도 활용 가능하다. 도로 온도가 33도 이상으로 상승할 경우, 염수 대신 물을 살포하여 도로 온도를 낮추는 방식으로 여름철 폭염 피해를 줄일 수 있다. 이를 통해 여름철 고온으로 인한 도로 열기를 줄여 환경 온도 조절과 더불어 시스템의 활용 범위를 넓힐 수 있을 것이다.

### VI. 결 론

본 연구에서 제안한 ATmega328 기반 자동화 결빙 방지 시스템은 기존 염수 살포 장치와 열선 시스템의 문제점을 해결하며, 저비용으로 구현 가능한 효율적인 대안을 제시한다. 실시간 데이터를 처리하는 EMA 필터와 데이터 안정성을 보장하는 버터워스 필터를 적용하여 시스템의 성능을 극대화하였으며, 다양한 도로 환경에서의 적용 가능성을 확인하였다. 향후 연구에서는 다양한 도로 환경에서의 추가적인 성능 평가 및 최적화를 통해 실질적인 개선을 기대할 수 있을 것이다.

### 참 고 문 헌

- [1] 도로결빙 교통사고 특성 분석, 2022년 도로교통공단 자료
- [2] 허경용, "ATmega328 프로그래밍", 제이펍
- [3] 이철희, "디지털 신호 처리", 한빛아카데미
- [4] McAndrew, Alasdair, "디지털 영상처리의 기초", 한티미디어

# 한국어 다중 감정 분석 및 공감 메시지 생성을 위한 BERT와 LLM 통합 접근법

김현수\*, 이경민\*, 정환일\*, 장선혜\*, 전태수\*\*

## An Integrated Approach for Korean Emotion Analysis and Empathy Message Generation Based on BERT and LLM

Hyun-soo Kim\*, Gyeong-min Lee\*, Hwan-Il Jung\*, Seon-hye Jang\*, and Tae-soo Jun\*\*

### 요약

본 연구는 한국어 기반 다중 감정 분석과 공감 메시지 생성을 위해 BERT와 LLM 모델을 결합한 통합 접근법을 제안한다. 개별 모델 성능 측정 후, 통합 모델 성능 측정을 진행하며, 개별 모델 실험 결과를 통해 통합 모델의 문제점 근거를 찾는다. 본 연구는 AICC와 감정 관리 솔루션 산업 발달에 따라, 향후 다중 감정 분석 및 공감 메시지 생성 시스템 개발에 참고할 방향성을 제공한다.

### Abstract

This study proposes an integrated approach that combines BERT and LLM models for Korean-based multi-emotion analysis and empathy message generation. After measuring the performance of each model individually, we assess the performance of the integrated model and identify its issues based on insights gained from individual model experiments. This research provides directional guidance for developing multi-emotion analysis and empathy message generation systems, especially as the AICC and emotion management solution industries continue to evolve.

### Key words

Multi-emotion Analysis, Message Generation, BERT, Large Language Model(LLM)

## I. 서론

최근 인공지능 기술의 발달로 인해 의료, 마케팅, 고객 서비스 같은 다양한 분야에서 사용자와의 상호작용 개선을 위해 정확한 감정 분석의 필요성

이 커지고 있다. 사용자에게 더 나은 경험과 몰입감을 제공하기 위해 감정과 상황을 반영한 응답 생성의 역할이 중요해졌다. 그러나 대부분의 연구는 감정 분석과 응답 생성의 일관성 유지를 목적으로 대규모 언어 모델(LLM)을 단독으로 사용하거나, 단

\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학부생, hutos@kumoh.ac.kr

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수, taesoo.jun@kumoh.ac.kr(교신저자)

순 긍정/부정 또는 단일 감정 분류에 그치는 경우가 많았다. [1][2]

모델 구조가 단순하고 작업 흐름 일관성을 유지할 수 있다는 점에서 LLM 단독 사용이 유리하다. 하지만, 모델을 호출하는 과정에서 메모리와 연산 자원을 더욱 많이 소비한다는 점, 감정 분류와 같은 명확한 라벨링 작업은 비교적 가벼운 BERT 모델이 특화되었다는 점에서 장점만 지녔다고 할 수 없다.

따라서, 본 연구에서는 다중 감정 분석과 모델 역할 분리, 메모리 경합을 목적으로 BERT와 LLM 통합 접근법을 제안한다. 대략적 모델 성능을 확인하고 통합 시 발생할 수 있는 문제를 파악하고자 BERT와 LLM을 각각 실험한 후에 통합 실험을 진행한다. 통합 접근법은 다음 순서로 진행된다.

BERT를 사용하여 사용자 입력 텍스트의 감정을 분노, 놀람, 행복, 공포, 혐오, 슬픔, 중립 7가지의 비율로 다중 감정 분석을 수행한다. 해당 결과와 텍스트를 LLM의 입력으로 사용하여 감정과 맥락을 고려한 공감 메시지를 생성한다.

해당 접근법은 구조가 복잡해지고 초기 설정이 부담될 수 있다는 단점이 있으나, 감정 분석과 메시지 생성 모델을 분리함으로써 메모리와 연산 자원의 효율성 향상을 강점으로 한다.

## II. 실험 방법

본 연구에서는 BERT와 LLM을 각각 테스트한 후, 두 모델을 통합하여 감정 분석과 공감 메시지를 생성하는 방식으로 실험을 수행한다.

### 2.1 BERT 실험 방법

실험에서는 `skt/kobert-base-v1`을 감정 분석 모델로 사용한다. [3] 이는 사전 학습된 모델로, 한국어 고유 언어 구조와 문법 및 감정 표현을 정확하게 학습하도록 설계되어 있다. 감정 분석을 위한 분류 작업에 최적화된 구조다. 한국어 고유의 어감을 잘 반영할 수 있어, 감정 분석 기본 성능도 기대할 수 있다.

감정 분석에 사용된 데이터셋은 AI Hub의 감성

대화 말뭉치와 감정 분류를 위한 음성 데이터셋으로, 총 57,968개의 데이터를 사용하였다. [4][5] 데이터는 모두 단일 감정 라벨을 포함하고 있으며, 훈련용과 테스트용을 8:2 비율로 분할하였다. 이를 통해 7가지 감정 패턴을 학습하여 감정 확률 분포를 예측하도록 한다.

성능 지표는 정확도(Accuracy), Macro F1-Score, Class-wise recall을 사용하였다. [6] 이러한 지표를 선택한 이유는 학습 추이를 확인하면서 클래스 불균형 문제를 종합적으로 평가하고, 특정 감정 클래스의 예측 성능을 개선하기 위함이다.

### 2.2 LLM 실험 방법

메시지 생성에는 `ElutherAI/polyglot-ko-3.8b` 모델을 사용하였다. [7] `Polyglot-ko`는 한국어에 특화된 LLM으로, GPT 계열 모델에 비해 경량화되어 있으면서도 한국어의 문법과 표현을 효과적으로 반영할 수 있다. 메시지 생성 정확도가 비교적 높을 것으로 기대할 수 있다.

학습 데이터는 AI Hub의 공감형 대화 데이터셋 중 친구 role 대화 데이터 4,799개를 사용하였다. [8] `BitsAndBytesConfig`, `PeftModel`, `LoraConfig`를 사용하여 양자화함으로써 메모리 효율을 높였다.

모델 수렴 정도를 평가하고자 손실값(Loss), 문장 생성의 자연스러운 정도 판단하기 위해 Perplexity, 응답성을 평가하고자 Inference Time, 자원 효율성 평가를 목적으로 GPU Memory를 모델 성능 평가 지표로 사용하였다.

### 2.3 BERT, LLM 통합 접근 실험 방법

앞서 학습된 BERT와 LLM을 결합하여 감정 분석 및 메시지 생성 통합 파이프라인을 구축하였다. 먼저 BERT가 입력 텍스트에 대한 감정 분석을 통해 7가지 감정에 대한 확률 분포와 대표 감정을 산출한다. 이 결과를 바탕으로 LLM이 메시지를 생성한다. 통합 모델은 감정 분포와 텍스트 맥락을 고려하여 적합한 메시지를 생성한다.

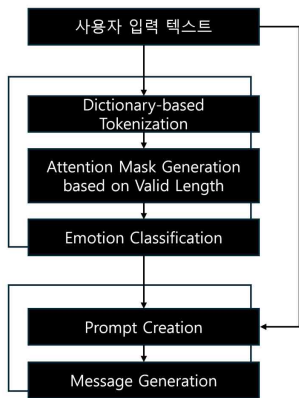


그림 1. 통합 파이프라인  
Fig. 1. Integrated pipeline

### III. 실험 및 결과와 분석

본 실험은 Google Colab 환경에서 NVIDIA L4 GPU(24GB 메모리) 및 CUDA 12.2 버전을 사용하여 수행하였다. 각 모델의 성능을 독립적으로 평가한 후, 통합 실험에서 성능과 최적화를 분석한다.

#### 3.1 BERT 실험 및 결과와 분석

학습 평균 응답 시간은 약 234.84ms로 측정되었다. 학습과 테스트 정확도 및 Macro F1-Score 간의

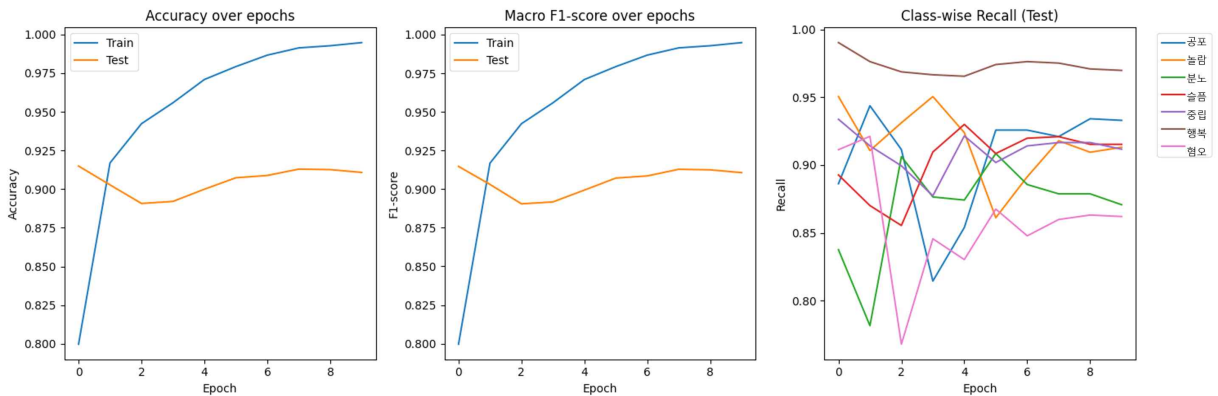


그림 2. BERT 성능 평가 지표  
Fig. 2. BERT Performance Evaluation Metrics

차이가 0.075~0.1 정도인 것을 보아, 과적합 가능성을 확인하였다. Class-wise Recall 결과, 혐오와 분노 감정에서 낮은 Recall을 보였다. 클래스별 성능 편차는 데이터 불균형이 주요 원인으로 추정되며, 향후 데이터 증강과 하이퍼파라미터 조정을 통한 성능 개선이 필요한 것으로 보인다.

#### 3.2 LLM 실험 및 결과와 분석

LLM 학습 결과, GPT-3나 GPT-4와 비교하여 Loss 값이 2.6488, Perplexity가 14.14로, 모델 크기가 2.7GB의 소형임을 고려할 때 응답 생성 정확도가 높고 비용 효율적이라고 평가할 수 있다. 하지만, 언어 모델링 측면에서는 다소 좋은 성능이라고 판단하기는 어렵다. 실시간 상호작용 시스템에 적용하여 사용자 경험 향상을 하고자 한다면, 해당 모델에 대하여 추가 학습, 하이퍼파라미터 최적화가 필요한 것으로 보인다.

표 1. LLM 학습 결과

Table 1. LLM Learning Outcome

Step	Loss	Perplexity
1130	2.6488	14.14
Inference Time		GPU Memory
598.72ms		5236.98MB

원기 내용을 입력하세요. (0 입력시 종료): 오늘 성능이 왜이리 별로니? 너 무슨 일 있지? 나 너무 걱정되고 스트레스 받아 일기에 대한 7가지 감정 : [['분노', '98.86%'], ['슬픔', '0.33%'], ['행복', '0.31%'], ['혐오', '0.21%'], ['공포', '0.17%'], ['놀람', '0.07%'], ['중립', '0.05%']] 공감 메시지 : 네가 그렇게 속상해 하는 거 보니까 내 마음도 안 좋다. 혹시 어디 불편한 데라도 있어? 기분은 좀 나아졌는지 모르겠네... 공감 메시지 생성 시간 : 17.82초

그림 3. 통합 테스트 결과  
Fig. 3. Integration test results



### 3.3 통합 접근 실험 및 결과와 분석

통합 테스트에서는 100자 입력 텍스트에 대하여 평균 응답 생성 시간이 약 18.18초 소요되었다. 이는 BERT와 LLM 통합 과정에서 최적화 문제가 있었음을 알 수 있다. BERT와 LLM 개별 실험에 근거하여 토큰화 중복, 모델 호출 중복, max\_length 제한, stop\_token 동작 오류로 인한 중복 응답 생성으로 인한 성능 저하가 원인이라고 추측된다.

## IV. 결론 및 향후 과제

본 논문에서는 사용자 텍스트 입력에 대해 skt/kobert-base-v1 모델로 다중 감정 분석을 수행하고, EleutherAI/polyglot-ko-3.8b 모델로 공감 메시지를 생성하는 통합 접근법을 제안하였다. BERT의 감정 분석에서 얻은 정확도, Macro F1-Score와 Class-wise recall은 실험 목적에 적합하다고 판단되나, 과적합 문제로 인해 일부 감정 클래스에서의 성능 저하가 확인되었다. LLM은 소형 모델로서 뒤쳐지지 않는 성능을 보이나, 실시간 상호작용 시스템으로 사용하기에는 추가 최적화가 필요하다. 따라서 해당 모델들을 고수한다면, 다중 감정 분석의 정확도 향상을 위해 더 많은 데이터 확보하고, 학습 epoch 증대해야 한다. 또한, 통합 과정에서 max\_length와 stop\_token 재설정 필요적이다.

하지만, small BERT와 BERT Tiny가 등장함에 따라 경량화된 모델을 사용하여 성능과 비용 효율성을 증대하거나, 고성능 프로세서 활용 가능 환경을 구축하여 보다 큰 LLM을 사용하는 것이 경제적이라고 판단된다. [9][10]

본 논문의 실험은 실용적인 성능을 보여주지는 못하였으나, 한국어 기반 BERT와 LLM 사용 예시를 제안하고, 감정 분석 및 공감 메시지 생성 통합 접근법의 방향성을 제시하는 근거 자료로 활용될 수 있을 것이다.

## 참고 문헌

[1] William Stigall and Md Abdullah Al Hafiz Khan,

"Large Language Models Performance Comparison of Emotion and Sentiment Classification", in ACM Southeast Conference. pp. 60-68, April 2024.

[2] Won-Min Lee and Byung-Won On, "Generating Emotional Sentences Through Sentiment and Emotion Word Masking-based BERT and GPT Pipeline Method", Journal of KIIT. Vol. 19, No. 9, pp. 29-40, Sep. 2021.

[3] SKTBrain, KoBERT, GitHub, <https://github.com/SKTBrain/KoBERT>

[4] 미디어젠, 감정 대화 말뭉치, AI-Hub, <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=&topMenu=&aihubDataSe=data&dataSetSn=86>

[5] KAIST 인공지능연구소, 감정 분류를 위한 대화 음성 데이터셋, <https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=263>

[6] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Introduction to Information Retrieval, Cambridge University Press. 2008.

[7] EleutherAI, polyglot-ko-3.8b, huggingface, <https://huggingface.co/EleutherAI/polyglot-ko-3.8b>

[8] 심심이(주), 공감형 대화, AI-Hub, <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=71305>

[9] Sangah Lee and Hansol Jang, "A Small-Scale Korean-Specific BERT Language Model", Journal of KIISE, Vol. 47, No. 7, pp. 682-692, Sep. 2020.

[10] Sanghyeok Boo and Kwangwon Song, "Korean Named Entity Recognition using BERT-Tiny", Journal of KCC, Vol 47, No. 1, pp. 1553-1555, Sep. 2020.

# 실시간 위치 기반 사진 공유 어플리케이션 아키텍처

김 승 민

## Real-time Location-Based Photo Sharing Application Architecture

Seungmin Kim

### 요 약

해당 논문에서 SNS 플랫폼 중 하나인 실시간 위치 기반 사진 공유 어플리케이션 아키텍처에 대해 간단히 이야기 한다. 아키텍처 안에 어떤 모듈들이 포함되어 있는지 이야기한다. (클라이언트-서버 관계) Amazon Web Service를 이용해 클라우드 환경(PaaS)에서 동작하는 서버는 Restful API 구축하고 이를 통해 클라이언트에게 서비스를 제공한다. 클라이언트에서 Google Map api를 이용해 LBS 형식의 서비스를 제공한다. 마지막으로 실시간 반영 방식에 대해 이야기한다.

### Abstract

In this paper, we briefly talk about the architecture of a real-time location-based photo sharing application, one of the SNS platforms. Talk about what modules are included in the architecture. (Client-server relationship) Servers operating in a cloud environment (PaaS) using the Amazon Web Service build a Restful API and provide services to clients through this. Clients use Google Mapi to provide services in LBS format. Finally, we will talk about real-time reflection methods.

### Key words

SNS, Restful API, Amazon Web Service, LBS

### 1. 서 론

현재 소셜 네트워크 서비스(이하 SNS라고 칭함) 어플리케이션은 국내 시장에 많은 비중을 차지하고 있다.[1] SNS 어플리케이션의 핵심 기능 중 하나는 실시간으로 정보를 공유하는 것이다. 주변에서 공유되는 정보를 접속 중인 사용자에게 요구되는 이벤트 없이 실시간 정보들이 반영되어야 한다. 본론에

선 이러한 요구 사항에 만족하는 시스템 아키텍처에 대해 말한다.

여기서 몇 가지 짚고 넘어가야할 개념적인 컨셉이 있다. 사진은 카메라로 촬영한 현실을 기록하는 기술이고, 이미지는 모든 종류의 그림을 의미한다. 사진보다 더 큰 개념이 이미지가 되겠다. 여기서 말하는 어플리케이션은 기존 SNS 어플리케이션과 달리 사진에 집중하고 이와 연계하여 여러 서비스를

---

\* 계명대학교 컴퓨터공학과, ak47001@naver.com

제공하는 SNS 플랫폼이다.

요구 사항에 대해 이야기 하자면 사진에 대한 위치를 알고 있다는 것이다. 접속중인 사용자 주변에 올라온 사진들을 실시간으로 볼 수 있으며 내가 올린 사진들도 주변 사람들에게 실시간으로 반영된다는 것이다. 그래서 LBS 형식의 서비스를 제공하는 어플리케이션이 된다.[2] 논문에서 다룰 주요 내용은 사진을 저장하고 어떻게 실시간 업데이트를 반영할 것인가이다. 전반적인 아키텍처와 그 안의 각종 모듈, 간단하게 API를 설명하고 마친다.[3]

- User Manager 모듈 : 사용자 로그인 및
- Photo Processor 모듈 : 사진 전처리
- Photo Manager 모듈 : 사진 데이터 가공
- Photo Store 모듈 : 사진 저장소 역할
- Chatting Scheduler 모듈 : 사용자 간 실시간 채팅
- Session Scheduler 모듈 : 서버와 클라이언트 간 스트림, 서로의 이벤트를 전달하기 위해 존재
- Folder Manager 모듈 : 사용자가 업로드한 이미지를 폴더 단위로 관리

## II. 실시간 공유 아키텍처

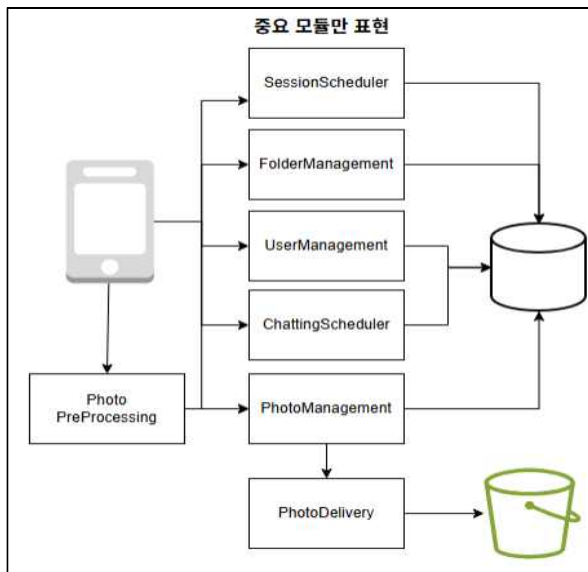


그림 1. 모듈 관계도  
Fig. 1. Module relation

실시간 위치는 두 단어로 분리하여 해석해야 된다. 실시간은 우리 어플리케이션에 접속 중인 사용자들이 서로가 공유한 사진을 사용자의 이벤트 없이 실시간으로 즉각 조회할 수 있음을 말하며 위치는 사진이 찍혔던 위치 즉, 해당 사진을 찍은 카메라의 위치가 된다. 사진에 대한 위치를 효과적으로 사용자들에게 보여주기 위해 구글 지도를 사용한다. 직관적으로 전달하기 위해 모바일 어플리케이션으로 구현하고 Flutter 프레임워크를 통해 Google API를 활용한다. 우리 어플리케이션에서 실시간 위치 정보를 바탕으로 사진에 정보를 주입한다. [4]

Session server는 클라이언트와 클라이언트 간의 연결 고리 역할을 수행하며 실시간 반영하게끔 하는 역할이다. 이로써 서버는 접속 중인 사용자의 실시간 위치를 추적할 수 있고 주변에서 발생한 사진 공유 이벤트를 다른 사용자들에게 즉각적으로 알릴 수 있게 된다.

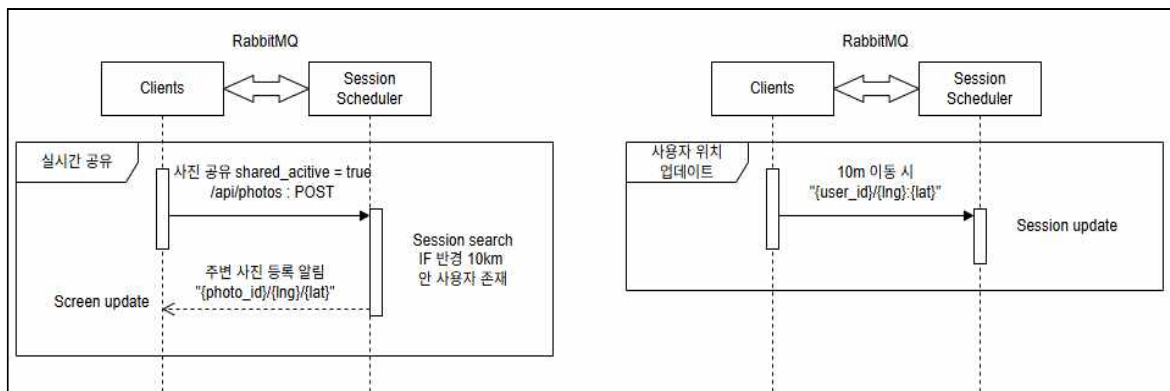


그림 2. 통신 흐름도  
Fig. 2. Communication Flow

### III. 결 론

결론적으로 우리는 실시간 위치 기반 사진 공유 플랫폼 전체적인 아키텍처에 대해 제시하고 추가로 실시간 반영에 대한 흐름을 제시한다. 실제 테스트 환경을 갖추고 AWS mysql과 연계하여 수십, 수백 만개의 이미지 데이터가 있을 경우 쿼리문 또는 서버 자체 로직을 어떻게 최적화하여 사용자들에게 응답을 할 것인지 생각해보아야 한다.

### 참 고 문 헌

- [1] 정지은, (2023.07.25), 한국 경제, <https://www.hankyung.com/article/2023072513741>
- [2] 김민정, 박영호. (2011). 안드로이드 기반 동적 위치관리 기능 구현. 한국정보처리학회 학술대회논문집, 18(1), 1501-1504.
- [3] Fielding, Roy Thomas. Architectural Styles and the Design of Network-based Software Architectures. Doctoral dissertation, University of California, Irvine, 2000.
- [4] 이준희, 김우생. (2010). Google Map API와 LBS를 이용한 소셜 네트워킹 사이트 설계 및 구현. 한국정보처리학회 학술대회논문집, 17(2), 381-384.

# 컴퓨터 비전을 활용한 자율주행 자동차

전혜린\*, 강성지\*, 이지현\*, 이원일\*\*

## A Self-Driving Car Using Computer Vision

Hye Lin Jeon\*, Sung Ji Kang\*, Ji Hyeon Lee\*, and Won Il Lee\*\*

### 요 약

자율주행 기술의 발전은 인적 요인에 의한 사고 감소와 교통 효율 향상에 기여하며, 이를 위해 주변 객체 인식 기능이 필수적이다. 본 논문에서는 교통 표지판, 신호등, 장애물 등 다양한 도로 환경에서의 객체 인식을 위해 데이터를 수집, 학습하여 정지, 주의, 좌·우회전 표지판 등의 인식을 통해 자율주행하는 자동차를 설계한다. 또한, 어두운 환경에서도 차선 인식이 가능하도록 영상처리를 통한 전조등 자동 제어 기능과 주행 중 과속 방지턱 인식을 통한 속도 제어를 통해 안전성을 높이고, GUI 시스템을 통해 주행 속도 조절과 정지 제어가 가능하도록 하여 사용자 편의성을 더한다. 이를 통해 본 논문에서는 자율주행 자동차가 복잡한 도로 환경에서도 효율적으로 주행할 수 있는 기반을 제공하고자 한다. 향후에는 자율주행 시스템의 안정성과 신뢰성을 더욱 향상시켜, 복잡한 도로 환경에서도 효율적이고 안전하게 주행할 수 있도록 지속적인 개선과 연구를 진행하고자 한다.

### Abstract

The advancement of self-driving car technology, supported by computer vision and artificial intelligence, contributes to reducing accidents caused by human factors and improving traffic efficiency. For this, object recognition capabilities are essential. In this study, we design an autonomous vehicle equipped with a Single Shot Detector (SSD) model to recognize various objects in diverse road environments, such as traffic signs, signals, and obstacles, by collecting and training data. Through the recognition of stop signs, caution signs, and left and right turn signs, the vehicle can navigate autonomously. Additionally, to enhance safety, the vehicle features automatic headlight control through image processing for lane recognition in low-light conditions, as well as speed control in response to speed bumps encountered while driving. A GUI system also allows users to adjust driving speed and control stopping, thereby increasing user convenience. This paper seeks to establish a foundation for the self-driving car to operate efficiently in complex road environments. In the future, we aim to continuously improve the stability and reliability of autonomous systems through ongoing research, ensuring safe and efficient driving even in complex road conditions.

### Key words

Self-Driving Car, Computer Vision, Single Shot Detector, Artificial Intelligence

---

\* 국립금오공과대학교 전자공학부(반도체시스템전공), heylin8282@naver.com, rktdjdwl112@naver.com, leejh10535@gmail.com

\*\* 국립금오공과대학교 전자공학부(반도체시스템전공), wilee@kumoh.ac.kr (교신저자)

※ 이 연구는 정부(교육부-산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0022161, 반도체전공트랙사업)

## I. 서 론

최근 자율주행 기술의 발전은 교통 안전성과 효율성의 혁신적 향상에 기여할 수 있는 핵심 기술로 자리 잡고 있다. 교통사고의 상당수는 운전자의 피로, 주의 산만, 판단 착오 등 인적 요인에 의해 발생하며, 이는 교통사고 감소와 더불어 교통 체증 완화, 연료 소비 절감 등의 이점을 기대하게 한다 [1]. 자율주행 기술은 이러한 이점들을 바탕으로, 스마트시티 구현과 친환경 교통 시스템 도입의 기반을 마련하고 있으며, 전 세계적으로 연구 및 실증이 활발히 진행되고 있다. 특히, 복잡한 도로 환경에서 주행하는 차량이 주변 객체를 인식하고, 주행 경로와 속도 조절을 효과적으로 수행하는 능력은 자율주행 시스템의 핵심 요소로 간주된다.

본 논문에서는 자율주행 차량의 주행 안정성을 높이기 위해 교통 표지판, 신호등, 장애물 등을 인식하는 시스템을 개발하였다. 도로 위 요소들을 데이터로 학습하여 정지, 주의, 좌·우회전 표지판 인식을 구현하고, 어두운 환경에서의 차선 인식을 위한 전조등 자동 제어 기능을 추가하였다. 또한, 과속방지턱 인식과 GUI 기반의 주행 속도 및 정지 제어 기능을 통해 사용자 편의성을 높였다.

## II. SSD 알고리즘을 활용한 자율주행

본 연구에서는 자율주행 자동차의 객체 감지 모델을 생성하기 위해 CNN(Convolutional Neural Network) 구조를 채택하고, GoogLeNet 기반의 SSD(Single Shot Detector) 알고리즘을 활용한다.

CNN은 이미지의 분류 및 특징 추출에 효과적인 신경망 구조로, 입력 데이터에서 특징을 추출하는 합성곱 계층(Convolutional Layer)과 추출된 데이터의 크기를 줄이는 풀링 계층(Pooling Layer)으로 구성된다[2][3]. 이러한 구조는 다양한 레이어를 통해 입력 데이터에서 특징을 추출하고 분류하는 데 강점을 가진다.

GoogLeNet의 SSD는 GoogLeNet 아키텍처를 기반으로 한 객체 감지 모델로, 인셉션 모듈을 활용하여 다양한 스케일의 특징을 효과적으로 추출하며, s

ingle shot 방식으로 객체를 동시에 탐지할 수 있는 장점을 갖고 있다[4]. 바운딩 박스 회귀와 클래스 예측을 통해 높은 정확도의 객체 감지가 가능하며, 빠른 처리 속도로 실시간 애플리케이션에 적합하다 [5].

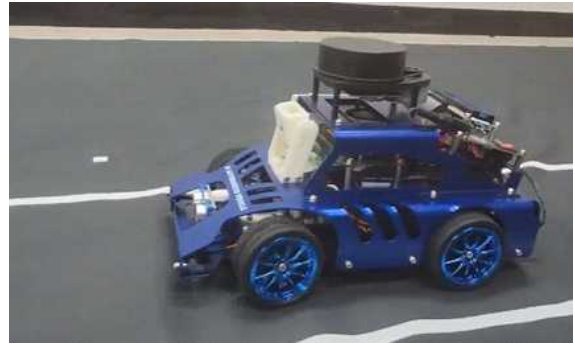


그림 1. 자율 주행 차량  
Fig. 1. Self-Driving Car

그림 1은 설계에 활용한 자율주행차로 차선 인식 및 객체 감지를 위한 Jetson Xavier 보드, 차량 운행을 위한 DC모터, 조향 제어를 위한 서보 모터, 자동차의 상태 표시 및 제어를 위한 LCD 모니터 등으로 구성되어 있다.

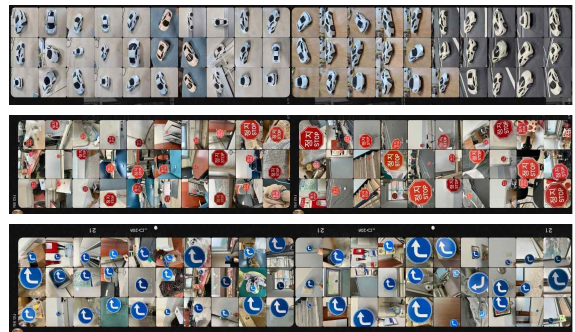


그림 2. 객체 데이터  
Fig. 2. Object data

```
Step 1: Loss/loss/loss=3.3746, Loss/classification_loss=0.7191 Loss/regularization_loss=0.0852, Loss/total_loss=4.1789, learning_rate=0.0040 (1.439 sec/step)
Step 2: Loss/loss/loss=3.3238, Loss/classification_loss=0.7228 Loss/regularization_loss=0.0852, Loss/total_loss=4.1318, learning_rate=0.0040 (1.887 sec/step)
Step 3: Loss/loss/loss=2.2166, Loss/classification_loss=0.6923 Loss/regularization_loss=0.0852, Loss/total_loss=3.9941, learning_rate=0.0040 (1.765 sec/step)
Step 148: Loss/loss/loss=0.1090, Loss/classification_loss=0.0671 Loss/regularization_loss=0.0852, Loss/total_loss=0.2583, learning_rate=0.0061 (1.902 sec/step)
Step 149: Loss/loss/loss=0.0674, Loss/classification_loss=0.0563 Loss/regularization_loss=0.0852, Loss/total_loss=0.2089, learning_rate=0.0061 (1.851 sec/step)
Step 150: Loss/loss/loss=0.0881, Loss/classification_loss=0.0651 Loss/regularization_loss=0.0852, Loss/total_loss=0.2385, learning_rate=0.0061 (1.863 sec/step)
Training done! 0:5:25
```

그림 3. 객체 감지 학습  
Fig. 3. Object Detection Training



그림 4. 데이터 학습 결과  
Fig. 4. Data Training Results

자율주행 자동차의 기본 주행을 구현하기 위해 정지, 주의, 좌·우회전 표지판, 신호등, 장애물(사람, 동물, 차)에 대해 각각 120개의 데이터를 수집하여 라벨링을 진행하였으며, Max Steps를 150, Batch Size를 24로 설정해 라벨링한 모델에 대한 학습을 진행시켰다.

자동차의 주행은 복합차로 주행을 목표로 하여 단일 차선 700개, 복합차선 500개의 데이터를 수집하고 라벨링 하여 학습을 진행하였다.

정지 표지판을 발견하면 멈추고, 횡단보도 표지판을 발견하면 10초 동안 속도를 낮춘 채 운행하도록 설정했다. 이와 함께 주행 중 적색등을 감지하면 정지하며, 정지 상태에서 녹색등을 감지하면 운행을 재개하도록 하였으며, 지시 표지판이 나타나면 지시 방향에 따라 좌·우회전을 하도록 설정하였다.

또한 동물이나 사람을 발견하면 즉시 정지하고, 자동차를 발견할 경우 정지 후 3초간 정적을 올리도록 했다. 안전한 주행을 위해 여러 차례의 테스트를 거쳐 이러한 기본 주행 기능을 완성했다.

또한, 어두운 환경에서 차선 인식의 어려움을 해결하기 위해 전체 픽셀의 평균값을 기준으로 전조등의 자동 제어 기능을 구현하여, 그림 5와 같이 어두운 환경에서도 주행할 수 있도록 구현했다.

주의 표지판 인식뿐만 아니라 그림 5와 같이 과속 방지턱 인식을 통해 과속 방지턱 앞에서 자동으로 속도를 감속하도록 하여, 운전자가 흔들림 없이 편안하고 안전하게 주행할 수 있도록 하였다.

마지막으로, 자동차의 GUI를 구현하여 현재 주행 속도를 표시하고, 사용자가 임의로 정지 및 속도 조절이 가능하도록 알고리즘을 추가하였다.

GUI 구현에 앞서, 효율적인 스레드 실행을 위해

파라미터 변경 및 함수 정의 등의 과정을 거쳐 프로그램의 안정성을 높였다.

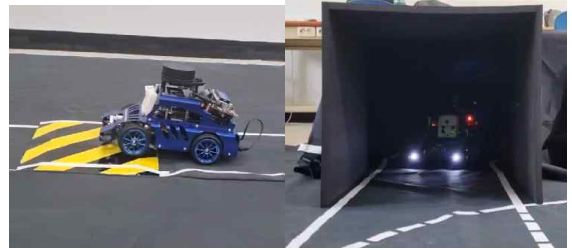


그림 5 좌-과속 방지턱, 우-터널 주행  
Fig. 5. Left-Speed Bump, Right-Tunnel Operation

### III. 결 론

본 연구에서는 효율적이고 신뢰할 수 있는 객체 감지를 위해 CNN 기반의 GoogLeNet SSD 알고리즘을 활용한 자율주행 자동차를 설계하였다. 다양한 주행 상황에서의 객체 인식 및 차선 인식 문제를 해결하고자 하였으며, 교통 표지판, 신호등, 장애물 및 차선 등에 대한 학습 및 주행 제어를 통해 자율주행 자동차가 다양한 도로 환경에서 안전하고 효율적으로 주행할 수 있는 기반을 마련하였다. 향후 연구에서는 객체 감지 정확도를 더욱 향상시키기 위한 다양한 데이터셋의 활용과 복잡한 교차로 및 다양한 주행 환경에서의 성능 평가를 포함하여 지속 가능한 발전을 도모할 계획이다.

### 참 고 문 헌

- [1] 안경환, 이상우, 한우용, “자율주행 자동차 기술 동향”, 전자통신동향분석 = Electronics and telecommunications trends, v.28 no.4, pp.35-44, (2013)
- [2] 임현국, “자율주행 차량 영상 기반 객체 인식 인공지능 기술 현황” 한국정보통신학회논문지, 25(8), 1117-1123. (2021)
- [3] Liu, Wei et al., "SSD: Single Shot MultiBox Detector," European Conference on Computer Vision (ECCV), (2016)
- [4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, "Deep Learning" (2016)
- [5] Liu, Wei et al(2016), Op. cit

# 학습 가능한 컨텍스트 벡터를 활용한 블라인드 올인원 영상 복원

이재훈\*, 김철환\*, 손창환\*\*

## Blind All-in-One Image Restoration Using Learnable Context Vectors

Jae-Hun Lee\*, Cheol-Hwan Kim\*, and Chang-Hwan Son\*\*

### 요약

고해상도 영상 복원 기술은 자율 주행 시스템에서 발생하는 다양한 열화 요인으로 인한 영상 품질 저하 문제를 해결하는 데 중요한 역할을 한다. 특히 비, 안개, 노이즈와 같은 외부 요인으로 인한 영상 손상은 자율 주행 시스템의 인식과 처리 성능에 직접적인 영향을 미치므로, 이러한 열화 요인을 효과적으로 제거하는 기술이 필수적이다. 본 연구에서는 여러 열화 요인을 동시에 처리할 수 있는 멀티모달 기반의 블라인드 올인원 복원 기법을 제안한다. 제안된 모델은 사전 학습된 CLIP 모델에서 생성된 학습 가능한 컨텍스트 벡터를 영상 복원 과정에 적용하여 다양한 열화 요인에 대한 복원 성능을 개선하였다. 실험 결과, 제안된 모델은 노이즈 제거, 비 제거, 안개 제거 작업에서 기존 모델에 비해 PSNR 및 SSIM 지표에서 우수한 성능을 나타냈다.

### Abstract

High-resolution image restoration technology plays an important role in solving the problem of image quality degradation caused by various deterioration factors in autonomous driving systems. In particular, image damage caused by external factors such as rain, fog, and noise directly affects the recognition and processing performance of autonomous driving systems, so a technology that effectively removes these deterioration factors is essential. In this study, we propose a multimodal-based blind all-in-one restoration technique that can simultaneously process multiple deterioration factors. The proposed model improves the restoration performance for various deterioration factors by applying the learnable context vector generated from the pre-trained CLIP model to the image restoration process. The experimental results show that the proposed model outperforms the existing models in PSNR and SSIM in noise removal, rain removal, and fog removal tasks.

### Key words

high-resolution image restoration, image degradation, blind restoration, all-in-one restoration model, learnable context vectors

---

\*국립군산대학교 소프트웨어학과, dlwogns410@naver.com, william36.naver.com

\*\* 국립군산대학교 소프트웨어학과(교신저자), cson@kunsan.ac.kr



## 1. 서 론

자율 주행 기술의 발전과 함께 고해상도 영상 처리는 안전하고 효율적인 운전 경험을 보장하는 데 필수적인 요소로 자리 잡고 있다. 그러나 차량이 주행하는 환경에서는 노이즈, 비, 안개와 같은 다양한 외부 요인으로 인해 영상 손상이 빈번히 발생한다. 이러한 문제를 해결하기 위해 영상 복원 기술의 중요성이 부각되고 있으며, 특히 여러 유형의 훼손을 동시에 처리할 수 있는 다중 열화 이미지 복원 모델의 필요성이 강조되고 있다.

이에 본 논문에서는 노이즈, 비, 안개와 같은 여러 훼손을 동시에 복원할 수 있는 올인원 영상 복원 모델을 제안한다. 특히, 사전 학습된 영상 및 텍스트 정렬 모델인 CLIP(Contrastive Language-Image Pretraining)[1]에서 학습된 이미지 인코더와 텍스트 인코더를 활용하여 동일한 쌍 간의 유사도를 최대화하고, 서로 다른 쌍 간의 유사도를 최소화함으로써 생성된 학습 가능한 컨텍스트 벡터를 활용한 멀티모달 기반의 블라인드 올인원 복원 기법을 소개하고자 한다.

블라인드 영상 복원은 열화의 유형과 강도에 대한 사전 정보 없이 손상된 영상을 복원하는 것을 의미하며, 이는 실제 운전 환경에서 발생할 수 있는 다양한 형태의 열화를 효과적으로 처리하는 데 필수적이다. 또한, CLIP의 멀티모달 특성을 활용한 학습 가능한 컨텍스트 벡터는 각 열화 유형에 특화된 복원을 가능하게 하여 자율 주행 시스템의 탐지 및 인식 성능을 크게 향상시킬 것으로 기대된다.

## II. 제안한 학습 가능한 컨텍스트 벡터를 활용한 블라인드 올인원 복원 기법

### 2.1 학습 가능한 컨텍스트 벡터

그림 1은 학습 가능한 컨텍스트 벡터를 생성하는 CoOp(Context Optimization) 모델이다. 이미지와 텍스트는 CLIP을 통해 사전학습된 각각의 인코더를 통과해 특징 벡터로 변환되고, 이를 코사인 유사도를 통해 비교하게 된다. 만약 매칭되는 쌍일 경우,

코사인 유사도를 최대화하고 매칭이 아닌 경우 최소화한다. 훈련 중, 사전 학습된 파라미터들은 고정시키고, 교차 엔트로피 손실을 최소화한다.

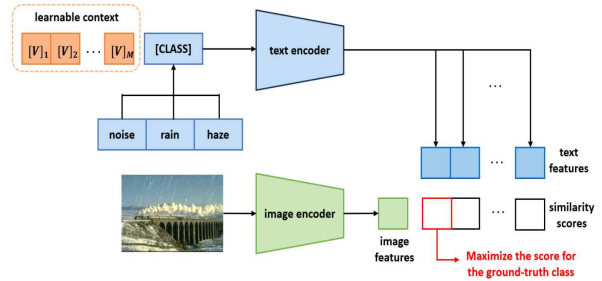


그림 1. 학습 가능한 컨텍스트 벡터를 생성하는 아키텍처  
Fig. 1. Architecture that generates learnable context vectors

학습 가능한 컨텍스트 벡터는 식 (1)의 클래스-특정 컨텍스트(CSC, Class-Specific Context)방식을 사용해 ‘노이즈’, ‘비’, ‘안개’ 각 클래스에 맞춤형 문맥 벡터를 학습하여 각 클래스에 특화된 시각적 표현을 생성한다.

$$[V_1^i | V_2^i | \dots | V_M^i] \neq [V_1^j | V_2^j | \dots | V_M^j] \quad (1)$$

for  $i \neq j$  and  $i, j \in \{1, \dots, K\}$

### 2.2 전체 파이프라인

그림 2는 제안한 모델 파이프라인이다. 열화된 입력 영상에서 컨볼루션을 통해 특징을 추출한다. 추출된 특징은 4단계 계층형 인코더-디코더 구조를 거치게 된다. 인코더는 고해상도 입력에서 시작하여 공간 해상도를 점진적으로 축소하고 채널 용량을 증가시켜, 저해상도 잠재 표현을 생성한다. 디코더는 해상도를 점진적으로 높이며, 저해상도를 고해상도로 변환한다. 이때, 프롬프트 블록과의 상호작용을 통해 열화 정보를 반영한 프롬프트가 각 디코더 레벨에 전달되어, 디코더는 열화 유형에 맞는 복원 작업을 수행한다. 프롬프트 블록은 열화된 영상을 복원할 때 복원을 유도하는 플러그인 모듈이며 PGM(Prompt Generation Module)과 PIM(Prompt Interaction Module)로 구성되어 있다. PGM은 먼저 입력 특징  $F_l$ 에 대해 글로벌 평균 풀링(GAP)를 적

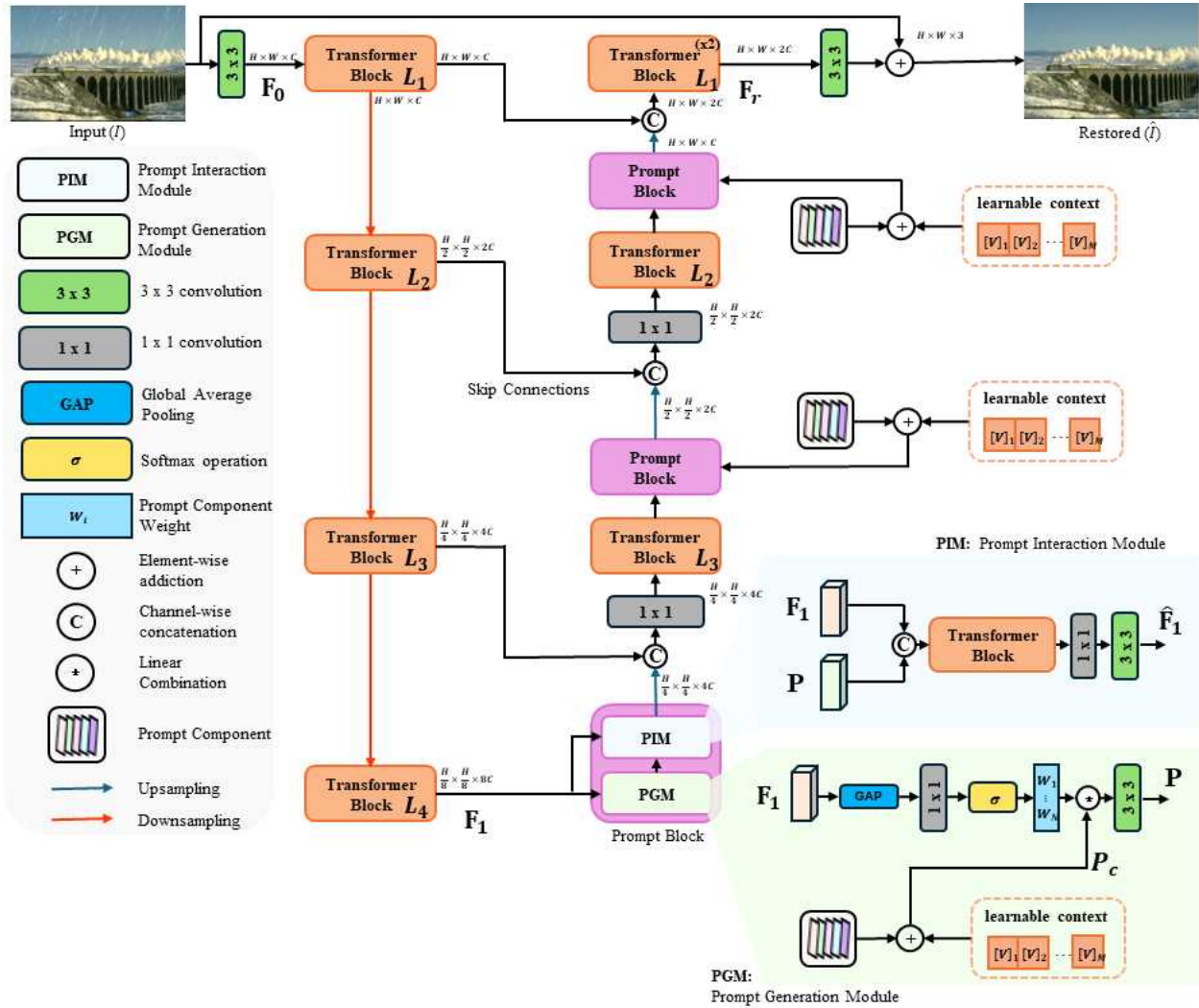


그림 2. 제안한 컨텍스트 벡터가 결합된 블라인드 올인원 복원 모델  
 Fig. 2. Proposed blind all-in-one restoration model combined with context vectors

용하여 특징 벡터를 생성한다. 그리고 1x1 컨볼루션을 통과한 후 소프트맥스 연산을 수행하여 프롬프트 가중치를 생성한다. 생성된 가중치는 이전에 학습된 학습 가능한 컨텍스트 벡터와 행렬 간 덧셈을 수행한 프롬프트 구성요소에 곱해지고 3x3 컨볼루션을 거쳐, 입력 영상 특성에 따라 동적으로 조정된 프롬프트  $P$ 를 생성한다. 프롬프트 구성요소는 학습 가능한 파라미터로, 입력 특징과 상호작용하여 열화 유형 정보를 포함한다. 본 논문에서는 PromptIR 모델의 프롬프트 구성 요소와 학습 가능한 컨텍스트 벡터를 활용해 프롬프트가 열화 상황별 정보를 보다 풍부하게 반영해 입력 특징에 대한 복원 성능을 극대화하였다. PIM은 입력 특징  $F_i$ 와 생성된 프

롬프트  $P$ 를 채널 차원에서 결합한 후 트랜스포머 블록을 통과시켜 최종 출력 특징  $\hat{F}_i$ 를 생성함으로써 열화 정보를 반영하고, 이를 통해 동적으로 복원 성능을 최적화하는 가이드 역할을 수행한다.

### III. 실험 및 결과

본 논문에서는 ‘노이즈’, ‘비’, ‘안개’ 클래스를 집중적으로 복원한다. 노이즈 제거는 BSD400과 WED 데이터셋 4,744장을 훈련에 사용하고, BSD68과 Urban100 168장을 테스트에 사용하였다. 노이즈는 손상 수준  $\sigma = 15, 25, 50$ 의 백색 가우시안 노이즈를 추가해 생성하였다.



(a)

(b)

(c)

그림 3. 결과 영상, (a) 노이즈 제거, (b) 비 제거, (c) 안개 제거, (첫 번째 행) 원본 영상, (두 번째 행) 열화된 영상, (세 번째 행) Airnet[2], (네 번째 행) PromptIR[3], (다섯 번째 행) 학습 가능한 컨텍스트 벡터만 학습, (여섯 번째 행) 제안한 모델.

Fig. 3. Resulting images, (a) Denoising, (b) Deraining, (c) Dehazing, (First row) Original images, (Second row) Degraded images, (Third row) AirNet[2], (Fourth row) PromptIR[3], (Fifth row) Learning only the learnable context vector, (Sixth row) Proposed model.

표 1. 정량적 평가

Table 1. Quantitative evaluation

PSNR/SSIM	Denoising $\sigma = 15$	Denoising $\sigma = 25$	Denoising $\sigma = 50$	Deraining	Dehazing
Airnet	33.91/0.9394	31.32/0.9044	28.02/0.8302	34.51/0.9635	27.35/0.9610
PromptIR	33.67/0.9584	31.10/0.9330	27.85/0.8818	35.35/0.9857	28.75/0.9645
PromptIR (context V)	33.95/0.9611	31.29/0.9362	27.98/0.8859	37.50/0.9909	30.22/0.9729
<b>Proposed Model</b>	<b>33.99/0.9617</b>	<b>31.33/0.9370</b>	<b>28.04/0.8874</b>	<b>37.83/0.9919</b>	<b>31.04/0.9780</b>

비 제거는 Rain100L 데이터셋에서의 원본, 비 영상 200쌍을 훈련에 사용하였고, 100개의 쌍을 테스트에 사용하였다. 안개 제거는 OTS 데이터셋 72,135장을 훈련에 사용하였고, SOTS 데이터셋 500장을 테스트에 사용하였다. 학습에 사용한 배치 크기는 32, 에폭은 50으로 설정하였다. 또한, 학습률은 0.0002로 설정하였고, 가중치 갱신을 위해 아담 옵티마이저 (Adam Optimizer)를 사용하였다. 화질 평가 지표로는 최대 신호 대 잡음 비(PSNR, Peak Signal-to-Noise Ratio)와 구조적 유사성(SSIM, Structural Similarity)를 사용하였다. PSNR은 원본 영상과 복원 영상의 화소 값이 얼마나 유사한지 평가하는 척도이고 구조적 유사성은 영상 구조나 텍스처의 복원 능력을 판별하는 척도이다.

그림 3은 기존의 모델과 제안한 모델을 사용하여 얻는 결과 영상이다. PromptIR(context V)모델은 기존 PromptIR 모델의 프롬프트 컴포넌트를 제외하고 학습가능한 컨텍스트 벡터를 추가하여 학습을 한 모델이다. 노이즈 영상은 제안한 모델의 흉바닥이 더 선명한 것을 확인할 수 있고 비 영상은 제안한 모델의 하늘에 빗줄기가 더 잘 제거된 것을 확인할 수 있다. 또, 안개 영상은 사람의 머리 부분의 안개가 더 잘 제거된 것을 확인할 수 있다.

표 1은 테스트 집합에 대해서 PSNR과 SSIM을 정량적으로 평가한 지표이다. 결과를 보면, 제안한 모델의 PSNR과 SSIM의 수치가 제일 높은 것을 확인할 수 있다. 이 결과는 학습 가능한 컨텍스트 벡터가 복원 성능을 향상하는데 효과적이었음을 의미한다.

#### IV. 결론

본 논문에서는 학습 가능한 컨텍스트 벡터를 결합한 블라인드 올인원 복원 기법을 소개하였다. 제안한 모델은 PromptIR에 비해 노이즈는 PSNR이 평균 0.336 정도 향상되었고, 비는 2.48이 향상되었다. 또한, 안개는 2.29가 향상되었다. 이를 통해 노이즈, 비, 안개와 같은 열화 유형을 효과적으로 복원할 수 있음을 입증하였다. 실험 결과, 학습 가능한 컨텍스트 벡터를 프롬프트 컴포넌트와 결합하여 각 열화 유형에 맞는 특징을 반영함으로써, 기존 복원 모델 대비 향상된 복원 성능을 확인하였다.

#### V. 참고문헌

- [1] A. Radford et al. "Learning transferable visual models from natural language supervision," International Conference on Machine Learning, PMLR 139:8748-8763, 2021.
- [2] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv, and X. Peng, "All-in-one image restoration for unknown corruption," in Proc. CVPR, pp. 17452 - 17462, Jun 2022.
- [3] V. Potlapalli, S. W. Zamir, S.Khan, and F.S.khan, "PromptIR: Prompting for All-in-One Blind Image Restoration" , arXiv:2306.13090, Jun. 2023.

# 생성형 인공지능을 활용한 네트워크 데이터 분석 및 시각화

장은서\*, 최권택\*\*

## Network data analysis and visualization using generative AI

Eunseo Jang\*, and Kwontaeg Choi\*\*

### 요 약

Graph DB는 데이터 간 관계를 직관적으로 표현할 수 있다는 장점이 있지만 쿼리 언어에 대한 이해도가 있는 사용자만 이용이 가능하다는 제약이 있다. 본 연구는 사용자의 자연어 입력을 Cypher로 변환하고, 시각화 및 해석 결과를 제공하여 비전문가도 Graph DB를 쉽게 활용하고 데이터 분석에 도움을 받을 수 있도록 한다. 시각화와 해석 결과를 제공하는 시스템을 통해 사용자는 복잡한 데이터 관계를 직관적으로 이해할 수 있으며 대화형 인터페이스를 통해 연속적인 데이터 탐색이 가능할 것이다. 이를 통해 Graph DB의 접근성을 높이고 Cypher에 대한 지식 없이도 데이터 분석과 의사 결정에 도움을 주고자 한다.

### Abstract

Graph DB has the advantage of being able to intuitively express relationships between data, but it has the limitation that it can only be used by users who have an understanding of the query language. This study converts the user's natural language input into Cypher and provides visualization and interpretation results so that even non-experts can easily use Graph DB and receive help in data analysis. A system that provides visualization and analysis results will allow users to intuitively understand complex data relationships and enable continuous data exploration through an interactive interface. Through this, we aim to increase the accessibility of Graph DB and help with data analysis and decision-making without knowledge of Cypher.

### Key words

Graph DB, Query language, visualization, Cypher, generative AI, natural language

### 1. 서 론

인터넷의 급격한 확장과 다양한 기술의 발달은 방대한 양의 데이터 확산을 가져옴과 동시에 점차 다양한 형태로의 변화를 일으켰다. 이러한 데이터는

복잡한 구조를 갖추고 있어 기존의 관계형 데이터 베이스(RDB)만으로는 효율적으로 처리하고 유의미한 관계를 파악하는 데 한계가 있다. 이에 따라 데이터 처리와 표현에 대한 새로운 접근이 요구되며 그 해결책 중 하나로 Graph DB가 주목받고 있는

\*강남대학교 소프트웨어융합학부, ccolornyong@gmail.com

\*\*강남대학교 ICT융합공학부, kwontaeg.choi@kangnam.ac.kr (교신저자)

실정이다[1].

Graph DB는 그래프 구조로 데이터를 저장하고 관리하는 시스템으로 노드(Node)와 엣지(Edge)를 통해 데이터 간의 관계를 명시적으로 표현한다. 이와 같은 특징으로 빠른 관계 탐색을 지원하고 높은 확장성을 제공하여 다양한 규모의 데이터를 유연하게 처리할 수 있도록 한다. 뿐만 아니라 데이터 간 관계를 직관적으로 시각화할 수 있도록 지원하여 데이터의 상호작용과 연결성을 쉽게 이해할 수 있다 [2].

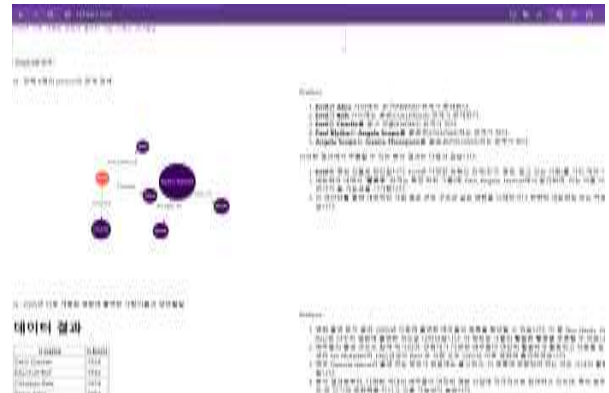
그러나, 이러한 Graph DB의 사용을 위해서는 Graph DB마다 사용되고 있는 언어들에 이해하고 다룰 수 있어야 하며 별도의 설치와 설정 과정이 필요하다는 단점이 있다. 이에 따라 몇 가지 Graph DB 활용 툴이 제공되고는 있지만 대체로 텍스트 형식의 아웃풋을 제공하기 때문에 Graph DB의 장점인 시각화 기능을 충분히 살리지 못하고 분석의 수준이 단순한 질의에 그치는 한계가 있었다.

본 논문에서는 생성형 인공지능을 활용하여 DB 언어를 모르는 일반 사용자도 Graph DB를 다루고 데이터 분석 및 의사 결정에 도움을 받을 수 있는 기술을 제안하고자 한다.

## II. 제안 방법

본 연구에서는 사용자가 입력한 자연어를 Neo4j에서 사용하는 Cypher 언어로 변환하여 데이터를 시각화하고 이를 사용자에게 해석과 함께 제공하는 방법을 제안한다. 자연어를 Cypher로 변환하는 과정과 시각화 결과를 자연어 기반으로 해석하는 과정에서 생성형 인공지능을 활용하며 정확하고 일관된 결과를 얻을 수 있도록 프롬프트 엔지니어링을 수행한다. Graph DB로는 Neo4j를 활용하며 생성형 인공지능 모델은 Chat-GPT 4를 활용한다. 또한 본 연구에서는 Graph DB에 익숙하지 않은 사용자도 해당 기술을 사용할 수 있도록 React와 Spring을 활용하여 인터페이스를 구성하였다. 시스템을 구성하는 단계는 크게 자연어의 Cypher 언어 변환, 쿼리 결과 시각화, 생성형 인공지능을 활용한 데이터 분석, 대화형 시스템과 같은 4가지로 이루어지며 다음

과 같다.



### 2.1 자연어의 Cypher 언어 변환

사용자가 입력한 자연어를 Cypher로 변환을 수행한다. 기본적으로 생성형 인공지능은 사용자의 의도를 파악해 Cypher 쿼리를 생성할 수 있으나 Neo4j에 대한 이해가 부족하고 SQL의 문법과 Cypher를 혼동하는 경우가 있다.

이를 해결하기 위해 프롬프트 엔지니어링을 수행하였으며 Cypher에 대한 명확한 정의를 제공하고 몇 가지 변환 예시를 추가하는 방식으로 쿼리 생성의 정확도를 높였다. 최종 결과 또한 JSON 형식의 포맷으로 제한하여 결과의 일관성을 유지하고 분석에 필요한 데이터를 구조화하여 사용자에게 제공할 수 있도록 하였다. 사용자가 자연어로 “5명의 person의 관계 검색”을 요청하면 시스템은 다음과 같이 Cypher로 해당 문장을 변환한다.

```
MATCH (n:Person)-[r]->(m) RETURN n, r, m LIMIT 5
```

사용자가 “2005년 이후 개봉된 영화에 출연한 사람 이름과 생년월일”을 요청하면 시스템에서는 다음과 같은 Cypher 언어로 변환을 수행한다.

```
MATCH (n:Person)-[r:ACTED_IN]->(m:Movie) WHERE m.released > 2005 RETURN n.name, n.born
```

### 2.2 Cypher 쿼리 결과 시각화

복잡한 구조의 관계형 데이터베이스(RDB)와 다르게 Graph DB는 노드와 엣지 간의 관계를 직관적으

로 표현할 수 있다. 본 연구에서는 이러한 Graph DB의 특성을 활용하여 사용자에게 시각화 된 답변을 제공하고 사용자가 직접 데이터 시각화에 관여할 수 있는 인터랙티브한 결과물을 제공하고자 하였다. 이용자가 요구하는 데이터의 특성에 따라 각 데이터 간 관계 구조 관련 요청이면 네트워크 그래프 형식, 이름이나 속성 등의 검색이면 테이블 형식으로 시각화가 진행된다.

특정 데이터에서 “5명의 perseon의 관계검색” 을 요청하면 그래프 형식으로 시각화가 진행되며 두 개의 노드 그룹이 생성되고 각 노드 간의 관계를 엮기를 통해 쉽게 이해할 수 있도록 명시된다.

“2005년 이후 개봉된 영화에 출연한 사람 이름과 생년월일” 을 요청하면 테이블 형식으로 시각화가 진행되며 해당 요구사항에 맞는 객체들이 이름과 생년월일 구조로 출력된다. 해당 시각화는 이용자가 요청한 데이터의 복잡도에 따라 좀 더 디테일하고 다양해질 수 있다.

### 2.3 Graph DB 결과 해석 및 분석

시각화된 결과물을 제공하는 것만으로는 사용자에게 충분한 데이터를 전달하기 어렵기 때문에 본 연구에서는 사용자에게 실질적인 인사이트 제공과 깊이 있는 이해를 돕고자 시각화 결과의 해석을 함께 제공한다. 생성형 인공지능을 활용하였으며 시각화된 데이터를 심도 있게 해석할 수 있는 출력을 생성하도록 특정 해석 포인트와 출력 형식을 지정하여 구체화하는 프롬프트 엔지니어링을 수행하였다. 첫 번째 예시인 “5명의 perseon의 관계 검색” 에 대한 시각화 해석에서는 연결 관계의 분류 설명을 통해 각 노드의 관계를 정리하였으며 그래프의 특이성을 찾아내어 주요 인물이라고 볼 수 있는 두 노드 그룹의 중점을 설명한다. 시각화에서 볼 수 있는 노드와 관계의 표현을 정의하였으며 추가적인 제안 사항을 제공하고 있다.

“2005년 이후 개봉된 영화에 출연한 사람 이름과 생년월일” 에 대한 시각화 해석에서는 단순한 테이블 설명 외에도 인기와 영향력에 관한 분석을 제안하며 추가적인 해석을 제공하였다.

### 2.4 대화형 시스템

사용자가 자연스러운 대화 형태로 데이터에 접근하고 분석할 수 있도록 챗봇 형식을 차용했다. 기존 Noe4j 브라우저와 같은 일반적인 Graph DB 도구는 쿼리 실행 후 그 결과만을 제공하며 작업 로그가 남지 않기 때문에 사용자가 분석 과정을 이해하고 일관성을 유지하기 어렵다는 단점이 있다. 본 연구에서 제안하는 챗봇 형식은 사용자와의 대화 기록을 보존하여 데이터 분석의 연속성을 확보하고 심화된 분석을 수행할 수 있도록 지원한다. 화면 구성을 보면 상단에 질의를 위한 텍스트 창이 위치하고 있고 하단으로 질문에 대한 답변이 쌓이는 형식으로 구현하였다.

## III. 결 론

본 연구에서는 자연어를 Cypher 쿼리로 변환하여 Graph DB 데이터를 시각화하고 해석하는 대화형 분석 시스템을 제안하였다. 생성형 인공지능과 프롬프트 엔지니어링을 활용하여 사용자가 복잡한 관계 데이터를 직관적으로 탐색할 수 있도록 하였으며, 대화 기록을 통해 분석의 연속성을 유지함으로써 심층적인 데이터 탐구를 가능하게 하였다. 이 시스템은 Graph DB의 접근성을 높이고 비전문가도 쉽게 데이터 분석을 수행할 수 있는 환경을 제공한다. 앞으로 더욱 다양한 분석 요구를 충족할 수 있는 확장 연구가 기대된다.

## 참 고 문 헌

- [1] Guia,J., Gonçalves Soares,V. and Bernardino,J. (2017) Graph databases: Neo4j analysis. In: Proceedings of the 19th International Conference on Enterprise Information Systems, Vol. 1, SCITEPRESS - Science and Technology Publications, Porto, Portugal. pp. 351-356.
- [2] Miller,J.J. (2013) Graph database applications and concepts with Neo4j. In: Proceedings of the Southern Association for Information Systems Conference, Vol. 2324, AIS, Atlanta, GA, p. 36.





# 지속적인 원예 활동을 돕는 허브식물 재배 보조 앱

이윤희\*, 임가영\*, 김차종\*\*

## Herb cultivation assistance application that helps with continuous gardening activities

Yunhee Lee\*, Gayoung Lim\*, and Chajong Kim\*\*

### 요 약

원예 활동의 활성화를 위한 기존 식물 케어 관련 연구 및 서비스는 지속적인 참여를 유도하는 요소가 부족했다. 본 논문에서는 재배 실패 확률을 낮추고 흥미 요소 및 동기 부여를 제공하여 지속적인 원예 활동을 유도하는 애플리케이션을 개발하였다. 본 애플리케이션을 직접 사용해 본 결과, 허브 재배 과정에서 흥미와 동기 부여를 지속적으로 유지할 수 있었으며, 캐릭터 성장 기능과 허브 활용 방법 추천 기능이 원예 활동에 대한 몰입도를 높이는 데 기여함을 확인할 수 있었다. 이러한 실용성과 실효성은 사용자들의 원예 활동 지속성을 높이는 데 중요한 역할을 할 것으로 기대된다.

### Abstract

Existing plant care-related research and services aimed at promoting horticultural activities lacked factors that encouraged continuous participation. In this paper, we developed an application designed to sustain horticultural activities by reducing the likelihood of cultivation failure and providing elements of interest and motivation. Through direct use of this application, it was confirmed that interest and motivation could be continuously maintained during the herb cultivation process, and that the character growth feature and the herb usage recommendation feature contributed to enhancing user engagement in horticultural activities. This practicality and effectiveness are expected to play an important role in improving the continuity of users' horticultural activities.

### Key words

herb, horticulture therapy, gamification, App Service

## 1. 서 론

일반적으로 원예 활동이 ‘주변에 식물을 두고 가꾸는 일’로 간주할 때, 원예치료는 주로 ‘활동’을 주요 치료 수단으로 하고, 이에 따른 ‘상호작용’과 ‘반응’을 주요 원리로 하는 원예 활동의 구조화를 통

해 치료 효과를 얻는 것을 목표로 한다[1].

원예 활동은 자연을 통해 정신적, 신체적 재활, 사회 상호작용, 지각 자극 등에 많은 도움을 주고 스트레스 감소와 정서 안정에 효과가 있어, 이를 보조하는 식물 케어 서비스 연구와 애플리케이션 개발이 이루어지고 있다[2].

\* 국립한밭대학교 컴퓨터공학과, 20211889@edu.hanbat.ac.kr, 20211925@edu.hanbat.ac.kr

\*\* 국립한밭대학교, 컴퓨터공학과 교수(교신저자), cjkim@hanbat.ac.kr

그러나 각종 센서가 요구되며 사용자와 식물 간 물리적으로 떨어진 환경을 상정하기 때문에 원예 활동을 목적으로 하는 일반적인 가정에 적용하기 어려운 한계가 있다. 이에 더해 식물에 대한 기능성 정보와 흥미를 유발하는 요소에 부족함이 있다. 지속적인 원예 활동 유도하기 위해서는 흥미 유발과 동기를 부여할 수 있는 요소를 고려해야 한다[3].

원예 작물에는 과일, 채소, 화훼작물, 그리고 허브와 같은 약용 식물 등이 포함된다. 특히 허브는 소규모 실내 공간에서 재배 가능하며 오감을 자극하는 특성을 가졌다. 또한, 허브는 압화, 포푸리, 드라이플라워 등의 공예작업뿐만 아니라 차로 우려내거나 요리 재료로 활용되는 등 그 이용 범위가 매우 넓다[4].

본 논문에서는 위와 같은 이점을 가진 허브를 중심으로, 원예 활동에 지속적인 관심을 기울이게 하는 방안을 분석하여 개발한 재배 보조 애플리케이션을 소개한다. 애플리케이션이 제공하는 다양한 기능을 통해 사용자가 허브 재배에 꾸준히 참여하도록 유도하며 최종적으로 애플리케이션의 사용자가 허브 재배를 통해 긍정적인 변화를 경험할 수 있도록 돕는 것을 목표로 한다.

## II. 본론

### 2.1 기능 설계

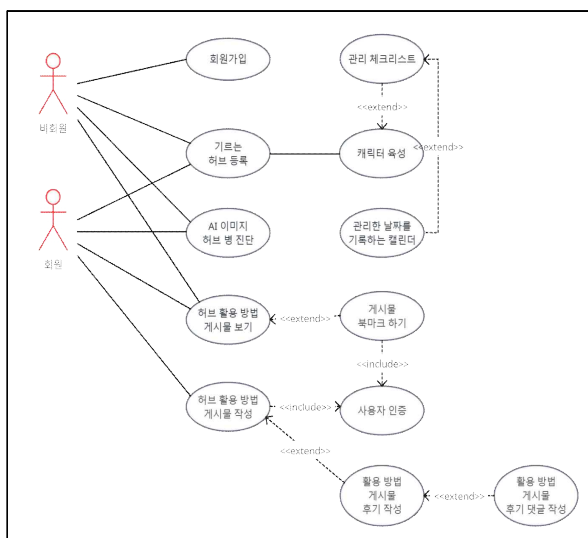


그림 1. 애플리케이션 유스케이스 다이어그램  
Fig. 1. Application use case diagram

#### 2.1.1 허브 재배 보조 기능

이 기능은 사용자들의 허브 재배를 돕는 것을 목적으로 설계되었다(그림 1). 허브 육성 정보를 제공하며, 관리 활동을 기록하는 체크리스트와 캘린더 기능을 포함하였다. 또한, 재배 과정에서 발생할 수 있는 병해충을 해결하기 위해 병 진단 기능을 추가하였다. 이러한 기본적인 재배 보조 기능은 허브 육성을 수월하게 하여 사용자가 식물 기르기에 어려움을 겪어 포기하는 상황을 방지하는 데 도움을 줄 것이다.

#### 2.1.2 캐릭터

사용자가 기르는 허브를 등록하면 캐릭터를 육성할 수 있도록 하였다. 각 캐릭터는 사용자가 관리하는 허브의 관리 빈도에 따라 성장하며, 허브의 효능과 관련된 스토리를 가지고 있다. 이 기능은 원예 활동과 애플리케이션에 대한 흥미를 유발하고, 사용자의 지속적인 참여를 촉진하기 위해 설계되었다.

#### 2.1.3 허브 활용 방법 추천 기능

허브를 활용할 수 있는 다양한 방법을 추천받고 공유할 수 있도록 탭을 설계하였다. 이를 통해 수확한 허브의 활용 방안을 제시하고, 허브의 기능성 정보를 충족하도록 하였다. 기본적으로 허브차와 포푸리 제작 방법을 제공하며, 사용자는 추가적으로 게시판을 통해 자신만의 허브 활용 방법을 공유할 수 있다. 해당 게시판이 활성화될수록 사용자 간 정보 교환이 활발해져 허브 육성과 활용 방법에 관심이 높아질 것으로 예상된다. 이 기능은 소통의 수단을 마련하여 지속적인 애플리케이션 사용을 유도하고, 허브 활용이 원예 활동의 동기 부여 요소로 작용할 수 있도록 돕는다.

## 2.2 기능 개발

#### 2.2.1 개발 환경 및 도구

프론트엔드 개발에는 Flutter를 선택하였다. Material Design을 기본으로 제공하여 UI 디자인 시간을 절감할 수 있었다. Flutter의 풍부한 위젯 지원은 일관성 있는 사용자 인터페이스 구현을 편리하게 하였다.

백엔드 서비스와 데이터 처리를 위해서는 Firebase를 사용하였다. Firebase Realtime Database를 통해 데이터를 실시간으로 저장하고 즉시 사용자에게 반영하는 기능을 구현하였다. 데이터는 Firebase의 Authentication 기능에서 제공하는 고유한 UID를 기준으로 각 사용자별로 관리된다. UID를 기반으로 사용자별로 북마크 한 게시물, 업로드한 게시물(posts), 그리고 후기를 구조화하여 저장하였다.

### 2.2.2 UI/UX 구현

홈 화면에는 사용자가 자주 사용하는 기능인 캐릭터, 관리 체크리스트, 허브 정보 확인 기능을 배치하였다. 특히 캐릭터는 홈 화면 중앙에 배치하여 사용자가 애플리케이션과 허브 관리에 몰입하도록 유도하였다.

체크리스트와 허브 정보는 다이얼로그 형태로 제공하여 화면 이동 없이 즉각적으로 확인할 수 있게 하였다. 사용자가 불필요한 화면 전환으로 느끼는 피로도를 줄이고, 더 나은 사용자 경험을 제공하도록 설계하였다.

허브 활용 정보 화면의 목록은 한눈에 쉽게 파악할 수 있도록, 제목과 중간 크기 이미지가 포함된 카드 형식으로 정보를 제공하였다. 카드를 선택하면, 해당 항목의 상세 설명을 확인할 수 있도록 하여, 사용자가 정보를 더 깊이 탐색할 수 있게 하였다.

### 2.2.3 결과물

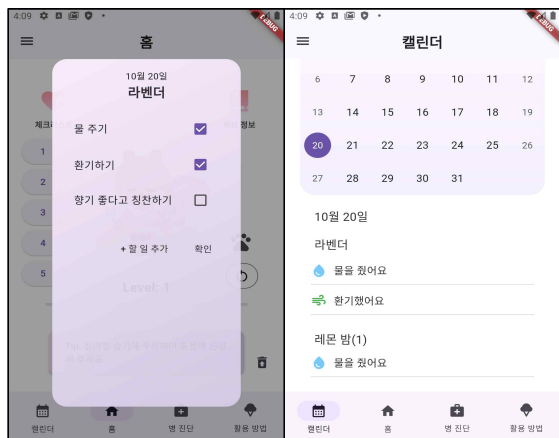


그림 2. 홈 화면의 관리 체크리스트 및 캘린더  
Fig. 2. Administration checklist and calendar on the main screen

사용자는 육성 정보 버튼을 통해 기르고 있는 허브의 관리 방법을 확인할 수 있다. 체크리스트에서는 ‘물 주기’, ‘환기하기’ 등의 관리한 항목을 체크할 수 있고, 과거의 완료된 항목은 캘린더 화면에서 볼 수 있다(그림 2). 또한, 구글 캘린더와 연동하여 관리 항목이 구글 캘린더에 자동으로 추가된다.

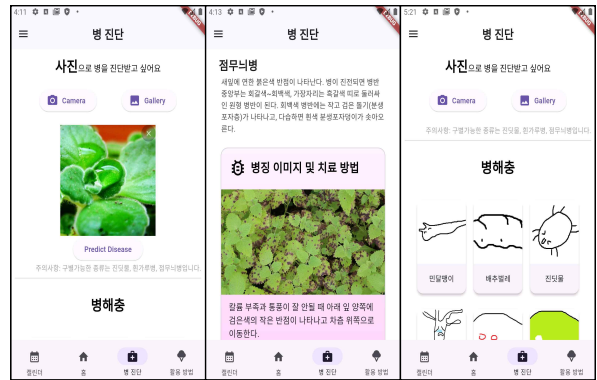


그림 3. 병 진단 화면 및 예시  
Fig. 3. Diagnostic screen and example

사진으로 진단하거나 카드 형식으로 제공되는 병해충 종류를 선택해서 상세 설명과 치료 방법을 확인할 수 있다(그림 3). 사용자가 사진을 올리면 AI가 이를 분석해 점무늬병, 진딧물, 흰가루병과 같은 병해충을 진단하고, 해당 병해충에 대한 적절한 치료 방법을 제시한다.

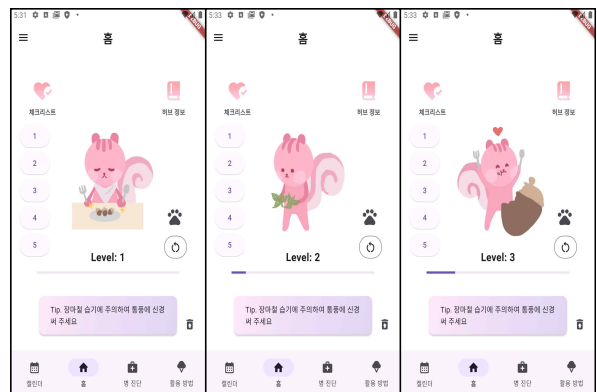


그림 4. 홈 화면의 캐릭터 성장  
Fig. 4. Growing characters on the home screen

캐릭터 제작에는 Rive 플랫폼을 사용하였다. 사용자의 허브 관리 빈도에 따라 캐릭터가 성장하며, 관리 활동을 완료할 때마다 경험치가 누적된다. 일정 수치에 도달하면 캐릭터가 진화하고, 총 3단계

진화를 거쳐 최종 단계에 도달하면 성장이 완료된다(그림 4). 사용자는 성장 완료된 캐릭터를 유지하거나 다른 캐릭터를 만나는 선택지를 고를 수 있다. 성장이 완료된 캐릭터는 도감에 등록되어 사용자가 관리한 허브와 캐릭터의 기록을 남긴다. 또한, 캐릭터는 터치 시 반응하도록 구현하여 사용자와의 상호작용을 강화하였다.

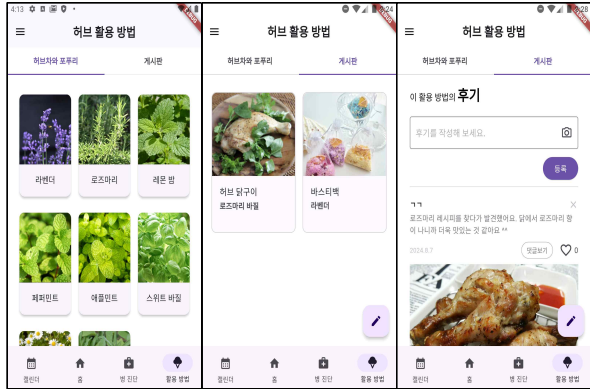


그림 5. 활용 정보 추천 화면

Fig. 5. Screen recommending how to utilize the herb

사용자가 허브를 활용하는 방법을 추천받고 공유할 수 있는 탭을 개발하였다. 앱 내에서 허브차와 포푸리 제작 방법을 제공하고 자신만의 허브 활용 방법을 게시판을 통해 공유할 수 있도록 하였다(그림 5). 게시판 활성화를 위해 다른 사용자의 활용 방법 게시물에 대한 후기 작성 기능, 해당 후기에 대해 질문 등을 남길 수 있는 댓글 기능을 제작하였다.

사용자는 게시글을 북마크 하여 ‘북마크’ 메뉴를 통해 언제든지 확인할 수 있으며, 게시글에 이미지를 첨부한 후기 댓글을 작성할 수 있다. 마음에 드는 후기에는 ‘좋아요’ 버튼을 눌러 호응할 수 있도록 하였다.

### III. 결 론

본 논문에서는 지속적인 허브 재배를 위해 보조 기능을 제공하는 애플리케이션을 개발하였다. 이 애플리케이션은 허브 재배를 돕고, 원예 활동에 지속적인 관심을 유도하며, 허브를 사용하는 다양한 활용법을 제시하는 다양한 기능을 포함한다.

애플리케이션을 직접 사용한 결과, 허브 재배 과정에서 흥미와 동기 부여가 지속적으로 유지되는 것을 확인할 수 있었다. 특히, 캐릭터 성장 기능과 허브 활용 추천 기능은 원예 활동에 대한 몰입도를 높이는 데 기여했으며, 이를 통해 사용자의 지속적인 참여를 유도하는 효과를 기대할 수 있었다.

또한, 사용자는 캐릭터를 통해 원예 활동, 원예치료에 참여함으로써 긍정적인 변화를 경험할 가능성이 있다. 다만, 병 진단 기능의 경우 학습 데이터 확보 과정에서 한계가 있어, 현재 진단 가능한 병해충의 종류가 세 가지뿐이라는 단점이 있다. 향후 데이터 확충을 통해 이 기능을 개선할 필요가 있다.

마지막으로, 애플리케이션이 제공하는 게시판 기능을 통해 사용자들이 서로의 경험을 공유하고 소통함으로써 사회적 고립감을 줄이고 사회적 교류를 촉진할 수 있다. 허브 활용 방법 제공 기능을 통해 사람들의 허브 사용이 증가한다면, 허브 씨앗, 영양제, 아로마 오일, 천연비누 등 허브와 관련된 다양한 상품의 수요로 이어지는 것을 기대할 수 있을 것이다.

### 참 고 문 헌

- [1] 곽혜란 and 장유진. (2011). 식물 윤리적 측면에서 본 원예치료 패러다임 발전방향에 대한 연구. 환경철학, 12, 65-90.
- [2] 곽혜란 이종석 Shannon E Jarrot Paula D Relf. "노인시설을 위한 효율적 프로그램으로서의 원예활동 효과검증 - 관찰법에 의한 평가 -." 인간 식물환경학회지 3.4 (2000): 7-15.
- [3] 박대환(Dea-Hwan Park),성대경(Dae-Kyung Sung),and 이부형(Boo-Hyung Lee). "현실과 가상 환경에서 동작하는 식물 재배 도우미." Proceedings of KIIT Conference 2017.6 (2017): 364-367.
- [4] 강점순,조동,최영환,안종길 ,and 손병구. "꽃과 허브식물을 이용한 원예치료적 효과." 農業技術開發研究所報 3.1 (1999): 187-207.

# 머신러닝을 통한 서울시 노인복지시설 수요 불균형에 영향을 미치는 주요 요인 분석

황호준\*, 강현규\*\*

## Machine Learning-Based Analysis of Demand-Supply Imbalance in Elderly Welfare Facilities in Seoul

Hojun Hwang\* and Hyunkyu Kang\*\*

### 요 약

본 연구는 초고령사회로 진입하고 있는 대한민국에서 노인복지시설의 수요-공급 불균형을 머신러닝 기법을 활용하여 분석하고자 하였다. 서울시 25개 자치구를 대상으로 노인인구, 의료보장 적용인구, 노령연금 수급자 수 등 다양한 사회경제적 변수를 수집하였다. 랜덤 포레스트 회귀 모델을 통해 주요 영향 요인을 도출한 결과, 의료보장 적용인구와 노령연금 수급자수가 시설 수요에 큰 영향을 미치는 것으로 나타났다. 이를 바탕으로 종합 필요도 점수를 산출하여 지역별로 시각화하였으며, 시설 부족 지역에 대한 우선적인 시설 확충 방안을 제시하였다. 본 연구는 데이터 기반의 접근을 통해 노인복지시설의 효율적인 자원 배분과 정책 수립에 기여하고자 한다.

### Abstract

This study aims to analyze the supply-demand imbalance of elderly welfare facilities in South Korea, which is entering a super-aged society, using machine learning techniques. We collected various socioeconomic variables such as the elderly population, medically insured individuals, and recipients of old-age pensions across 25 districts in Seoul. Using a Random Forest regression model, key factors affecting facility demand were identified. The results indicated that the number of medically insured individuals and recipients of old-age pensions significantly impact facility demand. Based on these findings, a comprehensive necessity score was calculated and visualized for each district, suggesting prioritized facility expansion in areas lacking adequate services. This study contributes to effective resource allocation and policy-making for elderly welfare facilities through data-driven approaches.

### Key words

Elderly Welfare Facilities, Demand-Supply Imbalance, Machine Learning, Random Forest, Optimal Location

### 1. 서 론

대한민국의 노인 인구는 전체 인구의 17.5%인

901만 8천 명에 달하며, 2026년에는 20.6%로 초고령사회에 진입할 것으로 예상된다[1]. 그러나 사회에서 노인은 부담을 주는 존재로 인식되어 자존감

\* 건국대학교 과학기술대학 컴퓨터공학과, exygnos@naver.com

\*\* 건국대학교 과학기술대학 컴퓨터공학과 교수(교신저자), hkkang@kku.ac.kr

저하와 사회적 고립을 초래하고 있다[2]. 특히, 의료 복지 형태의 요양 시설은 님비(NIMBY) 현상으로 인해 지역 내 수요-공급 불균형이 심화되고 있다[3].

최근 활동적 노후(active aging)와 지역사회에서의 노후(aging in community)에 대한 중요성이 부각되고 있으며[4], 노인들이 생활해 온 지역에서 지속 가능한 노후를 보내기 위해서는 지역사회 내 노인복지 시설의 설치와 운영이 필수적이다. 그러나 노인복지 시설의 수요와 입지에 대한 연구는 미흡한 실정이다. 이에 본 연구는 머신러닝 기법을 활용하여 서울시 25개 자치구의 노인복지시설 수요-공급 불균형 요인을 분석하고자 한다.

## II. 방법론

### 2.1 데이터 수집 및 전처리

본 연구는 서울시 25개 자치구를 대상으로 노인 복지시설의 수요-공급 불균형을 분석하기 위해 공공 데이터 포털에서 제공하는 다양한 데이터를 활용하였다. 수집된 데이터는 노인인구[5], 노령연금 수급자수[6], 의료보장 적용인구[7], 공공의료시설 접근성[8], 세수입[9], 사망자수[10] 등 사회경제적 지표를 포함한다.

데이터의 통합과 전처리는 Python의 pandas와 numpy 라이브러리를 사용하여 수행하였다. 결측치(missing values)와 이상치(outlier)를 처리하고, 변수 간 단위 차이를 제거하기 위해 scikit-learn의 MinMaxScaler를 활용하여 스케일링을 적용하였다.

표 1. 주요 변수 및 활용 데이터

Table 1. Key Variables and Data Utilized

변수	내용
일반적 특성	노인인구, 노령연금수급자수 등
경제	세수입
의료 접근성	공공의료시설 및 종합병원 접근성
기타	사망자수, 의료보장 적용 인구 등

### 2.2 변수 선정 및 수요-공급 비율 계산

독립 변수로는 노인인구, 노령연금 수급자수, 의료보장 적용인구, 공공의료시설 접근성 등을 선정하

였으며, 종속 변수로는 노인복지시설의 수요-공급 비율을 사용하였다. 수요-공급 비율은 다음과 같이 계산하였다:

$$\text{수요-공급 비율} = \frac{\text{노인복지시설수}}{\text{노인인구수}} \quad (1)$$

이를 통해 각 지역구별로 노인복지시설의 상대적 부족 정도를 평가하였다.

### 2.3 머신러닝 모델 구축

노인복지시설 수요-공급 불균형에 영향을 미치는 주요 요인을 파악하기 위해 랜덤 포레스트 회귀 모델을 구축하였다. 해당 모델은 결정 트리의 앙상블을 기반으로 하며, 각 트리는 데이터 셋의 서로 다른 샘플과 피처(feature)를 사용하여 독립적으로 학습함으로써 전체 모델의 예측 정확도를 높이고 과적합을 방지한다. 데이터는 훈련 세트(80%)와 테스트 세트(20%)로 분할하였으며, scikit-learn 라이브러리를 활용하였다. 모델의 성능은 루트 평균 제곱 오차(RMSE)를 사용하여 평가하였다.

### 2.4 변수 중요도 분석 및 시각화

랜덤 포레스트 모델의 feature\_importances\_ 속성을 활용하여 각 변수의 중요도를 추출하였다. 이를 바탕으로 변수별로 종합 필요도 점수를 계산하였으며, 점수는 다음과 같이 정의된다:

$$\text{종합 필요도 점수} = \sum_i (\text{스케일링된 변수 } i \times \text{변수 중요도 } i) \quad (2)$$

예를 들어, 두 변수 '노인인구'와 '의료 접근성'의 중요도가 각각 0.6과 0.4이고, 스케일링된 값이 각각 0.7과 0.9라면, 종합 필요도 점수는 다음과 같이 계산된다:

$$\text{종합 필요도 점수} = (0.6 \times 0.7) + (0.4 \times 0.9) = 0.78 \quad (3)$$

계산된 점수는 GeoPandas와 Folium 라이브러리를

사용하여 서울시 지도 위에 시각화하였다. 이를 통해 노인복지시설의 추가적인 필요성이 큰 지역을 한눈에 파악할 수 있도록 하였다.

### III. 분석 결과

#### 3.1 모델 성능 평가

랜덤 포레스트 회귀 모델의 성능은 테스트 데이터 세트를 대상으로 평가되었으며, 루트 평균 제곱 오차(RMSE)는 약 0.00057로 나타났다. 이는 모델이 노인복지시설의 수요-공급 비율을 비교적 정확하게 예측할 수 있음을 의미한다. 실제 값과 예측 값 간의 상관관계도 높게 나타나 모델의 적합도가 우수함을 확인하였다.

잔차(residuals) 분포 히스토그램을 통해 모델의 예측 오차를 분석한 결과, 대부분의 잔차가 0 근처에 집중되어 있어 예측값이 실제값과 잘 일치하고 있음을 알 수 있었다.

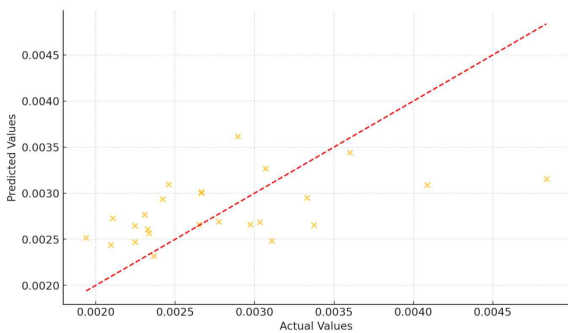


그림 1. 실제값과 예측값 그래프  
Fig. 1. Actual vs. Predicted Values Graph

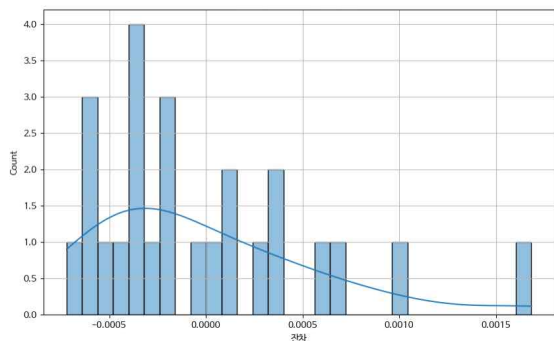


그림 2. 모델 성능 잔차 분포 히스토그램  
Fig. 2. Model Performance Residual Distribution Histogram

잔차는 모델 예측의 정확성을 평가하는 중요한 지표로, 예측값이 실제값에 얼마나 가까운지를 나타낸다. 그러나 일부 데이터에서는 오차가 크게 나타나 모델의 개선 여지가 있음을 시사한다.

#### 3.2 피쳐(feature) 중요도 분석

랜덤 포레스트 모델의 피쳐 중요도를 분석한 결과, 의료보장 적용인구, 노령연금 수급자수, 공공의료시설 접근성 등이 노인복지시설의 수요-공급 불균형에 가장 큰 영향을 미치는 요인으로 나타났다. 이는 건강 관리 서비스에 대한 접근성과 경제적 지원이 노인복지시설 수요에 중요한 영향을 미침을 의미한다.

구체적으로, 의료보장 적용인구는 노인들의 의료 서비스 이용 가능성을 나타내며, 이는 노인복지시설 이용과 직결된다. 노령연금 수급자수는 노인들의 경제적 여건을 반영하며, 경제적 지원이 시설 이용 결정에 중요한 요인으로 작용한다. 공공의료시설 접근성은 지역 내 의료 인프라의 수준을 나타내며, 접근성이 높을수록 노인복지시설의 수요도 증가하는 경향을 보였다.

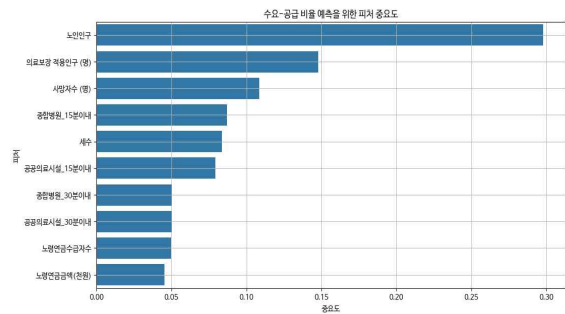


그림 3. 피쳐 중요도 바 그래프  
Fig. 3. Feature Importance Bar Graph

#### 3.3 종합 필요도 점수 시각화

각 변수의 중요도를 고려하여 종합 필요도 점수를 계산하고, 이를 서울시 지도 위에 시각화하였다. 그 결과, 대부분의 자치구에서 노인복지시설이 부족한 것으로 나타났으며, 특히 일부 지역에서는 시설 확충이 시급한 것으로 분석되었다.

시각화 결과에 따르면, 시설 부족 지역은 종합 필요도 점수가 높게 나타나 어두운 색으로 표시되었으며, 이는 노인복지시설의 추가 설치가 필요함을 의미한다. 반면, 종합 필요도 점수가 낮은 지역은 상대적으로 시설이 충분히 제공되고 있어 밝은 색으로 표시되었다.

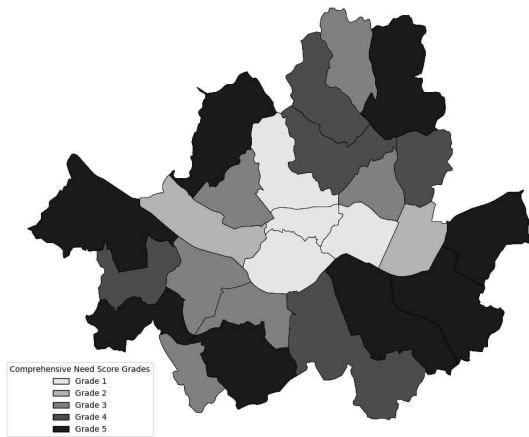


그림 4. 노인복지시설의 수요 시각화  
Fig. 4. Visualization of Demand for Elderly Welfare Facilities

### 3.4 결과에 대한 논의

본 연구의 결과는 노인복지시설의 수요-공급 불균형이 지역별로 상이하며, 이는 사회경제적 요인과 밀접한 관련이 있음을 보여준다. 특히, 의료보장 적용인구와 노령연금 수급자수가 높은 지역에서 시설 수요가 높게 나타났는데, 이는 해당 지역의 노인들이 경제적·의료적 지원을 통해 복지시설 이용에 대한 필요성과 접근성이 높아졌음을 의미한다.

또한, 공공의료시설 접근성이 높은 지역에서 노인복지시설의 수요가 증가하는 경향을 보였다. 이는 의료 인프라와 복지 인프라 간의 상호 보완적 관계를 시사하며, 노인들이 복지시설을 이용함에 있어 의료 서비스와의 연계성을 중요시한다는 것을 알 수 있다.

모델의 성능 평가 결과와 피쳐 중요도 분석을 종합하면, 머신러닝 기법을 활용한 본 연구의 접근 방식이 노인복지시설의 수요 예측과 최적 입지 선정에 효과적임을 확인할 수 있다. 이는 향후 정책 결정과 자원 배분에 있어 데이터 기반의 의사 결정이

필요함을 강조한다.

## IV. 결 론

본 연구는 머신러닝 기법을 활용하여 서울시 노인복지시설의 수요-공급 불균형을 분석하고 주요 영향 요인을 확인하였다. 그 결과, 노인인구와 의료보장 적용인구 등이 시설 수요에 큰 영향을 미치며, 데이터 기반의 접근이 자원 배분과 정책 수립에 효과적임을 확인하였다. 이에 따라 시설 부족 지역을 우선적으로 선정하여 노인복지시설을 설립하고, 운영의 실효성을 높이기 위한 평가가 필요하다.

향후 연구에서는 사회적·환경적 요인을 포함한 더 다양한 변수를 고려하고, 시계열 분석을 통해 장기적인 수요 예측을 진행할 수 있다. 또한, 본 연구의 방법론을 다른 지역이나 국가에 적용하여 노인복지시설의 최적 입지 선정에 활용할 수 있을 것이다. 이를 통해 초고령 사회에 대비한 실효성 있는 대응 방안을 마련하는 데 기여할 것으로 기대된다.

## 참 고 문 헌

- [1] 통계청. "2022 고령자 통계". Retrieved from [https://kostat.go.kr/board.es?mid=a10301010000&bid=10820&tag=&act=view&list\\_no=420896&ref\\_bid=](https://kostat.go.kr/board.es?mid=a10301010000&bid=10820&tag=&act=view&list_no=420896&ref_bid=), 2022.
- [2] 조수현, 정순돌. "사회위험수준인식과 노인낙인의 관계: 사회적 관계망의 조절효과." 보건사회연구, 39(2), 131-159, 2019.
- [3] 안소현. "노인복지시설의 사회적 기피현상 완화 계획안." 국내석사학위논문, 홍익대학교 대학원, 서울, 2020.
- [4] 김지윤, 이진학. "공간 최적화 방법을 이용한 서울시 소규모 노인복지센터의 최적 입지 연구: 공간 효율성과 형평성을 고려한 다목적 최적화 접근." 한국지도학회지, 23(3), 33-53, 2023.
- [5] 통계청, "인구총조사", 성, 세대구성 및 연령별 고령자 - 시군구, [https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT\\_1IN2010&conn\\_path=12](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN2010&conn_path=12), 2023.
- [6] 국민연금공단, "국민연금통계", 시군구별 노령연



- 금 세부 지급현황, [https://kosis.kr/statHtml/statHtml.do?orgId=322&tblId=DT\\_32202\\_B045&conn\\_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=322&tblId=DT_32202_B045&conn_path=I2), 2023.
- [7] 국민건강보험공단, "노인장기요양보험통계", 시·군·구별 자격별 의료보장 적용인구 현황, [https://kosis.kr/statHtml/statHtml.do?orgId=350&tblId=DT\\_35006\\_N003&conn\\_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=350&tblId=DT_35006_N003&conn_path=I2), 2023.
- [8] 한국교통연구원, "교통접근성지표", 의료시설 접근 가능 시설 수, [https://kosis.kr/statHtml/statHtml.do?orgId=444&tblId=DT\\_444001\\_006&conn\\_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=444&tblId=DT_444001_006&conn_path=I2), 2016.
- [9] 국세청, "국세통계", 2.1.3 지역별·세목별 세수 현황[2005~], [https://kosis.kr/statHtml/statHtml.do?orgId=133&tblId=TX\\_13301\\_A012&conn\\_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=133&tblId=TX_13301_A012&conn_path=I2), 2022.
- [10] 통계청, "인구동향조사", 시군구/성/연령(5세)별 사망자수(1997~), 사망률(1998~), [https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT\\_1B80A18&conn\\_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B80A18&conn_path=I2), 2023.

# IMU 센서와 다층 신경망을 이용한 실시간 낙상 감지 및 분류

유병일\*, 이승환\*, 문지환\*, 이은성\*, 이진윤\*, 이인수\*\*

## Real-Time Fall Detection and Classification Using IMU Sensors and Multilayer Neural Networks

Byeong-il Yu\*, Seung Hwan Lee\*, Ji Hwan Moon\*, Eun Seoung Lee\*, Jeon Yun Lee\*,  
and In-Soo Lee\*\*

### 요약

본 연구는 IMU 센서를 활용해 낙상 여부와 유형을 분류하는 딥러닝 모델을 제안한다. 다층 신경망을 기반으로 미낙상, 전방 낙상, 후방 낙상, 측면 낙상을 실시간으로 분류하는 시스템을 개발하였다. 테스트 결과, 모델은 대부분의 낙상 유형을 정확히 예측했으며, 특히 낙상의 시작과 종료를 실시간으로 감지하여 실시간 낙상 감지 시스템으로서의 가능성을 확인하였다. 제안된 모델은 실시간 낙상 감지 시스템 개발에 중요한 기초 자료로 활용될 수 있을 것이다

### Abstract

This study proposes a deep learning model that classifies fall occurrences and fall types using IMU sensors. A system was developed using a Multilayer Neural Network to classify non-falls, forward falls, backward falls, and sideways falls in real-time. The test results show that the model accurately predicts most fall types and successfully detects the start and end of falls in real-time, confirming its potential as a real-time fall detection system. The proposed model can serve as a crucial foundation for developing real-time fall prevention and detection systems.

### Key words

Fall Detection, Fall Classification, Multilayer Neural Network, Deep Learning, Time-series Data Classification, IMU

## 1. 서론

낙상은 고령자, 장애인, 기타 고위험군에서 흔히 발생하는 사고로, 신속하게 대응하지 않으면 심각한 부상이나 사망으로 이어질 수 있다. 특히 혼자 있는 상황에서 낙상을 적시에 감지하지 못하면 응급 처치가 지연되어 위험이 더욱 커진다. 이러한 이유로

실시간으로 낙상을 감지할 수 있는 시스템의 필요성이 점점 커지고 있다. 본 연구는 IMU 센서 데이터를 기반으로 딥러닝 모델을 활용해 낙상 여부와 유형을 정확히 분류하는 시스템을 개발하고, 이를 통해 실시간 낙상 감지 시스템의 구현 가능성을 모색하고자 한다.

\* 경북대학교 IT대학 전자공학부 / ybl3897@knu.ac.kr

\*\* 경북대학교 IT대학 전자공학부 교수 / insoolee@knu.ac.kr

## II. MNN을 이용한 낙상 감지 및 분류 모델

본 연구에서 사용한 딥러닝 모델은 Multilayer Neural Network[1]로, 그 구조는 그림 1과 같다.

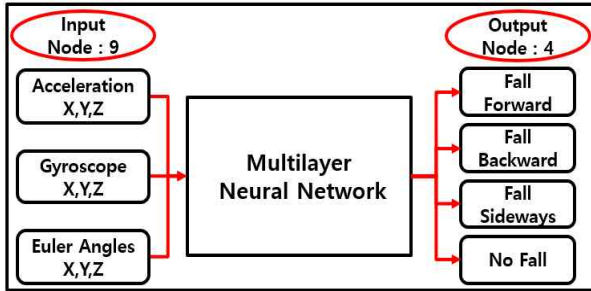


그림 1. 다층 신경망 기반 감지 및 분류 모델  
Fig. 1. Detection and Classification Model Based on Multilayer Neural Networks

입력 노드는 가속도(Acceleration), 자이로스코프(Gyroscope), 오일러 각(Euler Angles)에 대한 각 X, Y, Z 세 축의 데이터를 받아 총 9개로 설정한다. 출력 노드는 Fall Forward, Fall Backward, Fall Sideways, No Fall로 설정하여 낙상 유형을 예측한다. 2.1 - 2.4절에서는 데이터셋 구조, 데이터셋 전처리, 모델 학습, 그리고 테스트 결과 순으로 모델 구현 과정을 자세히 설명한다.

### 2.1 데이터셋 구조

본 연구에서 사용한 데이터셋은 KFall Dataset[2]에서 제공받은 IMU 센서[3]로 측정된 낙상 데이터셋이다. 그림 2는 KFall Dataset 구조를 보여준다.

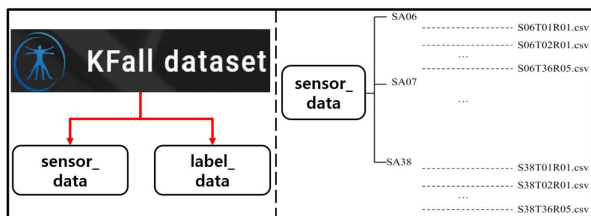


그림 2. KFall Dataset 구조  
Fig. 2. KFall Dataset Structure

데이터셋은 sensor\_data와 낙상 발생 구간을 분류한 label\_data로 구성된다. sensor\_data는 SA06부터

SA38까지(SA34는 제외) 총 32명의 사용자에게서 수집되었다. 각 사용자는 특정 Task ID에 따라 센서 데이터를 기록했으며, 이 Task ID는 미낙상과 다양한 낙상 유형(전방, 후방, 측면 낙상 등)을 나타낸다. sensor\_data의 각 행은 특정 시간대의 가속도, 자이로, 오일러각의 정보를 포함하며, label\_data는 각 Task ID에 대응하는 낙상 구간 정보를 포함한다.

### 2.2 낙상 유형 분류 및 라벨링 과정

본 연구의 딥러닝 모델 출력은 4개의 낙상 유형으로 구성된다. 따라서 각 Task ID는 해당 낙상 유형으로 분류된다. 그림 3은 Task ID에 따른 낙상 유형 분류를 보여준다.

<i>Fall Mapping:</i>	( 20, 23, 25, 28, 30, 31, 32, 33 : <i>Fall Forward</i> )
	( 22, 24, 26 : <i>Fall Sideways</i> )
	( 21, 27, 29, 34 : <i>Fall Backward</i> )
	( 1-19, 35, 36 : <i>No Fall</i> )

그림 3. Task ID에 따른 낙상 유형 분류  
Fig. 3. Classification of Fall Types by Task ID

Task ID 20, 23, 25, 28, 30-33은 전방 낙상, Task ID 22, 24, 26은 측면 낙상으로 분류된다. Task ID 21, 27, 29, 34는 후면 낙상으로, 나머지 Task ID는 미 낙상으로 분류된다.

라벨링 과정은 낙상 구간을 정확히 구분해 모델이 낙상 발생 여부와 유형을 학습할 수 있도록 하기 위해 필수적이다. 그림 4는 SA06의 Task ID = 30 데이터를 라벨링하는 과정을 보여준다.

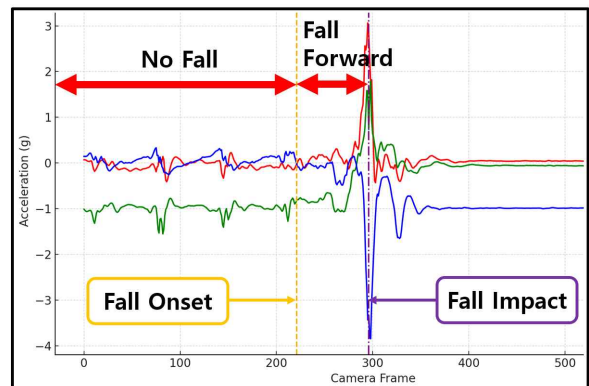


그림 4. SA06의 Task ID = 30 라벨링  
Fig. 4. Labeling of SA06, Task ID = 30

sensor\_data는 Fall\_Onset\_frame과 Fall\_Impact\_frame을 기준으로 낙상 구간과 미낙상 구간으로 나뉜다. Fall\_Onset\_frame은 낙상이 시작되는 시점을, Fall\_Impact\_frame은 낙상이 종료되는 시점을 의미한다. Fall\_Onset\_frame 이전의 데이터는 'No Fall'로, 두 프레임 사이 데이터는 해당 Task ID의 낙상 유형으로 라벨링된다. 이후 전처리된 데이터는 하나의 데이터프레임으로 병합되어 센서 데이터와 라벨이 결합된다. 모델 입력은 가속도계, 자이로스코프, 오일러각 데이터를 포함한 시간대별 센서 데이터로 구성되며, 모든 센서 값은 평균 0, 분산 1로 표준화된다.

### 2.3 모델 학습

본 연구에서는 4가지 낙상 유형을 분류하기 위해 Multilayer Neural Network를 사용하였으며, 모델 학습에는 sparse categorical crossentropy 손실 함수를 적용했다. 이 손실 함수는 실제 라벨과 예측 간의 차이를 계산해 모델의 정확도를 평가한다. 최적화에는 Adam 옵티마이저를 사용하여 학습률을 효율적으로 조정하고, 가중치 업데이트를 통해 학습 속도를 높였다.

모델의 과적합을 방지하고 성능을 평가하기 위해 데이터셋을 학습 데이터와 테스트 데이터로 분할했다. 그림 5는 데이터셋 분할 구조 및 학습을 보여준다.

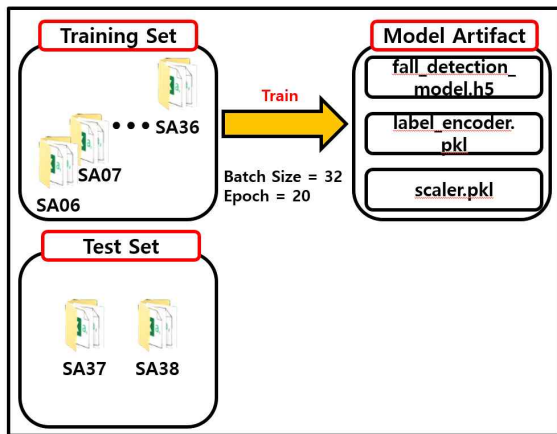
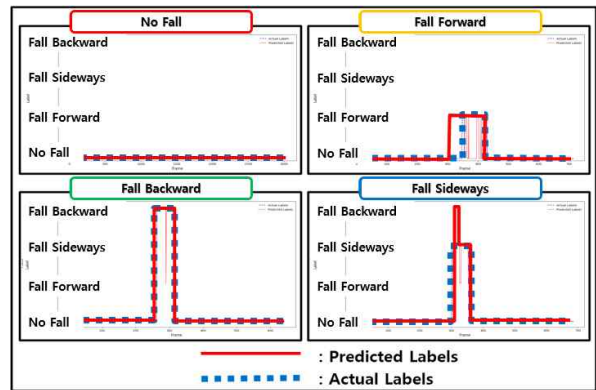


그림 5. 데이터셋 분할 구조 및 학습  
Fig. 5. Dataset Split Structure and Training

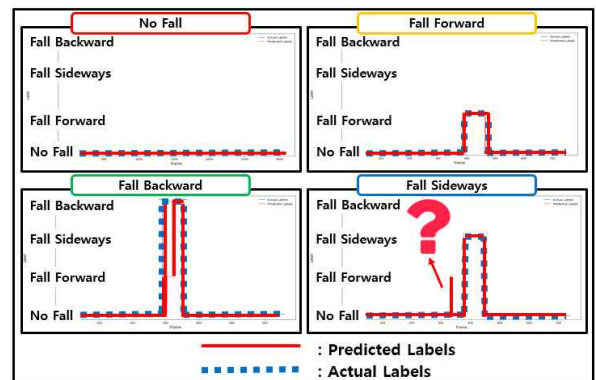
모델 학습에는 SA06부터 SA36까지 (SA34는 제외) 30명의 데이터를 사용했고, 평가 데이터로는 SA37과 SA38의 데이터를 사용해 성능을 검증했다. 학습은 Batch size는 32로, 20 Epoch 동안 진행했다. 학습 후 가중치는 저장되어 실시간 낙상 감지 시스템에 활용될 예정이다. 입력 데이터와 라벨은 LabelEncoder와 StandardScaler로 인코딩 및 표준화되었으며, 재사용을 위해 함께 저장했다.

### 2.4 테스트 결과

학습된 Multilayer Neural Network 모델을 SA37, SA38 테스트 데이터에 적용해 성능을 평가했다. 아래 그림 6의 (a)와 (b)는 각각 SA37과 SA38 테스트 데이터의 예측과 실제 라벨과 비교를 보여준다.



(a) SA37 테스트 데이터의 예측과 실제 라벨 비교  
(a). Comparison of Predictions and Actual Labels for SA37 Test Data



(b) SA38 테스트 데이터의 예측과 실제 라벨 비교  
(b). Comparison of Predictions and Actual Labels for SA38 Test Data

그림 6. SA37, SA38 테스트 데이터의 예측과 실제 라벨 비교

Fig. 6. Comparison of Predictions and Actual Labels for SA37 and SA38 Test Data

테스트 데이터는 StandardScaler로 표준화하고, LabelEncoder로 라벨을 수치형으로 변환한 후, 학습된 Multilayer Neural Network로 예측을 수행했다. 그림 6에서 모델은 미낙상, 전면 낙상, 후면 낙상을 정확하게 예측하고, Fall Onset Frame부터 Fall Impact Frame까지의 낙상 구간과 유형을 거의 동시에 분류해 실시간 예측 성능이 뛰어남을 보여주었다. 다만, 측면 낙상에서 약 350 프레임에서 전면 낙상 노이즈가 발생했고, 이를 분석하기 위해 데이터셋 수집 영상을 확인했다. 그림 7은 측면 낙상 중 발생한 전면 낙상 노이즈의 원인을 보여준다.

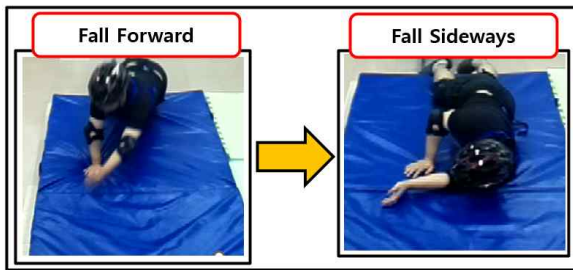


그림 7. 측면 낙상 중 발생한 전면 낙상 노이즈 원인  
Fig. 7. Cause of Fall Forward noise in Fall Sideways event

영상에서 실험자가 약 350 프레임에서 전면으로 먼저 넘어진 후 측면으로 넘어짐을 확인할 수 있었다. 따라서 모델의 예측 결과가 실제 상황을 잘 반영하고 있음을 알 수 있다.

### III. 결 론

본 연구에서는 IMU 센서 데이터를 기반으로 낙상 유형을 분류하는 딥러닝 모델을 개발하고 성능을 평가했다. 다층 신경망을 사용해 미낙상, 전방 낙상, 후방 낙상, 측면 낙상을 분류하는 모델을 제안했다. 테스트 결과, 모델은 대부분의 낙상 유형을 정확하게 예측하였으나, 측면 낙상 데이터에서 노이즈가 발생했다. 실험 영상 분석을 통해 노이즈 발생 원인을 파악함으로써, 모델이 실제 상황을 잘 반영하고 있음을 확인할 수 있었다.

특히, 모델은 낙상의 시작과 종료를 거의 실시간으로 감지하여, 실시간 낙상 감지 시스템으로의 적용 가능성을 보여주었다. 본 연구에서 제안한 모델

은 향후 실시간 낙상 감지 시스템 개발에 중요한 기초 자료로 활용될 수 있을 것이다.

### 참 고 문 헌

- [1] Z. Zhang, W. Ma, C. Zhang, and Y. Yu, "Improved Multilayer Neural Network for Activity Recognition Using Wearable Sensors," IEEE Access, vol. 9, pp. 140973-140986, 2021.
- [2] Yu X, Jang J, and Xiong S\*, 2021. A Large-scale Open Motion Dataset (KFall) and Benchmark Algorithms for Detecting Pre-impact Fall of the Elderly Using Wearable Inertial Sensors. Frontiers in Aging Neuroscience.
- [3] S. Abdelmoula, A. Chihi, A. B. Nassif, and J. Verrelst, "Inertial sensors technologies for navigation applications: state of the art and future trends," Satellite Navigation, vol. 3, no. 1, pp. 1-14, 2022.

# 역삼투막 공정에서 IoT 운영을 위한 모델링

김동현\*, 구경모\*, 임정훈\*, 이창민\*, 은주연\*, 정설영\*, 임환규\*\*

## Modeling IoT Operations on Reverse Osmosis Membrane Processes

Dong-hyun Kim\*, Kyung-mo Gu\*, Jeong-hoon Lim\*, Chang-min Lee\*, Ju-yeon Eun\*,  
Seol-Young Jeong\*, Hwan Kyu Lim\*\*

### 요약

현대 산업에서 물 정화와 해수 담수화는 필수적인 공정으로 자리 잡았다. 특히 RO(Reverse Osmosis) 시스템은 효율적인 수처리를 위해 널리 사용되고 있으며, 이 공정에서 필터의 상태를 모니터링하고 적절한 시기에 교체하는 것이 중요하다. 본 연구는 IoT(사물인터넷) 기술을 활용하여 RO 공정에서 필터 교체 시점을 자동으로 예측하는 모델을 개발하는 것을 목표로 한다. 기존의 수작업 기반의 관리 방식을 개선하고, 실시간 데이터 분석을 통해 공정 효율성을 극대화하며 운영 비용을 절감하는 방안을 제시한다.

### Abstract

In modern industries, water purification and seawater desalination have become essential processes. Particularly, Reverse Osmosis (RO) systems are widely used for efficient water treatment, and monitoring the condition of filters and replacing them at the right time is crucial. This study aims to develop a model that automatically predicts filter replacement timing in RO processes using IoT (Internet of Things) technology. It proposes an improvement to the traditional manual management methods by maximizing process efficiency and reducing operational costs through real-time data analysis.

### Key words

Reverse Osmosis Process, Feed Pressure, LSTM, LSTM-Transformer, Temporal Fusion Transformer

## 1. 서론

RO 시스템(역삼투 시스템)은 물 정화 기술 중 하나이다. 이 시스템은 고압을 이용해 물을 반투막

\*경북대학교 IT대학 컴퓨터학부, bomul0524@gmail.com

\*경북대학교 IT대학 컴퓨터학부, rnruda1078@gmail.com

\*경북대학교 IT대학 컴퓨터학부, jhl10815@naver.com

\*경북대학교 IT대학 컴퓨터학부, windo3969@knu.ac.kr

\*경북대학교 IT대학 컴퓨터학부, dmswndus2@knu.ac.kr

\*경북대학교 IT대학 컴퓨터학부 교수(교신저자), snowflower@knu.ac.kr

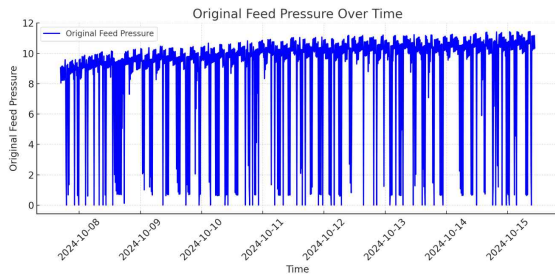
\*\* (주)아쿠아웍스 기술연구소장(교신저자), okhawnkvu@gmail.com

(semipermeable membrane)을 통해 통과시키며, 이 과정을 통해 물속의 불순물, 미생물, 이온 등을 제거한다. 결과적으로 매우 깨끗한 물을 얻을 수 있다. 기존 시스템의 경우 RO필터의 약품 세척 시기와 교체 시기를 사용자의 경험으로만 예측하여 필터를 더 사용할 수 있음에도 세척, 교체로 인한 비용 손실이 발생했다. 본 연구에서는 이러한 불필요한 비용 손실을 막을 수 있도록 정확한 시기를 예측하는 인공지능 모델링 개발에 관한 연구를 진행하였다. 모델로는 LSTM, LSTM-Transformer, TFT 3가지의 모델을 가지고 비교 분석을 통한 모델링 개발을 진행하였다.

## II. 데이터 수집 및 전처리

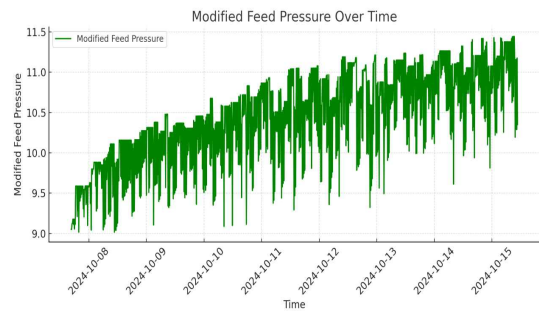
수질 데이터는 (주)아쿠아웍스에서 수집한 수질 데이터를 바탕으로 분석을 진행하였다. 수집한 데이터는 RO처리수 전기전도도, RO농축수 전기전도도, RO처리수 유량, RO 순환수 유량, RO 유입 압력 등이 1분 단위로 기록되어 있으며, 본 연구에서는 그 중 RO처리수 전기전도도와 RO유입 압력 값을 추출하여 사용하였다.

수질 관리 시스템의 플러싱(flushing)과정에서 발생하는 변동으로 인해 일시적인 급락 현상이 나타날 수 있다. [그림1]에서 볼 수 있듯이, 특정 구간에서 갑작스럽게 낮아지는 패턴이 반복적으로 관찰된다. 이러한 변동은 데이터 분석 및 모델링 단계에서 이상치로 작용하여 과적합(overfitting)을 유발할 수 있으며, 결과의 예측 성능 및 신뢰성을 저하시킬 수 있다. 따라서, 본 연구에서는 플러싱으로 인한 데이터의 왜곡을 방지하고자 일련의 전처리 과정을 진행하였다. [그림1]은 원본 데이터를 나타낸다.



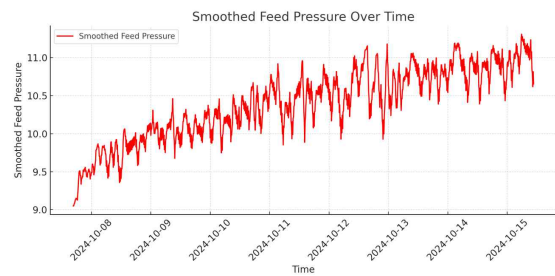
[그림 1] flushing 값이 포함된 원본 데이터

- 1) 이동평균 기반 낮은 값 대체  
데이터의 갑작스러운 하락을 완화하기 위해, 크기가 10인 슬라이딩 윈도우를 적용하여 이동 평균을 계산하였다. 이후 값이 평균보다 낮은 경우, 직전 3개의 데이터 중 최댓값으로 대체하였다.
- 2) Z-Score 기반 이상치 탐지 및 대체  
임계값을 2로 설정하여 평균에서 2 표준 편차를 벗어난 값을 이상치로 간주하였다. 이러한 이상치들은 직전 5개 데이터 중 최댓값을 대체하였다. [그림 2]는 1), 2) 전처리 결과 데이터를 나타낸다.



[그림 2] flushing 값을 이전 데이터의 대체

- 3) Data Smoothing 적용  
크기가 5인 이동 평균 스무딩 기법을 적용하여 데이터의 단기적 변동성을 감소시켰다. [그림 3]는 최종 전처리된 결과 데이터를 나타낸다.



[그림 3] 최종 전처리된 결과 데이터

## III. 모델 설계 및 성능 지표

전처리된 데이터를 증강하여 4개의 반복 사이클로 구성된 시계열 데이터를 사용하였다. 이를 고려하여 데이터의 일반화 성능을 높이기 위해 처음 3개의 사이클을 훈련 데이터로, 마지막 1개의 사이클을 테스트 데이터로 사용했다.

모델 설계에서, 각 모델은 이전 모델의 한계를 보완하며 시계열 데이터의 다양한 패턴을 학습할 수 있도록 발전시켰다. 본 연구에서는 LSTM, Sequential LSTM-Transformer, LSTM-Transformer, TFT(Temporal Fusion Transfomer) 모델을 순차적으로 적용시켰다.

Sequential LSTM-Transformer는 LSTM 구조에 Transformer 모듈을 추가하여 장기적인 패턴을 보완하였으며, 긴 시퀀스의 종속성 학습을 통해 성능을 개선하도록 설계되었다. Parallel LSTM-Transformer는 Sequential 모델과 달리 병렬 구조를 채택하여 학습 속도와 계산 효율성을 높였다.

TFT (Temporal Fusion Transformer)는 이전 모델의 장점을 종합적으로 결합하여 시계열 데이터의 전체적 패턴과 다양한 시점을 동시에 학습할 수 있도록 최종 모델로 사용되었다.

추가로, 각 모델에서 500의 시퀀스 길이를 설정하여 한 사이클 내에서의 변동 패턴을 충분히 반영할 수 있도록 구성하였다. 이러한 모델 구조 및 학습 설정은 시계열 데이터의 반복적인 패턴을 고려하여 과적합을 방지하고 일반화 성능을 높이는 데 중점을 두었다.

예측 모델의 성능 검증을 위해 현장에서 사용되는 수질 데이터의 20%를 예측 데이터로, 나머지 80%를 평가 데이터로 구분하였다. 모델의 정확도를 평가하기 위해 평균 제곱 오차(RMSE), Mean Absolute Percentage Error(MAPE), Coefficient of Determination ( $R^2$ )를 사용하였다.

#### 4.1 LSTM

본 연구의 첫 번째 모델로 기본 LSTM 모델을 사용하였다. 이 모델은 시계열 데이터의 순차적 종속성을 학습하는 데 중점을 두며, 양방향 LSTM 층을 통해 입력 시퀀스의 특징을 효과적으로 추출한다[1]-[3].

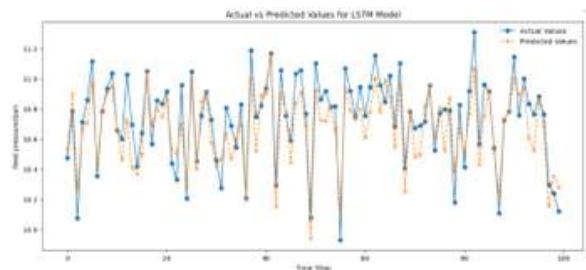
또한, Global Average Pooling 및 Global Max Pooling을 통해 LSTM 층에서 추출된 정보를 요약하여 최종 예측에 활용하였다. 학습 과정에서는 MSE 손실 함수를 사용하고, 최적화에는 Adam 옵티마이저를 적용하였다. [그림 4]는 LSTM의 구조이다.

LSTM 모델의 성능은 RMSE, MAPE,  $R^2$  지표로 평가하였으며, 이는 이후 모델들과의 비교 기준으로 활용된다. [그림 5]는 LSTM의 예측 결과 그래프이다.

Layer (type)	Output Shape	Param #	Connected to
input_layer_11 (InputLayer)	(None, 500, 1)	0	전체 화면을 종료하십시오
bidirectional_20 (Bidirectional)	(None, 500, 256)	133,120	input_layer_11[0][0]
layer_normalization_34 (LayerNormalization)	(None, 500, 256)	512	bidirectional_20[0][0]
bidirectional_21 (Bidirectional)	(None, 500, 128)	164,352	layer_normalization_3...
layer_normalization_35 (LayerNormalization)	(None, 500, 128)	256	bidirectional_21[0][0]
global_average_pooling1d_9 (GlobalAveragePooling1D)	(None, 128)	0	layer_normalization_3...
global_max_pooling1d_9 (GlobalMaxPooling1D)	(None, 128)	0	layer_normalization_3...
concatenate_13 (Concatenate)	(None, 256)	0	global_average_poolin... global_max_pooling1d_...
dense_50 (Dense)	(None, 64)	16,448	concatenate_13[0][0]
dense_51 (Dense)	(None, 1)	65	dense_50[0][0]

Total params: 314,753 (1.20 MB)  
Trainable params: 314,753 (1.20 MB)  
Non-trainable params: 0 (0.00 B)

[그림 4] LSTM 구조



[그림 5] LSTM 예측 결과 그래프

#### 4.2 Sequential LSTM-Transformer

Sequential LSTM-Transformer 모델은 LSTM 층과 Transformer 모듈을 결합한 하이브리드 구조로, 단기 및 장기 패턴을 동시에 학습할 수 있도록 설계되었다.

양방향 LSTM(Bidirectional LSTM) 층은 순차적 종속성을 학습하여 시계열 데이터의 특징을 효과적으로 추출하며, Multi-Head Attention 모듈은 장기적인 의존성을 반영하여 중요한 피처에 집중할 수 있도록 돕는다[4].

이 모델에서는 시퀀스 길이를 500으로 설정하여 주기 내 변동 패턴을 충분히 학습할 수 있도록 하였고, Layer Normalization을 적용하여 학습의 안정성을 강화하였다. [그림 6]은 Sequential LSTM-Transformer의 구조이다.



Layer (type)	Output Shape	Param #	Connected to
input_layer_5 (InputLayer)	(None, 500)	0	-
embedding_5 (Embedding)	(None, 500, 300)	6,000,000	input_layer_5[0][0]
bidirectional_10 (Bidirectional)	(None, 500, 256)	439,296	embedding_5[0][0]
layer_normalization_20 (LayerNormalization)	(None, 500, 256)	512	bidirectional_10[0][0]
bidirectional_11 (Bidirectional)	(None, 500, 128)	164,352	layer_normalization_20
layer_normalization_21 (LayerNormalization)	(None, 500, 128)	256	bidirectional_11[0][0]
multi_head_attention_5 (MultiHeadAttention)	[(None, None, 300), (None, None, None)]	132,228	layer_normalization_21, layer_normalization_21
layer_normalization_23 (LayerNormalization)	(None, None, 300)	600	multi_head_attention_5
global_average_pooling1d_4 (GlobalAveragePooling1D)	(None, 300)	0	layer_normalization_23
global_max_pooling1d_5 (GlobalMaxPooling1D)	(None, 300)	0	layer_normalization_23
concatenate_7 (Concatenate)	(None, 600)	0	global_average_pooling1d_4, global_max_pooling1d_5
dense_34 (Dense)	(None, 64)	38,464	concatenate_7[0][0]
dense_35 (Dense)	(None, 1)	65	dense_34[0][0]

Total params: 6,775,773 (25.85 MB)  
 Trainable params: 6,775,773 (25.85 MB)  
 Non-trainable params: 0 (0.00 B)

[그림 6] Sequential LSTM-Transformer 구조

### 4.3 LSTM-Transformer Parelal

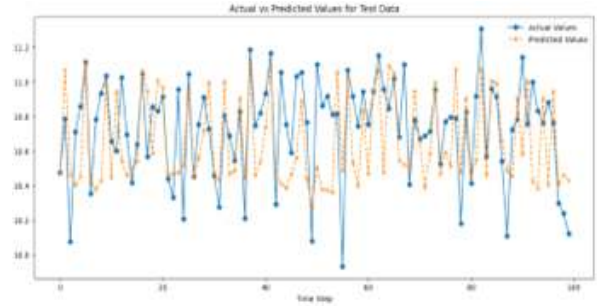
Sequential 모델에 이어, Parallel LSTM-Transformer 모델을 실험에 적용하였다. 이 모델은 LSTM과 Transformer 모듈을 병렬로 구성한 구조로, 두 가지 모듈이 서로 다른 방식으로 시계열 데이터를 학습하도록 설계되었다. [그림 7]은 LSTM-Transformer Parelal의 구조이다.

LSTM 부분은 양방향 LSTM 층을 통해 순차적 종속성을 학습하며, Layer Normalization을 적용하여 안정적인 학습이 이루어지도록 하였다. Transformer 부분은 Multi-Head Attention 모듈을 통해 중요한 피처를 학습하고, 장기적인 패턴을 반영하여 종속성을 강화한다.

Layer (type)	Output Shape	Param #	Connected to
input_layer_6 (InputLayer)	(None, 500)	0	-
embedding_6 (Embedding)	(None, 500, 300)	6,000,000	input_layer_6[0][0]
bidirectional_12 (Bidirectional)	(None, 500, 256)	439,296	embedding_6[0][0]
layer_normalization_24 (LayerNormalization)	(None, 500, 256)	512	bidirectional_12[0][0]
bidirectional_13 (Bidirectional)	(None, 500, 128)	164,352	layer_normalization_24
multi_head_attention_6 (MultiHeadAttention)	[(None, None, 300), (None, None, None)]	231,300	embedding_6[0][0], embedding_6[0][0], embedding_6[0][0]
layer_normalization_25 (LayerNormalization)	(None, 500, 128)	256	bidirectional_13[0][0]
layer_normalization_27 (LayerNormalization)	(None, None, 300)	600	multi_head_attention_6
concatenate_8 (Concatenate)	(None, 500, 428)	0	layer_normalization_25, layer_normalization_25
global_average_pooling1d_6 (GlobalAveragePooling1D)	(None, 428)	0	concatenate_8[0][0]
global_max_pooling1d_6 (GlobalMaxPooling1D)	(None, 428)	0	concatenate_8[0][0]
concatenate_9 (Concatenate)	(None, 856)	0	global_average_pooling1d_6, global_max_pooling1d_6
dense_40 (Dense)	(None, 64)	54,848	concatenate_9[0][0]
dense_41 (Dense)	(None, 1)	65	dense_40[0][0]

Total params: 6,801,229 (26.29 MB)  
 Trainable params: 6,801,229 (26.29 MB)  
 Non-trainable params: 0 (0.00 B)

[그림 7] LSTM-Transformer Parelal 구조



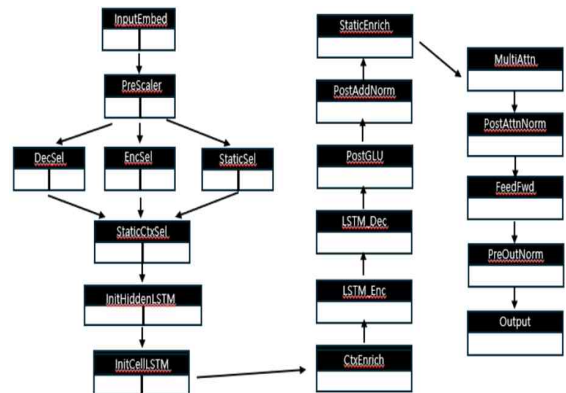
[그림 8] LSTM-Transformer Parelal 예측 결과 그래프

이 두 분기의 출력을 병합(Concatenate)한 후, Global Average Pooling과 Global Max Pooling을 적용하여 다양한 특징을 요약한 후 최종 예측을 수행한다. [그림 8]은 LSTM-Transformer Parelal의 예측 결과 그래프이다.

### 4.4 Temporal Fusion Transformer (TFT)

마지막으로, Temporal Fusion Transformer (TFT) 모델을 실험에 활용하였다. TFT 모델은 Temporal Fusion Transformer 모듈을 통해 과거와 미래의 시계열 정보를 융합하여, 다양한 시점에서 중요한 패턴을 효과적으로 학습하도록 설계되었다.

TFT 모델은 PyTorch Lightning을 활용해 정의되었으며, 모듈화된 구조 덕분에 효율적인 학습 및 검증 과정을 지원한다. 학습 과정에서는 Adam 옵티마이저와 Weight Decay를 적용하여 모델의 일반화 성능을 높이고 과적합을 방지하도록 구성하였다. [그림 9]는 Temporal Fusion Layer의 구조이다.



[그림 9] Temporal Fusion Transformer 구조

TFT의 핵심 요소는 여러 시점을 동시에 학습하는 Multi-Head Attention 및 Temporal Fusion Layer로, 이들은 단기 및 장기적 패턴을 동시에 반영하여 높은 예측 성능을 목표로 한다. 본 모델은 이전 모델들의 장점을 종합적으로 결합하여 시계열 데이터의 전체적인 패턴을 효과적으로 예측하기 위한 최종 모델로 활용되었다[5].

#### IV. 성능 분석

[표 1]의 비교 분석 결과를 바탕으로 판단하였을 때, LSTM-Transformer Parallel 모델이 가장 우수한 성능을 보였다. 이 모델은 낮은 RMSE와 MAPE 값을 기록하여 예측 오차가 적었고, R<sup>2</sup> 값이 0.99를 넘어 실제 데이터를 잘 설명할 수 있음을 보여주었다.

[표 1] 각 모델 성능 비교

Model	RMSE	MAPE (%)	R <sup>2</sup>
LSTM	0.15062	1.20054	0.72255
Sequential LSTM-Transformer	0.85348	5.82793	0.85924
LSTM-Transformer Parallel	0.19936	1.23025	0.99232
Temporal Fusion Transformer	4.4441	31.93467	-2.79307

이는 LSTM-Transformer Parallel 모델이 LSTM의 순차적 학습 능력과 Transformer의 병렬 처리 능력을 결합하여, RO 시스템의 복잡한 시계열 패턴을 효과적으로 학습한 결과로 해석할 수 있다.

#### IV. 결 론

본 연구에서는 RO 시스템의 유입 압력과 전기전도도를 예측하기 위해 LSTM, Sequential LSTM-Transformer, LSTM-Transformer Parallel, Temporal Fusion Transformer(TFT) 모델을 비교 분석하였다. 그중 LSTM-Transformer Parallel이 가장 뛰어난 성능을 보였다. 인공지능 모델을 통해 세척, 교체 시기를 알 수 있는 경우 운영자의 경험에만 의존하는 이전과 달리 보다 효율적이고 공정한 운영이 가능해진다.

#### 참 고 문 헌

1. Giridhar Maji, Soumya Sen, "Stock Price Prediction

Using LSTM on Indian Share Market", Proceedings of 32nd International Conference on Computer Applications in Industry and Engineering(2019): vol.63  
 2. 주형구, 임준목, "하수처리시설 수질인자의 머신러닝 예측 모델", 대한산업공학회지 제49권 제1회(2023): pp 95-106  
 3. 조은영, 최선홍, 장동우, "LSTM을 이용한 상수관망 내 실시간 유량 및 수질 데이터 예측", 한국수자원학회 2023년도 학술발표회(2023): pp 463-463  
 4. K. A. Althelaya, E. -S. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long-term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 2018, pp. 151-156  
 5. Xiaokang Hu, "Stock Price Prediction Based on Temporal Fusion Transformer", 3rd International Conference On Machine Learning, Big Data and Business Intelligence(MLBDBI)(2021)

#### 사사의 글

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음" (2021-0-01082)

# MTadGAN을 활용한 전기차 배터리팩 이상 예측 시스템

하창욱\*, 송현서\*, 유병일\*, 이재민\*, 이인수\*\*

## Anomaly Prediction System for Electric Vehicle Battery Packs Using MTadGAN

Chang Wook Ha\*, Hyeon Seo Song\*, Byeong-il Yu\*, Jae Min Lee\*, and In-Soo Lee\*\*

### 요 약

본 연구는 MTadGAN(다변량 시계열 데이터를 이용한 GAN 기반 모델)을 활용하여 전기자동차 배터리팩의 이상 여부를 실시간으로 예측하는 시스템을 제안하였다. 기존의 배터리 이상 감지 방식은 주로 사람이 데이터를 해석하고 이상 여부를 판단하는 데 의존했으나, MTadGAN은 모델이 자동으로 여러 변수를 동시에 분석하여 복잡한 이상 패턴을 더욱 정교하게 예측할 수 있다. 테스트 결과, 제안된 모델은 배터리 불량 여부를 높은 정확도로 예측하였으며, 이를 통해 전기자동차 배터리팩의 사고를 예방하는 데 중요한 기초 자료로 활용될 수 있음을 확인하였다.

### Abstract

This study proposes a system for real-time detection of anomalies in electric vehicle battery packs using MTadGAN, a GAN-based model for multivariate time series data. Traditional methods for detecting battery anomalies have relied on human interpretation of data to determine abnormalities. However, MTadGAN automatically analyzes multiple variables simultaneously, enabling more precise detection of complex anomaly patterns. The test results showed that the proposed model accurately predicted battery defects, confirming its potential as a crucial resource for preventing accidents in electric vehicle battery packs.

### Key words

Real-time Monitoring, Electric Vehicle Safety, Fault Detection, MTadGAN, Multivariate Analysis

## 1. 서 론

전기자동차의 보급이 확산됨에 따라 배터리 관련 사고가 빈번하게 발생하고 있다. 특히, 리튬이온 배터리의 화재 위험성은 전기자동차의 안전에 큰 위협이 되며, 배터리팩의 불량 여부를 조기에 감지하는 것이 필수적이다. 본 연구는 MTadGAN을 활

용하여 다변량 시계열 데이터를 분석함으로써 배터리팩의 이상 여부를 실시간으로 예측하는 시스템을 제안한다. 기존의 배터리 이상 예측 방식은 주로 사람이 수집된 데이터를 해석하고 주관적으로 이상 여부를 판별하는 데 의존했다. 반면, MTadGAN은 모델이 자동으로 여러 변수를 동시에 분석하고, 보다 정교하게 이상 패턴을 예측할 수 있어 기존 방

\* 경북대학교 IT대학 전자공학부 / hco0339@knu.ac.kr

\*\* 경북대학교 IT대학 전자공학부 교수 / insoolee@knu.ac.kr

법보다 신뢰성과 효율성이 높다. 본 연구에서는 모델이 감지한 이상치를 실시간으로 분석하고, 전기자동차 배터리팩의 불량 여부를 자동으로 판단해주는 시스템을 구축하여 전기자동차 배터리팩의 안정성을 강화하고 사고 예방에 기여하는 것을 목표로 한다.

## II. MTadGAN 모델을 이용한 배터리팩 불량 여부 판단

### 2.1.1 MTadGAN 개요

아래 그림 1은 MTadGAN의 구성을 보여준다.

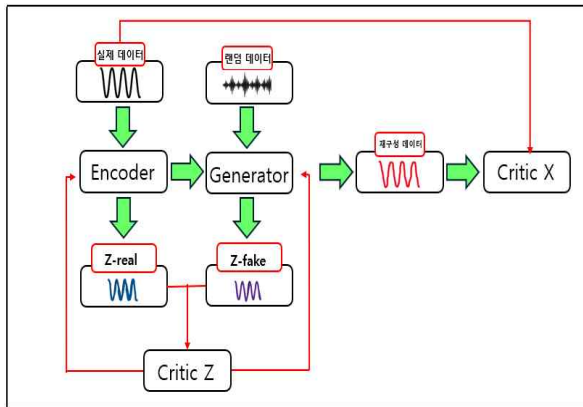


그림 1. MTadGAN 구성 요소  
Fig. 1. Components of MTadGAN

MTadGAN(Multivariate Time-series Anomaly Detection using Generative Adversarial Networks)[2], [3]은 다변량 시계열 데이터에서 이상치를 예측하는 GAN 기반 모델로, Generator와 판별자의 경쟁을 통해 비정상 패턴을 예측한다. Encoder는 데이터를 잠재 공간으로 변환하고, Generator는 이를 바탕으로 데이터를 생성한다. 판별자는 Critic X와 Z로 나뉘어 각각 데이터 간 유사성과 잠재 공간 차이를 평가한다.

### 2.1.2 손실 함수

재구성 손실[4]은 입력 데이터와 생성된 데이터 간 차이를 줄이고, 판별자 손실[6]은 실제와 생성된

데이터를 구별하는 능력을 평가하며, 두 손실의 균형을 통해 모델 성능이 개선된다.

### 2.1.3 성능 및 활용

MTadGAN은 시계열 데이터에서 정상 패턴과 이상 패턴을 효과적으로 구별하며, 다양한 분야에서 이상치 예측에 유용한 성능을 발휘한다.

## 2.2 데이터 전처리

실제 현장의 데이터는 의미 없는 값을 포함하거나 결측치, 또는 이상 데이터가 발생하여 데이터 품질이 떨어진다. 따라서 정확한 분석 결과를 얻기 위해 데이터 전처리를 실시한다. 아래 그림 2은 데이터 전처리 과정을 보여준다.

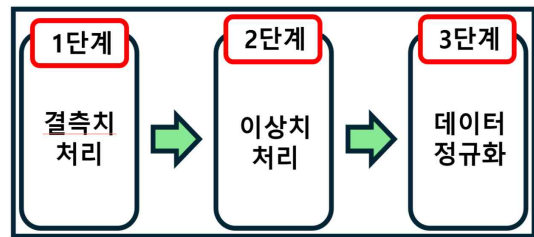


그림 2. 데이터 전처리 과정  
Fig. 2 Data Preprocessing Process

### 2.2.1 결측치 처리

데이터의 결측치를 처리하기 위해 null 값을 포함하여 결측이 발생한 행을 확인한다. 결측이 발생한 속성이 여러 개의 그룹으로 나누어지는 범주형인 경우, 해당 속성의 최빈값으로 결측치를 채우고, 수치형 속성은 결측치가 발생한 시점의 이전값과 이후값을 고려하여 보간법을 적용하여 처리한다.

### 2.2.2 이상치 처리

속성의 값이 일반적인 값과 편차가 큰 이상치[6]는 사분범위(inter-quartile range, IQR)를 이용하여 판별하고 제거한다. 아래 그림 3은 사분범위 모형을 보여준다.

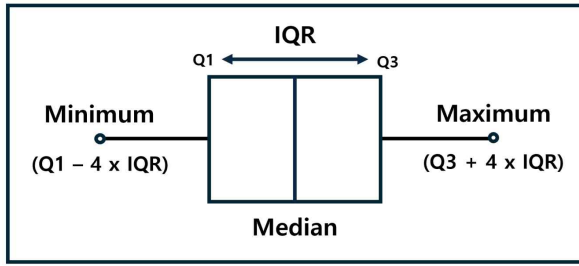


그림 3. 이상치 예측을 위한 사분범위(IQR) 모형  
Fig. 3 Inter-Quartile Range Model for Outlier Detection

IQR은 데이터의 75% 값인 3사분위수(Q3)와 25%의 값인 1사분위수(Q1)의 차로 정의되며, 변동성이 큰 데이터의 특성을 고려하여 속성의 값이 IQR의 4배 범위를 벗어나면 이상치로 간주해 값을 검출한다.

### 2.2.3 데이터 차원 축소 및 정규화

학습 알고리즘에서 데이터를 정확하게 모델링하기 위해 정규화를 실시한다. 먼저 다수의 셀 전압과 온도 데이터를 AI 모형에 입력하기 위해 주성분 분석(PCA)[7] 알고리즘을 적용하여 학습 데이터의 차원을 3차원으로 축소한다. 이후 축소된 데이터의 평균이 0, 분산이 1이 되도록 표준 정규화하여 데이터의 분포를 조정 한 후, 최대-최소 정규화를 통해 데이터값을 [0, 1] 범위로 변환한다.

## 2.3 모델 학습

아래 그림 4는 전체 학습 과정을 보여준다.



그림 4. MTadGAN 학습 과정  
Fig. 4. MTadGAN Training Process

전체 학습 과정[8]은 4단계로 구성된다. 아래 2.2.1절부터 2.2.4절은 각 단계별 학습 과정을 자세히 설명한다. 학습 데이터는 배터리 충전 데이터인 1000\_chg.csv 파일을 사용한다.

### 2.3.1 Critic X 학습

아래 그림 5는 Critic X 학습 과정을 보여준다.

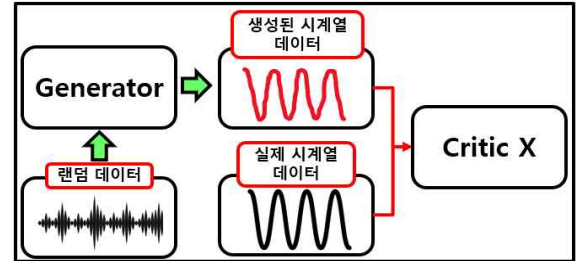


그림 5. Critic X 학습과정  
Fig. 5. Critic X Training Process

Critic X는 실제 시계열 데이터와 무작위 Latent Space에서 Generator가 생성한 시계열 데이터를 입력으로 받아 두 데이터를 비교하여 차이를 구분하는 역할을 한다. 이 단계의 목표는 Critic X가 실제 시계열 데이터와 생성된 시계열 데이터를 정확히 판별하도록 하는 것이다.

### 2.3.2 Critic Z 학습

아래 그림 6은 Critic Z 학습 과정을 보여준다.

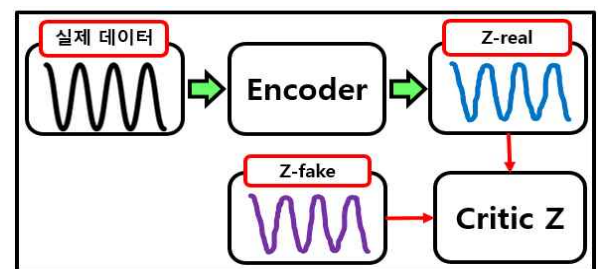
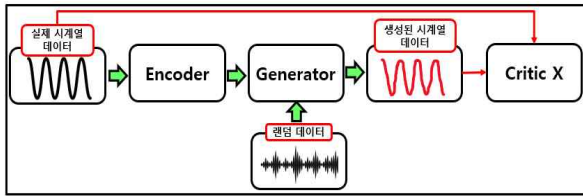


그림 6. Critic Z 학습과정  
Fig. 6. Critic Z Training Process

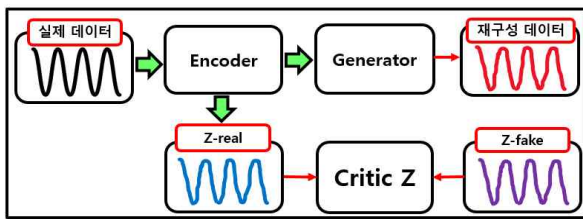
Critic Z는 실제 시계열 데이터를 인코딩한 Latent Space와 무작위로 생성된 Latent Space를 입력으로 받아 두 Latent Space 간의 차이를 구별하는 역할을 한다. 이 단계의 목표는 Critic Z가 실제 데이터에서 인코딩된 Latent Space와 무작위 Latent Space를 정확히 판별하도록 학습하는 것이다.

### 2.3.3 Encoder와 Generator 학습

아래 그림 (a)는 Critic X를 통한 Encoder와 Generator의 학습 과정을, 그림 (b)는 Critic Z를 통한 학습 과정을 보여준다.



(a) Critic X를 통한 Encoder와 Generator의 학습 과정  
(a) The training process of the Encoder and Generator through Critic X



(b) Critic Z를 통한 Encoder와 Generator의 학습 과정  
(b) The training process of the Encoder and Generator through Critic Z

그림 7. Critic X와 Critic Z를 통한 Encoder와 Generator의 학습 과정

Fig. 7. The training processes of the Encoder and Generator through Critic X and Critic Z

Encoder는 실제 시계열 데이터를 Latent Space로 변환하고, Generator는 무작위 Latent Space에서 시계열 데이터를 생성한다. Critic X는 실제 데이터와 생성된 데이터를 비교해 Generator가 더 정교하게 재생성할 수 있도록 피드백을 제공하며, Encoder는 이를 반영해 학습한다. Critic Z는 인코딩된 Latent Space와 무작위 Latent Space를 비교해 Encoder의 변환 적절성을 평가한다. Critic X와 Z의 피드백으로 Encoder와 Generator는 함께 최적화된다.

### III. 시뮬레이션 결과 및 고찰

테스트 데이터에도 PCA 변환을 적용하는 것은 일관성을 유지하고, 학습 시 파악한 데이터 구조를 테스트 데이터에 동일하게 반영하기 위함이다. 이를 통해 모델이 학습한 주성분 기반의 특성 공간에서 테스트 데이터를 평가하여 예측 성능의 신뢰성을 높일 수 있다. 아래 그림은 그 결과이다.

```
After PCA reduction df.shape = (5358, 4)
-----
   date    pca_1    pca_2    pca_3
0      1  21.169896  1.482040  0.943905
1      2  21.167226  1.482374  0.943312
2      3  21.333333  1.697547  0.584585
3      4  21.168563  1.483585  0.926521
4      5  21.167162  1.482097  0.943287
-----
```

그림 8. Test07\_NG\_dchg 테스트 데이터의 pca 차원 축소 결과

Fig. 8. The PCA dimension reduction results of the Test07\_NG\_dchg test data.

1000.csv파일을 통해 MTad-GAN 모델의 성능을 평가하기 위해 테스트 데이터 파일로는 Test07\_NG\_dchg.csv 파일을, anomaly 파일로는 Test07\_NG\_dchg.label.csv 파일을 이용한다. Test07\_NG\_dchg.csv 파일은 센서와이어 불량으로 인해 전압 이상을 보이는 데이터를 포함한다. 그림 8은 Test07\_NG\_dchg.csv 데이터 파일의 True Anomaly와 Predicted Anomaly를 나타낸다.

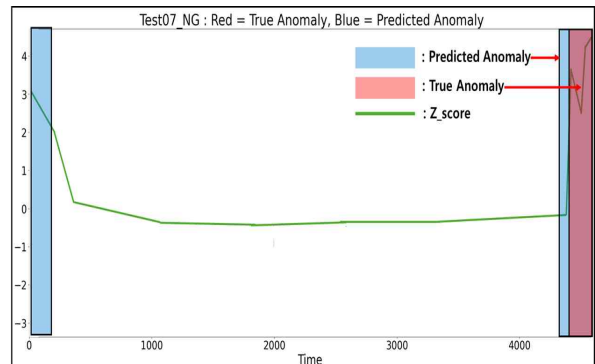


그림 9. Test07\_NG\_dchg 테스트 데이터로 예측한 이상치와 실제 이상치 비교

Fig. 9. Comparison of predicted anomalies with actual anomalies using the Test07\_NG\_dchg test data

테스트 데이터 또한 학습 데이터와 동일한 전처리 과정을 거친다. 이후 만들어진 MTadGAN 모델로 테스트 데이터에 대한 예측 계산을 수행한다. 그림 1의 파란색 박스는 Predicted Anomaly를 나타내고, 빨간색 박스는 True Anomaly를 나타낸다. 그림 1에서 볼 수 있듯이 PCA 성분 값이 비정상적인 점프를 보일 때 Anomaly가 발생되고, 이를 통해 예측한 값과 True Anomaly가 상당한 정도로 일치함

을 알 수 있다. 예측이 조금 벗어난 부분도 존재하는데, 이 부분은 학습 모형에 따라 예측 결과의 차이가 존재할 수 있기 때문이다.

#### IV. 결 론

본 연구에서는 MTadGAN 모델을 활용하여 전기자동차 배터리팩의 이상 여부를 실시간으로 예측하는 시스템을 제안하고 그 성능을 평가하였다. 테스트 결과, 제안된 모델은 대부분의 배터리 불량률 성공적으로 예측하였으며, 이를 통해 배터리팩의 안정성의 개선 가능성을 확인하였다. 본 연구에서 사용한 다변량 시계열 데이터 분석 기법은 배터리팩 내 수십 개에서 수백 개의 배터리셀에서 발생하는 다양한 변수(전압, 온도, 저항 등)를 효율적으로 다룰 수 있다. 특히 PCA와 MTadGAN을 적용한 이상 예측 시스템은 배터리셀 불량률 조기에 감지하여, 모듈 전체를 교체해야 하는 상황을 방지할 수 있다.

본 연구의 분석 결과는 전기자동차 배터리팩 제조 공정에서 품질 관리 및 모니터링 시스템을 개선하는 데 중요한 기초 자료로 활용될 수 있다. 향후 배터리셀과 모듈 생산에 특화된 제조 현장에서 본 모델이 적용된다면, 제품의 안정성을 높이고 사고 예방에 기여할 수 있을 것으로 기대된다.

#### 참 고 문 헌

- [1] 중소벤처기업부, Korea AI Manufacturing Platform(KAMP), CNC 머신 AI데이터셋, KAIST(UNIST, 이피엠솔루션즈), 2020.12.14., <https://kamp-ai.kr>
- [2] Houssam, Z., et al., "Efficient GAN-based Anomaly Detection," in Proc. of Conference, vol. X, pp. Y-Z, August 2018.
- [3] Li, D., et al., "Anomaly Detection with Generative Adversarial Networks for Multivariate Time Series," in Proc. of Conference, vol. X, pp. Y-Z, March 2019.
- [4] Goodfellow, I., Zaremba, W., et al., "Improved

Techniques for Training GANs," in Proc. of Conference, vol. X, pp. Y-Z, June 2016.

- [5] Li, D., Hu, G., and Spanos, C.J., "A data-driven strategy for detection and diagnosis of building chiller faults," IEEE Trans. on Systems, Man, and Cybernetics, vol. X, pp. Y-Z, November 2016.
- [6] H. Vinutha, B. Poornima, and B. Sagar, "Detection of outliers using interquartile range technique from intrusion dataset," in Information and Decision Sciences, Springer, pp. 511-518, March 2018.
- [7] Li, S., Wen, J., "A Model-based Fault Detection and Diagnostic Methodology based on PCA Method and Wavelet Transform," Journal of Applied Mechanics, vol. X, pp. Y-Z, May 2014.
- [8] Geiger, A., Liu, D., et al., "TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks," arXiv preprint, vol. X, pp. Y-Z, 2020.

# 응급실 혼잡도 모니터링을 위한 비침습적 랜덤 포레스트 기반 실시간 점유도 추정 시스템 설계

이재성\*, 이병찬\*, 임은혜\*, 정성원\*, 김영인\*

## Design of a Non-Invasive Real-Time Congestion Estimation for Emergency Department Congestion Using Random Forest-Based Occupancy Prediction

Jae-Seong Lee\*, Byeong-Chan Lee\*, Eun-Hye Im\*, Seong-Won Jeong\* and Young-In Kim\*

### 요약

응급실 혼잡도를 정확히 예측하면 환자 배정을 최적화하고 대기 시간을 줄이며 치료 효과를 높일 수 있으나, 기존 시스템은 실시간 성능과 정확성에서 한계를 지니고 있다. 또한, 전공의 대규모 사직으로 응급실 환자 미수용 사례 증가로 인해 신뢰할 수 있는 실시간 응급 의료 정보의 필요성이 더욱 강조되고 있다. 이러한 문제를 해결하기 위해 본 연구는 CO<sub>2</sub>, 온도, 소리 및 적외선 센서를 활용한 비침습적 실시간 혼잡도 추정 시스템을 제안하며, 개인정보를 보호하면서 응급실 점유율을 모니터링할 수 있도록 설계한다. 제안된 시스템은 다중 센서 데이터를 사용하여 더 정확한 점유율 추정과 혼잡도 예측을 수행하기 위해 랜덤 포레스트 모델을 사용한다. 실험 결과, 응급실 환경과 유사한 기존의 방 점유 데이터셋을 사용한 실시간 점유율 예측에서 우수한 인원수 추종이 가능함을 보였으며, 이 결과는 개인정보를 보호하면서도 효과적으로 응급실 혼잡도를 측정할 수 있음을 보였다.

### Abstract

Emergency Department(ED) congestion can optimize patient allocation, reduce waiting times, and improve treatment outcomes; however, current systems are limited in both real-time performance and accuracy. Additionally, the recent mass resignation of medical residents has heightened the need for reliable, real-time emergency medical information due to an increase in cases of ED patient rejection. To address these challenges, this study proposes a non-invasive, real-time congestion estimation system using CO<sub>2</sub>, temperature, sound, and infrared sensors designed to monitor ED occupancy while preserving patient privacy. The proposed system utilizes a Random Forest model to integrate multi-sensor data, enabling more accurate occupancy estimation and congestion prediction. Experimental results, derived from a real-time occupancy estimation using a dataset of room occupancy similar to an ED environment, demonstrate the system's high accuracy in tracking the number of occupants. These findings indicate that this approach effectively measures ED congestion while ensuring privacy protection.

### Key words

Emergency Department, Occupancy Estimation, Non-Invasive Sensor, Random Forest, Privacy Protection

---

\* 부산대학교 IT응용공학과



## I. 서론

지난 10년간 전 세계적으로 응급실(ED: Emergency Department) 과밀화 문제는 중요한 공중 보건 이슈로 대두되었으며[1], 특히 최근 들어 국내에서 발생한 전공의 사직과 파업으로 응급 환자들이 적시에 치료받지 못하거나, 응급실이 환자를 수용하지 못하는 사례도 빈번히 보고되고 있다. 응급 상황에서 구급대원이 응급실의 수용 가능 인원과 점유율을 정확히 파악하는 능력은 환자의 생명과 직결되므로, 이를 지원하는 머신 러닝 기반 연구가 활발히 진행 중이다[1]. 그러나 현재의 응급실 종합상황판 시스템은 대기 환자의 수를 정확히 반영하지 못하거나, 입력 지연으로 인한 한계가 있으며, 기존 점유율 예측 연구에서는 주로 침습적 센서를 사용하여 개인정보 침해 문제와 높은 비용 부담이 발생하는 문제가 있다[2].

응급실은 일반 대중이 의료 서비스를 접하는 주요 시작점이므로, 예측할 수 없는 수요로 인해 병원 운영과 관리에 어려움을 겪기 쉽다. 이러한 환경에서 ED 방문자 수의 정확한 추정은 자원 배치 최적화, 비용 절감, 공공 신뢰도 제고에 필수적이다[1]. 최근 ED 혼잡도와 점유율 예측을 위해 ARIMA, 선형 회귀(LR), 푸아송 회귀, 인공 신경망(ANN) 등 다양한 접근법이 사용되고 있으며, 이러한 모델은 다양한 요인에 따라 예측 성능이 달라진다[1]. ED 혼잡도를 효율적으로 예측하기 위해서는 현재의 ED 혼잡도를 구하여야 하며, 이를 통해 병원은 사전에 자원을 적절히 배분하여 환자 치료의 질을 개선할 수 있을 것이다. 또한, 현재 ED 혼잡도 추정에 기반한 예측은 단순히 자원 최적화를 넘어 환자 배정과 의료 서비스 효율성을 극대화하는 중요한 역할을 한다. 따라서 실시간 점유율 추정 시스템은 점유 상태에 따라 긴급 상황에서의 신속 대응을 가능하게 한다[2].

본 연구에서는 최소한의 비침습적 센서를 활용한 머신 러닝 기반 점유율 추정 시스템을 제안하며, 기존 침습적 센서 사용의 개인정보 침해와 비용 문제를 개선하고자 한다. 특히 온도, 소리, 적외선(PIR: Passive infrared) 및 CO2 센서의 데이터를 사용하는 방식으로 실시간 점유 상태를 더욱 정확하게 추정

할 수 있도록 설계하였다[2]. 이 시스템을 활용한다면 응급실 대기 시간 단축과 치료 효과 개선, 그리고 의료 자원의 효율적 사용에 기여할 수 있을 것으로 기대한다.

## II. 시스템 개요

### 2.1 시스템 기능

#### (1) 비침습적 실시간 혼잡도 추정 시스템

본 연구에서 제안하는 시스템은 그림 1과 같이 이산화탄소 농도, 온도, 빛 및 적외선 센서를 통해 수집한 데이터를 기반으로 응급실 내 실시간 혼잡도를 추정한다. 이러한 실시간 혼잡도 모니터링은 응급실 간 환자 분산을 최적화하여 대기 시간을 단축하고 치료 효율성을 향상할 수 있다. 이를 통해 환자가 적절한 곳에서 치료를 받을 기회를 높일 수 있으며, 혼잡도가 낮은 응급실로 안내함으로써 전체적인 응급 의료 서비스의 효율성을 높일 수 있다.

제안한 시스템의 주요 특징은 비침습적 센서만을 활용하여 공간 내 점유율을 추정한다는 점이다. 이는 기존의 침습적 방식과 달리 개인정보 보호를 보장하면서도 정확한 혼잡도 측정이 가능하다. 특히 의료 환경에서 요구되는 프라이버시 보호 요구사항을 충족하면서 지속적인 데이터 수집이 가능하다는 장점이 있다. 따라서 환자 및 방문객의 개인정보를 침해하지 않으면서 공간 내 인원 점유율을 추정하여 개인 정보 보호에 대한 윤리적 문제를 해결할 수 있다. 또한 환자의 신뢰를 높이고, 병원 내에서 데이터 수집을 지속할 수 있도록 한다.

### 2.2 제안 알고리즘

본 연구의 제안된 알고리즘은 그림 1과 같다. 이 알고리즘은 응급 상황 발생 시 구급대원이 환자에게 가장 적합한 응급실을 추천하는 것으로, 응급실 혼잡도와 사용자 위치라는 두 가지 기준을 적용한다. 현재 응급실의 혼잡도는 응급실 종합상황판의 지표를 참고하여 초록색(원활), 노란색(혼잡), 빨간색(포화) 세 단계로 분류된다. 시스템은 사용자의 현재 위치를 입력 값으로 받아 OHAT 권고 기준에 따라 80분 이내에 도달 가능한 응급실을 고려한다[4].

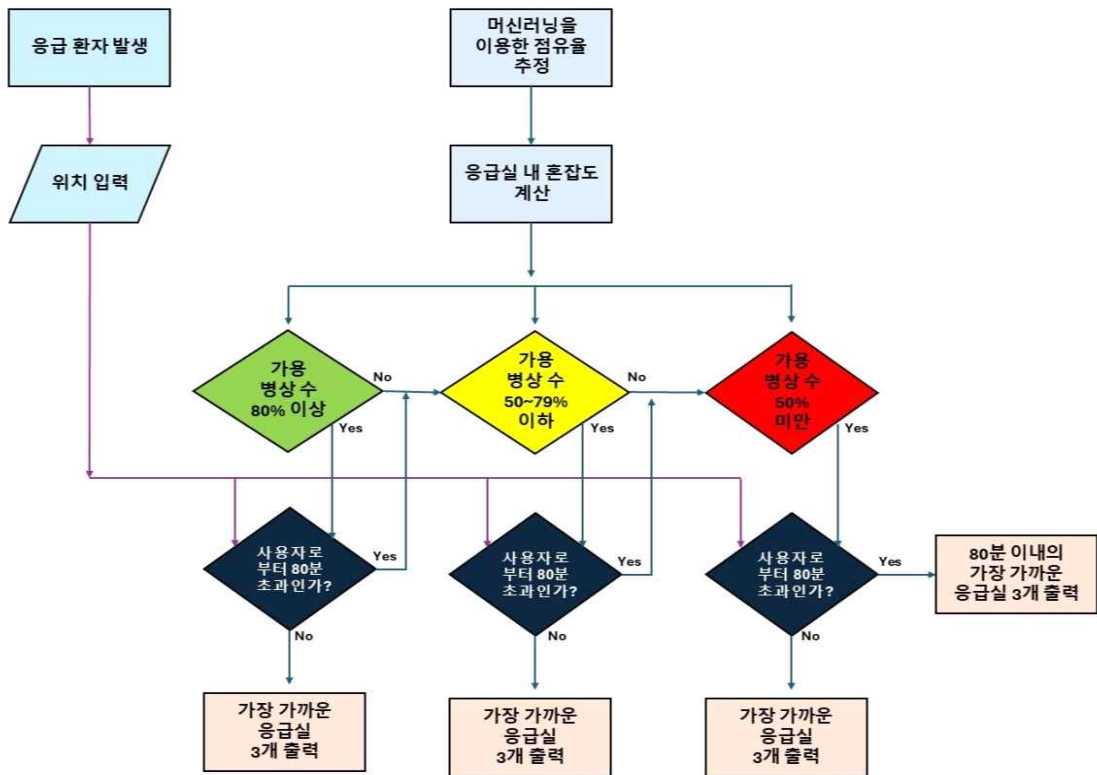


그림 1. 응급 환자 발생 시 점유율 추정치를 이용한 응급실 추천 알고리즘  
 Fig. 1. Emergency Department Recommendation Algorithm Using Occupancy Estimates in Emergency Cases

환자 배정을 위해 우선순위 기반 분류 접근법 (priority-based classification approach)을 사용하여 혼잡도가 낮은 응급실을 먼저 배정한다. 배정 순서는 원활 → 혼잡 → 포화 순으로 진행되며, 보건복지부의 응급의료체계에 따라 최적의 접근성을 보장하기 위해 1시간 이내(최대 100km) 응급실을 먼저 필터링한다. 주변에 원활한 상태의 응급실이 없으면, 가장 가까운 혼잡한 응급실로, 혼잡한 응급실도 없다면 포화 상태의 응급실로 차례대로 배정한다.

본 연구의 알고리즘은 응급실 혼잡도를 정확히 예측하기 위하여, 비침습적 센서를 이용하여 추정된 응급실 점유율을 중앙응급의료센터에서 제공되는 최대 병상수로 나누어 실시간 혼잡도를 추정하고, 추정한 혼잡도를 가용 상태에 따라 구분한다. 이 방법으로 개인정보 침해 없이 응급실의 혼잡도를 추정할 수 있다.

제안한 알고리즘을 적용한 시스템에서는 제안한 알고리즘과 사용자의 위치를 복합적으로 고려하여 환자가 적시에 치료받을 수 있는 응급실을 추천한다.

### 2.3 혼잡도 추정 모델 개발

본 논문에서는 실제 응급실 데이터를 수집하는데 어려움이 있어, 유사한 환경에서 방 점유 인원을 추정하는 [3]의 데이터셋을 사용하여 실험을 진행하였다. 이 데이터셋은 CO<sub>2</sub>, 온도, 소리, 적외선 센서를 이용해 제한된 실내 공간에서 수집된 데이터로 구성되어 있다.

실험은 해당 실내 공간을 응급실의 최소기능제품 (MVP: Minimum Viable Product)으로 가정하여 진행하였다. 첫 번째 단계에서는 비침습적 센서인 CO<sub>2</sub>, 온도, 소리, 적외선 센서를 이용하여 데이터를 수집한다고 가정하여 [3]의 데이터셋을 사용하였다. 두 번째 단계에서는 센서 데이터의 스케일을 조정하기 위해 표준화를 수행하였다. 다음으로, [3]과같이 기계 학습 알고리즘인 선형 판별 분석(LDA: Linear Discriminant Analysis), 이차 판별 분석(QDA: Quadratic Discriminant Analysis), 선형 커널을 사용한 서포트 벡터 머신, RBF 커널을 사용한 서포트 벡터

머신과 RF(Random Forest)를 적용하여, 10-겹 교차 검증 기법으로 각 모델의 정확도와 F1-스코어를 비교하였다. 그 결과, 실내 점유율을 추정하는 데 있어 RF 알고리즘이 가장 우수한 성능을 보여, RF 알고리즘을 선택하여 사용하였다.

### III. 결론

본 논문에서는 응급실 혼잡도 예측 성능을 높이기 위해 비침습적 센서 데이터를 활용한 랜덤 포레스트(RF) 모델을 제안하였다. 특히, CO2, 온도, 적외선 등의 비침습적 센서를 이용해 실시간으로 점유율을 예측함으로써 환자 대기 시간 감소와 의료 서비스 효율성 증대를 목표로 하였다.

본 연구는 실제 응급실 환경을 반영하기 어려운 한계를 극복하기 위해 실내 공간을 응급실의 최소 기능제품(MVP)으로 가정하고 실험을 진행했으며, 이 과정에서 다중 센서 데이터를 활용한 점유율 예측의 가능성을 확인하였다. 그러나, 실험 데이터는 실제 응급실이 아닌 유사 환경에서 수집된 것이므로, 향후 연구에서는 실제 응급실에서 수집한 데이터를 통해 RF 모델의 성능을 더욱 정밀하게 검증할 필요가 있다.

이를 통해 응급실 특유의 환경적 요인을 반영한 혼잡도 예측이 가능해질 것이며, 궁극적으로 더욱 신뢰도 높은 응급실 혼잡도 예측 시스템 개발에 기여할 수 있을 것이다. 이러한 추가 검증이 이루어진다면, 본 연구의 모델은 실시간 의료 자원 관리와 환자 배정 최적화에 있어 효과적인 도구로서 활용될 수 있을 것으로 기대된다.

### 참 고 문 헌

[1] E. Silva, M. F. Pereira, J. T. Vieira, J. Ferreira-Coimbra, M. Henriques, and N. F. Rodrigues, "Predicting hospital emergency department visits accurately: A systematic review," *Int. J. Health Plann. Manage.*, vol. 38, no. 4, pp. 904-917, Oct. 2023.

[2] P. Chaudhari, Y. Xiao, M. M. C. Cheng, and T.

Li, "Fundamentals, algorithms, and technologies of occupancy detection for smart buildings using IoT sensors," *Sensors*, vol. 24, no. 7, pp. 2123-2173, Mar. 2024.

[3] A. P. Singh, V. Jain, S. Chaudhari, F. A. Kraemer, S. Werner and V. Garg, "Machine learning-based occupancy estimation using multivariate sensor nodes," *IEEE Globecom Workshops*, pp. 1-6, Dec. 2018.

[4] Jang W. M., Lee J., Eun S. J., Yim J., Kim Y., and Kwak M. Y., "Travel time to emergency care not by geographic time, but by optimal time: A nationwide cross-sectional study for establishing optimal hospital access time to emergency medical care in South Korea," *PLoS ONE*, vol. 12, pp. 1-15, May 2021.

# 자연어 기반 SLD 자동 생성과 GIS 스타일링 자동화

최진형\*, 최권택\*\*

## Natural Language to SLD: Automating Style Generation for GIS

Jinhyeong Choe\*, and Kwontaeg Choi\*\*

### 요 약

기존의 지도 스타일링 방식은 매번 개별적으로 XML(SLD) 파일을 작성해야 하는 번거로움과 오류 가능성이 높았다. 이를 해결하기 위해 이 논문에서는 사용자의 자연어 입력을 기반으로 SLD 스타일 파일을 자동 생성하고 Geoserver에 적용할 수 있는 시스템을 제안한다. 이 시스템은 Python과 Geoserver의 REST API를 통해 XML(SLD) 파일을 자동 생성하며, 복잡한 SLD 구문을 알지 못하는 사용자도 쉽게 사용할 수 있다. 제안된 방법은 시각화 작업의 효율성을 높이고, GIS 데이터의 스타일링 일관성을 개선한다.

### Abstract

The conventional methods for GIS styling require the manual creation of XML (SLD) files, which is labor-intensive and prone to errors. This paper proposes a system that automates the generation and application of SLD style files based on user's natural language input. Using Python and Geoserver's REST API, the system automatically creates XML (SLD) files, allowing users without knowledge of complex SLD syntax to easily apply desired styles. The proposed method improves the efficiency of visualization and enhances consistency in GIS data styling.

### Key words

GIS styling, SLD automation, natural language input, Geoserver, visualization efficiency

## I. 서 론

GIS(Geographic Information System) 정보는 위치와 지리적 속성을 포함한 데이터를 다루고 분석하는 데 유용하며, 다양한 분야에서 공간 정보를 이

해하고 활용하는 중요한 역할을 한다. 이때 GIS에서 데이터를 표현하는 데 사용되는 주요 요소는 Shape와 Metadata이다. Shape는 지리적 객체의 위치와 형태를 점(Point), 선(Line), 다각형(Polygon) 등으로 나타내어 공간 구조를 정의하며, Metadata는 이러한

\*강남대학교 ICT융합공학부, wlsjud2414@kangnam.ac.kr

\*\*강남대학교 ICT융합공학부, kwontaeg.choi@kangnam.ac.kr(교신저자표시)

객체의 속성 정보를 담고 있다. 예를 들어 일반적인 지도에서는 도시 경계가 다각형으로 표현되는 Shape 정보와 때 각 도시의 인구, 면적 등의 특성이 포함된 Metadata로 이루어져 있다. 이러한 Shape와 Metadata가 결합된 데이터를 시각화하기 위해 XML 기반의 스타일링 언어인 SLD (Styled Layer Descriptor)를 활용해 다양한 시각적 규칙을 적용할 수 있다.

SLD는 지도 위의 객체를 색상, 크기, 투명도 등의 요소로 시각화하여 정보를 보다 직관적으로 표현할 수 있게 하지만, 개별 데이터에 맞게 SLD 파일을 매번 새로 작성해야 하는 문제점이 존재한다. 데이터마다 다른 스타일링 규칙을 코딩해야 하는 과정은 시간과 노력 소모가 클 뿐 아니라 오류 가능성도 높다. 또한, 표준화된 규칙이 없는 상태에서 반복적인 스타일링 작업을 수행하는 것은 비효율적이며, 지도 시각화 확장에 제약을 줄 수 있다.

본 연구는 이러한 문제를 해결하기 위해 사용자의 자연어 입력[1]을 통해 SLD 스타일 파일을 자동 생성하고 적용할 수 있는 시스템을 제안한다.

## II. 시스템 구성

본 연구에서는 사용자가 자연어로 스타일 변경 요청을 입력하면 2차원 GIS 데이터를 위한 시각적 스타일링이 자동으로 생성되도록 설계한 시스템을 제안한다. 이 시스템은 Geoserver와 Python 서버를 활용하여 REST API로 XML 기반의 스타일 파일을 자동 생성하고 적용하도록 구성되어 있으며, 주요 구성 요소는 사용자 입력 처리, 프롬프트 엔지니어링[2]을 통한 XML 생성, Geoserver 업로드 및 스타일 적용의 세 단계로 이루어진다.

### 2.1 사용자 입력 처리 및 프롬프트 엔지니어링

시스템의 첫 번째 단계는 사용자의 자연어 입력을 처리하는 모듈로, 사용자가 자유롭게 입력한 스타일링 요청을 수신한다. 사용자는 예를 들어 “주경계를 빨간색으로 채우고 노란색 테두리를 설정해주세요”와 같은 형식으로 스타일링 요청을 할 수 있으며, 이 요청은 시스템 내부에서 프롬프트 엔지

니어링 모듈로 전달된다. 프롬프트 엔지니어링과 XML 생성 모듈은 사용자로부터 전달된 자연어 요청을 해석하여 XML(SLD) 형식의 스타일 파일로 자동 변환하는 과정을 담당한다. 입력된 자연어는 GPT 기반의 자연어 처리 모델을 통해 분석되어 스타일링에 필요한 색상, 모양, 크기, 테두리 두께 등과 같은 요소를 추출하며, 이를 XML 규칙에 따라 적절히 저장하여 SLD 파일 형태로 변환한다. 이 XML 파일은 Geoserver에 적용될 준비를 마친 상태로 Python 서버를 통해 전송된다.

### 2.2 Geoserver 업로드 및 스타일 적용

Geoserver 업로드 및 스타일 적용 모듈에서는 생성된 SLD 파일이 Geoserver의 REST API를 통해 전송되며, 파일은 Python 서버를 통해 Geoserver의 스타일 폴더에 POST 요청으로 업로드된다. Geoserver는 서버에서 파일의 유효성을 검토한 후 XML 스타일을 지정된 GIS 데이터셋에 반영하여, 요청한 스타일이 기존 지도에 새로운 시각적 요소로 적용된다. 파일 전송이 성공적으로 이루어지면 Python 서버는 Geoserver로부터 성공 신호를 수신하고, 사용자에게 스타일이 반영된 지도의 업데이트된 모습을 확인할 수 있는 상태가 된다. 이 과정을 통해 사용자는 자신의 자연어 요청에 따라 새롭게 스타일링된 GIS 지도를 시각적으로 확인할 수 있다.

### 2.3 실험 내용

본 연구에서 제안한 시스템의 정확성과 유효성을 평가하기 위해 “폴리곤 색상은 빨간색, 테두리는 노란색, 테두리 굵기는 4”로 설정된 스타일링 요청을 실험에 적용하였다.

```
...
<Rule>
<PolygonSymbolizer>
  <Fill>
    <CssParameter name="fill">#FF0000</CssParameter>
  </Fill>
  <Stroke>
    <CssParameter name="stroke">#FFFF00</CssParameter>
    <CssParameter name="stroke-width">4</CssParameter>
  </Stroke>
</PolygonSymbolizer>
</Rule>
...
```

그림 2. 생성된 SLD 파일 예시  
Fig. 2. Example of Generated SLD File

실험 결과, 시스템은 해당 요청에 대한 SLD 스타일 파일을 정확하게 생성하고, 요청된 스타일을 지도에 잘 반영하였다. 그림1은 생성된 SLD 파일의 예시이다.

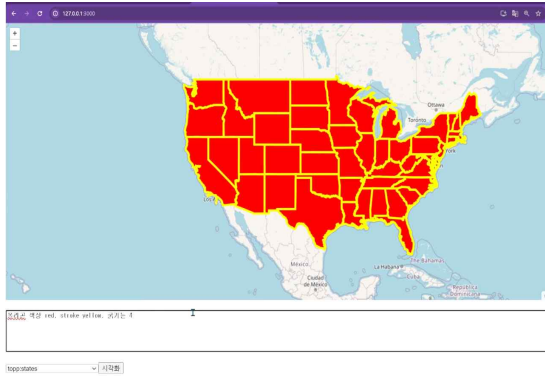


그림 2. 자연어 기반 SLD 적용 후  
Fig. 2. After Applying Natural Language-Based SLD

다만, 사용자의 단순한 스타일 요청에서는 높은 정확성을 보였으나, 스타일링 규칙이 복잡해질수록 정확도가 다소 낮아지는 경향이 관찰되었다. 예를 들어 여러 색상, 다양한 패턴 또는 특정 조건이 결합된 스타일을 자연어로 요청할 경우 일부 속성이 누락되거나 잘못 해석될 가능성이 있었다. 따라서 향후 연구에서는 복잡한 스타일링 요청에 대해 시스템의 정확성을 높이기 위한 알고리즘 개선과 추가적인 구문 검증 기능이 개선되어야 할 것이다.

#### 2.4 시스템의 한계와 향후 개선 방향

본 연구에서 제안한 시스템은 자연어를 통해 Geoserver 스타일링을 손쉽게 변경할 수 있는 장점을 갖추고 있다. 비전문가도 복잡한 SLD XML 구문 없이 원하는 스타일을 자연어로 요청할 수 있어 GIS 데이터 시각화의 접근성을 높이고, 다양한 사용자가 공간 데이터를 효과적으로 활용할 수 있도록 돕는다. 그러나 자연어의 모호성으로 인해 복잡하거나 추상적인 요청에 대해서는 정확한 XML 변환이 어려울 수 있는 한계가 존재한다.

또한, 현재 시스템은 기본적인 XML 스타일 구성만 지원하여 고급 GIS 스타일링(예: 복잡한 3D 시각화, 다중 스타일 적용 등)에는 제한이 있다.

Geoserver의 응답 속도 역시 XML 파일 크기와 네트워크 환경에 따라 영향을 받아, 실시간 반영이 지연될 가능성도 있다. 향후 연구에서는 이러한 한계를 보완하기 위해 복잡한 스타일 요청에 대한 알고리즘 개선과 구문 검증 기능 강화, 그리고 Geoserver와의 통신 효율성을 높이는 방안을 탐구할 필요가 있다.

### III. 결론

본 연구에서는 사용자의 자연어 입력을 통해 SLD 스타일 파일을 자동 생성하고 Geoserver에 적용하여 GIS 데이터 시각화를 자동화하는 시스템을 제안하였다. 이 시스템은 복잡한 XML 구문을 작성하지 않고도 지도 스타일을 변경할 수 있어 GIS 시각화 작업의 효율성을 높이며, 비전문가도 쉽게 사용할 수 있다는 장점을 가진다.

다만, 자연어의 모호성이나 복잡한 스타일링 요청에서 일부 속성이 정확히 반영되지 않는 한계가 있었다. 이를 보완하기 위해 향후 연구에서는 자연어 해석의 정밀성을 높이는 알고리즘 개선이 필요하다.

### 참 고 문 헌

- [1] Yu-Gyeong Hwang, Jin-Young Shin, and Seok-Jong Yu, "Dataset Creation and Visualization System based on Natural Language Prompt", The Journal of Korean Institute of Information Technology, Vol. 22, No. 4, pp. 23-28, 2024.
- [2] Myung-Suk Lee, "A Study on Prompt Engineering Techniques based on chatGPT", Proceedings of the Korean Society of Computer Information Conference, 대전, pp. 715-718, 2023-01-12.

# 딥러닝기반의 감정인식 및 IoT 기술을 이용한 스마트디퓨저 제작

김주성\*, 장원혁\*, 최우원\*, 황진영\*, 정진한\*, 이인수\*\*

## Implementation of Smart Diffuser Using Deep Learning-based Emotional Recognition and IoT Technology

Joosung Kim\*, Wonhyeok Jang\*, Woowon Choi\*, Jinyoung Hwang\*, Jinhan Jung\*, Insoo Lee\*\*

School of Electronic Engineering, Kyungpook National University, Republic of Korea

### 요 약

본 논문에서는 지난 논문에서 제안한 사용자의 감정상태를 인식하여 적절한 향을 낼 수 있는 딥러닝 기반의 스마트 디퓨저 시스템의 개선사항에 대해 기술한다. CNN을 이용한 감정인식모델의 정확도를 데이터 증강 기법을 통해 기존 대비 20% 이상 개선했다. 또한 IoT제어와 GUI를 추가한 시스템에 대해 요약한다. 본 논문에서는 스마트 디퓨저 시스템의 개선방향과, 데이터 증강기법에 대해 기술한다.

### Abstract

In this paper, we describe the improvements of the deep learning-based smart diffuser system that can recognize the user's emotional state and produce appropriate scents proposed in the last paper. The accuracy of the emotion recognition model using CNN has been improved by more than 20% compared to the previous one through the data augmentation technique. It also summarizes the system with IoT control and GUI added. In this paper, the improvement direction of the smart diffuser system and the data augmentation technique are described.

### Key words

smart diffuser, emotional analysis, CNN, deep learning, facial expression detecting, data augmentation, IoT

## 1. 서 론

현대 사회에서 스트레스와 감정 조절의 중요성이 강조됨에 따라 기억과 정서, 행동에 중요한 영향을 미치는 향기 기술에 대한 관심이 높아지고 있다. 이

전 연구에서는 딥러닝을 통해 사용자 감정을 인식하고 이에 맞는 향을 제공하는 스마트 디퓨저 시스템을 개발하여 감정 조절을 돕고자 했다. 그러나 해당 시스템에는 한계가 있었다. 감정 인식 모델이 다양한 표정을 정확히 해석하지 못해 인식 정확도가

\* 경북대학교 전자공학부 학부생

\*\* 경북대학교 전자전자공학부 교수 (교신저자), insoolee@knu.ac.kr

떨어지는 문제가 있었다. 특히 행복 감정과 분노 감정 간 인식정확도 차이가 존재하여 감정 분석의 균형 잡힌 결과가 나오기 어려웠다[1].

본 논문에서는 이러한 한계를 보완하여 감정 인식의 정밀도를 개선한 스마트 디퓨저를 제안한다. 더불어, 터치스크린과 애플리케이션을 통한 IoT 제어 기능을 추가하여 사용자 편의성을 높이고 개인 맞춤형 향기 경험을 제공할 수 있도록 설계를 강화했다.

## II. 감정인식 모델링 개선

### 2.1 데이터 증강기법

감정 인식 모델의 성능을 개선하기 위해 데이터 증강 기법을 사용하였다. 감정 이미지 데이터는 수집이 어렵고 제한된 양의 데이터만 확보할 수 있었기 때문에, 이를 보완하여 모델의 학습 효율을 높이기 위해 다양한 데이터 증강 기법을 적용하였다[2]. 본 연구에서는 본 데이터셋의 훈련 및 검증 이미지에 대해 회전, 좌우 반전, 명도 조정 등의 변형을 통해 데이터셋의 양과 다양성을 인위적으로 증가시켰다[3]. 그림 1은 다양한 데이터 증강 기법의 예시이다.

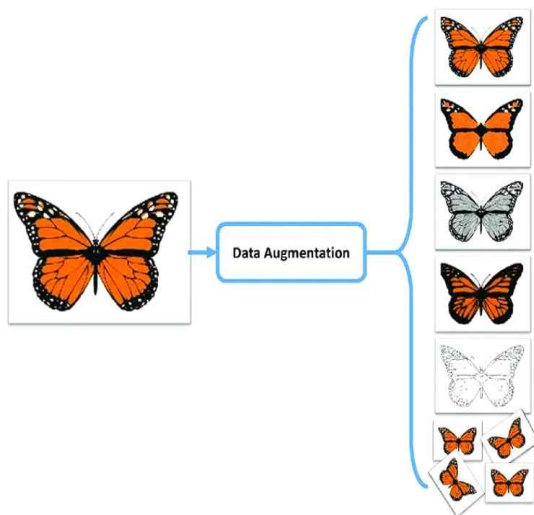


그림 1. 데이터 증강기법 예시  
Fig 1. Examples of data augmentation

#### 1. Rotation

이미지를 왼쪽으로 20도, 오른쪽으로 20도씩 회전하여 각기 다른 각도에서 발생할 수 있는 시각적 변화를 모델이 학습하도록 하였다. 이로 인해 감정 인식 성능이 특정 각도에 편중되지 않고, 다양한 상황에서도 높은 정확도를 유지할 수 있도록 도왔다.

#### 2. Flipping

이미지를 좌우로 뒤집어 데이터셋을 확장함으로써 모델이 학습할 수 있는 데이터 패턴을 더 많이 확보하였다. 이를 통해 동일한 감정이지만 다른 방향에서 관찰될 수 있는 다양한 이미지를 모델이 인식하고 학습할 수 있게 하였다.

#### 3. Lighting condition

이미지의 명도를 0.8배로 어둡게, 1.2배로 밝게 조정하여 여러 조명 조건에서도 감정을 정확히 인식할 수 있도록 학습시켰다. 이를 통해 감정 이미지가 밝거나 어두운 환경에서도 모델이 일관되게 감정을 판별할 수 있도록 데이터의 다양성을 확보했다.

#### 4. Scaling

이미지를 일정 비율로 확대하거나 축소하여 다양한 거리에서 촬영된 효과를 모델이 학습할 수 있도록 했다. 이를 통해 얼굴이 클로즈업되거나 멀리서 촬영된 경우에도 감정을 일관되게 인식할 수 있도록 학습했다.

#### 5. Gaussian Noise

이미지에 가우시안 잡음을 추가하여 다양한 노이즈 환경에서도 감정을 잘 인식할 수 있도록 훈련하였다. 실제 상황에서의 다양한 촬영 환경을 모델이 견딜 수 있도록 돕는 중요한 기법이었다

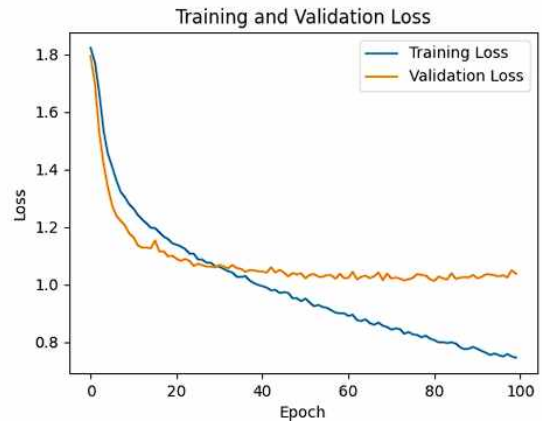
#### 6. Color Jittering

색조, 채도 등을 무작위로 변경하여 감정 이미지의 색상이나 채도에 관계없이 감정을 인식하도록 모델을 훈련했다. 이는 이미지에 다양한 조명과 색상 왜곡이 있을 때에도 모델의 일관성을 높여주었다.

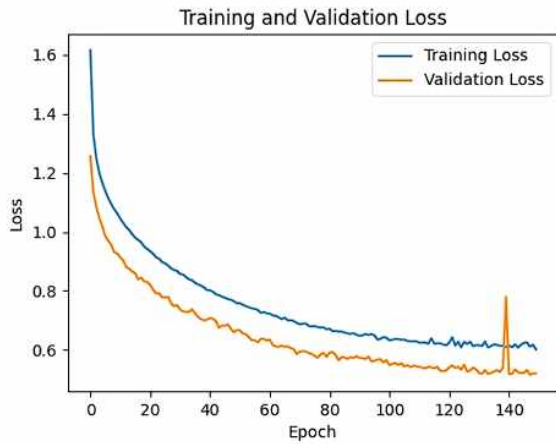


## 2.2 모델링 개선결과

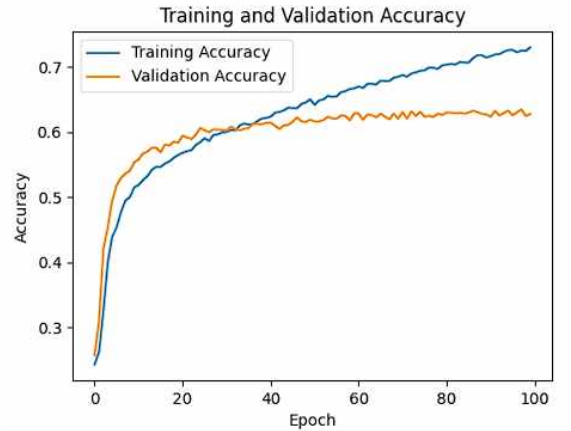
이러한 데이터 증강 기법들을 통해 원본 데이터의 양을 효과적으로 증가시키고, 모델이 학습할 수 있는 다양한 조건과 변형을 제공함으로써 모델의 일반화 능력을 개선하였다. 그 결과, 검증 데이터에 대한 정확도(val\_accuracy)가 기존의 63%에서 83%로 향상되었다. 이는 증강 기법을 통한 데이터 다양성 확보가 모델의 성능 향상에 큰 기여를 했음을 보여 준다[4][5]. 그림 2는 개선 전과 후 모델학습 그래프 차이를 나타낸다.



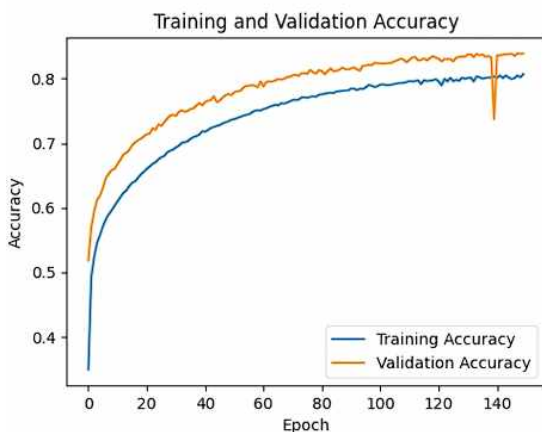
(c) 개선 전 훈련과 검증 loss  
(c) Training and validation loss before improvement



(a) 개선 후 훈련과 검증 loss  
(a) Training and validation loss after improvement



(a) 개선 전 훈련과 검증 accuray  
(a) Training and validation accuracy before improvement



(b) 개선 후 훈련과 검증 accuray  
(b) Training and validation accuracy after improvement

그림 2. 모델 학습 그래프 비교  
Fig. 2. Comparison of model learning graph

## III. 제안한 스마트 디퓨저

### 3.1 Firebase 기반 IoT제어

Firebase는 구글에서 제공하는 모바일 및 웹 플랫폼으로 실시간 데이터 베이스의 기능을 제공한다. Firebase Realtime Database는 클라우드 기반 NoSQL 데이터베이스로, 데이터 변경 사항을 실시간으로 반영하여 여러 클라이언트 간의 데이터 동기화를 가능하게 한다. 어플리케이션은 안드로이드 스튜디오 환경에서 Kotlin과 Jetpack Compose를 사용하여 개발되었으며 Firebase Realtime Database를 통해 스마트 디퓨저와 통신한다. 사용자는 시각적 인터페이스를 활용해 스마트 디퓨저로 동작 정보를 전송할 수

있고 어플리케이션은 남은 발향 시간, 온습도 데이터등의 변경 사항을 실시간으로 모니터링하여 디퓨저의 상태정보를 수신하고 이를 사용자에게 시각적으로 표시한다. 이와 같이 Firebase의 실시간 데이터베이스 기능을 활용하여 사용자 명령에 대한 스마트 디퓨저의 빠른 응답과 디퓨저 상태에 대한 실시간 모니터링이 가능하다[6].

### 3.2 동작 과정 요약

스마트 디퓨저 시스템은 사용자에게 상황에 맞는 다양한 경험을 제공하기 위해 터치스크린 모드와 어플리케이션 모드의 두 가지 조작 방식을 구현하였다. 두 모드 모두 자동 모드와 수동 모드를 지원하여 사용자가 감정 인식을 통한 맞춤형 향 분사(자동 모드)나 직접 원하는 향과 분사 시간을 설정하는 기능(수동 모드)을 사용할 수 있다.

터치스크린 모드는 사용자가 디퓨저를 직접 조작할 수 있도록 설계되었다. 디퓨저 전면부 터치스크린을 통해 손쉽게 모드 선택과 설정이 가능하며, 즉

각적인 피드백을 받을 수 있어 사용자 편의성을 향상시킨다.

반면, 어플리케이션 모드는 사용자가 외출 중이거나 디퓨저와 물리적으로 떨어져 있을 때에도 디퓨저를 원격으로 제어할 수 있게 한다. 스마트폰 어플리케이션을 통해 원하는 향과 분사 시간을 사전에 설정하고, 원격지에서도 편리하게 디퓨저를 조작할 수 있어 편의성을 높였다.

이러한 두 가지 조작 모드는 사용자의 다양한 상황과 필요에 대응하여 스마트 디퓨저의 활용성을 극대화한다. 실내에 있을 때는 터치스크린을 통해 직관적이고 빠르게 디퓨저를 제어하고, 외부에서는 어플리케이션을 통해 원격으로 디퓨저를 제어함으로써 시스템의 활용가능성을 높인다.

스마트 디퓨저의 전체 동작 흐름은 사용자 관점에서 다음과 같이 진행된다. 그림 3은 동작 플로우 차트를 나타낸다. 먼저, 사용자가 라즈베리파이에 저장된 실행 파일을 실행하면 초기 메인 화면에 터치스크린 모드와 어플리케이션 모드를 선택할 수 있는 버튼이 표시된다.

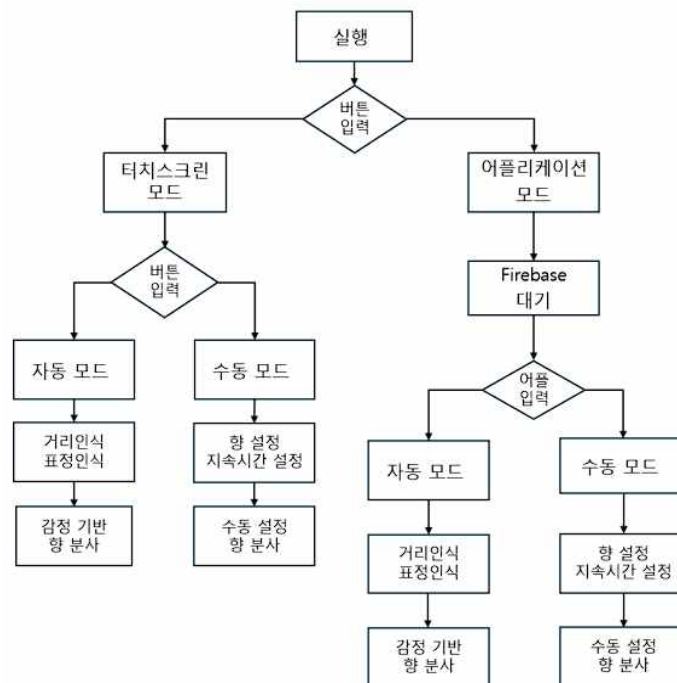


그림 3. 동작 플로우 차트  
Fig 3. Operating flow chart

사용자가 터치스크린 모드 버튼을 선택하면 자동 모드 또는 수동 모드를 선택할 수 있는 창이 나타난다. 자동 모드를 선택하면 초음파 거리 센서를 통해 사용자의 접근을 감지한 후, 카메라 모듈로 표정 인식을 수행하여 인식된 감정과 유사도에 기반한 향이 분사된다. 반면에 수동 모드를 선택하면 터치스크린 화면에서 향의 종류와 지속 시간을 선택하여 원하는 향을 분사한다.

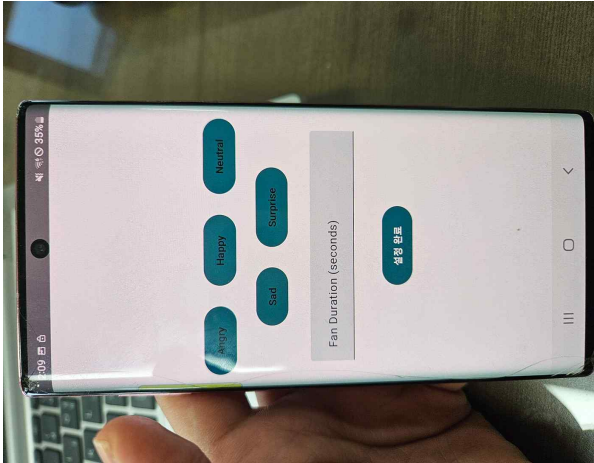


그림 4. 시스템 동작 화면  
Fig. 4. System operating display

사용자가 어플리케이션 모드 버튼을 선택하면 스마트 디퓨저는 파이어베이스 데이터베이스의 mode 값이 설정되기를 대기한다. 이때, 사용자가 어플리케이션을 통해 자동 모드 또는 수동 모드 버튼을 선택하면 mode 값이 "automatic" 또는 "manual"로 업데이트된다. 라즈베리파이에서 실행 중인 프로그램은 업데이트된 mode 값이 "automatic"인 경우 앞서

언급한 자동 모드 과정과 동일한 프로세스를 실행한다. 만약 업데이트된 값이 "manual"인 경우, 어플리케이션을 통해 원하는 향과 지속시간을 파이어베이스 데이터베이스에 업데이트하고 라즈베리파이에서 실행 중인 프로그램은 업데이트한 값을 읽어 설정한 향과 동작시간에 대응하는 향분사 동작을 수행한다.

그림 4는 실제로 구현한 라즈베리파이의 터치스크린 화면과 어플리케이션 화면을 나타낸다.

#### IV. 결 론

본 연구에서는 이전 논문에서의 한계를 극복하고, 감정 인식의 정밀도를 개선한 새로운 시스템을 제안하였다. 기존 연구에서 발생했던 감정 인식의 정확도 문제, 특히 행복과 분노 감정 간의 인식정확도 차이를 보다 개선할 수 있도록 데이터 증강 기법을 활용하였다. 이를 통해 사용자에게 더욱 신뢰할 수 있는 감정 기반 향기 경험을 제공할 수 있게 되었다.

또한, 본 연구에서는 IoT 기술을 활용하여 터치스크린과 어플리케이션을 통한 제어 기능을 추가함으로써 사용자 편의성을 크게 향상시켰다. 이러한 기능은 사용자가 자신의 감정 상태에 맞춘 맞춤형 향기를 선택하고 조절할 수 있는 기회를 제공하여, 향기 기술이 감정 조절에 미치는 긍정적인 영향을 극대화할 수 있도록 한다.

향후 연구에서는 감정 인식 모델의 지속적인 발전과 더불어, 다양한 사용자 환경과 감정 상태를 고려한 향기 조절 알고리즘의 개발이 필요하다. 이를 통해 더욱 풍부하고 개인화된 향기 경험을 제공하고, 궁극적으로 사용자의 삶의 질 향상에 기여할 수 있는 스마트 디퓨저 시스템으로 발전할 것으로 기대된다.

#### 참 고 문 헌

[1] W H Jang, J H Pyo, J S Kim et al. "Implementation of a smart diffuser based on deep learning and user's emotion recognition" Journal of

KIIT. Vol. 22, No. 10, Oct. 31, 2024

- [2] Shorten C, Khoshgoftaar T M, "A survey on image data augmentation for deep learning" Journal of Big Data, Vol. 6, No. 1, pp. 1-48, Jul 2019.
- [3] Wang, Jason, and Luis Perez. "The effectiveness of data augmentation in image classification using deep learning" Convolutional Neural Networks Vis. Recognit Vol. 11, No. 2017, pp. 1-8, 2017.
- [4] P Y Simard, D Steinkraus and J C Platt, "Best practices for convolutional neural networks applied to visual document analysis" Seventh International Conference on Document Analysis and Recognition , pp. 958-963, 2003
- [5] Krizhevsky A., Sutskever I and Hinton G E, "Imagenet classification with deep convolutional neural networks" Communications of the ACM, Vol. 60, No. 6, pp. 84-90, 2017.
- [6] LI, Wu-Jeng,. "JustIoT internet of things based on the firebase real-time database" 2018 IEEE International Conference on Smart Manufacturing, Industrial & Logistics Engineering, pp. 43-47, 2018.

# 자연어 기반 지리 정보 검색: 지식 기반 기법을 통한 쿼리 처리 향상

윤유빈\*, 최권택\*\*

## Natural Language-Based Geographic Information Retrieval: Enhancing Query Processing through Knowledge-Based Techniques

Yubeen Yun\*, and Kwontaeg Choi\*\*

### 요 약

기존 GIS 시스템에서는 CQL을 표준 질의 언어로 사용하여 GIS 데이터를 효과적으로 사용하고 있다. 그러나 CQL은 SQL 문법과 공간 연산을 결합한 형태로 설계되어 있어, 비전문가가 사용하기에는 어려움이 있다. 따라서 본 연구에서는 비전문가도 ChatGPT를 이용하여 GIS 데이터를 보다 쉽게 사용할 수 있는 효과적인 방법을 제시한다.

### Abstract

In existing GIS systems, CQL is used as the standard query language for effectively utilizing GIS data. However, CQL combines SQL syntax with spatial operations, making it challenging for non-experts to use. Therefore, this study proposes an effective method for enabling non-experts to access GIS data more easily through ChatGPT.

### Key words

GIS, CQL, Non-experts, ChatGPT, Query Language

### 1. 서 론

GIS(지리정보 시스템)는 지리 정보를 시각화하고 분석하는 중요한 도구이다. GIS 시스템에서는 다양한 데이터 필터링과 스타일링을 통해 지리적 정보

를 효과적으로 활용할 수 있으며, 이를 위해 주로 사용되는 쿼리언어로 CQL(Common Query Language)이 있다.

CQL은 OGC(Open Geospatial Consortium)에서 지리적 데이터 질의를 위해 개발된 텍스트 기반의 질

\*강남대학교 ICT공학부, ybhs020308@kangnam.ac.kr,

\*\*강남대학교 ICT융합공학부, kwontaegchoi@kangnam.ac.kr(교신저자)

의 언어로, SQL과 유사한 구문과 공간 연산을 결합하여 지리 데이터를 필터링하는 데 사용된다.

CQL을 통해 사용자는 속성 데이터나 지리적 위치를 기준으로 데이터를 필터링하고, 논리 연산자를 사용하여 복합적인 조건을 생성할 수 있다.

이러한 기능은 GIS 데이터에 대해 정교한 질의와 분석을 가능하게 하지만, 복잡한 구문과 공간적 개념은 비전문가가 GIS 데이터를 효과적으로 활용하기 어렵게 만든다. 비전문가를 위한 자연어를 SQL로 변환하는 연구는 이루어졌지만[1], GIS 분야에서의 자연어를 CQL로의 변환 연구는 상대적으로 부족하다. 이에 본 연구에서는 비전문가도 ChatGPT를 활용하여 GIS 데이터를 쉽게 사용할 수 있는 효과적인 방법을 제시함으로써 GIS의 접근성을 높이고, 누구나 지리 정보를 활용할 수 있는 환경을 구축하는 것을 목표로 한다.

## II. 시스템 개요 및 실험

ChatGPT가 지리 정보를 얼마나 잘 이해하고, 자연어를 문법에 맞는 CQL로 변환하며, 신뢰성 있는 지리 정보를 효과적으로 활용할 수 있는지를 실험한다. 사용자가 자연어로 입력한 질의를 ChatGPT가 CQL로 변환하고, GeoServer에서 데이터를 필터링하고 시각화하는 과정을 다룬다. 이 실험을 통해 ChatGPT의 지리 정보 처리 능력과 그 한계를 검토하고, 비전문가도 효과적으로 GIS를 활용할 수 있는 환경을 제공하고자 한다.

제안하는 시스템의 흐름을 그림1에 나타내었다. 이 시스템은 사용자가 자연어로 쿼리를 입력하면, 그 결과가 필터링된 지도로 출력되는 구조를 가진다. 이 전체 흐름은 크게 네 가지 단계로 구성된다.

1. 지도 출력 및 레이어 적용 - 리액트 환경에서 OpenLayers를 사용하여 기본 지도를 출력하고, GeoServer로부터 지도 정보와 관련된 레이어와 메타데이터를 적용한다. 이 과정은 시스템에서 지도를 준비하고 기본 데이터를 불러오는 단계이다.

2. 자연어 쿼리 입력 및 반환 - 사용자가 자연어로 쿼리를 입력한 후 전송하면, Flask 서버에서 ChatGPT가 이를 처리한다. 이 단계에서는 자연어로

작성된 명령어가 JSON형식의 CQL로 변환된다.

3. CQL 전달 및 데이터 필터링 - 변환된 CQL은 Flask서버에서 GeoServer로 전달되어, 해당 쿼리에 따라 데이터를 필터링한다. 이 과정에서 필터링된 지리 데이터가 반환된다.

4. 지도 업데이트 및 시각화 - 마지막으로, 반환된 데이터를 OpenLayers로 다시 전달하여 지도상의 데이터를 업데이트한다. 이 단계에서 필터링된 지리 데이터를 기반으로 최종 시각화된 지도가 사용자에게 출력된다.

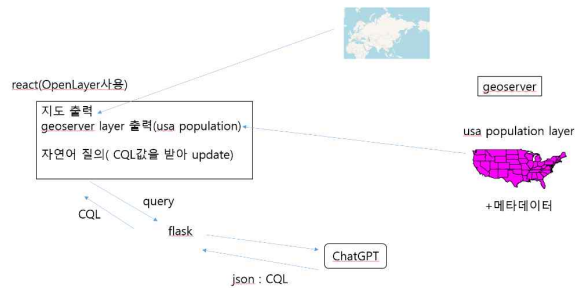


그림 1. 제안하는 시스템의 구조 흐름도 (자연어를 기반으로 한 CQL 생성 및 시각화 과정)

Fig. 1. Structural flow

실험을 통해 ChatGPT의 지리 정보 처리 능력을 평가하였다.

GIS 속성에 어떠한 것이 있는지 잘 알고 있는 사용자가 질의를 "인구 수가 10,000명 이상이며, 남성 인구가 여성 인구보다 많은 지역"으로 주었을 때는 PERSONS >= 10000 AND MALE > FEMALE와 같이 정확한 CQL 문법을 주었는데, 비전문가처럼 질의를 간단하게 "미국 서부 지역"으로 주었을 때는 SUB\_REGION = '서부'와 같이 실행이 안되는 CQL 문법을 주었다. 이를 통해 시스템이 사용자 질의 수준에 영향을 받는다는 점을 발견하였으며 GIS에 대한 이해도가 높은 사용자의 질의는 보다 명확하고 정확한 결과를 도출할 가능성이 높다는 것을 발견했다.

ChatGPT에게 "대중교통 이용 인구가 많은 주 혹은 미국 서부 지역"으로 질의를 주었을 때 결과는 SELECT \* FROM feature WHERE PUBTRANS = (SELECT MAX(PUBTRANS) FROM feature), SUB\_REGION = '서부'와 같이 CQL 문법이 아닌 SQL문법으로 생성되어 GIS에서 사용할 수 없게 되

었다. 이는 ChatGPT가 CQL 대신 SQL을 요구하는 것으로 오해한 결과였다. 이는 ChatGPT가 아직은 지리 정보에 대한 학습량이 충분하지 않음을 보여 준다. 또한, 질의를 "미국에 파이프가이즈(햄버거 가게)가 제일 많은 주", "미국에 파이프가이즈가 없는 주로 주었을 때 STATE\_NAME = 'California'와 STATE\_NAME NOT IN ('Virginia', 'Florida', 'Texas', 'Illinois', 'California', 'New York') 으로 각각 CQL 문법으로 받았으나, 파이프가이즈 매장이 제일 많은 주가 California 주였으나, 파이프가이즈가 없는 주에도 California 주가 있었다. 이는 ChatGPT의 정보 신뢰도가 떨어지는 것을 알 수 있었다.

이를 통해 ChatGPT의 CQL에 대한 이해 수준, 추론 능력, 그리고 사용자 질의를 처리 정확도가 결과에 큰 영향을 미친다는 것을 확인할 수 있었다. ChatGPT는 이런 부분에서 인문사회적인 정보 및 추론 능력은 점차 나아질 것으로 후에 정확도는 더 높아질 것으로 예상된다.

Engineers, 447-479.

### III. 결 론

본 연구에서는 비전문가도 자연어를 통해 복잡한 GIS 작업을 수행할 수 있는 환경을 제시함으로써 GIS의 접근성을 크게 높이고, 비전문가도 GIS 데이터를 효과적으로 활용할 수 있는 가능성을 확인하였다. 향후 연구에서는 프롬프트 엔지니어링을 통해 효과적이고 정확한 CQL을 얻기 위한 가이드라인을 제시하고, 다양한 질의 패턴에 따라 ChatGPT의 응답을 최적화하고, 비전문가가 더 쉽게 GIS 데이터를 활용할 수 있는 방법을 탐구할 것이다. 이를 통해 신뢰성 있고 실용성을 갖춘 GIS 활용 환경을 제공하는 데 기여하고자 한다.

### 참 고 문 헌

- [1] Yoon, H. S., Choi, J. W., Kim, J. S., Heo, J. H., & Kang, P. S. (2021). "SQL Conversion of Korean Natural Language Queries using Multilingual BERT." Proceedings of the Spring Conference of the Korean Institute of Industrial





# Block-Based Stochastic Computing 하드웨어 설계 및 자동화 검증 프레임워크

김태균\*, 서영교\*\*

## Design and Implementation of Block-Based Stochastic Computing and Its Automated Verification Framework

Taegyun Kim\* and Yeongkyo Seo\*\*

### 요 약

본 논문에서는 Stochastic Computing(SC) 기반의 Block-Based Stochastic Computing(BSC) 설계와 이를 검증하기 위한 자동화 프레임워크를 제안한다. 본 프레임워크는 C++로 구현한 BSC 소프트웨어 모델과 Verilog로 작성된 BSC RTL 설계의 결과를 동일한 입력 벡터를 이용해 벡터 매칭 방식으로 검증한다. 이를 통해 BSC 설계의 정확성을 효율적으로 평가하고, 소프트웨어와 하드웨어 간 결과 비교를 통해 BSC의 성능을 분석할 수 있다.

### Abstract

This paper proposes an automated verification framework for Block-Based Stochastic Computing (BSC) design based on Stochastic Computing (SC). The framework verifies the results of a BSC software model implemented in C++ and a BSC RTL design written in Verilog by applying identical input vectors and performing vector matching. This approach enables efficient evaluation of the accuracy of the BSC design and facilitates performance analysis of BSC by comparing results between the software and hardware implementations.

### Key words

Stochastic Computing (SC), Block-Based Stochastic Computing (BSC), Vector Matching, C++ Model, Verilog RTL

---

\* 인하대학교 정보통신공학과 Email: sheonkim01@inha.edu

\*\* 인하대학교 반도체융합전공 Email: yeongkyo@inha.ac.kr (교신저자)

※ 이 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음. (IITP-2021-0-02052)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대지능형반도체기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2021M3F3A2A01037531)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-PIM인공지능반도체핵심기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2022M3I7A2079155)

※ 본 연구는 IDEC에서 EDA Tool를 지원받아 수행하였습니다.

## I. 서 론

최근 컴퓨팅 시스템에서 효율적인 연산 처리가 중요한 과제로 부각됨에 따라, 확률적 연산 (Stochastic Computing, SC)과 블록 기반 확률적 연산 (Block-Based Stochastic Computing, BSC)이 주목받고 있다[1]. SC는 확률적 비트스트림을 사용하여 저전력, 소규모 연산이 가능하나 입력값이 커질수록 정확도가 낮아지는 한계가 있다. 이를 개선한 BSC는 입력값을 블록으로 분할해 계산의 효율성과 정확성을 높여, 저전력 요구가 높은 머신러닝 및 TinyML 등의 실시간 응용에서도 활용 가능하다[1].

본 논문에서는 C++로 구현된 BSC 연산 결과와 Verilog로 구현된 BSC 하드웨어 연산 결과를 자동으로 검증하는 프레임워크를 제안한다. 이를 통해 동일한 입력 벡터를 사용하여 두 시스템의 BSC 출력이 일치하는지 검증하고, BSC 설계의 신뢰성과 효율성을 평가하고자 한다.

## II. 확률적 연산과 블록 기반 확률적 연산

SC는 확률적 비트스트림을 통해 간단한 논리 연산으로 곱셈, 덧셈 등의 연산을 구현할 수 있어, 전력 소모와 하드웨어 비용이 적다는 장점이 있다. 그러나 입력값이 커질수록 정확도가 낮아지며, 다수의 비트가 필요한 경우 효율성이 떨어질 수 있다. 이를 개선한 BSC는 입력을 작은 블록으로 분할하여 연산을 수행함으로써 SC보다 더 높은 정확도와 효율을 제공한다[1]. 긴 비트스트림을 사용해야 하며, 이로 인해 연산 시간이 길어지는 단점이 있다. 이는 SC의 성능을 제한하는 중요한 요인으로 작용한다.

## III. BSC의 C++ 구현 및 Verilog 검증 프레임워크

본 연구에서는 BSC의 연산 알고리즘을 C++로 구현하고, 이를 검증하기 위해 Verilog로 동일한 BSC 알고리즘을 설계하였다. C++로 구현한 BSC는 [1]번 논문에서 제시된 기술을 적용하여, 무작위 입력 벡터와 각 연산의 결과를 텍스트 파일로 저장하여 검증에 사용할 수 있게 하였다. Verilog로 작성

된 BSC 하드웨어 또한 동일한 입력 벡터를 사용하여 계산한 결과를 저장하며, 이 결과를 C++로 작성된 검증 스크립트가 자동으로 비교하여 동일 입력에 대한 출력의 일치 여부를 확인한다. 이를 통해 소프트웨어 상의 BSC 구현이 하드웨어에서도 동일하게 작동하는지를 검증할 수 있다.

## IV. BSC의 구현 방법

Block-based Stochastic Computing(BSC)의 구현을 위해 C++ 및 Verilog를 활용하여 소프트웨어와 하드웨어 수준의 BSC 구조를 각각 설계하였다. 먼저 [1]번 논문의 기술을 바탕으로 C++ 코드를 통해 소프트웨어에서 BSC 동작을 시뮬레이션하고, 이를 Verilog RTL 설계로 옮겨 하드웨어 구현을 진행하였다. 이후 동일한 입력 벡터에 대해 소프트웨어와 하드웨어 간 결과를 비교하여 BSC의 정확성과 일관성을 검증하였다.

C++ 구현은 다음 절차를 따른다. 먼저 BSC에 필요한 벡터를 무작위로 생성하여 계산에 사용한다. 양수 벡터(POS)는 3개의 8비트 벡터로, 음수 벡터(NEG)는 2개의 8비트 벡터로 구성된다. 각 벡터의 비트는 난수로 무작위 설정되며, POS와 NEG 그룹 내 벡터 쌍마다 비트 단위 AND 연산을 수행하여 각 비트 위치에서 일치하는 1의 개수를 산출한다. 산출된 1의 개수는 양수 및 음수 그룹의 1의 총합인  $A_p$ 와  $A_n$ 으로 누적된다. 이후 양수와 음수 벡터의 1의 개수 차이를 계산하여 최종 결과를 결정한다. POS의 1의 개수가 NEG의 1의 개수보다 많으면 양수 비트스트림  $Sop$  이 선택되고, 반대의 경우 음수 비트스트림  $Son$ 이 최종 BSC 결과로 선택된다. 각 연산 후 산출된 BSC 결과와 기대값 간의 오차를 계산하여 BSC 성능을 평가하였다.

Verilog 설계는 C++ 구현의 논리를 하드웨어에 적용하여 소프트웨어 시뮬레이션 결과와 비교 검증이 가능하도록 구성하였다. 주요 모듈은 `bsc_mac`과 `bsc_add`로 구성되며, 각각 BSC 매크로 및 연산 기능을 담당한다. `bsc_mac` 모듈은 POS와 NEG 벡터를 입력으로 받아 각 사이클에서 `valid_out`, `sign`, `out` 신호를 통해 최종 BSC 연산 결과를 출력하며, 내부적으로 POS와 NEG 벡터를 비트 단위로 AND 연산하여 각 블록의 결과를 생성한다.

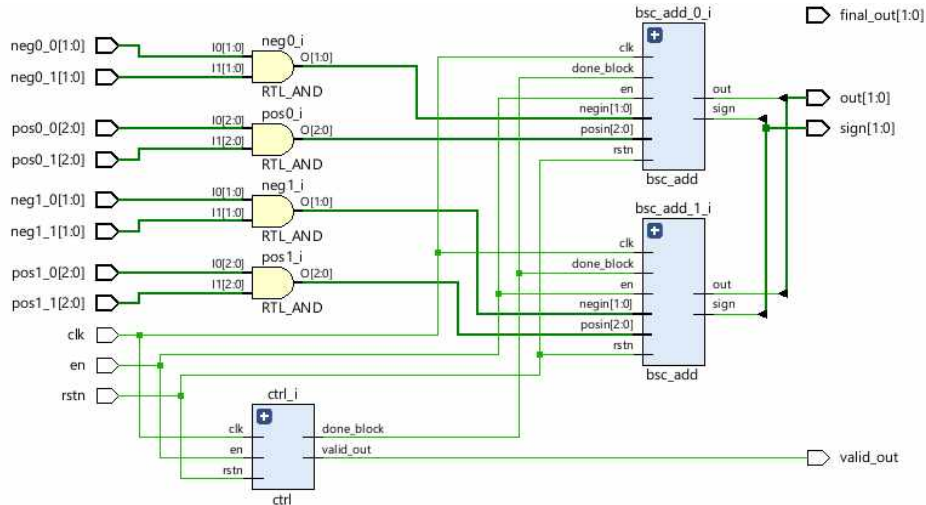


그림 1. Block-based Stochastic Computing(BSC) Schematic  
 Fig. 1. Block-based Stochastic Computing(BSC) Schematic

bsc\_add 모듈은 벡터 내 1의 개수를 누적하는 카운터를 사용해  $A_p$ 와  $A_n$ 을 계산하며,  $A_p - A_n$ 과  $A_n - A_p$ 의 차이를 통해 최종 출력 out을 생성한다. 이를 통해 POS 및 NEG의 비트 차이를 비교하고, 결과 비트스트림을 sign 신호를 기반으로 MUX를 통해 선택한다. 이와 같은 C++ 및 Verilog 구현을 통해 BSC의 계산 성능을 평가하였으며, 각 입력에 대해 소프트웨어와 하드웨어 간 동일한 출력을 산출함으로써 구현의 정확성을 검증하였다.

### V. BSC의 Vector Matching

Block-Based Stochastic Computing(BSC) 결과 검증을 위해 C++로 작성된 BSC 시뮬레이션 코드와 Verilog로 구현된 하드웨어 BSC 모듈의 결과를 비교하는 자동화 검증 프레임워크를 구축하였다. 동일한 입력 벡터에 대해 두 시스템에서 계산된 BSC 결과의 일치 여부를 확인함으로써 구현된 하드웨어의 정확성을 검증하는 방법을 제시한다.

먼저, C++로 구성한 BSC 시뮬레이션은 입력 벡터에 대한 BSC 연산 결과를 파일(BSC\_result.txt)로 저장한다. 이후, 동일한 입력 벡터를 Verilog로 작성된 BSC 모듈에 인가하여 Verilog 기반의 하드웨어 결과(BSC\_hard\_result.txt)도 동일한 형식의 파일로 저장한다. 이를 통해 동일한 입력 벡터에 대해 C++와 Verilog에서 산출한 두 출력을 쉽게 비교할 수

있는 파일 기반의 자동 검증 시스템을 구축하였다. 이 과정에서 C++ 코드로 두 결과 파일을 비교하는 프로그램을 작성하였으며, 이 프로그램은 각 입력 벡터에 대해 두 결과 파일이 일치하는지 확인하고, 불일치가 발생할 경우 해당 입력과 결과를 출력하여 신속한 문제 파악이 가능하도록 하였다. 이와 같은 벡터 매칭을 통해 하드웨어 설계가 시뮬레이션 결과와 동일하게 동작하는지 검증할 수 있으며, 이 과정은 하드웨어 BSC 설계 검증에 있어 중요한 자동화된 검증 수단이 된다.

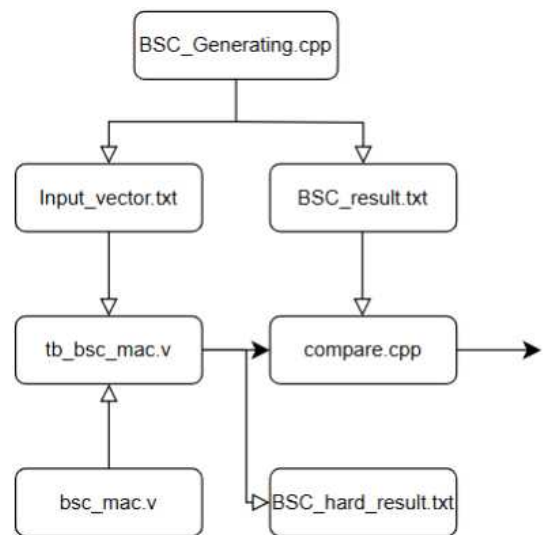


그림 2 C++을 이용한 Vector Matching 순서도  
 Fig. 2. Flow Diagram for Vector Matching using C++

Verilog Simulation으로부터 얻은 출력값을 C++ Simulation으로부터 얻은 출력값과 비교하는 Vector Matching을 수행함으로써 C++으로 구현된 BSC 모듈과 RTL 설계 간 동일성을 검증할 수 있다.

### 참 고 문 헌

[1] Y. Song, E. H. Sha, Q. Zhuge, R. Xu, Y. Zhang, B. Li, and L. Yang, "BSC: Block-based Stochastic Computing to Enable Accurate and Efficient TinyML," arXiv preprint arXiv:2111.06686, 2021.

### VI. 결 론

본 논문에서는 C++으로 구현된 BSC 결과와 Verilog 시뮬레이션 간의 벡터 매칭을 통해, 복잡한 검증 과정 없이 RTL 설계와 C++ 모듈의 동일성을 검증하는 자동화 프레임워크를 제시하였다.

그림 3는 "Input\_vector"를 생성하여 C++ 구현 결과와 Verilog BSC 시뮬레이션 결과에 대해 벡터 매칭을 수행한 결과를 보여준다. C++ 실행 시 입력 데이터의 크기를 인자로 설정할 수 있어, 다양한 크기의 벡터에 대해 유연하게 검증할 수 있다.

```
All bits match in BSC_results.txt and BSC_hard_result.txt.
All bits match in test_vectors.txt and read_vectors.txt.
```

그림 3. C++, Verilog 결과와 Vector Matching 실행 결과  
Fig. 3. Results of Vector Matching between C++ and Verilog Outputs

그림 4는 의도적으로 오류를 삽입하여 벡터 매칭에 실패한 경우를 나타낸다. 오류가 발생했을 때 C++과 Verilog 시뮬레이션 결과 및 그 차이를 표시함으로써, 실패한 줄과 해당 벡터를 신속하게 파악할 수 있도록 설계하였다. 이러한 기능을 통해 의도된 오류가 있는 경우에도 효과적으로 분석할 수 있으며, 오류 위치와 원인을 빠르게 확인할 수 있다.

본 연구에서 제안한 자동화 검증 프레임워크는 BSC 설계의 성능과 오류율을 효율적으로 평가할 수 있는 도구로, 다양한 크기의 입력 벡터에 대해 유연한 검증이 가능하다. 이 프레임워크는 새로운 BSC 설계를 도입할 때도 효과적인 검증 도구로 활용될 수 있으며, 향후 BSC의 성능 개선을 위한 기반을 제공하는 데 기여할 것으로 기대된다.

```
Difference line 2533: 11 != 10
Difference line 2534: 01 != 00
Difference line 2535: 11 != 10
```

그림 4 의도한 오류로 Vector Matching에 실패한 경우  
Fig. 4. Vector Matching Failure Case due to Intentional Error

# 2T FeFET을 이용한 Content Addressable Memory 설계

노규민\*, 서영교\*\*

## Design of 2T FeFET based Content Addressable Memory

Gyumin Noh\*, and Yeongkyo Seo\*\*

### 요약

강유전체를 활용한 Ferroelectric-FET(FeFET)은 고속 동작과 비휘발성의 특징을 가진다. FeFET의 Threshold Voltage 변화 특성을 이용하여 Content Addressable Memory(CAM)을 설계하고, 이렇게 FeFET을 통하여 CAM을 설계하는 것이 빠른 동작을 가질 수 있고 신뢰성 또한 높다는 것을 이 논문을 통하여서 입증한다.

### Abstract

Ferroelectric-FET(FeFET) using ferroelectric is characterized by high-speed operation and non-volatility. Designing a Content Addressable Memory(CAM) using the threshold voltage change characteristics of FeFET and designing CAM through FeFET in this way proves that it can have fast operation and is also reliable through this work.

### Key words

Ferroelectric-FET(FeFET), Content Addressable Memory(CAM), High-speed, reliability, Parallel operation

### 1. 서론

Content Addressable Memory(CAM)는 메모리 블록의 모든 항목에 대해 병렬로 검색을 수행하는 특수한 형태의 In Memory Computing 회로이다. 이러한 특징 때문에 router, associative memory, 머신러닝의 하드웨어 및 고성능 데이터베이스 등에 많이 사용된다. CAM의 경우는 2진법에 따라 0과 1로 상태를 구분하게 된다.

FeFET은 기존 MOSFET과는 다르게 강유전체를 사용하여 FET의 Threshold voltage를 조절할 수 있다. 전계를 인가하여 이러한 Threshold Voltage를 High  $V_{th}$  또는 Low  $V_{th}$  상태로 조절할 수 있고 이를 Program이라고 한다. FeFET은 Non-volatile 특성을 지니게 된다.

이러한 FeFET을 가지고 Content-Addressable Memory(CAM)을 설계하게 되면 빠른 동작과 Non-volatile 특성을 지니게 되고 2개의 Transistor를 이용하다 보니 적은 Transistor를 가지고 면적을 보

\* 인하대학교 전자공학과 nohyumin1@inha.edu

\*\* 인하대학교 반도체융합전공 yeongkyo@inha.ac.kr (교신저자)

※ 사사문구는 공간 부족으로 뒷 장에 첨부하였음.

다 효율적으로 활용할 수 있게 된다.

## II. 메모리 Cell 구조 및 회로 구현

이 논문에서는 Memory Cell을 그림 1과 같이 구현하였다. 강유전체의 분극을 활용하여 Gate에 임계값 이상의 큰 전계를 걸어주게 되면 분극을 만들어 주며 Program 한다. SL과 SLB의 전압 공급을 통하여 Program을 한다. 이러한 Program을 통해 해당 FeFET을 High Vth 또는 Low Vth상태 두 가지로 만들 수가 있고 Program 된 Cell은 전원이 꺼져도 그 상태를 유지하게 된다.

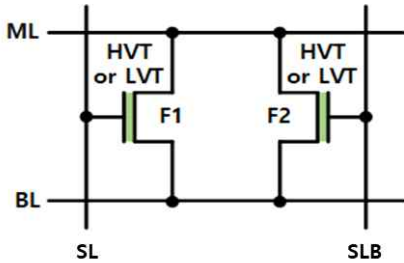


그림 1. FeRAM 셀 구조  
Fig. 1. FeRAM cell

X-peri의 경우는 SL의 전압을 Control 하는 SL Read Driver 와 Write Driver로 구성을 하였다.

Y-Peri의 경우는 ML과 BL의 Write Driver 그리고 ML의 Pre-charger 그리고 ML의 floating 상태와 Discharging 상태를 판단해주는 INV - NAND-INV 구조의 회로가 사용되었다. (그림2)

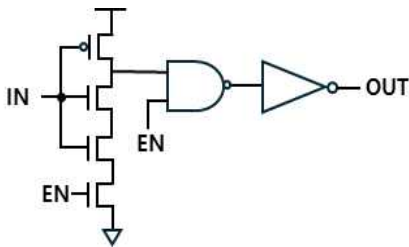


그림 2. ML 판정 회로  
Fig. 2. ML decision Circuit

그림2의 회로는 ML이 INPUT으로 들어오게 되고 Enable 신호에 맞게 켜지고 꺼진다. 16개의 row로

구성하였기에 16개의 회로가 동시에 동작하게 되고 CAM의 특성상 주소를 중복으로 가지고 있는 것이 불가능하니 15개 혹은 16개의 ML이 Discharging 된다. 그림2의 회로는 이러한 특징을 잘 반영하여 Switching Current와 Short circuit Current를 Switching 횟수와 NAND Gate를 통하여 최대한 제어한다는 장점을 가지게 된다.

앞에 기술한 Memory Cell과 Peri를 가지고 16x64 Memory Array를 구성하였다. 각각의 row는 64개 Cell들의 High Vth / Low Vth Program 상태에 따라서 값을 가지게 된다. Cell의 F1이 Low Vth, F2가 High Vth 상태이면 0이라는 값을 가지게 되고 반대로 F2가 Low Vth, F1이 High Vth가 되면 1이라는 값을 가진다.

SL이 0이고 SLB가 1인 경우를 Weight가 0인 경우 이고 반대로 SL이 0이고 SLB가 1인 경우를 Weight 1로 놓는다.

이렇게 한 ML이 64bit의 Data를 가지게 되고 SL의 입력 Data 즉 Weight가 64개의 Cell Data와 동일하면 Match, 반대로 하나라도 Mismatch 하면 ML이 Discharging되게 되는 특성을 가지게 되고 병렬적으로 16개의 ML에 대한 연산을 수행함으로써 빠른 동작이 가능하고 Match되는 row의 주소를 알 수가 있다.

## III. Match/Mismatch Simulation 결과

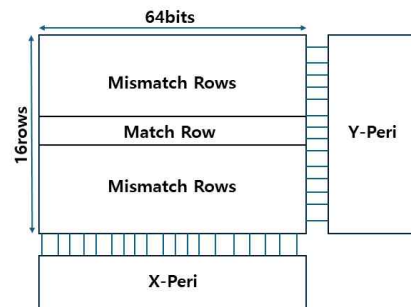


그림 3. CAM 구조  
Fig. 3. Structure of CAM

본 논문에서는 그림 3와 같이 64개의 Cell에 대하여 0개 가 Mismatch 될 때 즉 Match 상태, 그림

고 1개가 Mismatch 될 때 16개가 Mismatch 될 때, 64개 전부 Mismatch 될 때 이렇게 4가지의 Case에 대해서 Simulation을 보이겠다. 1개가 Mismatch 되었을 경우도 좋은 판정을 보여야 CAM으로써 정확한 동작을 증명할 수 있다.

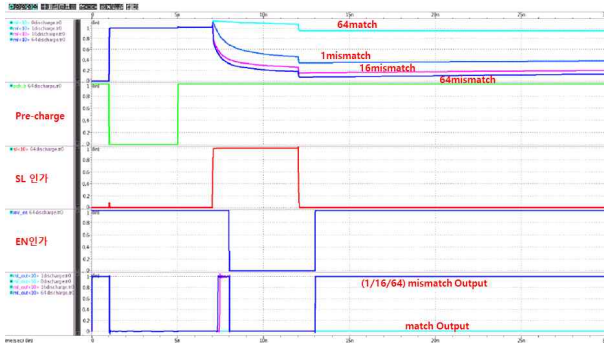


그림 4. 시뮬레이션 결과  
Fig. 4. Result of Simulation

Simulation 결과는 그림4와 같이 나왔다. 64 Cell 모두 Floating의 경우 즉 Match Case의 경우 ML이 유지되고 마지막 output 결과 또한 Enable 신호가 끝나고 0으로 유지된다. 그러나 1개~64개의 Cell이 discharging한 경우 즉 discharging Cell 수 만큼의 Cell이 mismatch된 경우는 ML이 discharging되고 Output 결과 또한 1로 유지가 되는 것을 알 수가 있다.

위의 Simulation 결과는 해당 2 Transistor FeFET Cell을 바탕으로 만든 CAM이 정상 동작을 하며 Enable 신호 또한 SL인가로부터 1ns만 되어도 판정할 수 있을 만큼 고속 동작을 시사한다. 또한 이러한 고속 동작에서 단 한 개의 Cell만 discharging 되어도 mismatch임을 구별 할 수 있는 CAM임을 입증하였다.

### 참 고 문 헌

[1] Bekdache, Omar, et al. "Scalable complementary FeFET CAM design." 2023 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2023.  
[2] Cai, Jiahao, et al. "Energy efficient data search

design and optimization based on a compact ferroelectric fet content addressable memory." Proceedings of the 59th ACM/IEEE Design Automation Conference. 2022.

[3] K. P. Gnawali, S. N. Mozaffari, and S. Tragoudas, "Low power spintronic ternary content addressable memory," IEEE Transactions on Nanotechnology, vol. 17, no. 6, pp. 1206-1216, 2018.  
[4] X. Yin, C. Li, Q. Huang, L. Zhang, M. Niemier, X. S. Hu, C. Zhuo, and K. Ni, "Fecam: A universal compact digital and analog content addressable memory using ferroelectric," IEEE Transactions on Electron Devices, vol. 67, no. 7, pp. 2785-2792, 2020.

### 사 사 문 구

※ 이 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음. (IITP-2021-0-02052)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대지능형반도체기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2021M3F3A2A01037531)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-PIM인공지능반도체핵심기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2022M3I7A2079155)

※ 이 연구는 IDEC에서 EDA Tool를 지원받아 수행하였습니다.

# 구미시 트램 도입의 기대효과 및 운용 비용 분석

배정석\*, 최윤지\*, 변혜정\*, 이희진\*\*

## Analysis of expected effects and operating costs of tram introduction in Gumi City

Jeong-Seok Bae\*, Yun-Ji Choi\*, Hye-jeong Byun\*, and Hee-jin Lee\*\*

### 요약

본 논문에서는 구미시에 트램을 도입 할 경우 트램 한 대를 운용하는 데에 드는 전기 요금을 산정한다. 이는 구미시 트램 도입에 투입되는 경제적 규모를 예측하여 구미시 트램 도입의 실재화 및 교통 인프라 발전에 기여할 것을 기대한다.

### Abstract

In this paper, when trams are introduced in Gumi City, the electricity cost for operating one tram is calculated. This is expected to contribute to the realization of the introduction of trams in Gumi City and the development of transportation infrastructure by predicting the economic scale invested in the introduction of trams in Gumi City.

### Key words

Tram, GUMI, Electric charges

## I. 서론

트램은 건설 및 운영 비용이 도시 철도에 비해 저렴하고, 도시의 이동성을 향상시키는 친환경적 대중교통이다.

본 논문에서는 구미시에서 트램을 운행 할 경우의 운용 비용을 구미시 시내 버스 운영 시간을 참조하여 계절별, 시간대별로 책정하였다. 한국전력공사가 제공하는 전기 요금표를 참고하였으며, 스웨덴 예테보리에서 운행 중인 M31 모델을 본 연구의 트

램 모델로 선정하였다.

이는 구미시에 트램을 도입 할 경우 운용에 투입되는 재화를 추정하여 실제 트램 도입 시 요구되는 경제적 규모를 예측할 수 있게 한다.

## II. 트램 모델 선정

본 연구에서는 구미시에 도입할 트램 모델로 스웨덴 예테보리에서 운행 중인 M31 트램을 선정하였다. M31 트램의 사양은 표[1]와 같다.

\* 국립금오공과대학교 전자공학부

\*\* 국립금오공과대학교 전자공학부 교수, 교신저자 email: jinlee@kumoh.ac.kr



표 1. M31 트램 사양

Table 1. M31 Tram Specifications

항목	사양
차량 모델	M31
차량 유형	노면전차
길이	30.7m
좌석 수	81개
입석 수	109개
무게	34.5톤
구동 방식	600~750V, 직류(DC)
최대 출력	300kW
최고 속도	70km/h

M31 트램은 600-750V의 직류 전원으로 구동되는 전기식 트램으로, 이는 현대 도시형 전철 시스템에서 널리 채택되는 전압 범위로 안정적인 전력 공급이 가능하다. 차량은 최대 300kW의 출력을 가지고, 최고 속도 70km/h에 도달할 수 있다. 또한, 트램의 최대 수용 인원이 190명이므로 도심 내 빠르고 효율적인 대중교통 수단으로 적합하다.

### III. 트램 운용비용

본 논문에서의 트램 운행 시간은 구미시 시내버스 운행 시간을 고려한 06시 ~ 23시 30분으로 설정하고, 출근 시간 07시 ~ 09시 및 퇴근시간 17시 ~ 19시를 RH로, 그 외의 시간을 평시로 구분한다. 또한 트램은 한국전력으로부터 3상 22.9kV를 수전 받아 변압기 및 정류기를 거친 뒤 직류로 바꾸어 공급받는 직류 급전시스템을 사용한다[1].

2024년 7월 1일 기준으로 시행된 한국전력공사의 전기요금표를 참조하여, 트램의 전력 요금은 고압 A(표준전압 3,300V, 66,000V 이하), 선택Ⅲ(전기사용시간이 월 500시간 초과), 산업용전력(을)의 요금제인 표[2]에 부합한다.

표[2], 표[3]을 참고한 계절별 트램 한 대의 운용비용은 여름철에는 84,316,950원, 봄·가을철에는 108,353,250원, 겨울철에는 108,388,800원으로 산정된다.

표 2. 산업용전력(을) 전기요금표

Table 2. Industrial power (b) electricity rate table

구분	기본 요금(1 0원/k W)	전력량요금(원/kWh)				
		시간대 (부하)	여름철 (6-8월)	봄·가을철 (3-5월, 9-10월)	겨울철 (11-2월)	
고압 A	선택 Ⅰ	722	경	99.5	99.5	106.5
			중간	152.4	122.0	152.6
			최대	234.5	152.7	210.1
	선택 Ⅱ	832	경	94.0	94.0	101.0
			중간	146.9	116.5	147.1
			최대	229.0	147.2	204.6
	선택 Ⅲ	981	경	93.1	93.1	100.4
			중간	146.3	115.2	146.5
			최대	216.6	138.9	193.4

표 3. 계절별, 시간대별 구분

Table 3. Classification by season, time zone

계절 시간대	여름철 (6-8월)	봄·가을철 (3-5월, 9-10월)	겨울철 (11-2월)
경부하	22:00-08:00	22:00-08:00	22:00-08:00
중간부 하	08:00-11:00	08:00-11:00	08:00-09:00
	12:00-13:00	12:00-13:00	12:00-16:00
	18:00-22:00	18:00-22:00	19:00-22:00
최대부 하	11:00-12:00	11:00-12:00	09:00-12:00
	13:00-18:00	13:00-18:00	16:00-19:00

### IV. 결 론

구미시에 트램을 도입 할 경우 트램 한 대를 이용하는 데에 여름철에는 84,316,950원, 봄·가을철에는 108,353,250원, 겨울철에는 108,388,800원의 전기 요금이 책정된다. 본 연구 결과가 구미시 트램 도입의 발판이 될 것을 기대한다.

### 참 고 문 헌

- [1] 김대원, 김민규, 최은서, 이희진, “전력공급 효율성을 고려한 구미시 트램 도입 방안 연구”, 한국정보기술학회, pp 103-112, 2024

# 컴퓨터 게임 환경에서 양방향 게이트형 순환 유닛 신경망을 이용한 뇌전도 기반 스트레스 인식

장윤원\*, 응웬부이응옥한\*\*, 김형국\*\*\*

## EEG Signal-Based Stress Recognition Using Bidirectional Gated Recurrent Unit in Computer Game Environment

Yun-Won Jang\*, Nguyen Bui Ngoc Han\*\*, and Hyoung-Gook Kim\*\*\*

### 요 약

본 논문에서는 컴퓨터 게임 환경에서 사용자의 스트레스를 인식하는 방식을 소개한다. 구현된 방식에서는 웨어러블 디바이스를 통하여 게임을 플레이하는 동안 발생하는 뇌전도 신호를 실시간으로 수집하고, 양방향 게이트 순환 유닛 신경망을 이용하여 뇌전도 신호로부터 스트레스를 인식한다. 실험 결과, 구현된 모델은 98.11%의 높은 인식 정확도를 달성하였다. 이 연구의 결과는 엔터테인먼트를 통해 사람들의 건강을 개선하는 앱, 게임 및 알고리즘을 개발하는 데 기여할 수 있다.

### Abstract

In this paper, we introduce a method to recognize user stress in computer game environments. In the implemented method, EEG signals generated while playing a game are collected in real time through a wearable device, and stress is recognized from the EEG signals using a bidirectional gated recurrent unit neural network. As a result of the experiment, the implemented model achieved a high recognition accuracy of 98.11%. The results of this study can contribute to the development of apps, games, and algorithms that improve people's health through entertainment.

### Key words

Electroencephalogram (EEG) signal, stress recognition, bidirectional gated recurrent unit(BGRU)

### 1. 서 론

뇌전도(Electroencephalogram, EEG) 신호는 뇌의 활동을 실시간으로 반영하는 생리적 신호로, 이는

인간의 무의식적 반응을 포착할 수 있어 주관의 개입과 속임수로 인한 오류를 최소화한다는 장점이 있다. EEG를 이용한 기존의 감정 인식 연구에서는 주로 동영상 시청, 음악 청취 등과 같은 시청각 자

\* 광운대학교, yunwonjang@kw.ac.kr

\*\* 광운대학교, nbngochan99@kw.ac.kr

\*\*\* 광운대학교, hkim@kw.ac.kr (교신저자)

※ 본 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2023R1A2C1006756).

극이 사용되어왔다. 그러나 최근의 연구들로부터 VR 환경에서 얻어진 생체 신호로 더욱 효과적인 감정 인식이 가능하고 인지 행동 치료에 3D 게임을 접목했을 때 기존의 방식보다 정신 건강 개선에 탁월한 효과를 보인다는 연구 결과가 도출되고 있다 [1,2].

본 연구에서는 컴퓨터 게임 중 측정된 플레이어의 뇌전도 신호를 양방향 게이트 순환 유닛(BGRU) 신경망에 적용하여 모델링한 후, 이를 러셀의 감정 원형모델을 기반으로 stress와 non- stress 두 가지 상태를 인식한다. 이를 통해 게임 환경에서 BGRU를 이용한 스트레스 인식 시스템의 효과를 확인하고, 게임이 스트레스 관리를 위한 디지털 치료제로 활용될 가능성에 대해 탐구하고자 한다.

## II. 제안된 방식

그림 1은 뇌파 기반 스트레스 인식 시스템의 전체 구조도를 나타낸다. 사용자가 컴퓨터 게임을 플레이하는 동안 발생하는 EEG 신호를 실시간으로 수집하고, 이를 BGRU 신경망 모델을 통해 학습을 수행한 후, 입력된 플레이어의 뇌전도 신호를 학습된 BGRU 모델을 통해 stress 상태와 차분함을 나타내는 non-stress 상태 두 가지 클래스로 인식한다.

구현한 EEG 신호 기반 스트레스 인식 BGRU 모델은 BGRU, batch normalization, dropout 그리고 dense layer 총 네 가지 핵심 레이어로 구성된다.

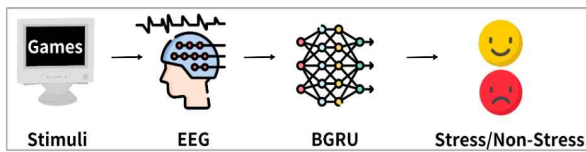


그림 1. BGRU 기반 뇌파 스트레스 인식 시스템 구조도  
Fig. 1. BGRU-EEG system structure

그림 2는 스트레스 인식 시스템에 사용된 BGRU의 구조도를 나타낸다. 이 모델은 EEG 데이터를 양방향으로 처리하며, 순방향 GRU는 시간 순서대로 뇌전도 입력 데이터  $x_t$ 를 분석하고 역방향 GRU는 미래에서 과거  $x_{t+1}, x_{t-1}$ 까지의 입력을 파악한다. 또한 각 단계에서는 현재의 입력과 이전 상태를 기반으로 은닉 상태를 업데이트하여, 시간에 따른 데

이터의 변화 양상을 파악하고, 이로부터 스트레스와 관련된 특징 패턴을 추출한다.

모델의 훈련은 Google colab의 T4 GPU 환경을 사용하여 이루어졌으며, batch size는 256으로 300회의 epoch에 걸쳐 성능을 최적화하였다. 학습 과정에서는 교차 엔트로피 손실과 Adam 최적화 도구가 사용되었으며 학습률은 0.001%로 설정되었다.

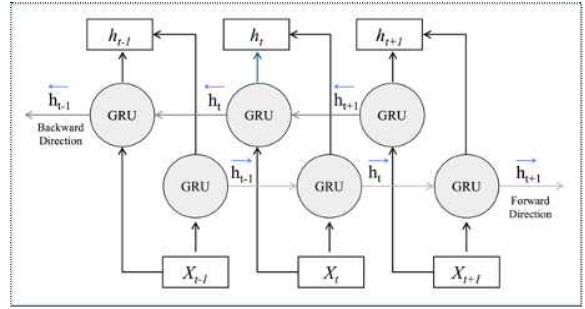


그림 2. BGRU 모델의 구조.  $x_t$ 와  $h_t$ 는 각각 BGRU의 입력 및 출력 임베딩이다.

Fig. 2. BGRU network structure.  $x_t$  and  $h_t$  is input and output embeddings of BGRU, respectively.

## III. 실험 및 결과

연구에 사용된 데이터셋은 차분함과 스트레스를 유발하는 게임을 플레이하는 6명의 참가자로부터 직접 EEG 데이터를 수집하였다. 데이터셋에 대한 구체적인 정보는 표 1에 제시되어 있다.

또한 연구에는 총 네 가지의 게임이 사용되었으며, 차분함을 유발하기 위해서는 하늘을 비행하고 자연을 탐색하는 Sky: Children of the Light와 티베트 신화를 배경으로 신비롭고 아름다운 그래픽을 자랑하는 MANDAGON을 선정하였으며, 반면에 스트레스를 유발하기 위해서는 버려진 장난감 공장에서 괴물들을 피해 살아남아야 하는 Poppy Playtime과 가정집에서 일어나는 공포스러운 현상들을 다룬 [Nightmare Files] Clap Clap을 선정하였다.

표 2에 나타난 실험 결과에 따르면, BGRU 기반 스트레스 인식 모델은 98.11%의 높은 스트레스 인식 정확도를 제시한다. 그리고 그림 3의 혼동행렬표에는 인식된 결과에 대한 True Positive(8381), True Negative (8326), False Positive(188), False Negative (133) 상세한 내용이 나타나있다.

표 1. 실험 데이터셋에 대한 개요

Table 1. Overview of our dataset

구분	설명
데이터 유형	EEG 신호
수집 방식	직접
피험자의 수	20대 남성 6명
EEG 녹화 장치	EPOC X - 14 Channel Wireless EEG Headset
EEG 녹화 시간	60초
자극	네 가지 게임 차분함 유발: Sky: Children of the Light MANDAGON 스트레스 유발: Poppy Playtime [Nightmare Files] Clap Clap
학습 샘플	13,620(80%)
테스트 샘플	3408(20%)

표 2. BGRU-EEG 모델의 테스트 정확도

Table 2. Test accuracy of BGRU-EEG model

구현된 시스템	테스트 정확도
BGRU-EEG	98.11%

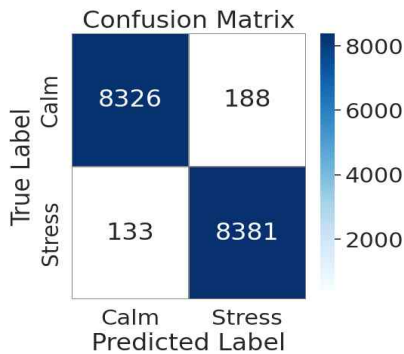


그림 3. BGRU-EEG 모델에 대한 혼동행렬표

Fig. 3. Confusion matrix of BGRU-EEG model

이를 바탕으로 모델의 정확도(98.11%), 정밀도(97.81%), 재현율(98.44%), F1 점수(98.12%)를 각각 측정된 결과, 모델이 지닌 우수한 인식 성능을 확인할 수 있었다.

#### IV. 결 론

본 논문에서는 BGRU 기반 스트레스 인식 시스

템을 통해 게임을 활용한 스트레스 관리 가능성을 탐구하였으며, 이 시스템은 EEG 신호 분류에서 높은 성능을 보여주었다. 향후 연구에서는 앱, 게임, 알고리즘 등 정신 건강을 위한 디지털 치료제 개발에 기여하는 것을 목표로 보다 많은 게임과 더 넓은 범위의 감정으로 이 시스템을 확장하고자 한다.

#### 감사의 글

본 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF- 2023R1A2C1006756).

#### 참 고 문 헌

- [1] J. Nam, H. Chung, and H. Lee, "A new terrain in HCI: Emotion recognition interface using biometric data for an immersive VR experience," arXiv preprint arXiv:1912.01177, Dec. 2019.
- [2] Merry, Sally N., et al. "The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial." Bmj 344, 2012.
- [3] J. X. Chen, D. M. Jiang, and Y. N. Zhang, "A Hierarchical Bidirectional GRU Model with Attention for EEG-Based Emotion Classification," IEEE Access, Vol.7, pp. 118530-118540, 2019.

# 에너지 효율적인 온칩 데이터버스 반전 기법을 위한 다수 근사치 연산 정확도를 높인 인코더 설계

이래인\*, 서영교\*\*

## Accuracy Enhanced Approximate Majority Encoder Design for Energy-Efficient On-Chip Data Bus Inversion Technique

Raein Lee\*, Yeongkyo Seo\*\*

### 요 약

DBI( Data Bus Inversion )은 스위칭 되는 데이터 수가 과반수를 넘으면 데이터를 반전시켜 데이터의 스위칭 수를 줄이는 데이터 인코딩 기법이다. DBI 인코더에는 스위칭하는 데이터를 카운트하기 위한 과반수-판정 회로가 존재하는데 이는 복잡성이 매우 높다. 오프칩 데이터 버스에서는 스위칭 비용이 매우 크기 때문에 복잡성이 매우 높은 회로를 사용하더라도 DBI에 의한 스위칭 절감 효과가 더 지배적이다. 하지만 오프칩과 비교하여 상대적으로 인터커넥트 커패시턴스 값이 작은 온칩 데이터 버스에서는 복잡성이 높은 회로를 사용하게 되면 DBI에 의한 스위칭 절감 효과보다 인코더 파워가 더 지배적이게 된다. 이러한 인코더 에너지 오버헤드때문에 온칩에서는 DBI가 적합하지 않다. 이 논문에서는 온칩 데이터 버스에 최적화된 DBI를 제시한다. 회로의 복잡성을 낮추기 위해 근사-과반수 판정 회로를 적용하였으며 근사 계산에 의해 스위칭 활동이 약간 증가하더라도 온칩 데이터 버스에 적합한 저지연, 저면적, 저전력을 달성을 목표로 한다.

### Abstract

Data Bus Inversion (DBI) is a data encoding technique that inverts data when the number of switching bits exceeds the majority, thereby reducing overall data switching activity. DBI encoders incorporate a majority-determination circuit that counts the switching bits, which can be highly complex. In off-chip data buses, where the cost of switching is significant, the reduction in switching achieved by DBI justifies using this complex circuit. However, for on-chip data buses, which generally have lower interconnect capacitance, the power consumed by the encoder becomes more significant, making conventional DBI less suitable. This paper proposes an optimized DBI for on-chip data buses. To reduce circuit complexity, we apply an approximate majority-determination circuit, aiming to achieve low latency, compact area, and low power, which are well-suited to on-chip data buses, even if there is a slight trade-off in switching activity due to the approximate computation.

### Key words

Data bus inversion, switching , encoding, majority, on-chip bus, approximation

---

\* 인하대학교 정보통신공학과 Email: rain2200@inha.edu

\*\* 인하대학교 반도체융합전공 Email: yeongkyo@inha.ac.kr (교신저자)

※ 사사문구는 공간 부족으로 뒷 장에 첨부하였음.

## 1. 서 론

시스템에서 데이터의 전송은 인터커넥트 커패시턴스의 충전 및 방전에 의해 많은 에너지가 소모된다. 이를 해결하기 위한 방법으로 DBI(Data Bus Inversion)가 채택되었다. DBI는 스위칭 되는 데이터의 수가 과반수를 넘으면 데이터를 반전시켜 데이터의 스위칭 수를 줄이는 데이터 인코딩 기법이다.

DBI 인코더의 주요 구성 요소로는 스위칭 수를 세를 과반수 판정 회로가 존재한다. 기존의 DBI에 사용되는 과반수 판정 회로는 매우 높은 파워가 소모되며 높은 지연시간과 큰 면적을 갖는다. 오프칩 데이터 버스는 스위칭 비용이 매우 크기 때문에 복잡도가 높은 과반수 판정 회로를 쓰더라도 DBI에 의한 스위칭 감소 효과가 더 지배적이다. 하지만 오프칩 데이터 버스와 비교하여 온칩 데이터 버스는 상대적으로 인터커넥트 커패시턴스 값이 매우 작기 때문에 스위칭 절감 효과보다 과반수 판정회로 의한 인코더 파워가 지배적인 에너지 오버헤드가 발생한다.

최근 온칩 데이터 버스에 따른 파워 소모가 증가함에 따라 온칩에 최적화된 DBI 연구가 필요할 것으로 예상된다. 이 연구에서는 에너지 오버헤드를 해결하기 위한 근사-과반수 판정 회로를 제안한다. 과반수 판정회로에 근사계산을 적용하여 에너지 오버헤드를 줄인 온칩 데이터 버스에 적합한 DBI를 제안한다.

## II. 데이터 버스 반전 기법 ( DBI )

그림 1에서와 같이 DBI는 인코더와 디코더로 이루어져 있다. 인코더에서는 이전데이터와 현재데이터의 스위칭 횟수를 카운트하여 스위칭 횟수가 절반을 넘으면 inversion flag을 1로 세워 현재 데이터를 반전시키고 0으로 세워진다면 현재 데이터를 유지시킨다. 인코딩된 데이터는 버스를 통해 송신부인 디코더로 데이터가 전송된다. 디코더에서는 전송받은 데이터를 inversion flag를 기반으로 다시 복원한다.

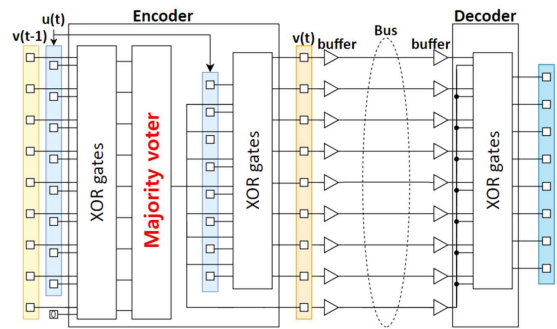


그림 1. Encoder와 Decoder로 이루어진 DBI 기본 구조

기존의 DBI 인코더에는 그림2의 논리 합성을 사용한 과반수 판정 회로(MAJ-SYN)를 사용하며 이는 복잡성이 매우 높다. 오프칩 데이터 버스는 스위칭 비용이 매우 높기 때문에 복잡성이 매우 높은 회로를 사용하더라도 DBI 기법에 의해 감소한 스위칭 파워 효과가 더 지배적이다. 하지만 오프칩 데이터 버스와 비교하여 상대적으로 인터커넥트 커패시턴스가 작은 온칩 데이터 버스에서는 기존의 DBI를 사용하게 되면 스위칭 감소 효과보다 복잡한 인코더에 의한 파워 소비가 더 지배적이게 된다.

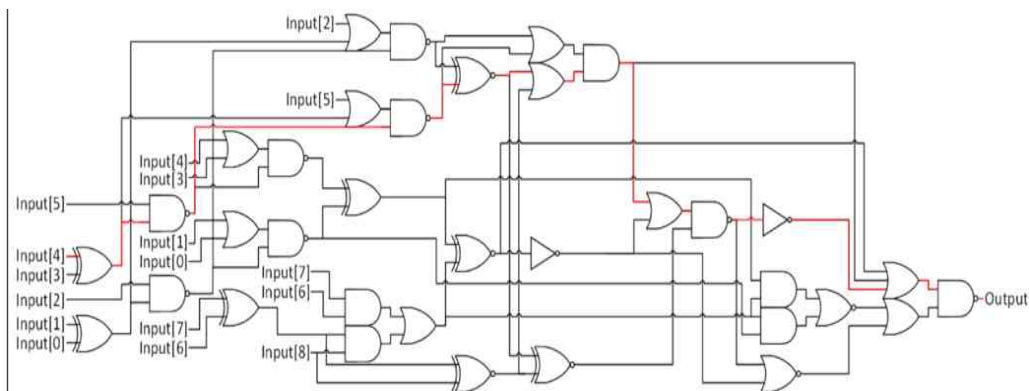


그림 2. 논리 합성을 사용한 9비트 과반수-판정 회로(MAJ-SYN)

따라서 기존의 DBI는 온칩 데이터 버스에 적합하지 않다. 온칩 데이터 버스에 적합하기 위해서는 인코더 회로에 의한 파워를 줄여 에너지 오버헤드를 해결해야 한다.

### III. 근사-과반수 판정 회로

온칩 데이터 버스에 최적화된 과반수 판정 회로를 찾기 위해 근사 방식을 도입하였다. 그림2(a)는 기존의 근사-과반수 판정 회로 AOI-AOI를 보여준다. AOI-AOI 구조를 사용하여 과반수 판정을 근사적으로 연산한다. AOI 두 개의 레벨로 회로를 설계하여 매우 낮은 Latency를 달성하였고 전력과 면적을 줄일 수 있다. 하지만 AOI-AOI 근사-과반수 회로는 정확도가 매우 낮은 문제점이 있다. 정확도가 낮을수록 잘못된 반전 플래그를 세우기 때문에 평균적인 스위칭 수가 증가 하게 된다. 따라서 이 연구에서는 AOI-AOI와 구조에서 NAND gate와 NOR gate만을 사용하여 가능한 모든 경우의 수 중 제일 높은 정확도를 갖는 모델을 모색한다.

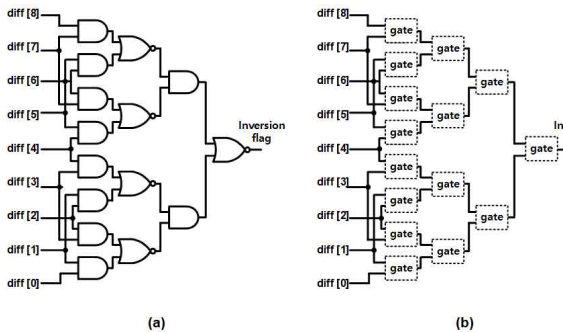


그림 2. (a) 근사 과반수 회수 AOI-AOI  
(b) AOI-AOI 구조

그림 2(b)는 AOI-AOI의 구조를 보여주며 총 15개의 gate가 사용된다. 15개의 gate에 NAND, NOR만 사용하여 가능한  $2^{15}$  개의 경우의 수 중 정확도가 제일 높은 회로를 찾는다. 그 결과 그림4 (a), (b)에서와 같이 15개의 gate가 모두 NAND gate 또는 NOR gate로만 이루어진 과반수 판정회로가 제일 높은 정확도를 보였다. 그림3 (a)는 NAND gate로만 이루어진 F0를 보여주며 그림3(b)는 NOR gate로만

이루어진 F1을 보여준다. 근사-과반수 판정회로 AOI-AOI는 9비트 입력의 가능한 경우의 수 512개에서 111개의 잘못된 판정을 하였으며 스위칭 활동을 10.39% 감소시킨다. 정확한 판정회로를 사용하는 DBI의 스위칭 감소 비율이 18.25%임을 감안하여 AOI-AOI는 낮은 감소비율을 갖는다. 그에 반면 F0,F1모델은 9비트 입력에서 가능한 경우 512개중 93개의 잘못된 판정을 하며 이는 스위칭 활동을 12.75% 감소시킨다.

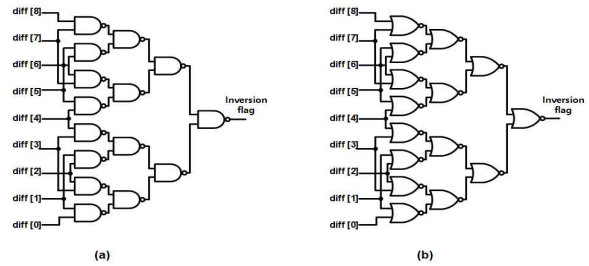


그림 3. (a) NAND 게이트로만 이루어진 제안된 회로 F0, (b) NOR 게이트로만 이루어진 제안된 회로 F1

F0,F1 모델은 AOI-AOI와 비교하여 구조는 같으나 스위칭 활동을 더 많이 감소시키며 트랜지스터 개수가 적은 NAND, NOR gate로만 이루어져있기 때문에 더 낮은 면적과 파워가 기대된다. 하지만 F0,F1모델은 여전히 93개의 높은 에러케이스를 발생시킨다. 그림4는 더 높은 정확도를 갖는 회로 FM을 보여준다. FM은 F0회로를 세 개를 병렬로 연결하였으며 각각 입력  $x, h_1(x), h_2(x)$ 를 갖는다.  $h_1(x), h_2(x)$ 은 입력을 골고루 받는 입력이며  $x$ 는 입력 diff[8:0]을 간략화하여 표현한 것이다.

$$x = (x_8, x_7, x_6, x_5, x_4, x_3, x_2, x_1, x_0)$$

$$h_1 = (x_4, x_3, x_5, x_8, x_0, x_6, x_2, x_1, x_7)$$

$$h_2 = (x_1, x_5, x_8, x_2, x_4, x_7, x_6, x_3, x_0)$$

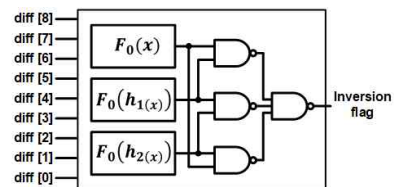


그림 4 .근사 과반수 회로 FM

FM은 9비트 입력의 전체 경우의 수 512개 중에서 39개의 에러케이스만 발생하며 스위칭 활동을 16.35% 감소시킨다. FM은 총 4개의 레벨로 이루어져있기 때문에 기존의 DBI와 비교하여 매우 낮은 지연시간을 갖는다.

### III. 결론

제안된 디자인과 MAJ-SYN, AOI-AOI는 상용 28nm 공정과 표준 라이브러리를 사용하여 설계 및 합성하였다. 표 1에서는 합성 결과인 지연시간, 면적, 작동 전력을 나타낸다. 그리고 Saving in switching activity는 DBI를 사용하지 않는 경우와 비교한 스위칭 감소 비율을 나타낸다.

표 1. Latency, Area, Power performance, Saving in switching activity of MAJ\_SYN, AOI-AOI, F0,F1,FM

	Latency (ns)	Area ( $\mu\text{m}^2$ )	Power ( $\mu\text{W}$ )	Saving in Switching activity (%)
MAJ_SYN	0.50	34.98	42.92	18.25
AOI-AOI	0.32	25.03	16.56	10.39
F0	0.22	23.86	16.00	12.75
F1	0.30	23.86	16.15	12.75
FM	0.32	39.86	29.35	16.35

제안한 회로 F0,F1은 MAJ-SYN와 비교하여 지연시간이 절반이상 줄어들었으며 면적과 전력이 각각 30%, 60% 이상 감소 하였다. 또한 AOI-AOI와 비교하였을 때도 더 낮은 지연시간과 면적, 작동 전력을 보여준다. FM은 많은 게이트가 사용되었기 때문에 MAJ\_SYN보다 더 높은 면적을 갖지만 전력은 31.6% 감소하였다. 또한 네 개의 게이트 레벨로 이루어져있기 때문에 여전히 낮은 지연시간을 보여준다.

이러한 저지연, 저면적, 저전력 특성을 갖춘 과반수 판정 회로는 인코더의 전력 소모를 감소시키며, 온칩 데이터 버스의 에너지 오버헤드를 효과적으로 해결하는 데 기여 한다.

### 참 고 문 헌

[1] Choudhary,J.;Balasubramanian,P.; Varghese, D.M; Singh, D.P.; Maskell, D. “Generalized majority voter design method for N-modular redundant

systems used in mission-and safety-critical applications.“ Computer 2019, 8,10

[2] Kon-Woo Kwon “Latency-Optimized Design of Data Bus Inversion”, electronics,2022,p1205,article 1205

[3] Kon-Woo Kwon “Majority Approximators for Low-Latency Data Bus Inversion ”, electronics, 2022, p3352,article 3352

[4] M.Madhu, V.Srinivasa Murty, “Dynamic Coding Technique For Low-Power Data Bus“, IEEE Computer Society Annual Symposium on VLSI(ISVLSI'03), 2003, p252-253

[5] Kon-Woo Kwon “Optimal bus coding for OR-chained buses”, IEICE Electronics Express, Vol.18, No.3, 1-4

[6] M. Gebhar, S.W.Keckler, B.Khailany, R.Krashinsky, W.J.Dally , “Unifying primary cache, scratch, and register file memories in a throughput processor” , 45<sup>th</sup> Annual IEEE/ACM International Symposium on Microarchitecture, 2012

### 사 사 문 구

※ 이 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음. (IITP-2021-0-02052)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대지능형반도체기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2021M3F3A2A01037531)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-PIM인공지능반도체핵심기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2022M3I7A2079155)

※ 이 연구는 IDEC에서 EDA Tool를 지원받아 수행하였습니다.



# 분산 GPU 클러스터 스케줄러 비교 분석

양 헤 일

## Comparative Analysis of Schedulers in Distributed GPU Clusters

Hyeil Yang\*

### 요 약

고성능 컴퓨팅 환경에서 GPU 자원의 효율적인 관리와 활용은 데이터 집약적인 연산 작업을 수행에 있어 매우 중요하다. 따라서, 적절한 GPU 스케줄링 전략은 자원 활용도를 최대화하고 전체 시스템 성능을 향상시키는 결정적인 역할을 한다. 본 연구는 Kubernetes 내에서 네가지 GPU 스케줄러를 비교하여 각 스케줄러의 장단점을 파악하는 것을 목표로 한다. 비교 대상이 되는 네 가지 스케줄러는 Kubernetes 기본 스케줄러, Posiedon, 그리고 Volcano이다.

### Abstract

Efficient management and utilization of GPU resources are crucial in high-performance computing environments, particularly for performing data-intensive computational tasks. Therefore, selecting an appropriate GPU scheduling strategy plays a critical role in maximizing resource utilization and enhancing overall system performance. This study aims to compare different GPU schedulers within Kubernetes to assess their strength and weakness. The schedulers under comparison are: Kubernetes Default Scheduler, Poseidon, Volcano.

### Key words

GPU Scheduling, Kubernetes, Resource Allocation, Kubernetes Scheduler

## I. 서 론

고성능 컴퓨팅과 인공지능 모델 훈련 작업의 증가로 인해, 대규모 클러스터 환경에서 GPU 자원을 효과적으로 관리하는 것이 매우 중요한 과제가 되었다[1]. 자원 할당 및 스케줄링은 전체 시스템 성능에 큰 영향을 미치며[2], 효율적인 스케줄러는 자원의 활용도를 높이고 작업의 처리 시간을 단축할 수 있다.

본 논문에서는 Kubernetes 환경에서 사용되는 네

가지 GPU 스케줄러를 비교한다. 비교 대상은 Kubernetes 기본 스케줄러, Poseidon, 그리고 Volcano이다. 본 연구의 목적은 각 스케줄러가 작업을 어떻게 처리하고 자원을 관리하는지를 분석하고, 각 알고리즘의 장단점을 논의하여 분산 환경에서의 최적의 스케줄링 전략을 도출하는데에 있다.

## II. 스케줄러 개요

- Kubernetes 기본 스케줄러

\*성균관대학교, 소프트웨어학과, vhveil@g.skku.edu

클러스터의 리소스 가용성을 바탕으로 작업을 적절한 노드에 배치한다[3].

• **Poseidon**

Poseidon은 Firmament 스케줄링 코어[4]를 통합하여, 그래프 기반 최적화 알고리즘을 사용한다. 이를 통해 자원 할당을 최적화한다.

• **Volcano**

고성능 작업을 위해 설계된 Kubernetes 스케줄러의 확장으로, 배치 처리, AI 등의 자원 집약적인 작업을 효과적으로 관리한다[5].

III. 스케줄러 비교

스케줄러를 비교하기 위해 총 2의 GPU가 탑재된 2개의 노드로 구성된 클러스터에서 실험을 진행하였다. 분산 학습 작업을 을위크로드로 사용하였고, 모든 스케줄러에 대한 테스트는 동일한 조건에서 진행하였다.

본 실험에서는 리소스 부족 상황에서 각 스케줄러가 어떻게 대처하는지를 평가하기 위하여 진행되었다. 이를 평가하기 위해, 스케줄러가 얼마나 많은 작업을 성공적으로 스케줄링하였는지를 나타내는 작업 완료 비율과 대기 시간을 측정하였다. 대기 시간은 스케줄러가 작업 처리를 시작할때까지 걸리는 시간을 의미한다.

• **Kubernetes 기본 스케줄러와 Poseidon**

표 1은 리소스 부족 상황에서 Kubernetes 기본 스케줄러와 Poseidon 스케줄러의 성능을 비교한다.

표 1. Kubernetes 기본 스케줄러와 Poseidon 스케줄러 비교

Table 1. Comparison between Kubernetes default scheduler and Poseidon

스케줄러	작업 완료 비율	총 대기 시간
Kubernetes Default Scheduler	100%	2분 12초
Poseidon	66.7%	무한대

Kubernetes 기본 스케줄러는 모든 작업을 성공적으로 스케줄링한 반면, Poseidon은 리소스 부족으로 일부 작업에 대해 스케줄링을 실패하였다. Poseidon은 현재 재스케줄링 기능을 지원하지 않기 때문에, 리소스 부족으로 실패한 작업은 다시 스케줄링 되지 않는다. 따라서 이에 대한 대기 시간을 무한대로 표시하였다.

• **Kubernetes 기본 스케줄러와 Volcano**

대규모 분산 학습 모델은 여러 GPU 사용을 필요로 하고, 이러한 GPU 사용을 요구하는 작업들은 동시에 스케줄링 되어야 한다. 이렇게 관련된 모든 작업 유닛을 동시에 스케줄링하는 것을 gang scheduling이라 한다[6].

Kubernetes 기본 스케줄러는 gang scheduling을 지원하지 않는다. Job 리소스의 parallelism 속성을 통해 한 번에 실행할 수 있는 파드의 최대 수를 지정할 수 있으나, 모든 파드가 동시에 시작되거나 동시에 리소스를 할당받는 것을 보장하지 않는다.

반면에, Volcano는 minAvailable 설정을 통해 gang scheduling을 지원한다. 지정된 모든 파드가 동시에 리소스를 할당받고 실행될 수 있도록 보류하는 기능을 통해, 모든 파드가 동시에 작업을 시작할 수 있도록 보장한다.

표 2는 Kubernetes 기본 스케줄러와 Volcano 스케줄러의 성능을 비교한 것이다. 실험은 작업의 요청량이 클러스터의 사용 가능 리소스를 초과하는 시나리오에서 진행되었다.

표 2. Kubernetes 기본 스케줄러와 Volcano 스케줄러 비교

Table 2. Comparison between Kubernetes default scheduler and Volcano.

스케줄러	작업 완료 비율	총 대기 시간	Gang scheduling 목표 부합 여부
Kubernetes Default Scheduler	100%	1분 33초	X
Volcano	100%	1분 30초	X
Volcano의 Gang Scheduling	0%	알 수 없음	O

이 실험을 통해 Kubernetes 기본 스케줄러는 순차적으로 작업을 실행하기 때문에 모든 파드를 성공적으로 스케줄링하였다. 이를 통해 Kubernetes 기본 스케줄러는 gang scheduling을 지원하지 않음을 확인할 수 있다.

Volcano는 gang scheduling을 활성화한 경우와 비활성화한 경우, 두가지로 나누어 실험을 진행하였다. Gang scheduling을 활성화하지 않았을 때, Volcano는 Kubernetes 기본 스케줄러와 유사하게 파드를 순차적으로 스케줄링한다. 이 경우, 모든 작업이 순차적으로 완료되며, 실행 시간은 Kubernetes 기본 스케줄러와 비슷하게 나타났다.

Gang scheduling을 비활성화한 경우, Volcano는 모든 파드가 동시에 리소스를 할당받고 실행될 수 있도록 작업 실행을 보류한다. 작업은 충분한 리소스가 확보될 때까지 실행이 지연된다. 이로 인해 실험에서는 작업 완료 비율이 0%로 나타났으며, 대기 시간은 측정 불가능하다.

#### IV. 결론

논문에서는 Kubernetes 환경에서 사용되는 세 가지 GPU 스케줄러: Kubernetes 기본 스케줄러, Poseidon, 그리고 Volcano 의 성능을 비교 분석하였다. 이러한 비교를 통해 각 스케줄러가 리소스 부족 상황과 분산 학습 모델 요구에 어떻게 대응하는지 확인하였다.

리소스 부족 상황에서, Kubernetes 기본 스케줄러는 사용 가능한 리소스를 이용하여 작업을 순차적으로 스케줄링한다. Poseidon은 스케줄링에 실패한 작업을 재스케줄링을 수행하지 않는다. Volcano는 gang scheduling을 사용하지 않는 경우, Kubernetes 기본 스케줄러와 유사하게, 작업을 순차적으로 스케줄링한다. 반면 Volcano 스케줄러의 gang scheduling을 사용하는 경우, 필요한 모든 리소스가 확보될 때까지 작업의 실행을 보류한다.

#### 참고 문헌

[1] Khallouli, Wael, and J. Huang, "Cluster resource scheduling in cloud computing: literature review

and research challenges," *The Journal of Supercomputing*, vol.78, no.5, pp.6898-6943, 2022.

- [2] Hussain, H., et al. "A survey on resource allocation in high performance distributed computing systems," in *Parallel Computing*, vol.39, no.11, pp.709-736, 2013.
- [3] Burns, B., et al. "Borg, omega, and Kubernetes," in *Communications of the ACM*, vol.59, no.5, pp.50-57, 2016.
- [4] Gog, I., et al, "Firmament:Fast, centralized cluster scheduling at scale," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp.99-115.
- [5] volcano-sh, "GitHub - volcano-sh/volcano: A Cloud Native Batch System (Project under CNCF)," GitHub, Sep.19, 2024.  
<https://github.com/volcano-sh/volcano> (accessed Oct.27, 2024).
- [6] D. G. Feitelson, "Packing schemes for gang scheduling," in *Proc. Workshop on Job Scheduling Strategies for Parallel Processing*, Springer, 1996, pp.89-110.

# OP-Amp 설계를 통한 555 Timer 구현

한정윤\*, 이은해\*, 이상원\*\*

## 555 Timer implementation with OP-Amp Design

Jeong Yun Han\*, Eun Hae Lee\*, and Sang Won Lee\*\*

### 요 약

555 timer는 PWM 생성, 타이머 기능, 신호 처리 등 다방면에서 사용되고 있다. 555 timer의 내부 회로는 전압 분배기, 비교기, 충전/방전 결정단, 방전단, 출력단으로 구성되어 있다. 해당 논문에서는 비교기 영역을 OP-Amp로 활용하여 재구성하는 방안에 대해 연구하였다. 이를 위해 555 timer의 동작 원리와 OP-Amp의 내부 회로를 분석하였으며, 더 나아가 555 timer를 활용하여 Duty Cycle을 자유롭게 조절할 수 있는 구형과 발생기를 구현하였다.

### Abstract

The 555 timer is widely used in various applications, including PWM generation, timing functions and signal processing. The internal circuit of the 555 timer consists of a voltage divider, comparators, charge/discharge control, a discharge stage and an output stage. Among them, this paper focused on reconfiguring the comparator section using OP-Amp. To achieve it, the operating principles of the 555 timer and the internal circuit of the OP-Amp are analyzed. Furthermore, a square wave generator was implemented to allows the free adjustment of the duty cycle.

### Key words

555 timer, OP-Amp, Duty Cycle, A Square Wave Generator

## I. 서 론

555 timer의 내부 회로는 Voltage divider, comparator(Threshold, Trigger) Latch, Output, Discharge로 이루어져 있다[1]. 해당 논문에서는 comparator를 OP-Amp를 활용하여 재구성할 수 있는 방안에 대해 연구하였다. 이를 위해 OP-Amp의 내부 회로를 5가지의 역할로 나누어 분석하였고, 임계 전압과 비교하여 NOR GATE를 활용한 Latch에 충

전과 방전을 결정할 값을 각각의 OP-Amp가 출력한다는 결과를 도출해내었다. 추가적으로, 기존 555 timer를 활용한 구형과 발생 어플리케이션에서의 듀티 제한 문제를 해결하기 위하여 다이오드를 추가하였으며, 이를 통해 커패시터의 충전 경로로 분리하여 듀티를 조절할 수 있도록 하였다.

## II. 555 timer 및 OP-Amp의 구성과 구형과 발생기의 동작 원리

\* 국립공주대학교 공과대학 전기전자제어공학부 전자공학 전공(dmsgo427@naver.com, jyhan1122@naver.com)

\*\* 교신저자 : 공주대학교 공과대학 전기전자제어공학부 전자공학 전공 (swlee@kongju.ac.kr)

## 2.1 555 timer의 동작 원리

그림 1은 555 timer의 내부 회로도이다. 타이머 내부에는 3개의 저항이 직렬로 연결되어  $V_{CC}$ 를 1/3 과 2/3으로 분할한다. 기준 전압은 Trigger(1/3  $V_{CC}$ )와 Threshold(2/3  $V_{CC}$ )를 감지하는 데 사용된다. 입력 신호가 1/3 또는 2/3  $V_{CC}$ 를 넘어서면, 비교기가 동작하여 Latch의 역할을 하는 파트를 제어한다. 이 비교기는 차동 증폭기 형태의 BJT 회로로, 입력 전압이 설정된 임계값을 넘었을 때 두 BJT 중 하나를 차단하고, 다른 하나를 활성화하는 방식으로 상태를 바꾼다. Latch의 상태에 따라 방전 트랜지스터가 켜지거나 꺼지며 커패시터가 방전된다. 이 과정은 파형의 듀티 사이클을 결정짓는 핵심적인 부분이 된다. 마지막으로 BJT가 출력단에서 증폭기 역할을 하여 안정적인 파형을 외부 회로로 전달한다.

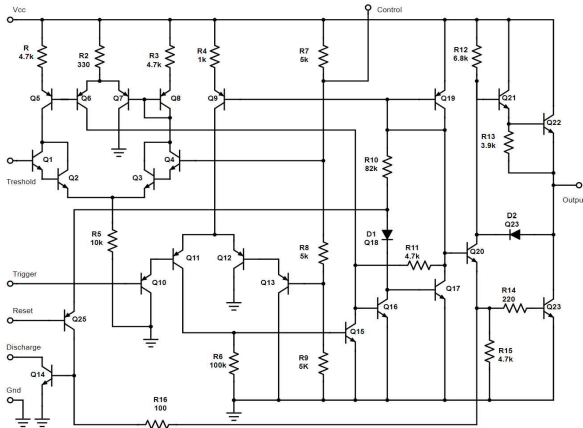


그림 1. 555 timer 내부 회로도

Fig. 1. 555 timer internal circuit diagram

## 2.2 OP-Amp의 동작 원리

OP-Amp는 그림 2와 같이 25개의 BJT와 11개의 저항으로 구성되어 있다. 이 회로는 표 1에서 표현되는 5가지 영역으로 분리할 수 있으며, Bias circuit, first stage, second stage, output stage 방향으로 전류가 흐르게 되며 입력 신호가 증폭되게 된다.

따라서, 555 timer의 Threshold 전압, Trigger 전압이 입력되면 OP-Amp를 지나 Latch에 적절한 출력이 되도록 한다.[2]

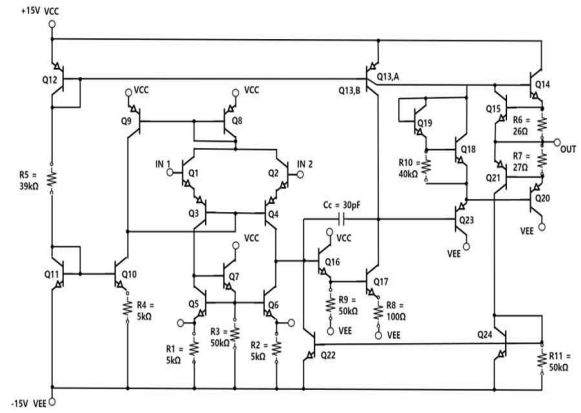


그림 2. OP-Amp 내부 구조 회로도

Fig. 2. OP-Amp internal structure circuit diagram

표 1. OP-Amp 영역별 설명

Table 1. OP-Amp area Description

영역	설명
Bias circuit	$I_{REF}$ 를 만드는 회로
First stage	- 차동 신호를 입력 받고, 전류 거울을 거쳐 단동 신호를 출력 - Level Shifter로 전압 크기가 작아짐
Second stage	Level Shifter로 낮아진 전압을 다시 증폭
Output stage	OP-Amp가 작은 출력 저항을 갖게 함
Short circuit	큰 전류가 흐를 때, IC 회로가 고장 나는 것을 방지하는 회로

## 2.3 구형파 발생기의 동작 원리

그림 3의 Charging Path를 보면, 충전 시간 동안  $R_1$ 과  $R_2$ 를 통해 커패시터를 충전한다.[3]

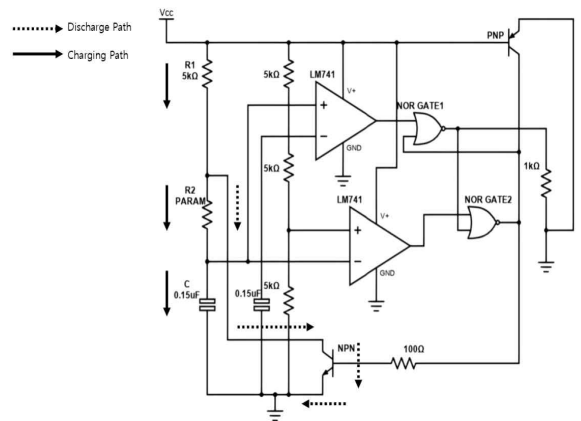


그림 3. 555 timer를 활용한 구형파 발생기 회로도

Fig. 3. Circuit diagram of square wave generator using 555 timer

그러나 Discharging Path를 보면, 방전 시간 동안 커패시터는  $R_2$ 만을 통해 방전된다는 것을 확인할 수 있다. 이는 식 (1), (2)로 나타낼 수 있으며, 충전 시간과 방전 시간을  $R_1$ ,  $R_2$ , 그리고 커패시터를 통해 바꿀 수 있음을 알 수 있다.

$$\tau = (R_1 + R_2)C \quad (1)$$

$$\tau = R_2C \quad (2)$$

이를 통해, 식(3), (4)와 같이  $T_1$ ,  $T_2$ 의 ON-time, OFF-time을 설정하여 듀티를 조절할 수 있다.

$$T_1 = (R_1 + R_2)C \ln 2 \approx 0.69(R_1 + R_2)C \quad (3)$$

$$T_2 = R_2C \ln 2 \approx 0.69R_2C \quad (4)$$

$$T = (R_1 + 2R_2)C \ln 2 \approx 0.69(R_1 + 2R_2)C \quad (5)$$

Duty Cycle은 전체 사이클, 식(5)과 ON-time의 비율로 나타낼 수 있으므로, 식 (6)과 같이 표현할 수 있다. 이에 해당 논문에서는  $R_2$ 의 값을 바꿔가며 Duty Cycle을 조절하였다.

$$Duty\ Cycle = R_1 / (R_1 + R_2) \times 100 (\%) \quad (6)$$

표 2에 따르면, 555 timer의  $R_2$ 를 증가시킬수록 Duty Cycle이 줄어드는 비율이 감소한다. 따라서,  $R_2$ 를 계속 증가시켜도 Duty Cycle은 최대 50%까지만 감소한다는 사실을 확인할 수 있다.

표 2. 555 timer Duty Ratio  
Table 2. 555 timer Duty Ratio

$R_1 / R_2$	Duty Ratio(%)	
	$T_1 / T$	$R_1 / (R_1 + 2R_2)$
1k/ 1k	84.5%	85.7%
1k/ 5k	65.8%	66.7%
1k/ 10k	59.3%	60%
1k/ 15k	56.6%	57.1%
1k/ 20k	55.3%	55.6%

### III. OP-Amp를 활용한 555 timer 회로 제안

앞서 설명한 555 timer의 내부 회로의 역할에 따라 소자를 대치하였다. 그림 4는 555 timer를 재구성한 내부 회로도이다. 2개의 OP-Amp와 2개의

NOR GATE, 저항, 커패시터, BJT로 구성되어 있다. Timer의 출력이 High가 되려면 Trigger의 입력이  $1/3 V_{cc}$ 보다 작아야 하고, 출력이 Low가 되려면 Threshold 전압이  $2/3 V_{cc}$ 보다 커야 한다.

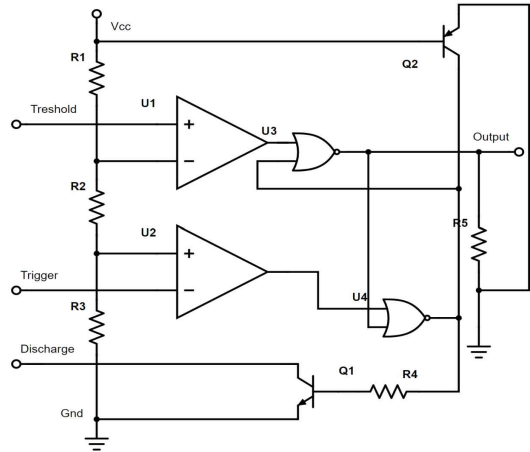


그림 4. OP-Amp를 활용한 555 timer 회로도  
Fig. 4. 555 timer circuit diagram using OP-Amp

예를 들어 입력 전압 = 5V, Threshold 전압 = 3V Trigger 전압 = 1V일 때, Latch의 reset = 0이 되고, set = 1이 된다. 이로 인해 출력 Q = 1이 되고, Q' = 0이 된다. 따라서, Latch가 set 되어 555 timer 출력이 High가 된다. 출력이 High가 된다는 것은 커패시터가 충전된다는 것을 의미하고, 반대로 Low가 되는 것이 방전되는 것을 의미한다.

## IV. 개선된 구형파 발생기 회로 제안

### 4.1 개선된 구형파 발생기 회로 구성

555 timer의 충전 시간과 방전 시간을 원하는 시간에 맞춰서 변경하려면 Duty Cycle의 조절이 필요하다. Duty Cycle을 조절하기 위해서는 timer의 시정수 값에 변화를 주면 된다. 즉, 시정수를 구성하는 ( $R_2$ ) 저항값에 변화를 주면 555 timer의 충·방전 시간을 조절할 수 있다.

앞선 실험의 결과로, 기존의 타이머 회로는 50% 이하의 듀티 사이클을 만들기 어렵다는 문제점이 있다는 것을 확인했다. 이에 그림 5와 같이  $R_2$ 에 다이오드를 추가하고, 충전 경로를 분리하는 방식을 제안한다.[4]

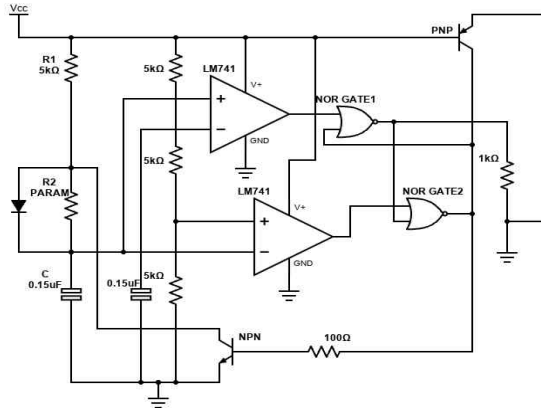


그림 5. Diode를 추가한 회로도  
Fig. 5. Circuit diagram with added Diode

$$T_1 = 0.69 R_2 C \quad (6)$$

$$T = (R_1 + R_2) C \ln 2 \approx 0.69 (R_1 + R_2) C \quad (7)$$

$R_2$ 에 다이오드를 추가하여 충·방전 회로를 분리함으로써 식(6), (7)을 도출할 수 있다. 이를 통해 제안하는 구형파 발생기의 Duty Cycle 식을 이끌어 낼 수 있었다.

$$Duty Cycle = \frac{R_1}{R_1 + R_2} \times 100(\%) \quad (8)$$

표 3은 다이오드를 추가한 회로의  $R_2$ 를 조절하여 식(8)과 실험 결과로 얻어낸 Duty Cycle 값을 비교하여 정리한 표이다. 그림 5의 회로를 분석하면, 충전 주기 동안 전류는  $R_1$ 와 다이오드를 통해 흐르고, 방전 주기 동안 전류는  $R_2$ 와 방전 트랜지스터를 통해 방전된다. 실험 결과를 식에 대입하여 비교한 결과, 듀티 사이클을 50% 이하로 낮출 수 있음을 증명할 수 있다.

표 3. 555 timer with Diode Duty Cycle  
Table 3. 555 timer with Diode Duty Cycle

$R_1 / R_2$	Duty Cycle(%)	
	$T_1 / T$	$R_1 / (R_1 + R_2)$
1k/ 1k	82.8%	83.3%
1k/ 5k	53.1%	50%
1k/ 10k	35.4%	33.3%
1k/ 15k	27.5%	25%
1k/ 20k	22.6%	20%

## 4.2 개선된 구형파 발생기 회로 구현

555 timer, OP-AMP의 동작 원리를 토대로 PCB 기판을 설계하여 구형파 발생기를 구현하였다. 입력단에 전압을 공급하면 이상적인 구형파가 발생하는 모습을 확인할 수 있다.

그림 6은 재구성한 555 timer에 Diode를 추가한 회로이다. 그림 7은 제안하는 구형파 발생기에 5V의 입력 전압을 공급했을 경우 형성되는 구형파의 모습이다. 신호의  $T_1$ 은 1.16 ms,  $T_2$ 는 1.2ms로 Total 2.72ms가 출력된다. 또한 커패시터 전압의 Peak to Peak는 1.88V, Output 전압의 Peak to Peak는 3.36V가 출력됨을 확인할 수 있었다. 또한 실험에서 확인된 주기를 변화시키기 위해 가변저항과 다이오드를 추가하여 Duty cycle을 50% 이하로 설정할 수 있다는 것을 표 3과 같이 알 수 있다.

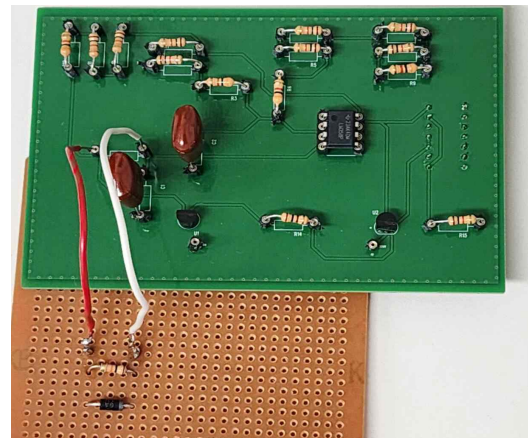


그림 6. 제안하는 구형파 발생기 PCB  
Fig. 6. Proposed square wave generator PCB

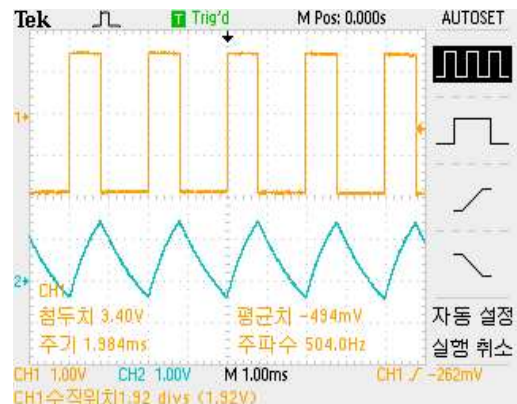


그림 7. PCB waveform  
Fig. 7. PCB waveform

## V. 결 론

해당 논문에서는 Bjt level의 555 timer의 각 역할을 분석하고 그 중 comparator를 OP-Amp를 활용하여 재구성하여 그 기능이 기존 모델과 같음을 증명하였다. 또한 555 timer를 바탕으로 구현한 구형과 발생기 회로에 Diode를 추가하여 충전 경로와 방전 경로를 분리하여 output 전압 듀티 사이클을 50% 미만으로 낮출 수 있는 가능성을 제시하고, 이를 실험을 통해 증명하였다.

## 참고 문헌

- [1] LM555 Datasheet. Retrieved from Texas Instruments. (n.d.). <https://www.ti.com/product/LM555>
- [2] Texas Instruments. (n.d.). AN-20 An Applications Guide for Op Amps. Retrieved from <https://www.ti.com/lit/an/snoa621c/snoa621c.pdf>
- [3] Md. Moyeed Abrar, Design and Implementation of Astable Multivibrator using 555 Timer <https://www.slideshare.net/slideshow/design-and-implementation-of-astable-multivibrator-using-555-timer/77459772>
- [4] Saunak Bhattacharya, Abhijit Kundu, Debraj Chakraborty, Angsuman Sarkar, Moumita Mukherjee, "Strain Modulated Asymmetrical Si/SiGe Superlattice p<sup>+</sup>-i-n<sup>+</sup> Switches for MMW Low-Loss Secure Communication Systems", IEEE Transactions on Device and Materials Reliability, vol.23, no.1, pp.14-26, 2023.
- [5] Camaron, N.(2021). Signal generation with 555 timer IC. In: Electronics Projects with the ESP8266 and ESP32. Apress, Berkely, CA



# 방사선 내성 SRAM들의 설계 면적과 안정성 비교 및 분석

이주헌\*, 서영교\*\*

## Area and Stability Comparison and Analysis of Radiation-Hardened SRAMs

Juheon Lee\* and Yeongkyo Seo\*\*

### 요 약

현대 컴퓨터 및 전자 장치의 확산과 함께 방사선 내성 SRAM의 필요성이 증가하고 있다. 방사선 내성 SRAM cell인 10T Quatro, 12T we-Quatro, 12T Dice 구조를 설계하고 시뮬레이션을 통해 비교 및 분석한다. Quatro와 we-Quatro의 면적은 동일하다. Dice에 비해 Quatro와 we-Quatro의 면적은 약 60%이다. 쓰기 실패 가속 환경에서 Quatro는 99.9%의 실패 확률을 보였고 Quatro는 5.5%, Dice는 11.6%이다. Soft error 복원력 비교 결과 Dice SRAM cell 이 가장 우수했다.

### Abstract

As modern computers and electronic devices become more widespread, the demand for radiation-hardened SRAM is growing. We designed and simulated radiation-hardened SRAM cells: 10T Quatro, 12T we-Quatro, and 12T Dice structures. Quatro and we-Quatro have the same area. Quatro and we-Quatro are approximately 60% smaller than Dice. In an accelerated write failure environment, Quatro exhibited a failure probability of 99.9%, while Quatro showed 5.5% and Dice showed 11.6%. When comparing soft error resilience, the Dice SRAM cell demonstrated the best performance.

### Key words

Radiation hardened memory, Quatro, we-Quatro, DICE, Area, Read & Write stability, Soft error resilience

---

\* 인하대학교 정보통신공학과 Email: juheonlee@inha.edu

\*\* 인하대학교 반도체융합전공 Email: yeongkyo@inha.ac.kr (교신저자)

※ 이 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음. (IITP-2021-0-02052)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대지능형반도체기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2021M3F3A2A01037531)

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-PIM인공지능반도체핵심기술개발(소자)사업 지원을 받아 수행된 연구임 (No. 2022M3I7A2079155)

※ 이 연구는 IDEC에서 EDA Tool를 지원받아 수행하였습니다.

## I. 서 론

현대 컴퓨터 및 전자 장치의 확산과 함께 방사선 내성 SRAM의 필요성이 증가하고 있다.

우주, 항공, 군사용 시스템에서는 방사능으로 인한 오류 발생 가능성이 높다. 지구 대기권을 벗어나거나 고고도에서 작동하는 장비는 방사선, 우주선 및 고에너지 입자에 노출되며, 이들은 전자 시스템에 SEU(Single Event Upset)를 일으킬 수 있다. 따라서 이러한 시스템에서 사용할 메모리, 특히 데이터 저장의 핵심인 SRAM에 방사선 내성이 필수적이다.

또한 의료 분야에서는 방사선 치료 장비가 암 환자 치료에 사용되며, 이러한 장비 내부의 컴퓨터 시스템 역시 방사선에 노출될 수 있다. 치료 과정 중 발생하는 방사선은 메모리 시스템에 영향을 미칠 수 있고 이는 치명적인 오류를 초래할 수도 있다.

이런 문제를 해결하기 위해 Quatro(그림2), We-Quatro(그림 3), Dice cell(그림 4) 등 Dual Interlocking 구조를 이용한 SRAM cell이 개발되었다. 이 논문에서는 Quatro, We-Quatro, Dice의 performance와 Area를 비교한다.

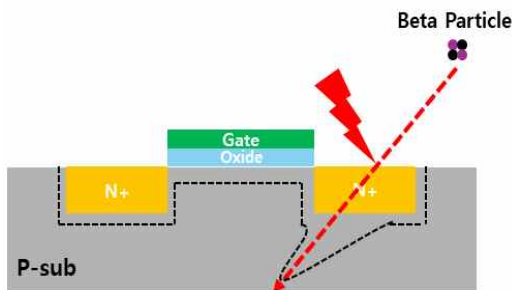


그림 1. 방사성 입자에 의한 Soft error 시나리오  
Fig. 1. Soft error scenario caused by radioactive particles

## II. 방사능 내성 SRAM cells 구조 및 동작 원리

Quatro, We-Quatro와 Dice cell은 모두 Dual interlocking 구조를 채택하여 soft error 복원력을 키운 SRAM이다.

예를 들어 Quatro cell의 데이터 저장 노드인 Q1에 1(high)가 저장되어 있다고 하자. Q1 위치에 방

사성 입자 충돌로 인해 순간적인 전류 발생으로 Q1의 전압도 순간적으로 감소하는 상황이라면, 직접적으로 영향을 받지 않는 Q4에서 Q3과 Q2의 전압을 유지하는 하도록 돕는다. 따라서 Q1의 전압이 변하는 것을 억제해준다.

we-Quatro는 기존 Quatro에 데이터 저장 노드인 Q3, Q4에 access transistor를 각각 추가하여 쓰기 성능을 높이는 구조이다. Access transistor를 추가했음에도 면적은 Quatro와 동일하여 큰 장점이 있다[1]. 따라서 soft error를 방지하는 원리는 Quatro와 동일하다.

Dice cell은 통상적인 6T SRAM 2개를 data node를 상호 보완적으로 연결한 구조이다. 비교적 연결이 간단하므로 회로도를 통한 분석이 쉽다. 하지만 layout을 하기에는 Quatro와 we-Quatro에 비해 연결이 복잡하여서 면적 크다는 것이 단점이다.

모든 SRAM의 설계는 삼성 28nm 공정을 이용하였다.

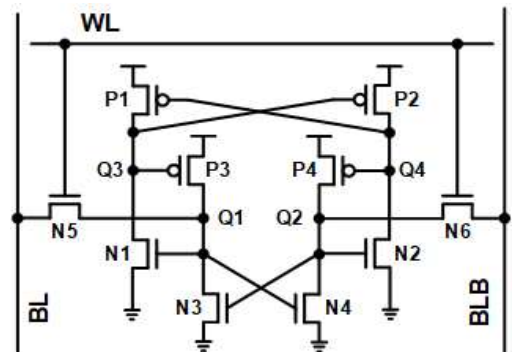


그림 2. 10T Quatro SRAM 회로도  
Fig. 2. 10T Quatro SRAM schematic

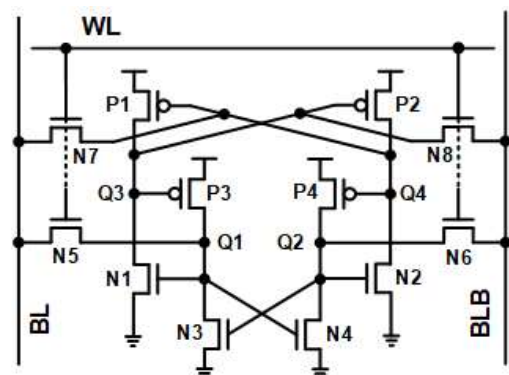


그림 3. 12T we-Quatro SRAM 회로도  
Fig. 3. 12T we-Quatro SRAM schematic

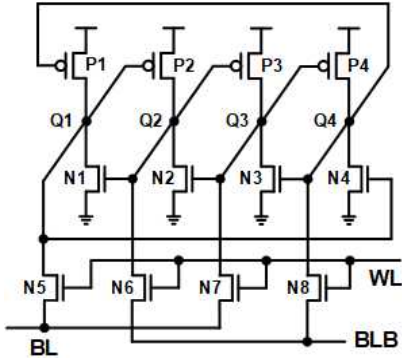


그림 4. 12T Dice SRAM 회로도  
Fig. 4. 12T Dice SRAM schematic

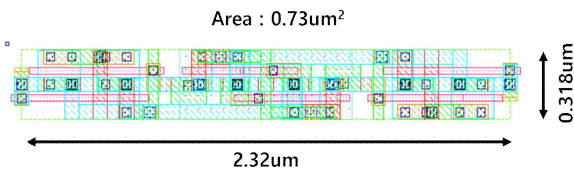


그림 5. Quatro SRAM 레이아웃 (면적: 0.73um²)  
Fig. 5. Quatro SRAM layout (Area: 0.73um²)

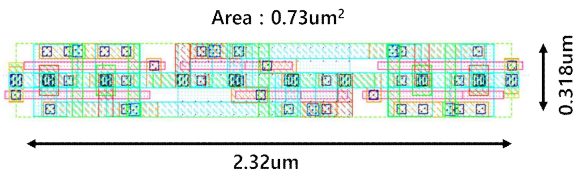


그림 6. we-Quatro SRAM 레이아웃 (면적: 0.73um²)  
Fig. 6. we-Quatro SRAM layout (Area: 0.73um²)

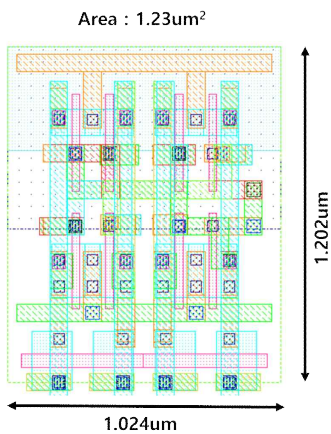


그림 7. DICE SRAM 레이아웃 (면적: 1.23um²)  
Fig. 7. DICE SRAM layout (AREA: 1.23um²)

### III. 쓰기 및 읽기 안정성

쓰기 안정성을 평가하기 위해 쓰기 실패 가능성을 공평하게 증가시킨 후 Monte Carlo simulation을

진행하였다. 쓰기 실패를 가속한 방법은 pull up PMOS의 width를 각 셀 모두 같은 크기만큼 증가시키고 시뮬레이션 온도를 섭씨 0도로 설정한 것이다. 또한 1V 였던 VDD를 0.6V로 낮춰보며 두 가지 경우를 모두 확인 하였다(그림 8). 총 10000번의 MC(Monte Carlo) simulation 결과 중 Quatro는 VDD가 1V와 0.6V일 때 모두 99.9% 실패하였고, We-Quatro는 VDD가 0.6V일 때 5.5%, DICE는 VDD가 0.6V일 때 11.6% 쓰기 실패하였다. 따라서 Quatro의 쓰기 안정성이 현저히 떨어지는 것을 확인할 수 있고 we-Quatro의 쓰기 안정성이 가장 좋은 것을 알 수 있다.

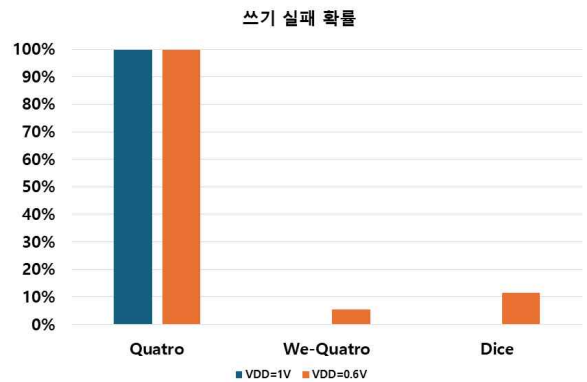


그림 8. 10000번 몬테카를로 시뮬레이션 쓰기 실패 확률  
Fig. 8. Write failure probability at 10000 MC simulation

읽기 안정성을 평가하기 위해 WL이 ON 되었을 때 BL과 BLB의 차이가 100mV(0.1VDD)가 될 때까지의 시간을 측정하였다(그림 9). 이 수치는 작을수록 좋은 것이다. Quatro의 access transistor 개수가 상대적으로 적기 때문에 읽기를 할 때 BL에 데이터를 전달해 주는 속도가 느린 것을 확인 할 수 있다.

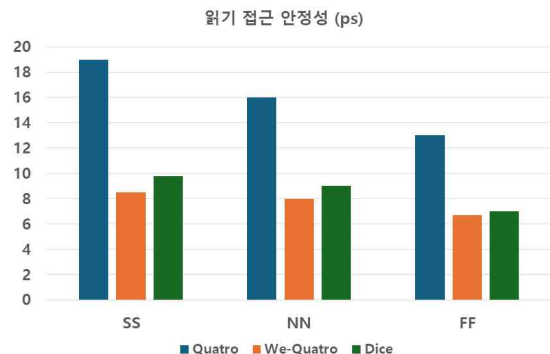


그림 9. 읽기 접근 안정성  
Fig. 9. Read access stability

읽기 안정성을 평가하는 지표 중 하나인 read SNM(Signal Noise Margin)은 읽기 접근 안정성과 trade off 관계에 있으며 측정 결과는 그림 10과 같다. Quatro의 Read SNM 이 가장 좋다는 것을 알 수 있다.

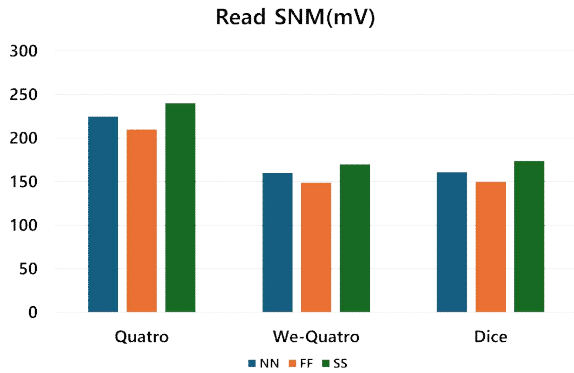


그림 10. Read SNM  
Fig. 10. Read SNM

#### IV. Soft error 복원력

Soft error 복원력은 인위적인 지수 전류원[3]을 통해 방사선에 의한 전류 모델을 데이터 저장 노드에 인가하여 측정하였다. soft error 발생하기 직전의 I peak를 측정하여 비교하였다. Quatro와 we-Quatro의 soft error 복원력은 비슷했고 Dice의 복원력이 가장 좋았다(그림 11). 그 이유는 구조의 차이에서 확인 할 수 있다. Quatro와 we-Quatro의 Q1 노드는 Q2, Q3의 pull down node인 N1과 N2에 직접적으로 영향을 주게된다. Q1이 0에서 1로 seu가 발생하는 상황이 발생하면 pull down N1과 N2가 동시에 켜지게 된다. 따라서 Q4에서의 soft error 보다 Q1이 취약하다.

Dice SRAM은 모든 데이터 저장 노드가 soft error 균등한 민감도를 갖는다. 이때 가장 취약한 상황은 한 데이터 저장 노드에서 1이었던 데이터가 0으로 바뀔 때이다. 하지만 이 경우 또한 pull up PMOS가 pull down NMOS보다 강해야지만 data가 flip된다. 따라서 Dice의 soft error 복원력 성능이 Quatro나 we-Quatro에 비해 좋은 것이다.

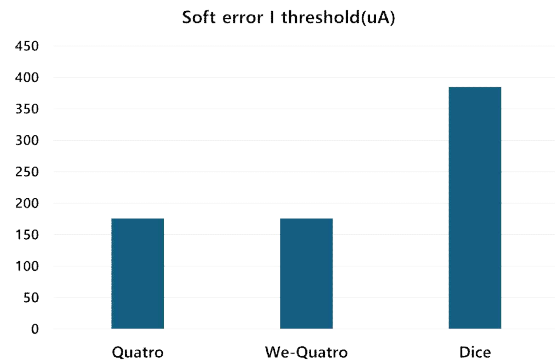


그림 11. Soft error 복원력  
Fig. 11. Soft error resilience

#### V. 결론

이 논문에서는 방사선 내성 SRAM cell인 Quatro, we-Quatro, Dice SRAM을 layout하여 면적을 비교하고, 시뮬레이션을 통한 성능 비교를 하였다.

면적 측면에서 Quatro와 we-Quatro의 면적은 Dice의 면적에 약 60%로 굉장한 이점이 있다. 하지만 Soft error에서 Dice는 높은 soft error 복원력으로 고신뢰성을 기대할 수 있다. 따라서 어떤 시스템에 사용되는지에 따라 적합한 SRAM cell이 달라질 것이다.

Soft error 복원력을 측정하기 위해 전류 모델을 사용했지만 더 정확한 분석을 위해서는 실제 chip에 방사성 입자를 방출하여 현실의 환경에서 어떻게 작용하는지 연구할 필요가 있다.

#### 참고 문헌

- [1] Le Dinh Trang Dang et al, "We-Quatro: Radiation-Hardened SRAM Cell With Parametric Process Variation Tolerance" in IEEE TRANSACTIONS ON NUCLEAR SCIENCE VOL. 64, NO. 9, SEPTEMBER 2017.
- [2] H. B. Wang et al., "An area efficient SEU-tolerant latch design," IEEE Trans. Nucl. Sci., vol. 61, no. 6, pp. 3660-3666, Dec. 2014.
- [3] R. C. Baumann, "Radiation-induced soft errors in advanced semiconductor technologies," IEEE Trans. Device Mater. Rel., vol. 5, no. 3, pp. 305-316,

Sep. 2005.

- [4] K. Agawa, H. Hara, T. Takayanagi, and T. Kuroda, "A bitline leakage compensation scheme for low-voltage SRAMs," *IEEE J. Solid-State Circuits*, vol. 36, no. 5, pp. 726-734, May 2001.
- [5] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 24, no. 12, pp. 1859-1880, Dec. 2005.



# Focal-Length를 이용한 유해 조류 거리 측정

권태현\*, 이영학\*\*

## Distance Estimation of Hazardous Birds Using Focla Length

Tae-Hyun Kwon\* and Yeung-Hak Lee\*\*

### 요 약

다수의 과수원이 유해 조류로 인해 피해를 입고 있다. 본 논문에서는 비전 AI를 이용한 유해 조류를 퇴출하기 위해 카메라와 새까지의 거리를 연구한다. 조류를 퇴치하기 위해 객체와의 거리를 측정하는 문제를 해결하고자 Focla-Length를 사용하였으며, 객체의 대푯값을 설정하며 고려할 점들과 그 해결 방법을 기술하였다

### Abstract

Many orchards are suffering damage from harmful birds. This paper studies the distance between the camera and birds to deter these harmful birds using vision AI. To solve the problem of measuring the distance to objects for bird control, we utilized focal length. We also discussed the considerations and solutions for setting representative values of the objects.

### Key words

Harmful Birds, Distance Measurement, Vision AI, Focal Length, Representative Value Setting

## I. 서 론

유해 조류가 과수원 농작물에 입히는 피해가 늘고 있다. 2023년 기준 피해액은 180농가에 2억 2000만 원 상당이며, 피해 규모 또한 2021년 0.59km<sup>2</sup>에서 2023년 0.88km<sup>2</sup>로 확대되었다. 농작물 피해의 50%가 조류에 의한 피해인 만큼 유해 조류를 퇴치하기 위해 퇴치 시스템이 필요하다[1]. 이 시스템은 조류를 탐지하여 퇴치 장치를 작동시키며 이때 조류가 충분히 가까이 왔을 때 퇴치 장치가 작동되어야만 효

율적인 퇴치를 할 수 있다. 이를 위해 Focal-Length를 이용해 조류와의 거리를 측정하는 이론적 방법을 제시한다.

## II. Focal-Length

Focla-Length(초점 거리)란 카메라의 렌즈와 이미지 센서 간의 거리를 의미한다. 초점 거리(d)와 이미지 센서에 맺힌 상의 크기(h)를 이용해 피사체와의 거리(D) 또는 크기(H)를 알 수 있다[2].

\* 국립안동대학교, teaggoggo@naver.com

\*\* 국립안동대학교 SW융합교육원, yhyi@anu.ac.kr(교신저자)

※ “본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음” (2019-0-01113)

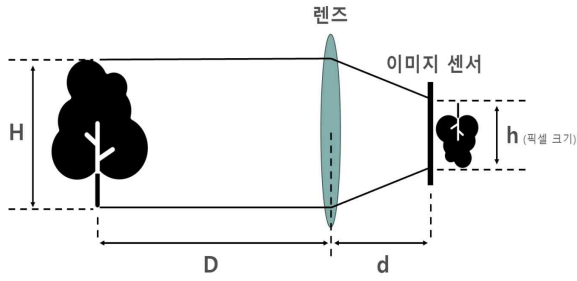


그림 1. Focal-Length의 기본 구성요소  
Fig. 1. Basic Components of Focal Length

Focal-length를 이용해 다음과 같이 피사체의 크기를 측정할 수 있다.

$$D = \frac{d}{h} \times H \quad (1)$$

그러나 위 방법을 이용해 피사체와의 거리를 구하기 위해서는 피사체의 크기를 알아야 한다. 이러한 문제를 해결하기 위해서 피사체의 크기에 대한 기준을 잡아 거리를 구하기로 한다.

### III. 실제 크기의 대푯값 설정

#### 3.1 대푯값 설정 기준

새를 탐지한 박스의 가로, 세로 길이는 새의 자세와 각도에 따라 매우 큰 편차가 발생한다. 따라서 이 연구에서는 박스의 대각선을 이용하기로 하였다. 참고로 객체 하나의 라벨링 박스에 대한 대각선의 길이는 다음과 같이 구할 수 있다.

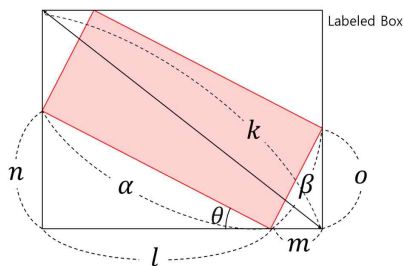


그림 2. 라벨링 박스의 대각선 길이 계산  
Fig. 2. Calculating the Diagonal Length of the Labeling Box

$$k = \frac{\sqrt{(l+m)^2 + (n+o)^2}}{\sqrt{(\alpha \cos \theta + \beta \sin \theta)^2 + (\alpha \sin \theta + \beta \cos \theta)^2}}$$

대각선을 이용해 실제 크기의 대푯값을 구한 결과, 대각선의 길이 분포가 가우시안 분포를 따름을 확인했다.



그림 3. 새 탐지 박스의 대각선 길이  
Fig. 3. Diagonal Length of Bird Detection Box

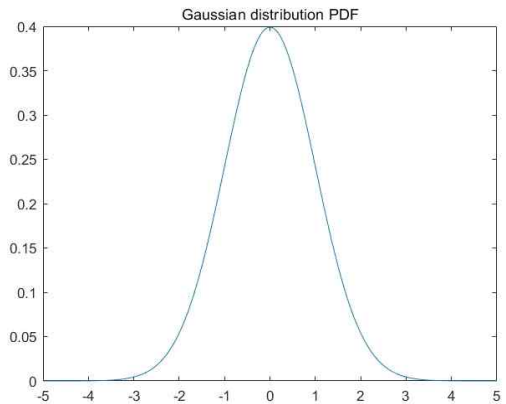


그림 4. 새 탐지 박스의 크기에 따른 그래프  
Fig. 4. Graph of Bird Detection Box Size

따라서 실제 크기의 대푯값으로서 가우시안 분포의 중앙값을 따르기로 한다.

#### 3.2 객체별 실제 크기 차이 문제 및 해결

새와 일정 거리가 되면 새를 쫓는 장치(폭죽 등 [3])가 작동한다. 그러나 새의 대푯값을 하나로 정하는 경우 새의 실제 크기에 따라 실제 거리에 차이가 생긴다. 실제 새의 크기는 10cm에서 1m가 넘는 종까지 다양하다. 그러나 모든 새를 종류에 따라 구분하는 것은 현실적으로 어렵다. 따라서 대푯값을 큰 새로 설정할지 작은 새로 설정할지에 대한 기준이 필요하다.



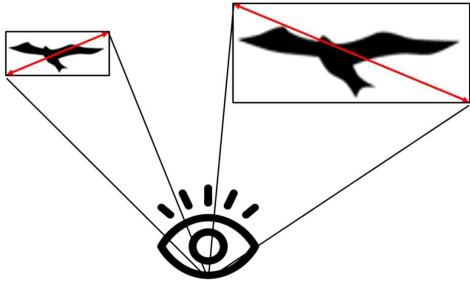


그림 5. 새의 크기에 따른 기준 설정  
Fig. 5. Criteria Setting Based on Bird Size

### 3.2.1 작은 새를 기준으로 하는 잡는 경우

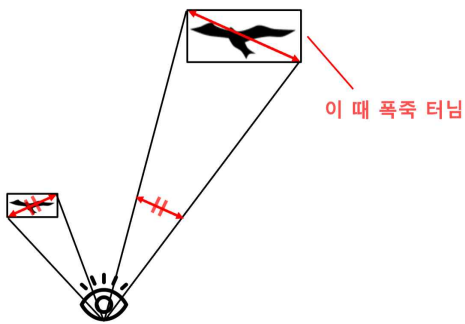


그림 6. 작은 새를 기준으로 설정한 경우의 장치 작동  
Fig. 6. Operation of the Device When Based on Small Birds

작은 새를 기준으로 설정하여 거리를 계산하는 경우, 큰 새가 멀리 있을 때 계산한 거리에 있다고 판단하여 장치가 작동한다.

### 3.3 큰 새를 기준으로 하는 잡는 경우

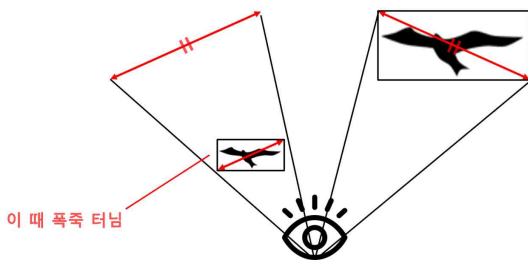


그림 7. 큰 새를 기준으로 설정한 경우의 장치 작동  
Fig. 7. Operation of the Device When Based on Large Birds

큰 새를 기준으로 설정하여 거리를 계산하는 경우, 작은 새가 가까이 있을 때 계산한 거리에 있다고 판단하여 장치가 작동한다.

## 3.4 비교분석 및 문제 해결

표 1. 기준 삼을 크기에 따른 분석

Table 1. Analysis Based on the Size Used as a Reference

기준 삼을 크기	반대 크기에 대한 장치 작동 거리	특징
작은 새	멀 때 작동	효과 미비 과하게 작동
큰 새	가까울 때 작동	기준 이상의 효과 작동 횟수 최소화

분석 결과 큰 새를 기준으로 삼아 대포값을 설정하고 거리를 계산해야 한다.

## IV. 결론

본 논문에서는 효율적인 유해 조류 퇴치를 위해 조류의 거리를 측정하는 방법을 제시하고, 객체의 대포값 설정에 대한 문제점과 해결 방법을 연구하였다. 대상이 되는 조류 크기의 대포값을 설정하여 Focla-Length를 이용함으로써 거리를 측정할 수 있다. 대포값 설정 문제에서 대상의 대각선을 측정하며, 큰 새를 조류 크기의 대포값으로 설정함으로써 객체별 크기 차이 문제를 해소하고 퇴치 장치 작동의 효율성을 높여 유해 조류 퇴치 시스템의 정확성과 효율성을 크게 향상시킬 것으로 기대된다.

## 참고 문헌

- [1] <https://www.newsjeju.net/news/articleView.html?idxno=400830> (accessed by October 9, 2024)
- [2] <https://www.edmundoptics.com/knowledge-center/application-notes/imaging/understanding-focal-length-and-field-of-view/> (accessed by October 9, 2024)
- [3] 김영호, 윤철구, 임상철, 손인창, 박재혁, 김동환, 김선규, “과수 조류피해 방지에 관한 연구”, 연구보고서, pp. 18-28, 2002.

# 노약자를 위한 스마트 보행기

문지환\*, 이은성\*, 이진윤\*, 유병일\*, 이승환\*, 이인수\*\*

## Smart Walker for the Elderly

Ji Hwan Moon\*, Eun Seoung Lee\*, Jeon Yun Lee\*, Byeong-il Yu\*, Seung Hwan Lee\*,  
and In-Soo Lee\*\*

### 요약

본 연구에서는 고령화 사회를 대비하여 노인들의 안전한 이동을 지원하는 스마트 보행기를 개발하였다. 제안된 보행기는 실시간 위치 추적, 위험물체 인식, 그리고 음성 기반 챗봇을 통해 노인의 편의성을 향상시킨다. 특히, 얼굴 인식 로그인 기능은 미래의 공유 모빌리티 사회에서 다수의 사용자가 보행기를 안전하게 공유할 수 있도록 설계되었다. 본 연구를 통해 노인들의 독립적인 생활을 지원하며, 보호자가 실시간으로 보행자의 상태를 확인하고 응급 상황에 신속히 대처할 수 있는 기술적 기반을 제공한다.

### Abstract

This study proposes a deep learning model that utilizes an IMU sensor to classify fall incidents and types. A system based on a Feedforward Neural Network was developed to classify non-falls, forward falls, backward falls, and side falls in real-time. Test results showed that the model accurately predicted most types of falls, especially detecting the start and end of falls in real time, confirming its potential as a real-time fall detection system. The proposed model can serve as an important foundation for the development of real-time fall prevention and detection systems.

### Key words

Smart walker, face recognition, real-time location tracking, hazard object detection, voice chatbot, application

## 1. 서론

고령화 사회가 가속화됨에 따라, 노인들의 안전하고 독립적인 생활을 지원하는 기술의 필요성이 증가하고 있다. 특히, 노인들의 이동을 돕는 보행기의 중요성이 더욱 부각되고 있다고 생각하여, 기존의 보행기 기능을 넘어선 노인들의 안전과 편의성

을 극대화할 수 있는 스마트 보행기를 개발하였다.

본 연구에서는 노인의 이동을 돕고, 실시간 위치 추적, 위험물체 인식, 음성 챗봇 등의 기능을 통합한 스마트 보행기를 개발하는 것을 목표로 하였다. 이를 통해 노인의 안전과 편의성을 높이며, 보호자가 실시간으로 보행자의 상태를 모니터링하고 긴급 상황에 빠르게 대처할 수 있도록 한다.

\* 경북대학교 IT대학 전자공학부 / dlwjsdbs@knu.ac.kr

\*\* 경북대학교 IT대학 전자공학부 교수 / insoolee@knu.ac.kr

## II. 스마트 보행기 구성

본 연구에서는 고령화 사회에 대비하여 노인들의 스마트 보행기를 개발하였다. 아래 그림 1은 스마트 보행기에 사용된 장비와 기능을 보여준다.

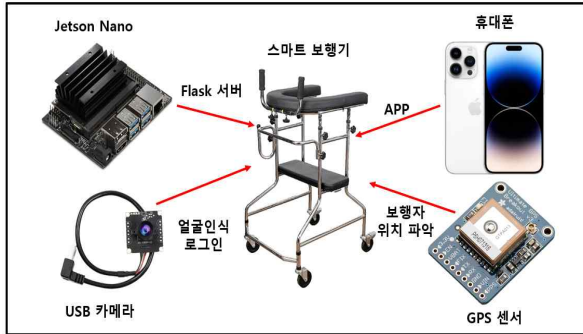


그림 1. 스마트 보행기의 구성  
Fig. 1. Components of the smart walker

### 2.1 애플리케이션

본 연구에서는 보행기의 다양한 기능을 유기적으로 연결하기 위해 애플리케이션을 제작하였다. 애플리케이션은 다양한 기능을 보행자와 보호자가 손쉽게 활용할 수 있도록 설계되었다. 또한, 사용자 경험을 고려해 직관적이고 간편한 인터페이스를 제공하여, 노인층도 쉽게 사용할 수 있도록 했다.

#### 2.1.1 Flask 서버

본 연구에서는 Python의 라이브러리인 Flask를 사용하여 서버와 클라이언트 간의 통신을 구현하였다. 아래 그림 3은 통신 과정을 보여준다.

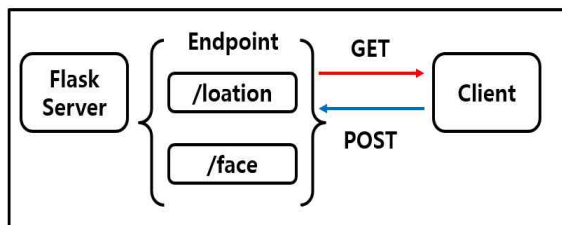


그림 2. 서버와 클라이언트의 통신 과정  
Fig. 2. Communication process between the server and client

엔드포인트는 서버가 클라이언트로부터 요청을 받고 처리하는 경로를 의미하며, 다양한 요청을 효

율적으로 처리하기 위해 엔드포인트를 사용하였다.

### 2.2 GPS를 이용한 보행자 위치 파악

본 연구에서는 보행자의 위치를 실시간으로 추적하여 보호자에게 제공하기 위해 GET 및 POST 방식을 사용하는 애플리케이션을 개발하였다. 아래 그림 3은 보행자의 위치를 확인하는 화면을 보여준다.



그림 3. 보행자의 위치 파악  
Fig. 3. Pedestrian position tracking

Flask 서버는 이러한 통신을 처리하며, 이를 처리하기 위해 /location 엔드포인트를 사용하였다. 보행자는 GPS 센서를 통해 수집한 위치 데이터를 POST 방식으로 서버에 전송한다. 보호자는 같은 엔드포인트에 GET 요청을 보내 실시간으로 보행자의 최신 위치를 조회할 수 있다.

이를 통해 보행자의 위치를 실시간으로 모니터링함으로써 보호자가 신속한 대응을 할 수 있도록 돕는다.

### 2.3 얼굴인식 로그인

본 연구에서는 OpenCV를 이용하여 사용자 얼굴인식을 이용한 로그인을 구현하였다.

이 기능은 고령화 사회에서 노인 인구의 증가에 따른 공유 보행기 사용을 염두에 두고 설계된 기능으로, 미래의 모빌리티 트렌드를 고려하였다. 얼굴인식을 통해 사용자가 보행기를 안전하고 편리하게 이용할 수 있도록 하여, 공유 보행기의 효율적인 사용과 보안성을 동시에 확보하였다.

### 2.3.1 얼굴 인식 학습

본 연구에서는 사용자의 얼굴을 학습하기 위해 얼굴을 검출, 인식, 학습하는 과정으로 구분하였다. 아래 그림 5는 얼굴 인식 학습 과정을 보여준다.

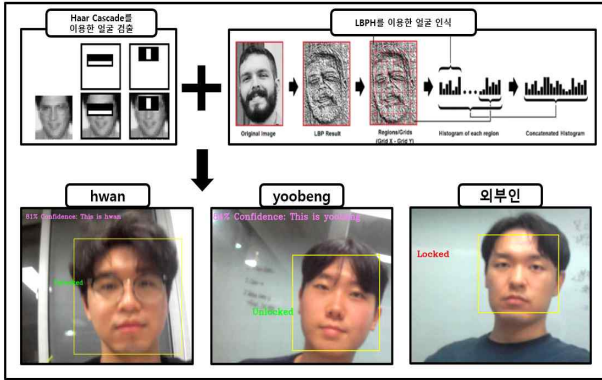


그림 4. 얼굴 인식 학습 과정  
Fig. 4. Facial Recognition Training Process

얼굴 검출을 위해 사용되는 OpenCV의 Haar Cascade 분류기는 이미지에서 얼굴이 있는 영역을 빠르게 탐지하는 역할을 한다. 이를 통해 다양한 각도와 조명에서 수집된 사용자 얼굴 이미지에서 얼굴을 정확히 검출한다.

얼굴이 검출되면, LBPH(Local Binary Patterns Histogram) 알고리즘을 사용하여 얼굴의 특징을 추출하고 인식한다. LBPH는 얼굴의 국소적인 패턴을 히스토그램으로 변환하여 얼굴의 특징을 벡터 형식으로 저장한다. 이는 얼굴의 특징을 비교하고 학습할 수 있는 중요한 데이터로 사용된다.

LBPH 알고리즘을 통해 추출된 얼굴 특징을 바탕으로 학습을 진행하고, 학습된 모델을 저장하였다. 학습된 모델은 얼굴을 인식할 수 있는 형태로 변환되며, 이후 얼굴 인식 요청 시 사용자의 얼굴을 정확히 인식할 수 있도록 한다.

### 2.3.2 앱 인벤터와 Flask 서버 간 통신

학습된 데이터는 Flask 서버에서 얼굴 인식 요청이 들어왔을 때 활용된다. 사용자가 얼굴 인식을 시도할 때, Flask 서버의 /face 엔드포인트로 GET 요청을 보내며, 서버는 카메라로 촬영한 이미지를 학습된 얼굴 데이터와 비교하여 성공 여부를 판단한

다. 얼굴이 성공적으로 인식되면 서버는 success 신호와 함께 사용자의 이름을 응답으로 전송한다.

애플리케이션은 이 데이터를 수신하여 사용자가 로그인에 성공했음을 처리하며, 만약 얼굴 인식에 실패할 경우 실패 알림을 출력한다. 이를 통해 얼굴 인식을 기반으로 한 보행기 사용이 안전하고 편리하게 이루어질 수 있다.

### 2.4 위험물체 인식

본 연구에서는 YOLOv5 모델을 사용하여 보행기 주변의 위험물체를 실시간으로 인식하는 기능을 구현하였다. 아래 그림 6은 위험 물체를 인식할때의 화면을 보여준다.



그림 5. 위험물체 인식  
Fig. 5. Structure of App

포트홀과 자동차 데이터는 각각 ROBOFLOW와 COCO를 사용해 학습했으며, threshold 값을 0.85로 설정하여 보행기의 일정 거리 내로 접근하는 위험물체를 감지한다. 또한, 젯슨 나노를 이용해 YOLO 모델이 7.5 FPS의 속도로 위험물체를 실시간으로 인식할 수 있다.

### 2.4.2 위험물체 인식 및 경고 시스템 연동

YOLOv5 모델이 위험물체를 인식하면, 해당 정보를 Flask 서버로 전송하는 시스템이 구축되었다. Flask 서버는 위험물체 인식 데이터를 실시간으로 챗봇에게 전달하며, 챗봇은 이를 바탕으로 보행자에게 이상 여부를 음성으로 묻는다. 보행자가 5분 이내에 응답하지 않거나 응급 상황이 발생했다고 판

단되면, Flask 서버는 보호자에게 경고 신호를 전송한다. 이 신호는 보호자의 앱으로 전달되며, 앱은 즉시 경고 알림을 표시하여 보호자가 상황을 빠르게 인지하고 대응할 수 있도록 설계되었다.

## 2.5 ChatBot

스마트 기기에 익숙하지 않은 노년층을 위해 보행기에 탑재될 음성 챗봇을 개발하였다. 아래 그림 6은 챗봇의 개요를 보여준다.

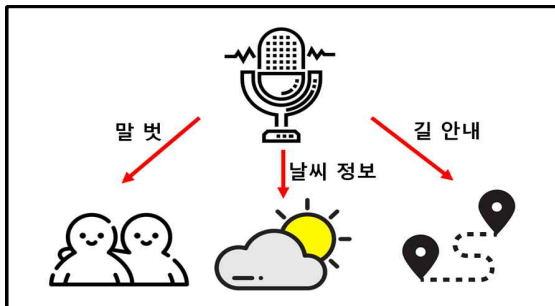


그림 6. ChatBot의 개요  
Fig. 6. Overview of Chatbot

간단한 음성 명령만으로 목적지를 안내받고, 현재 날씨 정보를 얻을 수 있다. 또한 홀로 지내는 독거노인에게 정서적 지원을 제공하고자 말벗 기능을 추가하였다. 음성 챗봇은 사용자에게 편의성과 더불어 친근한 상호작용을 통해 일상 속 동반자가 되어 줄 것이다.

### 2.5.1 Chatbot 구현 과정

음성 서비스를 위해 STT(Speech To Text)와 TTS(Text To Speech) 기능의 음성 챗봇을 개발하였다. 아래 그림 7은 챗봇의 구현 과정을 보여준다.

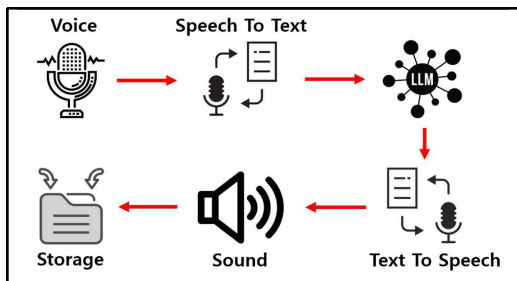


그림 7. 챗봇 구현 과정  
Fig. 7. Implementation Process of ChatBot

음성 명령을 내리면, OpenAI의 Whisper AI를 활용해 음성을 텍스트로 변환하고, 이 텍스트를 OpenAI의 GPT-4.0 Turbo API를 사용하여 답변한다. 데이터 저장을 위해 MongoDB Atlas를 사용하여 대화를 더욱 자연스럽게 관리하고, 사용자 맞춤형 상호작용을 제공한다. 아래 표1은 챗봇 구현에 사용된 주요 API 서비스를 보여준다.

표 1. 챗봇 구현에 사용된 주요 API 및 서비스  
Table 1. APIs and Services Used for Chatbot Implementation

기능	활용된 API 및 서비스
음성인식(TTS)	OpenAI Whisper AI
음성 합성(TTS)	Microsoft Azure TTS
대화 생성	OpenAI GPT-4.0 Turbo API
위치 정보 분석	Kakao Rest API
날씨 정보 제공	OpenWeather API
데이터 저장	MongoDB Atlas

### 2.5.2 프롬프트 엔지니어링

챗봇의 대화 톤과 말투를 자연스럽게 친근하게 조정하기 위해 프롬프트 엔지니어링이 적용되었다.

#### 2.5.2.1 역할(role) 설정

챗봇의 역할을 설정하여, 친근하고 공손한 말투를 유지할 수 있도록 했다. 가령, "모든 답변은 -요체로 대답해야 합니다."와 같이 명확한 지시를 포함하여 자연스러운 대화 흐름을 유지하도록 역할을 정의하였다.

#### 2.5.2.2 매개변수 설정

아래 그림 8은 실시간 사용자 응답을 통한 챗봇 대화 예시를 보여준다.

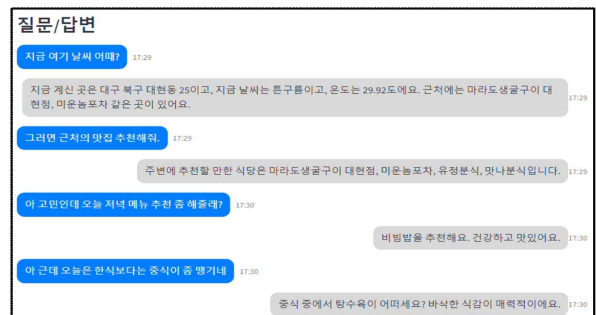


그림 8. 실시간 사용자 응답을 통한 챗봇 대화 예시  
Fig. 8. Real-Time User Responses in Chatbot

temperature(0.8), top\_p(1.0), frequency\_penalty(0.65) 등의 매개변수를 조정하여 챗봇의 응답 스타일을 최적화하였다. temperature(0.8)는 응답의 창의성을 적당히 유지하면서 예측 가능성을 높였고, top\_p(1.0)는 가장 높은 확률의 응답을 선택하게 했으며, frequency\_penalty(0.65)는 동일 단어의 반복을 적절히 억제하여 다양한 표현을 사용할 수 있도록 조정하였다.

### 2.5.3 네비게이션 및 날씨 정보 제공

서버에서 받아온 위도와 경도를 사용하여 Kakao의 Rest API를 통해 해당 위치의 정확한 주소를 파악할 수 있다. 또한, 현재 위치를 중심으로 주변의 주요 건물, 목표 지점까지의 거리, 이동해야 할 방향 등을 사용자에게 설명한다.

위치 정보는 OpenWeather API를 사용해 현재 날씨와 예보를 받아오는 데 활용된다. 이를 통해 사용자는 이동 시의 기상 상황을 실시간으로 파악할 수 있으며, 기상 변화에도 대비할 수 있다. API에서 얻은 데이터는 Microsoft Azure의 TTS 기능을 사용하여 'ko-KR-SunHiNeural' 목소리로 출력된다.

### 2.5.4 위험 신호 처리

위험물체 인식 기능을 통해 위험 상황이 감지될 경우, 서버에서 위험 신호를 전달받고 즉시 보행자의 안전을 확인한다. 보행자가 15초 이내 답변이 없거나 도움이 필요하다고 요청하면 서버에 위험 신호를 전송한다. 아래 그림 9는 길 안내 및 안부 확인에 대한 챗봇 대화 예시를 보여준다.

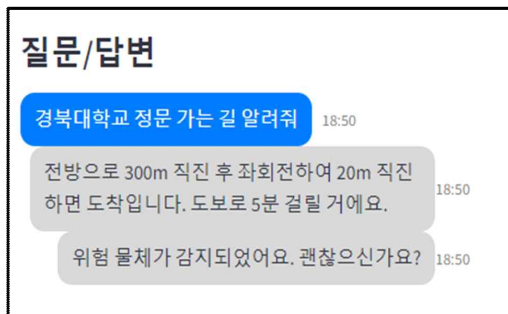


그림 9. 길 안내 및 안부 확인에 대한 챗봇 대화 예시  
Fig. 9. Directions and Safety Check in Chatbot

## III. 결 론

본 연구에서는 고령화 사회를 대비하여 노인들의 안전하고 편리한 이동을 지원하기 위한 차세대 스마트 보행기를 개발하였다. 제안된 보행기는 실시간 위치 추적, 위험물체 인식, 음성 기반 챗봇, 얼굴 인식 로그인 등 다양한 기능을 통합하여 보행자의 안전을 실시간으로 모니터링하고, 긴급 상황에 신속하게 대응할 수 있도록 설계되었다. 특히, 얼굴 인식 로그인을 통해 다수의 사용자가 보행기를 안전하게 공유할 수 있는 시스템을 구현함으로써, 미래의 공유 모빌리티 사회에 대비한 효율적인 사용 방안을 제시하였다.

이번 연구를 통해 제안된 스마트 보행기 시스템은 노인들의 독립적 생활을 지원하고, 보호자가 실시간으로 상태를 확인할 수 있는 기술적 기반을 제공함으로써 노인 복지 향상에 기여할 수 있을 것으로 기대된다. 추후 연구에서는 더 다양한 환경에서의 테스트를 통해 시스템의 정확성을 높이고, 추가적인 기능을 통합하여 사용자의 편의성을 더욱 강화할 예정이다.

## 참 고 문 헌

- [1] Chen, K., et al. "Smart walkers for the elderly: A survey." IEEE Sensors Journal 17.17 (2017): 5775-5784.
- [2] Zekavat, S. A., and Buehrer, R. M. "Handbook of Position Location: Theory, Practice and Advances." John Wiley & Sons, 2011.
- [3] Zhang, K., et al. "Joint face detection and alignment using multitask cascaded convolutional networks." IEEE Signal Processing Letters 23.10 (2016): 1499-1503.
- [4] Ultralytics YOLOv5, <https://docs.ultralytics.com/ko>
- [5] Radford, A., et al. "Robust speech recognition via large-scale weak supervision." arXiv preprint arXiv:2111.02861 (2021).

# 효율적 BIM을 위한 다중 센서 기반 Perception Torch

김현우\*, 권유진\*, 김재민\*, 이헌철\*\*

## Multi-sensor based Perception Torch for efficient BIM

Hyeonwoo Kim\*, Yujin Kwon\*, Jamin Kim\*, and Heoncheol Lee\*\*

### 요약

4차 산업 혁명을 중심으로 산업 전반에 디지털화와 자동화가 촉진되며 건축물의 생애주기 전반에 걸친 정보를 디지털화하는 BIM(Building Information Modeling)이 주목받기 시작했다. 본 논문은 3D LiDAR SLAM과 인공지능 기반 객체 탐지 모델을 이용하여 효율적 BIM을 위한 다중 센서 기반 Perception Torch를 제안한다. 제안하는 시스템을 통해 기존의 건축물 유지관리 시스템의 효율성이 높아질 것을 기대한다.

### Abstract

The Fourth Industrial Revolution has accelerated the digitalization and automation across various industries, highlighting the significance of Building Information Modeling (BIM), which digitalizes information throughout the entire life cycle of buildings. This paper proposes a multi-sensor-based Perception Torch utilizing 3D LiDAR SLAM and AI-based object detection models to enhance the efficiency of BIM processes. Through the proposed system, we anticipate significant improvements in the effectiveness of existing building maintenance systems.

### Key words

BIM, multi-sensor-based, perception, torch, 3D LiDAR SLAM, Object Detection, Object Localization

## 1. 서론

4차 산업 혁명은 다양한 혁신 기술을 기반으로 산업 전반에 디지털화와 자동화를 촉진했다. 특히 AEC(Architecture, Engineering & Construction) 분야에서는 건축물의 생애 주기 전반에 걸쳐 정보를 생성하고 관리하는 디지털 프로세스인 BIM(Building Information Modeling)이 주목받기 시작했다 [1].

본 논문에서는 SLAM (Simultaneous Localization

and Mapping) 및 객체인식을 통해 효율적 BIM을 위한 다중 센서 기반 Perception Torch를 제안한다. 3D LiDAR를 활용해 3D Point Cloud를 획득하고, RGB-D 카메라와 인공지능을 통해 건축물 내 객체를 인식하여 위치를 추정한다. 또한, Hand-held 디바이스를 제작함으로써, 이동체 로봇을 활용해 건물 내부 모델링을 진행할 때 발생할 수 있는 계단이나 좁은 복도에서의 충돌이나 낙하의 위험성 문제를 극복할 수 있다는 점에서 의미가 있다.

\* 국립금오공과대학교 전자공학부, hwkim@kumoh.ac.kr, yuwl@kumoh.ac.kr, woals3534@kumoh.ac.kr,

\*\* 국립금오공과대학교 IT융복합공학과, hlee@kumoh.ac.kr (corresponding author)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과로 수행되었음. (IITP-2024-RS-2024-00438430\*, 50%) (IITP-2024-RS-2024-00437190, 50%).

## II. 제안하는 시스템

전체적인 시스템 파이프라인은 그림 1과 같다. IMU가 내장된 LiDAR 데이터를 활용한 SLAM으로 지도를 작성하고, RGB-D 카메라의 데이터와 YOLO (You Only Look Once) [2] 모델을 활용한 객체인식으로 건축물 유지관리에 필요한 객체들의 위치를 추정하여 최종적으로는 객체가 표시된 3D Map을 얻을 수 있다.

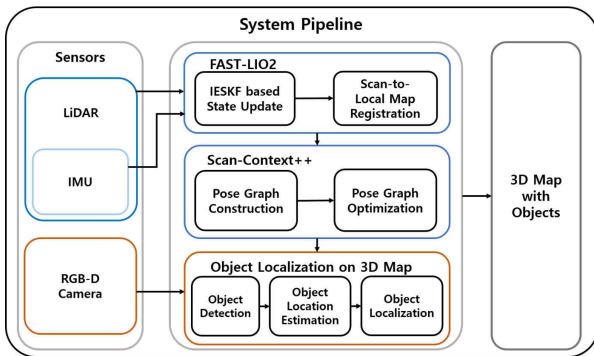


그림 1. 전체적인 시스템 구조도  
Fig. 1. Overall system architecture

실험에 사용된 온보드 PC는 Jetson AGX Orin Developer Kit이고, LiDAR는 Livox MID-360을 사용하였다. IMU는 해당 LiDAR의 내장 IMU를 사용하였으며, 객체 인식을 위해 사용된 카메라는 Intel RealSense Depth Camera D455이다. 그림 2는 자체적으로 제작한 Torch 디바이스의 구조이다. Torch 디바이스에 LiDAR와 RGB-D 카메라를 장착하여 진행하였고, 환경은 실제 건물이었다.

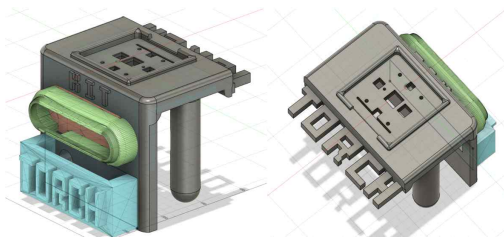


그림 2. Torch 디바이스 구조  
Fig. 2. Torch device structure

그림 3은 글로벌관 4층의 일부를 Mapping한 결과이며, 3D 지도 작성 결과를 보면 수평으로 정렬되어 있으며 안정적인 SLAM 결과를 보여준다. 또

한, 건축 관련 객체를 낮은 오차로 인식하여 이를 3D 지도 상에 정확히 나타낸 결과를 볼 수 있다.

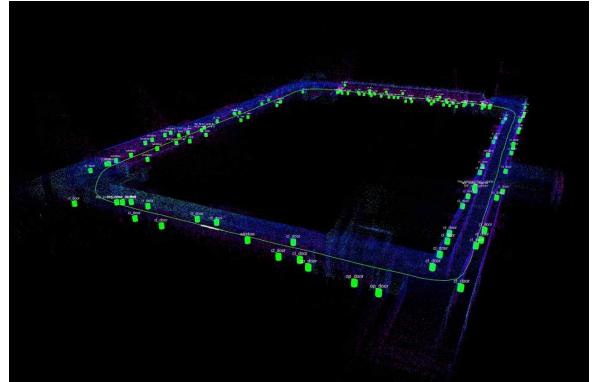


그림 3. 실제 건물에서의 SLAM 결과  
Fig. 3. SLAM Results in a Real Building

## III. 결론

본 논문에서는 효율적 BIM을 위한 다중 센서 기반 Perception Torch 시스템을 제안하였다. 제안된 방법은 3D LiDAR SLAM과 YOLO를 활용하여 효율적 BIM을 위해 객체 위치가 표시된 3D Map을 작성한다. 수작업으로 이루어지는 건축물 유지관리 방법에서 사용자가 별도의 수작업을 할 필요 없이 데이터가 수집되는 제안된 시스템을 통해 작업 효율을 높이고 시간 비용을 절감할 수 있을 것을 기대한다.

## 참고 문헌

- [1] SN. H. Ham, "A Study on Application of BIM(Building Information Modeling) to Pre-design in Construction Project", 2009. PhD Thesis. Hanyang University.
- [2] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).



# 다중 소형 로봇 시스템을 이용한 다중 객체 지도 작성 및 병합 기법

정영환\*, 백재욱\*\*, 최동규\*\*, 이헌철\*\*\*

## Multi-Object Map Building and Merging with a Multiple Small Robots System

YoungHwan Jung\*, Jaeuk Baek\*\*, Donggyu Choi\*\*, and Heoncheol Lee\*\*\*

### 요약

본 논문은 각각의 소형 로봇에서 얻은 객체 위치 정보들을 이용해 다중 객체 지도를 그리고 이 지도들을 병합하는 기법을 제안한다. 심층학습 기반 객체인식을 통해서 객체 정보들을 얻고 얻은 객체 정보들 중 객체 클래스를 기준으로 매칭 쌍을 형성하고 이 매칭 쌍들에서 최적의 변환을 찾아 지도 병합을 수행한다. 제안된 기법은 실제 다중 소형 로봇 시스템 기반 실험을 통해 성공적으로 수행됨이 확인되었다.

### Abstract

This paper proposes a technique for drawing multiple object maps and merging these maps using object location information obtained from each small robot. Obtain object information through deep learning-based object detection, form matching pairs based on object classes among the obtained object information, find the optimal transformation from these matching pairs, and perform map merging. The proposed method was tested and verified by experiments with real multiple small robot systems.

### 1. 서론

전장 환경에서 소형 로봇 시스템은 대형 로봇 시스템에 비해 저렴하고 간단하게 제작될 수 있다는 장점이 있다. 하지만 소형 로봇의 경우 무게의 제약으로 인해 용량이 작은 배터리를 사용할 수밖에 없어 운용 가능 시간이 짧고 저전력의 저성능센서를 사용해야만 한다는 단점이 있다. 운용 가능 시

간이 짧으므로 정보 수집 시간이 오래 걸리는 넓은 전장 환경에서는 다수의 로봇을 이용해서 각각의 환경 정보를 수집하는 것이 일반적이다. 그런데 저성능 센서는 다수의 노이즈로 인해 왜곡된 데이터를 많이 얻게 되는데 이로 인해 다수의 로봇이 각각 수집한 환경 정보에 오차가 발생하게 된다.

본 논문에서는 심층학습 기반 객체 인식을 통해 객체 정보를 얻고 객체의 클래스를 기반으로 하여

\* 금오공과대학교 전자공학부 (E-mail: 20171061@kumoh.ac.kr)

\*\* 한국전자통신연구원(ETRI) 국방ICT융합연구실 (E-mail: juback@etri.re.kr, dgchoi@etri.re.kr)

\*\*\* 금오공과대학교 전자공학부, IT융복합공학과 (E-mail: hcllee@kumoh.ac.kr)

※ 이 논문은 2022년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원(No. KRIT-CT-22-006-002, 초소형 군집로봇 상황/환경 인지 기술) 및 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업(IITP-2024-RS-2024-00437190, 50%)의 지원을 받아 수행된 연구임.

상관관계를 구성한 후 최적의 변환을 찾아 다중 로봇에서 수집한 지도를 병합하고 최근접 이웃 탐색 NN(Nearest Neighbor) 검사를 통해 병합의 성공률을 판단하여 최적의 병합을 찾는 방법을 제안한다.

## II. 제안하는 기법

다중 소형 로봇 시스템은 소형 로봇 3대와 중앙 PC로 구성되어 있으며 각각의 소형 로봇들은 고유한 로봇 ID를 가지고 있다. 각각의 로봇들은 서로 같은 공통의 객체도 보지만 서로 다른 동선으로 이동하거나 서로 다른 카메라 시야를 통해서 각각 공통 객체 이외에도 서로 다른 객체들을 보게 된다. 이때 로봇들은 심층학습 기반 객체 인식의 모델로 YOLOX(You Only Look Once X)를 사용하여 객체의 정보들을 얻는다. 이후 로봇 기준의 객체 위치를 얻기 위해서 2D SLAM(Simultaneous Localization and Mapping)을 수행한다. 이때 저성능 센서의 단점을 보완하고자 파티클 필터를 적용하여 단일 로봇에서 객체 위치추정의 성능을 높였다.

이렇게 로봇들은 각각 다중 객체의 정보를 얻고 네트워크 통신으로 연결된 중앙 PC에서 이 정보들을 수집한다. 전송되는 객체의 정보에는 위치 정보뿐만 아니라 객체의 클래스 정보 또한 포함된다. 이때 서로 다른 두 대의 로봇에서 보낸 객체 정보에서 같은 클래스를 가진 객체들의 수가 충분한 개수가 되면 이 같은 클래스를 기준으로 형성 가능한 모든 대응 매칭 쌍을 만든다. 그런 다음 기준이 될 로봇의 객체 위치 정보들을 소스로 지정하고 변환이 될 로봇의 객체 위치 정보들을 타겟으로 지정한다. 모든 매칭 쌍에 대해서 SVD(Singular Value Decomposition)을 사용하여 두 객체 정보들에 대한 최적의 변환을 찾고 이 최적의 변환을 타겟에 적용한다. 소스 객체와 변환된 타겟의 객체들을 최근접 이웃탐색 NN 검사 기반으로 병합에 성공했는지 판단하고 병합 성공률을 측정한다. 가장 높은 성공률을 보인 대응 매칭 쌍을 최적의 병합 경우로 판단하고 서로 병합된 지도를 출력하게 된다. 병합은 소스로 지정된 로봇 지도에 다른 두 대의 타겟 로봇 지도를 각각 병합해 먼저 두 개의 지도를 얻은 후 이 두 개의 지도를 최종 병합하여 병합된 지도

를 얻는다. 제안된 기법을 사용하여 실험한 결과 소스와 타겟1 소스와 타겟2를 병합해 얻는 두 개의 지도 및 최종 병합된 지도는 그림1, 그림2와 같다.

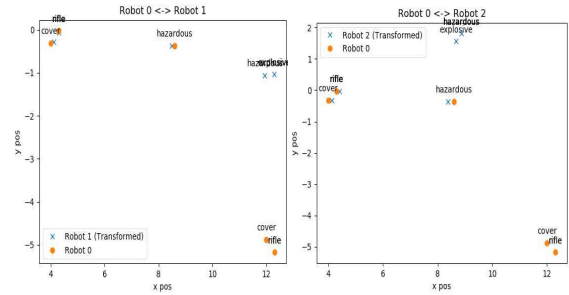


그림 1. 소스+타겟1, 소스+타겟2 병합 지도

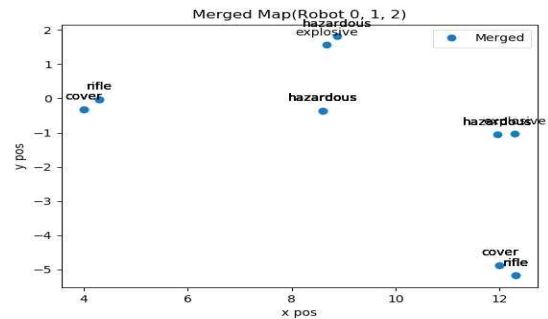


그림 2. 최종 병합된 지도

## III. 결 론

본 논문은 저성능 센서의 사용으로 노이즈가 포함된 제한된 정보를 얻는 상황에서 다중 소형 로봇에서 얻은 다중 객체 지도들을 병합하는 기법을 제안하였다. 향후에는 객체의 정보가 더 많은 복잡한 상황에서도 더 효율적으로 지도를 병합하는 방법을 연구하여 성능을 향상시키고자 한다.

## 참 고 문 헌

- [1] H. Lee and S. Lee, "Extended Spectra-Based Grid Map Merging With Unilateral Observations for Multi-Robot SLAM", IEEE Access, vol. 9, pp.79651-79662, 2021.
- [2] H. Lee, Y. Cho, and B. Lee, "Accurate map merging with virtual emphasis for multi-robot systems", Electronics Letters, vol. 49, no. 15, pp. 932-934, 2013.

# 머신러닝 기반 앙상블 모델 비교 및 소프트보팅을 이용한 전력 수요 예측

김도준\*

## Comparison of ensemble models based on machine learning and prediction of power demand using softvoting

Dojun Kim\*

### 요 약

현재 기후위기로 인한 극단적인 기후 현상에 따라 전력 수요의 변동성이 증가하고 있다. 이에 기후위기 문제를 해결하기 위해 재생에너지가 확대됨에 따라 전력망이 복잡해지고, 전력 공급이 일정하지 않아 수요 예측이 어려워지고 있다. 따라서 급격한 전력 수요 변동성에 대응하고, 전력 시스템의 안정적인 운영을 위해 정확한 수요 예측 연구가 필요하다. 본 논문에서는 RandomForest, XGBoost, LightGBM, CatBoost 네 모델을 통해 각 기상, 시간, 전력 수요 변수들을 학습시키고 일주일간 제주 전력 수요 예측을 한 후 오차율을 비교한다. 이후 오차율이 낮은 모델들을 SoftVoting 기법으로 결합해 다시 학습 및 예측 진행 후 기존 모델과 오차율 비교를 통해 가장 정확도가 높은 모델을 판단한다. 이를 통해 전력 수요의 정확도를 개선하여 전력 과잉 생산과 수요 급증으로 인한 문제를 예방하고 해결하는 데 토대가 되는 것을 목표로 한다.

### Abstract

The volatility of power demand is increasing due to extreme weather events caused by the current climate crisis. As renewable energy is expanded to solve the climate crisis, the power grid is becoming more complex and demand forecasting is becoming more difficult due to inconsistent power supply. Therefore, accurate demand forecasting research is needed to respond to the rapid variability of electricity demand and ensure stable operation of the power system. In this paper, four models, RandomForest, XGBoost, LightGBM and CatBoost, are used to learn each weather, time and power demand variable and compare their error rates after forecasting Jeju's power demand for one week. The models with the lowest error rates are then combined using the SoftVoting technique to learn and predict again, and the most accurate model is determined by comparing the error rates with the existing models. The aim is to improve the accuracy of power demand, laying the foundation for preventing and solving problems caused by overproduction and demand surges.

### Key words

Power Demand Prediction, Random Forest, XGBoost, LightGBM, CatBoost, SoftVoting

---

\*명지대학교 전기공학과, dojunwork@gmail.com

## 1. 서 론

전 지구적으로 전례 없는 기상이변 현상과 극한 기후로 전력 수요의 변동성이 더욱 커지고 있다. 이러한 기후 위기의 주요 원인인 화석연료 발전으로 인한 탄소 배출에 대응하기 위해 재생에너지가 확대되고 있다. [1] 하지만 재생에너지의 간헐성, 출력의 불안정성과 전력 공급과잉으로 인해 제주지역 같은 경우 재생에너지의 출력제어가 지속되며 논란이 되고 있다. [2] 또한 재생에너지와 같은 분산 에너지로 인해 전력망이 더욱 복잡해져 데이터 기반 전력 시스템의 필요성이 더욱 증대되고 있다. 이에 기후 위기로 인한 전력 수요의 급격한 변동을 해결하고 전력 시스템의 안정화를 위해 데이터 기반 전력 수요 분석과 정확한 전력 수요 예측 연구는 필수적이다.

본 연구에서는 머신러닝 기반 앙상블 모델인 RandomForest, XGBoost, LightGBM, CatBoost를 활용하여 여러 기상, 시간, 전력 수요 데이터를 학습시켜 일주일간의 전력 수요를 예측한다. 예측 결과를 통해 높은 성능을 보이는 모델을 SoftVoting 기법을 활용하여 결합 후 기존 모델과 비교해 더 정확한 모델을 판단한다. 이를 통해 전력 수요 예측 정확도와 전력 수급 안정성을 높일 수 있는 기반을 제공하고자 한다.

## II. 전력 수요 예측

### 2.1 데이터 수집 및 분석

전력 수요 예측을 위해서 표 1에 있는 데이터를 선정하여 수집했다. 전력 수요 예측의 기준 지역은 제주이며, 재생에너지의 비중이 높고 최근 출력제어가 빈번하게 발생한다는 문제가 있기에 전력 수요 예측의 연구와 적용이 필요하다고 생각되어 선정했다. 전력 수요 데이터는 공공데이터포털의 2022년부터 2023년간 한국전력거래소\_시간별 제주 전력 수요 데이터를 사용했으며, 기상 데이터의 경우 기상청 기상자료개방포털의 2022년부터 2023년까지의 제주 종관기상관측(ASOS) 자료를 사용했다.

표 1. 예측에 사용된 데이터

Table 1. Data used for prediction

데이터 분류	상세 관측 데이터
시간	연, 월, 시간, 연중 일수, 요일, 공휴일 여부
기상	기온, 일사, 풍속, 습도, 강수량
전력 수요	- 관측 시간의 전력 수요 - 7일 전 전력 수요 - 14일 전 전력 수요 - 21일 전 전력수요

제주 전력 수요 분포를 시간대, 월, 계절별로 분류해 전력 수요 변화를 먼저 분석하여 제주의 전력 소비 패턴을 확인했다. 이를 통해 전력 수요가 시간 변수에 큰 의존성을 보이는 것을 알 수 있다. 또한 연간 제주 전력 수요 변화 추이를 분석하여 기후위기로 인한 극한기후 현상과 밀접하게 연관되어 있을 것으로 예상하고 이후 기상과 전력 수요간 관계 분석을 진행했다. 그 결과 기온이 낮거나, 높을 때 전력 수요는 증가하는 경향을 보이는 것을 알 수 있다. 그림 1은 기온에 따른 전력 수요의 분포를 나타낸 것으로 기온과 전력 수요가 서로 연관관계를 가짐을 시각적으로 쉽게 파악해볼 수 있다.

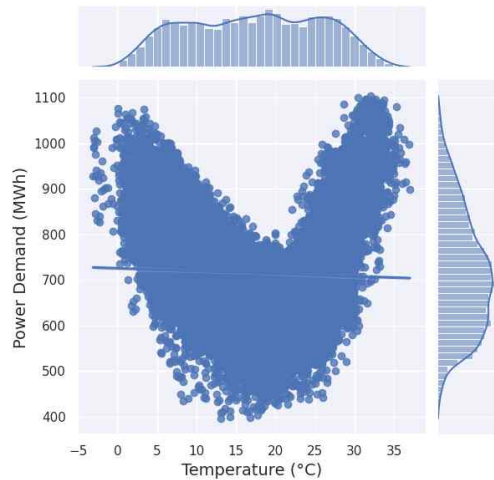


그림 1. 기온-전력수요 그래프

Fig. 1. Temperature-PowerDemand Graph

### 2.2. 모델 선정

본 논문에서는 먼저 RandomForest, XGBoost, LightGBM, CatBoost 네 가지 앙상블 모델을 이용하

여 제주 전력 수요 예측을 진행했다.

RandomForest 모델은 결정트리 기반의 앙상블 모델로 배깅(Bagging) 알고리즘을 사용하여 여러 개의 결정트리 모형으로부터 학습한 결과를 종합하여 하나의 결과를 도출하는 모델이다. [3] 이어서 언급할 세 모델은 배깅 알고리즘을 사용하는 RandomForest와 달리 그래디언트 부스팅 알고리즘에 기반한 모델이다. XGBoost(eXtreme Gradient Boosting) 모델은 트리 분할 과정에서 가지치기로 과적합을 방지하며, 최적의 구조를 찾아 성능을 높인다는 장점이 있으며, 메모리 최적화와 병렬처리로 대규모 데이터에서도 효과적이다. [4] LightGBM(Light Gradient Boosting Machine) 모델은 마이크로소프트사에서 개발했으며 리프 중심 트리 분할 방식을 사용하여 손실을 줄일 수 있는 쪽으로 리프를 분할해 성능을 높인다. 또한 빠른 학습이 가능하고 효율적인 메모리 사용으로 대용량 데이터 처리에도 적합하다. [5] CatBoost(Category Boosting) 모델은 범주형 데이터 처리에 특화됐으며 기본 하이퍼파라미터 최적화가 잘 되어 있어 기본적으로 우수한 성능을 보이고 과적합 방지를 위한 기법도 적용되어 있어 다른 데이터에서도 높은 성능을 유지한다. [6]

### 2.3. 모델 예측 비교 및 평가

본 논문에서는 먼저 RandomForest, XGBoost, LightGBM, CatBoost 각 네 모델을 사용하여 예측을 진행하였으며 수집한 데이터를 넣어 학습하는 과정에서 여러 번 GridSearch를 통해 최적의 하이퍼파라미터 조합을 찾아 튜닝하는 과정을 거쳤다. 이후 일주일간의 제주 전력 수요를 예측하여 정확도를 비교하였고 정확도가 높은 모델들을 성능향상과 과적합 방지를 위해 SoftVoting 기법을 사용해 결합하여 다시 학습 및 예측을 진행했다. 예측 결과를 기존 네 모델을 사용했을 때와의 각 오차를 비교하여 가장 성능이 높은 모델을 판단한다. 모델의 예측 정확도를 판단하는 지표로는 예측값과 실제값 간 절대적인 오차를 백분율로 표현하는 MAPE(Mean Absolute Percentage Error)를 사용하였으며 오차를 비율로 표현해 여러 모델의 성능을 직관적으로 비교할 수 있어 선택했다.

표 2. 예측 성능 비교 (MAPE 값)

Table 2. Predicted Performance Comparison (MAPE score)

모델	MAPE 값 (%)
RandomForest	3.24
XGBoost	2.57
LightGBM	2.37
CatBoost	2.39
SoftVoting	2.27

표 2는 각 모델의 학습 및 예측 후 MAPE 값을 비교하여 나타낸 것이다. SoftVoting을 제외한 각 네 모델의 예측 오차를 보면 RandomForest 모델이 상대적으로 높은 값을 보이며, 그 외 세 모델은 낮은 값을 보이는 것을 알 수 있다. 이에 표 2에서 보이는 SoftVoting의 MAPE 값은 상대적으로 예측 오차가 높은 RandomForest는 제외하고 비슷하게 낮은 MAPE 값을 보이는 XGBoost, LightGBM, CatBoost 모델을 선택하여 SoftVoting 기법으로 결합 후 예측한 결과로 기존 네 개의 앙상블 모델보다 더 높은 성능을 보이는 것을 확인할 수 있다.

그림 2는 본 전력 수요 예측 연구에서 가장 우수한 결과를 보이는 SoftVoting 모델의 일주일간 제주 전력 수요 예측 결과를 시각적으로 나타낸 것이며 회색 실선은 실제 수요, 검은색 점선은 모델의 예측값을 의미한다.

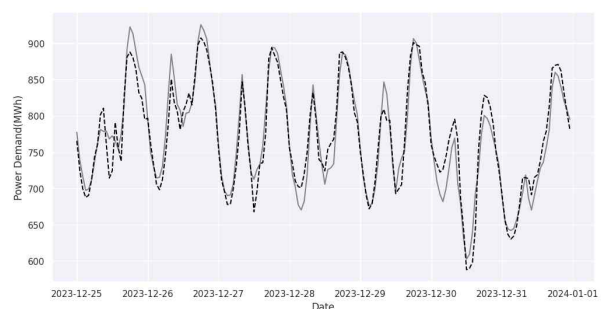


그림 2. SoftVoting 모델 예측값  
Fig. 2. SoftVoting Model Predicted Values

그림 3은 SoftVoting 이전 단독 앙상블 모델의 예측 중 가장 낮은 MAPE 값을 나타낸 LightGBM 모델의 입력 변수 중요도이다. 입력 변수 중요도는 각 변수의 트리 분할에서 사용 빈도를 바탕으로 평가되며, 이를 통해 모델의 예측에 중요한 역할을 하는

변수를 파악할 수 있다. 그림 3을 통해 대표적으로 연중 일수와 기온, 시간, 7일 전 전력 수요, 일사량이 모델의 주요 입력 변수임을 알 수 있다. 이를 통해 데이터 분석 단계에서 알 수 있었던 기온과 전력 수요의 밀접한 연관관계와 전력 수요의 시간의 의존성이 실제 모델에서 예측할 때도 주요 결정 요인으로 작용함을 확인할 수 있다. 이는 전력 수요의 예측 정확도를 높이기 위해서 해당 부분들을 주요하게 고려해야 함을 시사한다.

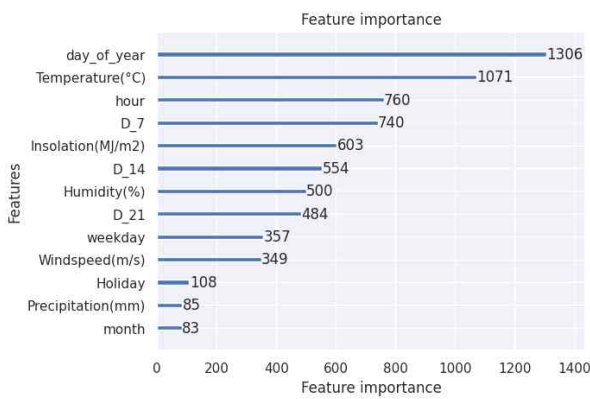


그림 3. 입력 변수 중요도 (LightGBM)  
Fig. 3. Feature Importance (LightGBM)

### III. 결 론

본 연구는 RandomForest, XGBoost, LightGBM, CatBoost 네 가지 모델의 일주일간 제주 전력 수요를 예측하고 그중 우수한 성능을 보이는 XGBoost, LightGBM, CatBoost 모델을 SoftVoting 기법을 이용해 결합하여 예측해 비교했다. 그 결과 기존 모델보다 세 모델을 SoftVoting 기법으로 결합한 모델이 가장 높은 성능을 보였으며, 전력 수요 예측에 적합함을 판단할 수 있다. 하지만 오차 값이 크게 차이가 나지 않는 것을 볼 때 RandomForest를 제외한 각 모델 간 정확도가 하이퍼파라미터의 튜닝과 데이터 분석의 목적과 특성에 따라 변화할 수 있으므로 적절한 모델을 선택하는 것이 중요하다.

데이터 분석과 각 모델의 입력 변수 중요도를 통해 전력 수요 예측에서 기온, 일사량과 전력 수요의 연관관계, 그리고 시간 요인에 영향을 받는 부분을 확인해볼 수 있었으며 앞으로 전력 수요 예측에서는 기상과 시간 데이터, 그리고 전력 수요 데이터를

주요 특징으로 적절히 분류해 선정하여 예측을 진행해야 한다.

본 연구를 통해 전력 과잉 생산으로 인한 출력제어나 수요 급증 등 수요와 공급의 불균형으로 발생하는 다양한 문제들을 전력 수요를 사전에 예측함으로써 이를 해결하는 데 기여하고 더 나아가 효율적인 전력제어를 통한 스마트그리드 산업의 발전과 더불어 불필요한 전력 사용을 줄여 기후위기 대응에도 도움이 되길 바란다. 또한 앞으로 더욱 정교한 전력 수요 예측과 전력 사용패턴 분석을 위한 하이퍼파라미터 튜닝, 모델 개선, 전력 수요 변수 분석 등 후속 연구가 지속되길 바라며, 본 논문이 안정적이고 효율적인 전력 공급을 위한 해당 연구의 기초 자료로 활용되기를 기대한다.

### 참 고 문 헌

- [1] 산업통상자원부, “재생에너지 3020 이행계획,” 2017.
- [2] 정선호, “제주지역 신재생 출력제어 최소화를 위한 플러스DR 제도 소개,” 전기의세계, vol. 71, no. 8, pp. 10-16, 2022.
- [3] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.
- [4] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, pp. 785-794, Aug. 2016.
- [5] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” Advances in Neural Information Processing Systems, vol. 30, pp. 3149-3157, 2017.
- [6] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: Unbiased Boosting with Categorical Features,” 32nd Conference on Neural Information Processing Systems, Montréal, Canada, vol. 31, pp. 6639-6649, Dec. 2018.

# 시각 장애인을 위한 개인 맞춤형 입력 지원 휴대용 점자 키보드

구진영\*, 김다진\*\*, 배준수\*, 엄승엽\*, 정아람\*, 이인수\*\*\*

## Personalized input support portable braille keyboard for the visually impaired

JinYeong Koo\*, Dajin Kim\*\*, Junsu Bae\*, Seungyeop Eom\*, Aram Jung\*, Insoo Lee\*\*\*

### 요 약

디지털 접근성은 모든 사회적 영역에서 필수적 요소로 자리 잡고 있지만, 시각 장애인에게는 여전히 큰 장벽이 존재한다. 본 논문에서는 LLM 기반의 개인화된 단어 추천과 SymSpell 알고리즘을 활용한 오타 교정 기능을 탑재한 개인 맞춤형 입력 지원 휴대용 점자 키보드를 제안한다. 이를 통해 시각적 피드백이 어려운 사용자에게 빠르고 정확한 입력 경험을 제공함으로써, 디지털 접근성 향상과 정보 격차 해소에 기여할 수 있을 것이다.

### Abstract

Digital accessibility has become an essential element across all social domains, yet significant barriers remain for visually impaired individuals. This paper proposes a personalized input-support portable braille keyboard equipped with LLM-based personalized word suggestion and SymSpell algorithm-powered typo correction features. By offering a fast and accurate typing experience to users who lack visual feedback, this study presents the potential to enhance digital accessibility and bridge the information gap.

### Key words

blind people, braille keyboard, word recommendation, LLM, typo correction

## 1. 서 론

디지털 기술이 모든 사회적 영역에서 빠르게 확산되면서, 정보 접근성의 중요성이 더욱 부각되고 있다. 특히 시각 장애인의 경우 디지털 접근성의 부족은 교육, 사회 활동, 그리고 경제적 기회에서 큰 장벽으로 작용하고 있다[1]. 스마트폰이나 태블릿과

같은 터치스크린 기반의 기기는 다양한 메뉴와 인터페이스가 시각적으로만 제공되기 때문에 시각 장애인의 입력 효율성이 떨어지며, 특정 기능을 정확히 찾고 선택하는 것이 쉽지 않다. 이러한 디지털 격차를 해소하기 위해 다양한 입력 방식이 제안되었으나, 대부분의 기존 점자 키보드는 부피가 크고, 휴대성과 편의성이 떨어지거나 단순 입력 장치에

\*경북대학교 전자공학부, \*\*경북대학교 컴퓨터학부, \*\*\*경북대학교 전자전기공학부 교수 (교신저자) insoolee@knu.ac.kr

불과하다는 문제가 있었다.

본 논문에서는 이러한 문제를 해결하기 위해 인공지능(AI) 기술을 활용한 초경량 휴대용 점자 키보드를 제안한다. 이 키보드는 한 손으로도 쉽게 입력이 가능하며, 개인화된 언어 체계를 반영하여 사용자 맞춤형 단어 추천과 오타 교정 기능을 제공하는 점에서 기존 키보드와 차별화된다. 특히, LLM(대규모 언어 모델) 기반의 단어 추천 기능과 오타 교정 기능은 사용자의 언어 패턴을 반영한 맞춤형 시스템으로 사용자의 입력 피로를 줄여주고, 쉽게 의사소통할 수 있도록 지원한다.

## II. 적용 기술

### 2-1 점역 알고리즘

점역 알고리즘은 점자를 한글로 변환한다. 한글 점자는 63개의 점자로 모든 한글을 표현하기 위해 각 글자의 앞, 뒤의 문자에 따라 같은 점형이라도 다른 문자를 표현한다. 이를 구현하기 위해 번역될 한글 유니코드 값을 64개 행을 가진 2차원 불규칙 배열로 관리하였다. 키보드로부터 입력받은 2진 배열 형태의 점자는 0~63까지의 정수 형태로 변환되고 각 값과 같은 인덱스의 배열 요소는 점형에서 번역될 수 있는 한글 유니코드 값들의 배열로 구성된다. 각 문자는 유니코드 값에 따라 초성, 중성, 종성, 약어 4개의 클래스 중 하나로 설정된다. 매 키보드 입력마다 정수 형태의 점형, 배열을 통해 변환된 한글 유니코드 값, 변환된 문자의 클래스로 구성된 구조체를 버퍼에 저장한다. 다음 입력에서 변환될 한글 유니코드 값은 입력된 점자와 같은 인덱스에 저장된 배열 요소 중에서 버퍼에 저장된 이전 입력 점형, 클래스를 조건으로 점자 규칙[2]에 따라 결정된다.

### 2-2 LLM과 Lang Chain RAG

LLM (Large Language Model)의 few-shot prompt와 RAG (Retrieval-Augmented Generation)를 결합하여 단어 추천 기능을 구현하였다. LLM은 사전 학습된 모델로, 사용자로부터 주어진 몇 가지 예시만으로도 (few-shot) 적절한 응답을 생성할 수 있다 [3]. RAG

는 정보 검색과 언어 모델을 통합하여, 모델이 외부 데이터베이스에서 관련 정보를 검색하고 이를 바탕으로 더 정확한 응답을 생성하도록 돕는 기법이다. RAG는 기존의 언어 모델들이 가지는 정보를 제한성을 극복하는 데 기여한다 [4]. Vector Database는 고차원 데이터를 저장하고 검색하는 데 최적화된 데이터베이스로, 유사한 맥락을 빠르게 찾을 수 있는 기능을 제공한다. 본 연구에서는 Chroma DB를 사용하여 사용자 입력 데이터베이스를 구축하였다. 또한, 다양한 자연어 처리 도구 통합 프레임워크인 LangChain을 활용하였다. 주요 문제는 실시간성이 중요한 키보드 시스템에서 단어 추천의 응답 시간이 느리다는 것과 인터넷 연결이 필요하다는 것이다. 이에 대한 대안으로 Trie 자료구조를 활용한 단어 추천 알고리즘을 제공할 예정이다. Trie 자료구조는 단어를 효율적으로 저장하고 검색하는 데 유용하다.

### 2-3 SymSpell

오타 교정 기능은 사용자 입력 오류를 자동으로 탐지하고 교정을 제안함으로써 텍스트 입력 경험을 개선하는 것을 목표로 한다. 본 연구에서는 SymSpell 알고리즘을 활용하여 빠르고 효율적인 오타 교정을 구현하였다. SymSpell [5]은 사전에 가능한 모든 오타 후보를 생성하여 저장하고, 편집 거리 기반 검색을 통해 빠르게 교정을 수행하며, 대규모 데이터나 높은 연산 자원을 요구하지 않아 휴대용 장치에서 사용하기 용이하다. 또한, 본 연구의 대상 언어인 한국어는 초성, 중성, 종성으로 구성된 자모 구조를 지니고 있어 단순한 문자 비교만으로는 오타 교정에 한계가 있다. 이를 해결하기 위해 자모를 초성, 중성, 종성 단위로 분해하고 각 단위에서 편집 거리를 계산하는 방식을 적용하였다. 이러한 접근법은 자모 단위에서의 유사성을 분석함으로써 한국어의 언어적 특성을 반영한 정밀한 교정을 가능하게 하며, 그 결과 오타 교정의 정확도와 신뢰성을 크게 향상시킨다.

## III. 제안한 점자 시스템 구현



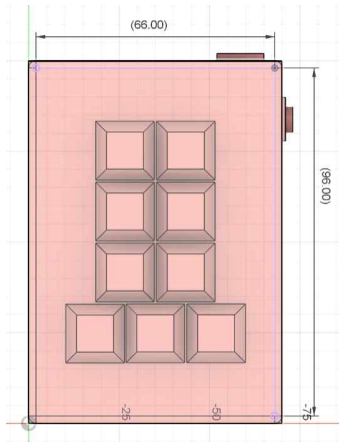


그림 1. 외형도  
Fig. 1. Outline view

그림 1은 점자 키보드의 외형 구조를 나타낸 것이다. 입력부는 6개의 점자 형태 키보드, 4개의 기능 키보드로 구성되며 출력부는 스피커로 구성된다. 전원 스위치를 통해서 전원을 켜면 건전지의 전압이 컨버터 (MT3608)를 통해서 알맞은 전압레벨로 변환되어 라즈베리파이 제로2W에 인가된다. 이후 키보드는 키 입력을 받는 대기 상태에 들어가게 되고, 사용자는 6개의 점자 모양 키보드를 이용하여 점자를 입력하게 된다.

그림 2는 키보드의 신호 처리 회로도로 Input에서 보내는 전기신호가 어떤 Output으로 나오는 지에 따라서 어떤 키가 입력되었는지 알 수 있다. 이 때 매트릭스 구조를 차용하여 키보드 회로를 구성하였다. 매트릭스 구조는 사용되는 스위치 개수 대비 적은 수의 케이블로 키 입력을 받을 수 있어 효과적이다. 또한, 매트릭스 구조 사용 시 발생할 수 있는 동시 입력 중복 문제인 고스트 현상을 방지하기 위해 각 스위치에 다이오드를 추가하였다[6].

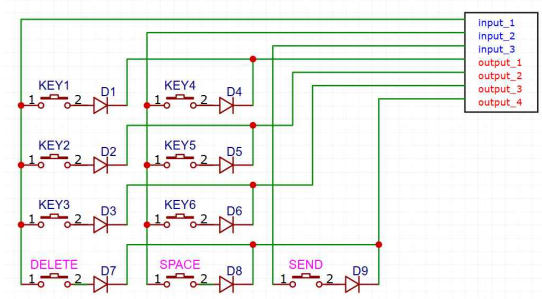


그림 2. 키보드 회로도  
Fig. 2. Keyboard circuits

점자 정보는 2진 배열로 입력되고 이를 점역 알고리즘을 통해 한글로 변환한다. Send 버튼 입력 시 변환된 문자가 사용자의 기기로 전송된다. 이 때 전송된 문자는 스피커를 통해 출력되어 사용자에게 어떤 단어가 전송되었는지 알린다.

문자는 USB HID report를 통해 전송된다. USB HID report descriptor는 USB HID Spec [7][8] 을 참고하여 표 1과 같이 작성하였다.

표 1. USB HID report descriptor  
Table 1. USB HID report descriptor

Usage (Keyboard), Collection (Application), Usage Page (Keyboard), Usage Minimum (224), ; LeftControl Usage Minimum (225), ; LeftShift Logical Minimum (0), Logical Maximum (1), Report Size (1), Report Count (2), Input (Data, Variable, Absolute), ; Modifier
2-bit
Report Size (6), Report Count (1), Input (Constant), ; Pad Report Size (8), Report Count (1), Input (Constant), ; Reserved byte Usage Minimum (4), ; a Usage Maximum (44), ; Spacebar Logical Minimum (0), Logical Maximum (40), Report Size (8), Report Count (6), Input (Data, Array), ; Key arrays (6 bytes)
End Collection

표 1의 report descriptor에 따라 report는 표 2와 같이 8-byte로 구성된다.

표 2. USB HID report  
Table 2. USB HID report

Byte	Description
0	Modifier 2-bit + Pad 6-bit
1	Reserved
2-7	Key code 1-6

문자 전송 및 삭제를 위한 Modifier와 Keycode 범 위만을 유효한 값으로 설정하였다. 각 문자에 맞는 report를 할당하여 글자를 전송한다. Single Board Computer를 사용하므로 장치를 Host가 아닌 USB HID device로 인식시키기 위해서 USB HID gadget을 설정해야한다. Linux 운영체제를 사용하여 Linux configs를 통해 USB HID Gadget을 설정하였고 설정 스크립트를 Systemd service로 등록하여 부팅 시 실행하도록 하였다.

SPACE키가 입력되면 Symspell 알고리즘을 이용한 오타 감지 프로그램이 실행된다. 입력 단어와 사전의 자모를 분해하고 편집거리를 비교함으로써 보다 정확한 교정 단어를 제안하도록 구성된다. 알고리즘을 통해서 오타가 감지되면 교정된 문자를 음성으로 출력하여 알리고 사용자는 오타 교정 여부를 선택한다.

그림 1의 제품 외형도 우측 상단에 위치하는 스위치를 통해서 단어 추천 기능을 사용할 수 있다. 점자가 입력된 상태에서 해당 스위치를 누르면 그림 3의 단어 추천 시퀀스가 동작하게 된다. 기존 입력되어 있는 단어를 Gemini API를 통해서 자동으로 완성시키고 해당 목록을 불러오게 된다.

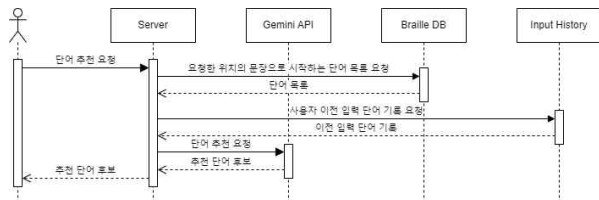


그림 3. 단어 추천 시퀀스 다이어그램  
Fig. 3. Word recommendation sequence diagram

이 때 Langchain RAG를 사용하여 기존에 사용자가 입력했던 단어들을 우선적으로 추천한다. 이처럼 단어 추천 기능은 사용자의 입력 습관과 자주 사용하는 어휘를 학습한다. 이를 통해 사용자는 타이핑 환경, 기기와 무관하게 일관된 맞춤형 추천 기능을 사용할 수 있다.

추천되는 단어 후보들은 차례대로 음성 출력되며 사용자는 출력되는 음성을 듣고 원하는 단어를 선택할 수 있다. 이 때 각 점자에 할당된 키 번호를 통해서 선택한다. 추천 단어를 선택하지 않으면 다

시 입력 대기 상태로 돌아가게 된다. 그림 4는 점자 키보드의 전체 순서도를 나타낸 것이다.

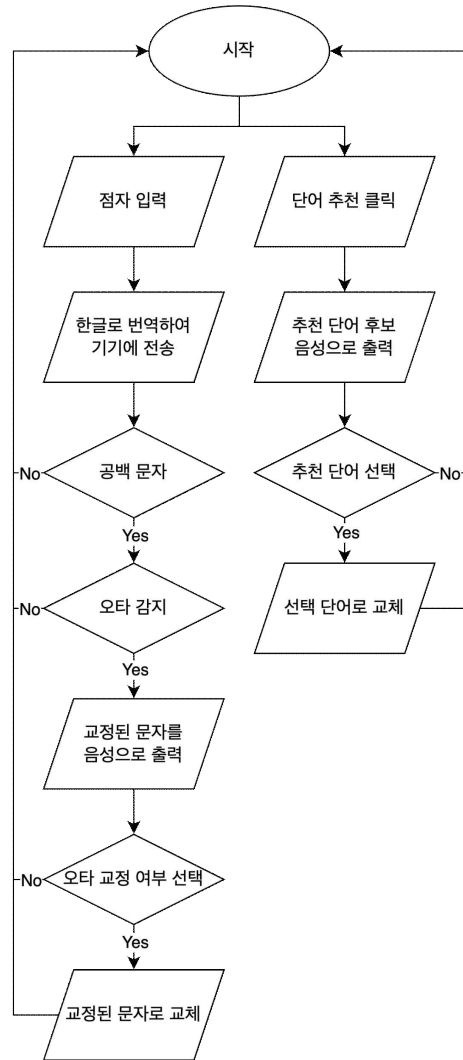


그림 4. 순서도  
Fig. 4. Flow chart

#### IV. 결론

본 논문에서는 시각 장애인의 디지털 접근성을 개선하기 위한 AI 기반 휴대용 점자 키보드를 제안한다. LLM(대규모 언어 모델) 기반 단어 추천 기능은 사용자의 입력 습관을 학습하여 개인화된 어휘 추천을 실시간으로 제공하며, 오타 교정 기능은 Symspell 알고리즘과 자모 분해를 결합해 빠르고 정확하게 오류를 탐지하고 교정한다. 이를 통해 본 연구의 키보드는 시각적 피드백 없이도 사용자에게

정확한 입력 경험을 제공한다. 또한, 단순한 입력 장치를 넘어 사용자의 언어 습관에 맞춘 지능형 타이핑 경험을 제공하는 도구로 자리 잡을 가능성을 보여주었다.

향후 연구에서는 더 많은 언어와 다양한 입력 환경에서의 확장성을 고려하여 시스템의 정확성과 실용성을 더욱 높일 수 있을 것이다. 본 연구의 결과는 시각 장애인의 정보 접근성을 증진시키고, 디지털 격차를 해소할 수 있을 것으로 기대한다.

## 참 고 문 헌

- [1] Hwang Ju-hee, "The Effect of Digital Divide on Life Satisfaction of Disabled People," 정보화정책, vol. 26, no. 3, pp. 53-68, 2019.
- [2] 국립국어원, 한국어 점자 규정, 국립국어원, 2024.
- [3] Tom B. Brown, "Language models are few-shot learners," arXiv preprint arXiv:2005.14165, 2020.
- [4] Patrick Lewis et al., "Retrieval-augmented generation for knowledge-intensive nlp tasks," arXiv preprint arXiv:2009.10793, 2020.
- [5] W. Wolfgarbe, "SymSpell: 1 million times faster spelling correction algorithm," GitHub repository, 2017.
- [6] "Understanding the Switch Matrix," Rec.Games.Pinball, Coinop Cauldron, May 14, 1996.
- [7] "Universal Serial Bus (USB) Device Class Definition for Human Interface Devices (HID) Firmware Specification," Version 1.11, May 27, 2001.
- [8] "HID Usage Tables for Universal Serial Bus (USB)," Version 1.5, January 26, 2024.

# 딥러닝을 활용한 리셀 스니커즈 가격 예측 및 SHAP을 통한 요인 분석

김은수\*, 김세연\*, 한규림\*, 김은혜\*

## Predicting Resale Sneaker Prices Using Deep Learning and Factor Analysis with SHAP

Eunsoo Kim\*, Seyeon Kim\*, Gyulim Hahn\*, and Eunhye Kim\*

### 요약

스니커즈 리셀 시장에서 가격은 희소성, 수요 등 다양한 요인에 의해 변동되며, 이러한 불확실성은 리셀 시장 참여자들에게 잠재적 위험 요소로 작용하고 있다. 본 연구에서는 Kream 플랫폼의 거래 데이터와 검색량 지표를 활용하여 스니커즈 리셀 가격을 예측하였다. LSTM과 GRU의 순차적 계층 구조를 활용한 딥러닝 모델을 구축하였으며, SHAP 분석을 통해 가격 변동 요인의 상대적 중요도를 파악함으로써 리셀 시장 참여자들의 의사 결정에 기여할 수 있을 것으로 기대된다.

### Abstract

In the sneaker resell market, prices fluctuate due to various factors such as scarcity and demand, and this uncertainty acts as a potential risk factor for market participants. This study predicts sneaker resell prices using transaction data and search volume indicators from the Kream platform. We constructed a deep learning model using a sequential layer structure of LSTM and GRU, and through SHAP analysis, we identified the relative importance of price fluctuation factors. This research is expected to contribute to decision-making for resell market participants.

### Key words

Sneaker Resell Market, Deep Learning, Sequential Hierarchical Structure, Price Prediction, SHAP

## 1. 서론

최근 리셀 시장은 한정판 스니커즈에 대한 수요 증가와 함께 빠르게 성장하고 있다. 스니커즈에 대

한 소비자들의 높은 관심으로 스니커즈 재테크가 활성화되었으며, 희소성이 있는 스니커즈를 구매한 후 더 높은 가격에 판매하여 이익을 얻는 소비자들이 늘고 있다. 이러한 리셀 문화는 전통적 소비 개념을 넘어 자산으로서의 가치에 기반한 새로운 거

\*서울여자대학교 데이터사이언스학과 학부생, lyra8@swu.ac.kr

\*서울여자대학교 데이터사이언스학과 학부생, syeon.kim77@gmail.com

\*서울여자대학교 데이터사이언스학과 학부생, onegyulim@swu.ac.kr

\*서울여자대학교 데이터사이언스학과 학부생, eunevesves@gmail.com

래 형태로 주목받고 있다[1].

국내 리셀 시장에서 Kream은 희소성 높은 스니커즈를 안전하게 거래할 수 있도록 지원하며 거래의 투명성과 안전성을 보장하기 위해 제품 검수 서비스를 제공하는 대표적인 리셀 플랫폼이다. 또한, 실제 거래 가격을 기반으로 한 시세 정보를 공개하여 시장 참여자들이 리셀 가격을 예측할 수 있도록 돕고 있다.

리셀 가격은 제품의 희소성, 시장 반응 등 다양한 요인에 따라 크게 변동할 수 있어, 가격 예측의 불확실성이 크다[2]. 불확실성은 리셀 시장 참여자들에게 잠재적 위험 요소로 작용하고 있으며, 리셀 스니커즈의 가격에 가장 많은 영향을 미치는 요인이 무엇인지 파악하는 것이 중요해지고 있다.

따라서 본 연구는 스니커즈 리셀 가격 변동의 주요 요인을 이해함으로써 참여자들이 더 합리적인 의사 결정을 할 수 있도록 도움을 주고자 한다. 구체적으로, 딥러닝을 활용해 검색량 지표와 거래내역과 같은 시계열적 요소를 반영하여 리셀 플랫폼에서 거래되는 스니커즈의 리셀 가격을 예측하고, SHAP(Shapley Additive Explanations)을 통해 가격에 영향을 주는 요인을 분석하는 것을 목표로 한다. 이를 통해 참여자 관심도를 반영한 주요 요인을 정교하게 파악하고, 모델의 예측력과 해석력을 강화하고자 한다.

## II. 연구 방법

Kream에서 인기순으로 정렬된 아디다스 상품 1,000개의 원시 데이터를 수집하였다. 데이터 셋은 상품 정보, 상품 거래 내역으로 구성되어 있으며, 총 370,745 건의 거래 데이터를 수집하였다. 상품 정보는 상품명, 모델 번호, 발매가, 발매일, 색상, 관심 상품 수, 스타일 수, url, 상품 번호, 거래 건 수, 구매 희망자 수, 판매 희망자 수로 구성되어 있고, 상품 거래는 모델 번호, 사이즈, 거래 날짜, 거래 가격, 배송 타입, url로 구성되어 있다. 스타일 수는 Kream에서 상품 구매자가 능동적으로 작성한 리뷰의 수를 의미한다. 또한, 네이버 오픈API를 활용하여 참여자의 관심도를 반영하는 검색량 지표인 스니커즈 상품명에 대한 검색량 데이터를 수집하였다.

수집된 데이터 셋은 모델의 고유번호를 기준으로 통합한 뒤, 한국 사이즈로 통일하고 배송 타입, 콜라보 여부 등의 속성을 추가하는 전처리를 진행하였다.

본 연구는 스니커즈 리셀 가격 예측을 위한 LSTM과 GRU의 순차적 계층 구조를 제안한다. LSTM(Long Short-Term Memory)은 입력, 출력, 망각 게이트를 통해 시계열 데이터의 중요 정보를 선별적으로 저장하고 업데이트하는 메모리 기반의 순환 신경망 구조이다[3]. GRU(Gated Recurrent Unit)는 업데이트 게이트와 리셋 게이트만을 활용하여 LSTM의 메모리 메커니즘을 단순화하면서도 효과적인 시계열 패턴 학습이 가능한 구조이다[4].

LSTM은 복잡한 장기 의존성을 포착하여 초기 특성을 추출하지만, 다수의 게이트로 인해 계산 비용이 높아 학습이 느려지는 단점이 있다. 반면, GRU는 업데이트 게이트와 리셋 게이트를 사용해 필요한 정보만 간결하게 유지하고 불필요한 정보를 효율적으로 제거하여 LSTM의 계산 복잡성을 줄이고 빠른 학습을 가능하게 한다. 이와 같은 계층적 특성 추출 과정은 각 모델의 장점을 극대화하며, 최종적으로 각 레이어의 가중치가 상호 보완적으로 조정되어, Dense 레이어에서 예측에 최적화된 특성으로 변환된다.

## III. 연구 결과

모델의 성능은 RMSE(Root Mean Squared Error), MAE(Mean Absolute Error), R<sup>2</sup>(R-Squared) 지표를 사용하여 평가하였다. 각 지표의 계산 방법은 다음과 같다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{1}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{3}$$

$y_i$  = 리셀가  
 $\hat{y}_i$  = 예측된 리셀가  
 $n$  = 데이터 포인트 수  
 $\bar{y}$  = 리셀가의 평균

RMSE(수식 1)는 예측된 리셀가와 리셀가 간의 차이의 제곱 평균을 취하여 전체적인 예측 오차 크기를 보여준다. MAE(수식 2)는 예측 오차의 절댓값 평균을 나타내며, 예측된 리셀가가 리셀가와 얼마나 일치하는지를 직관적으로 평가할 수 있다. R<sup>2</sup>(수식 3)값은 모델의 설명력을 나타내며, 리셀가의 변동성을 모델이 얼마나 잘 설명하는지 평가한다.

표 1. 모델 평가지표  
 Table 1. Model Evaluation

Evaluation Method	Value
MAE	13332.02
RMSE	16561.66
R-squared	0.9674

본 연구의 모델은 R<sup>2</sup> 값이 0.9674로, 모델의 높은 설명력을 입증하였으며, RMSE값은 16561.66, MAE 값은 13332.02로 예측 오차가 작음을 확인할 수 있었다.

모델의 해석력을 높이기 위해 SHAP(SHapley Additive exPlanations) 모델을 사용하여 각 특징이 모델 예측에 미치는 영향을 분석하였다. SHAP은 각 변수가 모델의 출력에 미치는 영향력을 정량화하여 직관적인 시각화를 가능하게 하며, 이를 통해 데이터의 의미를 보다 명확히 이해할 수 있도록 돕는다.

변수 중요도 분석에서 변수의 위치는 종속 변수(리셀 가격)에 대한 영향을 나타내며, 상위에 위치할수록 영향력이 큼을 의미한다. 또한, SHAP 분석 그래프에서 값이 중앙으로부터 오른쪽에 위치할 경우 양의 상관관계를, 왼쪽에 위치할 경우 음의 상관관계를 나타낸다. SHAP 분석의 색상은 빨간색일수록 변수의 영향력이 크고, 파란색일수록 영향력이 낮음을 의미한다[5].

그림 1은 SHAP 모델 결과를 시각화하여 변수 중요도를 파악한 것이다. 그 결과, released\_price(발매가), style\_count(스타일 수), sell\_count(판매 희망자

수), wish\_count(관심 상품 수), collaboration(콜라보 상품 비율:18.9%) 순으로 스니커즈 리셀가에 영향을 주는 것으로 확인되었다.

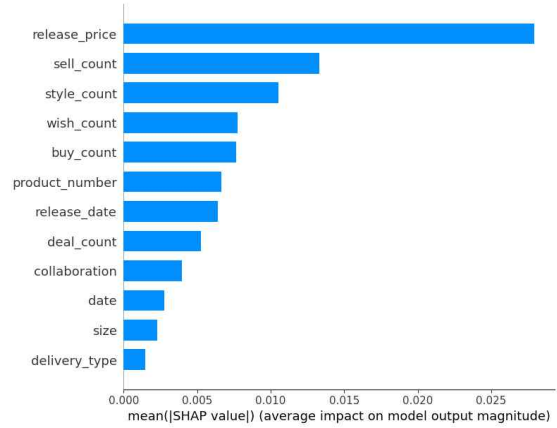


그림 1. 변수중요도 분석 결과  
 Fig. 1. Result of feature importance analysis

그림 2는 SHAP 분석 결과를 시각화한 것이다. 발매가는 스니커즈 리셀 가격에 가장 높은 SHAP value를 보였으며, 발매 가격이 높을수록 양의 상관관계를 가지는 것을 확인하였다. 참여자들이 해당 상품의 리셀 초기 가격을 설정할 때 의지할 수 있는 직접적인 지표가 발매가이기 때문에 리셀 가격이 형성될 때, 발매가가 가장 큰 영향을 미치는 요인임을 알 수 있다. 사용자들이 능동적으로 작성해야 하는 스타일 수가 많을수록 예측 가격에 양의 상관관계를 가지는 것을 확인하였다. Kream 플랫폼의 특성상, 스타일 등록은 별도의 리워드 없이 사용자가 자발적으로 사진과 글을 게시해야 하기 때문에 해당 상품에 대한 참여자의 관심을 나타내며, 이는 수요에도 영향을 주는 요인임을 알 수 있다.

판매 희망자 수가 적을수록 높은 리셀 가격을 가지는 것을 확인하였다. 판매 희망자가 많으면 과잉 공급으로 인해 스니커즈 리셀 가격이 하락할 수 있으며, 이는 리셀 시장에서 상품의 공급이 중요한 지표임을 알 수 있다. 반면, 구매 희망자 수가 많을수록 높은 리셀 가격을 가지는 것을 확인하였다. 구매 희망자 수는 실질적인 구매자 수요를 나타내는 직관적인 지표로, 수요가 많을수록 리셀 가격이 상승함을 의미한다.

발매일이 오래될수록 낮은 리셀 가격을 가지는

것을 확인하였고 최근에 발매된 경우, 높은 리셀 가격에 영향을 주는 것을 확인하였다. 이는 시간이 지남에 따라 제품이 인기를 잃거나 중고 시장에 충분히 공급되어 가격을 안정시킬 수 있음을 시사한다.

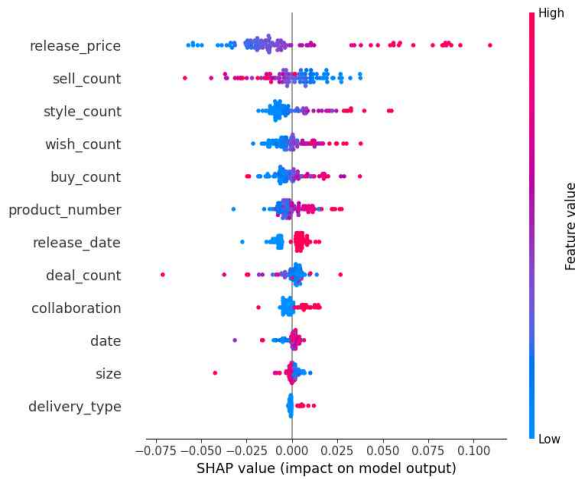


그림 2. SHAP 분석 결과  
Fig. 2. Result of SHAP analysis

### III. 결론 및 시사점

본 연구는 온라인 리셀 플랫폼 Kream에서 스니커즈 가격 예측을 위한 LSTM과 GRU의 순차적 계층구조 모델을 제안했다. 리셀 가격을 예측한 결과, 모델의  $R^2$  값이 0.9674로 높은 설명력을 입증했다. 또한, SHAP value를 통해 발매가, 스타일 수, 판매 희망자 수 등이 리셀 가격 예측에 중요한 영향을 미치는 것으로 확인되었다.

국내 리셀 시장이 커지고 있으며, 리셀 시장에 유입되는 참여자도 증가할 것으로 예상된다. 가격에 영향을 주는 요인 분석을 통해 참여자들에게 데이터 기반 의사결정이 가능하도록 돕는다. 향후 연구에서 다양한 외부 변수를 도입하여 리셀 시장의 복잡성을 더욱 심층적으로 분석한다면 모델의 예측력 및 해석력을 더욱 높일 수 있을 것으로 기대된다.

### 참고 문헌

[1] Juha Park, & Jaehoon Chun. (2023). The Meaning of Resell Activities Using the Online Second-hand Platform. *Journal of the Korean Society of*

*Clothing and Textiles*, 47(5), 822-838.

- [2] Yoon, H., & Kang, J. (2021). A Study on Resolving Barriers to Entry into the Resell Market by Exploring and Predicting Price Increases Using the XGBoost Model. *The Journal of Society for e-Business Studies*, 26(3), 155-174.
- [3] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.
- [4] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN Encoder-Decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*
- [5] Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *arXiv preprint arXiv:1705.07874*.

# 스마트미러를 활용한 ASD 아동의 감정 모니터링 시스템 개발

이도형\*, 임치현\*\*, 정현철\*\*\*, 박기태\*\*\*\*, 주종화\*\*\*\*\*

## Development of an Emotion Monitoring System for Children with ASD with Using a Smart Mirror

Do Hyeong Lee\*, Chi Hyeon Lim\*\*, Hyeon Cheol Jeong\*\*\*, Ki Tae Park\*\*\*\*,

and Jong Wha Ju\*\*\*\*\*

### 요약

자폐 스펙트럼 장애(ASD)를 가진 아동들은 감정 이해와 표현에 어려움을 겪어 사회적 상호작용과 의사소통에 제약을 받는다. 본 연구는 자폐 스펙트럼 장애(ASD) 아동을 위한 스마트 미러 기반 실시간 감정 모니터링 시스템을 제안한다. 이 시스템은 아동이 스마트 미러에서 자신의 이름을 선택후 사진을 촬영하면, EfficientNet 기반 AI 모델이 감정을 분석하여 친숙한 아이콘과 부엉이 캐릭터로 시각화된 피드백을 제공한다. 감정 분석 결과는 Firebase Cloud Messaging(FCM)을 통해 보호자에게 실시간으로 전송되며, 보호자는 전용 앱에서 주간·월간 감정 변화를 모니터링할 수 있다. 이를 통해 아동의 정서 발달을 지원하고 보호자가 손쉽게 아동 정서 확인을 할 수 있는 기반을 제공한다.

### Abstract

Children with Autism Spectrum Disorder (ASD) experience difficulties in understanding and expressing emotions, which limits their social interactions and communication. This study proposes a real-time emotion monitoring system based on a smart mirror for children with ASD. In this system, when a child selects their name on the smart mirror and takes a photo, an EfficientNet-based AI model analyzes the emotion and provides visual feedback through familiar icons and an owl character. The emotion analysis results are transmitted to caregivers in real-time via Firebase Cloud Messaging (FCM), allowing caregivers to monitor weekly and monthly emotional changes through a dedicated app. This system supports children's emotional development and provides an accessible foundation for caregivers to easily track and understand their child's emotional state.

### Key words

Autism Spectrum Disorder(ASD), Smart Mirror, Emotion Monitoring, Face Recognition, Gamification

---

\* 동국대학교 컴퓨터공학과, rrlfhtm2@dgu.ac.kr  
\*\* 동국대학교 컴퓨터공학과, lyc3004@dongguk.edu  
\*\*\* 동국대학교 컴퓨터공학과, startgm99@dgu.ac.kr  
\*\*\*\* 동국대학교 컴퓨터공학과, gitae9999@dongguk.edu  
\*\*\*\*\* 동국대학교 컴퓨터공학과, jwjoo@dgu.ac.kr (교신저자)

※ This research was supported by the Software-Centered University Project (2023-0-00049) and the University ICT Research Center Promotion Project (IITP-2024-2020-0-01789) of the Ministry of Science and ICT and the Institute for Information & Communications Technology Planning & Evaluation (IITP) in 2024.



## I. 서 론

자폐 스펙트럼 장애(ASD) 아동은 감정 인식과 표현에 어려움을 겪어 사회적 상호작용과 삶의 질에 큰 영향을 받는다. 이에 따라, ASD 아동의 감정 인식과 표현을 돕는 기술적 도구의 필요성이 강조된다. 기존 연구에 따르면, 다중양식 감정 훈련은 다양한 감각 자극을 결합해 감정 표현과 사회성 향상에 효과적이었으며[1], 증강현실(AR) 기반 중재는 동적 이미지를 통해 감정 표현 능력과 일반화에 긍정적인 영향을 미쳤다[2]. 또한, 시각적인 치료는 자기표현을 통해 정서적 안정과 긍정적인 자아감 형성에 기여한다[3]. 본 연구에서는 이러한 효과에 기반해 스마트 미러 기반의 실시간 감정 모니터링 시스템을 제안한다. 이 시스템은 얼굴 인식과 감정 분석으로 아동의 감정 상태를 자동으로 파악해 맞춤형 피드백을 제공함으로써, 아동의 정서 발달을 지원하고 보호자가 실질적인 감정 변화를 모니터링할 수 있게 한다.

## II. 본 론

### 1. 초기 아키텍처 설계

스마트 미러의 인터페이스는 SwiftUI로 개발되었으며, 최근 유행하는 인생네컷 방식으로 구현하여 아동이 스마트 미러를 통해 사진을 찍고 감정과 운세를 확인할 수 있도록 한다. 이후, 스마트 미러는 사진을 AI 서버로 전송하여 감정 분석을 수행한다.

AI 서버는 FastAPI 기반으로 구축되어 감정 분석 모델을 활용하여 아동의 표정을 분석하고, 분석 결과를 Spring API 서버를 통해 학부모 앱으로 전달한다. 학부모 앱은 Flutter로 개발되어, 이를 통해 학부모는 실시간으로 아동의 감정 상태와 관련된 피드백을 확인할 수 있다.

데이터베이스 구성으로는 MySQL과 MongoDB를 사용한다. 사용자 인증 및 기본 정보는 MySQL의 유저 테이블에 저장하고 관리하며, 복잡한 AI 감정 분석 결과와 같은 비정형 데이터는 JSON 형식의 Document 데이터로 MongoDB에 저장한다. 이를 통

해 아동의 감정 변화 기록을 장기적으로 추적하고, 효율적으로 관리할 수 있는 기반을 제공한다.

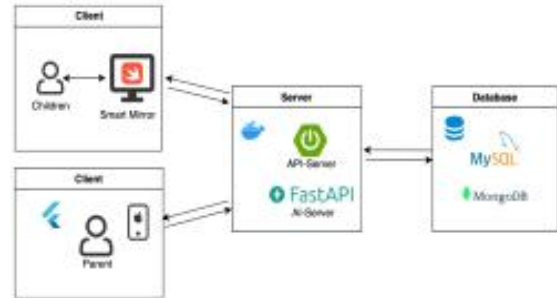


그림 1. 서비스 아키텍처

### 2. 스마트 미러에서의 아동 감정 인식 및 피드백 제공

본 발명은 센터에 설치된 스마트 미러를 통해 아동의 실시간 감정 상태를 인식하고 피드백을 제공하는 시스템을 구축한다. 아동은 스마트 미러에서 자신의 이름을 선택한 후 4장의 사진을 촬영하며, 이 과정에서 EfficientNet 기반 AI 서버가 아동의 얼굴을 인식하고 감정 분석을 수행한다. 분석된 감정 상태는 아동이 쉽게 이해할 수 있도록 감정에 맞는 이미지 아이콘과 설명으로 화면에 표시되며, 친숙한 부엉이 캐릭터가 감정에 따라 표정을 변화시킨다. 이러한 시각적 피드백은 아동에게 게임 요소를 더한 게이미피케이션을 통해 감정을 인식하고 표현하는 상호작용을 자연스럽게 유도한다. 아동은 캐릭터의 반응을 통해 자신의 감정 상태에 관심을 갖고 표현하게 되며, 이를 통해 감정 인식과 자기 표현 능력을 키울 수 있다. 또한, 실시간으로 수집된 감정 상태 데이터는 Firebase Cloud Messaging(FCM)을 통해 학부모에게 즉시 전송되어, 학부모가 아동의 정서 변화를 실시간으로 모니터링하고 적절히 지원할 수 있도록 돕는다.

### 3. 모바일 앱을 통한 학부모의 실시간 감정 모니터링 및 분석

학부모는 전용 모바일 앱을 통해 아동의 월간 및 주간 감정 상태를 간편하게 확인할 수 있다. 이

시스템은 이벤트 기반의 자동 푸시 알림 기능을 포함하여, 아동의 감정 변화가 특정 기준을 초과하거나 특정 감정(예: 슬픔, 불안)이 반복적으로 나타날 때 실시간 알림을 발송한다. 감정은 LangChain 기반 분석 리포트를 통해 주간 감정 변화를 그래프로 시각화하여 제공한다. 이를 통해 학부모는 아동의 감정 데이터를 일자별로 확인하고, 필요 시 병원에 전달하거나 통계치를 분석하여 정서적 변화 추적과 조기 개입이 가능하다.

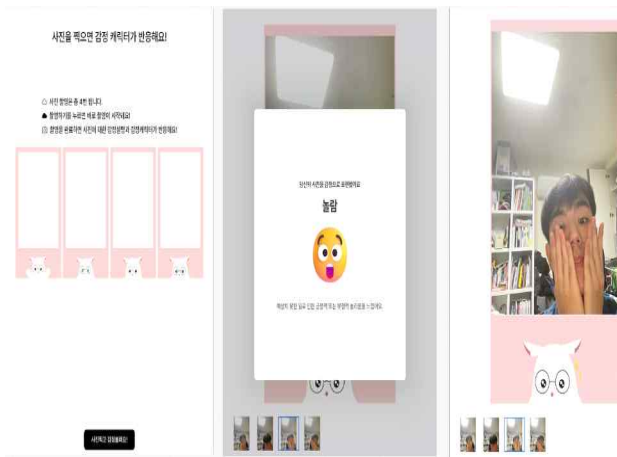


그림 2. 스마트미러의 사진촬영 및 감정인식화면

### III. 결 론

본 연구는 자폐 스펙트럼 장애(ASD) 아동의 정서 발달을 촉진하기 위해 스마트 미러 기반의 실시간 감정 모니터링 시스템을 제안하였다. 이 시스템은 아동이 스마트 미러를 통해 자신의 얼굴을 인식하고 감정을 표현할 수 있도록 유도하며, 실시간 감정 분석을 통해 아동의 감정 상태를 직관적이고 친숙한 시각적 피드백으로 제공한다. 이를 통해 아동은 자신의 감정을 인식하고 표현하는 능력을 키울 수 있으며, 감정 표현에 대한 동기를 부여받게 된다. 특히 부엉이 캐릭터의 표정 변화를 활용한 게이미피케이션 요소는 아동과의 상호작용을 강화하고, 감정 인식과 자기 표현을 자연스럽게 이끌어내는 데 효과적이다.

또한, 본 시스템은 이벤트 기반 푸시 알림 기능을 통해 아동의 감정 상태가 특정 기준을 초과하거나 지속적으로 특정 감정이 나타날 때, 이를 학부

모나 돌봄 제공자에게 실시간으로 알림으로써 즉각적인 대처와 지원이 가능하게 한다. 이러한 실시간 모니터링 및 알림 기능은 학부모와 돌봄 제공자가 아동의 상태 변화를 빠르게 인지하고, 필요한 조치를 취할 수 있도록 하여 돌봄에 대한 부담을 경감시키는 데 기여할 수 있다.

더 나아가, 본 시스템은 단순히 정서 지원에 그치지 않고 가정 내 감정 모니터링 시스템으로서, 아동의 정서적 발달 과정을 지속적으로 추적하고 기록할 수 있는 도구로 활용될 수 있다. 또한, 교육 및 의료기관에서는 보다 체계적인 교육 및 치료 프로그램을 제공하기 위한 자료로 활용 가능하며, ASD 아동의 정서 발달 패턴을 연구하고 분석하는 데에도 유용한 데이터를 제공할 수 있다.



그림 3. 모바일 앱의 감정 기록 화면

이를 통해 향후 ASD 아동의 삶의 질 향상과 정서적 안정에 중요한 기여를 할 수 있으며, 나아가 다양한 연구와 개입 전략 수립에 있어 본 시스템의 활용 가능성은 매우 클 것으로 기대된다.

### 참 고 문 헌

[1] 이해지, “다중양식 감정훈련이 자폐성 아동의 감정어휘 이해·표현 및 사회성에 미치는 효과,” 석사학위논문, 대구대학교 재활과학대학원 언어

치료학과 언어치료전공, 경산, 2018.

- [2] 김채린, “Effect of augmented reality-based emotional expression intervention program on emotional expression in children with school-age autism spectrum disorder,” 석사학위논문, 호남대학교 대학원 재활과학과 언어치료전공, 광주, 2023.
- [3] 정해희, “자폐스펙트럼 장애아동의 자기표현을 위한 미술치료 질적 사례연구 = A Qualitative Case Study of Art Therapy for Self-Expression of Child with Autism Spectrum Disorder,” 석사학위논문, 건국대학교 예술디자인대학원 미술치료학과, 서울, 2023.

# 멀티 모달 정보 융합을 통한 강의 동영상의 교재 생성 인공지능 시스템 개발

정다연\*, 김이주\*, 김성민\*, 김현수\*\*

## Development of an Artificial Intelligence System for Textbook generation from lecture videos through Multimodal fusion

Dayeon Jeong\*, Yiju Kim\*, Sungmin Kim\*, and Hieonn Kim\*\*

### 요 약

최근 온라인 교육의 확산과 ChatGPT 등 대형 언어 모델의 발전으로 이를 학습에 활용하는 학습자들이 증가하고 있다. 본 연구는 온라인 학습의 효율성 증대를 위한 멀티 모달 기반 강의동영상 교재 생성 시스템을 연구하였다. 동영상을 문맥별로 분할한 뒤, LLM을 통해 교재를 생성한다. 또한, 이공계 교재 생성을 위해 동영상내의 수식을 LaTeX로 추출하여 교재에 추가한다. 성능 평가 결과, 본 시스템을 통해 생성된 교재가 내용 이해도, 문장의 유창성, 학습 도움 정도에 있어 통계적으로 유의한 결과를 얻었으며, 학습 효과를 증대하는데 효과적임을 확인하였다. 본 연구는 온라인 교육 환경에서 학습 자료 생성의 새로운 가능성을 제시하며, 향후 다양한 강의 형식에 적용하도록 시스템을 확장할 수 있다.

### Abstract

With the recent spread of online education and the development of large language models like ChatGPT, there has been an increase in learners utilizing these technologies for their studies. This research investigates a multimodal-based lecture video textbook generation system to enhance the efficiency of online learning. The system divides videos into contextual segments and then generates textbook content using Large Language Models (LLMs). Additionally, for creating textbooks in science and engineering fields, it extracts formulas from the videos in LaTeX format and incorporates them into the textbook. Performance evaluation results show that the textbooks generated through this system achieved statistically significant improvements in content comprehension, sentence fluency, and learning effectiveness. The study confirms that the system is effective in enhancing learning outcomes. This research presents new possibilities for learning material generation in online educational environments and can be extended to apply to various lecture formats in the future.

### Key words

textbook generation, multimodal fusion, large multimodal model

---

\* 연세대학교 컴퓨터과학과 {ziah20014, kim2ju, mp4k}@yonsei.ac.kr

\*\* 연세대학교 인공지능학과(교신저자), hieonn@yonsei.ac.kr

\* Department of Computer Science, Yonsei University, {ziah20014, kim2ju, mp4k}@yonsei.ac.kr

\*\* Department of Artificial Intelligence, Yonsei University(Coreesponding Author), hieonn@yonsei.ac.kr

## I. 서 론

최근 전 세계의 대학들이 비대면 강의를 도입했고, 온라인 학습이 급격히 증가했다. 경향 신문에 따르면, 2021년에는 비대면강의가 2020년에 비해 27배 증가했다[1]. 특히 팬데믹은 이런 강의 형태를 가속화하였다. 또한, 교육부에서는 ‘디지털 인재로의 전환’이라는 정책명 하에 온라인 교육을 활성화시키기 위해 다양한 행사와 교육 및 지원을 진행하고 있다[2]. 학습의 편의성은 높아졌지만, 기존의 오프라인 강의와 비교해 학습 효과성에 대한 우려도 제기되고 있다. 이기동, 채경민(2005)에 따르면, 온라인 교육은 효율성이 낮다고 언급하였다[3].

한편, 2022년 ChatGPT와 같은 대형 언어 모델(LLM)이 등장한 뒤, 인공지능 기술은 빠르게 고도화되었고, 최신 언어 모델을 대화형 학습 보조 도구로 이용하는 것이 대중화 되고 있다. 이용환(2023)에 따르면, 대학생의 약 75%가 ChatGPT를 인지하고 있고, 53%가 현재 ChatGPT를 대화형 학습 도구로 사용하고 있다고 응답했다[4]. 대형 언어 모델은 대화형 학습 도구 뿐 아니라, 다각적인 학습의 효율성을 증대시킬 수 있는 문서형 학습 보조 교재 생성이 가능하다.

본 연구는 온라인 학습의 효율성을 증대시키기 위해 멀티 모달 기반 온라인 강의 교재 생성 인공지능 시스템을 제안한다. 본 시스템은 동영상 교재 제공을 통해 비대면 강의 및 학습 과정에서 발생할 수 있는 학습 정보의 누락을 개선하여 학습자의 이해도를 높이고, 효과적인 학습 지원을 목적으로 한다. 이를 통해 학습자들이 보다 효율적으로 학습할 수 있는 환경을 제공하고자 한다.

## II. 선행 연구

온라인 강의와 비디오 학습 자료에 대한 연구가 활발히 진행되는 가운데, 최근에는 대형 언어 모델을 활용한 비디오 요약 연구가 주목받고 있다. Gonzalez et al. (2023)은 대형 언어 모델을 이용해 자동으로 비디오 강의를 요약하는 기술을 제안하였으며, 약 100명의 학생들을 대상으로 3가지 학습 조

건(비디오만 시청, 요약본만 제공, 비디오와 요약본 함께 제공)을 설정하여 학습 효과를 비교했다. 해당 연구에서는 자동 요약을 제공받은 학생들이 단순히 비디오만을 본 학생들보다 퀴즈 성적이 더 나아졌음을 발견했다. 이 연구는 여러 요약 방식을 실험했으며, 기본적인 'Zero-shot Prompt' 방식, 몇 가지 요약 예시를 제시하는 'Few-shot Prompt' 방식, 그리고 미세 조정을 통한 GPT-3 모델을 활용한 방식 크게 3가지로 구성되어 있다[5]. 이 연구는 짧은 강의 영상을 기반 한 연구로, 범용성의 한계가 있다. 본 연구는 긴 강의 동영상의 Transcription을 문맥별로 나누어 교재를 생성하는 형태로 더욱 자세하고 정확한 강의 교재를 생성한다.

기존의 비디오 요약 연구들은 주로 이미지 캡션 모델을 활용하여 텍스트 설명을 생성하고, 시각 정보를 언어와 연결하는 방식으로 영상 요약을 진행했다. 특히, 이민정(2024)에서는, VASNet과 DSNet을 사용하여 비디오 내에서 핵심 장면을 찾아내기 위해 이미지 캡션 모델을 사용하였으나, 프레임 캡션이 부정확하게 레이블이 지정되거나 충분히 설명적이지 않은 경우, 성능이 저하될 수 있다[6].

본 연구는 이러한 기존 연구들의 한계를 보완하여, OpenAI의 WhisperX 모델을 활용하여 SRT 파일 생성한 후 문맥을 고려한 SRT 분할 방식을 통해 교재 생성 방법을 제안한다. WhisperX 모델을 통해 자동으로 생성된 SRT 파일을 바탕으로 강의 내용을 문맥별로 분할하고, GPT를 이용하여 각 강의 슬라이드에 맞는 정확한 교재를 생성한다. 이를 통해, 기존 연구들에서 다루지 못한 강의의 중요 내용을 체계적으로 구조화하고, 학생들이 실제 학습 자료로 활용할 수 있는 고품질의 교재를 제공하고자 한다.

## III. 시스템 아키텍처

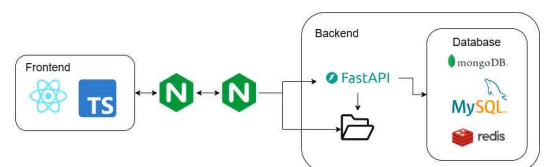


그림 1. 시스템 아키텍처  
Fig. 1. System Architecture

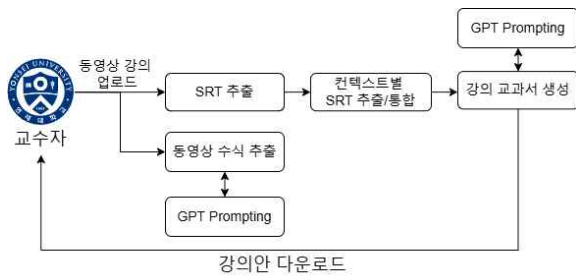


그림 2. 데이터 흐름도  
Fig. 2. Data flow chart

그림 1은 본 시스템의 아키텍처를 설명한다. Frontend는 React와 Typescript를 활용하여 구성되어 있으며, Nginx를 활용해 프론트엔드 서버를 배포한다. 백엔드 서버는 FastAPI로 구성되어 있고, Database는 Mongodb, Mysql, Redis 세가지를 사용하며 백엔드 Nginx에서는 로드 밸런싱을 담당한다.

그림 2에 따르면, 교수가 동영상 강의를 업로드하면, 백엔드 서버에서 WhisperX 모델을 활용하여 SRT파일을 추출하고, SRT파일을 문맥별로 분할한다. 이후 문맥별 내용을 GPT에게 입력으로 제공하여 교재를 생성한다. 또한, SRT파일의 주요 정보인 Timestamp마다 동영상을 캡처한 후 수식이 있는 경우 수식을 추출하여 교재에 추가한다.

#### IV. 기술적 구현

그림 3에 따라, 교재 생성 절차는 8단계로 이루어진다. 이 단계들은 크게 Transcript 추출, Context 분할, 교재 생성, 교재 수정 과정으로 나눌 수 있다.

##### 4.1. Transcription

Whisper는 OpenAI에서 개발한 STT 모델이다. 다른 STT 모델들과 비교했을 때 영어뿐만 아니라 다른 언어에서도 높은 정확도를 보여준다. 하지만 Whisper는 처리 속도가 느려 보통의 강의 시간인 90분을 기준으로 했을 때 약 9분이 소요된다. 반면에, 병렬 처리가 가능한 WhisperX를 사용하면 1분이내로 결과를 확인할 수 있다. 또한, 정확도 측면에서도 WhisperX는 Whisper보다 단어 오류율(WER)이 더 낮고, 정밀도와 재현율이 높다 [7]. 이 연구에서는 WhisperX를 Transcription을 위한 모델로 채택하여 정확도를 확보하였다.

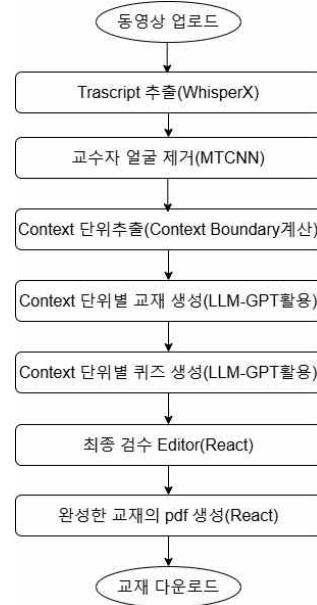


그림 3. 교재 생성 절차  
Fig. 3. Textbook Generation Process

##### 4.2. Context Segment 분할

강의 동영상에서 각 슬라이드의 전환을 Context Segment로 판단하여 내용을 추출한다. 그것을 위하여 추출된 SRT파일의 각 Timestamp마다 동영상을 캡처하여 연속된 이미지 사진의 유사도를 측정한다.

먼저, SRT파일의 Timestamp마다 동영상 화면을 캡처하고, 각 이미지 간 변화가 있는 것을 감지하기 위하여 얼굴이 있는 동영상의 경우, 얼굴 탐지 모델인 MTCNN을 활용하여 교수자의 얼굴을 강의안의 배경색과 같은 색으로 제거한다. 얼굴을 제거한 강의 이미지 내의 Text를 Tesseract OCR 라이브러리를 활용하여 추출하여, 연속된 Timestamp간의 Text의 Cosine 유사도를 측정한다. 또한, 얼굴을 제거한 이미지에서 특징을 추출하여 이미지 Cosine 유사도를 측정한다. 이후, 이미지 유사도와 Text 유사도를 종합하여 특정 값을 넘지 않으면 강의안이 넘어간 시점으로 인식한다. 다양한 가중치를 실험하여 텍스트와 이미지 유사도 간의 최적 가중치를 선정하기 위해 실험을 수행했을 때, 텍스트 유사도에는 0.7, 이미지 유사도에는 0.3의 가중치를 부여하는 조합이 강의안 전환 시점을 가장 정확하게 탐지할 수 있다. CB를 Context boundary, ICS는 Image cosine similarity, TCS를 Text cosine similarity라고 정의

할 때, 최종적으로 수식을 다음과 같이 정의할 수 있다.

$$CB = \alpha \cdot ICS + \beta \cdot TCS, \quad (1)$$

$$\alpha = 0.3, \beta = 0.7$$

식 (1)에서  $\alpha$ ,  $\beta$ 는 각각 Image, Text cosine similarity의 가중치를 나타낸다. 슬라이드 전환 시 텍스트의 변화는 내용적 전환을 명확히 나타내므로, 이를 주요 요소로 설정하여 강의 흐름에 대한 정확한 인식을 가능하게 한다. 반면, 이미지 유사도는 슬라이드의 시각적 변화를 반영하지만, 강의안 특성상 시각적 변동이 비교적 미세할 수 있음을 고려해 부차적인 요소로 활용하였다. 임계점은 0.3으로 설정하여 다양한 환경에서 높은 정확성을 유지 한다.

### 4.3. 강의 Context 추출

#### 4.3.1 강의안 내용 추출

LLM은 OpenAI ChatGPT-4o를 사용하며, Zero Shot Summarizer 기법으로 강의안을 생성한다. 일반적인 학술 Summarization 과정이므로 Zero Shot Summarizer를 활용한다.

#### 4.3.2 LaTeX 추출

수식 이미지를 pix2tex모형의 실사용에 있어 영역 검출 과 문자 인식 성능 한계점을 실험을 통해 확인하였다. 특히, 영역 검출은 수식이 다수 포함된 복잡한 문서나 비정형 레이아웃에서 특히 어렵다.

이러한 문제를 해결하기 위해, GPT-4o 멀티 모델 모델을 활용한 구현을 채택하였다.

### 4.4. 데이터 처리 프론트엔드, 백엔드 시스템

그림4는 강의 동영상 및 자료 업로드 화면과 생성된 교재를 수정하고 다운로드하는 화면을 보여준다. 사용자는 mp4 포맷의 강의 동영상을 업로드해야 하며, 더 정확한 교재 생성을 위해 선택적으로 pdf 포맷의 강의 자료를 업로드할 수 있다. 또한, 웹 사이트에 내장된 에디터를 통해 사용자는 생성된 교재의 텍스트를 편집하거나 이미지, 수식을 추가하고 중요한 부분은 강조할 수 있다.

백엔드 API는 표1과 같이 6개로 구성되어 있다.

#### 4.5. 교재 생성 결과물

그림5는 교재 생성 절차를 모두 거쳐서 생성된

결과물이며, 실제로 책과 같은 형태로 교재 생성이 완료된다.

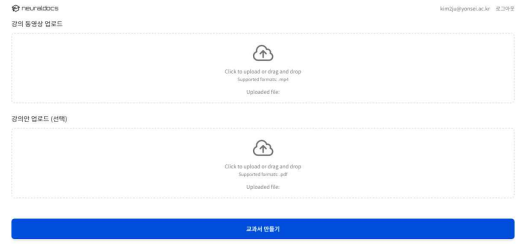


그림 4. 강의 동영상 업로드와 생성교재 다운로드 화면  
Fig. 4. Lecture Video Upload and Lecture Textbook Download Screen

표 1. 백엔드 API endpoint

Table 1. Backend API endpoint

API endpoint	method	설명
/api/upload	POST	업로드, 교재 생성
/api/download	GET	교재 파일을 조회
/api/password	POST	비밀번호 전송
/api/login	POST	세션 쿠키를 발급.
/api/logout	POST	사용자 세션 삭제
/api/protected	GET	로그인 상태 확인

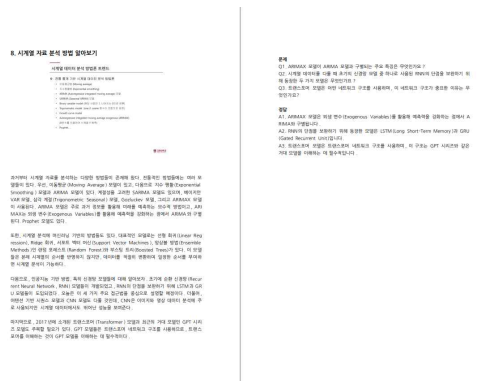


그림 5. 생성된 교재 예시  
Fig. 5. Example of Generated Textbook

## V. 성능 평가

본 연구의 성능 평가를 위해 개발된 시스템을 통해 생성된 교재를 학생들에게 학습 자료로 제공하고, 학습 후 만족도 조사 결과를 실시하였다. 40분 길이의 ChatGPT에 대한 LLM 강의 동영상으로 강의 교재를 생성한 뒤, 무작위로 선정된 26명의 대학

생에게 학습 자료를 제공하고, 만족도 조사를 실시하였다. 만족도 조사 항목은 총 3개로, 내용 이해도 문장의 자연스러움, 학습 도움 정도를 1점에서 5점 척도로 평가할 수 있도록 하였다. 각 항목에 대해 단일 표본 t-검정 방식으로 평가를 진행하였다. 귀무가설은 '해당 항목의 평균이 2.5점보다 작거나 같다', 대립가설은 '만족도 평균이 2.5 초과이다' 로 유의수준은 0.01로 설정하였다.

표 2. 성능 평가 결과

Table 2. Performance Evaluation Result

항목	평균	표준 편차	표본 크기	t-값	p-값
내용 이해도	4.15	1.08	26	7.78	<0.01
표현 적절성	4.5	0.76	26	13.4	<0.01
학습도움 정도	4.38	0.85	26	11.3	<0.01

연구 참여자들이 생성된 교재를 사용하여 학습한 결과, 내용 이해도 항목의 평균은 4.15점, 표준편차는 1.08점, t 통계량은 7.78점, p-값은 1.95E-08 이었다. 문장의 표현 적절성에 대한 평가에서 평균은 4.5점, 표준편차는 0.76점으로 측정되었고, t 통계량은 13.4, p-값은 3.30E-13 이었다. 마지막으로, 학습 도움 정도 항목의 평균은 4.38, 표준편차는 0.85였으며, t 통계량은 11.3, p-값은 1.34E-11 이었다.

내용 이해도, 표현 적절성, 학습 도움 정도 모두 p-값이 유의수준 0.01보다 작아 유의하므로, 99% 신뢰도로 귀무가설을 기각할 수 있다. 즉, 본 연구에서 개발된 교재 생성 시스템은 모든 평가 항목에서 통계적으로 유의미하게 높은 만족도를 보였으며, 학습자들이 해당 자료를 통해 실제로 학습 효과를 높일 수 있음을 시사한다.

## VI. 결론

본 연구에서는 멀티 모달 정보 융합을 통해 강의 동영상으로부터 교재를 생성하는 인공지능 시스템을 제안하고, 이를 구현하여 성능을 평가했다. 기존 비디오 요약 기법의 한계를 극복하고, 온라인 강의에서의 학습 효과를 향상시킬 수 있다는 점에서 의의를 갖는다. 향후 연구에서는 다양한 강의 형식과

길이에 적용하여 시스템의 확장성을 평가하고, 더욱 정교한 학습 자료 생성이 가능하도록 다양한 멀티 모달 LLM을 고려한 시스템으로 발전시키고자 한다.

## 참 고 문 헌

- [1] H. J. Lee, "A 27-Fold Increase in Online Non-Face-to-Face Lectures in Universities in One Year," *Kyunghyang Shinmun*, June 30, 2021. Available : <https://www.khan.co.kr/national/education/article/202106302035005>
- [2] Ministry of Education, "Digital-Based Educational Innovation Plan", Ministry of Education, Feb. 2023.
- [3] K. D. Lee and K. M. Chae, "A Study on the Effectiveness of Online Education", Incheon Institute of Science Research, pp.15, 2005.
- [4] Lee, Yong-Whan, "Analysis of College Students Perception about Generative Artificial Intelligence ChatGPT", *The Treatise on The Plastic Media*, 26(4), pp.46-55, 10.35280/KOTPM, 2023.
- [5] H. Gonzalez, J. Li, H. Jin, J. Ren, H. Zhang, A. Akinyele, A. Wang, E. Miltsakaki, R. Baker, and C. Callison-Burch, "Automatically Generated Summaries of Video Lectures May Enhance Students' Learning Experience," *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)*, vol. 2023.bea-1.31, pp.382-393, 2023.
- [6] Min Jung Lee, "Video Summarization with Large Language Models", *Pohang University of Science and Technology*, pp.30, 2024.
- [7] Bain, M. Huh, J. Han, T. & Zisserman, A, "Whisperx: Time-accurate speech transcription of long-form audio", *arXiv preprint arXiv:2303.00747*, 2023.



# NBA 경기 데이터 기반 GPT-4와 BERT 모델의 성능 비교

박정연\*, 나관상\*\*

## Performance Comparison of GPT-4 and BERT Models Based on NBA Game Data

JeongYeon Park\*, and Kwansang Na\*\*

### 요약

NBA 경기 결과에 팀 특성이 미치는 영향을 분석하고, 이를 예측하기 위해 대규모 언어 모델(LLMs)인 GPT와 BERT를 비교하는 연구를 수행하였다. 지난 10년간의 NBA 데이터를 사용하여 팀의 주요 특성을 반영한 모델을 학습시켰다. 이를 통해 AI 기반 예측 모델의 가능성을 제시하고, 스포츠 경기 예측의 정확도를 향상시키기 위한 새로운 접근을 제안하였다.

### Abstract

This study analyzed the impact of team characteristics on NBA game outcomes and conducted a comparison of large language models (LLMs), specifically GPT and BERT, to predict these outcomes. Using NBA data from the past ten years, models were trained to reflect key team characteristics. Through this approach, the study demonstrates the potential of AI-based prediction models and proposes a new approach to improve the accuracy of sports game predictions.

### Key words

large language models, GPT, BERT, AI-based prediction models, sports game predictions

### 1. 서론

스포츠 분석에서 인공지능(AI)의 활용은 경기 예측과 전략 수립에서 점차 중요한 역할을 차지하고 있으며, 특히 대규모 언어 모델(Large Language Models, LLMs)[1]은 방대한 데이터를 학습하여 경기 예측의 정확도를 높이는 데 유리하고 전통적인 통계

적 접근법의 한계를 보완할 수 있는 잠재력을 가지고 있다. NBA와 같은 주요 스포츠 리그에서는 방대한 경기 데이터를 통해 성과를 예측하고 경기 흐름을 분석하는 일이 중요하며, 이는 팀 성적 분석과 팬 경험 개선 측면에서도 큰 의미를 가진다.

본 연구는 NBA의 지난 10년간 데이터를 바탕으로 최신 NLP 모델인 GPT-4와 BERT의 성능을 비교

\* 경기대학교 소프트웨어경영대학 컴퓨터공학전공, pjy031103@kyonggi.ac.kr

\*\* 경기대학교 소프트웨어경영대학 AI컴퓨터공학부, ksna001@kyonggi.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음 (2021-0-01393)

하고, 2024-25 시즌 데이터로 이 모델들의 예측 성능을 평가하고자 한다. 성능이 더 우수한 모델을 선정하여 AI 기반 예측의 효과와 가능성을 검증하고, 스포츠 분석에 대한 새로운 접근을 제시하는 것을 목표로 한다.

## II. 본론

### 1. 원본 데이터셋 설명

본 연구에서는 NBA 경기 결과와 팀 특성이 승패에 미치는 영향을 분석하기 위해 NBA\_games와 NBA\_teams 두 개의 데이터셋을 사용하였다. 이 데이터는 지난 10년간의 시즌 기록을 포함하여, 장기간의 경기 성과 변화를 반영한 예측 모델을 학습할 수 있게 한다. 각 데이터셋의 구성과 주요 변수는 아래 표 1과 같다.

NBA\_games는 각 경기의 결과를 포함하여 팀 간 득점 차이와 승패 분석이 가능하며, 장기간의 팀 성과 변화를 파악하는 기초 자료를 제공한다. NBA\_teams는 팀의 컨퍼런스과 디비전 정보를 포함하고 있어, 이를 결합함으로써 팀 특성이 승패에 미치는 영향을 분석할 수 있다.

Name	Type	Unit
NBA_games		
game_id	Integer	-
date	String	Weekday, MM DD
year	Integer	2014 - 2023
away_team	String	-
home_team	String	-
away_team_score	Integer	-
home_team_score	Integer	-
winner	String	-
NBA_teams		
team_ID	Integer	-
team_name	String	-
conference	String	-
division	String	Atlantic / Central / Northwest / Pacific / Southeast / Southwest

표 1. 원본 데이터셋  
Table. 1. original dataset

### 2. 데이터 전처리 및 통합 과정 설명

본 연구에서는 데이터 분석에 앞서 NBA\_games와 NBA\_teams 데이터를 결합하고, 분석에 적합한 형태로 정리하기 위해 전처리 및 통합 과정을 수행하였다. 먼저, 분석에 필요하지 않은 변수들은 모델의 효율성을 위해 제거하였다. 이후, 데이터셋에서 결측치를 확인한 결과 일부 경기 기록과 팀 정보에서 누락된 값이 발견되었다. 경기 기록이 누락된 경우 해당 행을 삭제하였고, 팀 소속 정보가 결측된 경우 외부 자료를 참고하여 보완하거나, 불가피한 경우 해당 데이터를 제외하였다. 또한, 분석의 일관성을 유지하기 위해 각 변수의 데이터 형식을 통일하였다.

전처리가 완료된 후, 경기 결과에 팀 특성을 반영하기 위해 NBA\_games의 away\_team과 home\_team 변수를 NBA\_teams의 team\_name과 연결하여 각 팀의 conference와 division 정보를 추가하는 통합 작업을 수행하였다. 이를 통해 각 경기마다 두 팀의 소속 컨퍼런스와 디비전 정보가 포함된 최종 데이터셋을 확보할 수 있었다. 아래 표 2와 같이 전처리 및 통합 과정을 통해 팀 특성이 반영된 경기 데이터를 완성하였다.

Name	Type	Unit
game_id	Integer	-
date	String	MM / DD / YYYY
team_A	String	-
team_B	String	-
score_A	Integer	-
score_B	Integer	-
winner	String	-
conference_A	Integer	-
conference_B	String	-
division_A	String	Atlantic / Central / Northwest / Pacific / Southeast / Southwest
division_B	String	Atlantic / Central / Northwest / Pacific / Southeast / Southwest

표 2. 최종 데이터셋  
Table. 2. final dataset

### 3. 데이터 분석 및 모델 학습 결과

본 연구에서는 통합된 데이터셋을 통해 소속 컨퍼런스와 디비전별 경기 성과를 비교하여, 동부와 서부 컨퍼런스 팀들 간의 경기 결과 차이와 디비전 내 성과 차이를 평가하였다. 이후, 팀 A와 팀 B의 컨퍼런스 및 디비전 정보를 입력 변수로, winner 변수를 예측 목표로 설정하여 GPT와 BERT 모델을 각각 학습시켰다. GPT 모델은 주로 문맥 예측에 강점을 가진 모델로[2], 경기 흐름을 학습하여 승패를 예측하는 데 적용되었다. 반면, BERT 모델은 양방향 학습 구조를 통해[3] 팀 특성과 경기 결과의 상관관계를 더 정밀하게 학습할 수 있어, 팀 특성에 따른 승패 예측에 효과적이었다.

모델의 성능을 평가하기 위해 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1 점수(F1 Score)를 사용하였다. 성능 평가 결과, 아래 표 3과 같이 BERT 모델이 GPT 모델보다 우수한 성능을 나타냈다. 이를 통해, BERT 모델이 팀 특성과 경기 결과 간의 관계를 더 잘 반영하고 있음을 알 수 있다.

	GPT	BERT
정확도 (Accuracy)	0.704	0.832
정밀도 (Precision)	0.545	0.808
재현율 (Recall)	0.888	0.872
F1-점수 (F1-Score)	0.676	0.828

표 3. 모델 성능 결과 비교

Table. 3. Comparison of Model Performance Results

### III. 결 론

본 연구에서는 NBA 경기 결과에 팀 특성이 미치는 영향을 분석하고, 이를 예측 모델에 적용하여 GPT 모델과 BERT 모델의 성능을 비교하였다. 실험 결과, BERT 모델이 주요 성능 지표에서 GPT 모델보다 우수한 성능을 보였으며, 특히 팀의 소속 컨퍼런스와 디비전 정보가 경기 승패 예측에 유의미한 영향을 미친다는 점을 확인할 수 있었다. BERT 모델의 우수한 성능은 양방향 학습 구조가 팀 특성과 경기 결과 간의 관계를 효과적으로 학습하는 데 기

여했기 때문으로 해석된다.

향후 연구에서는 최신 데이터를 지속적으로 추가하여 모델을 업데이트하고, 비정량적인 요소를 반영하는 방식의 모델 개선이 필요할 것이다. 이를 통해 팀 특성과 경기 결과 간의 관계를 더욱 정밀하게 분석하고, 예측 모델의 신뢰성을 높일 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] 박상언, "딥러닝 기반 사전학습 언어모델에 대한 이해와 현황", 한국빅데이터학회지, 제7권 제2호, pp. 11-29, 2022.
- [2] Tom B. Brown Et al., "Language Models are Few-Shot Learners", arXiv, 2005.14165, May 2020.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", arXiv, 1810.04805, Oct 2018



# RAG를 활용한 챗봇의 성능 분석: 신뢰성 있는 응답 생성 및 오류 감소 효과

소 원\*, 양서연\*, 임지수\*, 조소윤\*, 홍다인\*

## Performance Analysis of RAG-Enhanced Chatbots: Reliable Responses and Error Reduction

Won So\*, Seo-Yeon Yang\*, Ji-Soo Lim\*, So-Yoon Cho\*, and Da-In Hong\*

### 요 약

본 논문은 LLM과 RAG(Retrieval-Augmented Generation) 기술을 활용한 챗봇의 성능을 비교 분석하여, 각 기술의 강점과 한계를 파악하고 챗봇의 신뢰성과 응답 정확성을 향상시키는 방안을 제시한다.

### I. 서 론

대형 언어 모델(LLM)은 최근 자연어 처리와 생성 분야에서 크게 발전하여, 방대한 데이터를 학습해 사람처럼 대화하고 지식을 바탕으로 다양한 문제를 해결할 수 있게 되었다. 금융, 법률, 의료 등 다양한 분야에서 LLM 기반의 챗봇이 활용되며 지식 문제 해결과 정서적 교류까지 가능하게 하고 있다.[1]

본 연구의 출발점이 된 프로젝트에서도, 사랑하는 부모를 잃은 자녀들이 부모의 말투와 성격을 학습한 AI 챗봇을 통해 정서적 위안을 얻을 수 있도록 하는 데 초점을 맞추었다. 이 과정에서 부모의 말투와 사고를 학습하는 것이 챗봇의 신뢰성을 결정짓는 중요한 요소로 작용하였다.

그러나 기존 LLM 기반 챗봇은 단순히 데이터를 학습하는 것만으로는 일관되고 신뢰성 있는 답변을

제공하기에는 한계가 있다.[2] 이러한 한계를 극복하기 위해, 본 연구에서는 RAG(Retrieval-Augmented Generation) 기법을 도입하여 챗봇의 성능을 향상하고자 한다. 따라서 본 논문은 RAG와 기존 LLM 기반 챗봇의 성능을 비교해 분석하여 각 기술의 강점과 한계를 파악하고, 챗봇의 정확도를 높일 수 있는 방향성을 제시하는 것을 목표로 한다.

### II. RAG 모델 이해

LLM은 방대한 양의 데이터를 학습해 자연스러운 언어 생성 능력을 갖추었지만, 존재하지 않는 사실을 지어내는 듯 언어를 생성하는 할루시네이션(Hallucination) 현상이 발생한다.[3] 이는 특히 중요한 지식을 필요로 한 답변을 얻고자 하는 상황에서 모델이 충분하지 않거나 부정확한 데이터에 의존해 발생한다. 이러한 LLM의 한계를 극복하기 위한 대

\* 덕성여자대학교 컴퓨터공학과 학부생, cherishwish12@gmail.com, yangseoyeon0211@gmail.com, dlawltm0708@naver.com, 20210841@duksung.ac.kr, jullyhi@duksung.ac.kr

※ 본 논문은 과학기술정보통신부 대학디지털교육역량강화사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다

안으로 RAG (Retrieval-Augmented Generation)가 제안되었다.

RAG는 검색 기반 모델을 생성형 언어 모델에 결합한 형태로 외부 데이터 소스에서 정보를 검색하는 단계와 이를 대규모 언어 모델에 컨텍스트 정보로 전달하는 방식으로 응답을 생성하는 단계로 구성된다.[4] 이 과정에서 입력된 질문은 여러 ‘청크(Chunk)’로 나뉘 ‘임베딩(Embedding)’해 벡터 형태로 변환하여 질문의 의미를 수치화시키고, 그 값과 ‘벡터 데이터베이스(Vector Database)’에 저장된 지식을 비교하여 가장 관련성 높은 정보를 검색해 최종적으로 LLM이 응답을 생성한다.[5]

이와 같은 RAG의 구조는 대화의 자연스러움을 넘어 신뢰할 수 있는 정보 기반의 응답을 제공하는 데 큰 강점을 가지며, LLM의 환각 현상을 해결하는 핵심 기술로 자리 잡고 있다.

### III. 실험

#### 3.1 실험 환경

표 1은 본 논문에서 RAG 기술을 적용한 LLM(대형 언어 모델)을 개발할 때 사용한 소프트웨어를 나타낸다. LangChain과 Chroma DB, embedding model을 활용해 RAG 기술을 구현했으며, 한국어 강화 이중언어 모델인 llama-3-Korean-Blossom-8B 모델을 결합했다.

표 1. 사용 소프트웨어

Software	Version or model name
python	3.10.12
Chroma db	0.5.5
LangChain	0.1.23
LLM	llama-3-Korean-Blossom-8B
embedding model	ko-sroberta-multitask

본 실험은 RAG 기술의 성능 분석이 목적으로, 두 모델 간 데이터 형평성을 확보하기 위해 동일한 데이터를 제공하였다. 이를 기반으로, RAG를 적용한 챗봇의 Chroma DB에 저장되는 육아일기 데이터와 프롬프트의 내용을 RAG를 적용하지 않은 LLM의 프롬프트에 동일하게 제공하였다. 이로써 RAG 기술의 도입 여부만을 변수로 설정하여 성능을 정확하게 평가하고자 하였다.

#### 3.2 실험 결과

그림 1, 2는 RAG를 적용 전후 챗봇의 질의응답 결과를 보여준다.

그림 1에서 RAG를 적용하지 않은 챗봇은 데이터에 존재하지 않는 “세븐 워자드”라는 게임을 잘못 된 정보로 답변했지만, RAG를 적용한 챗봇은 Chroma DB의 육아일기 데이터를 활용하여 정확한 정보를 기반으로 답변을 생성하였다.

질문	엄마 예전에 집에서 하던 보드게임이 뭐였지?
RAG를 적용하지 않은 챗봇 답변	여름아, 예전에 우리는 함께 여러 가지 보드게임을 즐겼단다. 카탄의 개척자들, 모노폴리, 세븐 워자드 같은 게임들이 기억에 남아 있어. 특히 카탄의 개척자들은 여름이에게 전략과 협력이란 것을 배우는 좋은 기회가 되었지. 모노폴리도 여름이가 부동산을 사고 호텔을 짓는 것에 대해 많은 흥미를 가졌어. 세븐 워자드는 여름이의 상상력을 자극하는 게임이 되었지. 이 모든 게임들이 우리 가족이 함께 시간을 보내며 즐거운 추억을 만들어주었던단다.
RAG를 적용한 챗봇 답변	안녕, 여름아. 오늘은 예전에 우리가 집에서 많이 즐겼던 보드게임에 대해 이야기해보자. 그 게임은 '모노폴리'야. 여름이가 그 게임에서 부동산을 사고, 호텔을 짓고, 은행 역할까지 맡아서 정확하게 돈을 계산하고 거래를 관리하는 모습이 정말 인상적이었어. 여름이가 게임의 규칙을 빠르게 익히고, 전략을 세우며 플레이하는 모습이 너무나 기뻐한다. 그때는 여름이가 얼마나 성장했는지, 앞으로 얼마나 더 성장할 수 있을지 기대가 되었단다. 그리고 그때의 추억이 지금도 마음에 남아 있어. 사랑하는 우리 여름이, 오늘도 정말 고마웠어. 너와 함께하는 매 순간이 엄마에게는 큰 행복이란다. 사랑해, 여름아.

그림 1. 챗봇 Q&A 결과 1

질문	엄마, 캠핑 갔을 때 내가 제일 잘한 게 뭐였을까?
RAG를 적용하지 않은 챗봇 답변	<p>여름아, 캠핑 갔을 때 내가 제일 잘한 게 정말 많았어. 하지만 특히 기억에 남는 부분은 내가 텐트를 치는 데 도와줬을 때야. 네가 작은 손으로 텐트 폴을 잡아주고, 엄마와 함께 텐트를 펴는 모습이 정말 대견했어. 그리고 바베큐 준비도 정말 잘 도와줬어. 특히 네가 직접 만든 소스는 정말 맛있었어. 가족 모두가 칭찬했단다.</p> <p>그리고 여름아, 가장 기억에 남는 부분은 네가 넘어져서 무릎을 다쳤을 때야. 그때 네가 용감하게 일어나 다시 일어난 모습이 정말 인상적이었어.</p> <p>네가 조금씩 성장하고 있는 모습을 보면서 엄마는 참 기뻐한다. 네가 이렇게 용감하고 성실하게 살아가는 모습을 계속 보여주길 바란다.</p>
RAG를 적용한 챗봇 답변	<p>여름아, 캠핑에서 네가 제일 잘한 것은 배드민턴이었어. 배드민턴을 치다가 넘어지면서도 용감하게 다시 일어서는 모습이 정말 멋지웠단다. 그 순간, 엄마는 네가 정말 강하고 용감한 아이임을 느꼈어.</p> <p>엄마는 언제나 네가 용기를 잃지 않고 끝까지 도전하는 모습을 보길 바란다. 사랑해, 우리 여름아.</p> <p>엄마는 네가 배드민턴을 더 잘해 보길 바란다. 엄마는 언제나 네 곁에서 응원할게. 사랑해, 우리 여름아.</p>

그림 2. 챗봇 Q&A 결과 2

그림 2에서는 RAG를 적용한 챗봇이 질문의 맥락을 더 잘 이해하고, 육아일기 데이터에 있는 부모님의 말투를 더욱 자연스럽게 재현하여 응답의 품질이 더 높았다.

#### IV. 결 론

본 연구를 통해 RAG 모델을 적용한 챗봇이 기존 LLM 기반 챗봇보다 오류 발생률이 낮고, 보다 정확하고 신뢰성 있는 답변을 제공할 수 있음을 확인하였다. RAG는 할루시네이션 현상을 억제하고, 외부 정보 검색을 통해 폭넓고 정확한 응답을 가능하게 한다. 특히, 사망한 부모를 대신하는 챗봇이라는 특수한 맥락에서, 육아일기 데이터를 활용한 검색 기법이 효과적으로 적용되어 신뢰성 있는 결과를 도출하였다. 이는 향후 RAG 기반 시스템이 추가적인 기술 접목을 통해 더욱 향상된 성능을 보일 가능성을 시사한다.

#### 참 고 문 헌

[1] 정천수. "LLM 애플리케이션 아키텍처를 활용한 생성형 AI 서비스 구현: RAG모델과 LangChain 프레임워크 기반." *지능정보연구* 29.4 (2023): 129-164.

[2] Meshram, Siddhant, et al. "Conversational AI:

Chatbots." *2021 International Conference on Intelligent Technologies (CONIT)*. IEEE, 2021.

[3] 이광우(Gwang-Wu Yi), and 김수균(Soo Kyun Kim). "국내 기업을 위한 RAG 구조 기반 질의 응답시스템 설계." *韓國컴퓨터情報學會論文誌* 29.7 (2024): 81-88.

[4] SHUSTER, Kurt, et al. Retrieval augmentation reduces hallucination in conversation. *arXiv preprint arXiv:2104.07567*, 2021.

[5] LEWIS, Patrick, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 2020, 33: 9459-9474.

# RAG 기반 LLM을 적용한 공지사항 챗봇 시스템 개발

고상희\*, 유승종\*, 이석현\*, 조현준\*, 최기영\*, 이동욱\*\*, 정설영\*\*\*

## Development of Announcement Chatbot System with RAG-based LLM

Sanghui Ko\*, Seungjong Yu\*, Seokhyun Lee\*, Hyunjoon Cho\*, Kiyeong Choi\*, Dongwook Lee\*\*, and Seol-young Jeong\*\*\*

### 요약

RAG(Retrieval Augmented Generation) 기반 시스템은 LLM(Large Language Models)에서 발생하는 여러 문제를 완화하며, 외부 데이터 소스를 실시간으로 검색하여 더욱 신뢰할 수 있는 답변을 생성한다. 본 논문은 경북대학교 컴퓨터학부의 공지사항을 대상으로 RAG기술을 활용한 질의응답 시스템을 제안한다. 공지사항 데이터를 크롤링하고 이를 벡터로 임베딩하여 검색할 수 있는 QA chatbot 시스템을 구축하였고, 실험 결과 해당 시스템은 기존의 챗봇 시스템에 비해 더 정확한 답변을 제공하는 것으로 나타났다.

### Abstract

The RAG-based system alleviates a number of problems arising from Large Language Models (LLM) and generates more reliable answers by searching external data sources in real time. This paper proposes a Q&A system using RAG (Retrieved Augmented Generation) technology for announcements in the Department of Computer Science at Kyungpook National University. We built a QA chatbot system that can crawl announcement data and embed it as a vector to search for it, and the experiment showed that the system provides more accurate answers than the existing chatbot system.

### Key words

rag, langchain, llm, qa system, chatbot, web crawling, vectorstore, embedding

---

\* 경북대학교 컴퓨터학부 학사과정, gsh244400@naver.com, bigbell999@knu.ac.kr, tjrgus0703@gmail.com, bluelol5478@gmail.com, rldud1237@knu.ac.kr

\*\* (주)데이터스트림즈 수석연구원, dwlee@datastreams.co.kr

\*\*\* 경북대학교 컴퓨터학부 교수(교신저자), snowflower@knu.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2021-0-01082)



## 1. 서론

정보의 빠르고 정확한 전달은 공공기관과 교육기관의 중요한 과제이다. 특히 학교의 공지사항은 학생들이 자주 확인하는 필수적인 정보이며, 이를 효율적으로 관리하고 제공하는 시스템의 필요성은 높아지고 있다. 이러한 맥락에서 자연어 처리 기술을 활용한 질문응답 시스템(QA system)은 공지사항에 대한 자동 응답과 정보 검색을 가능하게 한다. 인간과 기계 간의 상호작용이 점차 발전하면서, LLM을 기반으로 한 대화형 챗봇 시스템이 우리의 일상 생활에서 더욱 중요한 역할을 하고 있다[1]. 그러나 LLM 기반 시스템은 개인 정보 보안 문제, 정확도, 환각 현상 등으로 인해 여러 문제점이 발생할 수 있다[2].

이러한 문제를 해결하기 위해 RAG 모델이 주목받고 있다. RAG는 검색 기반 모델을 생성형 언어 모델에 결합하여 생성된 텍스트의 질과 정확성을 향상시키는 기술이다[3]. 이를 통해 LLM의 단점을 보완하고, 최신 정보에 대한 접근성을 높일 수 있다. 본 연구에서는 경북대학교 컴퓨터학부의 공지사항을 대상으로 RAG 모델을 적용하여 질문응답 시스템을 구축하고, 그 성능을 평가한다.

## II. 시스템 구성

이 시스템은 대학 웹사이트에서 공지사항을 스크래핑하여, 각 공지사항의 주요 텍스트 내용과 게시 날짜, URL과 같은 메타데이터를 추출한다. 이후 텍스트는 보다 정밀한 검색을 가능하게 하기 위해 작은 청크로 분할된다. 각 텍스트 청크는 두 가지 유형의 임베딩을 생성하여 인덱싱되며, 이는 다음과 같다.

1. 희소 임베딩: TF-IDF 벡터화를 통해 텍스트 청크의 희소 표현을 생성한다.
2. 밀집 임베딩: 임베딩 모델을 사용하여 텍스트의 밀집 벡터 표현을 생성한다.

임베딩 모델에서는 ollama와 upstage 두 가지 방식으로 접근하였다.

Ollama는 모델을 쉽게 배포하고 사용할 수 있는 환경을 제공하여, 사용자의 복잡한 설치나 설정 없

이도 쉽게 언어 모델을 활용할 수 있다. 하지만 Ollama는 로컬에서 실행되므로 사용자의 컴퓨터 성능에 직접적인 영향을 받는다. 따라서 로컬 시스템의 자원 한계로 인해 성능이 저하될 수 있다.

반면 Upstage는 한국어 이해와 생성에서 높은 정확도를 가지고 있으며, 한국어의 다양한 뉘앙스를 포착할 수 있는 능력이 뛰어나다. 또한 Upstage는 필요에 따라 경량화된 모델 버전을 제공하여, 낮은 자원에서도 빠르게 동작할 수 있도록 한다.

그림 1은 ollama와 Upstage의 임베딩 모델 성능 평가 결과를 나타낸다. Ollama 실행 시간은 492초, Upstage 실행 시간은 8초로 나타났다.

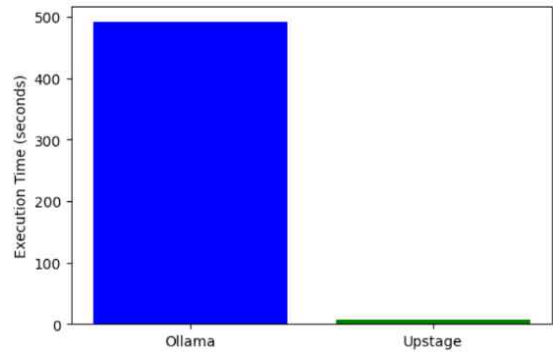


그림 1. ollama와 upstage 성능 비교

두 가지 방법을 비교한 결과, upstage 임베딩 모델이 본 프로젝트에 적합하다는 것을 알 수 있다.

이렇게 생성된 밀집 임베딩은 관련 메타데이터(텍스트 내용, URL, 날짜)와 함께 효율적인 유사도 검색을 위해 Pinecone 벡터 데이터베이스에 저장된다.

사용자가 쿼리를 제출하면 시스템은 다음 단계를 통해 이를 처리한다. 먼저 쿼리를 TF-IDF 벡터로 변환한 후, 저장된 문서 벡터와의 코사인 유사도를 계산하여 희소 검색을 수행한다. 이후 Upstage 모델을 통해 쿼리를 임베딩하고, Pinecone 인덱스에서 밀집 유사도 검색을 수행한다. 최종적으로 시스템은 희소 검색과 밀집 검색의 결과를 결합하여 가장 관련성 높은 문서를 식별하는 하이브리드 검색 접근 방식을 사용한다. 이를 통해 두 가지 검색 방법의 장점을 모두 활용하여 검색의 정확도를 향상시킨다.

그림 2는 본 시스템의 전체 구성도를 보여주는 다이어그램이다.

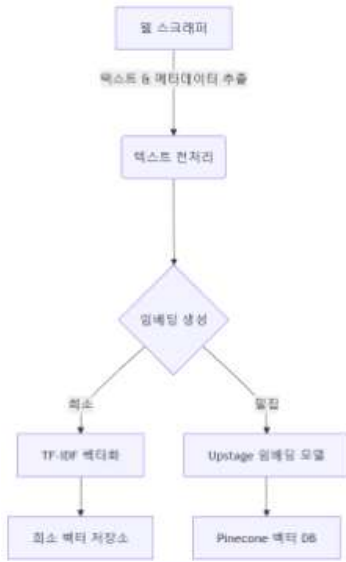


그림 2. 시스템 다이어그램

### III. 구현 결과

본 시스템과 기존 시스템의 성능을 비교하기 위해 동일한 질문 세트를 사용하여 다음과 같은 실험을 진행했다.

**전통적 질문응답 시스템:** 미리 학습된 사전 훈련 모델을 기반으로 질문에 대한 답변을 생성한다. 이 시스템은 추가적인 문서 검색 없이 모델 내에서 답변을 생성한다.

**RAG 기반 질문응답 시스템:** 질문이 들어오면 관련된 문서를 검색하고, 검색된 문서를 바탕으로 답변을 생성한다.

그림 3과 4는 각각 전통적 질문응답 시스템의 실험 결과와 RAG 기반 질문응답 시스템의 실험 결과를 나타낸다.

그림 3에서 RAG 모델을 적용하지 않은 챗봇은 사전 훈련된 지식에만 의존하기 때문에, 실시간으로 업데이트되는 정보에는 접근하지 못한다. 즉, 경북대학교 TOPCIT 시험 날짜나 튜터 관련 공지사항 같은 질문을 받았을 때, “해당 페이지를 방문해 최신 공지를 확인하세요.”와 같은 반응을 보인다. 최신 정보를 다루지 못하는 한계가 있어 사용자가 원하는 구체적인 정보를 제공하지 못하기 때문이다.

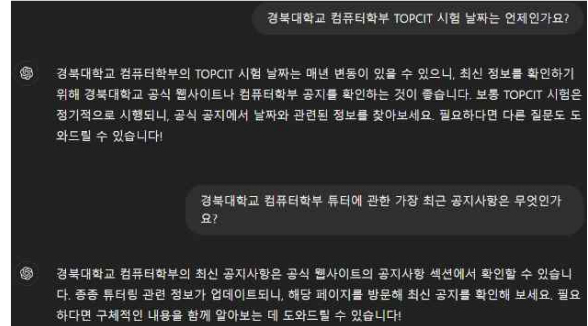


그림 3. RAG 모델을 적용하지 않은 챗봇



그림 4. RAG 모델을 적용한 챗봇

반면, RAG가 적용된 챗봇은 실시간으로 공지사항 데이터베이스에 접근하여 정보를 검색하고 제공할 수 있다. 사용자의 질문이 들어오면, 챗봇은 미리 수집된 경북대학교 컴퓨터학부 공지사항 데이터를 바탕으로 최신 정보를 검색한 후 적절한 답변을 제공한다. 예를 들어, 경북대학교 탐싯 시험 날짜나 수강 변경에 대한 정보에 대해 질문하면 챗봇이 공지사항에서 해당 정보를 찾아 정확한 날짜와 내용을 알려준다.

### IV. 결론

본 연구에서는 경북대학교 공지사항을 대상으로 RAG 기반 질문응답 시스템을 구축하고 그 성능을 평가하였다. RAG 모델은 기존의 LLM 기반 챗봇 시스템에서 발생할 수 있는 환각 문제를 줄이며, 검

색된 공지사항 데이터를 바탕으로 더욱 신뢰성 있는 답변을 생성할 수 있었다. 그러나 성능 향상을 위해 이미지 임베딩의 통합과 결과 결합 방법 개선이 필요하다. 특히, 이미지 임베딩을 통해 공지사항에 포함된 시각적 자료에 대한 이해를 높이고 검색 결과에 이미지 관련 정보를 포함시킬 수 있다[4]. 이를 위해 CLIP(Cross-Lingual Image Pretraining)과 같은 모델을 활용하여 텍스트와 이미지를 동시에 벡터화하는 방법이 유용할 것이다. 또한, 희소 검색과 밀집 검색의 가중치를 조절함으로써 쿼리의 특성에 맞춰 결과의 정확도를 향상시킬 수 있다. 향후 연구에서는 이러한 개선점을 실험하고 실제 성능을 측정하여 챗봇의 응답 정확성 및 속도를 더욱 향상시킬 수 있을 것으로 기대된다.

## 참 고 문 헌

- [1] 신성필. 초거대 AI 의 기반모델 (Foundation Model) 개념 및 표준화 동향. 한국통신학회지 (정보와통신), 2023, 40.6: 12-21.
- [2] 정효정(Alice Jung),송주현(Juhyun Song),서상훈(Sanghoon Seo),임진호(Jinhyo Lim),이현상(Hyunsang Lee),and 김동균(Dongkyun Kim). "RAG 기반 LLM 성능 평가 및 검증을 위한 LangChain 활용 RAGA 방법론 연구." 대한전자공학회 학술대회 2024.6 (2024): 2567-2572.
- [3] 조찬영(Chanyoung Jo),강성준(Seongjun Kang),and 정현준(Hyunjun Jun). "RAG기반 랭체인을 이용한 생성형 AI 챗봇 구현." Proceedings of KIIT Conference 2023.11 (2023): 460-463.
- [4] 전준형(Joon Hyoung Jun),김상철(Sang-Chul Kim), 김주철(Joo Chul Kim),and 윤성준(Seong Joon Yoon). "웹 애플리케이션 서버(WAS)에서의 검색 증강 생성(RAG) 기술을 이용한 지식 기반 QA 문제 해결." 한국통신학회 학술대회논문집 2024.1 (2024): 1109-1114.

# 외국 관광객을 위한 서울 관광 안내 지도 시스템

황재민\*, 송민석\*\*, 마지원\*\*, 이민희\*\*, 이규원\*\*\*, 이해연\*\*\*\*

## Seoul Tourism Guidance Map System for Foreign Tourists

Jae-min Hwang\*, Min-seok Song\*\*, Ji-won Ma\*\*, Min-hee Lee\*\*,  
Kyu-Won Lee\*\*\*, and Hae-Yeoun Lee\*\*\*\*

### 요약

서울은 대한민국의 수도이자 외국인들이 가장 많이 찾는 관광지 중 하나이다. 그러나 외국인들이 서울의 다양한 관광지를 검색하고 선택하는 과정에서 불편함을 겪는다. 이를 해결하기 위해 본 논문에서는 서울관광재단이 제공한 공공데이터를 활용한 영문 기반의 서울 관광 안내 지도 시스템을 개발에 대하여 설명한다. 이 시스템의 전반적인 UI는 영문으로 되어있고 사용자는 지도에서 관광지를 직관적으로 탐색하고, 카테고리별 필터링과 검색 기능을 통해 원하는 정보를 손쉽게 찾을 수 있다. 또한, 사용자 스크랩 기능을 통해 개별 관광지를 저장하고 관리할 수 있도록 설계되었다.

### Abstract

Seoul is the capital of Korea and one of the most visited tourist destinations by foreigners. However, foreigners experience inconvenience in searching and selecting various tourist destinations in Seoul. To solve this problem, this paper explains the development of an English-based Seoul tourism guidance map system using public data provided by the Seoul Tourism Foundation. The overall UI of this system is in English, and users can intuitively search for tourist destinations on a map and easily find desired information through filtering and search functions by category. In addition, it is designed to store and manage individual tourist destinations through the user clipping function.

### Key words

Foreign Tourists, Seoul, Website, Tourism map

## 1. 서론

서울은 대한민국의 수도이자 주요 관광지가 밀집한 지역으로, 많은 외국인 관광객들이 주로 서울을 방문한다. 비록 그림 1과 같이 코로나19로 인해 관

광객 수가 감소했지만, K-POP을 비롯한 K-콘텐츠에 대한 관심이 높아지면서 한국에 대한 세계적 관심은 여전히 증가하고 있다.

그러나 외국인들이 서울의 수많은 관광지를 검색하고 선택하는 과정은 쉽지 않다.

\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생, redmin852@naver.com

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생

\*\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수

\*\*\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수, haeveoun.lee@kumoh.ac.kr(교신저자)



그림 1. 연도별 서울 방문 외국인 관광객 [1]  
Fig. 1. Foreign tourists visiting Seoul by year

이와 같은 외국인의 불편함을 해결하기 위해 본 연구에서는 서울관광재단[2]에서 제공한 공공데이터를 활용하여 외국인 관광객을 위한 서울 관광 안내 지도를 구축하였다. 본 시스템은 외국인을 대상으로 영문 기반으로 설계되었으며, 직관적인 지도 상의 마커 기능을 통해 관광지를 시각적으로 탐색할 수 있다. 사용자는 카테고리 필터 및 키워드 검색을 통해 원하는 관광지를 쉽게 찾을 수 있으며, 관광지에 대한 상세 정보와 경로 안내를 제공하여 외국인들이 편리하게 여행 계획을 세울 수 있도록 돕는다. 또한, 스크랩 기능을 통해 관심 관광지를 저장하고 관리할 수 있는 맞춤형 서비스를 제공한다.

본 논문의 구성은 다음과 같다. 2장에서는 시스템 개발의 구체적인 절차와 주요 기능을 설명하고, 3장에서는 개발된 시스템의 결과를 나타낸다. 마지막으로 4장에서 결론과 향후 개선 방안을 제시한다.

## II. 서울 관광 안내 지도 개발

### 2.1 시스템 개요

본 시스템은 외국인 관광객을 대상으로 하며, 모든 UI는 직관적인 영문 인터페이스로 설계되었다. 사용자는 관광지를 쉽게 탐색하고 검색할 수 있도록 다양한 카테고리 필터와 필터 기능이 제공된다. 또한, 본 시스템은 그림 2와 같이 서버-클라이언트 구조로 구성되어 있어, 효율적인 데이터 처리가 가능하다. 클라이언트 측은 React를 이용하여 웹 애플리케이션의 인터페이스와 사용자 상호작용을 담당하며,

사용자 요청을 처리하여 결과를 보여준다. 서버 측은 Node.js와 Express를 기반으로 구현되었으며, MySQL 데이터베이스와 연결되어 공공데이터를 제공하고, 클라이언트의 요청에 따라 적절한 데이터를 반환하는 역할을 수행한다. 서버는 클라이언트로부터 검색 쿼리, 카테고리 필터 등의 요청을 받아 처리한다.

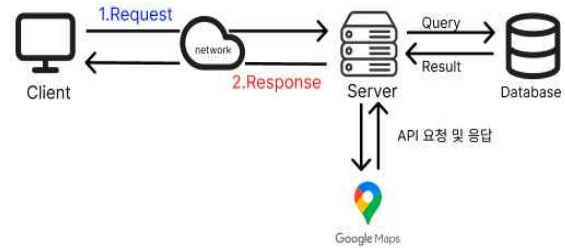


그림 2. 개발 시스템의 클라이언트-서버 구조  
Fig. 2. Client-Server Structure of Developed System

### 2.2 주요 기능 설명

서울관광재단에서 제공한 공공데이터의 위도 및 경도 정보를 기반으로, 각 관광지는 지도에 마커로 표시된다. 사용자는 해당 마커를 클릭하여 관광지 상세 정보를 확인할 수 있다. 사용자는 특정 키워드로 관광지를 검색할 수 있으며, 카테고리 필터를 통해 원하는 유형의 관광지(자연, 음식, 명소, 문화, 쇼핑)를 선택할 수 있다. 이 카테고리는 서울관광재단에서 제공한 공공데이터의 분류를 이용하였다. 사용자는 관광지 스크랩 기능을 이용하여 원하는 관광장소만 스크랩을 할 수 있고 사용자 마이페이지를 통해 스크랩 항목을 이용할 수 있다.

### 2.3 데이터 처리 및 활용

서울관광재단의 공공데이터는 JSON 형식으로 제공되었으며, 데이터의 전처리 과정을 통해 불필요한 정보를 사전에 걸러냈다. 이 데이터는 관광지의 상호명, 주소, 전화번호, 웹사이트, 운영시간, 운영요일, 휴무일, 영문명, 위도, 경도 정보가 포함되어 있어, 이를 활용해 지도 상에 관광지를 표시하고 카테고리화 하였다. 이용된 데이터의 유형별 분류는 표 1과 같다.

표 1. 서울관광재단의 관광지 공공데이터  
Table 1. Public Tourist Site Data of Seoul Tourism Foundation

데이터 이름	규모	갱신일
서울시 관광 음식	5487	매월 4일
서울시 관광 자연	606	
서울시 관광 명소	1914	
서울시 관광 문화	1137	
서울시 관광 쇼핑	786	
총합	9930	

## 2.4 기술 스택 및 구현 과정

관광지 정보는 MySQL 데이터베이스에 저장되며, Node.js와 Express를 통해 React 프론트엔드와 MySQL 데이터베이스 간의 통신을 구현하였다. 지도 API로는 Google Maps API[3]를 사용하였다.

### III. 시스템 개발 결과

그림 3은 개발한 서울 관광 안내 지도 시스템의 메인 화면으로, 직관적인 영문 인터페이스로 개발된 UI를 확인할 수 있다.

사용자는 지도 상의 마커 또는 관광지 목록에서 원하는 항목을 클릭하면, 그림 4에서 보이는 바와 같이 모달 창을 통해 해당 관광지의 상세 정보를 확인할 수 있다.

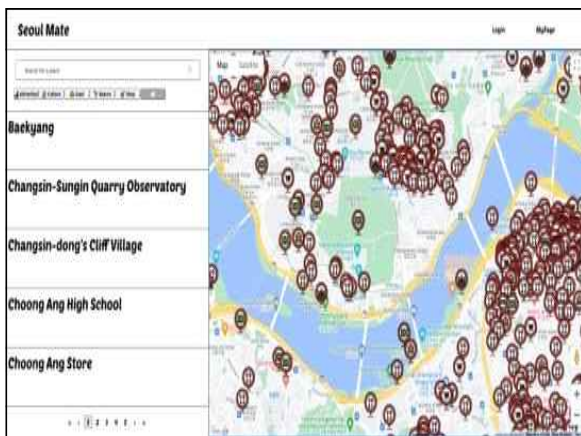


그림 3. 서울 관광 안내 지도 시스템 메인 화면  
Fig. 3. Seoul tourism guidance map system main screen



그림 4. 관광지 세부 정보 화면  
Fig. 4. Tourist Details Screen

또한, 그림 5에서 보이는 카테고리 필터 기능을 통해 사용자는 관광지 유형(자연, 음식, 명소 등)을 선택하여 손쉽게 원하는 관광지를 탐색할 수 있다. 이 기능은 사용자가 관심 있는 관광지만을 필터링하여 더 빠르고 편리하게 정보를 얻도록 돕는다.

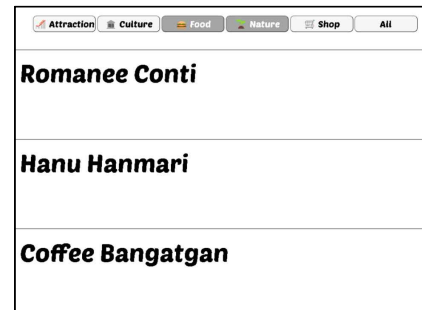


그림 5. 카테고리 선택 화면  
Fig. 5. Category selection screen

그림 6은 마이페이지에서 스크랩 기능이 구현된 화면을 보여준다. 사용자는 특정 관광지를 스크랩하여 저장할 수 있으며, 스크랩된 관광지 리스트에서 선택한 항목에 대한 상세 정보도 확인할 수 있다.

### III. 결 론

본 논문에서는 외국인 관광객을 대상으로 한 서울 관광 안내 지도 시스템을 개발하였다. 이 시스템은 서울관광재단에서 제공한 공공데이터를 기반으로, 외국인들이 서울의 다양한 관광지를 효율적으로

탐색하고 계획할 수 있도록 설계되었다. 특히 외국인 관광객이 직면하는 검색의 복잡성을 해결하고, 서울 여행을 더 쉽게 계획할 수 있도록 돕는 데 기여할 수 있다. 그러나 현재는 영어로만 서비스가 제공되고 있어 다양한 언어 지원이 필요하며, 사용자 맞춤형 관광지 추천 기능 등 향후 개선의 여지가 있다.

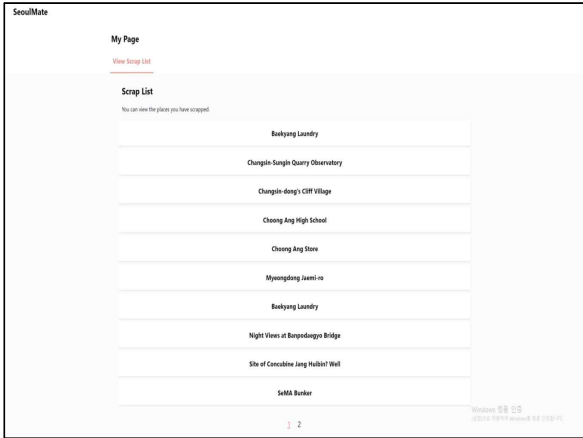


그림 6. 마이 페이지 화면  
Fig. 6. My page screen

## 참 고 문 헌

- [1] Korea Tourism Organization. "Yearly Statistics of Foreign Visitors to Korea.", <https://datalab.visitkorea.or.kr/datalab/portal/ts/getEntcnyFrngCust2Form.do>. Accessed [2024.10.20]
- [2] Seoul Open Data Plaza. "Seoul Tourism Public Data", <http://data.seoul.go.kr/>, 2023. Accessed [2024.10.20]
- [3] Google, "Google Maps JavaScript API", <https://developer.google.com/maps>, 2023. Accessed [2024.10.20]

# N-그램 및 임베딩 기반 표절 탐지 성능 비교

서동섭\*, 손경락\*, 박상우, 한동균, 박세진\*\*

## Performance Comparison of N-gram and Embedding-based Similarity for Plagiarism Detection

Dong-Seop Seo\*, Kyeong-Rak Son\*, Sang-Woo Park, Dong-Gyun Han, and Se-Jin Park\*\*

### 요 약

본 실험은 n-그램 기반 표절 탐지 기법과 임베딩 기반 유사도 계산 기법의 성능을 비교하고 두 기법의 차이점과 장단점을 분석한다. n-그램 기법은 텍스트의 표면적 유사도를 중점적으로 평가하는 반면 임베딩 기법은 문맥적 의미를 반영하여 문장 간의 의미적 유사성을 분석한다. 동일한 데이터셋을 사용한 실험 결과 n-그램 방식은 문장 구조 변화에 반응하는 반면, 임베딩 방식은 정교한 의미적 유사도 분석을 제공하여 표절 여부를 정확하게 판단할 수 있다는 것을 확인하였다. 특히 임베딩 방식은 문장의 구조 변화에도 불구하고 문맥적 유사성을 고려할 수 있어 표절 검사의 정확도를 높이는 데 유리함을 보여주었다.

### Abstract

This experiment compares the performance of n-gram-based plagiarism detection techniques and embedding-based similarity calculation techniques to analyze the differences, advantages and disadvantages of the two techniques. While the n-gram technique focuses on evaluating the superficial similarity of the text, the embedding technique analyzes the semantic similarity between sentences by reflecting contextual meaning. Experiments using the same data set confirmed that while the n-gram method responds to changes in sentence structure, the embedding method provides sophisticated semantic similarity analysis to accurately determine whether plagiarism exists. In particular, the embedding method showed that contextual similarities can be considered despite the structural changes in the text, which is advantageous for improving the accuracy of plagiarism tests.

### Key words

n-gram, embedding, similarity, semantic, plagiarism

## 1. 서 론

현대의 정보화 시대에서 텍스트 유사도 분석은 다양한 분야에서 활용되고 있으며, 그중 하나가 표절 검출이다. 특히 학계에서 표절 문제는 중요한 이슈로 인식되어, 표절을 검사할 수 있는 프로그램

개발이 이루어지고 있다. 이 중 카피킬러는 텍스트 유사도 분석을 기반으로 표절 여부를 확인해주는 도구로 널리 사용되고 있다. 카피킬러는 주로 n-그램 기반의 유사도 분석 기법을 활용하며 이는 n개의 단어로 분리하여 문장의 연속성을 파악하는 방법으로 널리 채택되고 있다.

\* 계명대학교 컴퓨터공학과, ehdtjq1278@naver.com, rudfkrlife@naver.com, ja513698@naver.com, hdg5639@naver.com,

\*\* 계명대학교 컴퓨터공학과, baksejin@kmu.ac.kr



그러나 문장 간의 단순한 n-그램 유사도만으로는 문맥적 유사성을 정확하게 검사하는 데 한계가 있다. 이를 보완하기 위해 최근 AI의 발전과 함께 임베딩 기법을 활용한 의미적 유사도 분석 방법이 표절 검사에 점진적으로 확대되고 있다. 임베딩 기법은 텍스트를 고정된 크기의 벡터로 변환한 후 그 벡터 간의 거리를 통해 문장 간의 의미적 유사도를 평가하는 방법이다. 임베딩 기반의 표절 검사는 기존의 표절 검사 방법과 달리 문장이나 문서의 의미적 유사성을 고려할 수 있어 보다 정교한 분석이 가능하다.

본 논문에서는 n-그램과 임베딩을 활용한 표절률 검사 결과를 비교하고 두 방법의 차이점 및 장단점을 분석한다.

## II. n-그램의 원리와 기술

n-그램은 n개의 단어로 구성된 시퀀스를 의미한다[1]. 이는 텍스트를 일정 길이의 단어 묶음으로 분류한 후, 각 단어 묶음이 다른 텍스트에서 얼마나 자주 동일하게 나타나는지를 계산하여 유사도를 측정하는 방법이다. 예를 들어, "본 논문은 표절률 실험 논문이다"라는 문장을 3-그램 방식으로 나누면 "본 논문은", "표절률 실험", "논문이다"와 같이 3개의 단어가 하나의 묶음이 형성된다. 이 묶음들이 다른 문장에서 동일하게 나타나는지를 비교하여 유사도를 계산하는 방식이다.

이러한 n-그램 유사도 분석을 기반으로 문서 간의 유사도를 수치화하여 표절 여부를 판단한다. 그러나 문장 내 전체 단어를 고려하는 모델과 비교할 때 n-그램 방식은 낮은 정확도를 보이며, 전체 문서에서 n개의 단어가 연속적으로 포함된 문장의 수가 적어지는 희소 문제와 같은 한계점이 존재한다.

## III. 임베딩 기법과 표절률 검사

임베딩은 문자를 벡터로 표현하는 방법 중 하나로, 수치형 데이터로 변환하여 표현하는 기법이다[5]. n-그램 방식과 달리 임베딩 기법은 문장이나 문서의 의미를 반영하는 벡터를 생성한다. 이러한 벡터는 문장 내의 문맥적 의미를 파악할 수 있게 해

주며, 겉으로는 다른 단어를 사용하더라도 동일하거나 유사한 의미를 가진 문장들을 보다 정확하게 비교할 수 있다[5].

벡터 사이의 코사인 유사도를 계산하여 두 문장이 의미상으로 얼마나 유사한지를 파악할 수 있다. 코사인 유사도는 두 벡터 간의 각도를 측정하는 방법으로, 측정값이 1에 가까울수록 두 문장이 유사하다고 판단한다.

임베딩을 활용한 표절률 검사는 기존의 n-그램 방식보다 의미적 유사성을 더욱 잘 반영할 수 있어 단순히 단어가 일치하지 않더라도 문장의 의미적 유사도를 기반으로 표절 여부를 판단할 수 있다. 이는 n-그램 표절 검사 방식이 인지하지 못한 부분이나 유사한 문장 표현에 대해서도 보다 정확한 분석을 가능하게 한다.

## IV. 표절률 실험

본 실험에서는 n-그램 기반 표절률 검사와 임베딩 기반 표절률 검사의 성능을 비교하기 위해 'Aihub'에 기재된 '논문 요약본'이라는 JSON 형식의 데이터셋을 다운받아 그 중 약 500개 정도의 요약본을 본 실험의 데이터셋으로 활용하였다. 이를 바탕으로 세 가지 다른 주제의 문장을 대상으로 실험을 진행하였으며, 대상이 되는 각각의 데이터는 아래 그림 1, 2, 3과 같다.

① 본 실험은 n-그램 기반 표절 탐지 기법과 임베딩 기반 유사도 계산 기법의 성능을 비교하고 두 기법의 차이점과 장단점을 분석한다. n-그램 기반은 텍스트의 표면적 유사도를 중점으로 평가하는 반면 임베딩 기법은 문맥적 의미를 반영하여 문장 간의 의미적 유사성을 분석한다. 동일한 데이터셋을 사용한 실험 결과 n-그램 방식은 문장 구조 변화에 반응하는 반면, 임베딩 방식은 정교한 의미적 유사도 분석을 제공하여 표절 여부를 정확하게 판단할 수 있다는 것을 확인하였다. 특히 임베딩 방식은 문장의 구조 변화에도 불구하고 문맥적 유사성을 고려할 수 있어 표절 검사의 정확도를 높이는 데 유리함을 보여주었다.

그림 1. 본 논문 요약

② 컴퓨터 교육은 실생활의 문제해결을 추구하면서도 문제해결의 필요에 의해 기능을 활용하기보다 유명한 회사의 응용프로그램의 기능적 요구에 의해 단순히 따라하면서 기능을 습득하도록 하는 행동주의적 방향으로 이루어지고 있다. 이에 본 연구는 좀 더 효과적으로 컴퓨터 소양교육을 하면서도 실제 생활과 괴리되지 않는 컴퓨터 교수-학습 모형을 개발하는데 그 목적을 둔다. 구성주의의 관점에서 학습자의 활동을 중시하고 자기 점검적 초민지를 이용하여 각 수업의 단계에 성찰을 구조화 하며, 보다 효과적으로 기능을 습득하고 활용할 수 있도록 하기 위한 문제 중심의 성찰적 협력학습 방법을 강구하여 모형을 개발하였다. 이를 실제 기능교육에 적용해 본 결과 본 모형을 적용한 집단이 그렇지 않은 집단에 비해 기능 학습의 성취도가 높았고, 상위집단, 중간집단, 하위집단의 성취 결과를 분석한 결과 상위 집단에 더 효과적임을 확인하였다.

그림 2. 데이터셋에서 발췌한 문장

③ 컴퓨터 교육은 현실 문제를 해결하려는 목표를 추구하면서도, 문제 해결을 위한 기능의 활용보다는 유명 기업의 응용 프로그램에 요구되는 기능을 단순히 따라하며 배우는 행동주의적인 접근으로 진행되고 있다. 본 연구는 이러한 한계를 극복하고, 실생활과 밀접하게 연계된 컴퓨터 교육 모형을 개발하는 것을 목적으로 한다. 이 모형은 구성주의 관점에서 학습자의 적극적인 참여를 강조하고, 자기 점검을 통해 성찰하는 초인지 전략을 수업 단계에 적용하여, 학습자가 더 효과적으로 기능을 습득하고 사용할 수 있도록 하는 문제 중심의 성찰적 협력 학습 방식을 제안한다. 이를 실제 기능 교육에 적용한 결과, 해당 모형을 사용한 학습 집단이 그렇지 않은 집단보다 기능 학습 성취도가 높았으며, 상위, 중간, 하위 집단을 분석한 결과 상위 집단에서 특히 더 큰 효과를 보였다.

그림 3. 그림 2에서 문장 구조만 수정한 문장

실험에서는 n-그램 방식과 임베딩 기반 방식을 활용하여 각 문장에 대한 표절률을 비교하였다. 그래프에서 파란색은 n-그램 방식에 의한 표절률을, 노란색은 임베딩 기반 방식에 의한 표절률을 나타내며, 표절률의 기준은 0.5 이하의 유사도는 “정상”, 0.5~0.7의 유사도는 “표절 의심”, 0.71 이상의 유사도는 “표절”로 분류하였다. 각 경우에 대한 실험 결과는 아래와 같이 제시된다.

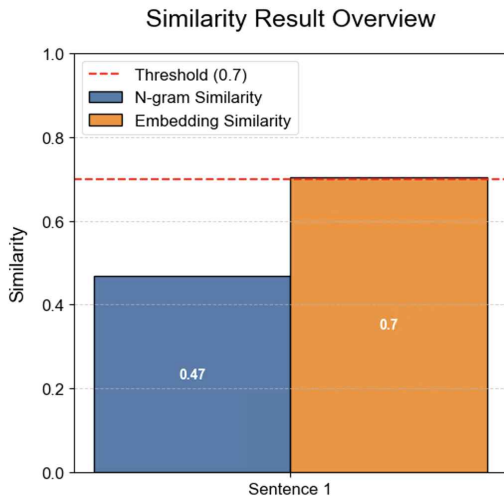


그림 4. 그림 1 문장에 대한 실험 결과

그림 4에서는 본 논문의 요약본을 대상으로 하였기에 비교적 낮은 n-그램 표절률과 임베딩 기반 표절률이 나타나는 것을 알 수 있다.

그림 5는 그림 2에 포함된 문장을 데이터셋 내의 문장과 비교한 결과를 나타내며, 이로 인해 표절률이 1.0(100%)으로 산출된 것을 확인할 수 있다. 이는 그림 2의 문장이 그대로 데이터셋에 존재한다는 것을 의미하며, 해당 문장은 기존 자료와 완벽히 일치하는 것으로 분석된다.

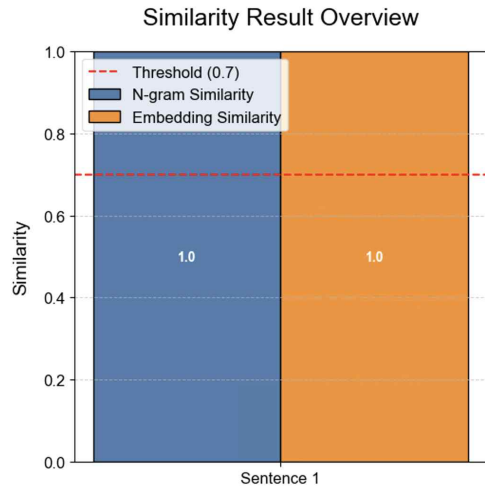


그림 5. 그림 2 문장에 대한 실험 결과

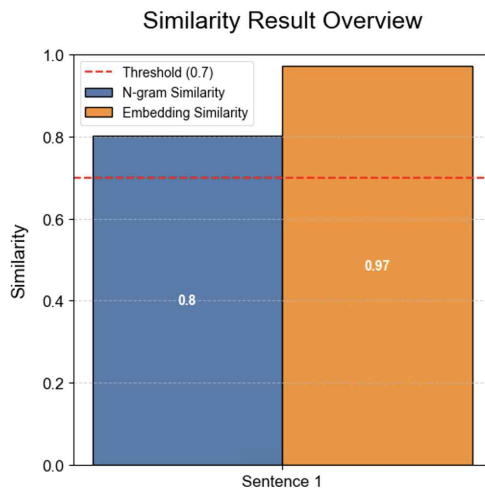


그림 6. 그림 3 문장에 대한 실험 결과

그림 6은 그림 3에 대한 표절률 검사 결과를 보여준다. n-그램 기반 표절 검사는 문장의 구조를 중심으로 비교하기 때문에, 문장 구조가 다를 경우 표절률이 감소하는 경향이 있어 표절률이 다소 낮은 것을 확인할 수 있다. 그러나, 임베딩 기반 표절 검사에서는 문장의 의미를 바탕으로 비교하기 때문에 구조적인 변형이 있더라도 동일한 의미를 가진 문장으로 인식된다. 따라서 그림 6에서도 임베딩 기반 표절률은 그림 5와 거의 동일하게 나타났다.

## V. 결론

본 실험에서는 n-그램 방식과 임베딩 기반 표절 검사의 성능 차이를 비교하였다. 결과적으로 n-그램

방식은 문장의 구조적 변화에 민감하게 반응하지만 임베딩 방식은 문장의 의미적 유사성을 반영하여 더 정확하게 표절 여부를 판단하는 데 유리함을 확인하였다. 특히 임베딩 방식은 문맥적 유사성까지 고려하여 단순한 형식 변화를 넘어선 표절을 탐지하는 능력을 보여주었다. 이러한 결과는 임베딩 기반 방식이 n-그램의 한계를 보완하며 더 정교한 표절 검사를 가능하게 함을 시사한다. 향후 연구에서는 다양한 데이터셋을 통해 임베딩 기법의 성능을 추가로 검증할 필요가 있다.

텍스트 임베딩 모델 연구. 서울대학교 대학원, 융합과학기술대학원 융합과학부(디지털정보융합 전공).

- [10] 이다빈, 최성필 "대용량 텍스트 자원을 활용한 한국어 형태소 임베딩의 모델별 성능 비교 분석" 정보과학회논문지 46.5 pp.413-418 (2019) : 413.

### 참 고 문 헌

- [1] 정이수, 김윤비, 조영래. (2023). 워드 임베딩 기반의 단백질 간 유사도 측정을 통한 단백질 상호작용과 약물-단백질 상호작용 예측. 정보과학회 컴퓨팅의 실제 논문지, 29(4), 178-183, 10.5626/KTCP.2023.29.4.178
- [2] 김승주. (2019). 논증적 글쓰기 과제의 자동 채점을 위한 N-gram 자질의 활용 가능성 탐색. 작문연구, (41), 37-62.
- [3] <https://wikidocs.net/33520>
- [4] <https://huggingface.co/jhgan/ko-sroberta-multitask>
- [5] 이정인, 안진희, 고경택, 김영석. (2023). 워드 임베딩(Word Embedding)을 활용한 최적의 키워드 추출 및 검색 방법 연구. 한국지반신소재학회 논문집, 22(2), 47-54.
- [6] Kim, C.-R., Park, J., & Park, D.-S. (2017). Anti-Plagiarism Systems for Korean Alphabets. *Advanced Science Letters*, 23(12), 12678-12682.
- [7] 김민범, 신주현. (조선대학교 산업기술융합대학원 소프트웨어융합공학과). 유의어 가중치를 적용한 논문 표절 유사도 측정 방법.
- [8] Maeng, J., Park, J. S., & Shon, J. G. (2019). Implementation of A Plagiarism Detecting System with Sentence and Syntactic Word Similarities. *KIPS Transactions on Software and Data Engineering*, 8(3), 109-114. <https://doi.org/10.3745/KTSDE.2019.8.3.109>
- [9] 김성현(서봉원). (2018). 단어 임베딩을 활용한

# 시각장애인을 위한 요리 보조 안경

편대희\*, 조원근\*, 박건우\*, 신동석\*, 이관형\*\*

## Cooking Assistance Glasses for the Visually Impaired

Daehee Pyun\*, Wongeun Jo\*, Geonwoo Park\*, Dongseok Sin\*, and Kwanhyeong Lee\*\*

### 요약

인간의 모든 감각 수용체의 70%에 해당하는 시각은 오감 중 가장 중요한 감각으로서 시각장애인들이 스스로 요리 하기 위해서는 많은 어려움이 있다. 본 연구에서는 시각장애인이 원활하게 요리할 수 있도록 하기 위해서, 주변환경을 인식할 수 있는 시각장애인을 위한 요리 보조 안경을 제작한다. 주변환경 인식 검출은 카메라로 촬영한 영상을 YOLO algorithm을 사용하여 처리하고, 시각장애인이 요리 보조 안경 착용 시 무게를 감소시키기 위한 방법을 제안한다.

### Abstract

Vision, accounting for 70% of all sensory receptors, is the most critical sense among the five senses. Visually impaired individuals face numerous challenges when cooking independently. This study aims to develop cooking assistant glasses for visually impaired individuals that can recognize the surrounding environment. The environmental recognition detection is processed using the YOLO algorithm on the footage captured by the camera. Additionally, we propose a method to reduce the weight of the glasses worn by visually impaired individuals.

### Key words

assistive device, object detection, conversational voice A.I.

## 1. 서론

보건복지부가 2020년 실시한 장애인 실태조사에 따르면 영양을 고려하여 식품을 골고루 섭취하는 시각장애인은 41.2%로, 일정한 시간에 식사하지 않는 시각장애인은 29.4%로 나타났다[1]. 이는 식자재의 라벨에 시각장애인을 위한 영양 정보 표시가 없고 장애로 인해 음식을 조리하는 데 큰 어려움을 겪기 때문에 시각장애인들이 대사 장애, 영양 불량

문제를 겪고 있음을 시사한다.

본 연구의 시각장애인을 위한 요리 보조 안경은 데이터베이스를 통해 음식 레시피를 제공하여 단계별로 세부적인 조리법을 안내한다.

연구방법은 카메라로 촬영한 주변 환경을 A.I.(Artificial Intelligence)가 처리하여 진행상황을 인지함으로써 사용자의 요구 없이도 자율적으로 필요한 정보를 제공한다. 안경의 착용 시간과 무게를 고려하여, 안경의 전원은 배터리로 공급하며 제품의

\* 대전대학교 휴먼로봇융합전공, diavillager@gmail.com, join0115@naver.com, ksspill@naver.com, habaik0147@naver.com,

\*\* 대전대학교 AI융합학부 교수, khlee@daejin.ac.kr

전체적인 모습은 안경과 충전기로만 구성되도록 하여 무게를 감소시키고 착용시간을 증가시키고자 한다. 또한 시각장애인이 원활하게 제품을 사용할 수 있도록 대부분의 기능은 마이크부의 음성 인식을 사용하여 특정 키워드가 인식될 때 동작하거나 정지한다.

표 1과 표 2는 보건복지부가 실시한 2020년 장애인 실태조사이다. 각 표의 이해를 돕기 위해 보고서의 실험군에서 시각장애와 청각장애를 선별하여 정리하였다. 표 1은 지난 일주일 동안 영양을 고려하여 식품을 골고루 섭취하였는지를, 표 2는 장애인의 식사 습관과 관련하여 일정한 시간에 식사하는가를 조사한 결과이다.

표 1. 지난 일주일 동안 영양을 고려하여 식품을 섭취하였는지 여부

Table 1. Whether food intake was considered for nutrition over the past week

(단위 : %, 명)

구분	시각장애	청각장애	전체
예	41.2	41.2	40.2
보통/가끔	50.3	47.0	47.1
아니오	8.6	11.8	12.7
계	100.0	100.0	100.0
전국추정수	252,702	384,668	2,622,950

표 2. 일정한 시간에 식사하는지 여부

Table 2. Whether meals are taken at a consistent time

(단위: %, 명)

구분	시각장애	청각장애	전체
먹음	66.5	71.7	67.8
가끔 불규칙	29.4	24.0	27.1
먹지 않음	4.1	4.3	5.1
계	100.0	100.0	100.0
전국추정수	252,702	384,668	2,622,950

## II. YOLO algorithm과 객체 검출

본 연구에서는 주변의 사물을 인식하기 위해서 객체 검출 YOLO algorithm을 사용한다. 보조 안경의 구성은 주변 사물을 인식하기 위한 카메라부, 착용자의 음성 지시를 수신하기 위한 마이크부, 상황

에 적합한 피드백을 착용자에게 전달하기 위한 스피커부, 그리고 YOLO algorithm의 객체 검지 상태를 산출할 수 있는 제어부로 구성된다. 제어부의 핵심인 MCU(Micro Controller Unit)의 크기가 안경 내부에 들어갈 수 있는 수준으로 제한되는 바, 이를 해결하기 위해 제어부를 안경 내부의 MCU (raspberry pi zero 2 w, 이하 raspberry pi)와 외부의 MCU(NVIDIA jetson nano, 이하 jetson nano)로 분리하여 구성하고 통신부의 무선 통신을 통해 MCU 간의 정보를 공유할 수 있도록 한다. 따라서 raspberry pi는 안경 내부의 제어를, jetson nano는 YOLO algorithm의 계산을 담당한다.

그림 1은 안경(glasses)과 거치형 충전기(charging stand)로 이루어진 제품 구성도이다.

안경은 raspberry pi 기반의 제어부(control unit), jetson nano와 무선으로 정보를 송수신하는 통신부(comm. unit), 사용자의 주변 환경을 촬영하여 제어부로 전송하는 카메라부(camera unit), 사용자의 음성을 녹음하여 제어부로 전송하고 제어부의 피드백을 음성으로 출력하는 스피커/마이크부(speaker/mic. unit), raspberry pi에 전력을 공급하는 배터리(battery)로 구성된다.

거치형 충전기는 jetson nano 기반의 제어부(control unit), raspberry pi와 무선으로 정보를 송수신하고 서버(server)에 필요한 레시피를 요청하는 통신부(comm. unit), yolo algorithm의 객체 검출을 실행하는 연산부(processing unit), 안경의 배터리를 충전하는 충전부(charging port)로 구성된다.

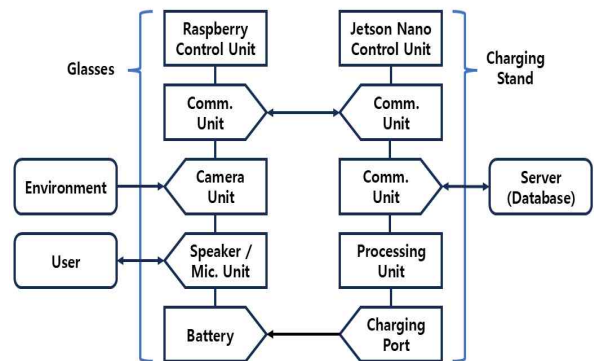


그림 1. 제품 구성도

Fig. 1. Product Configuration Diagram

그림 2는 안경의 정면에서 바라본 3D 도면이다. 카메라부(100)는 사용자의 정면을 촬영하기 위해 안경의 전면부에 위치한다. 제어부와 통신부, 배터리(200)는 안경 내부에 내장된다. 스피커/카메라부(300)는 원활한 음성 녹음 및 피드백 출력을 위해 사용자의 안면부와 가까운 좌우 측면부에 위치한다.

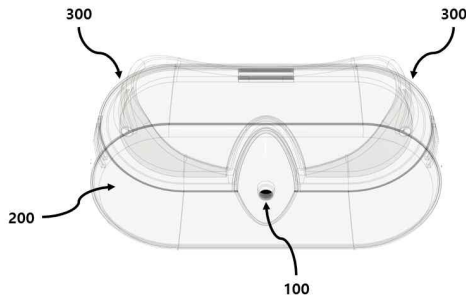


그림 2. 안경 3D 도면

Fig. 2. Glasses 3D Diagram

### III. 결 론

본 논문에서는 객체 검출 알고리즘을 사용하여 시각장애인의 요리를 보조할 수 있는 안경을 제안하였다. 이를 통해 착용자가 시각장애의 제약을 극복하고 주방에서 안전하고 자율적으로 요리를 만들 수 있게 하여 개인의 자존감을 높이고 더 활동적인 일상생활이 가능하다고 기대된다.

### 참 고 문 헌

- [1] Ministry of Health and Welfare, “2020 Survey on the Status of Disabled People”, pp. 259-260, December 2021.

# 주행자를 위한 전동킵보드 안전 시스템

서동현\*, 곽종수\*, 서방원\*\*

## Electric Scooter Safety System for Riders

Dong-Hyeon Seo\*, Jong-Su Kwak\*, and Bangwon Seo\*\*

### 요약

본 논문은 YOLOv5의 커스텀 학습기능을 통해 포트홀, 맨홀, 방지턱, 차량을 학습하고, YOLO와 OpenCV를 사용한 웹캠 객체인식을 이용해 장애물을 감지한다. 운전자가 알 수 있도록 진동과 소리로 경고를 주며, 장애물 종류에 따라 색상(Red, Green, Blue, Yellow)을 출력한다. 전방 카메라는 포트홀, 맨홀, 방지턱을, 후방 카메라는 차량을 인식한다. 이때 차량이 1미터 밖일 경우 소리와 색상 신호를 출력하고 1미터 내로 접근하면 추가 진동 경고로 위험을 알리는 장치이다.

### Abstract

This paper utilizes the custom training feature of YOLOv5 to detect potholes, manholes, speed bumps, and vehicles, using YOLO and OpenCV-based webcam object recognition to identify obstacles. It provides vibration and sound warnings to alert the driver and displays different colors (Red, Green, Blue, Yellow) depending on the type of obstacle. The front camera detects potholes, manholes, and speed bumps, while the rear camera detects vehicles. If the vehicle is more than 1 meter away, sound and color signals are triggered, and if it approaches within 1 meter, additional vibration warnings alert the driver of the danger.

### Key words

Webcam object detection, YOLO, Open-CV

## I. 서론

최근 전동킵보드를 이용하는 사람이 증가하면서 전동킵보드 사고율 또한 기하급수적으로 증가하였다. 그림 1과 그림 2는 각각 연도별 전동킵보드 교통사고 현황과 연령대별 전동킵보드 이용률 및 사고율을 나타낸다. 이 그림으로부터 전동킵보드 이용률과 사고율이 가장 높은 것은 20대인 것을 알 수 있다. 그림 3은 월별 전동킵보드 사고 증감율을 나

타내며, 2~10월의 사고 증감율은 양수이고 그중 가장 높은 시기는 2월이다. 또한 11~1월의 교통사고 증감율은 음수이고 가장 낮은 시기는 1월이다 [1]. 이를 근거로 전동킵보드를 가장 많이 이용하는 것으로 추정되는 20대 대학생들을 대상으로 전동킵보드 이용 시 불편했던 점들을 조사하였고, 해당 문제점들을 보완하고자 전동킵보드 안전 시스템을 구현하였다.

\* 국립공주대학교 전기전자제어공학부,

\*\* 국립공주대학교 전기전자제어공학부 (교신저자: seobw@kongju.ac.kr)



그림 1. 연도별 전동킥보드 교통사고 현황  
Fig. 1. Annual Electric Scooter Traffic Accident Overview

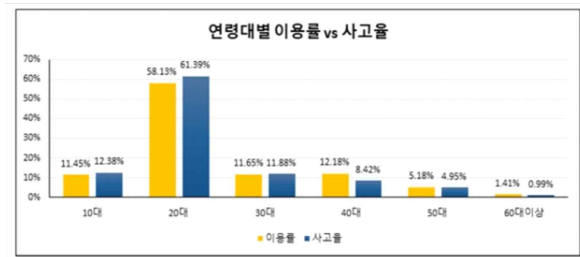


그림 2 연령대별 전동킥보드 이용률 및 사고율  
Fig. 2. Age Group-Based Electric Scooter Usage and Accident Rates

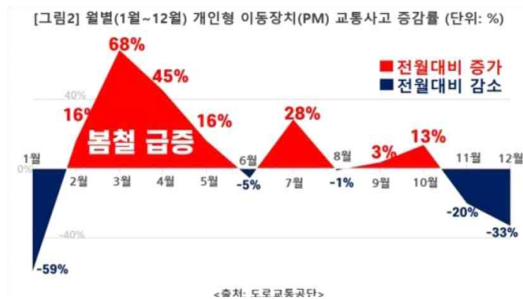


그림 3 월별 전동킥보드 사고 증감률  
Fig. 3. Monthly Changes in Electric Scooter Accident Rates

## II. 시스템 설계

### 2.1 전동킥보드 안전 시스템

전동킥보드 안전 시스템의 전체 시스템 구성도는 그림 4와 같다. 카메라(APC850 FHD), 프로세서(Raspberry Pi 4B), 디스플레이로 구성되었다. 이 시스템은 전방 및 후방 카메라를 활용하여 도로 위의 장애물과 차량을 인식하고, Raspberry Pi에 신호를 보낸다. 그 후 사용자가 알 수 있도록 손잡이 부분에 위치한 디스플레이, 진동, 스피커를 통해 위험신호를 전달한다.

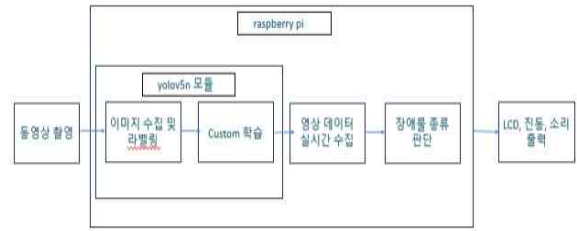


그림 4 전체 시스템 구성도  
Fig. 4. Overall System Architecture

이를 통해 운전자는 차도에서 킥보드를 주행할 때 수시로 뒤를 확인하며 경계할 필요 없이 장치에서 주는 신호만으로 후방차량의 접근 상태를 확인할 수 있으며, 야간 주행 시 전방의 포트홀, 맨홀, 방지턱과 같이 단차가 있어 사고를 유발할 수 있는 장애물을 감지하여 안전한 주행을 돕는다. 전체 시스템 동작 흐름도는 그림 5와 같다.

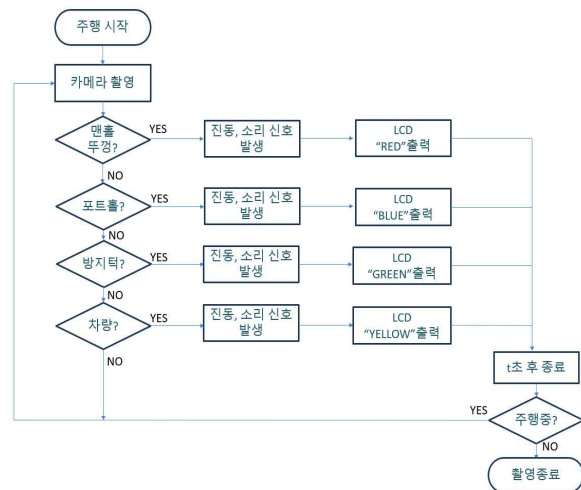


그림 5 시스템 동작 흐름도  
Fig. 5. System Operation Flowchart

### 2.2 웹캠 객체인식

포트홀, 맨홀, 방지턱, 차량 객체인식을 위해 Yolov5n 모델을 이용하여, 각각 약 3000개씩 총 11341장의 이미지로 custom 학습을 진행하였다. 동영상 frame단위로 cut하거나 roboflow 사이트를 참고하여 이미지를 수집하였고 labeling를 이용하여 라벨링 작업을 하였다.

그림 6은 딥러닝 후 얻은 Precision-recall curve 그래프다 [2]. x축은 재현율(Precision), y축은 정밀도(Recall)를 나타낸다. 오른쪽 상단에 적혀있는



mAP(Mean Average Precision)은 모델의 성능을 종합적으로 나타내며 1에 가까울수록 높은 성능을 보인다. 위 그래프는 포트홀 : 0.923, 맨홀 : 0.798, 방지턱 : 0.929, 차량 : 0.992로 정밀도와 재현율이 뛰어난 것을 확인할 수 있다. 마지막으로 그래프의 굵은 파란색 선은 모든 클래스를 종합한 결과로 대부분의 구간에서 높은 성능을 나타내는 것을 볼 수 있다.

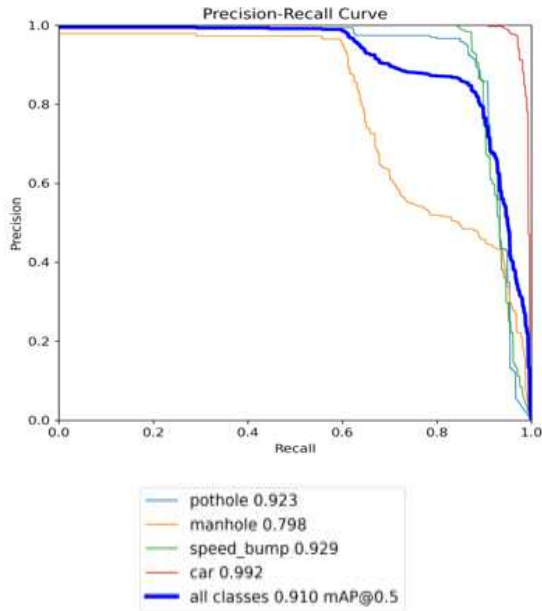


그림 6 신뢰도 재현율 그래프  
Fig. 6. Precision-recall curve graph

### 2.3 ELB060302를 이용한 PWM 제어

ELB060302 사운드 모듈과 스피커 앰프를 활용한 PWM 제어신호를 통해 소리를 출력하였다 [3]. 주파수와 Duty cycle을 제어하여 음향 신호를 발생시키는 방식으로 사용하였다. PWM 신호는 Raspberry Pi의 GPIO 핀을 사용하여 생성하며, 주파수와 듀티 사이클을 조정하여 다양한 소리를 출력할 수 있다. 본 장치에서는 1kHz 주파수와 50% 듀티 사이클로 PWM을 설정하여 일정한 경고음을 발생시켰다.

### 2.4 HC-SR04

HC-SR04는 TRIG와 ECHO 핀을 통해 작동한다. TRIG 핀에 짧은 펄스(10μs)를 보내면 초음파가 방출되고, 물체에 반사된 초음파가 센서로 돌아오면 ECHO 핀을 통해 그 시간을 기록합니다. 이 시간을

이용하여 거리 D를 구할 수 있다.

$$D = \frac{\text{시간} \times 343}{2} \quad (1)$$

수식 1에서 343m/s는 공기 중에서의 초음파 속도이며, 시간은 TRIG와 ECHO 사이의 신호 전달 시간이다. 또한 이 장치의 최대 측정거리는 5m인데, 본 실험에서는 1m 정도로 설정하였다.

### III. 실험 및 결과

최종 구현한 전동킥보드 안전 시스템은 그림 7, 8과 같다. 이 시스템은 딥러닝으로 학습된 가중치 파일을 사용하여 웹캠을 통한 실시간 객체 인식을 수행하며, 인식된 객체에 따라 주행자가 경고를 인지할 수 있도록 알람을 제공한다.



그림 7 포트홀, 맨홀, 방지턱에 대한 작동결과  
Fig. 7. Operation results for pothole, manhole, and speed bumps



그림 8 차량과 거리에 따른 신호 출력  
Fig. 8. Signal output by vehicle and distance

### III. 결 론

본 논문에서는 라즈베리파이 환경에서 YOLOv5n으로 학습된 가중치 파일을 사용하여 포트홀, 맨홀, 방지턱, 차량을 실시간으로 인식하고, 인식된 객체에 따라 주행자가 알 수 있도록 경보를 출력하는 시스템을 구현하였다. 이 시스템은 후방을 주시할 때 전방을 볼 수 없는 한계와 야간 주행 시 불편함을 해소하기 위해 고안된 것으로, 전동 킥보드뿐만 아니라 도로 옆을 달리는 자전거에도 적용이 가능할 것으로 보인다. 이를 통해 전동 킥보드 사고율을 줄이고, 운전자에게도 편안한 주행을 제공할 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] 이승현, 임영훈, 이승철, 이지훈, 오성현, “YOLOv5 기반 공유형 전동킥보드 법규위반 검출기법”, 대한전자공학회 학술대회, 제주, 2024년 6월.
- [2] 박창률, 이강민, 노형준, 김유강, “YOLO와 OpenCV 알고리즘을 이용한 안전 규제 위반 공유 전동킥보드 이용자 단속”, Proceedings of KIIT Conference, 제주, 2023년 11월.
- [3] 김윤태, 현규빈, 박찬정, “YOLOv5를 활용한 전동킥보드 탑승자 자동 단속 시스템 개발”, Proceedings of KIIT Conference, 제주, 2023년 11월.

# TPU소재를 이용한 인간-로봇 상호작용을 위한 로봇 손 설계

이민호\*, 김승환\*\*, 최영림\*\*\*, 김종욱\*\*\*\*

## Robotic hand design for human-robot interaction using TPU material

Minhoo Lee\*, Seunghwan Kim\*\* and Younglim Choi\*\*\*, Jongwook Kim\*\*\*\*

### 요 약

본 연구는 TPU(Thermoplastic Polyurethane) 소재의 유연성과 내구성을 활용하여 인간-로봇 상호작용(Human-Robot Interaction, HRI)에 적합한 로봇 손을 설계하고자 한다. TPU는 높은 유연성과 충격 흡수력을 갖춰, 로봇 손이 사람과 물체를 다룰 때 부드러운 촉각 피드백을 제공하여 안전성을 강화할 수 있다. 로봇 손과 팔의 출력 및 구동 시스템은 Arduino Mega 기반의 전력 제어 회로로 설계되었으며, 전원 관리 시스템과 통신 인터페이스를 통해 안정성을 높였다. 또한, 모터의 크기와 호환성 문제를 해결하고 부품의 배치를 조정함으로써 시스템의 내구성과 성능을 개선하였다. 이 설계는 인간과의 상호작용에서 안전하고 효율적인 로봇 손을 구현하는 데 중점을 두고 있다.

### Abstract

This study presents the design of a robotic hand optimized for Human-Robot Interaction (HRI) through the application of Thermoplastic Polyurethane (TPU), a material chosen for its exceptional flexibility and durability. TPU's high elasticity and impact absorption capabilities provide soft tactile feedback during interactions with humans and objects, thereby enhancing operational safety. The power output and actuation system of the robotic hand and arm are managed by an Arduino Mega-based power control circuit, integrating a power management system and communication interfaces to ensure system stability. Additionally, adjustments to motor compatibility and component layout were implemented to improve the durability and performance of the system. This design prioritizes the creation of a robotic hand that achieves safe and efficient interaction within human-centered environments.

### Key words

Thermoplastic Polyurethane(TPU), Human-Robot Interaction (HRI), Soft Robotic Hand, Tactile Feedback

---

\* 동아대학교 전자공학과, ku2634@naver.com, \*\*동아대학교 전자공학과, kshh7487@naver.com

\*\*\* 동아대학교 SW혁신센터, cyl88@dau.ac.kr(공동 교신저자),

\*\*\*\* 동아대학교 전자공학과, Email: kjwook@dau.ac.kr(공동 교신저자).

※ 본 연구성과물은 2024년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 (No. 과제 관리번호 : RS-2023-00242673 )

## I. 서 론

본 연구에서는 TPU 소재의 유연성과 내구성을 활용하여 인간-로봇 상호작용에 적합한 로봇 손을 설계하고자 한다. TPU는 기존의 로봇 손 설계에 사용되는 소재보다 높은 유연성과 충격 흡수력을 가지므로 안전하고 효과적인 상호작용이 가능하다[1].

로봇 손 설계에서 인간 손의 특성을 모방하려는 시도는 인간과의 상호작용을 자연스럽게 할 수 있는 로봇을 만드는 데 중요한 역할을 한다[2][3]. 그러나 기존의 철로 된 로봇 손 설계는 높은 정밀도를 제공하는 동시에 무겁고 충격에 민감하여 민첩성이나 안전성 측면에서 한계가 있다[4]. 이에 따라 유연성과 내구성을 겸비한 소프트 소재가 로봇 손 설계에 도입되고 있으며, TPU (Thermoplastic Polyurethane) 소재는 이러한 요구사항을 충족할 수 있는 소재로 주목받고 있다[5]. TPU의 또 다른 중요한 장점은 촉각 면에서의 뛰어난 특성이다. TPU는 실리콘이나 플라스틱과 비교했을 때 상대적으로 더 부드럽고 탄성이 좋아, 물체나 사람을 직접적으로 접촉할 때 더 자연스럽게 유사한 촉각 피드백을 준다[6]. 이러한 특성은 로봇 손이 물체를 섬세하게 다룰 때뿐 아니라 인간과의 안전한 상호작용을 위해 더욱 유리하게 작용한다. 본 연구에서는 TPU 소재를 활용한 로봇 손 설계를 통해 인간-로봇 상호작용에서의 안전성과 효율성을 높이는 방안을 모색하고자 한다. TPU 기반의 로봇 손은 휴머노이드 로봇 '엘리스(Alice)'[7]에 적용될 예정으로, 인간 손과 유사한 기능을 목표로 하고 있다.

2장에서는 기존 로봇손 설계와 비교 분석을 진행하고, 3장에서는 설계된 로봇 손을 설명한다.

## II. 배경 및 관련 연구

로봇 손 설계에서 소프트 소재와 철 혹은 알루미늄 등의 하드 소재의 비교 연구를 표 1에 정리하였다. 각 소재는 다양한 장단점을 지니고 있으며, TPU와 같은 소프트 소재가 HRI 관점에서 로봇 손 설계에 효과적일 수 있다.

표 1. 관련연구

Table 1. Related research

제목	요약
A Modular Bio-inspired Robotic Hand with High Sensitivity[8]	케이블 구동 방식과 생체모방 접근을 결합한 모듈형 로봇 손 설계 부드러운 실리콘과 견고한 뼈대를 결합해 충격 흡수와 안정적인 관절 위치 유지를 통해 인간 손의 조작 특성을 재현하는 데 중점
A practical 3D-printed soft robotic prosthetic hand with multi-articulating capabilities[9]	다중 관절을 구현한 3D 프린팅 소프트 로봇 의수를 개발하여 사용자 요구에 맞는 기능을 제공하려는 연구
Development and Grasp Stability Estimation of Sensorized Soft Robotic Hand[10]	그립 안정성을 평가할 수 있는 소프트 로봇 손을 개발하고, 다양한 객체에 대한 그립 상태를 머신러닝 기법을 통해 예측하는 방식

TPU의 또 다른 중요한 장점은 촉각 면에서의 뛰어난 특성이다. TPU는 실리콘이나 플라스틱과 비교했을 때 상대적으로 더 부드럽고 탄성이 좋아, 물체나 사람을 직접적으로 접촉할 때 더 자연스럽게 유사한 촉각 피드백을 제공한다[11].

## III. TPU 기반 로봇 손 설계 및 구현

### 1. 로봇 시스템의 전력 제어 회로로 설계

본 연구에서는 로봇 손 제어를 위해 전원 관리 시스템 및 통신 인터페이스를 포함한 로봇 시스템의 전력 제어 회로로 설계하였다. 시스템은 주요 제어 장치로 Arduino Mega (ATmega2560)를 사용하며, 서보 모터와 통신 모듈이 이를 중심으로 연결된다. 각 구성 요소와의 연결 방식은 그림 1과 같다.

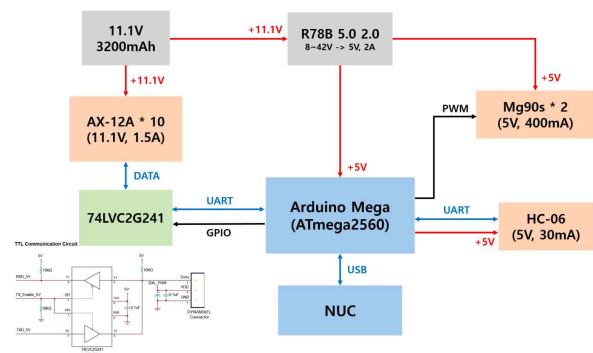


그림 1. 로봇 시스템의 전력 제어 회로로 설계  
Fig. 1. Designed as a power control circuit for robotic systems

전체 시스템의 전력은 11.1V 3200mAh 배터리에서 공급되며, R78B 5.0 2.0 전압 변환기를 통해 5V 전압으로 변환되어 저전압 장치에 안정적으로 제공된다. AX-12A 서보 모터는 배터리의 11.1V 전원을 직접 공급받아 고출력 동작을 수행하며, 74LVC2G241 TTL 통신 회로를 통해 데이터 신호를 전달받아 제어된다. 이 회로는 half duplex 방식의 AX-12A 모터와 Arduino Mega 간 안정적인 신호 전송을 위해 full duplex 통신 환경을 지원하는 역할을 한다. 또한, MG90s 서보 모터는 5V, 400mA 전력을 사용하고, PWM 신호를 통해 Arduino Mega로부터 직접 제어된다. 시스템 통신 인터페이스는 UART 프로토콜을 사용하여 실시간 데이터 전송 및 수신 가능하다. 또한, NUC와의 통신을 통해 각 모터의 관절을 제어할 수 있으며, HC-06 블루투스 모듈을 통해 무선으로도 제어할 수 있다.

## 2. 로봇 팔 구조 변경 모델링

기존 모델링[9]의 내부 구조는 본 연구에서 사용하고자 하는 모터의 크기 호환성 문제와 구동 범위에 따른 위치 문제 등으로 인해 안정성에서 한계를 보였다. 부품 배치와 구조물이 모터 간의 간섭 현상 등으로 실제 운영 시 성능이 저하될 가능성이 있었다. 이를 해결하기 위해 그림 2와 같이 설계에 몇 가지 주요 변경을 적용하였다.

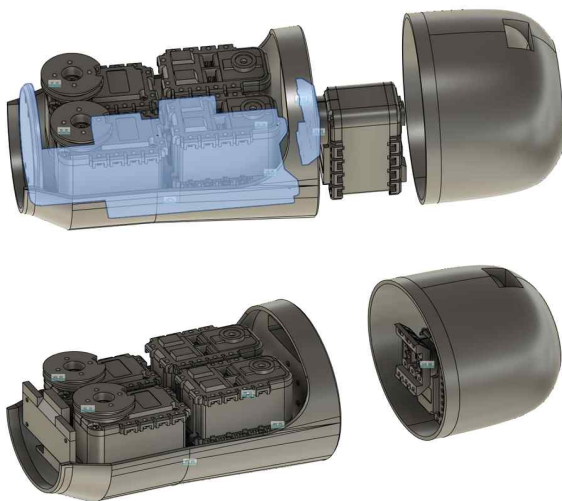


그림 2. 로봇 손 모델링  
Fig. 2. Robotic hand modeling

모터 크기에 맞춘 구조 재설계를 통해 모터와 내부 부품 간의 공간을 효율적으로 재배치하여, 부품이 안정적으로 고정될 수 있도록 내부 공간을 최대한 활용하였다. 고정 구조 강화를 통해 모터와 인접한 부품들의 진동과 충격에 대한 내구성을 높였다. 고정 구조물을 추가하고 보강함으로써 반복적인 모터 작동에도 부품 손상을 방지하도록 설계하였다. 또 무게 중심 조절을 통해 전체 시스템의 안정성을 높였다. 부품 배치를 조정하여 모터 구동 시 발생하는 간섭을 최소화하였다. 이러한 설계 변경은 모터 크기와 간섭 문제를 해결하여 시스템의 내구성과 성능을 향상시켰다.

## 3. 로봇 손과 팔 출력

로봇 손을 TPU(Thermoplastic Polyurethane)로 제작하여 인간-로봇 상호작용(HRI)에 적합한 특성을 구현하였다. TPU는 높은 유연성과 충격 흡수력을 제공하여, 인간과의 접촉 시 부드러운 촉각 피드백을 가능하게 한다. 로봇 팔의 경우 PLA를 이용하여, 손과 연결된 구동 구조는 높은 출력을 유지하면서도 가벼운 무게와 내구성을 확보하도록 설계되었다. 이 구조는 손가락의 정밀한 동작 제어를 지원하여, 인간과의 상호작용 환경에서 효율적이고 안전하게 사용할 수 있다.



그림 3. 구조 변경 된 로봇 손 출력  
Fig. 3. Structure-changed robot hand output

## III. 결 론

본 연구에서는 TPU 소재의 특성을 활용하여 인간-로봇 상호작용에 적합한 로봇 손을 설계하고 구현하였다. TPU는 기존 하드 소재와 비교해 더 높은 유연성과 충격 흡수력을 제공하여, 로봇 손이 사람

과 물체를 다룰 때 안전하고 자연스러운 촉각 피드백을 가능하게 한다. 이를 통해 사용자와의 상호작용에서 발생할 수 있는 위험성을 줄이고, 보다 부드럽고 안정적인 조작이 가능하도록 설계 하였다.

로봇 손 시스템 설계에서는 Arduino Mega를 중심으로 전원 관리 시스템 및 통신 인터페이스를 포함한 전력 제어 회로를 구성하여 안정성을 높였다. 또한, 모터 크기와 호환성을 고려한 모델링 설계를 통해 내부 구조를 최적화하고 부품의 배치를 조절하여 모터 구동 시 발생하는 진동과 간섭을 최소화 하였다. 이러한 설계 개선을 통해 로봇 손의 내구성과 성능을 향상시켰으며, 인간과 상호작용하는 다양한 환경에서의 적용 가능성을 입증하였다.

향후 연구에서는 TPU 소재의 촉각 피드백을 이용하여 사용자 연구를 진행하고자 한다.

## 참 고 문 헌

- [1] "TPU Material," HN New Material, [Online]. Available: <https://hnnewmaterial.com/ko/tpu-%EC%86%8C%EC%9E%AC/>. [Accessed: Nov. 2, 2024].
- [2] C.-G. Kim, E.-S. Kim, J.-W. Shin, and B.-Y. Park, "Real-time robotic arm control embedded system using hand gestures," J. Korea Inform. Technol. Soc., vol. 19, no. 11, pp. 45-53, 2021.
- [3] S.-S. Jang and T.-Y. Byun, "Design and implementation of a personal communicator based on embedded technology for remote device control," J. Korea Ind. Inform. Soc., vol. 16, no. 2, pp. 99-109, 2011.
- [4] "ETRI develops new tactile sensor technology for robot fingers," Yonhap News Agency, [Online]. Available: <https://www.yna.co.kr/view/AKR20240326098100063>. [Accessed: Nov. 2, 2024].
- [5] "Thermoplastic Polyurethane (TPU) Material Properties," Material Properties, [Online]. Available: <https://material-properties.org/ko/%ED%8F%B4%EB%A6%AC%EC%9A%B0%EB%A0%88%ED%83%84-%ED%85%8C%EB%A5%B4%EB%AA%A8%ED%94%8C%EB%9D%BC%EC%8A%A4%ED%8B%B1/>. [Accessed: Nov. 2, 2024].
- [6] Y. Choi, Y. Park, S. Lee, and J. Kim, "Design of a rehabilitation device for hand rehabilitation using TPU material," Proc. KIIT Conf., vol. 2023, no. 6, pp. 322-324, 2023.
- [7] "Robotic hand design and its application for rehabilitation," Arobot4all, [Online]. Available: <https://arobot4all.com/43/?bmode=view&idx=8668869>. [Accessed: Nov. 2, 2024].
- [8] C. Liu, A. Moncada, H. Matusik, D. I. Erus, and D. Rus, "A Modular Bio-inspired Robotic Hand with High Sensitivity," IEEE, April 2023.
- [9] A. Mohammadi, J. Lavranos, H. Zhou, R. Mutlu, G. Alici, Y. Tan, P. Choong, and D. Oetomo, "A practical 3D-printed soft robotic prosthetic hand with multi-articulating capabilities," PLOS ONE, vol. 15, no. 5, pp. e0232766, May 2020.
- [10] P. M. Khin, J. H. Low, M. H. Ang Jr., and C. H. Yeow, "Development and Grasp Stability Estimation of Sensorized Soft Robotic Hand," Frontiers in Robotics and AI, vol. 8, pp. 619-390, 2021.
- [11] "The Properties and Applications of TPU Plastic," Team Manufacturing, [Online]. Available: <https://www.team-mfg.com/ko/blog/tpu-plastic.html>. [Accessed: Nov. 2, 2024].

# 눈과 입모양 인식을 이용한 의사소통 보조 시스템

이서희\*, 이아현\*\*, 김원호\*\*\*

## A Communication Assistance System using Eye and Mouth Shape Recognition

Seo-hee Lee\*, Ah-hyeon Lee\*\*, Won-ho Kim\*\*\*

### 요 약

본 논문에서는 중증 환자들이 눈 깜빡임과 입 벌림 동작을 통해 의사소통할 수 있는 보조 시스템을 제안한다. 시스템의 기능은 사용자의 눈과 입모양의 움직임인식을 인식하여 사전 설정된 메시지를 LCD에 출력하며, 하드웨어 구성은 카메라, 프로세서, LCD 디스플레이로 설계하였다. 눈과 입모양의 중횡비를 기반으로 눈 깜빡임과 입 벌림 모양을 인식 처리하였다. 제작된 의사소통 보조기를 시험한 결과, 정면에서 1m 거리 내 촬영 시 평균 89%의 인식률을 보였으며, 인식된 눈과 입의 모양에 따라 메시지가 대부분 일치되게 출력함을 확인하였다.

### Abstract

In this paper, we propose an assistance system that allows severely ill patients to communicate through blink and mouth opening movements. The function of the system is to recognize the movement of the user's eyes and mouth and output a preset message to the LCD, and the hardware configuration is designed with a camera, processor, and LCD display. Based on the aspect ratio of the eye and mouth shape, the blink and mouth opening shapes were recognized and processed. As a result of testing the manufactured communication aid, it showed an average recognition rate of 89% when shooting within a distance of 1m from the front, and It was confirmed that most of the messages were outputted consistently according to the shape of the recognized eyes and mouth.

### Key words

Image processing, Face recognition, Assistive device, Message selection system

### 1. 서 론

중증 질환으로 인해 신체 움직임이 제한되거나 언어 구사가 어려운 환자들은 주변 사람들과의 의

사소통이 거의 불가능한 경우가 많다 [1]. 이러한 환자들에게 의사소통은 단순한 정보 전달을 넘어, 인간으로서의 존엄을 유지하고 사회 참여를 가능하게 하는 중요한 수단이다. 최근 다양한 보조기술이

\* 공주대학교 전기전자제어공학부, seohee183@gmail.com

\*\* 공주대학교 전기전자제어공학부, dkgus5446@naver.com

\*\*\* 지도교수, 공주대학교 전기전자제어공학부, whkim@kongju.ac.kr

개발되면서 환자들의 의사소통을 돕는 도구들이 등장하고 있지만, 대부분의 기기들은 고가이거나 복잡한 사용법으로 인해 실제 적용에 어려움이 있다 [2]. 특히, 매우 제한적인 신체 움직임만이 가능한 중증 환자들을 위한 보조 장치는 여전히 부족한 실정이다. 이를 개선하기 위하여 본 논문에서는 간단한 얼굴 움직임, 즉 눈 깜빡임과 입 벌림을 인식하여 의사소통을 도울 수 있는 시스템을 제안하고자 한다. 제안 시스템은 라즈베리파이 프로세서 보드와 웹캠을 활용하여 사용자의 눈과 입 움직임을 인식하고, 이를 통해 사전에 설정된 메시지를 선택해 사용자가 원하는 의사를 표현하거나 전달할 수 있는 기능을 제공한다.

## II. 시스템 설계

### 2.1 의사소통 보조기 시스템

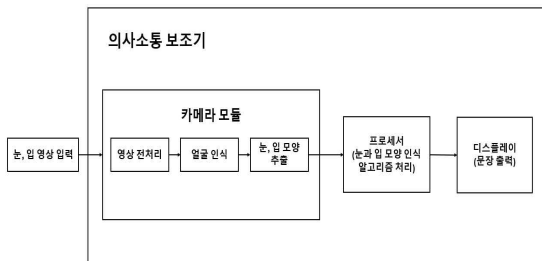


그림 1. 시스템 하드웨어 기능 구성도

Fig 1. Functional block diagram of system hardware

의사소통 보조기의 하드웨어 구성도는 그림 1과 같다. 의사소통 보조기는 카메라(APC850 FHD), 프로세서(Raspberry Pi 4B), 디스플레이(RPi LCD)로 구성되었다. 전체 시스템의 동작 흐름도는 그림 2와 같다.

본 시스템은 웹캠을 통해 눈과 입의 영상을 획득하고, 프로세서 내에서 눈과 입의 랜드마크를 추출해 눈 깜빡임과 입 벌림을 인식한다. 이후, LCD 디스플레이를 통해 눈 깜빡임과 입 벌림으로 사전 설정한 신호 체계(짧은 부호(.) 긴 부호(-))에 따라 특정 메시지를 선택할 수 있다. 또한, 사용자가 잘못된 입력을 했을 때 취소되는 기능을 추가하여 이를 신속하게 수정할 수 있도록 설계했다. 선택된 메시

지는 LCD 디스플레이를 통해 시각적으로 출력된다. 이를 통해 사용자는 눈 깜빡임과 입 벌림이라는 간단한 신체 동작만으로도 원하는 메시지를 정확하게 전달할 수 있다. 하드웨어 구성은 간단하면서도 손쉽게 의사소통이 어려운 중증 장애인에게 적용할 수 있으며, 복잡한 설정 없이 사용자의 의사소통을 지원할 수 있다.

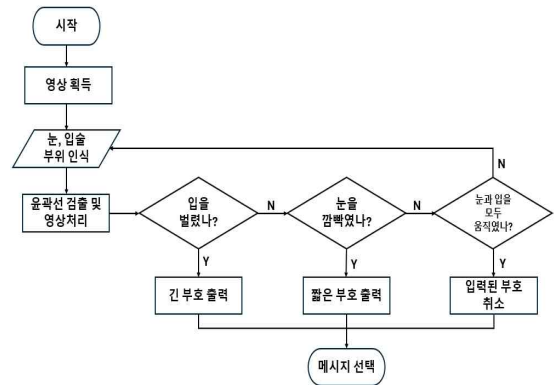


그림 2. 시스템 동작 흐름도

Fig 2. System Operation Flowchart

### 2.2 눈과 입 모양 인식 알고리즘

눈과 입의 움직임을 탐지하기 위해 dlib의 face detector와 shape predictor를 활용한다 [3]. NumPy를 통해 탐지된 좌표 간 거리를 계산하여 눈 깜빡임과 입 벌림 동작을 인식한다. 이 과정에서 눈 종횡비 (EAR, Eye Aspect Ratio), 입 종횡비 (MAR, Mouth Aspect Ratio) 계산 알고리즘을 이용하여 특정 좌표 간 거리를 계산한다 [4].

EAR 계산 수식은 식(1)과 같다. 그림 3과 같이 왼쪽과 오른쪽 눈의 수평 및 수직 거리를 각각 계산하여 EAR 값을 도출한다. MAR 계산 수식은 식 (2)와 같다. 그림 4는 MAR을 계산하기 위한 입모양 파라메타이다.

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2 \|p_1 - p_4\|} \quad (1)$$

$$MAR = \frac{\|p_3 - p_{11}\| + \|p_5 - p_9\|}{2 \|p_1 - p_7\|} \quad (2)$$

눈 깜빡임 동작은 짧은 부호(.), 입 벌림 동작은 긴 부호(-)로 변환되어 출력된다. 중증 환자를 위한



의사소통 보조기이므로 짧은 부호(.)는 기분에 관한 메시지를, 긴 부호(-)는 상태에 대한 메시지를 출력하여 사용자가 메시지를 쉽게 기억하고 사용할 수 있도록 하였다. 또한, 입력이 잘못된 경우 눈과 입을 순서 상관없이 모두 한 번 이상 움직이는 동작으로 입력을 취소할 수 있도록 하였다.

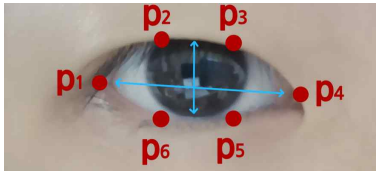


그림 3. EAR 계산 파라메타  
Fig 3. EAR calculation parameter

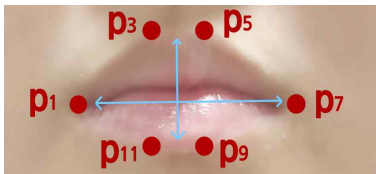


그림 4. MAR 계산 파라메타  
Fig 4. MAR calculation parameter

### 2.3 실험 및 결과

최종 구현된 의사소통 보조 시스템 사진은 그림 5와 같다. 웹캠을 통해 얼굴 영상을 촬영하고, LCD에 메시지를 출력하도록 제작하였다. 라즈베리파이 프로세서 보드를 내장하여 영상처리 알고리즘이 실행되며, LCD에는 웹캠 영상, 변환된 부호, 선택된 메시지가 표시된다.



그림 5. 제작된 시스템 사진  
Fig 5. Implemented system photo

시스템 작동 결과, 눈 깜빡임과 입 벌림 동작이 원활하게 인식되었으며, 입력된 부호에 따라 메시지

가 문제없이 출력됨을 확인하였다. 그림 6은 ‘---’ 입력을 통해 ‘Uncomfortable’이라는 문장을 출력한 모습이다. 웹캠으로부터 0.5m, 1m, 1.5m 떨어진 거리에서 얼굴의 정면, 우측, 좌측을 각각 100번씩 촬영하여 실험한 결과, 표 1과 같은 인식률이 도출되었다. 정면을 1m 이내에서 촬영할 때는 높은 인식률을 보였으나, 1m를 초과하거나 측면에서 촬영할 경우 인식률이 저하되는 경향을 보였다.

표 1. 실험 데이터 세트

Table 1. Experimental data sets

	0.5m	1m	1.5m
정면	90%	88%	52%
우측	48%	50%	48%
좌측	52%	48%	46%

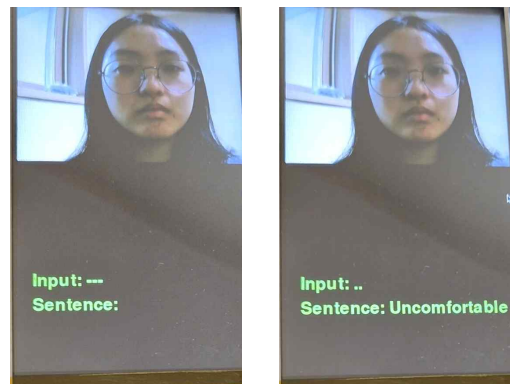


그림 6. 시스템 작동 결과

Fig 6. System Operation Results

### III. 결 론

본 논문에서는 라즈베리파이 환경에서 dlib를 기반으로 눈 깜빡임, 입 벌림 동작을 인식하여 의사소통을 돕는 시스템을 제안하고 구현하였다. 동작 결과, 눈과 입의 움직임이 정확하게 인식되었으며, 이에 따라 메시지가 정상적으로 출력되었다. 이 시스템은 의사소통이 어려운 중증 환자들이 자신의 의사를 효과적으로 표현할 수 있게 하며, 의료진 및 보호자와의 의사소통을 원활하게 하고 환자들에게 심리적 안정을 제공할 것으로 기대된다.

### 참 고 문 헌

- [1] 박은영, “ALS 환자의 의사소통 어려움 및 지원 요구에 관한 질적 연구”, 석사학위논문, 이화여자대학교, 2015.
- [2] 전은서, 김채연, 신애리, 전영은, 원동욱, “근위 축성 측색 경화증 환자를 위한 눈 움직임 검출 기반 모바일 플랫폼 의사소통 시스템”, 한국정보과학회 학술발표논문집, pp. 1701-1703, 2021.
- [3] 정영욱, 이영재, “dlib 라이브러리를 이용한 얼굴 인식 기반 출퇴근시스템”, 정보기술학회 하계학술대회 논문집, pp. 707-712, 2021.
- [4] Youwei Lu, “Real-time eye blink detection using general cameras: a facial landmarks approach”, International Science Journal of Engineering & Agriculture, vol.2, No.5, 2023.

## 캡스톤 프로젝트 알카이브에 관하여

제갈건<sup>3)</sup>, 김병현<sup>4)</sup>, 이상원<sup>5)</sup>, 조준현<sup>6)</sup>, 최준혁<sup>7)</sup>, 한유준<sup>8)</sup>

### About The Capston Project ALChive

JegalGeon, KimByeongHyeon, LeeSangWon, JoJunHyeon, ChoeJunHyeok, and HanYuJun

#### 요 약

기존의 주류관련 어플리케이션들은 정보만을 제공하거나 안드로이드, 아이폰 환경 중 하나에만 서비스를 제공하는 등 여러 부족한 부분이 존재했다. 또, 대표적인 주류만을 소개한다는 특징이 존재했는데, 이를 보완하여 전통주, 무알콜 주류 등 상대적으로 알려지지 않은 주류들을 포함해 사용자가 다양한 주류를 보다 쉽게 접할 수 있도록 하는 ALChive 프로젝트를 진행하게 되었다. ALChive는 Archive와 Alcohol의 합성어로 다양한 주류의 정보를 저장한다는 의미를 가진다. 해당 논문에서는 어떠한 배경으로 해당 프로젝트를 진행하게 되었는지와 해당 프로젝트를 통해 얻을 수 있는 이점, 그리고 해당 프로젝트의 구조에 대하여 다룬다.

#### Abstract

Existing liquor-related applications had several deficiencies, such as providing only information or providing services only to one of Android and iPhone environments. In addition, there was a feature that only representative alcoholic beverages were introduced, and by supplementing this, the ALChive project, which made it easier for users to access a variety of alcoholic beverages, including traditional alcoholic beverages and non-alcoholic alcoholic beverages, was carried out. This paper deals with the background of the project, the benefits that can be gained from it, and the structure of the project.

#### Key words

Alchive project, Traditional liquor&Non-alcoholic&Custom Alcohol recomemdaton service, AI-based recommendation system, Cloud computing

## 1. 서 론

최근 주류 시장은 전통주와 무알콜 주류와 같은 다양한 주류군에 대한 수요가 증가하고 있으며, 이에 대한 정보를 제공하는 서비스가 필요해지고 있다. 농림축산식품부의 발표에 따르면 2015년 409억 원 규모였던 전통주 시장은 2022년 약 1,629억 원으로 급격히 성장하였으며, 이는 다양한 주류 정보를 제공하는 서비스의 필요성과 경쟁력이 증가하고 있다는 것을 나타낸다. 본 프로젝트는 이러한 배경에

3) 계명대학교, man25816@naver.com

4) 계명대학교, kbo102142@gmail.com

5) 계명대학교, swl0511@naver.com

6) 계명대학교, wpqlks7@naver.com

7) 계명대학교, chlwnsgur0716@gmail.com

8) 계명대학교, a01049777206@gmail.com

서 전통주와 무알콜 주류의 접근성을 높이고자 하는 목표로 시작되었다.

표 1. 전통주 시장 규모 및 수출액

Table 1. The size and export value of the traditional liquor market

년도	시장 규모(단위 원)	수출액(단위:달러)
2015	409억	1,727만
2017	400억	1,766만
2019	531억	1,736만
2020	627억	1,822만
2021	941억	2,352만
2022	1,629억	2,466만

## II. 본 론

본 프로젝트에서는 구글 렌즈와 커스텀 검색 API 르르 이용하여 사용자가 촬영한 술의 정보를 효율적으로 검색하고 제공하는 어플리케이션을 제안한다. 이 어플리케이션은 술의 텍스트 추출, 유사 이미지 검색, 개인화 된 추천 시스템, 술 제조 과정 및 원산지 정보 제공, 그리고 실시간 시세 정보 기능을 통합하여 사용자 경험을 극대화한다. 술의 종류와 정보는 방대하나, 사용자들이 원하는 정보를 신속하게 찾는 것은 어려운 문제이다. 본 연구의 목적은 구글 렌즈와 커스텀 검색 API를 통해 사용자에게 쉽고 편리하게 술의 정보를 제공하는 어플리케이션을 개발하는 것이다.

### 가) ALChive의 주요 기능

- 1) 술 이미지 촬영 및 텍스트 추출 : 사용자의 술의 이미지를 촬영하면, 텍스트 정보를 자동으로 제공한다.
- 2) 술 제조 과정 및 원산지 정보 : 각 술의 제조 과정과 원산지 및 술에 대한 상세 정보를 제공한다.
- 3) 실시간 시세 기능 : 술의 현재 시세 정보를 제공하고 가격 변동 알림 기능을 제공한다.
- 4) 주류 인식 및 정보 제공 시스템 : 카메라를 통해 술의 병이나 라벨을 촬영하면 AI가 텍스트, 로고, 이미지를 분석하여 정보를 제공한다.

5) AI 기반 개인별 칵테일 추천 : 사용자가 보유한 재료와 선호도에 따라 맞춤형 칵테일 추천한다.

6) 전통주 레시피 제공 : 양주 뿐만 아니라 전통주를 기반으로 한 새로운 칵테일 레시피를 제공하여 차별화 된 서비스를 제공한다.

### 나) 경쟁 서비스와 차별성

기존 서비스는 주로 칵테일 정보 제공에 그치는 반면, ALChive는 AI를 통해 사용자가 보유한 재료를 기반으로 레시피도 추가로 제공한다. 한국의 전통 술문화와 현대적인 칵테일 문화를 융합하는 점에서 독창적이다.

### 다) 시장성 및 사업화 가능성

전통주 시장은 꾸준히 성장하고 있으며, 20-40대 고객층의 수요가 증가하고 있다. ALChive는 이러한 잠재 시장에 접근할 수 있는 효과적인 서비스로, 글로벌 확장성도 갖추고 있어 사업화 가능성이 높다.

### 라) 시스템 아키텍처

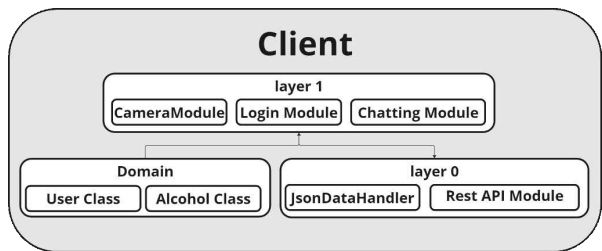


그림 1. 클라이언트 아키텍처

Fig. 1. Client Architecture

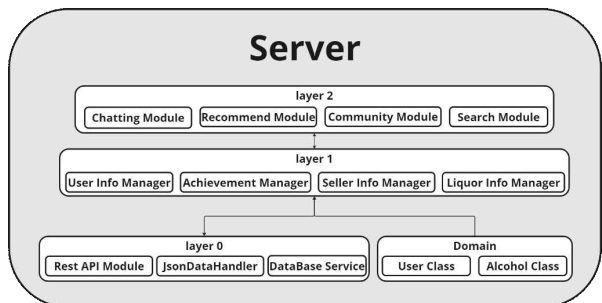


그림 2. 서버 아키텍처

Fig. 2. Server Architecture

## 참 고 문 헌

- [1] 농림축산식품부, "글로벌 '케이-푸드(K-Food)' 시장 확대를 위한 제 4차 식품산업 진흥 기본 계획 발표," 2020

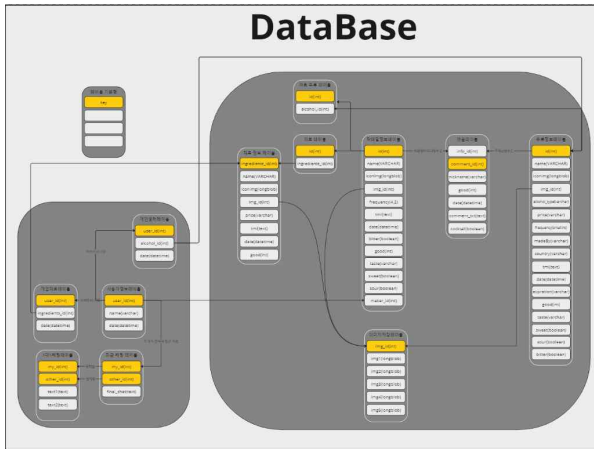


그림 3. 데이터베이스 아키텍처  
Fig. 3. DataBase Architecture

## III. 결 론

본 프로젝트에서는 다양한 주류에 대한 접근성을 높이고, 사용자 맞춤형 주류 경험을 제공하기 위해 ALChive 프로젝트를 설계하고 구현하였다. ALChive는 기존 주류 관련 어플리케이션의 한계를 보완하고, 구글 렌즈와 커스텀 검색 API를 활용하여 사용자가 직접 촬영한 주류 이미지를 기반으로 텍스트를 추출하고, 유사 주류를 검색하여 정보와 시세를 제공함으로써 사용자 경험을 극대화하였다. 더 나아가, AI 기반의 개인화된 주류 추천과 전통주를 포함한 맞춤형 칵테일 레시피 추천을 통해 기존 주류 서비스와 차별화된 가치를 제공하였다.

ALChive는 단순한 주류 정보 제공을 넘어 한국의 전통주와 현대의 주류 문화를 융합한 차별화된 사용자 경험을 창출함으로써 사용자들이 폭넓은 주류 경험을 할 수 있도록 돕는다. 특히, 전통주의 매력을 재발견할 수 있는 기회를 제공함으로써 사용자들이 다양한 주류를 경험할 수 있는 플랫폼으로 자리잡을 수 있을 것으로 기대된다. 이러한 ALChive의 기능은 한국 전통주 시장의 성장과 함께 주류 산업의 새로운 변화를 이끌어갈 잠재력을 보유하고 있으며, 향후 글로벌 시장으로의 확장을 통해 전 세계 소비자들에게 차별화된 서비스를 제공할 가능성도 갖추고 있다.

# 무선 통신 기반 수질 데이터 수집 및 시각화

강명준\*, 이상준, 양진모, 임종찬, 나연수, 윤승재, 김민서, 이현서

## Water Quality Data Collection and Visualization using Wireless Communication

Myeongjun Kang\*, Sangjun Lee, Jin-Mo Yang, Jong-Chan Im, Yeon-Soo Na, Seung-Jea Yun, Min-Seo Kim, Hyun-Seo Lee

### 요약

본 연구는 무선 통신 기술을 활용한 실시간 수질 데이터 수집 및 시각화 시스템을 제안한다. 제안된 시스템은 수질 데이터를 실시간으로 모니터링하고, 녹조 발생 가능성을 머신러닝 알고리즘으로 예측하여 사용자에게 경고 알림을 제공한다. 실험 결과, 제안된 시스템은 높은 데이터 정확도와 실시간 대응성을 보여 수질 오염 문제 해결에 효과적임을 확인하였다. 이 연구는 환경 보호와 공중 보건을 위한 실질적인 방안을 제시하며, 다양한 응용 분야에서 적용 가능성을 검토하였다.

### Abstract

This study proposes a real-time water quality data collection and visualization system using wireless communication technology. The proposed system monitors water quality data in real-time and predicts the likelihood of algal bloom occurrences using a machine learning algorithm. In addition, it provides users with alert notifications to enable prompt response. Experimental results show that the system is effective in addressing water pollution issues due to its high data accuracy and real-time responsiveness. This research offers a practical solution for environmental protection and public health improvement and explores the applicability of the system in various fields.

### Key words

Water Quality Monitoring, Wireless Communication, Algal Bloom Prediction, Real-time Data Collection, Machine Learning, Environmental Protection

## 1. 서론

### 1.1 연구 배경

수질 오염은 전 세계적으로 환경과 인간 건강에 심각한 영향을 미치는 주요 문제 중 하나이다. 특히, 녹조 발생은 수질 오염의 대표적 사례로, 물의

부영양화와 기온 상승 등의 복합적인 요인으로 인해 발생한다. 녹조는 수중 생태계에 독성 물질을 방출하여 생태계를 파괴하고, 수질을 악화시켜 인간과 동물의 건강에 직접적인 영향을 미친다.

기존의 수질 모니터링 방법은 주로 화학적 분석과 현장 수질 검사를 기반으로 하며, 이는 시간 소

\*국립공주대학교 (myeng07@naver.com)

모적이고 비용이 많이 드는 문제점이 있다. 또한, 고정된 위치에서만 데이터를 수집할 수 있어 실시간 모니터링과 신속한 대응에 한계가 있다. 따라서 실시간으로 수질 데이터를 모니터링하고 분석할 수 있는 효율적인 시스템이 필요하다.

### 1.2 연구 목적

본 논문은 무선 통신 기술을 기반으로 수질 데이터를 실시간으로 수집하고 시각화하는 시스템을 제안한다. 이 시스템은 녹조 발생 가능성을 예측하여 신속한 대응을 가능하게 하며, 환경 보호와 공중 보건을 개선하는 것을 목표로 한다.

## II. 연구의 필요성

### 2.1 녹조 문제의 심각성

녹조는 수중 생태계에 치명적인 영향을 미치며, 특히 남조류에 의한 녹조는 독성 물질을 방출하여 수질 오염을 악화시킨다. 국내 하천과 호수에서는 매년 남조류 경보가 발령되고 있으며, 이는 지역 사회와 환경에 심각한 부정적 영향을 미친다. 예를 들어, 낙동강에서 발견된 신경독소 BMAA는 알츠하이머와 루게릭병을 유발할 수 있는 물질로, 미국 환경보호청(EPA) 기준의 490배에 달하는 독성을 가지고 있다.



Fig. 1. 국내 낙동강 녹조 발생 상황 (24.08)

### 2.2 기존 기술의 한계

현재 수질 모니터링 방법은 비용이 많이 들고, 신속한 대응이 어렵다는 문제가 있다. 기존 대부분의 모니터링 시스템은 고정형 센서를 사용하여 특정 위치에서만 데이터를 수집할 수 있으며, 이는 넓은 수역에 대한 지속적이고 포괄적인 모니터링에 한계를 드러낸다. 이로 인해 실시간 데이터 수집 및

녹조 발생 예측이 어려워 신속한 대응이 불가능할 뿐만 아니라, 대응이 지연될 경우 수질 악화로 이어질 위험이 높다.

또한, 기존 모니터링 시스템은 현장 인력의 투입이 필수적인 경우가 많다. 예를 들어, 일정한 주기로 수질을 검사하거나, 녹조 문제를 해결하기 위해 현장 점검이 필요할 때 인력을 동원해야 한다. 이는 인력과 시간의 비용을 증가시킬 뿐만 아니라, 현장 작업 중 발생할 수 있는 안전 문제 또한 고려해야 한다. 특히, 넓은 범위에서 발생하는 녹조 문제의 경우, 인력 투입을 통한 대응은 물리적, 경제적 한계에 부딪힐 수 있다.

## III. 시스템 설계

### 3.1 시스템 개요

본 연구에서 제안하는 무선 통신 기반 수질 데이터 수집 및 시각화 시스템은 다음과 같은 요소로 구성된다.

#### 1) 센서 네트워크

pH, DO, 수온 등의 다양한 수질 센서를 사용하여 데이터를 실시간으로 수집한다.

#### 2) 통신 모듈

무선 통신 기술을 사용하여 센서 데이터가 중앙 서버로 전송되도록 한다.

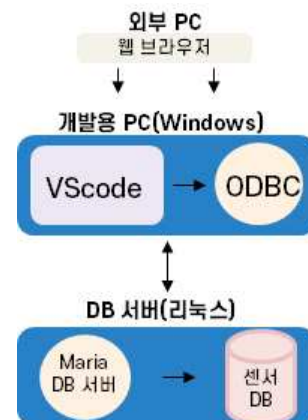


Fig. 2. 무선 통신 기술

3) 데이터 처리 서버

중앙 서버는 수집된 데이터를 분석하고, 머신러닝 알고리즘을 통해 녹조 발생 가능성을 예측한다.

4) 웹 기반 인터페이스

사용자에게 실시간 데이터와 분석 결과를 시각적으로 제공하며, 이상 상황 발생 시 경고 알람을 제공한다.

3.2 주요 기능

실시간 데이터 수집 및 전송 : 센서를 통해 수집된 데이터는 무선 통신 모듈을 통해 중앙 서버로 전송된다.

1) 데이터 분석 및 시각화

서버는 수집된 데이터를 분석하고, 머신러닝 모델을 통해 녹조 발생 가능성을 예측한다.

2) 경고 알람 시스템

예측된 녹조 발생 가능성이 높은 경우, 사용자에게 경고 알람을 제공하여 신속한 대응을 유도한다.

1) 수질 센서

pH, DO, 수온을 측정할 수 있는 센서를 사용하여 실시간 데이터를 수집한다. 센서는 수면 위에 위치한 자율제어 장치에 탑재된다.

2) 통신 모듈

수집된 데이터를 서버로 전송할 수 있는 블루투스 및 와이파이 모듈을 사용한다.

3) 데이터 처리 장치

아두이노와 같은 마이크로컨트롤러를 사용하여 데이터를 실시간으로 처리하고 전송한다.

4) 전력 공급 장치

배터리를 이용하여 지속적으로 전력을 공급한다.

4.2 소프트웨어 구성

IV. 시스템 설계

4.1 하드웨어 구성

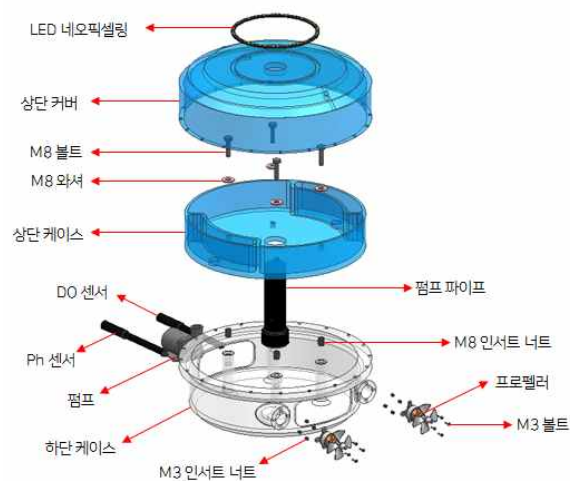


Fig. 3. 하드웨어 구성도



Fig. 4. 모니터링 시스템 페이지

1) 데이터 수집 및 전송 모듈

센서로부터 수집된 데이터를 아두이노를 통해 실시간으로 서버로 전송한다.

2) 데이터 전처리 모듈

수집된 데이터의 노이즈 제거, 이상치 처리 등을 수행하여 분석에 적합한 데이터를 준비한다.

3) 녹조 예측 알고리즘

랜덤 포레스트 회귀 모델을 사용하여 녹조 발생 가능성을 예측한다.



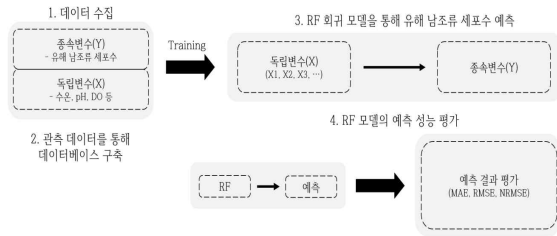


Fig. 5. 녹조 예측 알고리즘

#### 4) 실시간 알림 시스템

예측된 결과가 임계값을 초과할 경우, 사용자에게 실시간 알림을 제공한다.

### V. 실험 설정 및 결과

#### 5.1 실험 설정



Fig. 6. 실험 진행 장소

장소: 공주대학교 인근 호수

기간: 2024년 6월 ~ 8월

장비: 자율제어 장치, 수질 센서(PH, DO, 수온), 통신 모듈(블루투스, 와이파이)

방법: 자율제어 장치를 통해 수질 데이터를 수집하고, 수집된 데이터를 서버로 전송하여 머신러닝 모델을 통해 녹조 발생 가능성을 예측하였다.

#### 5.2 결과 분석

데이터 수집 정확도 : 센서로 수집된 데이터의 정확도는 평균 95% 이상을 기록하였다.

주행 정확도 : 정확한 위치로 기기가 이동하였다.

실시간 대응성 : 시스템은 녹조 발생 가능성을 사전에 예측하고 알림을 제공하여 신속한 대응이 가능했다.

### VI. 결론 및 향후 연구 방향

본 논문은 무선 통신 기술을 활용하여 실시간으로 수질 데이터를 수집하고 분석하는 시스템을 제안하였다. 제안된 시스템은 기존의 수질 모니터링 방법의 한계를 극복하고, 실시간 데이터 수집과 분석을 통해 녹조 발생을 예측하여 환경 보호와 공중 보건에 기여할 수 있다. 향후 연구에서는 대규모 수역에서의 시스템 적용 가능성을 평가하고, 다양한 오염 물질을 감지할 수 있는 추가 센서의 통합을 계획하고 있다. 또한, 고급 전력 절약 알고리즘을 개발하여 시스템의 에너지 효율성을 향상시킬 예정이다.

#### 참고문헌

- [1] Roger S. Pressman "Software Engineering A Practitiners' Approach" 3rd Ed. McGraw Hill
- [2] 환경부. 전국 녹조경보제 운영지점 및 발령 현황 2022
- [3] 미국 미네소타대학 연구팀. 네이처 커뮤니케이션스 (Nature Communications)
- [4] Smart Water Quality Monitoring with IoT Wireless Sensor Networks
- [5] A system for monitoring water quality in a large aquatic area using wireless sensor network technology
- [6] 해양적응형무선센서네트워크기반의수중환경모니터링시스템



# YOLO 알고리즘을 활용한 시각장애인 보행 보조 장치

공일총\*, 김민준\*\*, 변재휘\*\*\*, 서형주\*\*\*\*, 이관형\*\*\*\*\*

## Walking assist device for the visually impaired using YOLO algorithm

Gong Il Chung\*, Kim Min Joon\*\*, Byeon JaeHwi\*\*\* Suh Hyung Joo\*\*\*\*, and Lee Kwanhyeong\*\*\*\*\*

### 요 약

일반적으로 시각장애인은 보행 보조 장치인 지팡이를 이용하여 보행에 도움을 받고 있다. 현재 시각장애인이 사용하는 지팡이는 손 사용과 객체 인식에 대한 불편한 점이 있다. 이와 같은 단점을 보완하기 위해서, 본 연구에서는 객체를 인식할 수 있는 시각장애인용 보행 보조장치와 지도를 제작한다. 객체 인식은 You Only Look Once (YOLO) 알고리즘을 이용하여 보행 보조 장치를 제작하고, 지도제작은 TMap Software Development Kit (SDK) 를 사용한다.

### Abstract

In general, visually impaired people are assisted in walking by using a walking stick, which is a walking assistance device. Currently, the walking stick used by the visually impaired has inconvenience in hand use and object recognition. To compensate for these shortcomings, this study creates a walking assistance device and a map for the visually impaired that can recognize objects. Object recognition uses the You Only Look Once (YOLO) algorithm to produce a walking assistance device, and TMap Software Development Kit (SDK) is used for mapping.

### Key words

Walking Aid, YOLO, Object Recognition

## I. 서 론

최근 인공지능의 발전으로 객체 인식 알고리즘이 발전하고 있다. 객체 인식 알고리즘이란 이미지 또는 영상 내의 객체를 식별하기 위한 알고리즘을 나타낸다. 일반적으로 객체 식별 알고리즘은

Regions with Convolutional Neural Networks features (R-CNN)과 You Only Look Once (YOLO) 알고리즘이 있다. 시각장애인이 지팡이를 사용하여 대중교통을 이용할 때 한쪽 손을 사용할 수 없다. 그래서 본 연구에서는 기존 시각장애인의 지팡이의 단점을 보완한 시각장애인용 보행보조장치를 개발한다.

\* 대진대학교 휴먼·로봇융합전공, 20191427@daejin.ac.kr,

\*\* 대진대학교 휴먼·로봇융합전공, joon340@naver.com(교신저자)

\*\*\* 대진대학교 휴먼·로봇융합전공, shjshj1225@gmail.com(교신저자),

\*\*\*\* 대진대학교 휴먼·로봇융합전공, bb7133@naver.com(교신저자),

\*\*\*\*\* 대진대학교 AI융합학부 교수, khlee@daejin.ac.kr

## II. 보행 보조를 위한 객체 인식 알고리즘

본 논문에서는 객체 인식 알고리즘으로 YOLO 알고리즘을 사용한다. YOLO 알고리즘이란 2015년에 발표된 객체 인식 알고리즘으로 물체 검출시 이미지를 추출하여 처리한다.

R-CNN 알고리즘은 선택적 검사를 이용하여 후보영역을 추출하고, 합성곱 신경망 (CNN, Convolution Neural Network) 을 사용하여 매개 영역에서 하나의 고정 길이의 특징 벡터를 추출하고, 해당 특징 벡터를 서포트 벡터 머신에 입력하여 분류한다. [1] 그러나 YOLO 알고리즘은 그림 1과 같이 한 번의 패스에서 전체 영역의 탐지가 가능하여 속도가 빠르다. YOLO 알고리즘은 2024년 3월 기준으로 버전 YOLOv1 ~ YOLOv9까지 공개되었다. YOLO 알고리즘은 Jetson Orin Nano를 이용하여, Linux 환경 중 하나인 Ubuntu에서 실행하였으며 Ubuntu버전은 22.04버전을 사용한다.

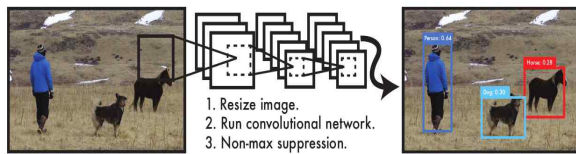


그림 1. YOLO(You Only Look Once) 알고리즘의 동작과정[2]

Fig. 1. The operation process of the YOLO algorithm[2]

YOLO 알고리즘의 객체 인식은 Jetson Orin Nano를 이용하여 영상정보를 수신한다. 수신된 영상정보는 YOLO 알고리즘에 의해 영상 내의 물체를 합성곱 신경망 (CNN, Convolutional Neural Network) 을 기반으로 그리드 (Grid) 에서 분할한 영상을 신경망에 통과시킨 뒤, 바운딩 박스 (Bbox, Bounding box) 나 예측 등을 통한 기법으로 최종 출력물을 생성한다. 이미지로 목록을 만든 후, 만든 내용을 불러와서 카메라로 인식되고 있는 물체 중 목록에 있는 것들을 비교하여 생성된 출력내용을 기반으로 탐지된 객체를 구분한다. 구분된 객체는 횡단보도와 사람, 자동차 등의 정보를 스피커 등 음성 출력장치로 전달되어 사용자에게 음성정보로 전달된다.

## III. 객체 인식 알고리즘 활용을 위한 application 개발

Application 개발은 Android Studio를 사용하였고 기능으로서는 지도와 실시간 영상 촬영을 할 수 있다. 지도 제작은 TMap Software Development Kit (SDK)를 이용한다. 지도와 application, YOLO모델 결합은 Software Development Kit (SDK)를 Android Studio에서 처리한다. Android Studio에 TensorFlow Lite 종속성과 TFLite 모델 파일을 추가한 후 TFLite Interpreter를 사용해 모델을 불러오고 객체 탐지를 수행하여 YOLO 모델에 연결한다. 학습시킨 YOLO 모델을 Open Neural Network Exchange(ONNX)방식을 사용하여 TensorFlow Lite 포맷으로 변환한다.

지도 기능은 Global Positioning System (GPS) 를 이용하여 사용자의 위치를 탐지하고 사용자의 현재 위치를 지도상에 표시한다. 사용자가 목적지를 음성 또는 타자로 입력하고 검색 버튼을 누르면 목적지가 설정된다. 현재 위치에서 목적지까지의 경로를 찾고, 해당 경로는 지도 위에 선의 형태로 표시된다. 그림 2는 보행 보조 장치의 동작을 설명하는 블록도이다.

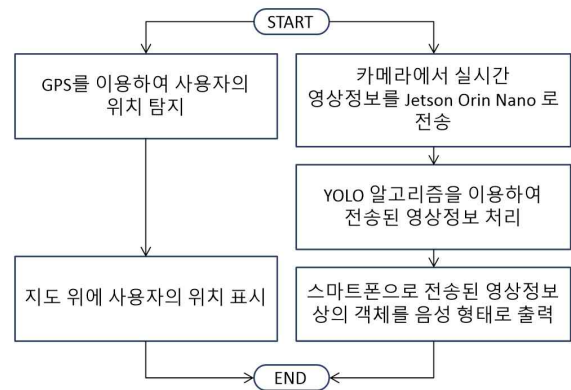


그림 2. 보행 보조 장치의 동작과정

Fig. 2. Operation Process of the Walking Assistance System

실시간 영상 촬영기능은 스마트폰의 카메라를 이용하여 실시간으로 촬영된 영상 정보를 Jetson Orin Nano에 전송한다. Jetson Orin Nano에 전송된 영상정보는 YOLO 알고리즘에 의해 객체별로 구분되어 다시 스마트폰으로 전송된다. 스마트폰으로 전송된

영상정보 상의 객체 정보는 스마트폰을 통해 음성  
의 형태로 출력하게 된다.

#### IV. 결 론

본 논문에서는 시각장애인 손 사용과 객체인식에  
대한 불편함 해결을 위해 객체 인식 알고리즘인  
YOLO 알고리즘을 이용하여 보행 보조 장치를 제작  
하였다. 제작된 보행 보조 장치는 시각장애인의 보  
행을 보조하고 주변 환경의 객체를 인식할 수 있어  
안전하게 보행을 할 수 있다. 또한, 지도 기능을 사  
용하여 목적지까지의 이동을 편리하게 보행을 할  
수 있다.

#### 참 고 문 헌

- [1] 추신. "RPN과 R-CNN 공유를 통한 효율적 실시간 객체 검출 시스템." 국내박사학위논문 배재대학교, 2018. 대전
- [2] <https://jamiakang.github.io/2017/06/18/you-only-look-once-unified-real-time-object-detection/>

# YOLOv5를 이용한 안전 보행 가이드 시스템

임정환\*, 한형규\*\*, 김원호\*\*\*

## A Safe Walking Guide System using YOLOv5

Jeong-Hwan Lim\*, Hyeong-Gyu Han\*\*, Won-Ho Kim\*\*\*

### 요 약

본 논문은 시각장애인을 위한 안전 보행 가이드 시스템을 제안하고 설계 및 구현에 관해 기술한다. 시스템은 인도의 점자 블록과 장애물을 영상 인식하여 사용자에게 경로를 알려주는 기능을 제공한다. 시스템의 하드웨어는 지팡이, 스위치, 프로세서, 스피커, 카메라, 전원으로 구성되며, 영상 인식 알고리즘은 YOLOv5를 기반으로 점자 블록 패턴을 인식하여 경로를 안내하는 기능을 구현하였다. 제작된 시스템을 시험한 결과, 제시된 요구 기능을 만족하였으며 실용성을 확인하였다.

### Abstract

This paper proposes a safe walking guide system for the visually impaired and explains the design and implementation. The system provides image recognition for Braille blocks and obstacles in India to inform users of the route. The hardware of the system consists of a cane, a switch, a processor, a speaker, a camera, and a power source, and the image recognition algorithm recognizes the Braille block pattern based on YOLOv5 to guide the route. As a result of testing the manufactured system, it satisfies the suggested essential functions and confirms its practicality.

### Keywords

Digital image processing, Hough transform, Deep learning, TTS, YOLOv5

## 1. 서 론

점자 블록과 흰지팡이는 시각장애인에게 중요한 안내 역할을 한다. 하지만 점자 블록의 구분과 장애물 회피에 어려움이 있어 추가적인 보행 지원 기술

이 필요하다. 특히 최근 늘어난 키보드의 사용량에 따라 무차별적인 주차된 키보드가 증가하고 있다. 이러한 보행 환경을 반영하여 본 논문에서는 시각장애인이 안정적으로 보행할 수 있도록 영상신호 처리와 딥러닝 기반의 보행 가이드 시스템을 제안

\*\*\* 공주대학교 전기전자제어공학부 전자공학전공, whkim@kongju.ac.kr (교신저자)

\*\* 공주대학교 전기전자제어공학부 전자공학전공 학부생, hans0192@naver.com

\* 공주대학교 전기전자제어공학부 전자공학전공 학부생, 1203321@naver.com

한다. 제안하는 시스템은 카메라로 획득한 영상 데이터를 실시간 분석하여 점자 블록의 패턴을 인식하고, 직진과 방향 전환 및 장애물을 감지하여 안내하는 알고리즘을 제공한다.

## II. 보행 가이드 시스템 설계

### 2.1 하드웨어 설계

그림 1은 시스템의 하드웨어 구성도이다.

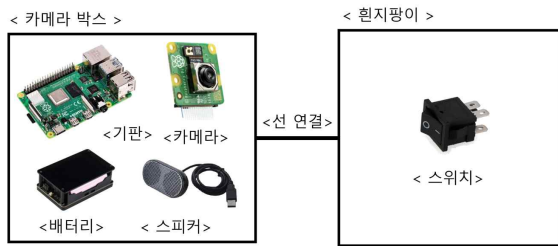


그림 1. 보행 안내 시스템의 하드웨어 구성  
Fig 1. Hardware configuration of walking guide system

시스템은 크게 흰 지팡이와 카메라 박스로 나뉘며 카메라 박스는 라즈베리파이 프로세서 보드, 스피커, 카메라, 전원장치가 장착되고, 흰 지팡이에는 스위치가 장착되어 사용자가 쉽게 전원을 작동시킬 수 있도록 설계하였다.

### 2.2 안전 보행 안내 알고리즘 설계

그림 2는 시각장애인을 위한 안전 보행 알고리즘의 흐름도이다.

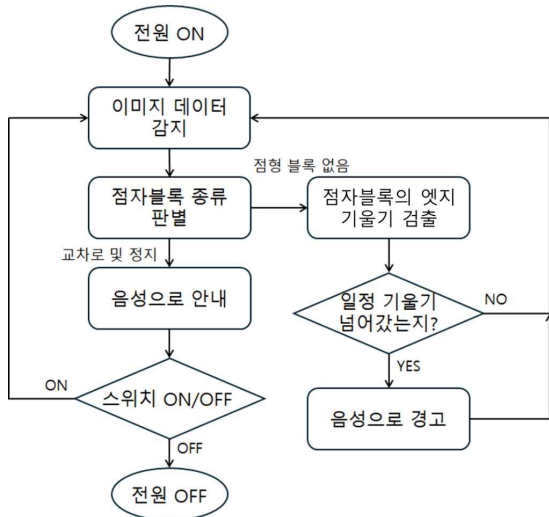


그림 2. 보행 안내 알고리즘 흐름도  
Fig 2. Flow chart of walking guide algorithm

알고리즘의 상세 동작 설명은 다음과 같다. 먼저 전원이 들어오면 카메라에서 실시간 이미지 데이터를 받고, 객체 검출 실행한다. 점형 블록이 인식되는 경우 점자 블록의 엣지 검출을 중지하며, 인식되지 않는 경우 엣지 검출을 실행한다. 이러한 과정을 진행하여 점자 블록의 형태에 따른 안내와 직진을 제대로 하고 있는지에 대한 정보를 실시간으로 안내하여 사용자 앞의 점자 블록 상황을 인지하고 대비할 수 있는 도움을 준다.

### 2.3 점자 블록 및 장애물 객체 검출 기능

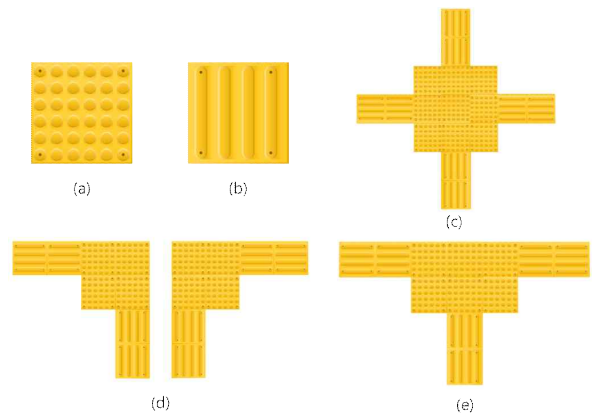


그림 3. 점자 블록 라벨링 대상 예시  
Fig 3. Example of braille block labeling targets

객체 검출은 그림이나 영상 데이터의 사람, 차량과 같은 객체의 위치를 찾을 수 있는 영상처리 기술이다. 이 기술은 공간에 특정 물체가 있는지를 판별하는 객체 감지와 해당 물체가 무엇에 해당하는지를 판단하는 객체 분류로 나눌 수 있으며 본 작품은 점자 블록이 다양하게 조합된 형태가 무엇인지, 장애물이 점자 블록에 존재하는지 판단하므로 객체 분류에 속한다 [1]. 객체 검출을 실시간으로 진행해 주는 모델은 YOLO(You Look Only Once)를 사용하였으며, 타 모델들과 비교적 모델 경량화 및 최적화 기능을 통해 높은 정확도와 더불어 작은 모델 크기를 유지하는 장점이 있다. [2]

점형 블록과 선형 블록은 블록마다 라벨링을 하였으며, 장애물을 제외한 객체들은 중심에 점형 블록을 가지고 선형 블록이 연결되는 개수로 구성됨으로 +자 교차로는 그림 2(c)처럼 모든 방향인 4개, 좌·우회전은 그림 2(d) 2방향, T자 교차로는 3개. 길

끝은 1개로 모든 교차로는 점형 블록과 선형 블록의 조합을 중심으로 라벨링 하였다. 장애물들은 점자 블록 위에 있는 개체만을 선택하여 위험하지 않은 장애물들은 인식되지 않도록 라벨링 하였다.

표 1. 각 객체 간 데이터 개수와 정확도  
Table 1. Number and accuracy of each object

객체	데이터 개수	정확도
우회전	252	76.6%
좌회전	264	70.0%
직•우회전	302	80.0%
직•좌회전	343	83.3%
T자 교차로	257	100.0%
+자 교차로	126	63.3%
길 끝	178	66.7%
장애물	1387	100.0%

라벨링 작업은 표 1의 객체와 같이 각 개체를 선택하여 라벨링 하였으며, 라벨링 데이터를 이용해 객체 검출을 위한 딥러닝을 진행했다. 데이터 학습 횟수는 100회로 진행했으며, 학습 결과는 그림 4과 같다. 모델이 점차 학습하면서 오차율이 점점 줄어드는 모습을 확인할 수 있으며, 이는 데이터의 정확도 증가를 의미한다.

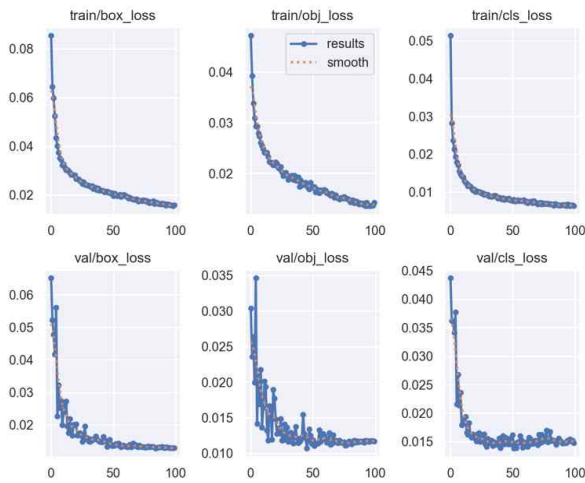


그림 4. 데이터 학습 결과  
Fig 4. Results of data training

표 1의 정확도는 딥러닝을 진행하고 얻어낸 추정치 데이터를 기반으로 시뮬레이션을 진행한 결과이며, 객체별로 이미지 30장씩 인식하였을 때, 얼마나 정확하게 검출하였는지 나타내는 지표이다. +자 교차로와 길 끝은 라벨링 하여 참고자료로 사용된 데이터 개수가 적어 정확도가 상대적으로 적은 수치가 나오는 것을 볼 수 있다. 다른 객체들도 전반적으로 참고자료 데이터의 개수가 적어질수록 정확도 또한 낮아지는 경향을 보여주고 있다.

## 2.4 점자 블록 에지 검출 기능

시각장애인이 점자 블록 위를 보행할 때 방향을 잘못 잡은 경우, 블록을 벗어나는 일을 방지하기 위하여 점자 블록 에지 검출 기능을 구성하였다. 에지 검출 기능은 그림 5의 흐름도처럼 순서대로 진행되도록 설계하였다.

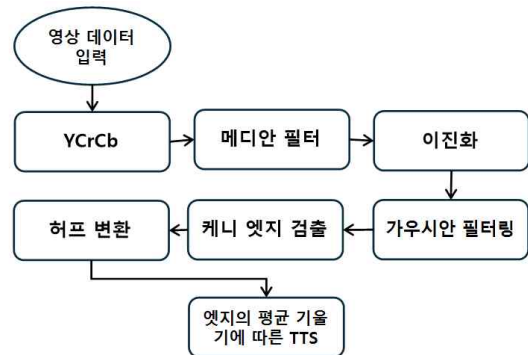


그림 5. 점자 블록 직선 검출 알고리즘  
Fig 5. Braille block lane detection algorithm

먼저 RGB는 삼원색을 사용하고 YCbCr는 빛과 색깔 정보를 같이 담고 있기에 일반적으로 사용하는 RGB를 대신하여 YCbCr를 사용하여 좀 더 섬세하게 색을 구별할 수 있도록 하였다. 메디안 필터는 점자 블록의 에지를 강조하고 주변 환경에 대한 잡음 또한 제거하는 효과가 있어 추출하기 쉽게 도와 줄 수 있다. 이진화 처리는 노란색 성분 범위를 설정하여 특정 값을 흑 또는 백으로 구분 짓는다. YCbCr 색 공간 중 Cr 성분을 이용해서 노란색 범위를 구하려면 식 (1)의 계산이 필요하다. 노란색은 R = 255, G = 255, B = 0이므로, Cr값은 148.7로 나타낼 수 있다.



$$Cr = 0.500R - 0.419G - 0.081B + 128 \quad (1)$$

그러나 점자 블록은 순수 노란색이 아니고 노란색 계열이므로 노란색 성분 범위를 100~145로 잡으면 이진화 처리를 더욱 효과적으로 할 수 있다. 다음으로 이진화된 이미지에 2차원의 가우시안 필터링을 거친다. 가우시안 필터는 이진화 데이터의 잡음을 제거하고, 회전에 대해서 대칭인 특성이 있어 케니 에지 검출이 효과적으로 점자 블록의 에지 검출이 가능하다. [3] 이후 확률적 허프 변환을 사용하여 검출된 에지에서 특정 길이의 직선들을 모아 기울기를 측정하고, 수집한 기울기들의 평균을 잡아 특정 기울기에 도달하면 사용자에게 경보를 제공한다.

### III. 실험 및 결과

그림 6은 제작된 안전 보행 시스템 사진이다. 이를 이용하여 안전 보행 가이드의 기능을 시험하였다.



그림 6. 제작된 안전 보행 가이드  
Fig 6. Implemented safe walking guide

점자 블록이 있는 실제 인도 구역에서 실험하였으며, 그림 7과 같이 선형과 점형 블록은 블록마다 라벨이 적용되고, 직 좌회전과 T자 교차로 같은 특정 객체 구역은 따로 라벨이 적용되며, 음성으로 안내가 되는 것을 확인하였다.

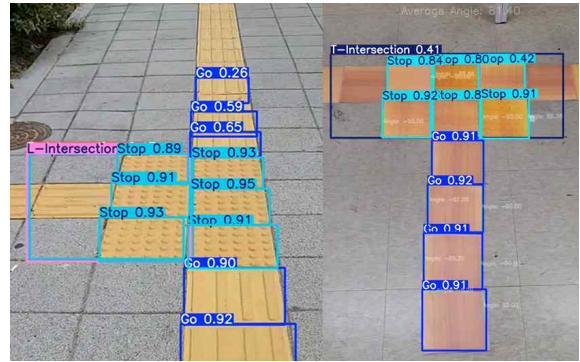


그림 7. 보행 안내 알고리즘 처리 결과 화면  
Fig 7. Processing result screen of walking guide algorithm

### IV. 결론

본 논문에서는 점자 블록 객체 검출과 점자 블록 에지 검출 기능을 사용하여 시각장애인의 보행 안내 알고리즘을 설계하고 구현하였다. 제작된 작품은 점자 블록이 놓인 길로 다니며 보행자의 앞의 상황을 카메라로 인식하며 길을 안내하거나 위험한 상황을 경고하여 좀 더 안전한 보행을 보조하여 시각장애인들의 외부 활동의 증가할 것으로 기대된다. 향후 연구에서는 더욱 정확도가 높은 객체 인식을 위하여 참고자료로 사용되는 데이터 세트를 더욱 수집하고, 점자 블록의 에지 검출을 더욱 효과적인 방법으로 할 수 있는 방법을 연구할 계획이다.

### 참 고 문 헌

- [1] Liu, L., Ouyang, and W., Wang, X, "Deep Learning for Generic Object Detection: A Survey", Int J Comput Vis 128, 261-318, Sept, 2019
- [2] Junekoo Kang, Valentin Bajenez, Somyeong Ahn, Minwoo Sung, and Youngseok Lee, "A Method to Enhance the Accuracy of Braille Block Recognition for Walking Assistance of the Visually Impaired: Use of YOLOv5 and Analysis of Vertex Coordinates", Journal of KIISE, 49(4), pp. 291-297, Apr, 2022.
- [3] Alasdair McAndrew, and TaeHyo Kim, 매트랩을

이용한 디지털 영상처리의 기초, 한티미디어,  
2016.

- [4] Hee-jin Kim, Jae-hyuk Yoon, and Soon-kak Kwon, "Deep Learning Braille Block Recognition Method for Embedded Devices", Journal of the Korea Industrial Information Systems Research, 28(4), pp. 1-9, Aug, 2023
- [5] Seong-Hyeon Park, and Young-Gyun Kim, "A Study of Detecting Curved Lane by Hough Transform for Autonomous Driving", 한국정보과학회 학술발표논문집, pp. 2104-2106, Dec, 2017
- [6] Gi-Seok Kim, Jin-Wook Lee, and Jae-Soo Cho, "Study on Effective Lane Detection Using Hough Transform and Lane Model", INFORMATION AND CONTROL SYMPOSIUM, pp. 34-36, May, 2009

# 자율주행 차량의 도시 교통 혼잡 완화를 위한 최적 신호 제어 시스템 연구

조성빈\*, 나관상\*\*

## Optimization of Traffic Signal Control System for Urban Congestion Mitigation with Autonomous Vehicles

SungBeen Cho\*, and Kwansang Na\*\*

### 요 약

본 연구는 자율주행 차량의 실시간 데이터를 활용하여 도시 교통 혼잡을 완화할 수 있는 최적 신호 제어 시스템을 제안한다. 자율주행 차량이 보급될 경우 기존 고정 신호 체계는 한계에 부딪히며, 혼잡과 대기 시간이 늘어날 수 있다. 이를 해결하기 위해 실시간 신호 제어 알고리즘을 개발하여 교차로의 교통 신호를 최적화하고, 기존 시스템과의 비교 평가를 통해 효과성을 검증하고자 한다. 이 연구는 자율주행 차량 상용화 및 교통 인프라 효율성 개선에 기여할 것이다.

### Abstract

This study proposes an optimized traffic signal control system to alleviate urban traffic congestion by utilizing real-time data from autonomous vehicles. As autonomous vehicles become more prevalent, fixed-pattern signal systems may become increasingly limited, leading to potential increases in congestion and wait times. To address this, a real-time signal control algorithm is developed to optimize intersection signals, and its effectiveness is evaluated by comparing it with conventional systems. This research aims to contribute to the commercialization of autonomous vehicles and improve the efficiency of urban traffic infrastructure.

### Key words

sAutonomous Vehicles, Traffic Signal Optimization, Urban Traffic Congestion, Real-Time Data, Intelligent Transportation Systems (ITS), Signal Control Algorithm, Smart City Infrastructure

### 1. 서 론

자율주행 기술은 미래의 스마트 시티와 교통 문

제 해결에 중요한 역할을 합니다[1]. 그러나 자율주행 차량의 효과적인 도입을 위해서는 교통 신호와 같은 기존 인프라와의 통합이 필수적입니다[2]. 현

\*경기대학교 융합과학대학 컴퓨터공학부, sb11210@kyonggi.ac.kr

\*\*경기대학교 소프트웨어경영대학 AI컴퓨터공학부, ksna001@kyonggi.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음 (2021-0-01393)

재 고정된 패턴의 신호 시스템은 자율주행 차량이 생성하는 실시간 데이터를 반영하지 못하여 교통 체증과 비효율을 유발할 가능성이 큼니다. 이에 따라 본 연구는 자율주행 차량의 데이터를 실시간으로 반영하여 도심 교통 혼잡을 완화하는 신호 제어 시스템을 제안하고자 합니다.

## II. 본론

### 1. 데이터셋 및 전처리

본 연구에서는 자율주행 차량의 실시간 데이터를 활용하여 교차로의 신호 제어를 최적화합니다. 이를 위해 교통 데이터셋을 수집하여 모델 학습에 필요한 형태로 전처리하였습니다. 데이터 전처리 과정에서 불필요한 변수를 제거하고, 결측치를 보완하여 분석의 신뢰성을 높였습니다. 각 교차로의 신호 변경 빈도, 차량 흐름, 시간대에 따른 혼잡도 등을 반영하여 데이터셋을 구성하였습니다. 이러한 접근 방식은 교통 신호의 적시 조정을 가능하게 하며, 과거 연구에서도 이러한 데이터 활용의 중요성이 강조되었습니다[3].

### 2. AI 기반 신호 제어 알고리즘

본 연구에서는 교차로에서 실시간 교통 상황을 반영할 수 있도록 인공지능(AI) 기반 최적화 알고리즘을 개발하였습니다. 이 알고리즘은 자율주행 차량의 위치, 속도, 방향 데이터를 학습하여 교통량 변화에 실시간으로 대응할 수 있습니다. 특히, 자율주행 차량이 통과하는 시점을 예측하여 신호 변경 시간을 동적으로 조정함으로써 대기 시간을 줄이고 교통 효율을 높이는 데 중점을 두었습니다[1].

### 3. 모델 성능 평가

개발된 신호 제어 시스템의 성능을 검증하기 위해 다양한 교통 상황을 시뮬레이션하였습니다. 시뮬레이션에서는 기존 고정 신호 시스템과 AI 기반 동적 신호 시스템을 비교하였으며, 정확도(Accuracy), 정밀도(Precision), 재현율(Recall) 등의 성능 지표를 사용하여 각 시스템의 효율성을 평가하였습니다. 평가 결과, AI 기반 신호 제어 시스템이 대기 시간과 교통 혼잡 완화에 있어 우수한 성능을 보였습니다[2].

## III. 결 론

본 연구는 자율주행 차량의 실시간 데이터를 활용한 교통 신호 최적화 시스템을 통해 도시 교통 혼잡을 줄이는 방안을 제시합니다. 실시간 데이터를 반영한 동적 신호 제어와 AI 기반 최적화 알고리즘을 통해 자율주행 차량의 운행 효율을 높이고, 도시 교통의 지속 가능한 발전에 기여할 수 있을 것입니다. 향후 연구에서는 실제 도심 환경에서의 시스템 적용과 효과 검증을 통해 실용성을 높이고자 합니다.

## 참 고 문 헌

- [1] Zhang, Y., & Zhao, J. (2023). Integrating Autonomous Vehicles (AVs) into Urban Traffic: Simulating Driving and Signal Control. *Applied Sciences*, 14(19), 8851.
- [2] Wang, X., et al. (2023). Research on Multi-Objective Optimization Models for Intersection Crossing of Connected Autonomous Vehicles With Traffic Signals. *IEEE Transactions on Intelligent Transportation Systems*.
- [3] Jing, P., Huang, H., & Chen, L. (2017). An Adaptive Traffic Signal Control in a Connected Vehicle Environment: A Systematic Review. *Information*, 8(3), 101.

# 3축 직교 좌표 로봇을 이용한 AI 기반 스마트 분리수거 시스템 개발

이윤승\*, 문영민, 박승규, 최성원, 신진호\*\*

## Development of an AI-Based Smart Recycling System Using a 3-Axis Cartesian Robot

Yoonseung Lee\*, Youngmin Moon, Seunggyu Park, Seungwon Choi, and Jin-Ho Shin\*\*

### 요약

본 논문에서는 3축 직교 좌표 로봇을 이용한 AI 기반 스마트 분리수거 시스템을 개발한다. 개발된 시스템은 AI 기반 영상 인식 기술과 센서들을 이용하여 폐기물을 자동으로 분류하고 재활용 가능 자원을 효율적으로 이용하는 것을 목표로 한다. NVIDIA TensorRT detectNet을 젯슨 나노에 내장한다. 내장된 TensorRT detectNet을 활용하여 투입 객체는 인식되며, 깊이 카메라와 직교 좌표 로봇을 이용해 분리 작업을 수행한다. 제안한 분리수거 시스템은 물체 투입 후 직교 좌표 로봇, 기계학습 및 초음파 세척 방법을 통하여 전체 분리수거 과정을 자동적으로 수행한다. 특히 자동 초음파 세척 과정을 통하여 재활용품 품질을 개선할 수 있다.

### Abstract

In this paper, we develop an AI-based smart recycling system using a 3-axis Cartesian robot. The developed system aims to automatically classify wastes and to efficiently use recyclable resources using an AI-based image recognition technology and sensors. We embed NVIDIA TensorRT detectNet into Jetson Nano. The input objects are recognized using the embedded TensorRT detectNet, and the waste separation work is performed using a depth camera and a Cartesian robot. The proposed recycling system automatically performs the total recycling process through a Cartesian robot, machine learning, and ultrasonic cleaning method after inputting objects. Especially, it can improve the quality of recyclable materials through an automatic ultrasonic cleaning process.

### Key words

Recycling System, AI, 3-Axis Cartesian Robot, Ultrasonic Cleaning, Image Recognition

## I. 서론

현대 사회에서 폐기물 처리와 자원 재활용의 비효율성은 중요한 환경 문제로 대두되고 있다. 플라

스틱 재활용률이 낮고, 잘못된 분리수거는 자원 낭비와 재활용품 품질 저하를 초래한다. 특히, 이물질과 라벨 제거는 재활용 과정에서 큰 도전 과제로, 제거가 제대로 이루어지지 않으면 재활용품은 폐기

\* 동의대학교 전자공학과 학부생, Busan, Korea. (dlbstmt0315@naver.com) (주저자)

\*\* 동의대학교 전자공학과 교수, Busan, Korea. (jshin7@deu.ac.kr) (교신저자)

※ 이 논문은 동의대학교 산학협력력 선도대학 육성사업(LINC 3.0)의 지원을 받아 수행되었습니다.

된다. 이를 해결하기 위해 본 논문에서는 AI 기반 스마트 분리수거 시스템을 제안한다. 딥러닝 기반의 영상 인식 기술, 센서 및 초음파 세척 기술을 결합하여 폐기물의 자동 분류와 이물질 제거를 통해 재활용 효율성을 높인다. 개발된 시스템은 공공장소와 산업 현장에서 분리수거의 정확성과 효율성을 높이고, 자원 관리 비용 절감과 재활용률 향상에 이바지할 것으로 기대된다.

## II. 스마트 분리수거 시스템

### 1. 시스템 구성 요소

[그림 1]과 같이 젯슨 나노(Jetson Nano), STM32 보드(F103RB,446RE), 엔코더 모터, 깊이 카메라 (Depth Camera), 웹캠(Webcam), 3축 직교 좌표 로봇, 초음파 센서, 수위 센서, 리미트 스위치, 스텝 모터(Stepper Motor), 전원공급장치로 구성하였다.

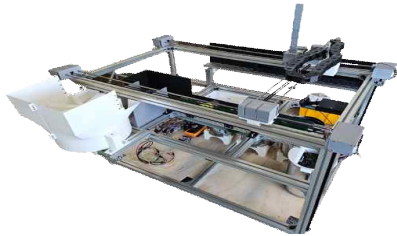


그림 1. 개발된 스마트 분리수거 시스템  
Fig. 1. The developed Smart recycling system

### 2. 시스템 동작 원리

#### 2.1 객체 인식과 2차원 좌표점 추출

젯슨 나노에 있는 카메라가 폐기물을 감지하면 NVIDIA TensorRT DetectNet 기반의 딥러닝 모델을 통해 젯슨 나노가 실시간으로 인식한 물체의 데이터를 처리하고, 깊이 카메라로 물체의 거리를 측정한다. DetectNet은 객체 검출을 수행하는 딥러닝 모델로, 입력된 영상에서 폐기물의 위치와 종류를 자동으로 식별하는 역할을 담당한다.

DetectNet이 인식한 물체를 OpenCV를 이용하여, 컨베이어 벨트를 중심으로 둔 후, 물체가 왼쪽에 있을 경우의 좌표점과 오른쪽 끝에 있을 경우의 Y축의 좌표를 계산한다. 이때 객체 인식은 젯슨 나노에 연결된 웹캠은 Y축 좌표, 깊이 카메라로는 X축 좌표를 도출하여 2차원 좌표점(X,Y)을 추출한다.

#### 2.2 로봇 제어를 이용한 분류 작업

스텝 모터로 3축 직교 좌표 로봇을 구동시켜 객체 인식된 물체를 이동시킨다. 본 3축 직교 좌표 로봇에서 스텝 모터의 정확한 제어를 위해 타이밍 벨트, 풀리, 감속비, 스텝당 각도 등 다양한 요소를 고려하여 [1]에서 제시된 수식을 이용한다.

카메라로부터 목표 좌표값(X,Y)을 전달받아 X,Y축의 스텝 모터로 목표 좌표값에 도달하도록 제어한다. Z축은 목표 좌표값을 따로 받지 않고 리미트 스위치가 물체를 감지할 때까지 Z축 스텝 모터를 구동한다. X축, Y축을 동시에 제어를 한 후에 Z축 모터를 구동한다. 이때 Z축 모터는 리미트 스위치가 컨베이어 벨트에 닿지 않도록 제어한다.

#### 2.3 세척 과정

분류된 재활용품은 초음파 세척기로 이물질을 제거하여 품질을 유지한다. [그림 2]는 깊이 카메라와 웹캠이 폐기물을 인식하고 스텝 모터와 엔코더 모터가 협력하여 분류하는 과정을 보여준다. 최종적으로 분류된 폐기물은 세척된다.

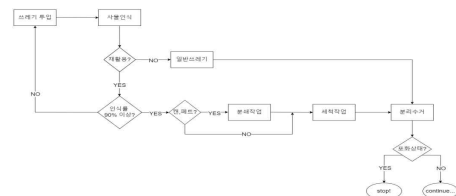


그림 2. 스마트 분리수거 시스템의 동작 흐름도  
Fig. 2. Flow chart of a smart recycling system

## III. 실험 결과

### 3.1 객체인식 성능평가

캔, 플라스틱, 병과 같은 폐기물의 실시간 검출 성능을 평가하였다. 젯슨 나노의 GPU 가속을 통해 최적화된 DetectNet 모델은 직접 촬영한 플라스틱 병 ( 500장 ), 캔 ( 300장 ), 병 ( 200장 )을 데이터셋으로 활용하였다. [그림 3]과 같이 프레임당 평균 79%의 mAP(Mean Average Precision)을 달성하여 임베디드 환경에서도 높은 정확도를 보였다. 각 클래스별 AP(Average Precision)는 캔은 88.2%, 플라

스틱 87.4% , 유리병 61.6%으로, 유리병의 검출 성능이 상대적으로 낮아 추가적인 데이터 증강 및 모델 개선의 필요성을 확인하였다. 데이터셋을 더 늘리면 훨씬 인식률을 확대할 수 있다.

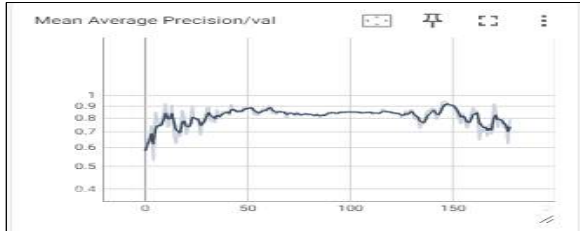


그림 3. 유리병, 캔, 플라스틱에 대한 mAP 결과  
Fig 3. mAP results for bottle, can, and plastic.

### 3.2 스텝 모터 제어 실험 결과

본 3축 직교 좌표 로봇 시스템은 3개의 스텝 모터가 장착되어 제어된다. 스텝 모터에 대한 이동거리를 제어한 실험 결과는 [그림 4]와 같다.

이동거리 오차를 보정하기 위하여 캘리브레이션 과정을 실시하였다. 이동거리에 대한 목표값을 50cm로 설정하여 100번의 실험 데이터를 수집하고 이를 식 (1)에 의해 평균한 보정값 0.13cm를 산출했다. [그림 4]는 목표 이동거리, 보정 전 이동거리, 보정 후 이동거리를 나타내었다.

$$\text{보정값} = \frac{1}{100} \sum_{n=1}^{100} (|E_n - T|) \quad (1)$$

$E_n$ : 이동거리 측정값  $T$ : 목표값(50cm)  $n$ : 실험 횟수

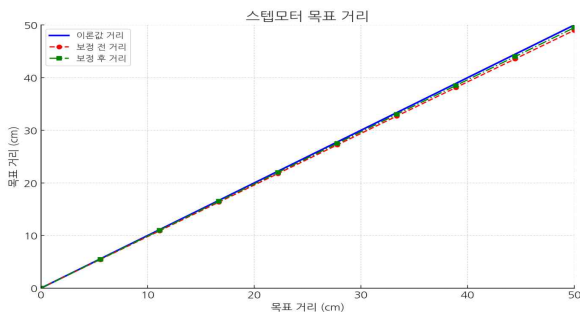


그림 4. 스텝모터 이동거리 제어 결과  
Fig. 4. Result of Step Motor Movement Distance Control

스텝 모터 제어 프로그램에 보정값을 적용하여 실시간으로 이동거리 오차를 최대 3mm까지 줄였다. 이를 통해 정밀 제어가 가능해졌다. 이러한 방식은 특히 물체의 무게나 외부 조건으로 인해 발생

할 수 있는 다양한 변수에 대응할 수 있었다. 캘리브레이션 과정을 많이 수행할수록 훨씬 정밀한 제어가 가능하다.

### 3.3 초음파 세척 실험 결과

본 연구에서는 폐기물 품질 유지를 위해 초음파 세척력 변화를 연구하는 실험을 수행하였다. 플라스틱, 캔, 유리병을 대상으로 세척 효과를 비교하였다. 세척력 판단기준은 정밀측정장비가 없어 육안으로 판단하였다. 상온에서 세척실험을 진행하였으며 세척시간을 30초, 1분 30초, 3분으로 조정하였다. 세척시간이 최대 3분이 적절하다고 판단하였다[2]. [표 1]과 같이 캔은 30초 이상 세척시 깨끗해졌으며 플라스틱과 유리병은 1분 30초 이상부터 세척력이 좋은 것으로 나타났다.

표 1. 초음파 세척 실험 결과

Table 1. Experimental result for ultrasonic cleaning

시간 / 세척 대상	30초	1분 30초	3분
캔	세척력이 충분함	30초와 차이 없음	30초와 차이 없음
플라스틱	세척력이 부족함	30초보다 세척력이 개선	1분 30초와 차이 없음
유리병	세척력이 부족함	30초보다 세척력이 개선	1분 30초와 차이 없음

## IV. 결 론

본 논문에서는 AI 기반 스마트 분리수거 시스템을 제안하였다. 제안된 시스템은 딥러닝 기반 영상 인식과 센서를 활용하여 폐기물을 자동으로 분류하며, 초음파 세척 기술을 통해 이물질들을 제거하며 재활용 품질을 크게 향상시켰다. 개발된 시스템은 기존의 분리수거 방식보다 정확성과 효율성을 높이며, 재활용품의 품질을 유지하는 데 이바지할 수 있다. 본 시스템을 통해 공공장소 및 산업 현장에서 지속적인 자원 관리의 가능성을 제시하였다.

## 참 고 문 헌

- [1] <https://blog.naver.com/newspacekor/30167060631>
- [2] 박종수, “초음파를 이용한 고추세척기를 개발”, 한국농업기계학회 1996년도 학술발표대회 논문집, pp. 85-89, June, 1996.

# 대학 챗봇 활용의 인식 및 효율성 평가

서동섭\*, 한동균\*, 박상우\*, 손경락\*, 박세진\*\*

## Awareness and Efficiency Evaluation of University Chatbot

Dong-Seop Seo\*, Dong-Gyun Han\*, Sang-Woo Park\*, Kyeong-Rak Son\*, and Se-Jin Park\*\*

### 요 약

본 논문은 대학 내 OpenAI 기반 챗봇 도입에 대한 학생들의 인식과 사용 의향을 분석하는 데 목적이 있다. 설문조사를 통해 학생들은 챗봇이 학습 지원 및 행정 업무에서 효율적이고 편리하다고 평가했으며 특히 반복적인 업무에 대한 신속한 응답 제공에 긍정적인 반응을 보였다. 그러나 일부 학생들은 챗봇의 사용이 학습 의지 저하와 AI 의존에 따른 문제 해결 능력 저하를 우려했으며, 개인정보 보호와 데이터 보안 문제도 제기되었다. 본 연구는 이러한 우려를 해결하기 위한 보완책과 기술적 개선이 필요하다고 제안하며 대학과 OpenAI 간의 협력 방안을 모색하는 데 기초 자료를 제공할 수 있다.

### Abstract

The purpose of this paper is to analyze students' perceptions and intentions to use OpenAI-based chatbots in universities. Through the survey, students evaluated chatbots as efficient and convenient in learning support and administrative tasks, and they responded positively to providing quick responses to repetitive tasks. However, some students were concerned that the use of chatbots would reduce their willingness to learn and their ability to solve problems due to AI dependence, and privacy and data security issues were also raised. This study suggests that complementary measures and technical improvements are needed to address these concerns and can provide basic data for seeking cooperation between universities and OpenAI.

### Key words

chatbot, student perceptions, data security

## I. 서 론

디지털 시대의 발전으로 인해 일상생활 속뿐만 아니라 교육 환경도 빠른 변화가 보이고 있으며 특히 AI 기술이 급격하게 발전함으로써 교육 현장에 AI 기술이 도입되고 있다. 이는 학생들이 수동적인 학습자에 머물러있지 않고 능동적으로 참여하고 상

호작용할 수 있는 교육 환경이 조성되고 있다는 것을 의미한다. 많은 AI 기술 중 OpenAI의 챗봇을 활용한 학습이 많아지고 있다. OpenAI 기반 챗봇은 대화형 서비스로 알려져 있고 수많은 데이터 기반으로 신속한 답변 제공 능력을 갖추고 있다. 실제로 숙O여대에서는 국내 최초로 대학 자체 챗봇을 개발하여 서비스를 시작하였다. 따라서 본 논문에서

\* 계명대학교 컴퓨터공학과, ehdtjq1278@naver.com, hdg5639@naver.com, ja513698@naver.com, rudfkrlife@naver.com,

\*\* 계명대학교 컴퓨터공학과, baksejin@kmu.ac.kr



는 대학 내 챗봇을 도입하여 행정 서비스 및 교육 현장에 어떠한 변화가 생기는지, 도입 가능성과 학생들의 인식과 사용 의향을 분석함으로써 향후 OpenAI와의 협력 방안을 제시하는 것을 목표로 한다.

## II. 챗봇의 유래

챗봇의 유래는 1996년에 개발된 엘리자라는 프로그램으로부터 시작되었다. 엘리자는 자연어 대화를 컴퓨터와 가능하게 하는 프로그램으로써 MIT의 MAC 타임 셰어링 시스템에서 구현되었으며 사용자들이 프로그램의 언어 능력을 지속해서 개선할 수 있도록 설계되었고 사용자로부터 입력된 문장을 분석하고 적절한 응답을 생성하는 방식으로 동작함으로써 대화형 컴퓨터의 출발점이 되었다[1]. 엘리자의 독창성은 입력된 문장을 규칙적으로 분석하고 미리 학습된 패턴에 따라 응답을 생성함으로써 컴퓨터가 이해하는 것처럼 보이도록 설계된 것이다. 엘리자로부터 후속적으로 개발된 인공지능 챗봇의 발전에 중요한 기반을 마련했다.

## III. 챗봇 기술

챗봇 기술은 크게 두 가지로 구분할 수 있는데 규칙 기반과 AI 기반이다[2]. 규칙 기반은 챗봇은 질문 의도를 파악하고 사전에 정의된 키워드와 유사성 분석하여 응답을 생성한다. 이는 1차원적인 응답을 생성한다는 뜻이다. 이러한 기술은 간단하고 특정 상황에서는 유연하게 쓰일 수 있지만 복잡한 대화에서는 한계가 보이며 사용자 만족도에 부정적인 미칠 수 있다[3]. 또 다른 기술은 AI 기반 기술이다.

AI 기반 기술은 머신러닝과 자연어 처리 기술을 활용하여 자연스럽고 적응력 있는 대화를 제공하여 사용성을 향상할 수 있다. 이 부분에서 사용자가 요구하는 질문을 파악하여 개인화된 서비스를 제공하는 것에 강점이 있다. 자연어 처리는 머신러닝을 사용하여 문장의 의미와 구조를 파악하는 기술이며 챗봇의 성능을 높이는 데 중요한 역할을 한다[4].

## IV. 설문조사 결과

이번 설문조사는 OpenAI 기반 챗봇의 사용 경험과 편리성, 그리고 대학 내에서의 활용 가능성에 대한 학생들의 인식을 조사하기 위해 진행하였다. 총 70명의 학생을 대상으로 진행하였고, 설문조사 결과에 따르면 대학 내 챗봇 활용이 긍정적인 반응을 얻고 있음이 확인되었다.

그림 1, 2, 3의 결과에 따르면 대부분 학생이 OpenAI 챗봇을 사용한 경험이 있으며, 다수의 응답이 챗봇이 사용하기 편리하고 학습에 도움이 된다고 평가하였다.

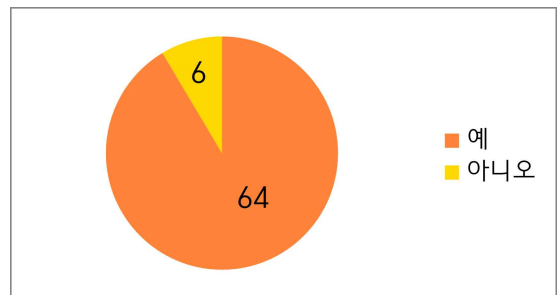


그림 1. OpenAI 기반 챗봇 사용 경험 조사  
Fig. 1. Experience using an OpenAI-based chatbot

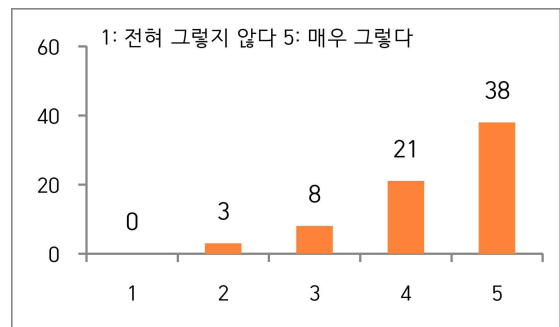


그림 2. OpenAI 기반 챗봇의 편리성 평가  
Fig. 2. Convenience of an OpenAI-based chatbot

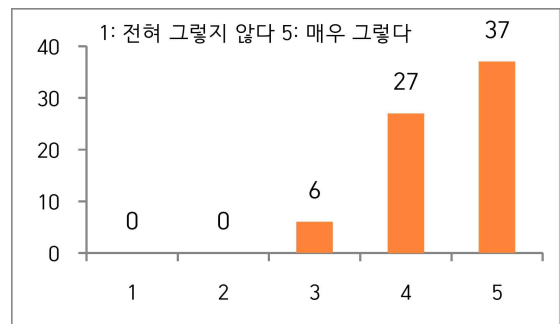


그림 3. OpenAI 기반 챗봇의 유용성 평가  
Fig. 3. Usefulness of OpenAI-based chatbots

그림 4, 5의 결과에 따르면, 학습 지원 목적으로 챗봇을 사용하는 것에 대해 58명(81.7%)의 응답자가 긍정적인 반응을 보이며 13명(18.3%)의 응답자가 중립 혹은 부정적인 의견을 내었다. 하지만, 행정 서비스의 경우에는 27명(38.5%)의 응답자가 중립적이거나 부정적인 견해를 밝히며 챗봇을 활용한 행정 서비스에 대해 소극적인 견해를 밝히고 있다.

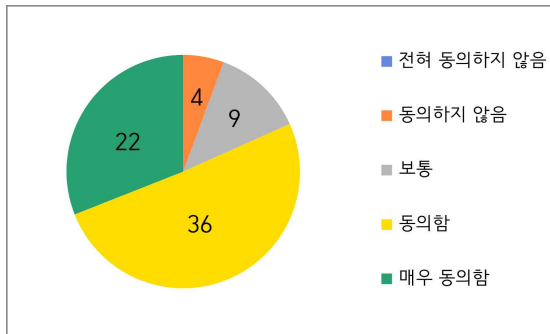


그림 4. 대학 내 학습 지원을 위한 OpenAI 챗봇 활용에 대한 학생들의 인식

Fig. 4. Students' perceptions of using OpenAI chatbots to support learning in universities

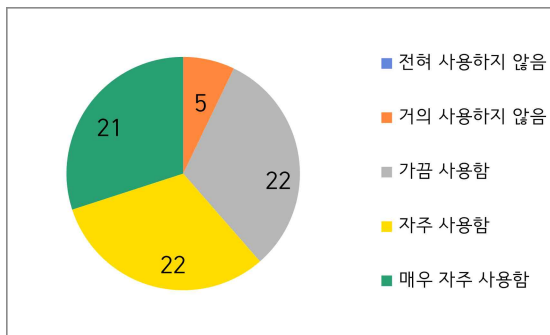


그림 5. OpenAI 챗봇을 활용한 행정 서비스 도입에 대한 학생들의 이용 의향

Fig. 5. Students' Intent to Use the Introduction of Administrative Services Using OpenAI Chatbot

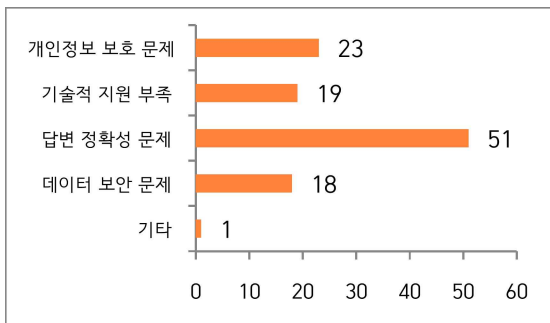


그림 6. 대학 내 OpenAI 챗봇의 사용 시 우려

Fig. 6. Concerns over the use of OpenAI chatbots in universities

그림 6의 결과에 따르면 학생들은 학습 지원 혹은 행정 서비스를 위한 대학 내 OpenAI 챗봇 서비스에 대하여 데이터 보안 문제와 개인정보 보호에 대해 우려를 나타내고 있으며 답변의 정확성 문제를 가장 많이 지적하였다. 다른 우려의 경우 AI에 대한 지나친 의존은 학습 의지와 창의성 저하를 초래할 수 있으며, 학생들의 문제 해결 능력을 떨어뜨릴 가능성을 제시하였다. 따라서 과제나 학습 활동에서 챗봇의 사용을 적절히 제한하거나, 추가 학습을 위한 지원 기능 마련이 필요하다는 의견을 제시했다. 또한 기술적 정확성 개선의 필요성도 언급되었으며, 특히 오픈AI의 답변 정확도를 지적하며 챗봇의 실효성에 의문을 제기하였고, 이를 해결하기 위한 보완책이 필요하다는 응답도 있었다. 그럼에도 불구하고, 행정 업무와 같은 반복적인 정보 제공 업무에서는 챗봇이 효율적일 것이라는 긍정적인 평가가 있었다.

## V. 결론

이번 설문조사를 통해 대학 내 챗봇에 대한 인식과 사용 의향을 파악하였다. 결과에 따르면, 많은 응답자가 학습 지원에서 챗봇을 사용하는 것에 긍정적인 반응을 보였고, 이를 통해 학습을 보조하고 필요한 정보를 효율적으로 제공할 수 있음을 확인할 수 있었다. 그리고 간단한 행정 서비스나 정보 탐색에 빠른 응답을 제공할 수 있는 부분에 긍정적인 반응을 보였다.

그러나 일부 응답자들은 학습 의지와 창의성 저하에 대한 우려를 보였으며, 챗봇에 대한 과도한 의존이 문제 해결 능력을 떨어뜨릴 수 있다는 의견도 있었다. 또한 개인정보 보호, 데이터 보안 그리고 답변의 정확성에 대해 지적하였다. 이것을 해결하기 위해 기술적인 개선과 함께 학습 지원의 범위를 적절히 조절하는 등의 보완책이 필요하다.

따라서 대학 내 OpenAI 기반 챗봇의 도입은 학습과 행정 지원의 효율성을 높일 가능성이 있지만, 이전 서술한 우려를 반영한 접근이 필요하다. 본 연구는 이러한 대학 내 챗봇 도입의 가능성을 탐색하고 학생들의 인식을 분석함으로써 향후 대학과

OpenAI 간의 협력 방안을 모색하는 데 중요한 기초 자료를 제공할 수 있을 것이다.

### 참 고 문 헌

- [1] Weizenbaum, Joseph. "ELIZA—a computer program for the study of natural language communication between man and machine", *Communications of the ACM* 9.1 (1966): 36-45.
- [2] 권오균, 신용태. (2024). AI 챗봇 시스템 주요 기술 동향 분석을 통한 민원 처리시스템 개선. *인터넷전자상거래연구*, 24(4), 49-62.
- [3] Ahrens, H., & S. W. (2020). Users' experiences with chatbots: findings from a questionnaire study. *Quality and User Experience*, 5(1), 1-18.
- [4] Zerouani, H., Haqiq, A., Bounabat, B. (2023). Chatbots Technology and its Challenges: An Overview. In: Kacprzyk, J., Ezziyyani, M., Balas, V.E. (eds) *International Conference on Advanced Intelligent Systems for Sustainable Development. AI2SD 2022. Lecture Notes in Networks and Systems*, vol 637. Springer, Cham. [https://doi.org/10.1007/978-3-031-26384-2\\_6](https://doi.org/10.1007/978-3-031-26384-2_6)
- [5] 유현미. (2024). 심리기술 훈련 상담 챗봇 설계를 위한 기초연구 : 전문가 델파이 기법. *한국콘텐츠학회논문지*, 24(9), 625-632, 10.5392/JKCA.2024.24.09.625
- [6] 박휘동, 이연준. (2024). 생성형 인공지능 챗봇 (Generative AI Chatbot)의 텍스트 움직임이 안구 운동에 미치는 영향 연구. *Archives of Design Research*, 37(4), 181-197, 10.15187/adr.2024.08.37.4.181
- [7] 류지수, 김연재, 이제환. (2024-07-11). 생성형 AI 기반 학교폭력 상담 챗봇 개발. *한국컴퓨터정보학회 학술발표논문집*, 제주.
- [8] 황기인. (2020). 한국형 심리상담 챗봇 서비스 모형 개발을 위한 기초연구 : 해외 사례분석 및 국내 선호도 조사를 중심으로 [석사학위논문, 숙명여자대학교]. <http://www-riss-kr.kims.kmu.ac.kr/link?id=T15485115>

# 스마트 캐디로봇 “쿨로” 개발

남상기\*, 한준희\*, 김성수\*, 이관형\*

## Development of Smart Caddy Robot “Coolro”

sanggi Nam\*, junhee Han\*, seongsu Kim\*, and kwanghyung Lee\*

### 요 약

본 논문에서는 자율 주행 기술과 인공지능(AI)을 결합하여 설계된 \*\*골프 캐디 로봇 "쿨로"의 설계 및 구현에 대해 논의한다. 쿨로는 골프장에서 플레이어를 따라다니며 GPS를 기반으로 실시간 기상 정보를 제공하고, MediaPipe를 활용한 골프 자세 분석 기능을 통해 사용자의 스윙 동작을 실시간으로 분석하고 교정한다. 또한, Firebase를 통해 사용자의 경기 데이터를 실시간으로 저장하고 관리하며, 모바일 애플리케이션과 연동하여 사용자가 자신의 경기 기록을 조회하고 분석할 수 있도록 한다. 본 연구는 쿨로의 자율 주행, AI 기반 동작 인식, 실시간 데이터 처리 및 저장 기능을 소개하고, 이를 통해 골프 경기를 효율적으로 보조할 수 있는 가능성을 탐구한다. 본 논문은 쿨로가 골프 경기에서의 사용자 경험을 크게 향상시키며, 향후 스포츠 산업에서 자율 주행 로봇의 응용 가능성을 넓힐 수 있음을 제안한다.

### Abstract

In this paper, we discuss the design and implementation of “Coolro, a golf caddy robot” designed by combining autonomous driving technology and artificial intelligence (AI). Coolro follows players on a golf course, provides real-time weather information based on GPS, and analyzes and corrects the user's swing motion in real time through the golf posture analysis function using MediaPipe. In addition, it stores and manages the user's game data in real time through Firebase, and allows the user to view and analyze their game records by linking with a mobile application. This study introduces Kullo's autonomous driving, AI-based motion recognition, and real-time data processing and storage functions, and explores the possibility of efficiently assisting golf games. This paper suggests that Coolro can significantly improve the user experience in golf games and expand the application potential of autonomous driving robots in the sports industry in the future.

### Key words

Golf Caddy Robot, Autonomous Driving, MediaPipe, Weather API, Firebase, AI Posture Analysis

### 1. 서 론

골프는 전 세계적으로 널리 사랑받는 스포츠로, 기술적 숙련도와 전략적 사고가 결합된 경기이다.

특히, 골프 경기는 플레이어가 각 홀을 공략하면서 클럽 선택, 바람의 방향, 경기장의 지형 등을 고려해야 하며, 이를 위해 캐디의 역할이 매우 중요하다. 캐디는 플레이어의 클럽을 운반하고, 경기 전략

\*대진대학교, hinoonyaso@naver.com, \*대진대학교, wogkswns@naver.com, \*대진대학교, tjdtntjgh19@naver.com

을 제공하며, 다양한 변수들을 고려해 플레이어를 보조한다. 하지만 캐디의 고용 비용이나 인력 부족 문제 등으로 인해 많은 플레이어들이 캐디의 도움 없이 경기를 진행해야 하는 경우가 늘어나고 있다.

이러한 문제를 해결하기 위한 대안으로 골프 캐디 로봇이 주목받고 있다. 특히, 기술의 발전에 따라 자율 주행 기술과 인공지능(AI)을 결합한 로봇들이 등장하면서, 경기 중 자동으로 플레이어를 따라다니고, 경기 상황에 필요한 다양한 정보를 제공할 수 있는 로봇들이 개발되고 있다. 본 논문에서는 골프 캐디 로봇 "쿨로(Coolro)"의 설계와 구현에 대해 논의하고자 합니다. 쿨로는 자율 주행 기술을 기반으로 플레이어를 따라다니며, GPS를 이용한 실시간 기상 정보 제공, AI 기반 골프 자세 분석, 골프 스코어 기록 및 관리 등 다양한 기능을 통해 플레이어의 경기 경험을 극대화하는 것을 목표로 하고 있다.

본 연구의 목적은 쿨로의 기술적 구성 요소를 분석하고, 각 기능의 동작 원리 및 구현 방법을 제시하며, 실용성을 평가하는 것이다. 쿨로의 도입으로 골프 경기가 어떻게 혁신될 수 있는지를 논의하고, 향후 스포츠 산업에서의 자율 주행 및 인공지능 기술의 응용 가능성에 대해 고찰하고자 한다.

## II. 본 론

쿨로의 핵심 기능 중 하나는 “자율 추종 주행 (Autonomous Follow Driving)”이다. 이 기능은 플레이어의 위치를 실시간으로 인식하고, 자율적으로 경로를 계산하여 플레이어를 따라다니는 것이다. 이를 위해 카메라와 TFT 센서를 사용하여 사람을 인식하고, 주변 환경을 파악한다.

쿨로는 MediaPipe를 사용해 플레이어의 골프 스윙동작을 실시간으로 인식하고 분석한다. MediaPipe는 구글에서 개발한 오픈소스 라이브러리로, 인간의 관절을 인식하고 자세를 분석하는 데 최적화된 도구다. 카메라를 통해 플레이어의 동작을 캡처하고, MediaPipe의 Pose Estimation 기능을 통해 33개의 주요 관절 포인트를 추적하여 스윙 동작을 분석한다.

쿨로는 GPS를 이용해 사용자의 현재 위치 정보

를 수신하고, 이를 기반으로 OpenWeatherMap API와 같은 외부 API에서 실시간 기상 정보를 수신한다. 이 데이터는 Flutter 기반의 모바일 애플리케이션을 통해 사용자가 쉽게 확인할 수 있으며, 기온, 바람의 방향, 풍속 등 경기 전략에 중요한 기상 정보를 직관적으로 제공한다. 기상 데이터는 매 분마다 업데이트되며, 네트워크 상태에 따라 효율적으로 캐싱하여 제공한다.

쿨로는 Firebase를 통해 사용자 데이터를 관리한다. Firebase Realtime Database를 이용해 사용자의 골프 스코어, 자세 분석 데이터, 녹화된 비디오 등을 실시간으로 저장하고, 이를 기반으로 경기가 끝난 후에도 자신의 데이터를 분석할 수 있도록 지원한다. 또한 Firebase Authentication을 통해 사용자는 자신의 계정에 로그인하여 데이터를 안전하게 저장하고 조회할 수 있다.

저장된 데이터는 추후 분석을 위해 모바일 애플리케이션에서 조회할 수 있으며, 이를 통해 사용자는 자신의 경기 기록을 추적하고, 향후 경기를 위한 전략을 개선할 수 있다.

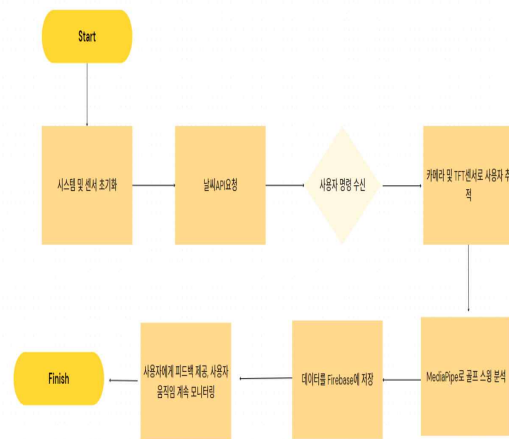


그림 1. Coolro의 알고리즘 동작 과정

쿨로가 동작하면 센서들을 초기화 한 후에 날씨 API를 요청한다. 다음으로는 사용자의 명령을 수신하고, 카메라 및 TFT센서를 이용해서 사용자를 추적하여 따라간다. 사용자가 스윙 후에 MediaPipe로 스윙 자세를 분석한 다음 그 데이터를 Firebase에 저장한 후 사용자에게 피드백을 제공해준다.

### III. 결 론

쿨로는 캐디 역할을 대체하거나 보완할 수 있는 혁신적인 로봇으로, 골프 산업에서의 기술적 진보를 대표한다. 특히, AI와 자율 주행 기술의 결합을 통해 플레이어의 경기 경험을 혁신적으로 향상시키며, 향후 스포츠 산업에서 로봇 기술의 응용 가능성을 크게 확장할 수 있다.

### 참 고 문 헌

- [1] Zhang, Z., Sun, H., & Zhang, G. (2020). A review of autonomous driving technologies for intelligent vehicles. *Journal of Automotive Engineering*, 34(2), 1-14.
- [2] King, M., & Wang, Y. (2018). AI-powered sports analysis: Applications in golf swing analysis. *Sports Technology Journal*, 19(4), 122-135.
- [3] OpenCV Team. (2021). OpenCV Documentation.
- [4] MediaPipe Team. (2020). MediaPipe: A framework for building multimodal applied machine learning pipelines.
- [5] Google Developers. (2021). Flutter Documentation.
- [6] Firebase Team. (2020). Firebase Documentation: Realtime Database.
- [7] Kim, J., & Lee, H. (2019). Autonomous follow robots for outdoor applications. *International Journal of Robotics Research*, 27(6), 235-250.

# 블랙박스 영상 기반 명암 변화에 강인한 방지턱 도색 벗겨짐 정도 예측

손준형\*, 김성준\*\*, 김대환\*\*\*

## Prediction of Speed Bump Paint Wear Level Invariant to Lighting Changes Based on Dashcam Video

JunHyeong Son\*, SeongJun Kim\*\*, and Daehwan Kim\*\*\*

### 요 약

과속방지턱의 유지관리는 도로 안전에 중요하지만, 현행 점검 방식은 인력 소모가 크고 효과적인 관리가 어렵다. 본 연구에서는 블랙박스 영상을 활용하여 명암 변화에 강인하게 과속방지턱의 도색 벗겨짐 정도를 자동으로 예측하는 시스템을 제안한다. 제안한 시스템은 외부 환경에서 발생하는 명암 변화와 조도 차이에 의한 도색 벗겨짐 인식의 어려움을 해결하기 위해 색상전이 및 색조 보정 기법을 적용하여 영상을 정규화한다. 또한 딥러닝 기법을 이용하여 도색 상태를 분석하고 벗겨짐 정도를 예측하도록 한다. 다양한 조건 속 과속방지턱 이미지를 실험하여 도색 벗겨짐 정도를 육안으로 확인 가능하도록 제시한다.

### Abstract

The maintenance of speed bumps is crucial for road safety; however, current inspection methods are labor-intensive and lack efficiency in effective management. This study proposes an automated system that leverages dashboard camera footage to robustly predict the degree of paint wear on speed bumps, even under varying illumination conditions. The proposed system normalizes the image by applying color transfer and color tone correction techniques to solve the difficulty of recognizing paint peeling due to the difference in contrast and illuminance occurring in the external environment. Additionally, deep learning techniques are employed to analyze the paint condition and predict the degree of wear. We experiment with speed bumps images under various conditions and present them so that the degree of peeling of the paint can be visually confirmed.

### Key words

Speed bump, Paint wear prediction, dashCam video, Road safety, Addressing manpower issues

---

\* 울산대학교 IT융합전공 thswngud123@naver.com

\*\* 울산대학교 IT융합전공 wns9001@naver.com

\*\*\* 울산대학교 IT융합전공 daehwankim@ulsan.ac.kr

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학 지원을 받아 수행되었음.  
(2024-0-00025)

## 1. 서 론

도로와 교통 시설물의 유지보수는 도로 안전에 중요한 요소이다. 실제로 도색이 벗겨진 과속방지턱이나 도로 파손 등은 많은 사고의 원인이 되고 있다. 특히 뉴스 기사와 보도 자료에 따르면, 과속방지턱의 도색이 벗겨지거나 도로 결함으로 인해 발생하는 사고가 꾸준히 발생하고 있는 상황이다.

이러한 문제를 해결하기 위해서는 체계적이고 효율적인 도로 관리 시스템이 필요하지만, 정부기관의 도로 및 교통 시설물 관리 공무원 수가 매우 부족하여 유지보수에 어려움을 겪고 있다. 따라서 소수의 인원으로 과속방지턱의 유지보수를 하기 효율적으로 하기 위한 자동화 방법이 필요하다.

본 논문에서는 블랙박스 영상을 활용하여 과속방지턱의 도색 상태를 자동으로 분석하고, 도색 벗겨짐 정도를 예측하는 방법을 제안한다. 블랙박스 영상은 도로 데이터를 실시간으로 손쉽게 수집할 수 있어 효율적인 도로 관리에 유용하며, 이를 통해 유지보수 작업의 정확성과 효율성을 높일 수 있다.

하지만 블랙박스 영상은 외부의 다양한 명암 변화와 조도 차이가 심한 특징이 있다. 이는 도색 벗겨짐 정도를 정확하게 인식하는데 어려움을 제공한다. 이를 해결하기 위해 색상전이 및 색조 보정 기법을 적용하여 영상을 정규화하고, 이러한 외부 환경 변화에도 강인한 예측 방법을 제안한다.

## II. 제안 방법

명암 변화에 강인한 과속방지턱 도색 벗겨짐 정도 예측 시스템 구조도는 그림 1과 같다. 도색 벗겨짐 예측 시스템은 블랙박스 영상을 입력으로 크게 3단계로 이루어져 있다. 과속 방지턱 재구성, 이미지 명암 보정 단계와 도색 벗겨짐 정도 예측 단계이다.

첫 번째, 과속방지턱 이미지 재구성은 블랙박스 영상 속 과속방지턱이 검출된 여러 프레임을 하나의 과속방지턱 이미지로 통합하는 단계이다.

두 번째, 이미지 명암 보정 단계는 과속방지턱 이미지 재구성 단계에서 만들어진 이미지를 색상전이와 색조보정을 하여 명암에 강인하도록 만드는

단계이다.

마지막으로 도색 벗겨짐 정도 예측 단계는 명암에 강인해진 과속방지턱 이미지를 클러스터링을 통해 색상을 분류한 후 색상 데이터를 학습하여 도색 벗겨짐 정도를 예측하는 단계이다.

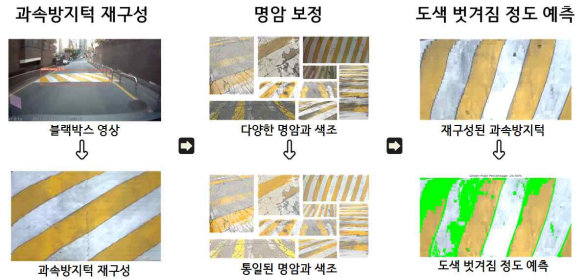


그림 1. 명암 변화에 강인한 도색 벗겨짐 예측 구조도

### 2.1 과속방지턱 이미지 재구성

과속방지턱 이미지 재구성 단계는 블랙박스 영상에서 과속방지턱 이미지를 사용자가 보다 쉽게 육안으로 확인할 수 있도록 하며, 한 프레임만으로는 과속방지턱 전체를 파악하기 어려운 문제를 해결하고자 한다. 이를 위해 과속방지턱의 그룹화, 왜곡 보정 및 원근 변환, 이미지 스티칭, 배경 제거 단계를 순차적으로 수행한다.

#### 2.1.1 과속 방지턱 그룹화

과속방지턱은 요철 구조이기 때문에 블랙박스 영상의 단일 프레임으로는 과속방지턱의 전체부분을 확인하기 어렵다. 따라서 단일 프레임이 아닌 여러 프레임을 하나의 이미지로 통합하는 과정이 필요하다. 블랙박스 영상 내에서 과속 방지턱이 검출된 프레임들과 마지막 프레임을 기준으로 후속 7프레임을 추가로 저장한다. 이후, 연속된 프레임 번호를 가진 프레임들을 각각의 과속방지턱 그룹으로 그룹화하는 과정을 거친다.



그림 2. 과속방지턱 그룹화



### 2.1.2 왜곡보정 및 원근변환

블랙박스는 넓은 시야각을 확보하기 위해 광각 렌즈를 사용하므로, 접선 왜곡 및 방사 왜곡이 발생한다. 이를 해결하기 위해 영상 시각화 과정을 적용한다. 블랙박스 렌즈의 170도 광각으로 인한 왜곡을 보정하기 위해 K1 값을 -0.55로 설정한다.

$$x_{corrected} = x * (1 + k1 * r^2 + k2 * r^4 + k3 * r^6) + 2 * p1 * x * y + p2 * (r^2 + 2 * x^2)$$

$$y_{corrected} = y * (1 + k1 * r^2 + k2 * r^4 + k3 * r^6) + p1 * (r^2 + 2 * y^2) + 2 * p2 * x * y$$

수식 1. 왜곡보정을 위한 수식

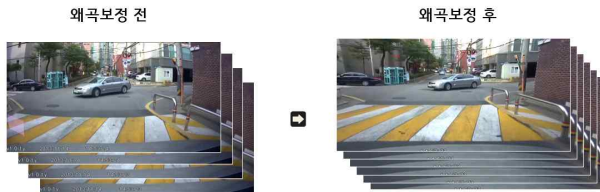


그림 3. 왜곡보정 전, 후

이후, 과속방지턱의 도색 상태를 일관성 있게 판단하기 위해 상면에서 관찰한 이미지처럼 만들 필요가 있기 때문에 원근 변환을 적용한다.

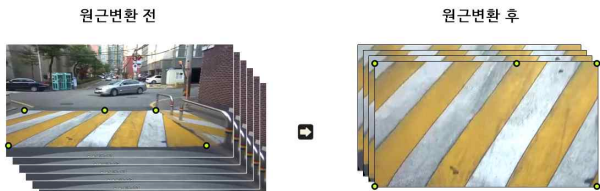


그림 4. 원근변환 전, 후

### 2.1.3 과속방지턱 이미지 스티칭

왜곡 보정과 원근 변환이 적용된 과속방지턱 그룹들을 각각 하나의 이미지로 스티칭하기 위해 이미지의 각 프레임에서 프레임의 하단을 ROI로 설정하여 크롭한 후, 세로 방향으로 스티칭하였다.

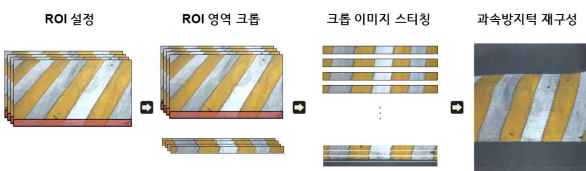


그림 5. 과속 방지턱 이미지 스티칭 과정

### 2.1.4 재구성 이미지 배경 제거

재구성된 과속 방지턱 이미지에서 과속 방지턱 부분을 제외한 나머지 영역은 도색 벗겨짐 예측에 필요하지 않으므로 배경을 제거한다. 캐니 엣지 (Canny edge) 검출 기법 [3] 을 활용하여 기울기 값이 '0' 에 가까운 엣지만을 탐지하였으며, 최종적으로 Max X 및 Max Y 값을 남겨 과속 방지턱의 경계를 추출하였다.



그림 6. 배경제거 과정

## 2.2 이미지 명암 보정

재구성된 과속방지턱 이미지마다 명암 차이에 따른 색상 변화가 발생하여, 도색 벗겨짐 판단에 오류를 발생시킬 수 있다. 따라서 이를 해결하기 위해 색상 전이와 색조 보정 과정을 거쳐 명암에 강인하도록 한다.

### 2.2.1 색상 전이

조도와 같은 외부 환경의 영향으로 블랙박스 영상의 명암이 달라지기 때문에 동일한 과속 방지턱 이미지도 다른 색상으로 판단될 수 있다. 따라서 일관성 있는 판단을 위해 이미지가 명암 변화에 강인하도록 만드는 과정이 필요하다. 각 이미지를 LAB 값으로 변환한 다음 타겟이미지와 소스이미지를 설정한 후 타겟이미지의 LAB 분포와 소스이미지의 LAB 분포를 같게 함으로써 해결하였다[4].

두 이미지를 LAB 색상 공간으로 변환한 후, 각 채널(L, A, B)에 대해 평균과 표준편차를 계산한다. 이후 소스 이미지의 각 픽셀 값에서 소스 이미지의 평균을 빼고, 타겟 이미지의 표준편차로 나눈 후 타겟 이미지의 평균을 더하여 픽셀 값을 조정한다. 이 방법을 통해 소스 이미지의 색상 분포를 타겟 이미지와 일치시킴으로써, 다양한 조명 조건에서 이미지 간의 색조와 명암을 통일할 수 있다.

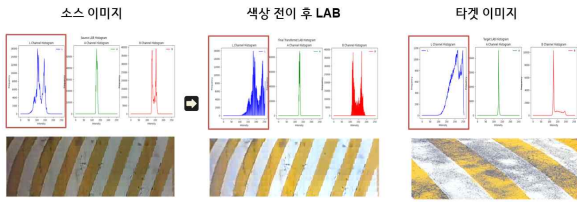


그림 7. 색상 전이에 따른 lab값 변화 비교

### 2.2.2 색조 보정

하나의 이미지 내에서도 명암 차이가 발생하기 때문에 동일한 색상이 다른 색상으로 분류되는 문제가 생긴다. 따라서 추가적인 색조 보정 단계를 수행한다.

K-Means 클러스터링 기법 [5]을 활용하여  $K=2$ 로 설정한 후, 이미지를 주황색과 흰색으로 분류하였다. 이후, 이미지를 HSV 색상 공간으로 변환하여 H 값(색상 값)을 통일하는 과정을 진행하였다.

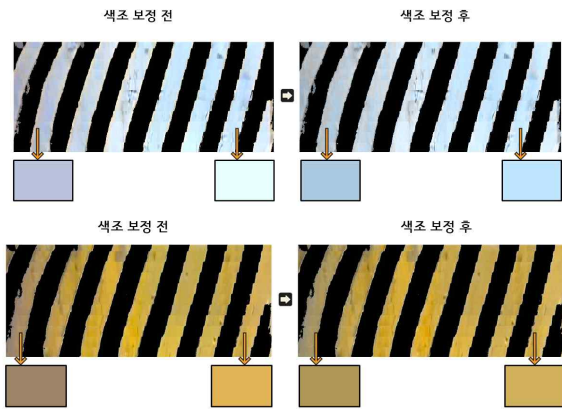


그림 8. 색조 보정 전, 후

## 2.3 도색 벗겨짐 정도 예측

과속방지턱 이미지를 색상별로 세그멘테이션하여 색상 데이터를 생성한 후, 해당 색상이 도색이 벗겨진 색상인지 여부를 판단하고 최종적으로 과속방지턱 도색의 벗겨짐을 평가한다.

### 2.3.1 색상 데이터 생성

과속방지턱 이미지는 주황색과 흰색으로 구성되어 있으며, 각 색상에서 도색 벗겨짐의 상태가 다르게 나타난다. 이에 따라, White, White\_dirty, Orange,

Orange\_dirty의 4개의 색상 라벨을 기준으로 한 이미지를 세그멘테이션하여 여러 색상 이미지로 분류한다. K-Means 클러스터링 기법을 사용해서 도색 벗겨짐을 구분할 수 있도록 한 이미지를 충분한 K개의 색상 이미지로 세분화 한다. 세분화된 세그멘테이션 이미지의 색상 픽셀을 랜덤으로 배열하여 사각형 형태의 색상 이미지를 생성한다.

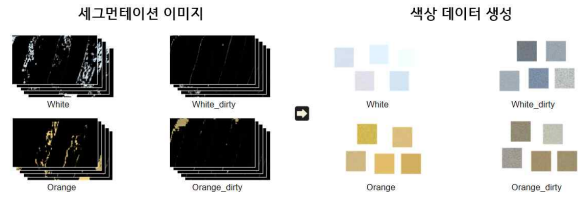


그림 9. 색상 데이터 생성

### 2.3.2 색상 데이터 학습

생성된 색상 데이터를 White, White\_dirty, Orange, Orange\_dirty의 네 가지 색상 라벨로 학습 하기 위해, 원본 이미지와 세그멘테이션된 색상 이미지를 육안 검토를 통해 직접 비교하고 네 가지 색상 라벨로 분류한다. 이후 분류된 색상 데이터를 바탕으로 학습을 진행한다.

## III. 실험

블랙박스 영상 속 과속방지턱을 검출하기 위해 YOLOv5 [2] 를 사용한다. 과속방지턱 이미지 데이터 셋은 8:2 비율로 Train 셋과 Validation 셋으로 분류한다. Roboflow 출처 데이터를 사용한다. 에폭은 30, 배치 사이즈는 32로 설정하여 학습한다.

표 1. 과속방지턱 이미지 데이터 셋

	Training	Validation	Total
Speed bump	4,303	1,076	5,379

색상 데이터 학습 과정에서는 DNN을 사용한다. 각 라벨에 해당하는 데이터셋은 8:2 비율로 Train 셋과 Validation 셋으로 분류한다. 사각형 색상 데이터의 크기는 128\*128로 설정한다. 에폭은 50, 배치 사이즈는 32로 설정하여 학습한다.

표 2. 색상 데이터 셋 (size : 128\*128)

	White	White_dirty	Orange	Orange_dirty
Color data	601	565	631	598

평가할 입력 영상을 색상 데이터 생성 과정과 동일하게 처리하여 색상 데이터를 생성한 후, 학습된 도색 벗겨짐 예측 모델로 평가한다. 모델의 평가 예시는 그림 10과 같다.

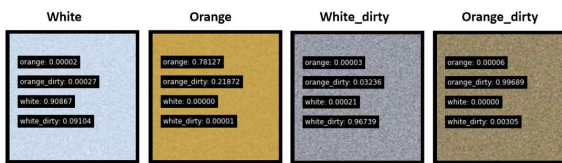


그림 10. 도색 벗겨짐 예측 모델을 통한 색상 데이터 분류

White\_dirty와 Orange\_dirty로 분류된 색상 데이터를 가진 세그먼테이션 이미지는 도색이 벗겨진 부분으로 간주하여 녹색 픽셀로 표시하며, 최종 결과는 그림 11, 그림 12, 그림 13과 같다.

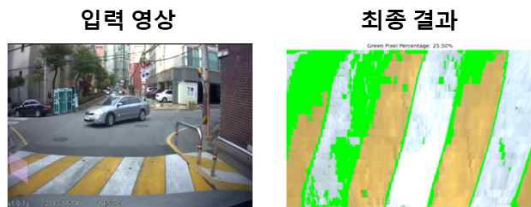


그림 11. 평범한 과속방지턱 이미지



그림 12. 도색 벗겨짐이 심한 과속방지턱 이미지

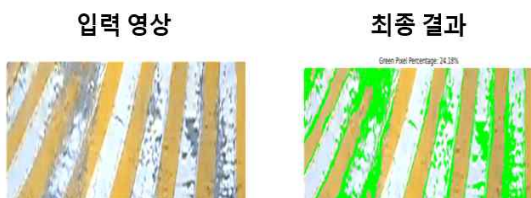


그림 13. 블랙박스 영상이 아닌 일반 과속방지턱 이미지

## IV. 결론

다양한 명암 조건에서도 블랙박스 영상을 통해 과속방지턱 도색의 벗겨짐 정도를 효과적으로 예측할 수 있음을 육안으로 확인하였다. 이러한 기술은 도로 시설물 관리의 자동화와 효율화에 중요한 도구가 될 수 있으며, 도색 벗겨짐을 조기에 감지함으로써 사고 예방과 유지보수 비용 절감에 기여할 것으로 기대된다.

향후 연구에서는 과속방지턱 재구성 과정에서 스티칭을 더욱 자연스럽게 개선하기 위한 다양한 방법을 시도할 계획이다. 또한, 블랙박스 영상뿐만 아니라 일반 과속방지턱 이미지에서도 우수한 성능을 보이는 만큼, 다양한 영상 기기로 확장하여 적용할 예정이다.

## 참고 문헌

- [1] <https://roboflow.com>.
- [2] <https://github.com/ultralytics/yolov5>.
- [3] Canny, J. (1986). A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679-698.
- [4] Reinhard, E., Adhikhmin, M., Gooch, B., & Shirley, P. (2001). Color transfer between images. IEEE Computer Graphics and Applications, 21(5), 34-41.
- [5] MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, 281-297.
- [6] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436-444.

# 자율주행 차량을 위한 실시간 보행자 행동 예측 시스템 개발

고보경\*, 나관상\*\*

## Development of a Real-Time Pedestrian Behavior Prediction System for Autonomous Vehicles

Bo-gyeong Ko\*, Kwan-sang Na\*\*

### 요 약

자율주행 기술은 교통사고 감소와 편의성 향상 등 다양한 이점을 제공하는 기술로 빠르게 발전하고 있다. 특히, 보행자와의 안전한 상호작용은 자율주행 기술 상용화에 중요한 요소다. 기존 보행자 인식 시스템은 보행자의 행동을 예측하지 못해, 자율주행 차량이 즉각적으로 대응하기 어려운 한계가 있다. 본 연구는 보행자의 자세와 움직임을 실시간으로 분석하여 행동을 예측함으로써 자율주행 차량이 보행자의 의도를 사전에 파악하고 선제적으로 대응할 수 있는 시스템을 제안한다. 또한 이 시스템은 MS COCO와 PIE 데이터셋을 사용하여 다양한 보행자 행동을 학습했으며, LSTM과 MobileNetV2를 결합한 경량 모델을 통해 자율주행 차량의 제한된 하드웨어에서 효율적으로 작동한다. CUDA 병렬 처리와 메모리 최적화를 통해 실시간 성능을 강화했으며, 실험 결과 다양한 도심 환경에서 높은 정확도로 보행자 행동을 예측함을 확인했다.

### Abstract

Autonomous driving technology is rapidly advancing, providing various benefits such as reducing traffic accidents and enhancing convenience. Safe interaction with pedestrians is a crucial factor for the commercialization of autonomous driving technology. Traditional pedestrian recognition systems struggle to respond quickly because they cannot predict pedestrian actions. This study proposes a system that analyzes pedestrians' posture and movements in real-time to predict their behavior, allowing autonomous vehicles to proactively respond to pedestrians' intentions. The system utilizes the MS COCO and PIE datasets to learn diverse pedestrian behaviors and combines LSTM with a lightweight MobileNetV2 model to operate efficiently within the limited hardware of autonomous vehicles. Through CUDA parallel processing and memory optimization, the system achieves enhanced real-time performance, and experimental results confirm that it accurately predicts pedestrian behavior in various urban environments.

### Key words

Autonomous Driving, Pedestrian Interaction, Behavior Prediction, Real-Time Analysis, LSTM, MobileNetV2

---

\*경기대학교 AI컴퓨터공학부 인공지능전공, gobk0507@gmail.com

\*\*경기대학교 AI컴퓨터공학부, ksna001@kyonggi.ac.kr

※ “본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음” (2021-0-01393)

## 1. 서 론

자율주행 기술은 교통사고 감소, 편의성 향상, 에너지 절약 등 다양한 사회적 이점을 제공할 수 있는 기술로 주목받고 있으며, 빠르게 발전하고 있다 [1]. 자율주행 차량의 핵심 기능 중 하나는 주변 환경을 인식하고 안전한 주행 경로를 계획하는 능력이다. 이러한 인식 기술 중에서도 특히 도로 위 보행자와의 상호작용은 자율주행 기술이 실질적으로 상용화되는 데에 중요한 요소로 작용한다.

보행자는 도로 위의 예측하기 어려운 요소 중 하나로, 보행자의 의도나 행동을 예측하지 못할 경우 자율주행 차량이 즉각적으로 대응하지 못해 사고로 이어질 가능성이 존재한다. 기존의 보행자 인식 시스템은 보행자가 차량 주변에 있는지 여부만을 감지하여 경고를 주거나 차량을 멈추게 하는 방식으로 작동한다[2]. 이러한 접근 방식은 보행자의 행동을 예측하지 못하고 단순히 인식에만 의존하므로, 보행자의 행동에 미리 대응하지 못한다는 한계를 가진다.

따라서, 보행자의 의도를 예측하여 차량이 보다 선제적으로 대응할 수 있는 시스템의 개발이 필수적이다. 예를 들어, 보행자가 길을 건너려는 의도를 가진 경우 자율주행 차량이 해당 보행자와의 안전 거리를 유지할 수 있도록 경로를 수정하거나 속도를 조절하는 것이 필요하다. 본 연구는 이러한 요구를 충족시키기 위해 보행자의 자세와 움직임을 실시간으로 분석하고, 보행자의 행동을 예측하여 차량이 빠르게 대응할 수 있도록 하는 예측 시스템을 개발하는 데에 목적을 둔다. 특히, 이 시스템은 자율주행 차량의 제한된 하드웨어 환경에서도 효과적으로 작동할 수 있도록 경량화된 모델을 사용하며, 복잡한 도심 환경에서도 높은 정확도로 보행자의 행동을 예측할 수 있도록 설계되었다.

본 연구의 주요 기여는 자율주행 차량의 안전성을 강화하는 예측 기반 대응 시스템을 제안하는 것이다. 이를 통해 자율주행 차량은 보행자의 행동을 보다 능동적으로 파악하여 사고 위험을 줄일 수 있으며, 안전하고 효율적인 주행을 가능하게 한다. 본 논문에서는 이러한 시스템의 설계, 구현, 및 평가

과정을 자세히 다루며, 연구의 효용성과 실용성을 입증하고자 한다.

## II. 알고리즘 설명 및 구현 방법

본 연구에서는 보행자의 행동을 예측하기 위해 딥러닝 기반 모델을 활용하며, 각 주요 단계는 다음과 같이 구성된다.

### 2.1 데이터 수집 및 전처리

보행자의 움직임 데이터의 효과적인 수집과 전처리를 위해 MS COCO 데이터셋과 PIE(Pedestrian Intention Estimation) 데이터셋을 사용하였다. 이 데이터셋은 다양한 자세와 행동을 포함하며, 도심 환경에서의 보행자 행동을 학습하는 데 적합하다[3]. 데이터 수집 과정에서는 보행자의 움직임을 다양한 시점에서 캡처하여 시간에 따른 행동 변화를 반영할 수 있도록 하였다. 전처리 단계에서는 보행자의 신체 부위 좌표와 자세 특징을 추출하고, 데이터의 다양성을 확보하기 위해 데이터 증강 기법(예: 회전, 크기 조절, 잡음 추가 등)을 적용하였다. 이러한 전처리 과정은 모델이 다양한 상황에서 보행자의 행동을 정확하게 인식하고 예측할 수 있도록 돕는다.



그림 1. MS COCO dataset의 샘플 이미지

Fig. 1. Samples of annotated images in the MS COCO dataset

### 2.2 행동 예측 모델

보행자 행동 예측을 위해 Long Short-Term Memory(LSTM) 네트워크를 기반으로 한 Recurrent Neural Network(RNN) 모델을 사용하였다. LSTM은 시간에 따른 순차적 데이터를 효과적으로 처리할 수 있어, 보행자의 과거 행동 정보를 바탕으로 향후 행동을 예측하는 데 유리하다[4].

제안된 LSTM 모델의 구조는 다음과 같다:

- 입력층: 보행자의 신체 부위 좌표와 자세 특징이 벡터 형태로 입력된다.

- 은닉층: 두 개의 LSTM 층으로 구성되며, 시간 축상의 데이터를 학습하여 행동 예측에 필요한 정보를 추출한다.
- 출력층: 소프트맥스 활성화 함수를 사용하여 보행자의 다음 행동(예: 정지, 횡단)을 예측하는 확률 값을 출력한다.

이 모델을 통해 보행자의 자세와 움직임 데이터를 실시간으로 분석하고, 보행자가 취할 행동을 예측할 수 있다.

### 2.3 모델 경량화 및 최적화

자율주행 차량의 제한된 하드웨어 자원에서 원활하게 작동할 수 있도록, MobileNetV2와 같은 경량화된 모델을 LSTM 모델과 결합하였다. MobileNetV2는 딥러닝 모델의 파라미터 수를 줄여 모델의 연산량을 낮추는 한편, 높은 예측 정확도를 유지할 수 있도록 설계되었다[5]. 본 연구에서는 MobileNetV2의 구조를 LSTM 모델과 통합하여, 실시간 예측 성능을 강화하였다. 이를 통해 자율주행 차량의 GPU 자원에서 효율적으로 작동할 수 있는 경량 모델을 구축하였다.

## III. 실시간 처리 및 시스템 최적화

보행자의 행동을 실시간으로 예측하고, 자율주행 차량이 즉각적으로 반응할 수 있도록 CUDA 병렬 처리와 메모리 최적화 기법을 적용하였다. CUDA 병렬 처리는 모델의 연산 지연을 줄이고, 실시간 처리가 가능하도록 지원한다. 메모리 최적화 기법을 통해 GPU 자원 사용을 최소화하며, 모델의 연산 효율을 극대화하였다. 실험 결과, 제안된 모델은 초당 30프레임의 속도로 실시간 처리가 가능하며, 자율주행 차량의 실시간 응답 요구 사항을 만족하였다.

## IV. 결론

본 연구는 자율주행 차량의 안전성을 향상시키기 위한 실시간 보행자 행동 예측 시스템을 제안하였다. 본 시스템은 보행자의 자세와 움직임을 실시간으로 분석하여 행동을 예측함으로써, 자율주행 차량

이 보행자의 의도를 파악하고 사전에 대응할 수 있도록 설계되었다. 제안된 시스템은 기존 보행자 인식 시스템이 단순히 인식 후 반응하는 방식에서 한 단계 발전하여, 보행자의 행동을 예측함으로써 자율주행 차량이 더욱 능동적으로 반응할 수 있는 기반을 제공한다.

또한, 본 연구에서는 자율주행 차량의 하드웨어 제약 조건을 고려하여 경량화된 MobileNetV2-LSTM 모델을 도입하였으며, 이를 통해 자율주행 환경에서의 실시간 처리가 가능하도록 최적화하였다. 실험 결과, 제안된 시스템은 다양한 도심 환경에서 보행자의 행동을 높은 정확도로 예측할 수 있음을 확인하였으며, 실시간 처리 성능 또한 자율주행 차량의 응답 요구 사항을 충족하였다. 이러한 성과는 본 시스템이 자율주행 차량의 안전성을 강화하는 데 실질적인 기여를 할 수 있음을 시사한다.

향후 연구에서는 더욱 다양한 도로 환경과 보행자 행동 패턴을 반영한 데이터셋을 구축하여 모델의 예측 정확도를 높이는 것이 필요하다. 또한, 제안된 시스템의 성능을 복잡한 도심 환경뿐만 아니라, 다양한 기후 및 조명 조건에서 검증함으로써 실용성을 더욱 강화할 계획이다. 더불어, 보행자뿐만 아니라 차량, 자전거 등 도로 위 다양한 객체와의 상호작용을 고려한 확장된 예측 시스템 개발을 통해 자율주행 차량의 안전성과 신뢰성을 한층 높이는 것이 목표이다.

이와 같은 연구는 자율주행 기술의 상용화에 필수적인 안전 요소로 작용할 것이며, 자율주행 차량이 보행자와 안전하게 상호작용할 수 있는 환경을 구축하는 데 기여할 것이다.

## 참 고 문 헌

- [1] 임현국, "자율주행 차량 영상 기반 객체 인식 인공지능 기술 현황", 한국정보통신학회논문지, vol. 25, no.8, pp. 1117-1123, 2021.
- [2] Q. Zhu, M.-C. Yeh, K.-T. Cheng and S. Avidan. "Fast human detection using a cascade of histograms of oriented gradients", Proc. CVPR,

October 2006.

- [3] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár. "Microsoft COCO: Common Objects in Context", May 2014.
- [4] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory", Neural Computation, vol. 9, no. 8, November 1997.
- [5] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov and Liang-Chieh Chen. "MobileNetV2: Inverted Residuals and Linear Bottlenecks", Proc. CVPR, pp. 4510-4520, Jan 2018.

# 802.11 be 우선순위 기반 Multi-RU 할당에서 기아 문제 해결을 위한 유동적 자원 범위 제한 할당 기법

노현우\*, 김선명\*\*

## A Dynamic Resource Range Allocation Scheme for Resolving Starvation in Priority-Based Multi-RU Allocation in 802.11be

Hyeonu Roh\* and Sunmyeng Kim\*\*

### 요약

802.11 be 표준에서는 Multiple-Resource Unit(MRU)를 허용하는 시도를 하였다. Quality of Service(QoS)를 제공하기 위한 우선순위에 따른 MRU 할당은 각 STA(Station)이 보내는 트래픽에 대해 요구 속도에 대한 만족도와 트래픽 유형에 따른 우선순위를 바탕으로 MRU를 할당한다. 이 방법은 우선순위가 낮은 STA이 자원을 지속적으로 할당 받지 못하여 전송 자체를 시도하지 못하는 기아 문제가 발생한다. 본 논문에서는 트래픽 유형에 따른 우선순위 기반으로 총 자원에서 자원의 할당 가능 범위를 할당하며 만족도에 따라 트래픽 유형별 우선순위에 따른 자원 사용 가능 범위를 유동적으로 움직여 기아 문제를 해결하는 기법을 제안한다. 제안된 방법은 일정 만족도 수준 이상인 유형에서 최소 만족도를 충족시키지 못하는 유형에게 사용 가능 자원 범위를 양도하는 방식으로 작동되기에 기존 우선순위 방식의 QoS를 지킬 수 있으며 기아 문제에서 기존의 방법보다 개선된다.

### Abstract

The 802.11be standard has made an attempt to allow Multiple-Resource Unit (MRU). The allocation of Multiple RUs based on priority aims to provide Quality of Service (QoS) by assigning Multiple RU based on the satisfaction of the required rate and the priority of the traffic type sent by each STA (Station). This method can lead to starvation issues, where lower-priority STAs fail to receive continuous resource allocation, preventing them from even attempting transmission. In this paper, we propose a technique that allocates the range of resources based on priority according to traffic types, dynamically adjusting the available resource range based on satisfaction levels to address the starvation problem. The proposed method operates by transferring the range of available resources from traffic types that do not meet minimum satisfaction levels to those that exceed a certain satisfaction threshold, thereby maintaining the existing QoS priority scheme and improving upon the starvation issues seen in previous methods

### Key words

802.11, starvation, WIFI7, QoS.

---

\* 국립금오공과대학교 컴퓨터소프트웨어공학과, nhw1124@kumoh.ac.kr

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과, sunmyeng@kumoh.ac.kr



## 1. 서론

IEEE 802.11 be 표준에서는 효율적인 스펙트럼 사용을 위한 더 많은 유연성을 제공하기 위해 MRU 할당을 도입하였다.

이는 IEEE 802.11 ax에서 단일 RU 할당과는 다르게 RU의 크기에 따라 소형 RU와 대형 RU로 분류한다. MRU 할당 방법을 위한 RU 결합은 소형 RU끼리 결합하며 대형 RU끼리 결합한다.

QoS 제공을 위한 우선순위에 대한 MRU 할당 방법은 각 STA마다 필요한 MRU를 정하고 각 STA의 트래픽 유형과 예상속도와 요구속도 사이의 우선순위에 따라 할당 순위를 정하고 각 STA에게 할당할 RU의 크기를 정하고 RU들의 논리적 위치를 정하는 방법[1]이 제안되었다.

하지만 해당 방식은 우선순위에 집중하고 있기 때문에 기아 문제가 발생할 수 있다. 이때 기아 문제란 우선순위가 높은 트래픽이 전송 기회를 독차지하여 우선순위가 낮은 트래픽의 전송률이 현저하게 떨어지는 문제를 말한다.

본 논문에서는 이러한 기아 문제를 해결하되 QoS를 지키기 위해 트래픽 유형별 우선순위 기반 유동적 자원 범위 제한 할당 기법을 제안한다.

## II. 트래픽 유형별 우선순위 기반 유동적 자원 범위 제한 할당

이 논문에서는 트래픽 유형별로 우선순위를 정하여 세 가지 분류로 나눈다 (표 1).

표 1. 트래픽 유형별 우선순위  
Table 1. Priority by traffic type

우선순위 집합	트래픽 유형	우선순위
A	Voice, Video	1
B	Best Effort	2
C	Background	3

초기 사용가능 자원 범위는 세 분류에 대하여 트래픽 유형별 우선순위에 대해 사용 가능 자원 범위를 분할한다 (그림 1).

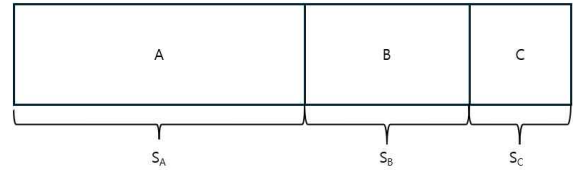


그림 1. 초기 자원 범위 분할  
Fig. 1. Initial resource range allocation

일정한 전송기간이 지난 후 각 유형별 만족도를 조사한다.(1)

$$Q_t = \frac{\sum(tr/requiretr)}{count(STA \in type t)} \quad (1)$$

$tr$ 은 전송속도를  $requiretr$ 은 요구 전송속도를 의미한다.

만족도에 따라 자원의 범위가 증가와 감소하게 되는데 현재 만족도만 고려할 때 변동성이 커져 극적인 자원 범위 이동 문제가 발생 할 수 있다. 이러한 문제를 해결하기 위해 과거의 만족도와 현재의 만족도에 따른 가중치 만족도를 조사하게 된다.(2)

$$Qw_t = (1 - \alpha) * Q_{last} + \alpha * Q_t \quad (2)$$

$Qw_t$ 은 과거 측정했던 만족도의 평균과 현재 측정된 만족도에 대해 가중치에 따라 계산한 가중치 만족도이다.

유형별 가중치 만족도와 최소 만족도, 최대만족도에 따라 자원 범위를 감소시키는 유형과 자원 범위를 증가시키는 유형으로 나눈다.  $H$ 와  $M$ 은 각각 최대 만족도, 최소 만족도를 의미한다. (3), (4)

$$Qw_t > H + dev_H \quad (3)$$

$$Qw_t < M - dev_M \quad (4)$$

$H$ 와  $M$ 만으로 최대 만족도에 대한 기준을 정할 경우 자원범위가 지속적으로 작은 값에 의해 움직임 가능성이 있어 편차에 대한 값으로 기준에 대한 여유 값을 가지도록 한다.

$dev_H$ 와  $dev_M$ 은 만족도가 증가, 감소 범위에 해당할 때  $Qw_t$ 와  $H, M$ 의 편차에 대해 과거 편차와 현재 편차에 대한 가중치 편차값이다. 다음과 같

은 식으로 산출한다 (5), (6)

$$dev_H = (1 - \delta_H) * dev_H + \delta_H * (Q - H) \quad (5)$$

$$dev_M = (1 - \delta_M) * dev_M + \delta_M * (M - Q) \quad (6)$$

자원 범위를 증가 시키는 유형은 다음과 같은 식에 따라 범위를 정한다 (7)

$$Qw_t > H + dev_H : \quad (7)$$

$$X = \beta * count(Q > H + dev_H)$$

$\beta$ 는 자원 범위의 이동이 얼마나 공격적이게 정해질지에 대한 임의의 값이다.

X는 총 나눠 줄 수 있는 자원 범위의 양이 된다.

자원을 늘리는 유형은 개수에 따라 최소 만족도에 대해 유형별 만족도가 얼마나 만족하고 있는가를 가중치로 설정한다. x값을 가중치에 따라 최소 만족도를 충족하지 못한 범위가 나눠 할당 받게 된다 (8), (그림 2).

$$Qw_t < M - dev_M : \quad (8)$$

$$S_n = S_n + X * \frac{(M - dev_M) - Qw_t}{\sum((M - dev_M) - Qw_t)}$$

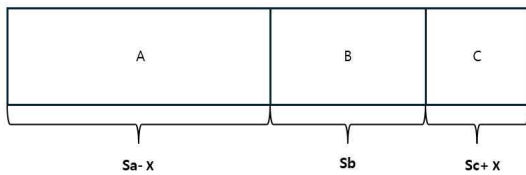


그림 2. 자원 범위 재분할 이후 예시  
Fig. 2. Example After Resource Range Reallocation

### III. 결 론

본 논문에서는 우선순위 할당 시 기아문제 발생에 대한 해결방안을 제안하였다. 제안된 방법은 과거와 현재 만족도의 움직임에 따른 사용 가능 자원 범위 제한을 통해 QoS를 지키며 우선순위가 낮은 트래픽에게 최소한의 보장을 지원하기에 기존의 방법보다 기아 문제 발생이 줄어든 결과를 보여준다.

### 참 고 문 헌

- [1] J. Pan, R. He, Y. Lin, and X. Fang, "Optimal MRUs Allocation Mechanism Based on User Priority for Wi-Fi 7 Network", GLOBECOM 2022 - 2022 IEEE Global Communications Conference\*, 1-6 Dec. 2022.

# TOPST NPU 보드를 활용한 실시간 차량 운전자 얼굴 인식 시스템 구현

송정헌\*, 이윤서\*, 이종호\*, 최혜림\*, 현종배\*\*

## Implementation of Real-Time Driver Face Recognition System Using TOPST NPU Board

Jeongheon Song\*, Yoonseo Lee\*, Jongho Lee\*, Hyerim Choi\*, and Jongbae Hyun\*\*

### 요 약

본 논문에서는 TOPST의 NPU(Neural Processing Unit) 보드를 활용한 운전자 얼굴 인식 시스템 구현 방법을 제안한다. YOLOv5 모델과 ArcFace 알고리즘을 이용한 얼굴 인식 알고리즘을 TOPST NPU 환경에 맞춰 수정하였는데, 두 모델을 NPU HW에 최적화하여 얼굴 검출 및 인식을 수행하였으며 결과 처리 및 차량 제어를 AP 보드에서 처리하도록 구현하였다. 그 결과, 제안된 시스템은 기본적인 얼굴 인식 기능을 수행하면서도 전력 소비 측면에서 이점을 보였다. 비록 처리 속도와 정확도 면에서 개선의 여지가 있으나, NPU 활용을 통한 임베디드 환경에서의 얼굴 인식 시스템 구현 가능성을 확인하고, 발전 방향을 찾을 수 있었다.

### Abstract

In this paper, we propose a method for implementing a driver face recognition system using TOPST's NPU (Neural Processing Unit) board. The face recognition algorithms using the YOLOv5 model and ArcFace algorithm were modified for the TOPST NPU environment, and the two models were optimized for the NPU HW to perform face detection and recognition, and the result processing and vehicle control were implemented to be handled by the AP board. As a result, the proposed system showed benefits in terms of power consumption while performing basic face recognition functions. Although there is room for improvement in terms of processing speed and accuracy, we have confirmed the feasibility of implementing a face recognition system in an embedded environment by utilizing NPUs and found a direction for further development.

### Key words

Neural Processing Unit, NPU, face recognition system, TOPST

---

\* 경북대학교 IT대학 컴퓨터학부, fight0037@knu.ac.kr, roiyeon@gmail.com, hopaper2367@naver.com, chr121910@gmail.com

\*\* 휴켄, jbhyn@huconn.com

※ "본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음" (2021-0-01082)

※ 이 연구는 과학기술정보통신부 및 정보통신산업진흥원의 "지역산업 SW인재양성 기반조성 사업" 지원으로 수행되었음.

## I. 서 론

최근 차량 도난 사고가 지속적으로 발생하고 있으며, 기존의 키 기반 보안 시스템만으로는 이를 완벽하게 방지하기 어려운 실정이다. 특히 복제된 키나 전자장치를 이용한 도난 시도가 증가하면서, 생체 인식과 같은 추가적인 보안 수단의 필요성이 대두되고 있다. 얼굴 인식 기술은 비접촉식이면서도 높은 보안성을 제공할 수 있는 솔루션으로 주목받고 있으나, 차량 환경에서의 실용적인 구현을 위해서는 연산 자원의 한계, 전력 소비, 인식 속도 등 여러 과제가 존재한다.

본 연구에서는 TOPST의 NPU 보드를 활용하여 차량용 운전자 얼굴 인식 시스템을 구현하고, 실제 차량 환경에서의 적용 가능성을 검토하였다. YOLOv5와 ArcFace 모델의 주요 연산을 NPU에서 처리하도록 구현하고, 결과 처리 및 차량 제어 기능을 AP보드에서 수행하는 방식을 채택하였다. 이를 통해 시동 제어와 연동하여 등록된 운전자만에 차량을 운행할 수 있도록 하는 보안 시스템의 기초를 마련하고자 하였다.

이를 통해 차량 도난 방지를 위한 저전력 얼굴 인식 시스템 구현 방법론을 제시하고, TOPST NPU 보드를 활용한 임베디드 환경에서의 딥러닝 모델 실행 최적화 방안에 대하여 기술하였으며 NPU와 AP 보드의 효율적인 작업 분배 방식을 제안하였다. 또한 실제 차량 환경에서의 한계점 분석 및 개선 방향을 도출함으로써, 비록 현재 단계에서는 성능상의 제약이 있으나 향후 하드웨어의 발전과 최적화 기법의 개선을 통해 실용적인 차량 보안 시스템으로 발전할 수 있는 가능성을 보여주었다. 특히 기존의 키 기반 보안 시스템과 결합하여 이중 보안을 구현할 경우, 차량 도난 방지에 효과적으로 기여할 수 있을 것으로 기대된다.

## II. 관련 연구

### 1. 임베디드 환경의 얼굴 인식 기술

얼굴 인식 기술이란 이미지나 영상의 얼굴을 검출하고 분석하여 누구인지 신원을 식별하는 기술이

다. 얼굴이라고 판단되는 영역을 찾아 객체를 검출한다. 검출된 얼굴을 데이터베이스에 저장된 이미지와 대조하여 일치하는지 판단한다.

임베디드 시스템은 특정한 기능만을 수행하도록 제작된 컴퓨터 시스템으로 마이크로프로세서를 내장하고 있는 단일 보드 마이크로 컴퓨터이다. 실시간 계산을 행하는 데 있어 제약을 가지므로 단순하고 최적화된 알고리즘을 활용해야 한다.

### 2. ArcFace

ArcFace 모델은 각도 마진 기반 손실함수 (angular margin loss)를 적용해 차별적인 특징을 학습하는 데에 중점을 둔 얼굴 인식 알고리즘이다. Softmax 분류기를 사용하는 다른 알고리즘과 달리 추가적인 마진 값을 주어 같은 클래스 내에서는 비슷하게 임베딩(Embedding)이 형성되고, 서로 다른 두 클래스 사이에는 더 큰 격차를 만든다. 각도에 직접적으로 페널티를 주는 방식으로 다양한 각도와 조명 조건에서도 얼굴을 인식할 수 있게 한다.

### 3. YOLOv5-Face

YOLOv5-Face는 YOLOv5를 얼굴 탐지에 최적화한 버전으로, 일반 객체 탐지 모델보다 얼굴 인식 성능이 뛰어나다. 추가적인 랜드마크 회귀 헤드를 탑재하여 얼굴 랜드마크를 탐지하며, Wing Loss 함수를 적용해 랜드마크 위치의 정확도를 높였다.

전통적인 Haar 특징 기반 얼굴 탐지 방식과 비교할 때, YOLOv5-Face는 정확성, 유연성, 속도 면에서 현저한 장점을 가진다. Haar 특징은 사전에 정의된 고정 패턴을 사용하여 정면 얼굴을 효율적으로 탐지하지만, 얼굴 각도, 조명 변화, 배경 복잡성에 취약하다는 한계가 있다. 반면 YOLOv5-Face는 딥러닝을 통해 다양한 얼굴 특징을 학습함으로써, 다양한 환경과 각도에서도 높은 정확도로 얼굴을 탐지할 수 있다.

## III. 시스템 구현

### 1. 임베디드 시스템 개발환경

본 논문에서는 TOPST의 AP보드와 NPU 보드를

기반으로 시스템을 구현하였다. AP보드는 시스템의 전반적인 제어와 통신을 담당하며, NPU보드는 딥러닝 모델의 추론을 전담하여 처리한다. 두 보드는 PCIe 인터페이스를 통해 연결되어 실시간 데이터 교환이 가능하도록 설계되었다. [표 1]은 AP보드의 하드웨어 사양을, [표 2]는 NPU보드의 사양을 보여준다.

[표 1] AP보드 하드웨어 사양

<b>CPU</b>	ARM® Cortex-A72 (1.6 GHz) & Cortex-A53(1.45 GHz)
<b>CPU</b>	ARM® Cortex-R5 (300MHz)
<b>GPU</b>	IMG PowerVR 9XTP(169 GFLOPS)
<b>Memory</b>	4GB / 8GB
<b>Storage</b>	Micro SD Card Slot
	eMMC 5.1 8GB / 16GB / 32GB
<b>Network</b>	Gigabit Ethernet
<b>Camera</b>	MIPI CSI2 2-lane 1ch, MIPI CSI2 4-lane 1ch

[표 2] NPU보드 하드웨어 사양

<b>CPU</b>	ARM® Cortex-A53(1.45 GHz)
<b>NPU</b>	1.0GHz, 8 TOPS
<b>Memory</b>	4GB LPDDR4X
<b>Storage</b>	32GB eMMC 5.1 Flash Storage, 8MB SNOR
<b>Camera IN</b>	MIPI CSI2 2-lane
<b>Video OUT</b>	MIPI DSI 2-lane

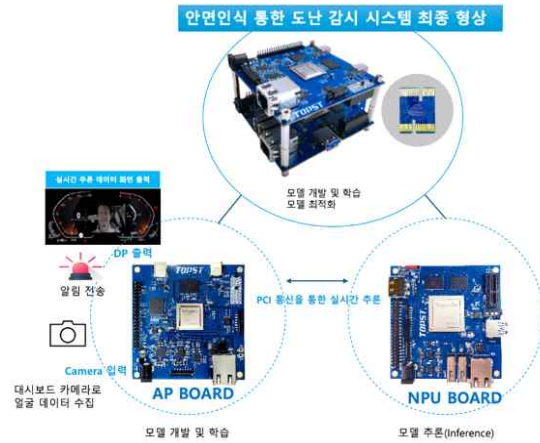
## 2. 시스템 구조

크게 세 부분으로 나누어 시스템을 구현하였으며 [그림 1]과 [그림 2]를 통해 전체 시스템의 흐름을 확인할 수 있다.

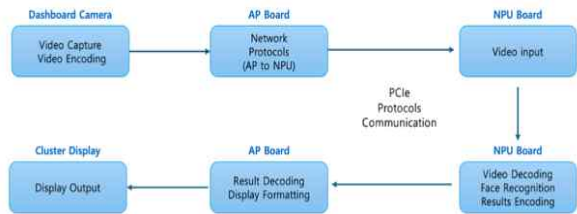
영상 획득 및 전처리 파트에서는 차량 대시보드 카메라를 통한 비디오 캡처 및 인코딩을 처리하고, AP보드에서 초기 영상 처리를 수행한다.

얼굴 검출 및 인식 파트에서는 NPU보드에서 YOLOv5-Face 기반 얼굴 검출이 실행되고 ArcFace 알고리즘을 통한 얼굴 특징 추출 및 인식을 수행한다. 이후 결과를 인코딩하고 AP보드로 전송한다.

후처리 및 디스플레이 파트에서는 AP보드에서 수행되는 결과 처리 및 차량 제어를 다룬다. 클러스터 디스플레이를 통한 결과 출력까지 수행한다.



[그림 1] 시스템 최종 형상



[그림 2] 시스템 소프트웨어 구조

또한 시스템의 효율적인 운영을 위해 [표 3]과 같이 작업을 분배하였다.

[표 3] AP보드와 NPU보드 간 작업 분배

AP보드	카메라 제어 및 영상 획득
	네트워크 프로토콜 처리
	디스플레이 출력 처리
	시스템 전반의 제어 및 관리
NPU보드	YOLOv5 모델 실행
	ArcFace 특징 추출
	얼굴 인식 결과 생성

## IV. 성능 분석

다음 [그림 3] 은 운전자의 얼굴을 인식하는 과정이며, 그 결과로 [그림 4]와 같이 나타난다. 이후 차량의 시동이 걸리는 결과가 [그림 5]와 같이 나타난다.

시스템의 각 단계별 처리 시간을 측정한 결과는 [표 4]와 같다. 전체 파이프라인의 end-to-end 지연시간은 평균 478.3ms로 측정되었으며, 이는 초당 약 2.1프레임의 처리 속도를 의미한다.



[그림 3] 운전자를 인식 중인 클러스터 UI



[그림 4] 운전자 인식 완료한 클러스터 화면



[그림 5] 차량 시동 결과

[표 4] 시스템 단계별 처리시간 측정 결과

		평균 시간
영상 획득 및 전처리	카메라 캡처	33.3ms (30fps)
	전처리 및 인코딩	25ms
	AP-NPU 보드 간 전송	30ms
얼굴 검출 및 인식	YOLOv5 얼굴 검출	150ms
	ArcFace 특징 추출	120ms
	결과 인코딩 및 전송	125ms
후처리 및 출력	디스플레이 출력	20ms

이러한 처리 결과를 분석해 본 결과, 임베디드 시스템의 하드웨어적 제약과 해당 NPU가 처리할 수 있는 모델의 제약에도 불구하고 효과적인 얼굴 인식이 가능함을 확인할 수 있었다. 얼굴 인식 부분을 Haarcascade 알고리즘으로 대체하여 AP 보드에서 처리한 것과 비교했을 때 평균 처리 시간은 큰 차이가 없지만, 정확도 측면에서 향상을 보였다. 이는 NPU가 시스템의 전반적인 성능 개선에 긍정적인 영향을 미쳤음을 알 수 있다.

결과적으로, 평균 78.5%의 준수한 인식률을

달성하였으며, 이는 실용적인 얼굴 인식을 구현하는데 충분하다. 추가적으로, 모델의 경량화와 최적화를 통해 에너지 효율성을 높였으며, 이는 임베디드 시스템에서 중요한 요소이다. 기술이 발전함에 따라 하드웨어 성능이 향상된다면 더욱 높은 인식률과 안정적인 시스템 구현이 가능할 것으로 기대된다.

## V. 결론

논문에서는 TOPST NPU 보드를 활용하여 차량용 실시간 운전자 얼굴 인식 시스템을 구현하고 그 성능을 검증하였다. YOLOv5 모델과 ArcFace 알고리즘을 NPU에서 구동하고, 후처리 및 차량 제어는 AP 보드에서 처리하는 방식으로 시스템을 설계하였다. 실험 결과, 등록된 4명의 운전자에 대해 다양한 조명 조건에서 테스트를 진행했을 때 평균 78.5%의 인식률을 보였으며, 초당 2.1프레임의 처리 속도를 달성하였다.

현재 많은 차량에서 IVI(In-Vehicle Infotainment) 시스템 등이 널리 사용되고 있다. 이러한 시스템에 NPU를 연동함으로써 객체 탐색 등 다양한 기능을 활용할 수 있다. 본 논문에서 제안한 얼굴 인식 시스템은 임베디드 환경에서의 적용 가능성을 보였으며, 차량에서 NPU를 사용하여 다양한 응용을 실현할 수 있음을 제시하였다. 이는 자동차뿐만 아니라 모바일 환경 또는 다양한 로봇에서 NPU의 활용 가능성을 보여준다. 향후 저전력, 고성능이 요구되는 엣지 디바이스 환경에서 NPU의 활용도가 더욱 증가할 것으로 기대된다.

## 참고 문헌

- [1] Qi, Delong, et al. "YOLO5Face: Why reinventing a face detector." *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2022
- [2] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, Stefanos Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition", *IEEE Transactions on Pattern*

Analysis and Machine Intelligence, vol. 44, no.  
10, pp. 5962-5979, 2022

- [3] 여준기, 신경선, 한진호, 박찬, 김태현, 권영수,  
"AP-NPU 임베디드 시스템에서의 Yolo v3 응용  
최적화", 2021년 대한전자공학회 추계학술대회  
논문집, pp. 196-198, 2021





# 공공데이터를 활용한 재료 기반 레시피 검색: 요~레시피

이지강\*, 이정우\*\*, 손현\*\*, 정재운\*\*, 이규원\*\*\*, 이해연\*\*\*\*

## Ingredient-based Recipes Search using Public Data: Yo-recipe

Ji-gang Lee\*, Jung-woo Lee\*\*, Hyun Son\*\*, Jae-un Jeong\*\*, Kyu-Won Lee\*\*\*  
and Hae-Yeoun Lee\*\*\*\*

### 요 약

요즘 증가하는 1인 가구의 경우, 다양한 식재료를 구매하여 요리하는 데 금전적 부담이 크다. 이에 본 논문에서는 현재 보유한 식재료를 효율적으로 활용하고자 하는 사용자를 위해 공공데이터 기반의 요~레시피 시스템을 제안한다. 이 시스템은 1인 가구, 알레르기 환자, 특정 조리 스타일 선호자, 영양 관리 희망자를 주요 대상으로 한다. '요레시피'는 사용자의 보유 식재료를 기반으로 맞춤형 레시피를 제공하며, 식재료 낭비 감소, 알레르기 대응, 영양 관리 등 다양한 요구를 충족시킨다. 본 시스템을 통해 사용자의 건강한 식생활과 경제적인 식재료 관리를 지원할 것으로 기대된다.

### Abstract

.In the case of the increasing number of single-person households these days, the financial burden is high in purchasing and cooking various ingredients. Therefore, this paper proposes a public data-based Yo-recipe system for users who want to efficiently utilize their current ingredients. This system mainly targets single-person households, allergy patients, preferences for specific cooking styles, and those who wish to manage nutrition. Yo-recipe provides customized recipes based on the user's own ingredients and meets various demands such as reduction of food waste, allergic response, and nutritional management. It is expected to support users' healthy eating habits and economic food ingredient management through this system.

### Key words

Public data, Personalized recipe, Recipe recommendation, Ingredient management, Single-person household

## 1. 서 론

최근 1인 가구의 증가로 인해 식생활 관리가 중요한 사회적 이슈로 대두되고 있다. 1인 가구의 경우 통계청이 발표한 '2023 통계로 보는 1인 가구'

통계에 따르면 그림 1과 같이, 2017년에 28.6%였던 1인 가구가 2022년에는 34.5%로 지속적인 상향추세를 보이고 있다. 이러한 1인 가구의 식생활 만족도는 다인 가구에 비해 낮은 것으로 알려져 있는데, 1인 가구의 경우 아침 결식률과 패스트푸드 및 가공

\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생, wlrkd4095@gmail.com

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생

\*\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수

\*\*\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수, haeveoun.lee@kumoh.ac.kr(교신저자)

식품의 섭취 빈도, 식사 형태가 불규칙하거나 영양소 섭취가 부족한 편 이러한 문제점이 1인 가구의 식생활 만족도를 감소시킨다[2].



그림 1. 연도별 1인가구 추이[1]  
Fig. 1. Changes in single-person households by year[1]

기존의 만개의 레시피[3]와 같이 이러한 1인 가구에 도움이 될 요리 레시피들을 제공하는 사이트들이 있지만, 자신이 보유하고 있는 식재료를 다중으로 넣어 요리가 가능한 레시피를 제공하는 기능은 제공하지 않고 있다.

이와 같은 재료 기반 레시피 제공을 위해 본 연구에서는 사용자가 보유한 식재료를 기반으로 적합한 요리법을 추천하는 ‘요~레시피’ 시스템을 제안한다. 본 시스템은 공공데이터를 활용하여 사용자가 입력한 보유 식재료를 바탕으로 가능한 요리 목록과 간단한 조리법을 글과 이미지로 제공하여 목적에 맞는 요리를 할 수 있다. 또한, 자주 먹게 되는 레시피의 경우 스크랩 기능을 통해 해당 레시피를 저장하고 관리 가능한 서비스를 제공한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 재료 기반 레시피 시스템 ‘요~레시피’ 개발의 절차와 주요 기능을 설명하고, 3장에서는 개발된 시스템의 결과를 나타낸다. 4장에서는 결론과 확장성을 제시한다.

## II. 재료 기반 레시피 시스템

요~레시피 시스템은 기본적으로 사용자가 보유하고 있는 재료에 초점을 둔 레시피 제공 시스템이다. 사용자는 자신이 보유하고 있는 재료를 콤마(,)를 기준으로 입력하며, 알레르기성 재료를 제외할 수

있으며, 조리 스타일과 요리 컨셉을 지정하여 레시피를 조회할 수 있다.

조회하여 나온 레시피들에 대해 음식 이미지를 클릭시 상세페이지로 연결이 되며, 상세페이지에선 레시피의 재료 목록, 영양 정보, 레시피 팁, 요리 과정이 존재한다. 제목 옆에 있는 좋아요 버튼을 클릭하게 되면 마이페이지 내에 레시피가 저장된다.

### 2.1 사용한 공공 데이터 상세정보

본 연구에서 사용된 공공데이터는 식품안전나라에서 제공하는 조리식품의 레시피 정보이다[4]. 이 데이터셋은 총 1,130개의 레시피 정보를 포함하고 있으나, 사용 가치 없음으로 판단된 7개의 레시피 정보를 제외한 총 1,123개의 레시피를 사용한다. 각 레시피는 표 1과 같이 주요 필드로 구성되어 있다.

각 주요 필드의 내용은 다음과 같다. 일련번호: 각 레시피의 고유 식별자. 메뉴명: 요리의 이름. 예: “칼륨 듬뿍 고구마죽“ 조리 방법: 해당 요리의 주요 조리 기술 예: “끓이기“ 요리 종류: 요리의 카테고리 예: “후식“ 중량 및 영양 정보: 1인분 기준의 중량과 열량, 탄수화물, 단백질, 지방, 나트륨 등의 영양 정보. 이미지 경로: 요리의 완성된 모습을 보여주는 이미지 파일의 URL을 제공한다. 소형과 대형 이미지가 각각 제공된다. 재료정보: 요리에 필요한 재료와 그 양을 상세히 기술한다. 만드는 법: 요리의 조리과정을 단계별로 설명한다. 최대 20단계까지 제공하며 만드는 법 이미지로 각 조리 단계에 해당하는 이미지의 URL을 제공한다.

### 2.2 시스템 구조 및 특징

본 시스템은 1인 가구를 대상으로 한다. 사용자는 레시피를 쉽게 검색할 수 있도록 다양한 필터 기능을 제공한다. 본 시스템은 그림 2와 같이 서버-클라이언트 구조를 가진다.

클라이언트 측은 JavaScript(EJS)를 이용하여 동적이며 반응형인 웹 페이지를 구현하였다. 사용자의 요청을 서버 측에 전송하고 서버에서 받아온 결과를 시각적으로 보인다. 서버 측은 Node.js기반의 Express.js 프레임워크를 사용하여 클라이언트의 요청

을 처리하고, 비즈니스 로직을 수행하며, 데이터베이스와 연결되어 필요한 데이터를 제공한다. 데이터베이스 MySQL은 공공데이터를 저장하고 있으며, 클라이언트의 요청에 따라 서버와 상호작용하며 결과값을 반환한다.

표 1. 공공데이터 상세 정보  
Table 1. Public-data details

데 이 터 상 세 정 보	일련번호	ex) 17
	메뉴명	칼륨 듬뿍 고구마죽
	조리방법	끓이기
	요리종류	후식
	종량(1인분)	-
	열량	205
	탄수화물	35
	단백질	3
	지방	6
	나트륨	68
	해쉬태그	-
	이미지경로 (소)	http://www.foodsafetykorea.go.kr/uploa dimg/cook/10_00017_2.png
	이미지경로 (대)	http://www.foodsafetykorea.go.kr/uploa dimg/cook/10_00017_1.png
재료정보	고구마죽 고구마 100g(2/3개), 설탕 2g(1/3작은술), 찹쌀가루 3g(2/3작은 술), 물 200ml(1컵), 잣 8g(8알)	
만드는법 1~20	1. 고구마는 깨끗이 씻어서 껍질을 벗기고 4cm 정도로 잘라준다. . .	
만드는법 이미지 1~20	http://www.foodsafetykorea.go.kr/uploa dimg/cook/20_00017_1.png . . .	

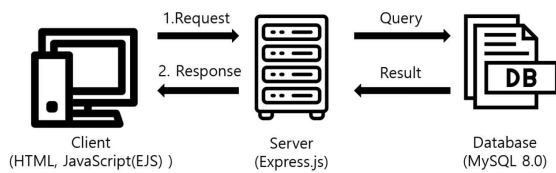


그림 2. 개발 시스템의 구조  
Fig. 2. Structure of Developed System

### III. 시스템 개발 결과

그림 3은 개발한 요~레시피 시스템의 메인 화면으로 재료, 빼고 싶은 재료를 콤마(,)를 기준으로 다중 입력하고, 조리 스타일, 요리 컨셉을 선택한 뒤 조리하기 버튼을 클릭하여 필터링된 결과를 볼 수 있다.

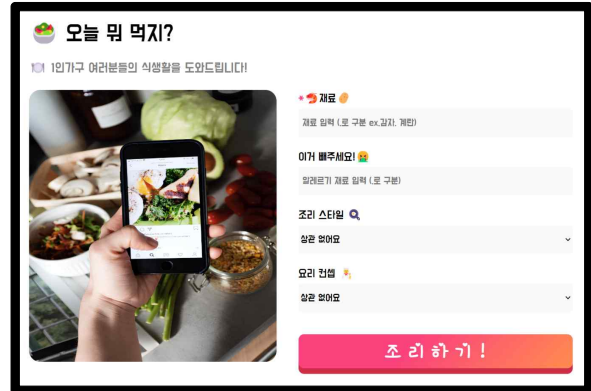


그림 3. 요~레시피 시스템 메인화면  
Fig. 3. Main-screen of Yo~recipe system

그림 4는 입력의 예시와 입력을 기반으로 필터링되어 나온 음식 목록들을 보인다.

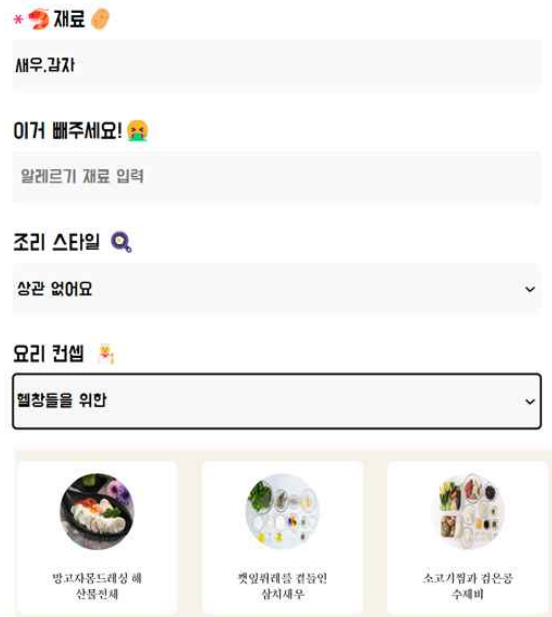


그림 4. 필터링된 레시피 화면  
Fig. 4. Filtered Recipe Screen

그림 5에서는 그림 4의 필터링된 음식 목록들 중 마음에 드는 이미지를 클릭하였을 때 연결되는 레시피 상세페이지로, 재료목록과 레시피 팁, 영양 정보와 사진과 같이 순차적인 조리과정을 보인다. 추가적으로 페이지의 설명을 보고도 이해가 안될 수도 있기에 해당 요리 이름을 기반으로 유튜브, 네이버, 만개의 레시피의 페이지에 연결되도록 기능을 제공하여 편의를 제공한다.



그림 5. 레시피 상세페이지  
Fig. 5. Recipe Details Page

#### IV. 결 론

본 논문에서는 1인 가구의 효율적인 식재료 관리와 건강한 식생활을 위한 요~레시피 시스템을 제안하였다. 공공데이터를 활용한 이 시스템은 사용자의 보유 식재료를 입력받아 맞춤형 레시피를 추천하고, 상세한 조리 방법을 제공한다. 제안된 방법은 식재료 낭비를 줄이고 영양 균형을 고려한 식단 구성을 가능하게 하여, 기존의 레시피 검색 방식보다 1인 가구의 실질적인 요구를 더욱 효과적으로 충족시킨다. 그러나 적은 공공데이터 크기로 인해 필터를 많이 넣게 되면 조건을 만족시키는 레시피 검색 결과가 나오지 않는 제약이 일부 존재한다. 향후 사용자 데이터 축적을 통해 개인화된 영양 관리 및 식재료 구매 조언 등으로 서비스를 확장할 수 있을 것으로 기대된다.

#### 참 고 문 헌

[1] KOSTAT, [https://kostat.go.kr/board.es?mid=a10301010000&bid=10820&list\\_no=428414&act=view&mainXml=Y](https://kostat.go.kr/board.es?mid=a10301010000&bid=10820&list_no=428414&act=view&mainXml=Y)

[2] H.-J. Lee, "Understanding Food Consumption Characteristics of Single-Person Households by Age Group", Korean society of consumer policy and education, 2022, Vol. 18, No.3, pp. 1-28.  
[3] "10000 Recipe", <https://www.10000recipe.com/>  
[4] FOODSAFETYKOREA, "Recipe DB for Cooked Food", [https://www.foodsafetykorea.go.kr/api/openApiInfo.do?menu\\_grp=MENU\\_GRP31&menu\\_no=661&show\\_cnt=10&start\\_idx=1&svc\\_no=COOKRCP01](https://www.foodsafetykorea.go.kr/api/openApiInfo.do?menu_grp=MENU_GRP31&menu_no=661&show_cnt=10&start_idx=1&svc_no=COOKRCP01)

# 반려 동물 관리를 위한 통합 애플리케이션: 펫토피아

손초희\*, 방유림\*\*, 변아영\*\*, 서영은\*\*, 이해연\*\*\*

## Integrated Application for Pet Management: Petopia

Cho-Hui Son\*, A-yeong Byun\*\*, Yu-Rim Bang\*\*, Young-Eun Seo\*\*, and Hae-Yeoun Lee\*\*\*

### 요약

매년 반려동물을 기르는 가구가 증가하고 있으며, 많은 사람들이 반려동물을 인생의 동반자로 여기며 함께 살아가고 있다. 이에 따라, 반려동물의 건강 관리와 생활을 지원하는 다양한 시스템들이 개발되고 있다. 그러나 현재의 시스템들은 대부분 건강 관리와 미용 관리에 한정되어 있으며, 반려동물의 전반적인 생활을 통합적으로 관리하기는 어려운 실정이다. 본 논문에서는 반려동물의 의료, 미용, 라이프 등 다양한 기능을 제공하는 통합 관리 시스템 “펫토피아”를 제안한다. 이를 통해 보호자들이 반려동물의 건강 상태뿐만 아니라 일상생활까지 체계적으로 관리할 수 있는 효과적인 솔루션을 제시하고자 한다.

### Abstract

The number of households with pets has been increasing every year, with many people considering pets as life companions. Consequently, various systems have been developed to support the health and daily care of pets. However, most current systems are limited to health and grooming management, making it difficult to comprehensively manage the overall well-being of pets. This paper proposes an integrated management system, petopia, that provides various features, including medical care, grooming, and life management for pets. Through this system, we aim to offer an effective solution that allows pet owners to systematically manage not only the health status of their pets but also their daily lives.

### Key words

Pet Management, Integrated System

## 1. 서론

2023년 한국 반려동물 보고서에 따르면, 2022년 말 기준으로 반려동물을 기르는 가구인 '반려 가구'가 전년 대비 2.8% 증가한 것으로 나타났다[1,2]. 이

를 통해 많은 사람들이 반려동물을 인생의 동반자로 여기며 함께 살아가고 있음을 알 수 있다. 이에 따라 애견 미용실, 애견 카페, 애견 장례식장 등 많은 사업이 주목 받고 있으며 반려동물을 위한 다양한 앱이 존재한다. 그러나 현재 대부분의 반려동물

\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생, schabc8436@kumoh.ac.kr

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생

\*\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수, haeveoun.lee@kumoh.ac.kr(교신저자)

관련 앱은 미용, 의료 등 특정 분야에만 집중된 서비스만을 제공하는 경우가 많다.

본 논문에서는 반려동물의 의료, 미용, 라이프 등을 통합적으로 관리 할 수 있는 펫토피아 반려동물 통합 애플리케이션 개발에 대해서 설명한다. 또한 반려동물 통합 애플리케이션의 활용 가능성에 대해서도 기술한다.

본 논문의 구성은 다음과 같다. 2장에서는 시스템 개요와 주요 기능을 간략히 설명한다. 3장에서는 개발된 시스템의 결과를 나타내며, 마지막으로 4장에서는 결론 및 향후 활용 방안을 제시한다.

## II. 반려동물 통합 애플리케이션 시스템

### 2.1 펫토피아

펫토피아는 반려인들이 많이 이용하는 서비스인 의료, 미용, 라이프 서비스를 종합적으로 제공하는 통합 애플리케이션이다. 이 시스템은 Spring Boot와 Flutter를 사용하여 개발되었으며, 메인 UI는 (그림 1)과 같다.



그림 1. 펫토피아  
Fig. 1. Petopia

반려동물의 이상 행동과 먹으면 안되는 음식 정보를 제공하고 다이어리를 통한 건강 관리, 음악을

위한 놀이 및 훈련, 병원 및 미용실 지도 제공, 반려동물 등록증 제공의 기능을 수행한다.

### 2.2 기능 설명

펫토피아는 다양한 기능을 제공한다. 기능은 크게 의료 관리, 미용 관리, 라이프 관리가 있다.

의료 관리 기능을 통해 사용자는 서울의 각 구별로 병원 목록을 확인할 수 있으며, 카카오 지도를 활용하여 병원의 정확한 위치를 파악할 수 있다. 피부병 증상 부위를 사진찍어 AI를 통해 간단하게 진단받아 볼 수 있으며, 비슷한 환경에서 자란 반려견들의 질병을 조회하여 반려견이 걸릴 수 있는 질병을 미리 예방할 수 있다. 또한, 예방 접종 기록을 체계적으로 관리할 수 있어 반려동물의 건강 관리를 돕는다.

미용 관리 기능에서는 서울 지역 내 구별 미용실 목록을 제공하며, 마찬가지로 카카오 지도를 통해 정확한 위치 확인이 가능하다. 또한 펫스널 컬러 진단 기능을 통해 반려동물과 어울리는 색상을 알아볼 수 있다.

라이프 관리 기능은 반려동물에 대한 유용한 팁을 제공하며, 반려동물의 일상을 기록할 수 있는 다이어리 기능을 포함한다. 또한 펫 필터 기능을 통해 반려동물에게 간단한 필터를 적용하여 사진을 찍을 수 있으며, 펫 스피커 기능을 통해 반려동물과 놀이를 하거나 훈련할 수 있다.

### 2.3 데이터베이스 설계

본 시스템은 MySQL 데이터베이스 관리 도구를 사용하였고 (그림 2)에 ERD 구조도를 도시하였다.

사용자의 정보를 담고 있는 users, 반려동물의 정보를 담고 있는 pet, 다이어리의 정보를 담고 있는 diary, 접종 기록의 정보를 담고 있는 shot\_records, 다이어리의 약 정보를 담고 있는 diary\_medicine, 약 정보를 담고있는 medicine, 품종의 정보를 담고 있는 species, 동물의 종류 정보를 담고 있는 pet\_type, 반려동물 팁 정보를 담고있는 tip, 놀아주기 정보를 담고 있는 play로 이루어져 있다.

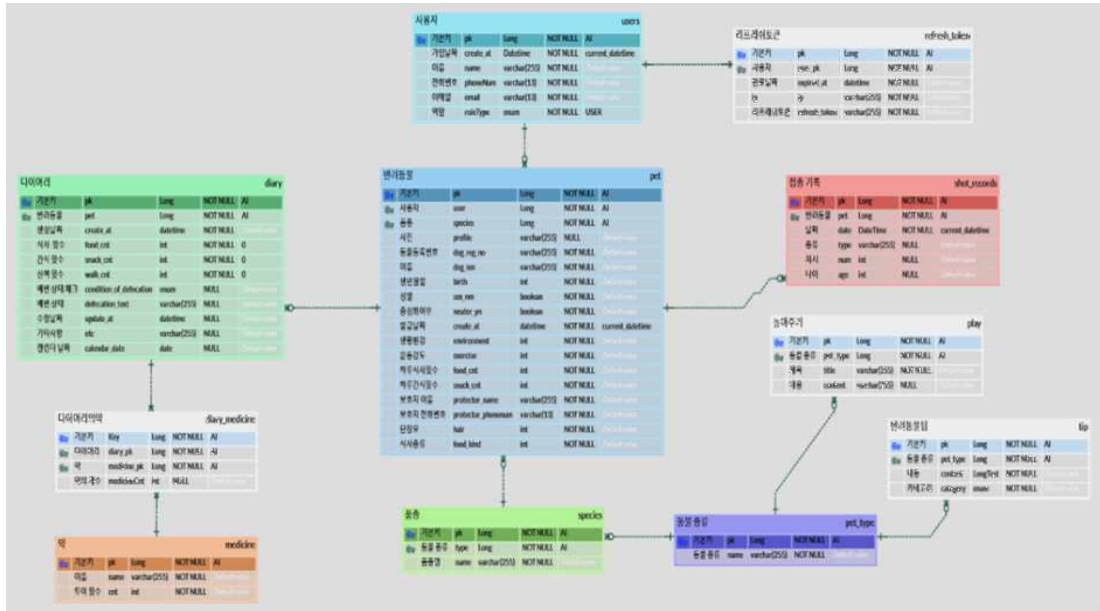


그림 2. 펫토피아의 ERD  
Fig. 2. ERD of Petopia

### III. 시스템 개발 결과

### 3.2. 기록 관리 기능

본 절에서는 펫토피아 애플리케이션의 각 기능에 대하여 개발한 결과에 대하여 설명한다.

달력에서 원하는 날짜를 선택하여 등록된 반려동물별로 예방접종 기록과 다이어리를 관리할 수 있다. 기록 관리 기능 UI는 (그림 4)과 같다.

#### 3.1 지도 조회 기능

서울의 각 구별 병원과 미용실 목록을 제공하며, 각 항목에는 병원이나 미용실의 정확한 주소와 연락처가 포함되어 있다. 또한, 카카오 지도를 통해 해당 장소의 위치를 시각적으로 확인할 수 있다. 지도 조회 기능 UI는 (그림 3)과 같다.



그림 3. 지도 조회 기능  
Fig. 3. Map Viewing Feature



그림 4. 기록 관리 기능  
Fig. 4. Record Management Feature

### 3.3. 진단 기능

반려동물과 유사한 환경에서 자란 다른 반려동물들의 질병 정보를 조회할 수 있으며, 해당 반려동물의 정보와 함께 질병명, CRP, IgG, AFP, 스트레스 수치를 확인할 수 있다. 또한 피부병 증상 부위의 사진을 통해 간단히 진단받고, 반려동물의 사진을 통해 어울리는 색상 3가지를 확인할 수 있다. 진단 기능의 UI는 (그림 5)와 같다.

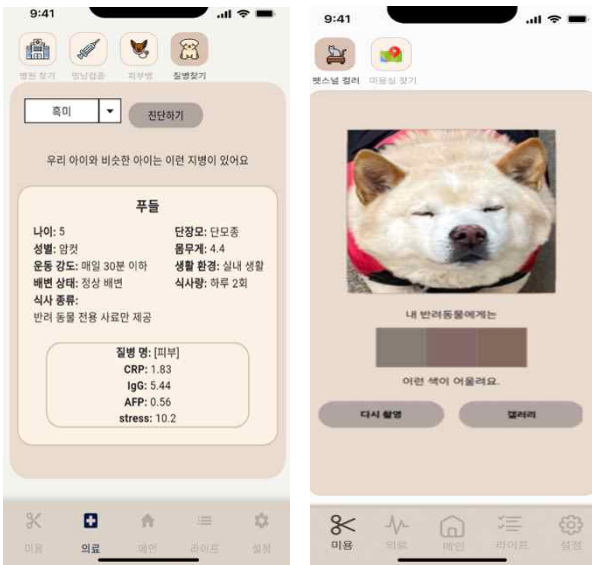


그림 5. 진단 기능  
Fig. 5. Diagnosis Feature

## IV. 결론 및 향후 활용 방안

본 논문에서는 반려 동물 통합 애플리케이션인 펫토피아 개발에 대해 설명하였다. 펫토피아는 근처의 병원과 미용실의 위치를 알려주고, 반려동물의 예방 접종 기록과 다이어리 기록을 쉽게 관리 할 수 있으며 펫 스피커 기능을 통해 반려동물을 훈련 시키거나 놀아 줄 수 있다.

이 시스템을 통해 반려동물 보호자들은 병원과 미용실 위치 확인, 예방 접종 기록 관리, 훈련과 같은 다양한 관리 작업을 한번에 처리할 수 있으며, 이를 통해 반려동물의 건강과 삶의 질 향상에 기여 할 수 있을 것으로 기대된다. 특히 의료 관리 기능을 통해 예방 접종 기록을 체계적으로 관리함으로써

써 반려동물의 건강을 보다 세밀하게 모니터링할 수 있다. 또한 가까운 미용실과 병원을 쉽게 찾을 수 있어 보호자의 시간과 노력을 절약할 수 있다. 더불어 펫 스피커 기능을 통해 반려동물 훈련을 손쉽게 진행할 수 있어, 보호자와 반려동물 간의 유대감이 더욱 강화될 것으로 기대된다. 이처럼 펫토피아는 반려동물 보호자들에게 통합적이고 편리한 관리 도구로서 실질적인 도움을 제공할 수 있을 것이다.

## 참 고 문 헌

- [1] KB 경영연구소, “2023 한국 반려동물 보고서”, 2023.06, pp. 18
- [2] 농림축산부, 2022년 동물보호에 대한 국민의식조사 결과 발표, 2022.02.02., pp. 01



# 국내 잉글랜드 프리미어리그 시청자를 위한 오락 애플리케이션 제안

이윤지\*, 최인영\*\*, 박현진\*\*, 김윤수\*\*, 이다은\*\*, 박수이\*\*, 최종인\*\*

## Proposal for an Entertainment Application for England Premier League Fans in South Korea

YunJi Lee\*, InYoung Choi\*\*, HyeonJin Park\*\*, YoonSu Kim\*\*, DaEun Lee\*\*, SuE Park\*\*, and JongIn Choi\*\*

### 요약

최근 국내에서 프리미어리그의 인기가 증가하면서, 해외 축구를 즐겨 시청하고 경기 관람에 대한 애착이 있는 10-30대 여성 팬층이 증가하는 추세이다. 이에 본 연구는 국내 프리미어리그 시청자들을 위한 오락 콘텐츠를 제공하는 애플리케이션을 제안한다. 이를 위해 해외 축구 시청 경험이 있는 여성을 대상으로 온라인 설문 조사와 심층 인터뷰를 실행했다. 프리미어리그 시청자가 축구 경기를 보다 재미있게 즐길 수 있는 공간을 연구의 중요 가치로 설정했다. 따라서 프리미어리그 경기에 대한 사용자의 애착도와 다양한 감상을 표현할 수 있는 수단을 제공하는 애플리케이션을 제안하고자 한다. 이는 프리미어리그 관람의 환경적 제약을 극복하고, 단순 경기 시청을 넘어선 즐거운 경기 관람 경험 개선에 기여할 것으로 기대된다.

### Abstract

Recently, the popularity of the Premier League has been increasing in South Korea, leading to a growing demographic of female fans in their teens to 30s who enjoy watching overseas soccer and have a strong attachment to game viewing. In response, this study proposes an application that provides entertainment content for domestic Premier League viewers. To achieve this, online surveys and in-depth interviews were conducted with women who have experience watching overseas soccer. The study emphasizes the importance of creating a space where Premier League viewers can enjoy soccer games more engagingly. Therefore, we propose an application that offers a means for users to express their attachment to Premier League matches and share various impressions. This is expected to help overcome the environmental limitations of watching the Premier League and contribute to enhancing the overall viewing experience beyond simple game watching.

### Key words

프리미어리그, 축구, 스포츠 경기 관람, 오락, 심박수, 축구 일기

---

\* 서울여자대학교 소프트웨어융합학과, yoonji01ee@gmail.com,

\*\* 서울여자대학교 디지털미디어학과, cla2rde1une@gmail.com, hyeon0150@gmail.com, yoonsu1717@gmail.com, leedaun07@naver.com, spark44@swu.ac.kr, funtech@swu.ac.kr(교신저자)

※ 본 연구는 교육부 및 한국연구재단 대학혁신지원사업으로부터 지원받은 연구임(2024)

## 1. 서 론

최근 Over-the-top(이하 OTT) 플랫폼 외 다양한 플랫폼을 통해 스포츠 콘텐츠가 공급되면서 대중들의 스포츠 시청에 대한 접근이 매우 편리해졌고, 이에 따라 스포츠에 대한 관심이 커지고 있다. 이러한 스포츠 중에서 꾸준히 인기를 끌고 있는 콘텐츠는 잉글랜드 프리미어리그(EPL: English Premier League)이다. 현재 202개 국가에서 6억 이상의 인구가 시청하고 있으며[1] 축구 통계 매체인 'Opta'가 매긴 세계 축구 상위 20개 리그 순위 중 잉글랜드 프리미어리그가 세계 최고의 리그로 뽑혔다[2].

프리미어리그가 각 중계권에서 시청자를 모으기 위한 주요 수단으로 인식되면서, 스포츠 경기 중계권에 대한 방송사 간 경쟁이 치열해지고 있다[3][4]. 이렇듯 프리미어리그는 스포츠 시장에서 충분히 매력 있는 콘텐츠라 할 수 있다.

잉글랜드 프리미어리그는 타 리그보다 한국 선수들의 활약이 돋보이기에 국내 시청자들의 관심도 또한 높다. 2021-2022시즌 프리미어리그를 중계한 SPOTV는 프리미어리그 구단의 모든 경기를 생중계 하며, 타 리그와 달리 높은 시청률을 보였다. 다음은 SPOTV 2021-2022시즌 프리미어리그 시청률이다 [1].

표 1. SPOTV 2021-2022 프리미어리그 시청률  
Table 1. SPOTV 2021-2022 Premier League ratings

방송 날짜	시청률(%)
2022년 2월 13일	2.137
2022년 5월 1일	2.763
2022년 5월 13일	1.596
2022년 5월 15일	4.844

이는 해외 프로 축구 리그인 잉글랜드 프리미어 리그에 대한 국내 시청자 및 수요가 확보되어 있다는 것을 입증한다. 본 논문에서는 이러한 국내 프리미어리그 시청자들을 위한 오락 콘텐츠를 제공하는 애플리케이션을 제안하려고 한다.

## II. 연구 주제

프리미어리그 경기 국내 시청자는 한국과 영국

사이의 물리적 거리로 인해 직접 관람이 어려우며, 실시간 중계를 시청하는 경우 시차로 인해 새벽 시간대에 경기를 시청하게 된다. 또한 현재 서비스되는 해외 축구 관련 플랫폼 중 국내 시청자를 타겟으로 하는 모바일 플랫폼은 존재하지 않는다. 이러한 요인들로 인해 프리미어리그 경기 국내 시청자는 경기 관련 콘텐츠를 즐기는 데에 있어 한계점이 존재한다.

프로 축구의 시청 만족도에 영향을 미치는 요인을 분석한 자료[5][6]에 따르면 시청자의 스포츠 관여도를 높이는 부가적인 오락 요소들의 제공이 전체적인 시청 만족감에 긍정적인 영향을 미친다. 또한 프로 축구 경기 관람자의 지각된 가치는 관람 만족도에 긍정적인 영향을 미치고, 지각된 서비스 가치 향상을 위해서는 오락 이벤트 등 경기 결과 외에 시청자가 즐거운 시간을 보내는 것이 더 중요하였다.

이러한 관점에서 부가적인 오락 요소들의 제공을 통해 향상되는 시청자의 관람 만족도가 국내 시청자들의 물리적 거리로 인한 시청 만족도 하락을 완화할 수 있는 요소가 될 거라 추론하였다.

25-35세대들의 최근 트렌드는 '라이프 로깅'이며, 일상의 모든 행동과 정보를 기록으로 남기려는 기술의 움직임이 개인 스마트 기기를 활용하여 개인이 24시간 생활 동안 벌이는 행동 및 신체 상태의 변화를 자동으로 측정, 기록하고 상세히 분석해 주는 '라이프 트래킹'으로 진화하고 있다[7]. 특히 최근 늘어난 여성 축구 팬들의 경우 스포츠 구단과 선수를 향한 강렬한 애정을 보이는 스포츠 팬 문화가 형성되면서, SNS를 통해 다른 팬들과 소통하고, 경기 관련 콘텐츠를 소비, 혹은 직접 생산하려는 사용자의 비율이 높았다[8].

연구 조사를 바탕으로 국내 프리미어리그 시청자의 온라인 중계 시청 만족감을 개선하여, 좋아하는 스포츠 경기를 더 재미있게 즐길 수 있는 공간을 제공하는 서비스를 구현하였다.

해당 프로젝트에서는 사용자가 체험할 수 있는 다양한 경기 관련 오락 콘텐츠를 제공하며, 사용자가 프리미어리그 경기에 대한 자신의 애착도와 감상을 표현할 수 있다. 사용자의 애정도를 애플리케이션 내부 비주얼 요소로 수치화하고, 개인 스마트

위치로 측정된 심박수를 통해 사용자의 경기 중 감정 흐름 기록을 제공한다.

### III. 사용자 조사

본 프로젝트의 구현 단계에 들어가기 전 사용자 가치 및 상세 기능 도출을 위한 사용자 조사를 진행하였다.

#### 3.1 정량 조사 결과

해외 축구 경기 시청 경험이 있는 10-30대 여성 112명을 대상으로 구글 폼을 활용한 온라인 설문 조사를 진행하였다. 다음 그림 1-5는 설문자의 응답 결과 중 주요 항목을 차트로 표현한 결과이다.

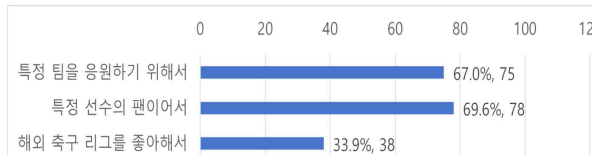


그림 1. 해외 축구 경기 시청 목적

Fig. 1. Motivations for watching international football matches

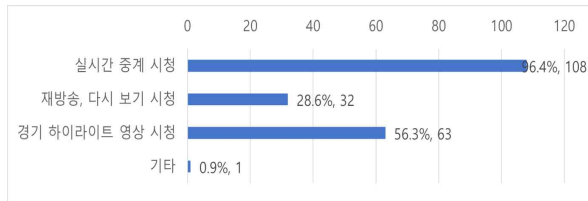


그림 2. 해외 축구 경기 시청 방식

Fig. 2. Methods of watching international football matches



그림 3. 해외 축구 경기 관람 환경(동반인 여부)

Fig. 3. Usual Companions of watching international football matches

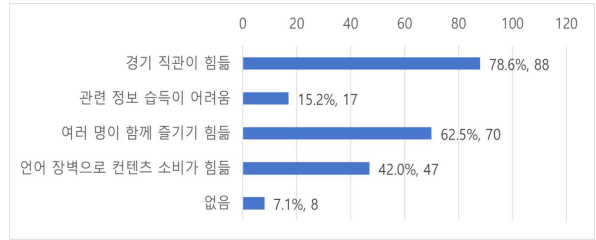


그림 4. 해외 축구 경기 관람 중 느낀 심리적 거리감 및 장벽의 원인

Fig. 4. Causes of psychological distance and barriers experienced while watching international football matches

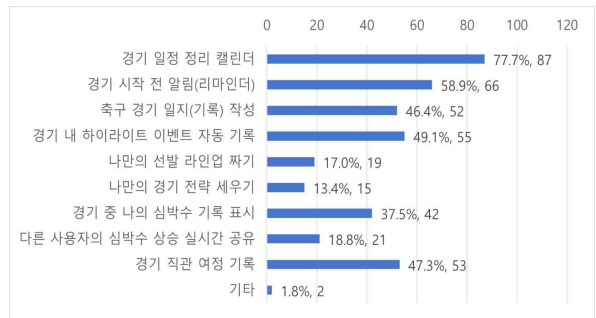


그림 5. 프로젝트에서 제공될 기능 선호도

Fig. 5. Preferences for planned features in the project

#### 3.2 정성 조사 결과

타겟 사용자층의 특성을 통해 선정한 페르소나와 유사한 해외 축구 시청자 2명과 대면 인터뷰를 진행하였다. 인터뷰이 A는 다양한 SNS를 통해 축구 관련 콘텐츠를 작성하고 다른 축구 팬과 소통하며 여러 축구 물품을 구매하는 여성 헤비 시청자(24세, 해외 축구 경기 시청 2년 차)였고, 인터뷰이 B는 어릴 때부터 해외 축구 경기를 시청해 왔으나 단일 SNS로 지인과 소통하며, 축구 관련 품목을 소비하지는 않는 여성 라이트 시청자(22세, 축구 경기 시청 20년 이상)였다.

두 인터뷰이는 경기 중 스마트워치를 통해 심박수를 측정하고 그에 대한 통계 리포트를 제공받는 기능인 '나의 심박수 기록'과 '다른 사용자와의 심박수 공유 기능'에 큰 관심을 보였다. 더불어 경기 내의 하이라이트 이벤트를 자동으로 기록해 보여주는 '이벤트 하이라이트 기능'을 통해 경기를 보면서 놓쳤을 것 같은 부분을 다시 확인할 수 있어 경기 시청 후 응원하는 팀의 팬으로서 몰입도가 향상될 것이라고 응답하였다.

## IV. 프로토타입

### 4.1 프로토타입 제작 방법

본 프로젝트의 주요 타겟층으로 설정한 사용자는 프리미어리그 경기를 시청하고 애정하는 10-30대 여성이다. 본 타겟층은 iOS 운영체제의 사용 비중이 높다[9]. 이에 주요 타겟층을 겨냥하여 iOS 애플리케이션을 플랫폼으로 선정했다.

디자인은 '생생한', '활기찬'을 키워드로 선정하여 스포츠 경기의 역동적인 분위기가 연출되도록 디자인했다. iOS 애플리케이션을 개발하기 위해서 SwiftUI 언어와 Springboot 를 사용했다.

### 4.2 서비스 목표 및 사용자 가치

사전 조사 내용을 바탕으로 국내의 프리미어리그 시청자들이 여러 가지 축구 재미 요소를 체험하고, 자신의 취미를 능동적인 태도로 즐기며 심리적인 행복감을 얻도록 만드는 것을 서비스의 목표로 설정했다.

사용자에게 더 나은 사용자 경험을 제공할 수 있는 기능을 도출하기 위해서 사용자 가치를 선정했다. 첫째, '즐거움 추구'다. 국내 잉글랜드 프리미어리그 시청자 285명을 대상으로 진행한 설문조사에

서 시청 동기 3가지 중 '오락 추구'가 5점 만점에 4.21로 가장 높다[10]. 이는 오락적 재미와 흥미를 얻기 위해 시청하는 경향이 높다는 것을 의미한다. 이를 고려해 즐거움 추구를 가장 중요한 가치로 정했다. 둘째, '흥분 표현'이다. 다음 가치는 타겟 사용자층에서 두드러지게 나타나는 라이프 로깅 트렌드를 반영하여, 사용자가 경기를 시청할 때 느꼈던 강렬한 감정과 선수에 대한 애정을 기록하고, 다른 사용자와 공유하고 싶어 하는 욕구를 충족시키고자 한다. 셋째, '맞춤형 정보'이다. 사용자의 개인화된 정보 획득 욕구를 반영하여 정보 탐색 과정의 편리성을 향상하고자 한다. 이러한 사용자 가치를 기반으로 다양한 기능 아이디어를 구체화하였다.

### 4.3 기능 아이디어

앞서 선정한 사용자 가치를 모두 만족하는 기능을 도출했다. 기능에 대한 자세한 설명은 아래 [표 2]와 같다.

### 4.4 프로토타입 구현 결과

본 연구의 타겟 사용자는 다음과 같은 흐름으로 애플리케이션을 사용한다. 축구 경기 시청 전 경기 예측하기 콘텐츠 활동을 수행하여 다른 사용자와 자신의 예측 결과를 비교한다.

표 2. 사용자 가치에 따른 기능 리스트  
Table 2. Feature list based on user value

심박수 타임라인	사용자가 스마트 워치를 착용한 채 경기를 시청하면, 경기 타임라인 화면에서 경기 이벤트와 사용자의 실시간 심박수를 매치하여 보여준다.
심박수 통계	홈팀 팬, 원정팀 팬과 사용자의 최저·평균·최고 심박수를 꺾은선 그래프로 나타내서, 사용자가 경기를 얼마나 재미있게 즐기고 있는지를 상대 지표로 확인할 수 있다.
선발 라인업 및 우승팀 예측	경기 시작 전, 선발라인업(포메이션, 선수 라인업)과 우승 팀(골 득점 여부) 예측을 하며, 경기에 대한 설렘과 기대감을 오락으로 해소하도록 돕는다.
축구 일기	경기 시청 후, 경기를 시청할 때 느꼈던 감정과 인상적인 순간을 기록하고 공유하며, 자신의 정체성을 강화한다. 특히, 경기에 대한 자신만의 MOM(Man of the Match)을 선정하여 일기 작성에 재미를 더한다.
레벨 시스템	축구에 대한 사용자의 애정을 앱 내 점수로 나타낸다. 경기 예측과 일기 작성 등의 참여로 점수를 획득할 수 있으며, 이는 사용자가 자신의 행동에 대하여 보상을 받으며 즐거움을 느끼고 행동을 반복해서 수행하도록 만든다.
선호하는 팀 선정	웹 크롤링을 통해 경기 일정 및 경기 중계 내용을 수집한다. 사용자가 선호하는 축구팀을 지정하면 관련 이벤트를 선별적으로 필터링하여 제공한다. 또한 축구 일기 화면에서 사용자가 선호하는 팀의 축구 일기를 우선으로 노출한다.

경기 중 사용자는 스마트 워치를 통해 심박수를 측정하고, 경기 하이라이트 중계와 심박수 정보를 동기화하여 제공받는다. 마지막으로 경기 시청 후 심박수의 데이터를 기반으로 축구 일기 기록을 수행한다.

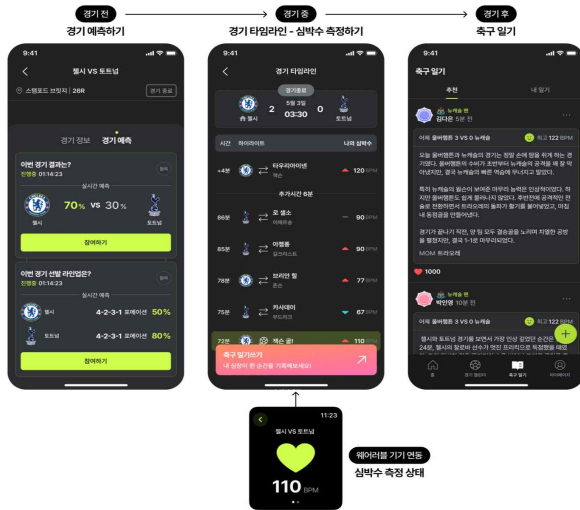


그림 6. 축구 경기 전, 중, 후 애플리케이션 이용 시나리오 GUI

Fig. 6. Scenario GUI for using applications before, during and after a football game

## V. 결론

본 연구는 잉글랜드 프리미어리그 국내 인기 상승에 따른 비대면 시청 환경에서의 관중 만족도 향상을 위한 애플리케이션을 제안한다. 본 애플리케이션을 통해 기대하는 바는 다음과 같다. 첫째, 다양한 경기 관련 오락 콘텐츠를 제공함으로써, 프리미어리그 시청자들의 능동적인 참여를 유도하고 몰입을 극대화하는 역할을 수행한다. 둘째, 프리미어리그 경기에 대한 사용자의 애착도와 다양한 감상을 표현할 수 있는 수단으로, 사용자의 애정도를 수치화한 비주얼 요소와 심박수를 통한 사용자의 감정 흐름 기록 등을 제공한다. 이를 통해 비대면 스포츠 관람 문화의 질적 향상을 기대한다.

## 참고 문헌

[1] Jeremy Wilson. "Premier League is world's favourite league". *The Telegraph*, 06 Nov 2007, [https://www](https://www.telegraph.co.uk/sport/football/2325057/Premier-League-is-worlds-favourite-league.html)

[www.telegraph.co.uk/sport/football/2325057/Premier-League-is-worlds-favourite-league.html](https://www.telegraph.co.uk/sport/football/2325057/Premier-League-is-worlds-favourite-league.html). accessed 28 Oct 2024.

[2] Ryan Benson. "The Strongest Leagues in the World: What the Opta Power Rankings Say". *Opta Analyst*, 14 Oct 2024, <https://theanalyst.com/2024/10/strongest-leagues-world-football-opta-power-rankings>. accessed 28 Oct 2024.

[3] 한국콘텐츠진흥원. *2019년 해외 콘텐츠시장 분석*. 한국콘텐츠진흥원, 2020.

[4] 최효정. "쿠판플레이, 英 프리미어리그 2025-26 시즌부터 6년 독점 중계권 따내". *조선비즈*, 2024년 4월 28일, <https://biz.chosun.com/distribution/channel/2024/04/28/JO52TQNV7RASBEYGJVC4IMXHKE>. 2024년 10월 30일 접속

[5] 안지형. "뉴미디어 스포츠 중계방송 시청이 시청만족 및 직접관람의도에 미치는 영향", 국내 석사 학위논문 이화여자대학교 대학원, 2012.

[6] 손룡, 최효규. "프로축구의 관중요인이 지각된 가치, 관람만족 및 재 관람의도에 미치는 영향", *기업경영리뷰* 1.1 pp. 131-146(2010) : 131.

[7] 김신혜. "라이프로그를 활용한 문화생활 큐레이션 서비스 디자인 연구 - JOYLOG 애플리케이션을 중심으로 - : A Study on Service Design for Cultural Activity Curation through Life-logging with Emphasis on App JOYLOG.", 국내석사학위논문 서울대학교 대학원 디자인학부(시각디자인 전공), 2015.

[8] 조유빈. "'디지털 네이티브' Z세대, 기업이 주목하는 Z트렌드". *시사저널*, 2022년 7월 7일, <https://www.sisajournal.com/news/articleView.html?idxno=241404>. 2024년 10월 29일 접속.

[9] 닐슨코리아 클릭. "Mobile App: IOS vs. Android 5월 이용행태 비교 분석". *Nielsen*, 2019년 6월 27일, <https://www.koreanclick.com/insights/newsletter-view.html?code=topic&id=532>. 2024년 10월 27일 접속.

[10] 정서영. "잉글랜드 프리미어리그 경기 국내 시청자 집단의 특성에 관한 연구.", 국내석사학위논문 건국대학교 언론홍보대학원, 2020.

# 걸음걸이에 의한 스마트폰 소유자 인증 모델의 iOS에서 변환

강동환\*, 최지우\*\*, 강태원\*\*\*, 최상일\*\*\*

## iOS Implementation of smartphone owner authentication model by gait in iOS

Donghwan Gang\*, Jiwoo Choi\*\*, Taewon Kang\*\*\*, and Sangil Choi\*\*\*

### 요 약

본 연구는 걸음걸이 데이터를 활용한 iOS 환경에서의 개인 인증 시스템의 구현 방법론을 제시한다. CNN 모델은 개개인의 걸음걸이를 식별할 수 있도록 다양한 걸음걸이 데이터를 학습하고, 스마트폰 소유자의 걸음걸이를 iOS 기기 내에서 등록하기 위해 Core ML을 통해 모델을 변환하였다. 향후 연구에서는 iOS 환경에서 사용자 등록 및 인증 과정을 구현하여, 걸음걸이 특징을 반영한 사용자 인증 시스템을 완성할 수 있을 것으로 기대한다.

### Abstract

This study presents a methodology for implementing a personal authentication system using gait data on an iOS platform. A CNN model is trained on various gait data to identify individuals based on their walking patterns. The model is converted using Core ML to enroll the smartphone owner's gait on iOS devices. Future research will focus on implementing user enrollment and authentication processes within the iOS environment, ultimately completing a gait-based authentication system.

### Key words

Gait, Authentication, Convolutional neural network, iOS, Smartphone

## 1. 서 론

늘어나는 정보량과 이와 관련해 발생하는 보안 문제의 심각성에 따라 정보 보안의 중요도가 증가하고 있다. 보안 문제에 대응하기 위한 전통적인 방식으로 비밀번호와 패턴 방식이 존재한다. 하지만 이러한 전통적인 방식이 사용자 편의를 추구함으로

인해 보안 약화의 가능성이 제기되고 있다[1]. 이를 개선하기 위해 지금까지 지문, 홍채, 안면 인식 같은 개인이 가지는 고유한 신체적 특성을 사용하여 보안을 강화해왔으며[2], 이에 따라 이제는 데이터 자체의 보안이 중요시되고 있다[3]. 따라서 본 연구는 이처럼 탈취가 힘든 인간의 걸음걸이를 기반으로 하는 스마트폰 인증 시스템[4]을 iOS 디바이스

\* 국립강릉원주대학교, gangdonghwan980@gmail.com

\*\* (주)인터포, cephular@gmail.com,

\*\*\* 국립강릉원주대학교, twkang@gwnu.ac.kr, \*\*\*강릉원주대학교, schoi@gwnu.ac.kr(교신저자)

(iPhone)에 구현하는 방법을 제시한다. 본 시스템은 새로 측정된 적법한 사용자의 생체 데이터를 외부로 전송하지 않고, 기기 내 재학습을 통해 안전하게 인증하는 것을 목표로 한다. 논문의 구성은 다음과 같다. 2장은 이전 연구 [4]를 통해 안드로이드 기반의 사용자 보행 인증 시스템과 모델을 살펴본다. 3장은 iOS에서의 모델 변환을 서술한다. 마지막 4장에서는 결론을 기술하고 논문을 마무리한다.

## II. CNN 기반 스마트폰 걸음걸이 인증 시스템

### 2.1. 시스템 정의 및 기능

시스템의 정의와 구성은 이전 연구 [4]와 동일하다. 스마트폰 걸음걸이 인증 시스템(이하 인증 시스템)이란, 인간의 걸음걸이 데이터를 CNN 기반 딥러닝 모델에 입력하여 스마트폰에 접근하려는 사람의 신원을 확인하는 안드로이드 기반 사용자 인증 시스템이다. 걸음걸이 데이터란, 스마트폰에 내장된 가속도계와 각속도계를 통해 수집된 데이터이다.

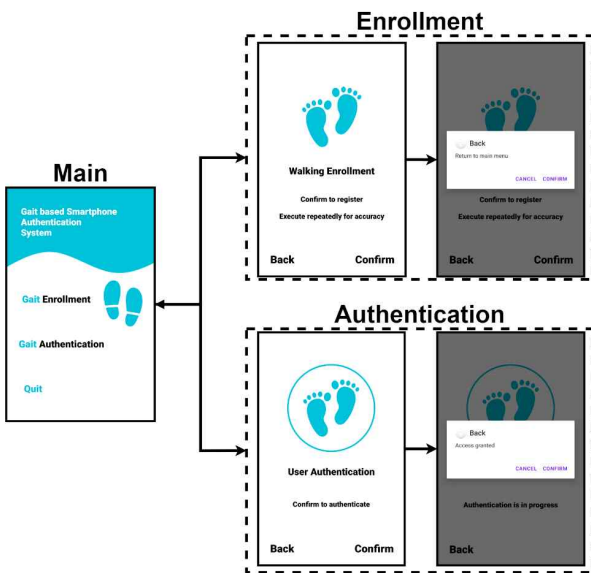


그림 1. 인증 시스템 개요

Fig. 1. Overview of the Authentication System

인증 시스템의 기능은 그림 1과 같이 등록 (Enrollment), 인증(Authentication) 두 가지로 구분된다. 등록 과정에서는 적법한 사용자의 걸음걸이 데이터를 수집하고, 전처리한 후 CNN 기반 사전 학습 모델의 기기 내 재학습을 통해 사용자를 등록한

다. 인증 과정에서는 스마트폰에 접근하려는 사람의 걸음걸이 데이터를 수집하고, 전처리한 후 재학습된 모델에 입력하여 인증 여부를 판단한다.

### 2.2. 변환 전 사전 학습 모델

사전 학습 모델은 TensorFlow 라이브러리를 통해 구성 및 학습된 걸음걸이 인증 모델을 의미하며, 이후 3장에서 apple에서 제공하는 API인 CoreMLTools를 통해 사용 가능한 형태로 변환된다. 사전 학습 모델의 구조는 이전 연구 [4]에서 사용한 모델과 동일하며, 그 세부 구조는 표 1과 같다. 사전 모델의 학습에는 Shimmer IMU 센서와 스마트폰을 통해 다양한 위치에서 수집한 29명의 데이터와 새로운 사용자 등록을 위한 더미 데이터를 사용하였다. 더미 데이터는 기존 사용자 데이터와 무관한 값인 0으로만 구성된 1.8초간의 가속도 및 각속도 데이터를 나타낸다.

표 1. 변환 전 모델 구조

Table 1. Model Structure Before Conversion

No	Layer (type)	No	Layer (type)
1	Conv1D	7	MaxPooling1D
2	MaxPooling1D	8	Dropout
3	Dropout	9	Conv1D
4	Conv1D	10	MaxPooling1D
5	MaxPooling1D	11	Flatten
6	Conv1D	12	Dense

## III. 사전 학습 모델 변환

### 3.1 모델 변환

iOS 네이티브 앱을 개발할 때, 기기 내 학습을 위해서는 Apple에서 제공하는 프레임워크인 Core ML을 사용해야 한다. TensorFlow나 PyTorch와 같은 라이브러리를 사용하여 학습된 모델은 Core ML에서 사용하기 위해 요구 사항에 맞게 변경해야 한다. 변환한 모델은 “.mlmodel” 또는 “.mlpackage” 형식으로 저장할 수 있으며, 기기 내 학습을 위해서는 “.mlmodel” 형식으로 변환해야 한다. 그림 2는 모델 변환을 위한 코드를 보여주며, 표 2는 변환된 모델의 구조를 나타낸다.

```
import coremltools as ct

output_name = "classLabel"
class_labels = [str(i) for i in range(30)]
input_tensor = ct.TensorType(shape=(1, 360, 1), name='conv1d_input')

coreml_model = ct.convert(
    model=keras_model, source="auto",
    inputs=[input_tensor],
    classifier_config=ct.ClassifierConfig(class_labels=class_labels),
    convert_to='neuralnetwork'
)

coreml_model.save('GaitAuthentication.mlmodel')
```

그림 2. 모델 변환 코드  
Fig. 2. Model Conversion Code

표 2. 변환 후 모델 구조  
Table 2. Model Structure After Conversion

No	Layer (type)	No	Layer (type)	No	Layer (type)
1	transpose	17	activationRelu	33	squeeze
2	expandDims	18	expandDims	34	transpose
3	convolution	19	transpose	35	expandDims
4	squeeze	20	pooling	36	convolution
5	transpose	21	transpose	37	squeeze
6	activationRelu	22	squeeze	38	transpose
7	expandDims	23	transpose	39	activationRelu
8	transpose	24	expandDims	40	expandDims
9	pooling	25	convolution	41	transpose
10	transpose	26	squeeze	42	pooling
11	squeeze	27	transpose	43	transpose
12	transpose	28	activationRelu	44	squeeze
13	expandDims	29	expandDims	45	reshapeStatic
14	convolution	30	transpose	46	innerProduct
15	squeeze	31	pooling	47	softmaxND
16	transpose	32	transpose		

### 3.2 모델 업데이트 설정

모델이 기기 내 학습을 지원하려면, 가중치가 변경 가능하도록 설정해야 한다. 설정 조건은 세 가지이다. 첫째, 업데이트가 가능한 레이어 타입이어야 하며, 업데이트 가능한 레이어의 타입은 "innerProduct", "convolution"으로 제한된다[5]. 여기서 "innerProduct"레이어는 Fully connected layer를 의미한다. 둘째, 마지막 활성화 함수 타입이 "softmax"여야 한다. 셋째, 출력층에부터 업데이트 설정 레이어까지 모든 레이어는 오류 역전파(backpropagation) 학습이 가능한 레이어여야 한다.

TensorFlow로 학습된 사전 학습 모델에서는 마지막 활성화 함수로 softmax를 사용했으나, 표 2를 보면, 변환된 모델에서는 활성화 함수 타입이 "softmaxND"

로 변환되었음을 알 수 있다. 이에 따라, 출력층 직전의 활성화 함수 "softmaxND"를 삭제하고 활성화 함수 "softmax"로 대체했다. 이후, 업데이트 가능한 레이어를 선택하고, 손실 함수(loss function), 옵티마이저(optimizer), 학습률(learning rate), 배치 크기(batch), 에포크(epoch) 등의 설정을 완료하면 iOS 기기에서의 학습이 가능해진다.

## IV. 결론

본 연구는 걸음걸이 데이터 기반 CNN 인증 모델을 iOS 디바이스에서 학습할 수 있도록 변환하는 방법론을 제시하였다. 향후 연구에서는 iOS 환경에서의 사용자 등록 및 인증 과정을 구현하여, 걸음걸이 특징을 반영한 사용자 인증 시스템을 완성할 수 있을 것으로 기대된다. 또한 이전 연구에서 개발된 안드로이드 기반 CNN 걸음걸이 인증 시스템과의 정확도를 비교 분석할 수 있을 것이다.

## 참고 문헌

- [1] A.P. Umejiaku, P. Dhakal, and V.S. Sheng, "Balancing Password Security and User Convenience: Exploring the Potential of Prompt Models for Password Generation", Electronics, no. 10:2159, 2023.
- [2] S. Sett and H. Gupta, "A Biometric Security Model for The Enhancement of Data Security", 2024 11th ICRIITO, pp. 1-5, 2024.
- [3] W. Yang, S. Wang, J. Hu, G. and Zheng, "Security and Accuracy of Fingerprint-Based Biometrics: A Review", Symmetry, vol. 11, no. 141, 2019.
- [4] J. Choi, S. Choi, and T. Kang, "Smartphone Authentication System using Personal Gaits and CNN", Journal of KIIT, vol. 20, no. 10, pp. 71-82, 2022.
- [5] coremltools API Reference, v8.0b1, (2024, Oct), [https://apple.github.io/coremltools/\\_modules/coremltools/models/neural\\_network/builder.html](https://apple.github.io/coremltools/_modules/coremltools/models/neural_network/builder.html).



# 노이즈 합성을 통한 게임 맵 생성

원민철\*, 김완규\*, 이승교\*, 이부형\*\*

## Game Map Generation Using Noise Combination

Won MinCheol\*, Kim WanKyu\*, Lee SeungGyo\*, and Boo-Hyung Lee\*\*

### 요약

본 논문에서는 게임에서 실제 플레이어가 활동하는 가상공간인 게임 맵을 구현하고 구현된 맵이 개발자의 의도에 따라 적합하게 생성되었는지를 분석한다. 먼저 게임맵을 구현하기 위해 노이즈기반 맵생성기법을 사용하는데 본 논문에서는 기본적인 퍼린 노이즈와 원하는 지형을 생성하기 위한 노이즈를 결합한 후 메쉬맵 생성 과정을 거쳐 2D 노이즈를 3D지형으로 변환하여 최종 환경의 맵을 완성한다. 그 후 사논지수를 이용해 지형분포와 평균경사등을 체크하여 최종 완성된 맵을 분석한다. 실험에서 노이즈합성을 통해 3D 메시맵을 구현하였으며 지형속성분류 그래프를 통해 맵 내 지형속성이 제대로 구현되었는지 확인하였으며 원하는 맵이 완성됨을 보였다.

### Key words

Game map, Unity, Perlin noise, Performance Evaluation

## I. 서론

게임 맵(Map)은 플레이어가 실제로 활동하는 게임의 가상 공간을 의미한다[1]. 맵은 캐릭터의 이동 경로와 환경 요소, 지형, 그리고 장애물을 포함하며, 게임의 스토리와 플레이 경험에 필수적인 배경을 제공한다. 이러한 맵을 수작업으로 제작하는 경우, 높은 품질을 보장할 수 있지만 개발 시간과 비용이 많이 들기 때문에 대부분 절차적 생성 알고리즘을 이용하여 자동적으로 생성하므로써 제작효율을 향상시킨다[2].

노이즈 맵(Noise Map)은 절차적 맵 생성 알고리즘에서 가장 기반이되는 개념으로, 맵 생성의 무작위성과 지형간의 자연스러운 변화를 구현하지만

인접 점 간 변동이 크기 때문에 연속적이거나 자연스런 지형 표현에 적합하지 않다.

본 논문에서는 노이즈맵을 변형시킨 퍼린노이즈와 지형을 생성하기 위한 노이즈들을 합성하여 최종 3차원 맵을 완성한 후 결과를 분석한다.

## II. 노이즈 맵(Noise Map)

노이즈 맵은 절차적 맵 생성의 가장 기반이 되는 요소로 그림 1과 같이 normal noise로 표현한다. 그림 1의 노이즈는 인접한 셀과 변화가 크기 때문에 연속적인 지형을 표현하기에 부적절하므로 그림 2와 같이 퍼린노이즈를 사용하여 자연스러운 맵을 구현한다.

\*국립공주대학교 공과대학 소프트웨어학과, mnmkjin2896@gmail.com,  
\*\*국립공주대학교 공과대학 소프트웨어학과 교수, bhl1998@kongju.ac.kr(교신저자)

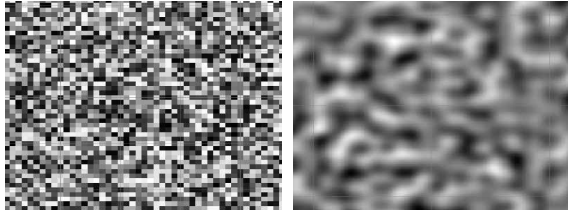


그림 1. Normal Noise      그림 2. Perlin Noise

펄린 노이즈의 주파수와 진폭을 조정함으로써 다양한 환경에 맞는 지형을 생성할 수 있지만 좀 더 정교하거나 다양한 지형을 구현하기 위해서는 노이즈 합성기법을 사용할 수 있다. 그림 3은 Cone Noise로 중간에서는 밝고 가장자리로 갈수록 어두워지는 특징을 갖는 노이즈이다. 그림 2의 펄린 노이즈에 그림3의 노이즈를 합성하면 그림 4와 같은 펄린 노이즈가 변화된 합성 노이즈 맵을 만들 수 있다.

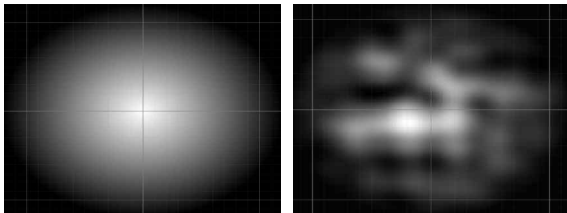


그림 3. Cone Noise      그림 4. 합성 Noise

### III. 맵생성 및 분석

본 장에서는 2장에서 제안한 노이즈 합성을 통해 게임 맵을 구현하였다. 그림 5는 노이즈 합성전의 그림 2의 펄린 노이즈맵에서 얻어진 값을 활용하여 1차적으로 구현한 맵이다. 펄린 노이즈 맵 각 점의 지형 속성과 고도 정보를 시각적 요소로 변환한 후 사전에 정의된 범위에 따라 색상으로 구분하여 산, 바다, 모래 및 눈(snow)등으로 표현한다. 그림 6은 그림 5의 맵에 그림 3의 cone 노이즈를 합성한 결과이다. cone noise특성에 따라 가장자리로 갈수록 노란색인 모래와 파란색인 물이 확장됨을 보인다. 그림 7은 그림 6의 합성 맵 내 지형속성을 색상별로 분류한 그래프이다.

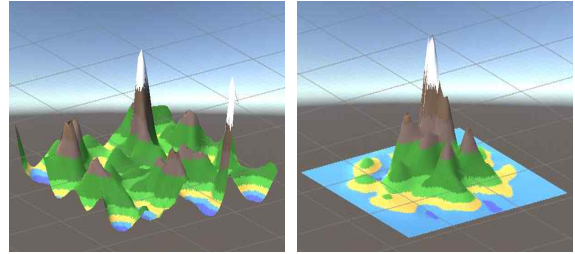


그림 5. 노이즈합성후 맵      그림 6. 노이즈합성후 맵

그림에서 숲을 표현한 진한녹색, 연한녹색이 50%이상, 바위를 표현한 갈색이 10%, 노란색으로 표현한 모래가 14%를 차지함을 알 수 있다.

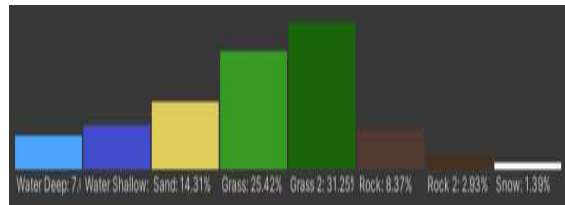


그림 7. 맵 내 지형속성분류 그래프

이와 같이 지형 속성 분류그래프를 이용하여 개발자 의도에 따라 지형속성이 제대로 구현되었는지 확인할 수 있으며 다양한 노이즈와 합성하므로써 지형속성을 원하는 대로 변경시킬 수 있다.

### IV. 결론

본 논문에서는 노이즈맵을 변형시킨 펄린노이즈와 지형을 생성하기 위한 cone 노이즈를 합성하여 최종 3차원 맵을 완성한 후 생성된 맵 내 지형속성과 지형속성의 점유율을 확인할 수 있었다. 이와 같이 노이즈합성을 통한 맵을 구현함으로써 인디 게임 개발자들이 더욱 풍부하고 몰입감 있는 게임 환경을 구축하는 데 기여할 수 있을 것이다.

### 참고문헌

- [1] <https://github.com/Auburn/FastNoiseLite>
- [2] 조별희, 이부형 “다양한 지형을 가진 다각형 랜덤 자연맵 생성 방법”, 2022한국정보기술학회 추계 종합학술대회 논문집, pp. 677-679, 2022

# 위치 기반 지역 모임 연결 웹 서비스

정재우\*, 송재훈\*, 김상욱\*, 박기범\*, 김진호\*, 안진호\*\*

## Location-based Local Club Connection Web Service

Jaewoo Jung\*, Jaehoon Song\*, Sangwook Kim\*, Gibum Park\*, Jinho Kim\* and Jinho Ahn\*\*

### 요약

기존의 커뮤니티 플랫폼들은 사람들을 찾는 편의성을 제공하였지만, 게시물 속에서 원하는 내용들을 일일이 찾아야 하는 불편함이 존재했다. 이 논문에서는 기존의 문제점들을 개선한 위치 기반 모임 서비스를 소개한다. 이를 통해 사람들이 반복적이고 소모적인 사회에서 벗어나, 진정으로 의미 있는 사회적 교류와 자기 계발에 참여할 수 있는 환경을 제공하고자 한다.

### Abstract

Existing community platforms have offered the convenience of finding people but presented the inconvenience of having to search individually through posts for desired content. This paper introduces a location-based gathering service that addresses these issues. The aim is to provide an environment that allows individuals to escape repetitive and exhausting social interactions, facilitating meaningful social engagement and personal development.

### Key words

location-based service, social interaction, filtered result, personal suggestion, manageability

### 1. 서론

1991년 이후로 본격적인 지방자치화 이루어지면서 주민자치의식의 함양 및 지역사회에 대한 관심이 높아짐으로 인해 주민 참여에 대한 관심 또한 증대되었다. 또한 90년대 후반부터는 한국 전역에 PC와 초고속 인터넷이 보급되면서 인터넷 이용자 숫자가 증가하게 되었고, 사회적 상호작용 공간으로서 사이

버 공간의 역할이 점점 커지고 있다.[1] 현재도 많은 사람들이 인터넷을 통해 보이지 않는 연결망으로 구축된 온라인 커뮤니티를 통해 지속적인 사회적 관계를 맺으며 이러한 관계를 통해 사회적 지지를 교환하고 있다.

하지만 한국보건사회연구원 조사결과에 따르면 50% 이상의 국민이 지역갈등이 심각하다고 인식하고 있

\* 경기대학교 AI컴퓨터공학부 {awfjol2008,dlffpqms4,sangwook\_88,rlqja530,jinho829}@kyonggi.ac.kr

\*\* 경기대학교 AI컴퓨터공학부 교수, jhahn@kyonggi.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음 (2021-0-01393).

다. 현대 사회에서 지역 사회의 연대감은 점차 약해지고 있다는 것이다. 국가 차원 뿐만 아니라 지역 단위에서 사회 통합은 중요한 의미를 지니며 지역은 주민의 생활 환경으로 이들의 삶의 질과 행복의 근간이 된다.[2] 이러한 이유로 지역 커뮤니티 활성화를 위한 통합 플랫폼 '모동(모여봐요, 동네에서)' 서비스를 기획하게 되었다.

## II. 서비스 효용

본 연구에서 제안하는 서비스는 현대인의 다방면적 자기 개발 욕구를 충족시키기 위해 기획된 지역 기반 취미 모임 검색 플랫폼이다. 기존의 플랫폼들이 단순한 키워드 검색 및 게시글 형식에 의존하여 불편함을 초래하는 반면, 본 서비스는 다양한 필터링 기능을 제공하여 사용자 경험을 개선하였다.

지역 기반 검색: 사용자의 현재 위치를 기반으로 근거리에서 진행되는 모임을 우선적으로 보여주어 편리성을 극대화 한다.

다양한 옵션: 취미 종류, 날짜, 연령대, 성비, 회비 등 세부적인 조건을 설정할 수 있어 사용자가 원하는 모임을 정확하게 찾을 수 있다.

맞춤형 추천: 사용자의 검색 및 참여 이력을 분석하여 선호에 맞는 모임을 추천한다.

원스톱 모임 관리: 모임 개설자들이 손쉽게 모임을 모집하고 관리할 수 있는 기능을 제공한다. 일정 관리, 참가자 확인, 공지사항 전송 등이 통합된 하나의 시스템을 통해 이루어진다.

이러한 기능들은 사용자들이 자신의 시간과 노력을 절약하며, 보다 효율적으로 자기 개발 활동을 할 수 있도록 돕는 것을 목표로 한다. 궁극적으로, 이 플랫폼은 지역 내 다양한 커뮤니티를 활성화하고 사람들 간의 연결을 강화함으로써 사회적 가치를 창출할 것으로 기대된다.[3]

## III. 기술적 구현

해당 연구에서는 많은 사용자의 모임 및 동아리 개

설과 참가가 가능하도록 하는 서비스를 안정적으로 유지하기 위해, 높은 동시접속 요청 및 사용자 제어, DB 테이블 조회를 효율적으로 처리하고자 한다. 이를 위해 로드 밸런싱과 데이터베이스 파티셔닝을 사용하여 클라이언트 트래픽을 분산시키고 데이터베이스 접근을 최적화한다.

### 3.1 Blue-Green 배포를 통한 무중단 배포

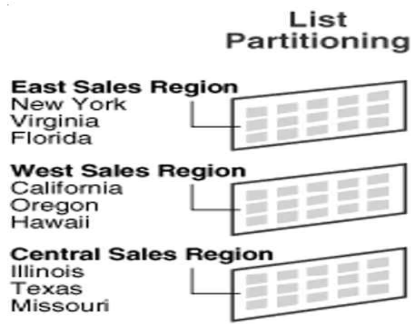
본 서비스는 blue와 green 두 개의 Spring Boot 서버를 Docker의 컨테이너를 통해 독립적으로 관리하며, 새 버전은 spring-green에 배포 후 서버 안정성 테스트 또는 헬스 체크를 거쳐 spring-blue에서 spring-green으로 기존 트래픽을 전환한다. 이를 통해 사용자는 서버가 재배포 되더라도 서비스를 중단 없이 이용할 수 있게 된다.

### 3.2 로드 밸런서를 통한 트래픽 분산

Web Application Server에 대한 높은 동시접속 요청을 효율적으로 처리하고 안정적인 서비스를 제공하기 위해 AWS 로드 밸런서를 통해 사용자 요청을 각기 다른 서버에 고르게 분산한다. 이 로드 밸런서는 실시간으로 사용자 트래픽을 분산시켜 서버 과부하를 방지하며, 다수의 사용자가 동시에 접근할 때도 성능 저하 없이 안정적인 응답을 제공할 수 있게 된다.

### 3.3 DB 파티셔닝

데이터베이스는 MySQL을 기반으로 하며, 사용자 및 그룹 데이터를 기준으로 파티셔닝하여 저장한다. [그림 1]과 같은 List 파티셔닝은, 물리적으로 지정된 값 목록에 따라 데이터가 분할된다. 이를 통해 조회하려는 인스턴스의 접근 속도가 최적화되며, 다수의 요청을 병렬로 처리할 수 있어 처리할 데이터의 양이 많은 환경에서도 높은 성능을 유지할 수 있게 된다.



[그림 1] 리스트 파티셔닝  
[Fig 1] List Partitioning

[4] Chun, S. (2008). Sociology of Self-Help : What Alternatives Can We Advise Other Than Self-Help? Korean Journal of Cultural Sociology, 17(2), 103-140.

#### IV. 결 론

오늘날 현대인은 다각적인 자기 개발과 사회적 연결망을 강화하는 데 큰 관심을 보이며 자기개발은 일시적인 유행을 넘어 하나의 사회 현상으로 자리매김하였다.[4] 그러나 현재의 플랫폼들에는 사용자들이 적절한 활동을 찾고 참여하는 데 있어 여러 불편함이 존재한다. 본 연구에서 제안하는 시스템은 위치기반 검색, 필터링 옵션, 원스톱 모임관리를 통해 불편함을 해소하여 사람들 간의 네트워크를 강화하여 지속 가능한 사회적 관계를 형성하는 데 기여할 것으로 보인다.

#### 참 고 문 헌

[1] Park, Y. ri, Koh, J. hwan, Ahn, H. june, Seo, C. wan, & Kim, G. han. (2009). Development of a Web-Based PPGIS Prototype for Community Reeneration Project Support. Korea Spatial Information Society, 17(2), 159-169.

[2] Cha, M., Lim, E., Kim, H., Yun, Y., Lee, H., Kang, S., Jeon, J., & Park, S. (2011). Strategies to Enhance Social Cohesion for Regional Development: Developing Workable Indicators (2nd ed., Issue 22). Korea Research Institute for Human Settlements, KRIHS.

[3] Martin, K. (2020). Strategies to Enhance Social Cohesion for Regional Development: Developing Workable Indicators.

# 생각모아: 중학교 여성 청소년의 정서적 안정을 위한 인지행동치료 기반 AI 생각 기록 일기 앱 서비스

이윤아\*, 김진영\*\*, 서가은\*\*, 이정연\*\*, 박수이\*\*, 최종인\*\*

## Collecting Thoughts: An AI-Based Cognitive Behavioral Therapy Thought Recording Diary App Service for the Emotional Stability of Female Adolescents in Middle School

YoonA Lee\*, JinYoung Kim\*\*, GaEun Seo\*\*, JeongYeon Lee\*\*, SuE Park\*\*, and JongIn Choi\*\*

### 요 약

최근 10대 우울증과 불안장애 환자가 증가함에 따라 청소년의 부정적 정서 개입의 필요성이 커지고 있다. 이에 중학교 여성 청소년의 정서적 안정을 지원하는 인지행동치료(CBT) 기반 앱 서비스 '생각모아'를 제안한다. 생각모아는 사용자가 하루 동안의 상황, 사고, 감정을 기록하고, 이러한 사고를 분석하여 인지적 오류와 대안적 사고를 제공함으로써 자기 성찰을 돕는다. DTR과 인지적 오류 이론을 검토해 서비스의 페르소나, 가치, 기능, 컨셉, 유저 시나리오를 도출하였으며, 프로토타입을 중학생 사용자 대상으로 테스트하였다. 본 연구는 중학생을 위한 CBT 기반 서비스라는 새로운 관점을 제안하고, AI를 활용한 인지 재구성에서 차별성을 지닌다.

### Abstract

Recently, there has been an increase in cases of depression and anxiety disorders among adolescents, highlighting the growing need for interventions targeting negative emotions. In response, we propose an AI-based cognitive behavioral therapy (CBT) app service called 'Collecting Thoughts' to support the emotional stability of female middle school students. This app allows users to record their daily situations, thoughts, and feelings, and analyzes these inputs to provide insights into cognitive errors and alternative perspectives, thereby facilitating self-reflection. The study reviews the Dysfunctional Thought Record (DTR) and theories of cognitive errors to derive the app's persona, value, feature list, concept, and user scenarios, and a prototype was tested with middle school students. This research presents a new perspective on CBT-based services tailored for middle school students and distinguishes itself through the use of AI in the cognitive restructuring process.

### Key words

중학생, 청소년, 인지행동치료(CBT), 역기능적 사고 기록지(DTR), 인지적 오류, 인공지능(AI), 정서적 안정

---

\* 서울여자대학교 교육심리학과, thisisprofita@gmail.com,

\*\* 서울여자대학교 디지털미디어학과, etesoleil@naver.com, rkdmstdkdkssu@gmail.com, leeje30400@gmail.com, spark44@swu.ac.kr, funtech@swu.ac.kr(교신저자)

## I. 서론

### 1.1 연구 배경 및 목적

최근 정서적 어려움을 경험하는 10대 청소년이 증가하고 있다. 건강보험심사평가원의 최근 5년(2017~2021)간 우울증과 불안장애 진료현황 분석 결과에 따르면, 10대 우울증 환자 수는 2017년 3만 273명에서 2021년 5만 7587명으로 90.2% 증가했고, 불안장애 환자 수는 2017년 1만 7763명에서 2021년 3만 1701명으로 78.5% 상승했다[1]. 청소년기의 우울과 불안은 청소년의 학교생활부적응(교사관계, 교우관계, 학습활동관계, 학교규칙준수관계)에 부정적인 영향을 주며[2], 우울과 불안이 클수록 자살 생각도 증가하는 것으로 나타났다[3].

청소년의 부정적 정서에 대한 효과적인 개입 방안 중 하나는 인지행동치료(Cognitive Behavioral Therapy: CBT)이다. CBT는 부적응적 감정과 행동이 부적응적 사고에서 비롯된다고 보며, 사고를 변화시킴으로써 부적응적 감정을 변화시키는 것을 목표로 한다[4]. 최근 사용자가 생각을 기록할 수 있는 CBT 관련 앱이 출시되고 있지만, 성인을 대상으로 개발되어 청소년이 사용하기에는 한계가 있다.

### 1.2 연구 범위 및 방법

본 연구는 중학생 청소년이 정서적으로 안정될 수 있도록 돕는 CBT 기반 사고 기록 서비스 ‘생각모아’를 제안한다. CBT의 인지재구조화는 사고, 감정, 행동 간의 관계를 파악하고, 내담자가 가진 부정적 자동 사고와 인지적 오류 유형을 확인하는 기법이다. 인지재구조화는 내담자가 부정적 자동 사고의 내용을 현실적이고 타당한 것으로 대체시키고, 그에 따른 감정과 행동의 변화를 살펴보게 한다[4]. 생각모아는 사용자가 자신의 상황, 생각, 감정을 분류하여 기록하게 하고, AI를 활용해 사용자의 부정적 자동 사고 속 인지적 오류 유형과 이에 대한 대안적 사고를 제시한다. 이에 본 연구에서는 CBT와 중학생 청소년에 대한 문헌 조사를 진행하고, 이를 바탕으로 페르소나와 사용자 가치를 설정하였다. 이후 사용자 시나리오를 토대로 중학생 청소년 1인에게 사용성 평가와 인터뷰를 진행하였다.

## II. 이론적 배경

### 2.1 역기능적 사고기록지

인지치료의 효과를 얻기 위해선 상황, 생각, 감정의 연결을 만드는 것이 중요하다. 그러나 내담자는 자신의 문제를 논의할 때 이 세 가지 요소를 명확하게 분리하여 인식하기 어려워하는 경우가 많다[5]. 이때, CBT의 핵심 요소인 역기능적 사고기록지(Dysfunctional Thought Record: DTR)를 활용하면 내담자가 상황, 사고, 감정을 효과적으로 분리하는 방법을 배울 수 있다[6].

DTR은 내담자가 자신의 인지와 그에 대한 반응을 평행하게 기록함으로써 인지를 검토하고 수정할 기회를 제공한다. 이를 통해 내담자는 자신의 인지와 그에 따른 합리적인 반응을 기록하고, 감정, 행동, 인지에 선행하는 사건의 구체적인 내용을 정리할 수 있다. DTR은 인지치료의 보조 치료로서 유용하며, 지속적이고 반복적인 훈련을 위한 활동 과제로 자주 활용된다[6].

### 2.2 인지적 오류 유형

내담자가 부적응적 감정을 경험할 때, 그들의 사고에는 인지적 오류가 나타날 수 있다. 인지적 오류란 현실을 제대로 지각하지 못하거나 그 의미를 왜곡하는 사고의 왜곡을 의미한다[3]. 이러한 인지적 오류를 인식하고 교정하는 방법을 가르치는 것은 내담자가 근거 기반의 융통성 있는 정보 처리 능력을 개발하는 데 도움을 준다. 현장의 치료자는 내담자에게 인지 왜곡의 목록을 제공하고, 그들의 사고에서 자주 나타나는 왜곡을 찾도록 유도하는 방식을 사용할 수 있다[4].

## III. 사용자 조사

청소년 초기의 중학생은 급격한 신체적 변화와 정서적 변화로 인해 작은 자극에도 강하게 반응하는 불안정한 모습을 보인다[7]. 불안, 화, 공포와 같은 정서에 대한 조절 능력은 초등학교에서 중학교로 올라갈수록 떨어지기 때문에 중학생은 정서조절에 더 큰 어려움을 겪을 수 있다[8].

따라서 청소년이 부정적 정서를 느낄 때 나타나는 인지적 오류를 조절하는 것은 중요하다. 중학생의 스트레스를 가정환경, 친구, 학업, 학교 및 교사의 4가지 하위영역으로 분석한 결과 가정환경, 친구, 학업 스트레스와 인지적 왜곡의 조절효과가 유의미하게 나타났다[9]. 따라서 본 연구는 정서 조절에 어려움을 겪고, 인지적 오류로 인해 부정적인 정서를 느끼는 중학생 여성 청소년을 주요 사용자로 선정해 표 1과 같이 사용자 페르소나를 설정했다.

표 1. 사용자 페르소나

Table 1. User Persona

기본 정보	나서진, 여성, 중학교 1학년, 13세
사고	한 가지 안 좋은 일이 있으면 그것만 생각하느라 좋은 것들은 생각하지 못한다.
정서	감정 기복이 심해 최근 정서 조절에 어려움을 겪고 있다.
스토리	반 배정과 학업 문제로 인한 스트레스가 많다. 친구들의 사소한 말에도 상처를 받지만 소심한 성격 탓에 혼자 속상해한다.
목표	사소한 일에도 쉽게 우울해지는 것 같아 힘들고, 감정을 잘 관리하고 싶다.

#### IV. 서비스 컨셉 및 사용자 시나리오

사용자 조사 및 페르소나를 바탕으로 선정한 사용자 가치는 ‘정서적 안정’이다. 정서적 안정이란 사용자가 정서적으로 안정되어 있고자 하는 욕구를 지님을 의미한다. 해당 사용자 가치를 달성하기 위해 표 2의 핵심 기능을 설계하였다.

표 2. 핵심 기능 리스트

Table 2. Core Feature List

하루 기록	상황 적기, 생각 적기, 감정의 심한 정도 선택, 감정의 종류 선택의 4단계를 통해 하루를 기록한다.
하루 요약	사용자가 상황, 생각, 감정을 혼합하여 적은 경우 시가 각각을 분류한 결과와 그 이유를 제공한다.
인지적 오류 유형	시를 활용해 사용자의 생각이 12가지 인지적 오류 유형에 해당할 경우, 해당 유형과 그 이유를 알린다.
대안적 사고	시를 통해 인지적 오류에 해당하는 생각을 적응적으로 수정한 결과를 제공한다.

서비스의 가치를 전달하기 위한 서비스 컨셉은 다음과 같다. 서비스의 이름인 생각모아는 일상 속 생각을 모은다는 의미를 지닌다. 생각모아는 우주를 배경으로 하여, 사용자의 불안정한 정서가 안정적으로 변화되는 과정을 두렵지만 포근한 우주에 비유한다. 그림 1은 서비스를 이끄는 메인 캐릭터인 ‘모아’이다. 모아는 포근한 말투로 사용자의 앱 사용 과정을 안내하고, 분석 결과를 제공한다.

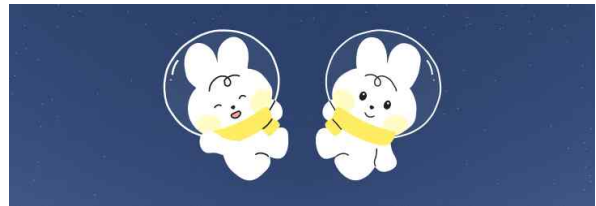


그림 1. 메인 캐릭터  
Fig. 1. Main Character

사용자의 생각 속 12가지 인지적 오류 유형은 12가지의 행성 캐릭터로 표현하였다. 인지적 오류 유형의 특성을 기반으로 캐릭터의 이름과 설명을 쉽게 서술하여 각 유형에 대한 청소년의 이해도를 높이고자 하였다. 그림 2는 인지적 오류 캐릭터들의 이미지이고, 표 3은 인지적 오류 캐릭터에 대한 설명이다.

흑백성	재앙성	외면성	느낌성
이름성	과장성	부분성	궁예성
일반화성	내탓성	해야해성	어둠성

그림 2. 인지적 오류 캐릭터 이미지  
Fig. 2. Images of Cognitive Error Characters



표 3. 인지적 오류 캐릭터 설명

Table 3. Descriptions of Cognitive Error Characters

인지적 오류	캐릭터명 및 설명
흑백논리	[흑백성] 모든 일을 두 가지로만 나눠서 생각하게 해요.
재앙화	[재앙성] 별다른 이유 없이 미래를 부정적으로 생각하게 해요.
긍정적인 면의 평가절하	[외면성] 자신의 좋은 면을 못 보고 스스로를 낮춰 생각하게 해요.
감정적 추론	[느낌성] 자신의 느낌이 틀림없는 사실이라고 생각하게 해요.
명명하기	[이름성] 자신이나 다른 사람에게 부정적인 결론이 담긴 이름을 붙이게 해요.
과장 및 축소	[과장성] 부정적인 것은 훨씬 크게, 긍정적인 것은 훨씬 작게 생각하게 해요.
정신적 여과	[부분성] 전체가 아닌 한 가지 작은 부분을 계속 생각하게 해요.
독심술	[궁예성] 다른 사람의 생각을 마음대로 넘겨짚게 해요.
지나친 일반화	[일반화성] 몇 가지 일로 모든 일에 부정적인 결론을 짓게 해요.
자기 탓	[내탓성] 모든 일을 자신의 탓으로 돌리게 해요.
당위 진술	[해야해성] 자신이나 다른 사람이 반드시 어떡해야 한다고 정해두게 해요.
터널 시야	[어둠성] 어떤 상황의 부정적인 면만을 보게 해요.

서비스 컨셉을 바탕으로 생각모아의 사용자 시나리오를 작성하였다. 생각모아에 가입하면 메인 캐릭터인 모아가 채팅 방식의 온보딩을 통해 상황, 사고, 정서 간의 관계를 설명한다. 사용자가 하루에 한 번씩 하루 동안의 상황, 사고, 감정의 심한 정도, 감정의 종류를 기록하면 AI가 해당 내용을 요약하여 제공한다. 이후, 사용자가 기록한 사고를 분석하여 사용자에게 나타나는 인지적 오류 유형을 알려주고, 사용자의 사고를 적응적인 대안적 사고로 변경하여 제공한다. 마지막으로, 사용자는 감정의 심한 정도가 개선되었는지 기록하고, 변화된 상황, 사고, 감정의 흐름을 확인한다.

## V. 사용성 평가

사용자 시나리오를 바탕으로 중학교 1학년 여학생 1인을 대상으로 2024년 10월 28일 사용자 테스트를 진행했다. 화면 UI 구성 툴인 Figma로 제작한

화면과, ChatGPT API를 활용해 사용자가 하루를 기록하고 인지 왜곡 유형을 확인하는 과업을 수행하게 하며 인터뷰를 진행하였다.

테스터는 홈 화면의 캐릭터를 보고 호감을 보였다. 우주 컨셉의 배경 화면을 본 뒤엔 새벽에 적고 잘 것 같다고 하면서 다음과 같은 반응을 보였다.

“귀여워요! 다른 애들도 좋아할 것 같아요. 새벽 감성 느낌? 새벽에 적고 잘 것 같은 느낌이에요.”

그림 3은 사용자의 상황과 생각을 기록하는 화면이다. 상황을 기록하는 화면에서는 어떤 일이 있었는지 적게 되고, 생각을 기록하는 화면에서는 그때 들었던 생각을 작성하게 된다. 제안 앱의 테스터는 사용 이후 다음과 같은 반응을 보였다.

“오늘 학교에서 어떤 일이 있었는지 쓸 것 같아요. 학교에서 속상했던 일을 쓸 것 같아요.”, “그냥 그때 들었던 것을 쓸 것 같아요.”

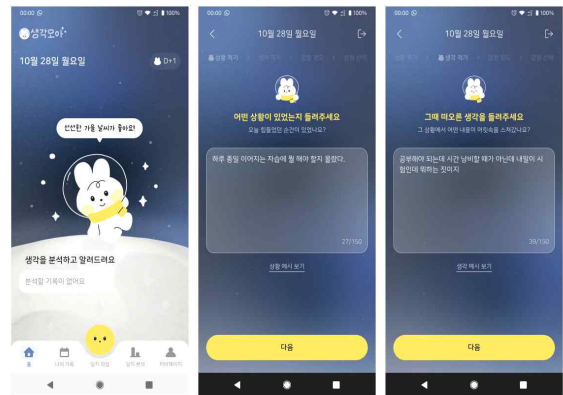


그림 3. 상황 및 생각 기록 화면

Fig. 3. Situation and Thought Recording Screen

그림 4는 사고 기록을 분석해 인지적 오류 유형을 보여주는 화면이다. 테스터는 당시의 생각을 돌아볼 수 있고, 자신을 위한 해결 방법을 알 수 있다는 점에서 긍정적인 반응을 보였다. 다만 방법을 알려줄 때 따뜻한 문구도 있었으면 좋겠다며 다음과 같은 반응을 보였다.

“그냥 내가 생각하고 보통 하던 생각인데 ‘내가 이렇게 생각했구나.’라고 알게 되는 것 같아요. 보고 아차 싶었어요. 그리고 방법을 알려주니까 좋아요. 그런데 저는 (MBTI) F라서 위로나 따뜻한 말도 있으면 좋을 것 같아요.”

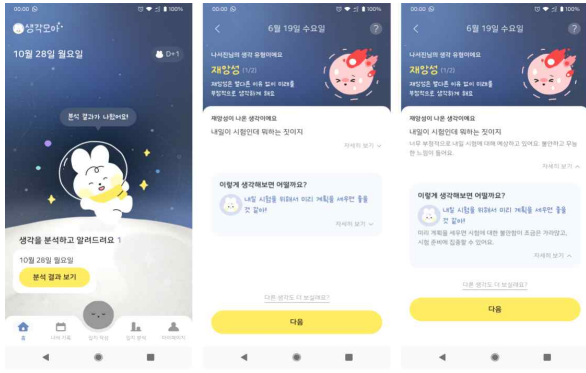


그림 4. 인지 오류 유형 화면  
Fig. 4. Cognitive Error Types Screen

## VI. 논의 및 결론

본 연구는 중학생 청소년이 겪는 정서적 어려움을 해결하기 위해 CBT 기반 AI 사고 기록 일기 앱 서비스 생각모아를 제안하였다. 이를 위해 DTR과 인지적 오류 유형 이론을 검토하고 문헌 연구 결과에 기반해 서비스의 페르소나, 가치 및 기능, 컨셉, 유저 시나리오를 도출하였다. 최종적으로 제작된 프로토타입을 활용해 중학생 사용자를 대상으로 테스트를 진행하였다.

본 연구는 중학생에게 맞춰진 CBT 기반 서비스라는 새로운 관점을 제안했다는 의의를 지닌다. 특히 AI를 인지재구조화를 위한 사고 분석 과정에 활용했다는 점에서 기존의 CBT 기반 앱과 차별성을 지닌다. 또한, 사용자 테스트에서 긍정적인 평가를 받았다는 점은 정서적 어려움을 겪는 청소년에게 서비스가 효과적으로 활용될 수 있음을 시사한다.

본 연구의 제한점과 향후 제언은 다음과 같다. 첫째, 본 연구의 사용자 테스트의 경우 중학생 1인을 대상으로 이루어져 테스트 결과를 모든 중학생 청소년에게 일반화하기 어렵다. 후속 연구에서는 테스트 대상자의 수를 확장하고, 서비스 이용 후 사용자의 부적응적 정서가 실제로 개선되었는지 구체적으로 측정할 필요가 있다. 둘째, AI가 분석한 인지적 오류 유형과 대안적 사고 결과에 대한 정확성을 확인하지 못하였다. 후속 연구에서는 AI 분석 결과를 전문가의 검토나 실제 사례와 비교해 검증한다면, 서비스의 신뢰성을 높일 수 있을 것이다.

## 참 고 문 헌

- [1] 건강보험심사평가원. (2022년 6월 24일). *최근 5년(2017~2021년) 우울증과 불안장애 진료현황 분석* [보도자료].
- [2] 이해경, 김수연, & 김혜원. (2013). *남녀 특성화 고등학생들의 진로의사결정유형, 진로결정자기 효능감, 자아존중감 및 우울불안이 학교생활부적응에 미치는 영향*. 청소년복지연구, 15(4), 129-156.
- [3] 장신재, & 김희수. (2019). *청소년의 우울·불안이 자살생각에 미치는 영향: 자아존중감의 매개효과를 중심으로*. 한국사회복지조사연구, 63, 169-191.
- [4] 김준경, 이수연, 이윤주, 정종진, & 최응용. (2010). *상담의 이론과 실제*. 학지사.
- [5] Neenan, M., & Dryden, W. (2011). *인지치료에 대해 알고 싶은 100가지* (이종호 외, Trans.). 학지사.
- [6] 최란, & 박중규. (2010). *스트레스 반응과 역기능적 태도 개선을 위한 대학생 대상 인지 재구조화 훈련의 효과*. 인지행동치료, 10(2), 27-42.
- [7] 홍미정. (2015). *중학생이 지각한 부모애착과 학교생활적응의 관계 : 정서조절곤란의 매개효과* [석사학위논문, 서강대학교].
- [8] 신희영. (2006). *초중학생의 정서조절능력과 전략유형 연구* [석사학위논문, 서울대학교].
- [9] 김성희. (2010). *청소년 스트레스 및 인지적 왜곡과 자살생각과의 관련성* [석사학위논문, 아주대학교].

※ 본 연구는 교육부 및 한국연구재단 대학혁신지원사업으로부터 지원받은 연구임(2024)

# AI 인물 필터링 적용 위치 기반 소셜 미디어 앱 개발

허재연, 정우영, 박세진

## AI-Powered Person Detection in a Location-Based Social Network Platform

Heo Jaeyeon, Jung Wooyeong, Se-Jin Park

### 요약

본 연구는 소셜 미디어의 인플루언서 중심 구조로 인한 부작용을 해결하기 위해, 딥러닝 기반 인물 필터링과 위치 기반 서비스를 결합한 새로운 SNS 플랫폼을 제안한다. AI 기술로 인물 중심 콘텐츠를 제한하고, 사용자의 위치 기반으로 일상적인 사진들을 공유하는 구조를 통해, 건강한 소통과 지역 커뮤니티 형성을 도모한다.

### Abstract

This study proposes a new SNS platform that combines deep learning-based person filtering and location-based services to address the negative effects of influencer-centric social media structures. By using AI technology to limit person-focused content and implementing a system that shares everyday photos based on user location, the platform aims to promote healthy communication and foster local community building.

### Key words

Location-based Social Network, AI image Filtering, Non-person Content, Real-time Photo Sharing, Shared Folder System

## 1. 서론

2024년 기준, 전 세계에서는 초당 61,400장, 매일 53억 장의 사진이 촬영되며, 140억 개의 이미지가 소셜 미디어에 공유되고 있다. 이러한 디지털 시대에 사진은 단순한 기록을 넘어 개인의 삶을 풍요롭게 하는 필수적인 도구로 진화했다. Nicholls와 SANE(2022) 사진 촬영 활동은 창의성, 기억력, 의사결정력 등의 인지적 능력을 향상시키며, 일상의 순간을 포착하고 공유하는 과정을 통해 사회적 연결

감을 증진시킨다고 언급했다. 그러나 현대 사회에서 SNS에 공유되는 사진은 단순히 긍정적인 결과만을 유저들에게 가져오지 않는다.

기존 소셜 미디어 플랫폼들의 인플루언서 중심 구조가 가진 부정적 특성을 탈피하기 위해, 본 연구는 인공지능 기반 이미지 분석과 위치 기반 서비스를 결합한 새로운 플랫폼을 제안한다. 딥러닝 기술을 활용하여 인물 중심 콘텐츠를 필터링하고, 사용자의 위치를 기반으로 근거리에서 공유된 일상적인 사진들을 우선적으로 노출하는 구조를 통해 대규모

\* 계명대학교 행정학과, ght010522@gmail.com

\*\* 계명대학교 컴퓨터공학과, uyeong7181@gmail.com

팔로워 기반의 인플루언서 생태계 형성을 제한한다.

이를 통해 본 연구 기반 SNS 어플리케이션은 소셜 미디어의 사회적 연결성을 유지하면서도, 같은 지역의 사용자들이 일상적 경험을 자연스럽게 공유하고 공감할 수 있는 환경을 제공한다. 초지능사회에서 기술이 단순히 효율성과 편의성을 넘어, 인간의 진정성 있는 소통과 공감능력을 증진시키는 도구로 활용될 수 있음을 보여준다.

## II. 기존 SNS 현황 및 한계점

현재 이미지 공유 및 위치 기반 서비스 분야에서 대표적인 플랫폼으로 인스타그램(Instagram)과 구글 지도(Google Map)가 있다.

인스타그램은 이미지 기반 SNS의 선두주자로서, 피드 공유, 실시간 스토리, 팔로잉 시스템을 통해 사용자 간 활발한 상호작용을 이끌어낸다. 특히 최근 도입된 릴스 기능과 탐색 기능은 새로운 콘텐츠 발견을 용이하게 하며, 개인과 기업 모두에게 효과적인 마케팅 플랫폼으로 자리잡았다.

구글 지도는 위치 기반 서비스의 대표적 예시로, 실시간 위치 공유, 거리 뷰, 장소 리뷰 및 사진 공유 등의 기능을 제공한다. 특히 사용자가 특정 장소에 사진을 업로드하고 리뷰를 작성할 수 있는 기능은 위치 정보와 시각적 콘텐츠의 효과적인 결합을 보여준다.

이전에 언급한 인스타그램, 구글 맵과 같은 시중의 SNS는 몇 가지 한계점을 가지고 있다. 인스타그램의 경우, 과도한 인물 중심의 콘텐츠로 인한 사회적 비교와 심리적 부담, 광고성 콘텐츠의 범람, 그리고 실제 위치 기반의 즉시성이 부족하다는 점을 들 수 있다. 구글 맵은 위치 정보와 이미지의 결합이라는 측면에서 혁신적이거나, 실시간성과 소셜 네트워킹 측면에서는 제한적이며, 공유된 이미지의 접근성이 상대적으로 낮다는 한계가 있다.

SNS가 주요 의사소통 방법 중 하나로 자리잡은 이래로 소셜 미디어 플랫폼의 인플루언서 중심 콘텐츠는 꾸준히 여러 부작용을 초래하고 있다. Al-Ansi 외 연구진(2023)의 중동 5개국 청소년 1,612명을 대상으로 한 연구에서 따르면 소셜 미디어 인플루언서들이 청소년들의 사회적, 윤리적, 건강적

측면에 부정적인 영향을 미치는 것으로 나타났다. 특히 청소년들은 인플루언서를 롤모델로 삼고 모방하려는 경향(평균 3.88/5.0)이 높았으며, 폭력적이고 공격적인 행동을 학습하게 되는 문제(평균 3.90/5.0)가 가장 심각한 것으로 조사되었다. 또한 인플루언서 콘텐츠에 장시간 노출될 경우 불안과 심리적 압박(평균 3.32/5.0), 사회적 고립감(평균 3.29/5.0), 우울감(평균 3.09/5.0) 등의 부작용이 나타났으며, 가족과의 유대관계 약화와 종교적 의무 소홀 등 전통적 가치관의 훼손도 우려되는 것으로 밝혀졌다. 자아정체성이 형성되는 청소년기에 인플루언서를 무비판적으로 롤모델로 삼는 현상이 자존감 저하와 우울증 증가의 주요 원인이 될 수 있다고 지적했다.

이처럼 인플루언서 중심의 SNS 소비로 인해 발생하는 행태를 조사한 결과, 현대 사회의 구성원들에게 부정적인 영향을 끼치고 있음을 알 수 있다.

이러한 한계점들을 고려할 때, 기존 서비스들의 장점을 통합하고 한계점을 보완하는 방향으로 SNS의 변화가 필요함을 시사한다. 이에 본 연구에서는 위치 기반의 실시간성과 AI 필터링을 통한 건강한 콘텐츠 제공을 결합한 새로운 형태의 서비스를 제안한다.

## III. 사용자 인터페이스 설계

본 연구에서 제안하는 플랫폼은 기존 소셜 미디어의 한계점을 보완하고 사용자 중심의 경험을 제공하기 위해 차별화된 인터페이스 설계를 적용하였다. 플랫폼의 주요 인터페이스는 크게 지도 기반 메인 화면, 위치 기반 피드 시스템, 공유 폴더 시스템, 프로필 및 성취 시스템으로 구성된다.

해당 플랫폼의 핵심인 지도 기반 메인 화면은 사용자의 공간적 맥락을 우선시하는 특징을 갖는다. 사진이 촬영된 위치를 지도상에 직관적으로 표시하고, 섬네일 형태의 이미지를 오버레이하여 시각적 정보를 효과적으로 전달한다. 하단 중앙에 배치된 위치 아이콘을 통해 사용자는 직관적으로 위치 기반 탐색을 수행할 수 있으며, 확대/축소 기능을 통해 원하는 범위의 콘텐츠를 탐색할 수 있다. 특히 미니멀한 UI 구성을 통해 사용자가 콘텐츠에 집중할 수 있도록 설계되었다.

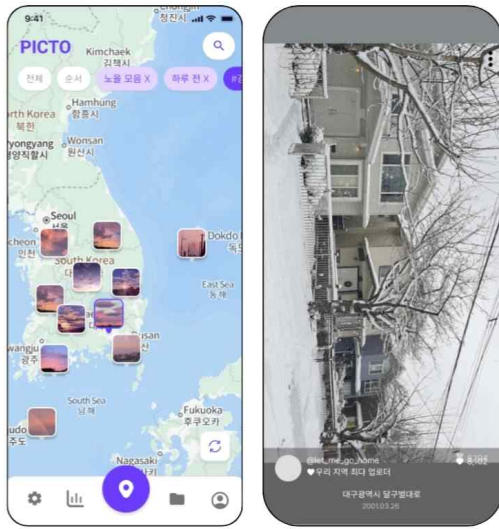


그림 1. 메인 지도화면 및 피드 화면  
Fig. 1. Main map and feed screen

위치 기반 피드 시스템은 기존 SNS의 시간순 피드나 알고리즘 기반 추천 시스템과는 차별화된 접근 방식을 취한다. 각 게시물에는 위치 정보가 강조되어 표시되며, 좋아요 기능을 통한 제한적 상호작용만을 허용한다. 주목할 만한 점은 댓글 기능을 과감히 제외함으로써 익명성으로 인한 부작용을 최소화하고, 위치 정보를 통한 자연스러운 연결성 확보에 중점을 두었다.

공유 폴더 시스템은 협업과 커뮤니케이션을 위한 새로운 접근방식을 제시한다. 팀 기반의 콘텐츠 관리가 가능하도록 체계적인 파일 구조를 제공하며, 폴더 내 채팅 인터페이스를 통해 목적 지향적인 대화가 가능하다. 이는 폐쇄적 커뮤니티 형성과 선택적 공유를 통한 프라이버시 보호를 가능하게 한다.

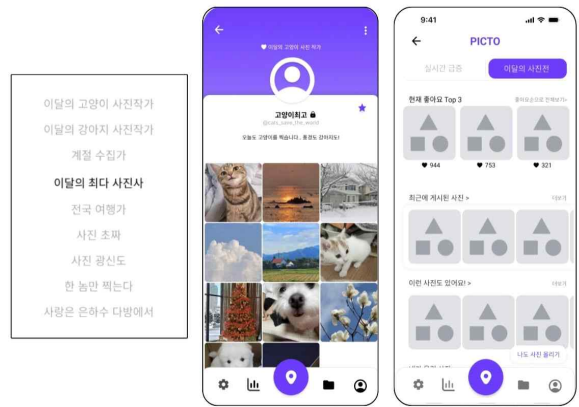


그림 3. 칭호 선택 페이지 및 대회 화면  
Fig. 3. Screen of title selection and contest

프로필 및 성취 칭호 시스템은 게이미피케이션 요소를 통해 건강한 참여를 유도한다. 사용자의 활동에 기반한 칭호 시스템을 도입하여 개인의 성취감을 고취시키고, 정기적인 사진 대회와 같은 이벤트를 통해 커뮤니티 참여를 자연스럽게 유도한다.

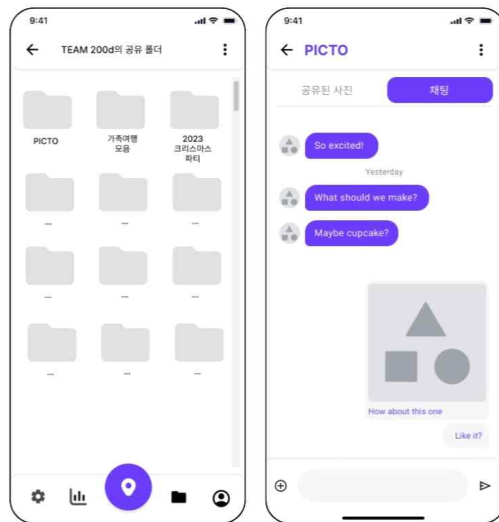


그림 2. 공유폴더 화면 및 공유폴더내 채팅 활성화 화면  
Fig. 2. Screen of shared folder and chatroom of shared folder

#### IV. 결 론

본 연구에서 제안한 위치 기반 Non-Person AI 필터링 사진 공유 SNS 어플리케이션은 기존 SNS의 한계를 극복하고 사용자들에게 더 건강한 소셜 미디어 경험을 제공할 것으로 기대된다. AI 기반 인물 필터링 시스템을 통해 불필요한 사회적 비교를 방지하고, 위치 기반 서비스를 통해 지역 커뮤니티와의 자연스러운 연결을 촉진할 수 있을 것이다.

향후 연구에서는 제안된 시스템의 실제 구현과 사용자 테스트를 통해 시스템의 효과성을 검증하고, AI 필터링 알고리즘의 정확도 향상과 처리 속도 개선을 위한 추가적인 연구를 진행할 계획이다. 본 연구는 현대 사회에서 요구되는 건강한 소셜 미디어 플랫폼의 새로운 방향을 제시했다는 점에서 의의가 있다.

참 고 문 헌

- [1] Nicholls, K. (2023). Can taking photos help your mental health? happiful.
- [2] SANE. (2021). Photography has the power to improve our mental wellbeing.
- [3] Al-Ansi AM, Hazaimh M, Hendi A, Al-Hrinat J, Adwan G. How do social media influencers change adolescents' behavior? An evidence from Middle East Countries. Heliyon. 2023 Apr 29;9(5):e15983. doi: 10.1016/j.heliyon.2023.e15983. PMID: 37168884; PMCID: PMC10165405.

# CBTC 시스템에서 응급 상황 대응을 위한 자원할당

강운제\*, 김선명\*\*

## Resource Allocation for Emergency Response in CBTC Systems

Yunje Kang\* and Sunmyeong Kim\*\*

### 요 약

Communications-based train control(CBTC)와 GSM-R은 철도 신호 및 제어 시스템의 핵심 기술로 무선 통신을 통해 실시간으로 열차 간의 정보를 교환 하여 제어를 수행하는 기술이다. CBTC, GSM-R은 서로 다른 통신 표준이지만 CBTC는 근거리 통신, GSM-R은 원거리 통신을 할 수 있다는 장단 때문에 현재 열차들에서 혼용해서 사용되기도 한다. 본 논문에서는 GSM-R의 차세대 모델인 Future Railway Mobile Communication System(FRMCS)를 기반으로 한 CBTC 환경을 가정한다. FRMCS는 원거리 통신이 가능하여 음성 데이터, 제어 데이터를 열차끼리 교환하고 승객들에게 wifi를 제공하는 것이 특징이다. CTBC는 승객용 WiFi, 센서 네트워크, 열차 내 외부 통신을 포함한 다양한 데이터를 통합적으로 관리하는 상황을 가정하였으며 본 논문에서는 데이터를 우선순위에 따라 나누고, Uplink에서 RU를 할당하는 방법과, Downlink에서 긴급 상황 시 발생할 수 있는 제어 시나리오에 대해 제시한다. 이를 통해 실시간으로 데이터를 우선순위에 따라 분배하고, 상황에 맞는 최적의 자원을 할당하여 시스템의 효율성과 안정성을 높이는 방법을 제안하였다.

### Abstract

Communications-Based Train Control (CBTC) and GSM-R are key technologies in railway signaling and control systems that perform control by exchanging real-time information between trains using wireless communication. Although CBTC and GSM-R follow different communication standards, they are sometimes used together in current train systems due to their complementary strengths, with CBTC handling short-range communication and GSM-R enabling long-range communication. This paper assumes a CBTC environment based on Future Railway Mobile Communication System (FRMCS), the next-generation model of GSM-R. FRMCS supports long-range communication, allowing trains to exchange voice and control data while also providing Wi-Fi services to passengers. The assumed scenario involves CBTC managing diverse data, including passenger Wi-Fi, sensor networks, and both internal and external train communications. In this paper, we propose methods for dividing data by priority, allocating Resource Units (RUs) in the uplink, and presenting control scenarios for downlink communication in emergency situations. Through this, we suggest a method to distribute data in real-time according to priority and to allocate resources optimally depending on the situation, thereby enhancing the system's efficiency and stability.

### Key words

shape recovery, shape from focus, artificial neural network, depth estimation, focus measure

---

\* 국립금오공과대학교 컴퓨터소프트웨어공학과, k1robot@kumoh.ac.kr,

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과, sunmveng@kumoh.ac.kr

## I. 서 론

철도 시스템에서의 안전과 효율성은 열차의 운영을 제어하는 다양한 통신 기술에 의해 크게 좌우된다. 최근 Communication Based Train Control (CBTC)[1]은 열차와 지상 시스템 간의 통신을 통해 실시간으로 운행 정보를 교환함으로써 안전한 운영을 보장하고 있다. 이러한 시스템은 특히, 열차 간의 거리를 보다 정확하게 측정하고 관리함으로써 운행 간격을 최소화할 수 있어, 철도 운영의 효율성을 극대화한다.

그러나 CBTC와 같은 시스템에서 여러 센서와 승객용 Wi-Fi 시스템이 동일한 네트워크 인프라를 공유할 때 응급 상황이 발생할 경우, 열차의 안전 운영을 위한 센서 데이터는 우선적으로 처리되어야 하지만, 승객 데이터를 포함한 다른 트래픽들과 충돌이 발생할 가능성이 있다. 이러한 충돌을 효과적으로 관리하지 않으면 중요한 센서 데이터의 전송이 지연되거나 손실될 수 있으며, 이는 열차 안전에 치명적인 결과를 초래할 수 있다.

본 논문에서는 CBTC와 같은 철도 통신 시스템에서 센서 데이터와 승객 데이터를 관리하기 위한 자원 할당 기법을 제안한다. 특히, 응급 상황[2]에서 센서 데이터의 우선순위를 동적으로 조정하고, 트래픽 충돌을 최소화하는 방식으로 네트워크 자원을 효과적으로 분배하는 방법에 중점을 둔다. 또한, 센서 전용 프레임과 승객용 프레임 사이클을 유지하면서도 네트워크의 안정성을 확보하고, 응급 상황 대응 속도를 향상시키기 위한 메커니즘을 논의한다.

본 논문은 철도 운행 중 발생할 수 있는 다양한 응급 상황에 대한 네트워크 자원 할당 방식을 다루며, 이를 통해 열차 운행 제어와 관련된 통신의 신뢰성과 안정성을 개선하고자 한다.

## II. 본 론

센서 데이터를 '센서'와 '승객용 Wi-Fi'로 분류하고, 센서 데이터의 우선순위를 1, 승객 STA의 우선순위를 2로 설정한다. 단, 센서에서 이상 값을 감지하여 응급 상황이 발생할 경우 해당 센서의 우선순

위는 0으로 설정된다.

센서들에게 모든 Resource Unit(RU)를 할당하는 Trigger Frame(TF) 사이클을 TF-S, 승객들에게 모든 RU를 할당하는 TF 사이클을 TF-P라고 정의한다. Access Point(AP)는 스케줄링을 통해 모든 센서로부터 데이터를 수신한 후, 승객 Station(STA)들에게 RU를 Uplink OFDMA-based Random Access (UORA) 방식으로 할당하는 과정을 하나의 주기로 삼는다. 응급 상황이 발생할 경우, TF-S에서는 모든 센서가 이미 RU를 할당받았기 때문에 AP가 즉각적으로 대응할 수 있다. 그러나 TF-P에서는 이상을 감지한 센서가 RU를 할당받지 못해 AP에게 데이터를 전송하지 못할 가능성이 존재한다. 이를 방지하기 위해, 응급 상황 발생 시 해당 센서의 back-off timer 값을 감소시키고, EDCA 환경에서 우선순위가 가장 높은 Voice 트래픽의 back-off timer인 DIFS보다 짧게 설정하여 신속하게 AP에게 데이터를 전달할 수 있도록 한다.

그림 1은 센서 데이터를 위한 스케줄링 된 사이클 TF-S와 승객 STA에게 경쟁 방식으로 RU를 할당하는 TF-P의 동작 방식을 보여준다. TF-S에서는 센서들이 사전에 할당된 RU를 사용하여 데이터를 전송하며, 모든 센서들의 데이터를 수신한 후 ACK 패킷이 전송된다. 반면, TF-P에서는 여러 STA들이 경쟁을 통해 RU를 할당받아 데이터를 전송하며, 모든 STA들의 데이터를 수신한 후 ACK 패킷이 전송된다.

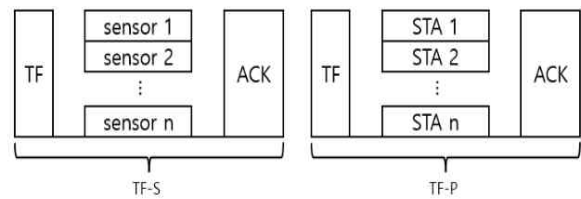


그림 1. TF-S, TF-P 구성  
Fig. 1. composition of TF-S, TF-P

그림 2는 모든 센서들의 데이터를 수신한 후, 주기를 수행하는 과정을 보여준다. 여러 TF-S 주기가 진행된 후, 승객 STA들에게 RU를 할당하는 TF-P 주기가 순차적으로 진행됨을 나타낸다.



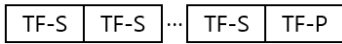


그림 2. TF-S, TF-P 주기  
Fig. 2. TF-S, TF-P cycle

AP가 응급 상황 데이터를 수신하게 되면, TF-P 주기 내 모든 작업을 중단하고 Downlink에서 승객 STA에게 데이터가 전송되지 않도록 하는 제어 프레임을 전송한다. 이로 인해 우선순위 2번으로 설정된 승객 STA에게 내려가는 모든 Downlink 전송이 멈추고, 응급 상황이 발생한 센서와의 통신에 집중하게 된다. 이때 센서가 제공하는 데이터는 Onboard System을 통해 다른 열차로 제어 메시지가 전달되며, 상황에 따라 감속 또는 가속을 통해 열차 운행이 제어된다.

운행 제어는 감지된 센서 유형에 따라 대처 방식이 달라진다. 예를 들어, 화재 상황에서는 화재의 정도에 따라 즉시 선로에 정차하거나, 가까운 승강장으로 이동하여 승객을 하차시킬 수 있다. 응급 상황이 발생하면 열차는 이를 관제 센터에 알린 후, 관제 센터로부터 제어 메시지를 수신하여 대응 조치를 수행하게 된다. 그러나 화재 감지 센서들이 지속적으로 AP에 데이터를 전송하려 시도할 경우, 다른 이상 감지 센서들의 데이터를 AP가 수신하지 못하는 상황이 발생할 수 있다.

TF-P 주기에서 여러 센서가 데이터를 전송하려 시도할 때, 트래픽 충돌이 발생하는 문제를 해결하기 위해, 우선순위가 0인 데이터를 수신한 후 해당 센서에 일정 시간 동안 타이머를 부과하는 방식을 제안한다. 이를 통해 동일한 센서로부터의 반복적인 데이터 전송을 방지하고, 다른 센서에서 발생한 이상 데이터를 원활하게 수신할 수 있도록 한다.

이러한 상황에서도 TF-P를 완전히 제거하고 TF-S만을 이용하여 통신하는 방식이 가능하지만, TF-P를 제거할 경우 네트워크 안정성이 저하될 위험이 있으며, 시스템 오버헤드 또한 증가할 수 있다. 따라서 네트워크의 안정성을 유지하고 오버헤드를 최소화하기 위해 TF-P와 TF-S 사이클을 함께 유지하는 방식을 제안한다.

### III. 결 론

철도 시스템 내에서 센서와 승객용 Wi-Fi 트래픽이 동일한 네트워크 인프라를 공유할 때, 응급 상황에서 센서 데이터의 우선순위를 동적으로 조정하고, 승객 데이터와의 충돌을 최소화하는 방법이 필수적이다. 이를 위해 TF-S(센서 전용 프레임)와 TF-P(승객 전용 프레임) 사이클을 활용한 통신 구조를 제시하였고, 우선순위에 따라 센서 데이터와 승객 데이터를 관리하여 응급 상황에서 네트워크의 안정성과 신뢰성을 높이기 위한 자원 할당 기법을 제안하였다.

### 참 고 문 헌

- [1] IEEE, "IEEE Recommended Practice for Functional Testing of a Communications-Based Train Control (CBTC) System," IEEE Std 1474.4-2011, 2011, 3
- [2] Korea Railroad Research Institute, "Development of Train Autonomous-Driving Control Core Technology," Korea Railroad Research Institute, 2017, 48-50

# KM-BERT와 XGBoost 기반의 지능형 인공지능 의료 서비스 추천 시스템 설계

함양훈\*, 이성민, 김민수, 곽찬희\*\*

## Design of an Intelligent Korean Medical Service Recommendation System based on KM-BERT and XGBoost

Yanghoon Ham\*, Seongmin Lee, Minsu Kim and Chanhee Kwak\*\*

### 요 약

제한적인 의료 자원에도 불구하고, 상급병원을 선호하는 현상은 의료계의 오래된 숙제이다. 이를 완화하고자 본 연구에서는 환자의 증상을 서술한 텍스트를 기반으로 상병코드를 예측하여 적절한 의료 기관의 수준을 제안하는 추천 시스템을 개발하고자 하였다. 특별히 본 연구에서는 트랜스포머 기반의 Bert 모델과 XGBoost 기반의 모델을 병렬적으로 구성한 앙상블 모델을 제안하여, 우수한 성능의 추천 시스템 모델 구축을 시도했다. 본 연구를 통해 의료 자원 사용의 최적화 및 한국어 기반 의료 추천 시스템 활용 확대를 기대한다.

### Abstract

Despite limited medical resources, the phenomenon of preferring advanced hospitals remains an ongoing challenge in the healthcare sector. To alleviate this issue, this study aims to develop a recommendation system that suggests appropriate levels of medical services by predicting diagnostic codes based on textual descriptions of patients' symptoms. Specifically, this research proposes an ensemble model that combines a transformer-based BERT model with an XGBoost-based model in parallel, attempting to build a recommendation system model with superior performance. This study anticipates optimizing the use of medical resources and expanding the application of Korean-language medical recommendation systems.

### Key words

recommendation system, ensemble learning, medical service recommendation

### I. 서 론

현대의료 환경은 외래진료의 과다, 특정 의료 기관의 과잉 이용, 그리고 임상 의사 수 부족 등의 문제를 겪고 있다. 이에 따라 의료서비스 접근성과 질이 저하되고 자연스럽게 고객만족도가 떨어지게

된다. 김효정(2013) 외는 이러한 상황을 보고 의료 기관들은 자원을 효율적으로 활용하지 못하는 등 국민에게는 경제적·시간적 부담으로, 중소병원·의원에는 경영난으로 그 결과가 나타나고 있다고 하였다[1]. 가벼운 증상임에도 상급병원으로 외래진료

\* 강남대학교 인공지능융합공학부 데이터사이언스전공, Email eric9898@kangnam.ac.kr

\*\* 강남대학교 인공지능융합공학부 데이터사이언스전공 조교수, Email chk@kangnam.ac.kr

를 받으러 오는 환자들은 자신의 증상에 대한 정보가 부족하며, 자신의 증상이 어떤 진료과에 해당하는지 판별하기도 쉽지 않다.

이러한 문제를 해결하고자 본 연구에서는 환자가 일상생활에서 느끼는 주관적 증상을 텍스트로 입력받아 상병코드를 중분류까지 예측하고 그에 맞는 진료과를 추천하는 모델을 개발하였다. 세부적으로는 트랜스포머 인코더 기반의 NLP 모델인 BERT와 Tree 기반 앙상블 모델인 XGBoost를 조합하여 최적의 구조를 도출하고자 하였다. 본 연구에서 제안한 모델을 활용한다면 의료지식이 부족한 의료서비스 이용자들의 의료 접근성을 강화하고, 제한된 의료 자원 최적화를 기대한다.

## II. 모델 설계

### 2-1. 데이터셋

데이터셋은 국민건강보험공단에서 제공하는 진료내역정보 데이터를 사용하였다. 해당 데이터셋은 2023년 1,048,576건에 대한 환자 기본정보 및 진료내역으로 구성된 데이터이다. 진료내역 상병코드를 KM-BERT 학습에 알맞은 텍스트 데이터로 변환하기 위해 건강보험심사평가원에서 제공하는 상병마스터 데이터를 활용하여 상병코드와 한글 병명을 일치하여 변환하였다. 환자가 입력하는 증상 설명 텍스트를 적합한 상병코드로 분류하기 위하여 상병코드별 증상을 나타내는 텍스트 데이터를 위키피디아와 네이버 의료 지식백과를 이용하여 추가하였다. 필요 이상으로 세분화되어있는 상병코드는 병합했고, 텍스트로 유추할 수 없는 중증 상병코드는 제거하였다.

표 1. km-BERT 학습데이터 예시  
Table 1. km-BERT Train Data Example

적용 모델	km-BERT
데이터 개수	882
구성	상병코드
	한글명
	증상 설명

표 2. XGBoost 학습데이터 예시  
Table 2. XGBoost Train Data Example

적용 모델	XGBoost	
데이터 개수	1,048,576	
구성	성별코드	성별
	연령대코드	연령대
	시도코드	지역
	진료과목코드	진료과
	주상병코드	주된 증상
	부상병코드	추가 증상
	총처방일수	약 처방 일수

### 2.2 방법론

본 연구에서는 한국어 의료 추천 시스템을 개발하기 위해 KM-BERT와 XGBoost를 결합한 멀티모달 앙상블 모델 구조를 제안한다. 해당 모델은 두 개의 병렬 트랙으로 구성되어 있으며, 이러한 접근 방식은 최근 개인 신용 예측 분야 등에서 뛰어난 성능을 보였다[2]. 해당 연구에서는 BERT 기반의 NLP 모델과 Transformer 모델을 결합하여 텍스트 데이터와 테이블 데이터를 효과적으로 처리하는 방법을 제시하였으며, 이는 본 연구의 의료 추천 시스템에도 적용할 수 있다.

본 연구에서 제안하는 모델은 환자의 증상(텍스트 데이터)과 환자 정보(테이블 데이터)를 병렬적으로 처리함으로써 최적의 진단 및 치료 방법을 제시할 수 있도록 설계되었다. 이러한 멀티모달 접근 방식은 다양한 유형의 데이터를 동시에 처리하여 정보의 풍부함을 극대화하고, 이를 통해, 진단 및 치료의 정확성을 향상하는 것을 목표로 한다. 선행 연구에서도 다양한 데이터 유형의 동시 처리가 성능 향상에 기여함을 입증한 바 있어, 본 연구의 모델 또한 이를 통해 효과적인 의료 추천 시스템을 구축하고자 한다.

### 2-2. 모델 구조 개요

Fig 1은 모델의 구조 도식화 그림이다.

Track-1은 KM-BERT 기반 증상 분석 트랙의 역할을 한다. Transformer Encoder 구조를 활용한 KM-BERT 모델을 사용하여 사용자가 입력한 증상 텍스트를 분석한다. 이 모델은 고려대학교 의료 빅

데이터 연구소에서 사전 학습된 KM-BERT를 활용하며, 약 1억 1천 6백만 단어로 이루어진 한국어 의료 코퍼스를 기반으로 사전학습이 수행되었다[3]. 해당 코퍼스는 의료 교과서, 건강 정보 뉴스, 의료 연구 논문 등으로 구성되어 있으며, 다양한 유형의 한국어 의료 문서에서 추출된 데이터로 구성되어 있다.

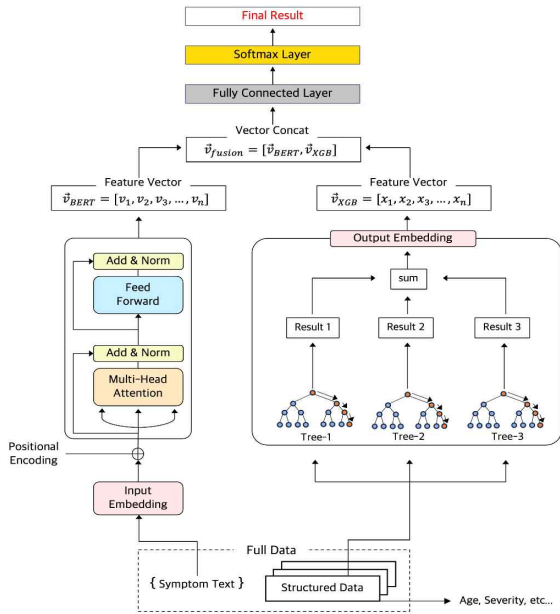


그림1. 모델 아키텍처 시각화  
Fig 1. model architecture visualization

Track-2는 환자의 정보인 정형 데이터를 처리하며 Tree 기반의 앙상블 모델인 XGBoost를 활용한다. 여기에는 환자의 나이, 성별, 진료 기록 등과 같은 환자 개인 정보가 포함된다. XGBoost는 KM-BERT가 분석하지 못할 수 있는 추가적인 패턴이나 상관관계를 포착하여 보다 정교한 예측을 가능하게 할 수 있다.

두 트랙의 결과는 단순 벡터 결합을 통해 결합된다. 결합된 벡터는 Fully Connected Layer와 Softmax Layer를 통해 최종적으로 다중 클래스 분류 결과를 산출한다.

### 2-3. 모델 학습 및 하이퍼파라미터 최적화

KM-BERT 모델의 학습 과정에서 주요 하이퍼파라미터인 Learning Rate, Batch Size, Epoch 수를 최적화하였다. Learning Rate 최적화를 위해서는 Cosin

e Annealing 스케줄러를 사용하였으며, 초기 학습률은 2e-5로 설정하고, 훈련이 진행됨에 따라 점진적으로 감소시켜 모델이 안정적으로 수렴할 수 있도록 하였다. Batch Size는 16과 32중에서 최적의 값을 탐색하였으며, Epoch 수는 3에서 10 사이로 설정하여 학습을 진행하였다.

XGBoost 모델의 하이퍼파라미터는 Optuna 라이브러리를 사용하여 최적화하였다. Optuna를 통해 learning\_rate(0.01-0.3), max\_depth(3-10), n\_estimators, subsample, colsample\_bytree 등의 다양한 하이퍼파라미터 조합을 탐색하여 최적의 모델을 도출하였다. 최적화 과정에서는 fl\_weighted 스코어를 사용하여 모델의 성능을 평가하고 개선하였다.

### 2-4. 성능 평가 및 교차 검증

본 연구에서는 Accuracy, Precision, Recall, F1-score와 같은 성능 평가 지표를 사용하여 모델을 평가하였다. 특히, 다중 클래스 분류 문제의 특성을 고려하여 각 클래스에 대한 분류 성능을 상세히 분석하였다. 데이터셋은 Train-test split 비율을 7:3으로 설정하여 학습과 테스트에 사용하였다. 또한, 모델의 일반화 성능을 평가하고 과적합을 방지하기 위해 5-fold 교차 검증을 수행하였다. 이를 통해 모델이 다양한 데이터 샘플에 대해 안정적인 성능을 유지하도록 하였다.

## III. 결론

향후 연구에서는 더 다양한 데이터셋을 확보하고, 다른 안정적인 기술을 도입하여 성능을 더욱 개선이 가능할 것으로 기대한다. 또한, 임상 데이터를 활용한 검증을 추가적으로 진행된다면 모델의 전반적인 성능의 향상은 물론 실제적 활용 가능성 또한 강화할 수 있을 것이다. 나아가 본 연구의 결과를 바탕으로 한국어 의료 추천 시스템의 활용 영역을 확대할 것으로 기대한다.

## 참고 문헌

[1] 김효정, 김영훈, 김한성, 우정식, 오수진, Hyo-Jeong Ki

- m, Young-Hoon Kim, Han-Sung Kim, Jung-Sik Woo, Su-Jin Oh. "외래 본인부담률 인상이 상급종합병원과 종합병원 외래 의료이용에 미친 영향" 보건행정학회지 23.1 (2013): 19-34.
- [2] K.Choi, S. YOO and J. Lee, "Incorporating BERT-based NLP and Transformer for An Ensemble Model and its Application to Personal Credit Prediction", Korean Institute of Smart Media, vol. 13, No. 4, pp. 9-15, April 2024
- [3] Kim, Y., Kim, JH., Lee, J.M. et al. A Pre-trained BERT for Korean Medical Natural Language Processing. Sci Rep 12, 13847 (2022). <https://doi.org/10.1038/s41598-022-17806-8>
- [4] Wonjun Park, Inchan Kim, Younghan Jung, Kyungsoo Kim, Byoung-woo Oh, "의료 격차 집단을 위한 증상 기반 병원 추천 시스템" Proceedings of KIIT Conference (2023.06): 991-995
- [5] Seohee Lee, Juyoung Kang, "환자의 주관적 증상 텍스트에 대한 진료과목 분류 모델 구축" 한국빅데이터학회 학회지 6.1 (2021): 51-62



# 자전거 이용자들을 위한 편의 서비스: 달리자전거

임나경\*, 강채희\*\*, 송채빈\*\*, 허겸\*\*, 이규원\*\*\*, 이해연\*\*\*\*

## Convenience Services for Cyclists: DalliBike

NaGyeong Im\*, Chaehui Kang\*\*, Chaebin Song\*\*, Gyeom Heo\*\*, Kyu-Won Lee\*\*\*,  
and Hae-Yeoun Lee\*\*\*\*

### 요 약

현재 자전거 애플리케이션은 필요한 기능들이 분산되어 있어 사용자에게 불편함을 초래하고 있다. 이를 해결하기 위해 자전거 전용 지도, 위험 요소 신고, 스톱워치 및 주행 달력, 게시판 기능을 통합한 달리자전거 애플리케이션을 제안한다. 달리자전거는 포트홀 등 위험 요소를 신고할 수 있는 알고리즘을 포함하여, 자전거 주행 중 발생할 수 있는 위험을 공유하고, 자전거 시설 및 정보를 쉽게 확인할 수 있는 마커 기능을 제공한다. 또한, 주행 기록과 랭킹 시스템을 통해 사용자가 자신의 주행 데이터를 관리하고 다른 사용자들과 경쟁할 수 있도록 설계되었다. 이 시스템은 사용자 간 소통을 촉진하고 자전거 이용률을 높이는 데 기여한다.

### Abstract

Currently, bicycle applications are causing inconvenience to users as necessary functions are distributed. To solve this problem, we propose a DalliBike application that integrates bicycle-only maps, reporting risk factors, stopwatches and driving calendars, and bulletin board functions. DalliBike application includes an algorithm that can report risk factors such as potholes, etc., and provides a marker function that can easily check bicycle facilities and information. In addition, the driving record and ranking system is designed to allow users to manage their driving data and compete with other users. The system contributes to facilitating communication between users and increasing bicycle utilization.

### Key words

Bicycle application, Integrated features, Risk factor reporting, User communication, Bicycle usage rate

## 1. 서 론

도시화와 환경 보호 의식 증가로 자전거 이용이 증가하고 있음에도 불구하고 자전거 이용에 관련된

불편함을 호소하는 의견이 여전히 존재하고 있다. 많은 사용자는 대여소의 위치를 찾기 어렵고, 도로 위에서의 안전한 주행을 보장하기 위한 서비스가 부족하다는 의견을 제기하고 있다. 또한, 자전거를

\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생, lwknk@kumoh.ac.kr

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 학생

\*\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수

\*\*\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과 교수, haeveoun.lee@kumoh.ac.kr(교신저자)

보관하고 주차할 수 있는 시설의 위치 정보 부족으로 인해 불편을 겪는 경우도 많다.

따라서 사람들은 지나가는 길에 위험한 요소가 있는지 확인하거나 자전거 관련 동호회를 찾기 위해 여러 카페에 가입한다. 그러나 각 카페에서 제공되는 정보가 상이하고 업데이트가 느리며, 때로는 정보를 찾기 어려운 경우도 많아서 사용자들은 원하는 정보를 찾기 위해 여러 곳을 찾아봐야 하는 번거로움이 있다.

이미 시장에 이러한 문제들을 해결하기 위해 각각의 기능을 가진 여러 애플리케이션이 있다. 하지만 이 모든 기능을 한 곳에서 편리하게 이용할 수 있는 통합된 애플리케이션이 존재하지 않아 사용자들이 여러 애플리케이션을 다운로드하고 관리해야 하는 번거로움이 있다.

이에 대한 해결책으로써, 본 논문에서는 자전거를 즐기는 모든 개인부터 팀 단위까지의 라이더들을 대상으로, 안전하고 편리한 라이딩을 위한 정보와 소통의 공간을 제공하는 달리자전거 애플리케이션 서비스의 개발에 대하여 설명한다.

본 논문의 구성은 다음과 같다. 2장은 달리자전거 서비스의 개발에 대하여 설명하고, 3장에서는 개발된 애플리케이션의 결과를 제시한다. 마지막으로 4장에서 결론과 향후 개선 방안을 설명한다.

## II. 달리자전거: 자전거 이용자들을 위한 편의 서비스

달리자전거 애플리케이션은 사용자의 편의를 위하여 자전거 사용자들을 위한 자전거 전용 지도와 게시판, 주행 달력으로 이루어져 있다. 먼저 자전거 전용 지도는 크게 자전거 시설 마커 표시, 위험 요소 신고, 스톱위치 기능으로 이루어져 있다. 기본적인 기능으로는 자전거 도로 표시 및 현재 위치를 표시하여 사용자의 편의성을 향상시켰다.

표 1에는 애플리케이션의 지도에 사용된 데이터 이름과 크기를 정리하였다[1].

표에 제시된 내용은 공공데이터로 자전거 주행 시 사용할 수 있는 시설들을 마커로 표시하여 자전거를 이용자들이 필요로 하는 정보를 제공하였다.

표 1. 사용된 공공데이터 세트  
Table 1. Experimental data sets

데이터 이름	데이터 크기 (행)
공기주입기	1,648
판매 수리점	102
숙박업소	2,559
자전거 대여소	1,402
자전거 보관소	16,132

자전거 전용 지도에서 포트홀 혹은 공사정보 신고기능을 위하여 그림 1과 같이 위험 요소 신고 알고리즘을 설계하였다.

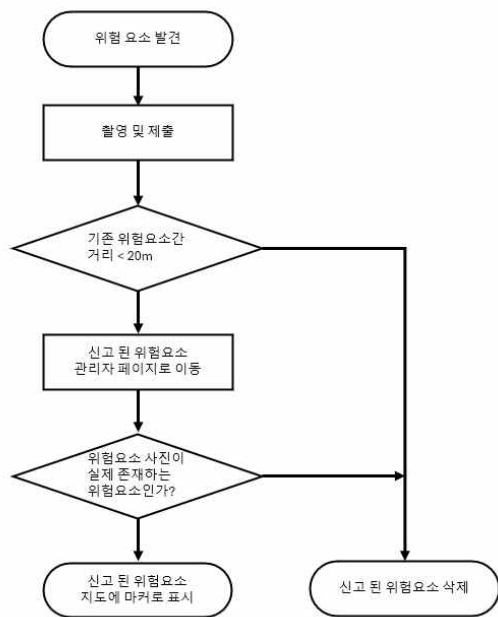


그림 1. 위험 요소 신고 알고리즘  
Fig. 1. Risk Factor Reporting Algorithm

동일 위험 요소의 판단 기준은 신규 제보와 기존 제보 간의 거리로 설정하였으며, 이를 계산하기 위해 하버사인 공식을 활용하였다[2]. 하버사인 공식은 구면에서 두 점 간의 대원 거리를 계산하는 방법으로, 다음과 같은 수식으로 표현된다.

$$d = 2r \cdot \arcsin \left( \sqrt{\sin^2 \left( \frac{lat_2 - lat_1}{2} \right) + \cos(lat_1) \cdot \cos(lat_2) \cdot \sin^2 \left( \frac{lon_2 - lon_1}{2} \right)} \right) < 0.020$$



위 수식을 사용하여 신규 제보의 위치랑 기존 제보의 위치 사이의 대원 거리( $d$ )가 20m 이내인지 확인하여 처리한다. 처리가 완료된 위험 요소는 관리자 페이지를 거쳐 자전거 전용 지도에 마커로 표시하여 자전거 도로 위에서의 안전한 주행을 도모하였다.

자전거 전용 지도의 스톱워치 기능은 하루에 자전거로 주행하는 시간을 측정할 수 있고, 서버의 자정이 지나가면 이전 기록들은 저장이 되고, 다시 0초부터 시작하게 된다. 저장된 기록들은 주행 달력에 표시된다. 주행 달력에는 월별 기록과 총 주행 시간을 확인할 수 있고, 랭킹 시스템이 도입되어 있다. 랭킹 시스템은 다른 사용자들과 경쟁하여, 한 달간 총 주행 시간 가장 높은 3위까지를 볼 수 있다. 랭킹 시스템은 사용자들 간의 경쟁을 도입시켜 자전거 이용률을 높인다.

마지막으로 게시판 기능은 여러 카테고리로 분류하여 사용자들 간의 소통을 원활하게 한다. 또한, 게시판을 통하여 모임 가질 수 있게 하여 자전거 이용률을 높인다.

### III. 달리지전거 개발 결과

본 절에서는 개발한 달리지전거 애플리케이션의 주요한 화면을 소개한다. 그림 2는 자전거 지도에 마커를 표시한 화면이다.

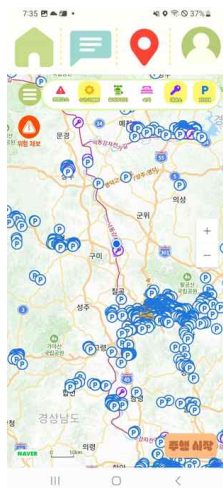


그림 2. 자전거 지도 마커 표시 화면  
Fig. 2. Bike Map Marker Display Screen

해당 지도를 보면, 자전거 길과 중앙의 사용자 현재 위치를 볼 수 있다. 또한, 화면 상단의 시설 메뉴를 선택한 시설의 마커가 뜨게 된다. 위의 화면에서는 수리 판매점과 대여점, 자전거 거치대의 마커가 표시된 것을 볼 수 있다.

그림 3은 자전거 지도 위험 요소 신고 화면으로 화면 상단의 메뉴 버튼 아래에 위험 요소 신고 버튼을 누르게 되면 화면 하단의 메뉴가 나오게 된다. 하단의 메뉴에서 위험 요소가 포트홀인지 공사정보인지 선택을 한 후 사진 버튼을 눌러 현재 위치에 있는 위험 요소 사진을 찍고 제보 완료 버튼을 누르게 되면 위험 요소 신고를 완료할 수 있다.

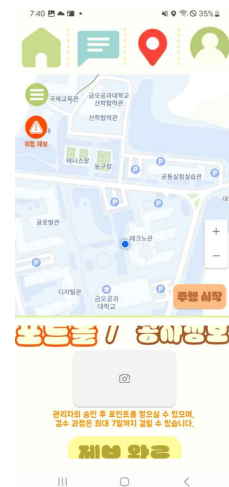


그림 3. 자전거 지도 위험 요소 신고 화면  
Fig. 3. Bicycle map hazard report screen

자전거 지도의 주행 시작 버튼을 누르게 되면 주행 스톱워치를 실행할 수 있다. 그렇게 기록된 주행 시간은 그림 4와 같은 주행 달력 화면에서 확인할 수 있다.

주행 달력 화면을 보면, 해당 날짜의 주행 시간을 달력 아래에서 확인할 수 있고, 해당 월에 대한 랭킹과 사용자의 랭킹을 확인할 수 있다.

그림 5은 게시판 화면으로, 게시판에 쓰여 있는 게시글을 볼 수 있다. 게시글 하나를 선택하게 되면 '좋아요' 버튼과 댓글을 쓸 수 있다. 그리고, 전체 게시판 옆의 버튼을 선택하였을 때 카테고리를 선택할 수 있고, 하단의 연필 모양 버튼을 누르게 되면 게시글을 쓸 수 있다.



그림 4. 주행 달력 화면

Fig. 4. Driving calendar screen



그림 5. 게시판 화면

Fig. 5. bulletin board screen

기록 외에도 주행한 이동 거리를 기록하는 기능으로 확장하고, 사용자가 평균 속도 및 최고 속도를 확인할 수 있는 기능을 추가할 예정이다. 또한, 커뮤니티 게시판에 게시글 검색기능을 도입하여 사용자가 원하는 게시글을 더욱 쉽게 찾을 수 있도록 개선할 계획이다.

## 참 고 문 헌

- [1] 공공데이터포털, “자전거 시설”, <https://www.data.go.kr/index.do>
- [2] C. Robusto, “The cosine-haversine formula”, *The American Mathematical Monthly*, Vol. 64, No. 1, pp. 38-40, 1957.

## IV. 결 론

본 논문에서 제안한 통합 자전거 애플리케이션 달리자전거는 자전거 이용자들이 직면하는 여러 불편함을 해결하기 위한 혁신적인 접근법을 제공한다. 자전거 전용 지도, 위험 요소 신고기능, 스톱위치 및 주행 달력, 그리고 게시판 기능을 통합함으로써 사용자들은 필요한 정보를 한 곳에서 쉽게 접근하고, 안전한 주행 환경을 조성할 수 있다. 이러한 서비스는 자전거 이용자들의 편의성을 향상시키고, 자전거 이용률을 높일 수 있다.

차후 연구에서는 자전거 지도 기능을 주행 시간

# AWS Rekognition을 활용한 AI 얼굴인식 자동 주문 기능 모듈 개발

한진호\*, 양지웅\*, 박성렬\*, 정태훈\*, 김재수\*

## AWS Rekognition-based AI Facial Recognition Automatic Ordering Function Module Development

Jinho Han\*, Jiung Yang\*, Seongryeol Park\*, Taehoon Jeong\* and Jaesoo Kim\*\*

### 요 약

기존의 키오스크에서의 작동방식은 사용자로 하여금 기기를 조작하도록 하여 상호작용을 수행하는 '완전히 수동적인 방식'으로 진행된다. 그렇기 때문에 사용자의 키오스크 사용 능숙도는 주문 처리 과정에서의 회전율에 영향을 끼치게 되고, 이는 사용자의 키오스크 사용을 불편하게 만드는 요소 중 하나가 된다. 이에 AWS rekognition을 활용하여 AI기반의 얼굴인식 자동 주문 기능을 구성하게 되었다. 이 논문에서는 신속한 주문 처리 및 사용자 편리성을 실현하고 디지털 취약계층의 디지털 기기 진입장벽을 해소하기 위해 AI기술을 이용하여 사용자를 인식하고 추천 메뉴를 제안해 줌으로써 사용자 편의성을 증진시키는 방식에 대해 다룬다.

### Abstract

The existing kiosk operation method proceeds in a 'completely passive way' in which the user operates the device to perform interaction. Therefore, the user's familiarity in using the kiosk affects the turnover rate of order processing, which is one of the factors that make it inconvenient to use the kiosk. As a result, AWS rekognition was used to configure an AI-based face recognition automatic order function. In this paper, in order to pursue rapid order processing and convenience and to solve the barriers of digital vulnerable groups to entry into digital devices, we deal with ways to increase user convenience by using AI technology to recognize users and suggest recommended menus.

### Key words

kiosk, face recognition, aws Rekognition, suggest memu

## 1. 서 론

사물의 이미지나 영상을 통해 객체 탐지 기능을 제공해 주는 Aws Rekognition은 사용자가 객체를 탐

지할 수 있게 해주는 서버 기반의 컴퓨터 비전 기술 서비스이다. AWS Rekognition은 구현 시에 딥러닝 모델을 구축할 필요가 없고 성능의 정확도가 일정 수준 이상 보장되어 있다는 점에서 편리성이 뛰

\* Kyungpook National University, hanjinho01@naver.com

+ kjs@knu.ac.kr

어난 장점을 갖고 있어 다양한 사업 분야에서 활발히 사용되고 있으며 범용적인 활용성을 가진다. [1] 일반적인 CNN기반의 네트워크 모델은 스마트폰, 임베디드 보드와 같은 소형기기에서의 계산 자원의 제약으로 딥러닝을 적용하기 어려운 한계에 직면한다.

$$G_{ij} = (F * X)(i, j) = \sum_{m=0}^{F_H-1} \sum_{n=0}^{F_w-1} F_{m,n} X_{(i-m), (j-n)}$$

[2] CNN연산에서의 행렬로 표현되는 이미지의  $i$  번째 행,  $j$  번째 열의 픽셀을 구하는 수식은 다음과 같다.  $i$  번째 행,  $j$  번째 열의 픽셀인  $G_{ij}$ 는 다음과 같이 원본 이미지  $X$ 와 커널  $F$ 간의 합성곱 (Convolution) 연산으로 얻을 수 있다. 이 식을 통해 이미지 처리 과정에 주요 사용되는 CNN의 진행과정에서 상당한 연산량이 필요함을 알 수 있다. 그렇기에 키오스크와 같은 저사양 기기에서 객체 탐지 딥러닝을 적용하는 것은 성능 개선 및 구현에 있어 분명한 제약 사항이 있고, 이에 서버 영역에서 이미지를 처리하여 키오스크로 데이터를 보내는 방식으로 키오스크의 얼굴인식 기능을 구현하게 되었다. 해당 기능 구현에 AWS Rekognition을 활용한다.

는 인물일 경우 return 값으로 받아온 JSON 데이터에서 등록된 Face Id를 추출한다. 그 후 Face Id를 key로 사용하는 서버 내 데이터베이스에서 해당 인물이 선호하는 메뉴(가장 많이 시킨 메뉴 sorting)를 기기에 다시 반환해준다. 등록이 되어있지 않은 인물의 경우 IndexFaces()로 새로운 Face Id를 부여한 후 DetectFaces()를 통해 성별, 연령을 추출하여 추천하는 메뉴를 다시 기기에 반환해준다. 이 과정을 나타낸 흐름도는 그림 2와 같다.

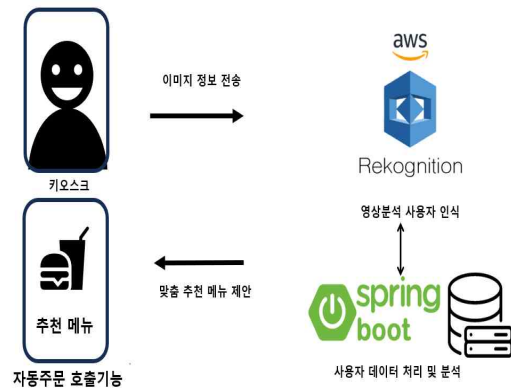


그림 1. AI 얼굴인식 자동 주문 기능 모듈 구성도  
Fig. 1. Configuration diagram of AI face recognition automatic ordering function module

## II. AI 얼굴인식 자동 주문 모듈 구현

본 논문에서 제안하는 키오스크에서의 얼굴인식 자동 주문 기능 모듈은 그림 1과 같다. 그림 1에서의 이미지 정보 입력 부분은 안드로이드 기반의 키오스크 영역과 스프링부트로 구성된 서버의 영역으로 구성되어 있다.

그림 1에서 키오스크의 영역은 사용자로부터 이미지 정보를 제공 받아 비트맵 형식으로 저장한 후 생성한 비트맵을 String형식으로 전환하여 서버에 전송한다. 서버 영역은 기기로부터 전송받은 String 데이터를 다시 bitmap 형식으로 전환하여 이미지 분석 processing을 진행한다.

AWS Rekognition api의 SearchFacesByImage()를 호출하여 이미지 데이터의 인물이 기존의 컬렉션에 등록된 인물인지 파악한다. 등록이 되어 있

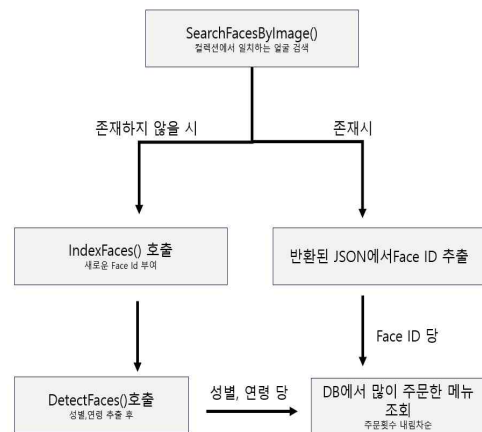


그림 2. AI 얼굴인식 자동 주문 기능 흐름도  
Fig. 2. Flow chart of AI face recognition automatic ordering function module

다시 기기 영역에서는 반환받은 데이터를 바탕으로 사용자에게 추천 메뉴를 주문 할 것인지 안내하는 방식으로 AWS Rekognition을 활용한 AI 얼굴인식 자동 주문 기능 모듈을 구현할 수 있다.

### III. 결 론

본 논문에서는 AWS Rekognition을 활용하여 사용자에게 추천 메뉴를 안내해주는 키오스크 시스템을 제안하였다. 제안된 방법은 주문과정에서의 사용자의 편의성을 증진시켜주고 디지털 취약 계층의 디지털 키오스크 사용의 불편함을 개선할 수 있을 것이다.

#### 감사의 글

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(2021-0-01082)

### 참 고 문 헌

- [1] 이한해솔, “A Study on Optimization of Light-weight Deep Learning for Object Detection”, February 2020.
- [2] 지능형인프라기술연구단, "딥러닝 모델 경량화 기술 November 2020.

# VSS-LMS를 이용한 FPGA 기반 적응형 필터 설계 및 구현

이정빈\*, 이종현\*, 박지섭\*, 정재진\*\*

## FPGA-Based Design and Implementation of an Adaptive Filter Using VSS-LMS Algorithm

Jeongbin Lee\*, Jonghyeon Lee\*, Jiseop Park\*, and Jaejin Jeong\*\*

### 요 약

적응형 필터는 다양한 환경에서 변화하는 신호를 실시간으로 처리하기 위해 필터 계수를 동적으로 조정하는 기술로, LMS(Least Mean Squares) 알고리즘은 간결한 구조와 계산의 효율성 덕분에 널리 사용되어 왔다. VSS(Variable Step Size) 알고리즘은 LMS 알고리즘의 수렴 속도를 높이기 위해 학습 속도를 동적으로 조절하는 기법이다. 본 논문에서는 이러한 LMS 알고리즘과 VSS 알고리즘을 이용해 FPGA 기반 적응형 필터를 설계했다. FPGA를 활용해 신호 처리 성능을 분석한 결과, VSS 알고리즘이 LMS 필터의 성능을 개선하는 데 도움이 되는 것을 확인할 수 있었으며, 이를 통해 향후 다양한 신호 처리 분야에 효율적으로 적용될 수 있을 것으로 기대된다.

### Abstract

An adaptive filter is a technology that dynamically adjusts filter coefficients to process signals in real-time across various environments. The LMS (Least Mean Squares) algorithm has been widely used due to its simple structure and computational efficiency. The VSS (Variable Step Size) algorithm is a technique that adjusts the learning rate dynamically to improve the convergence speed of the LMS algorithm. In this paper, an FPGA-based adaptive filter was designed using both LMS and VSS algorithms. The analysis of signal processing performance using FPGA demonstrated that the VSS algorithm helps enhance the performance of LMS filters. This suggests that it can be efficiently applied in various signal processing fields in the future.

### Key words

adaptive filter, LMS algorithm, Variable Step Size, FPGA implementation, signal processing

### I. 서 론

LMS(Least Mean Squares) 알고리즘은 구조가 간단하고 계산 효율성이 뛰어나, 주로 실시간 신호

처리 시스템에서 잡음 제거, 음성 인식, 채널 균등화 등에 활용된다[1]. 그러나 고정된 스텝 사이즈로 인해 성능에 한계가 있으며, 스텝 사이즈가 작으면 느리게 수렴하고, 크면 발산할 위험이 있다. 이를

\* 국립금오공과대학교 전자공학부 전자 및 전파 전공 (학사과정), {wjdqsl9931, p2849, whdgus335} @naver.com

\*\* 국립금오공과대학교 전자공학부 전자 및 전파, jeong@kumoh.ac.kr

해결하기 위해 VSS(Variable Step Size) 알고리즘이 도입되었으며, 신호 특성에 따라 스텝 사이즈를 동적으로 조정하여 LMS의 수렴 속도와 안정성을 향상시키고 안정적인 필터 성능을 제공한다[2].

FPGA(Field Programmable Gate Array)는 이러한 알고리즘을 구현하는 데 있어 다양한 장점을 제공한다. FPGA는 높은 병렬 처리 능력을 통해 빠르고 효율적인 신호 처리를 가능하게 하며, 실시간 시스템에서 낮은 지연 시간과 높은 처리량을 제공한다. 또한, FPGA는 유연한 설계 변경이 가능하여 알고리즘의 수정 및 최적화를 쉽게 수행할 수 있다.

본 논문에서는 LMS 알고리즘과 VSS 알고리즘을 결합한 FPGA 기반 적응형 필터 설계를 제안하며, 이를 통해 다양한 신호 처리 환경에서 성능 최적화를 목표로 한다. LMS 및 VSS-LMS 알고리즘을 FPGA 상에서 구현함으로써, 실시간 신호 처리 환경에서 발생할 수 있는 성능 저하 문제를 개선할 수 있으며, 신호 변화가 급격한 환경에서도 더 빠르고 안정적인 수렴을 보장할 수 있다.

## II. VSS-LMS 알고리즘

LMS 알고리즘은 출력 신호와 목표 신호 간의 오차를 줄이기 위해 필터 가중치를 실시간으로 업데이트한다. 목표 신호와 필터 출력 간의 오차를 계산하고, 그 오차의 제곱 평균을 최소화하는 방식으로 가중치를 갱신한다. FIR 필터와 결합하여 성능을 높일 수 있지만, 필터 계수를 실시간으로 계산하고 자동으로 업데이트하는 하드웨어가 추가되므로 구현의 복잡도가 증가한다.

LMS 알고리즘에서 출력 신호  $y(n)$ 은 입력 신호와 필터 가중치의 곱의 합으로 계산된다. 이를 수식으로 표현하면  $y(n)$ 은 각 입력 신호  $x_i(n)$ 에 해당하는 가중치  $w_i(n)$ 과의 곱을 합산한 값으로, 다음과 같이 수식(1)으로 나타낼 수 있다.

$$y(n) = \sum_{i=0}^N w_i(n) \cdot x_i(n) \quad (1)$$

여기서,  $w_i(n)$ 은 각 탭의 가중치를 나타내며,  $x_i(n)$ 은 각 탭에서 입력된 신호이다.  $N$ 은 탭의 길

이를 나타내며, 오차  $e(n)$ 은 다음과 같은 수식(2)을 따른다.

$$e(n) = d(n) - y(n) \quad (2)$$

이 오차는 LMS 알고리즘의 학습 과정에서 필터 계수를 업데이트하는 데 사용된다. 필터 계수는 다음과 같은 수식(3)으로 업데이트가 진행된다:

$$w_i(n+1) = w_i(n) + \mu \cdot e(n) \cdot x_i(n) \quad (3)$$

여기서  $\mu$ 는 학습률을 나타내며, 오차  $e(n)$ 와 입력 신호  $x_i(n)$ 에 따라 필터의 계수가 조정된다.

$w_i(n)$ 은  $n$ 번째 샘플에서  $i$ 번째 탭의 가중치를 의미하며,  $n+1$ 은 다음 샘플에서 업데이트된 가중치를 나타낸다.

VSS 알고리즘은 고정 스텝 사이즈로 사용되는 LMS 알고리즘의 문제를 해결하기 위해 설계되었다. 이는 입력 신호의 특성에 따라 학습률을 동적으로 조정한다. 신호 변화가 클 때 학습률  $\mu$ 을 높여 빠르게 수렴하도록 하며, 오차가 줄어들면 학습률을 낮추어 안정성을 유지한다. 이를 통해 다양한 환경에서 안정적인 성능을 유지할 수 있다.

학습률  $\mu$ 는 아래와 같은 수식(4)으로 표현된다 [3].

$$\mu(n) = \alpha \mu(n-1) + \gamma e(n)^2 \quad (4)$$

이 수식을 통해 학습률은 실시간으로 오차  $e(n)$ 의 크기에 비례하여 조정된다. 본 논문에서 설계된 적응형 필터의 구성도는 그림 1과 같다.

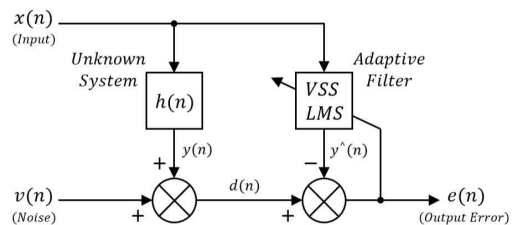


그림 1. VSS-LMS 알고리즘 적응형 필터 시스템 구성도  
Fig. 1. Adaptive filter system using the VSS-LMS algorithm

### III. FPGA에서의 알고리즘 구현

#### 3.1 FPGA에서의 알고리즘 구현 및 검증

FPGA에 구현하기에 앞서, LMS 알고리즘의 성능을 테스트벤치 시뮬레이션으로 검증하였다. MATLAB으로 생성한 입력 신호  $x(n)$ 과 목표 신호  $d(n)$ 를 기반으로 시뮬레이션을 진행하여, FPGA상의 실제 동작을 모사하였다. 데이터셋은 학습률  $\mu = 0.05$ 에서 2000개의 샘플로 LMS 필터 성능을 평가하며, 필터의 탭 길이는 4로 설정되었다. 필터링 과정에서 잡음 신호는 표준편차 0.01의 가우시안 노이즈로 설정되었다.

$\alpha = 0.95$ ,  $\gamma = (1 - \alpha)$  파라미터를 사용하였으며, 오차  $e(n)$ 와 입력 신호 크기에 비례해 학습률을 실시간으로 계산하고, 빠르고 안정적인 성능을 제공한다.

#### 3.2 실험 결과 및 분석

그림 2는 LMS 알고리즘과 VSS-LMS 알고리즘의 MSD 값을 시각화한 그래프이다. 여기서 MSD(Mean Square Deviation)는 필터 계수  $w$ 와 최적의 가중치  $w_{opt}$  사이의 차이를 제공하여 평균낸 값으로, 필터가 최적의 가중치에 얼마나 근접하게 수렴하는지를 평가하는 지표로 사용된다. MSD 값이 낮을수록 오차가 적어 필터가 원하는 신호에 더 근접하게 수렴함을 의미한다.

VSS-LMS 필터는 초기 수렴 속도에서 LMS 필터보다 빠르게 수렴하며, 학습률을 동적으로 조정해 변화가 많은 환경에서도 안정적인 성능을 유지하였다. 반면, LMS 필터는 고정된 학습률로 인해 수렴 속도가 느리고, 잡음이 많은 구간에서는 VSS-LMS에 비해 더 큰 오차를 보였다. 학습 후반부에서는 두 알고리즘 모두 비슷한 수렴 경향을 나타내지만, 최종적으로 VSS-LMS가 LMS보다 낮은 오차율을 유지한다.

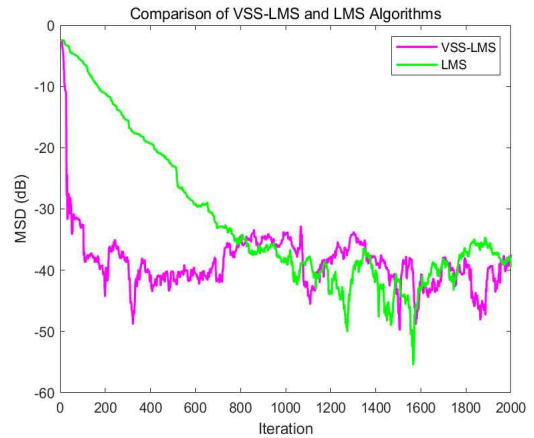


그림 2. LMS 필터와 VSS-LMS 필터의 성능 비교  
Fig. 2. Performance Comparison of LMS and VSS-LMS Filters

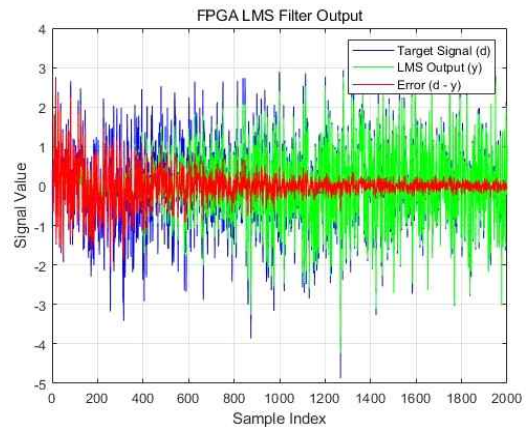


그림 3. FPGA에서 출력한 LMS 오차 데이터  
Fig. 3. LMS error data output from the FPGA

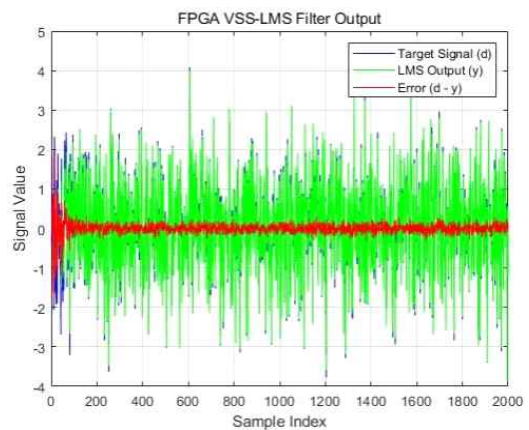


그림 4. FPGA에서 출력한 VSS-LMS 오차 데이터  
Fig. 4. VSS-LMS error data output from the FPGA



그림 3과 그림 4는 FPGA에서 처리한 출력 오차 데이터를 UART 통신으로 전송받아 시각화한 결과를 보여준다. 그림 3은 FPGA에서 기존의 LMS 필터를 구현하여 출력한 오차이며 그림 4는 VSS 알고리즘이 도입된 LMS 필터의 오차이다.

FPGA 출력 결과, VSS-LMS 알고리즘은 약 150개의 샘플 후에 안정적으로 수렴한 반면, LMS 알고리즘은 약 800개의 샘플 후에 수렴하였다. 결론적으로 두 알고리즘의 출력값을 비교한 결과, VSS-LMS는 초기 오차를 급격히 줄이며 LMS 알고리즘보다 빠른 수렴 속도를 보여주었다. 이를 통해 VSS-LMS 알고리즘을 적용함으로써 더욱 향상된 수렴 속도를 확인할 수 있었다.

## V. 결 론

본 논문에서는 LMS와 VSS 알고리즘을 결합한 FPGA 기반 적응형 필터를 설계하고, 그 성능을 수치적으로 비교하였다. 실험 결과, VSS-LMS 알고리즘은 LMS 알고리즘에 비해 빠르게 초기 오차를 줄이며 안정적으로 수렴하는 성능을 보여주었다. VSS-LMS 알고리즘은 학습률을 동적으로 조절함으로써 LMS 알고리즘을 사용했을 때 보다 더 빠르고 효율적인 수렴 속도를 제공하며, 초기 단계에서의 오차 감소가 더욱 우수함을 확인할 수 있었다. 이를 통해 VSS-LMS 알고리즘을 적용함으로써 수렴 속도가 크게 향상되는 장점을 확인할 수 있었다.

이러한 결과는 향후 잡음 제거, 음성 신호 처리, 통신 시스템의 채널 균등화 등 고속 신호 처리 요구가 높은 응용 분야에 유용하게 활용될 수 있다.

## 참 고 문 헌

[1] S. S. Godbole, P. M. Palsodkar and V. P. Raut, "FPGA Implementation of Adaptive LMS Filter", Proceedings of SPIT-IEEE Colloquium and International Conference, vol. 2, pp. 226-229, 2011.

[2] 우홍채, "결정함수 가변스텝 LMS 알고리즘", 융합신호처리학회 논문지, vol. 12(2), pp. 128-132,

2011.

[3] Liu. Yuqing, et al, "implementation of LMS Algorithm Based Adaptive Filter Using Verilog HDL", 2023 IEEE 16th International Conference on Electronic Measurement & Instruments (ICEMI). IEEE, pp. 463-467, August 2023.

[4] M. Li, L. Li and H. Tai. "Variable Step Size LMS Algorithm Based on Function Control,"Circuits Systems & Signal Processing, vol. 32, No. 6, pp, 3121-3130, 2013.

[5] A. Rosado-Muñoz, M. Bataller-Mompeán, E. Soria-Olivas, C. Scarante, and J. F. Guerrero-Martínez, "FPGA Implementation of an Adaptive Filter Robust to Impulsive Noise: Two Approaches," IEEE Transactions on Industrial Electronics, vol. 56, no. 3, pp. 860-870, June 2009.

[6] H. Oba, M. Kim and H. Arai. "FPGA Implementation of LMS and N-LMS Processor for Adaptive Array Applications" 2006 International Symposium on Intelligent Signal Processing and Communications, pp. 485-488, 2006.

[7] S.J. Visser, A.S. Dawood, J.A. Williams, "FPGA based real-time adaptive Filtering for space applications," Proceedings, IEEE International Conference on Field-Programmable Technology, Pages: 322-326, 2002.

# YOLOv10 기반 저조도 환경에서 객체 인식 성능 향상을 위한 이미지 전처리 기법 비교

이성욱\*, 이용진\*, 김진재\*, 지승한\*\*

## Comparison of Image Preprocessing Techniques for Improving Object Detection Performance in Low-Light Environments Using YOLOv10

Seongwook Lee\*, Yongjin Lee\*, Jinjae Kim\*, Seunghan Ji\*\*

### 요 약

자율주행 차량의 안전성을 강화하기 위해 저조도 및 악천후 조건에서의 객체 인식 성능 개선이 지속적으로 해결해야 할 중요한 과제로 다뤄져 왔다. 본 연구에서는 이러한 환경에서 객체 인식 성능을 향상하기 위해 YOLOv10 모델을 활용하였으며, 다양한 이미지 전처리 기법을 적용하고 그 효과를 분석했다. 히스토그램 평활화, CLAHE, 감마 보정 등의 기법을 사용하여 객체 인식 성능을 개선할 수 있는 방법론을 제시하며, 이를 통해 저조도 환경에서의 객체 인식 정확도를 향상하기 위한 방법을 제시하고 분석한다.

### Abstract

In order to enhance the safety of autonomous vehicles, improving object detection performance in low-light and bad weather conditions has been continuously addressed as an important task to be solved. In this study, the YOLOv10 model was used to improve object detection performance in this environment, and various image preprocessing techniques were applied and their effects were analyzed. We present a methodology to improve object detection performance using techniques such as histogram equalization, CLAHE, and gamma correction, and through this, we present and analyze methods to improve object detection accuracy in low-light environments.

### key words

Autonomous Vehicles, YOLOv10, Object Detection, Low-light Conditions, Image Preprocessing

## 1. 서 론

자율주행 기술의 발전과 함께 자율주행 차량(AV)의 안전성은 중요한 연구 주제가 되어가고 있다.

컴퓨터 비전 기술은 이러한 차량의 주행 환경 인식, 객체 탐지, 차선 추적 등의 기능을 수행하는 데 핵심적인 역할을 한다. 특히 야간, 눈, 비, 안개와 같은 이러한 저조도 조건에서의 자율주행 성능을 개

\* 대림대학교 메카트로닉스공학과. [solee890@naver.com](mailto:solee890@naver.com)

\*\* 대림대학교 메카트로닉스공학과. [jsh2012@daelim.ac.kr](mailto:jsh2012@daelim.ac.kr)

※ 이 연구는 과학기술정보통신부의 재원으로 한국지능정보사회진흥원의 지원을 받아 구축된 “데이터명”을 활용하여 수행된 연구입니다. 본 연구에 활용된 데이터는 AI 허브(aihub.or.kr)에서 다운로드 받으실 수 있습니다.

선하기 위한 연구는 아직 부족한 상황이다[1]. 최근 발표된 연구에 따르면, 자율주행 차량은 일반적인 주행 상황에서는 인간 운전자 차량에 비해 사고 발생 확률이 낮지만, 새벽과 황혼, 그리고 회전과 같은 특정 조건에서는 사고 발생 확률이 크게 증가하는 것으로 나타났다. 예를 들어, 새벽이나 황혼과 같은 저조도 조건에서는 자율주행 차량의 사고 발생 확률이 인간 운전자 차량보다 5.25배 더 높았으며, 회전 시에는 1.98배 더 높은 확률로 사고가 발생하는 것으로 나타났다[2]. 이러한 연구 결과는 자율주행 시스템의 안전성을 개선하기 위해 열악한 조명 조건에서의 객체 인식 성능을 향상하는 것이 필요함을 시사한다. 본 연구에서는 YOLOv10 모델을 활용하여 야간, 눈, 비, 안개와 같은 저조도 환경에서 촬영된 도로 주행 이미지를 학습시키고, 전처리 기법들이 객체 인식 성능에 어떤 영향을 미치는지를 평가하여, 최적의 성능을 도출하는 기법을 분석한다.

## II. 이미지 전처리 기법

### 2.1 Histogram Equalization

히스토그램 평활화(Histogram Equalization)는 컴퓨터 비전 분야에서 널리 사용되는 기법이다. 입력 이미지의 픽셀값들을 조정하여 이미지의 명암 분포를 고르게 만드는 기법이다. 이를 통해 이미지의 전반적인 대비가 증가하고, 저조도 이미지에서 디테일이 보다 명확하게 나타나게 된다. 해당 기법은 식(1)과 같이 정의된다.

$$I_{eq}(x,y) = \frac{L-1}{M \times N} \sum_{i=0}^{I(x,y)} h(i) \quad (1)$$

위 식에서  $I(x,y)$ 는 입력 이미지의 픽셀값을,  $I_{eq}(x,y)$ 는 평활화된 이미지의 픽셀값을 나타낸다.  $M \times N$ 은 이미지의 전체 픽셀 수이며,  $L$ 은 이미지에서 표현할 수 있는 최대 픽셀값을 의미한다. 이 기법을 통해 저조도 이미지에서도 더 명확한 디테일을 얻을 수 있다. 그림 1에서 원본 이미지(좌측)와 히스토그램 평활화가 적용된 이미지(우측)를

비교하면, 원본 이미지에서는 밝기 차이가 뚜렷하지 않아 세부 사항이 잘 드러나지 않지만, 히스토그램 평활화를 적용한 이미지에서는 명암 분포가 고르게 되어 선명한 세부 사항을 확인할 수 있다. 그러나 모든 이미지에서 동일한 효과를 기대할 수는 없으며, 색상 왜곡이나 노이즈가 증가하는 부작용이 발생할 수 있다. 특히, 컬러 이미지에 적용할 경우, 각 채널에서의 비일관성으로 인해 색상 변화가 발생하고, 결과적으로 이미지 품질이 저하될 수 있다[3]. 이러한 한계를 인식하고 상황에 맞는 적용이 중요하다.

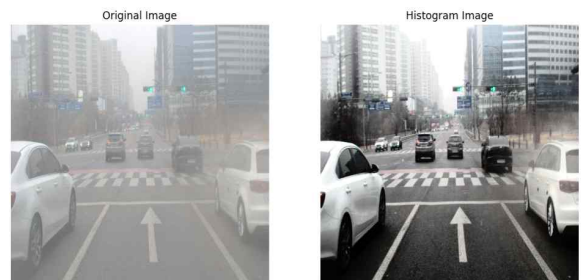


그림 1. 원본 이미지와 히스토그램 평활화 적용 이미지 비교

Fig 1. Comparison between the original image and the Histogram Equalization image

### 2.2 CLAHE

CLAHE(Contrast Limited Adaptive Histogram Equalization)는 과도한 명암비 증가와 노이즈 증폭과 같은 문제를 해결하기 위해 기존 히스토그램 평활화 기법을 개선한 이미지 대비 향상 기법이다. 이 기법은 전체 이미지를 여러 개의 작은 타일로 나눈 후, 각 타일에서 히스토그램을 계산하여 명암비를 조정한다. 이러한 과정에서 특정 임계값을 설정하여 히스토그램을 클리핑(Clipping)으로써, 명암비가 과도하게 증가하는 것을 방지하고, 노이즈 증폭도 억제할 수 있다. 조정된 명암비 값은 타일 간의 경계에서 부드럽게 연결되도록 인접한 타일들과 보간되며, 이를 통해 이미지 전체에서 자연스럽게 일관된 명암비 전환이 이루어진다[4]. 그림 2의 CLAHE 적용 결과를 보면, 로컬 대비가 조정되어 이미지의 세부 사항이 더욱 선명하게 표현되는 것을 확인할 수 있다. CLAHE의 가장 큰 장점은 로컬 대비를 효과적으로 조정하여 이미지 세부 사항을 더욱 뚜렷

하게 만드는 기능이다. 이는 의료 영상 처리나 고르지 못한 조명 조건에서 캡처한 영상에 특히 유용하다[5]. 그러나 이 기술은 타일 분할 및 보간 과정으로 인해 계산 복잡도가 증가한다는 단점이 있다.



그림 2. 원본 이미지와 CLAHE 적용 이미지 비교  
Fig 2. Comparison between the original image and the CLAHE image

### 2.3 감마 보정

감마 보정은 이미지의 밝기와 대비를 조정하기 위한 기법으로, 입력 값과 출력 값 사이의 비선형 변환을 사용하여 이미지의 시각적 품질을 향상한다. 이 기법은 픽셀값의 비선형 변환을 통해 이미지의 밝기를 조정하며, 어두운 부분의 디테일을 더욱 명확하게 표현할 수 있도록 한다. 특히, 저조도 환경에서도 시각적 선명도를 개선하는 데 효과적이다 [6]. 해당 기법은 식(2)과 같이 정의된다.

$$I_{Out}(x,y) = 255 \times \left( \frac{I_{in}(x,y)}{255} \right)^{\frac{1}{r}} \quad (2)$$

위 식에서는 보정된 출력 이미지의 픽셀값이며, 원본 입력 이미지의 픽셀값이다. 감마 값( $r$ )은 이미지의 밝기 조정을 결정하는 파라미터이다. 값이 1보다 작을 때는 어두운 부분이 더 밝게 표시되며, 1보다 클 때는 밝은 부분이 더 어둡게 표현된다. 이를 통해 사용자는 이미지의 밝기를 세밀하게 조정할 수 있으며, 저조도 이미지에서는 감마 값을 조절하여 어두운 부분의 디테일을 강화할 수 있다. 그림 3의 감마 보정 적용 결과는 이미지의 명암 대비와 디테일에 미치는 영향을 시각적으로 보여준다. 감마 보정의 주요 장점은 저조도 이미지에서 디테일을 개선하고 명암 대비를 조절하여 더 선명한 이미지를 제공할 수 있다는 점이다. 그러나 감마 보

정은 컬러 이미지에서 색상 왜곡을 초래할 수 있으며, 어두운 영역에서 노이즈가 더 두드러질 수 있다[7].



그림 3. 원본 이미지와 감마 보정 적용 이미지 비교  
Fig 3. Comparison between the original image and the Gamma Corrected image

### III. YOLO 모델 개요

YOLO(You Only Look Once)는 2016년에 발표된 객체 검출 모델로, 객체의 위치를 실시간으로 파악해야 하는 자율 주행과 같은 프로세스에 적합한 모델 중 하나다. 기존의 객체 검출 모델들은 2-stage detection을 사용하는 반면, YOLO는 분류(Classification)와 위치 추정(Localization)을 동시에 수행하는 1-stage detection 방식을 채택하여, 실시간 객체 검출이 가능하도록 설계되었다[8].

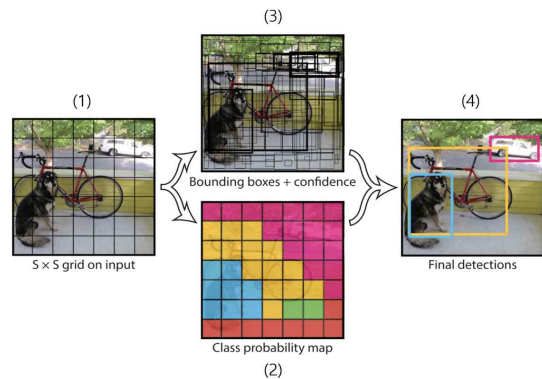


그림 4. YOLO 모델을 통한 물체 탐지 과정  
Fig 4. Object detection process using YOLO model

본 연구에서는 작성 시점 최신 버전인 YOLOv10 모델을 사용했다. YOLOv10은 새로운 훈련 전략과 모듈 도입으로 성능이 향상되었다. 특히, YOLOv10은 NMS(Non-Maximum Suppression)를 대체하는 새로운 훈련 전략인 consistent dual assignments를 도입하여, 중복 객체 감지 문제를 해결하면서도 추론

속도를 대폭 개선했다. 또한, 대형 커널 합성과 부분적 셀프 어텐션(Partial Self Attention, PSA) 모듈의 도입으로 모델의 전반적인 성능이 강화되었으며, 소형 모델에서도 높은 정확도를 유지할 수 있게 되었다. YOLOv10의 소형 모델은 YOLOv8과 비교하여 약 0.5 mAP의 성능 향상을 보이며, 매개변수가 2배 이상 줄어들었음에도 불구하고 더 높은 성능을 기록하였다[9]. 이러한 개선 사항들은 YOLOv10이 YOLOv8에 비해 성능과 효율성에서 더욱 진보된 모델임을 입증한다.

#### IV. 데이터셋 및 전처리

본 연구에서는 저조도 환경에서의 이미지를 수집하였으며, 그중 야간, 황혼, 안개, 눈, 비 등 악천후 상황에서 촬영된 이미지들을 선별하여 총 5개의 객체(차량, 횡단보도, 차선, 가드레일, 표지판)를 선정하고 직접 polygon 기법을 사용해 데이터 라벨링을 수행했다. 총 3,127장의 1920\*1080 크기의 이미지 데이터셋을 구축했으며, 데이터 증강(Data Augmentation)을 적용하여 총 12,508장의 이미지로 증강했으며 이미지 크기는 유지했다. 이후, 앞서 제안한 전처리 기법들을 각각의 데이터셋에 적용하여, 총 4개의 데이터셋을 준비했다. 하이퍼 파라미터는 원본 데이터셋 학습 기준 가장 성능이 높았던 기준으로 설정했다. 표 1은 학습에 사용된 주요 하이퍼 파라미터 설정을 나타낸다.

표 1. 하이퍼 파라미터 설정

Table 1. Hyperparameter Configuration

Epoch	100
Batch size	32
Optimizer	AdamW
Patience	5

원본 데이터셋과 전처리 기법을 적용한 3개의 데이터셋은 학습(train), 검증(valid), 테스트(test) 비율을 동일한 조건인 7:2:1로 구성하였으며, 이미지는 640 \* 640 크기로 축소하여 학습에 사용했다. YOLOv10 모델의 경우, 컴퓨팅 자원을 고려하여 YOLOv10s를 선택했고, 과적합을 방지하기 위해 조기 종료

(Early Stopping) 기준인 patience를 5로 설정하여 학습을 진행했다.

#### V. 결과

YOLOv10 모델의 저조도 환경에서의 객체 인식 성능을 평가하기 위해 다양한 이미지 전처리 기법을 적용한 실험 결과는 다음과 같다.

표 2. 실험 결과

Table 2. Experiment results

기법	mAP50	mAP50-95
원본	0.769	0.558
Histogram	0.823	0.671
Clahe	0.926	0.813
감마 보정	0.87	0.746

표 2의 실험 결과에서 확인할 수 있듯이, CLAHE 기법을 적용한 모델은 mAP50에서 0.926, mAP50-95에서 0.813로 가장 우수한 성능을 기록했다. 이러한 결과는 CLAHE가 전역적인 대비 조정 대신 로컬 영역에서의 대비를 세밀하게 조정하여 이미지의 세부 사항을 더 명확하게 표현했기 때문으로 판단된다. 특히, 저조도 환경에서는 어두운 영역의 디테일을 보강하는 것이 중요하며, CLAHE는 노이즈 발생을 억제하면서 밝기 변화에 민감하게 반응하여 디테일 손실을 효과적으로 최소화했다. 감마 보정 기법도 비교적 높은 성능을 보여 mAP50에서 0.87, mAP50-95에서 0.746을 기록했다. 이는 비선형 밝기 조정을 통해 어두운 영역의 디테일을 강조하는 데 효과적이었다. 하지만 감마 값이 적절하게 설정되지 않으면 이미지의 어두운 부분이 지나치게 밝아지거나, 밝은 부분이 과도하게 어두워져 색상 왜곡과 시각적 불균형이 발생할 수 있다. 이러한 이유로 감마 보정의 성능이 CLAHE보다 낮은 것으로 해석된다. 히스토그램 평활화 기법은 mAP50에서 0.823, mAP50-95에서 0.671로 성능이 개선되었지만, 다른 기법들에 비해 성능이 다소 낮았다. 이 기법은 전역적인 명암비 조정을 통해 이미지의 대비를 향상시키지만, 일부 영역에서 노이즈가 증가하거나 과도한 대비가 발생할 수 있다. 이러한 특성은 이미지의 세부 정보가 중요한 저조도 환경에서 부정적인 영

향을 미치며, 최적의 성능을 보장하지 못하는 원인으로 작용할 수 있다. 한편, 전처리를 하지 않은 원본 데이터의 경우 mAP50에서 0.769, mAP50-95에서 0.558로 나타났다. 이는 전처리 기법이 성능 향상에 얼마나 기여하는지를 보여주는 기준점이 되며, 저조도 환경에서 이러한 기법의 적용이 필수적임을 강조한다.

## VI. 결론 및 향후 연구

본 연구에서는 YOLOv10 모델의 저조도 환경에서의 객체 인식 성능을 향상시키기 위해 CLAHE, 감마 보정, 히스토그램 평활화 등 다양한 이미지 전처리 기법을 적용하고 그 효과를 비교했다. 연구 결과 CLAHE 기법이 로컬 대비를 세밀하게 조정하고 노이즈를 효과적으로 억제하여 가장 우수한 성능을 발휘했다. 이를 통해 CLAHE 기법이 저조도 환경에서의 객체 인식 성능 향상에 유용한 전처리 기법임을 확인했다. 향후 연구에서는 다양한 저조도 환경, 객체 크기, 복잡한 배경 조건에서의 성능을 추가적으로 평가하여 모델의 일반화 성능을 더욱 강화할 필요가 있다. 또한, CLAHE, 감마 보정, 히스토그램 평활화를 결합한 하이브리드 전처리 기법을 개발하는 가능성도 모색해 볼 수 있으며 저조도 환경에서 객체 인식 성능을 향상시킬 수 있는 이미지 전처리 기법을 제시함으로써 관련 분야의 추가 연구에 기여할 수 있을 것이다.

## 참 고 문 헌

[1] D. Kumar and N. Muhammad, "Object detection in adverse weather for autonomous driving through data merging and YOLOv8," *Sensors*, vol. 23, no. 8471, pp. 1-29, Oct. 2023.

[2] M. Abdel-Aty and S. Ding, "A matched case-control analysis of autonomous vs human-driven vehicle accidents," *Nature Communications*, vol. 15, no. 4931, pp. 1-9, May 2024.

[3] B. S. Rao, "Dynamic histogram equalization for contrast enhancement for digital images," *Applied*

*Soft Computing Journal*, vol. 89, pp. 1-11, Jan. 2020.

[4] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355-368, 1987.

[5] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, and K. E. Muller, "Contrast-limited adaptive histogram equalization: speed and effectiveness," *Proc. of the First Conf. on Visualization in Biomedical Computing*, pp. 337-345, 1990.

[6] S. Rahman, M. M. Rahman, M. A. Al-Wadud, G. D. Al-Quaderi, and M. Shoyaib, "An adaptive gamma correction for image enhancement," *EURASIP Journal on Image and Video Processing*, vol. 2016, no. 35, pp. 1-13, 2016.

[7] Y. Wang, Z. Liu, J. Liu, S. Xu, and S. Liu, "Low-light image enhancement with illumination-aware gamma correction and complete image modelling network," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 13128-13137, 2023.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, 2016.

[9] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, and G. Ding, "YOLOv10: Real-time end-to-end object detection," *arXiv preprint arXiv:2405.14458*, 2024.

# ADHD 청소년 환자를 위한 디지털 놀이치료 알고리즘

곽윤석, 김도연, 박정수, 안진호

## Digital Play Therapy Algorithm for ADHD Adolescent Patients

Yunsuk Kwak, Doyeon Kim, Jeongsoo Park, and Jinho Ahn

### 요약

최근 ADHD(주의력 결핍 과잉 행동 장애)는 청소년과 성인들 사이에서 신경발달 장애로서 큰 영향을 미치고 있는 문제입니다. 특히, 청소년기에서의 조기진단과 치료가 중요한데, 이는 성인기에서의 증상 악화를 예방하고 사회적 어려움을 줄일 수 있기 때문입니다. 최근 연구들은 두뇌 훈련 게임이 ADHD 환자의 인지 능력 향상에 효과적일 수 있음을 제시하고 있습니다. 예를 들어, 기억력과 집중력을 향상시키는 기술을 적용한 게임들은 사용자의 인지 기능을 지원하며, ADHD 관리에 새로운 접근법을 제공할 수 있습니다. 본 연구는 ADHD 환자를 대상으로 한 디지털 놀이치료 추천 알고리즘을 제안하며, 이를 통해 개인 맞춤형 치료 방법의 필요성을 강조하고자 합니다.

### Abstract

Recently, ADHD (attention deficit hyperactivity disorder) is a problem that has a significant impact as a neurodevelopmental disorder between teenagers and adults. In particular, early diagnosis and treatment in adolescence are important because it can prevent worsening symptoms in adulthood and reduce social difficulties. A recent study shows that brain training games can be effective in improving cognitive performance in ADHD patients. For example, games with technology that improves memory and concentration can support users' cognitive functions and provide a new approach to ADHD management. This study proposes a digital play therapy recommendation algorithm for ADHD patients, and through this, we would like to emphasize the need for personalized treatment methods.

### 1. 서론

최근 ADHD(Attention Deficit Hyperactivity Disorder)는 전 세계적으로 청소년과 성인에게 큰 영향을 미치는 신경발달 장애로 주목받고 있습니다.

ADHD 환자의 약 57%가 청소년기에 해당하며, 이 중에서도 10~14세 환자는 전체의 33.9%를 차지하고 있습니다. 이는 ADHD의 조기진단 및 치료가 중요한 이유 중 하나입니다. 청소년기에 적절한 치료를 받지 못할 경우, 성인 ADHD로 이어질 가능성이 높

\*경기대학교 AI컴퓨터공학부, {lukeyun,dodoy21,0902jungsu}kyonggi.ac.kr

\*\*경기대학교 AI컴퓨터공학부 교수, jhahn@kyonggi.ac.kr(교신저자)

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음 (2021-0-01393).

으며, 이는 사회생활에서의 다양한 어려움을 초래하여 부정적 감정을 유발하고 치료에 대한 저항성을 증가시킬 수 있습니다. 따라서 ADHD 환자에 대한 조기 치료의 중요성은 더욱 부각되고 있습니다. 특히, 두뇌 훈련 게임은 이 연령대의 환자들에게 큰 흥미를 불러일으킬 수 있는 효과적인 치료 방법 중 하나로 평가받고 있습니다. 두뇌 훈련 게임은 인지 능력을 향상시키고, 기억력과 집중력을 강화하는 데 큰 역할을 할 수 있습니다.

## II. CBR기반 추천모델

본 연구에서는 청소년기 ADHD 환자를 대상으로 하는 게임 추천 시스템을 설계하기 위해 Case-Based Reasoning(CBR, 사례 기반 추론) 기법을 활용합니다. CBR은 경험을 바탕으로 문제를 해결하는 인공지능 기법으로, 이 경우에는 ADHD 환자가 어떤 게임에서 잘 반응하는지를 사례로 기록하고 이를 바탕으로 새로운 추천을 수행합니다.

사례 수집: ADHD 환자와 정상 대조군의 디지털 놀이치료 사용 데이터를 수집합니다. 각 사례에는 환자의 특성 (나이, 성별 등), 사용한 게임의 특성 (게임 유형, 게임 플레이 기록) 및 성과 (인지 능력 향상 여부)이 포함됩니다.

사례 기반 추론 모델 설계:

사례 유사성 평가: 새로운 환자의 특성과 기존 사례의 특성을 비교하여 유사도를 평가합니다.

사례 검색: 유사한 사례들을 찾아내고, 이들의 게임 사용 패턴을 분석하여 새로운 환자에게 적합한 게임을 추천합니다.

사례 적용: 추천된 게임을 통해 환자의 인지 능력을 향상시킬 수 있는지 평가합니다.

평가 및 결과 분석: 추천된 놀이치료가 환자의 인지 능력 향상에 미치는 영향을 평가하고, CBR 모델의 효과성을 분석합니다.

CBR 기반의 디지털 놀이치료 추천 시스템은 ADHD 환자의 개인 맞춤형 치료를 지원할 수 있으며, 이를 통해 인지 능력을 향상시키는 데 기여할 것으로 기대됩니다. 또한, 본 연구는 개인화된 치료 접근법의 가능성을 제시하고, 디지털 놀이치료가 ADHD 관리에 새로운 해결책으로서의 역할을 확대할 수 있는 방법을 탐구할 것입니다.

접근법의 가능성을 제시하고, 디지털 놀이치료가 ADHD 관리에 새로운 해결책으로서의 역할을 확대할 수 있는 방법을 탐구할 것입니다.

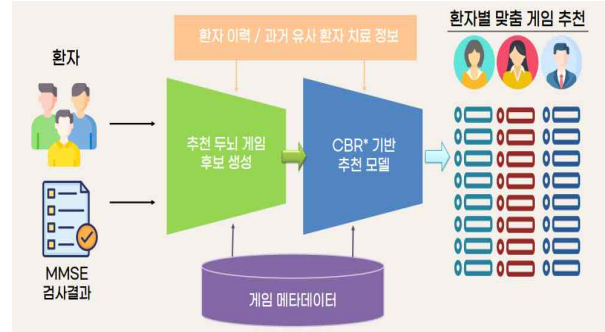


그림 1. 전체 시스템 구조  
Fig 1. Overall System Architecture

## III. 결 론

본 연구는 ADHD 청소년기 환자를 위한 디지털 놀이치료 추천 알고리즘을 제안하여, 조기 진단과 맞춤형 치료의 중요성을 강조합니다. 기억력 및 집중력 훈련 게임을 통해 ADHD 환자의 인지 능력을 향상시키고, 개인별 맞춤형 접근을 통해 치료의 효과를 극대화할 수 있음을 확인하였습니다. 이러한 접근은 환자들의 삶의 질을 향상시키는 데 기여할 수 있으며, 향후 연구를 통해 더욱 발전된 치료 방법을 제시하고자 합니다. CBR 기반의 디지털 놀이치료 추천 시스템은 ADHD 환자의 개인 맞춤형 치료를 지원할 수 있으며, 이를 통해 인지 능력을 향상시키는 데 기여할 것으로 기대됩니다. 또한, 본 연구는 개인화된 치료 접근법의 가능성을 제시하고, 디지털 놀이치료가 ADHD 관리에 새로운 해결책으로서의 역할을 확대할 수 있는 방법을 탐구할 것입니다.

## 참 고 문 헌

- [1] 안은영, “뇌활동 증진을 위한 손가락 운동용 기 능성 게임 장치 및 콘텐츠 개발,” \*멀티미디어 학회논문지\*, vol. 15, no. 11, pp. 1384-1390, Nov. 2012.
- [2] 최인욱, “기억술을 이용한 두뇌 훈련 게임 연



- 구,” \*산업디자인학연구\*, vol. 12, no. 1, pp. 149-158, Mar. 2018 .doi: 10.37254/ids.2018.03.43.15.149.
- [3] 이창조, 김미혜, “뇌졸중 환자의 언어/인지 재활 치료를 위한 기능성 게임,” \*한국엔터테인먼트 산업학회논문지\*, vol. 5, no. 1, pp. 73-80, Mar. 2011. doi: 10.21184/jkeia.2011.03.5.1.73.
- [4] 김정은, “의료건강 기능성 게임의 현황과 가능성에 대한 조망,” 2010.
- [5] 최은선, 최다빈, 윤나라, 유다연, 박지수, “ADHD 치료를 위한 아이트래킹 게임 기반 집중력과 기억력 향상 연구,” \*한국정보처리학회 학술대회논문집\*, vol. 30, no. 2, pp. 1016-1017, 2023.
- [6] 안상균, “뉴로피드백 훈련이 초등학생의 주의력과 학습성취동기에 미치는 영향,” \*한국산학기술학회 논문지\*, vol. 12, no. 12, pp. 5525-5530, Dec. 2011.
- [7] D. S. Kim and C. O. Choi, \*Electroencephalogram\*. Seoul: Korean Medical Book Publishing Company, 2001.
- [8] P. J. M. Prins et al., “Does Computerized Working Memory Training with Game Elements Enhance Motivation and Training Efficacy in Children with ADHD?” \*Cyberpsychology, Behavior, and Social Networking\*, vol. 14, no. 3, pp. 115-122, Mar. 2011.

# 데이터베이스 내 프라이버시 보호 기법: 미션 어플리케이션 내 차분 프라이버시 적용 방안 연구

조병욱, 김정희, 이준민, 김성훈, 박지선, 박세진

## Privacy Protection Techniques in the Database: A Study on the Application of Differential Privacy in Mission Applications

Byeong-uk Jo, Jeong-hui Kim, Jun-min Lee, Seong-hoon Kim, Ji-Seon Park, Se-Jin Park

### 요 약

DB에서는 프라이버시 보호를 위해 여러 가지 프라이버시 기법이 적용되는데 보통은 접근을 제어하거나 암호화하는 방식 등이 있다.

이들 중 암호화에는 차분 프라이버시(Differential Privacy, DP)라는 기법이 존재하는데, 이 기법은 DB에 노이즈를 추가해 탈취자가 DB의 내용이 어떤 사용자의 정보인지 추정하기 힘들게 하는 기법으로 본 논문에서는 이 기법들의 장단점을 비교, 분석하여 각 DP모델이 미션 어플리케이션 내에서 각 기능에 적합한지를 알아보고자 한다.

### Abstract

In DB, various privacy techniques are applied for privacy protection, and there are usually methods of controlling or encrypting access.

Among these methods, there is a technique called differential privacy (DP), especially in encryption, which adds noise when storing DBs to make it difficult for attackers to identify a specific individual with DB data. Accordingly, by comparing and analyzing the advantages and disadvantages of differential privacy, we want to find out which DP model is advantageous to apply within the mission application.

### Key words

differential privacy, noise, encryption, case

## I. 서 론

현대사회에서는 개인정보 유출이라는 문제에 대

해서 늘 고민해왔다. 더불어 지금 사회는 정보화  
시대이기에 많은 정보를 담기위해 각각의 기업, 정  
부, 기관, 개인이 데이터베이스(DB)에 각자의 업무

에 맞는 정보를 저장해 왔고 이중에는 개인의 정보가 저장되어 있으며 이런 정보를 보호하기 위한 문제는 이 시대의 핵심과제 중 하나이다.

자취자를 위한 미션인증 어플리케이션에서는 그에 필요한 능력들을 향상시키기 위해 사용자가 스스로 미션을 생성하기도 하지만 사용자끼리 서로에게 미션을 할당하고, 인증하는 시스템을 지원한다. 이런 시스템 속에는 개인의 생활모습과 같은 정보가 포함할 수밖에 없고 이에 대한 정보의 유출문제를 해결하기 위해서는 강력한 프라이버시 보호기법이 요구된다.

DB의 정보보호를 위해 여러가지 보호 기법들이 사용되고 있는데, 대부분 접근 제어를 하거나 암호화를 통해 보호한다. 본 논문에서는 암호화 기법의 한 종류로 DB 저장시에 노이즈를 추가하여 보호하는 차분 프라이버시를 비교, 분석을 통한 장단점을 기반으로 미션 어플리케이션에 대한 적용 방안을 제시하고자 한다.

## II. 차분 프라이버시

차분 프라이버시(Differential Privacy, DP)는 2006년 Dwork[3]에 의해 처음 제안된 개념이다. 이 기법은 원본 데이터를 중앙 서버에 저장한 후, 해당 데이터에 의도적으로 노이즈를 추가하여 개인정보를 보호한다. 노이즈를 생성하는 방법으로는 라플라스(Laplace), 가우스(Gaussian), 지수(Exponential) 메커니즘[4] 등이 사용된다. DP는 개별 데이터의 식별을 어렵게 하며, 데이터베이스 내 특정 항목의 추가나 변경이 전체 통계 결과에 미치는 영향을 최소화하는 데 목적이 있다. 그러나 DP는 원본 데이터를 서버에 저장하는 방식이기 때문에, 이후에 논의할 지역 차분 프라이버시(Local Differential Privacy, LDP)나 분산 차분 프라이버시(Distributed Differential Privacy, DDP)와 비교했을 때 데이터의 활용도는 높지만 보안 측면에서는 다소 취약할 수 있다.

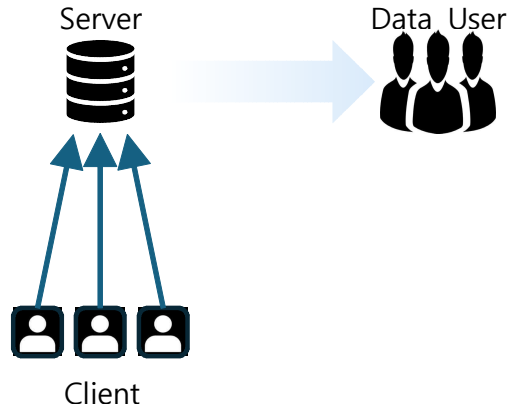


그림 1. DP 방식 접근  
Fig. 1. DP based approach

2. 지역 차분 프라이버시(Local Differential Privacy, LDP) : LDP는 개인이 노이즈를 추가해 데이터를 서버에 보낸다.[2] 이로 인해서 중앙 서버에 민감한 정보가 그대로 노출되지 않아 보안성을 강화한다. 또 중앙 서버를 신뢰할 필요가 없어진다. 하지만 노이즈를 개별로 추가하기 때문에 전체적인 데이터의 정확성이 떨어진다.

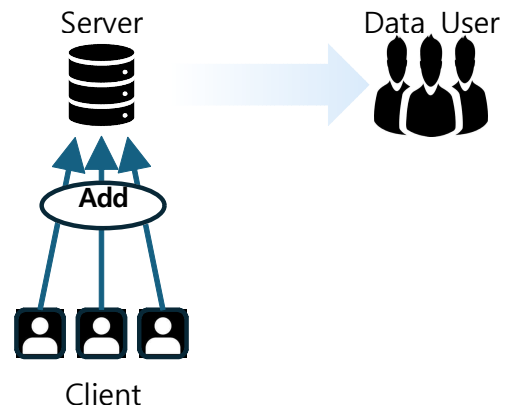


그림 2. LDP 방식 접근  
Fig. 2. LDP based approach

3. 분산 차분 프라이버시(Distributed Differential Privacy, DDP) : 분산 차분 프라이버시는 LDP와 같이 개인이 데이터에 노이즈를 추가하는 방식은 같지만 노이즈의 양을 조절해 데이터의 정확성을 높인다. 이때 낮아진 보안성을 보완하기 위해 데이터들을 암호화한다.[2]

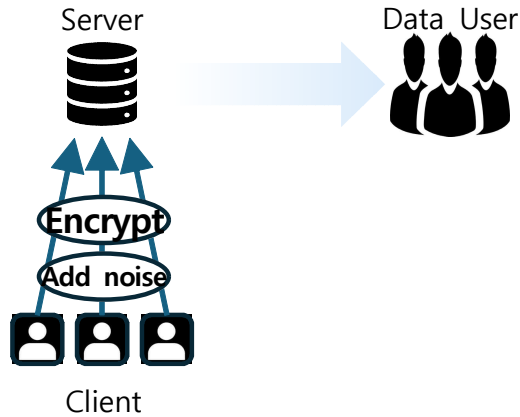


그림 3. DDP 방식 접근  
Fig. 3. DDP based approach

#### 4. 미션 인증 어플리케이션 적용 방안

본 논문에서 제시한 모델을 실제 어플리케이션에 적용하기 위한 예시로 1인 구성 세대의 생활 습관 형성을 도와주는 미션 어플리케이션을 들 수 있다. 이 어플리케이션은 사용자 데이터를 보호하면서도 유용한 기능을 제공하기 위해 서버와 클라이언트 측에 차등 프라이버시 모델을 결합한 방식을 활용할 수 있다.

서버 측에서는 미션 성공률, 참여 빈도와 같은 집계 데이터를 통해 사용자들에게 통계적으로 유용한 정보를 제공하는 동시에, 개별 사용자의 개인정보를 보호할 수 있는 차등 프라이버시(DP) 기법을 적용한다. 예를 들어, 전체 사용자의 보유 미션 목록이나 미션 일정과 같은 데이터를 보호하기 위해 서버에 노이즈를 추가해 분석을 수행하며, 이를 통해 사용자가 개인적으로 수행한 특정 미션의 데이터가 노출되지 않도록 한다.

클라이언트에서는 어플리케이션 사용자의 미션 인증 사진이나 미션 피드백 데이터를 서버로 전송하기 이전에 로컬 차분 프라이버시(Local Differential Privacy, LDP)를 이용해 데이터를 보호하게 된다. 이는 어플리케이션 사용자의 디바이스에서 먼저 노이즈를 추가하는 방식이며, 서버로 전송되는 과정에서 데이터가 사용자 개인의 정확한 선호 정보를 포함하지 않도록 하여 데이터 수집 과정에서의 프라이버시를 더욱 강화할 수 있는데 특히, 어플리케이션에서 제공하는 위치 기반 미션의 경우 어플리케이션 사용자가 특정 장소에서 미션을 완료

했을 때, 위치 정보에 분산 차분 프라이버시(Distributed Differential Privacy, DDP)를 적용하여 암호화된 상태의 데이터를 서버에 전송하게 된다. 위의 처리과정을 거친 위치 정보는 서버에서 집계 분석에 활용되며 위치 기반 추천 기능에 사용될 수 있지만, 어플리케이션 사용자의 정확한 위치 정보는 드러나지 않는다. 이처럼 본 어플리케이션은 여러 차분 프라이버시 모델들을 활용하여 어플리케이션 사용자의 프라이버시를 보호함과 동시에 데이터의 유용성을 최대한 보존하며, 이를 통해 어플리케이션 사용자에게 유의미한 통계 데이터와 그에 기반한 미션 추천 기능을 제공할 수 있다.

### III. 결 론

본 연구에서 제안한 차분 프라이버시 기법들의 미션 인증 어플리케이션 적용 방안은 이론적 분석을 바탕으로 한 것으로, 실제 구현 및 사용자 테스트를 통한 검증이 필요하다. 그러나 각 기법의 특성을 고려할 때, 다음과 같은 잠재적 이점을 예상할 수 있다.

데이터 보호 측면에서 DP, LDP, DDP의 적용으로 사용자들에게 더 높은 수준의 데이터 보안을 제공함으로써, 앱에 대한 신뢰도를 높일 수 있을 것으로 기대된다. 또한 차분 프라이버시 기법들을 적절히 조정하여 적용함으로써, 데이터의 유용성을 크게 손상시키지 않으면서도 프라이버시를 보호할 수 있을 것으로 예상된다. 마지막으로, 개인정보 보호에 관한 각종 규제와 법규를 준수하는 데 도움이 될 것으로 보인다.

그러나 이러한 이점들을 실제로 검증하고 최적화하기 위해서는 다음과 같은 후속 연구 및 실험이 필요하다.

먼저, DP 모델들을 적용한 실제 앱을 개발해야 할 것이다. 그 후 사용자 그룹 대상으로 테스트를 진행해 사용성과 효과를 평가한다. 또, DP 모델 적용 전후의 데이터 유용성과 서비스 품질을 비교·분석하고 여기서 최적의 노이즈 설정을 도출해야 한다. 마지막으로, 이러한 방식의 잠재적 보안 취약점을 식별하고 이런 취약점에 대한 대응 방안을 제

시해야 한다.

이러한 후속 연구를 통해 본 논문에서 제안한 DP 기반 미션 인증 어플리케이션 모델을 더욱 실용적으로 활용할 수 있을 것으로 기대된다. 이는 결과적으로 사용자의 프라이버시를 보호하면서 미션 관리와 습관 형성을 효과적으로 지원하는 서비스 개발에 기여할 수 있다.

[7] H. C. Chang, H. Kim, H. R. Park, K. S. Chun and H. H. Lee, "Gaeinjeongbo DB Gwalligisurui Boan Yogusahang Yeongu(A Study on the Security Requirements of Personal Information DB Management Technology)", KIISC, vol.18, no.2, pp.76-86, 2008

## 참 고 문 헌

- [1] S. H. Yoon, "Study on the Technical Security Factor for the Implementation of Secure DB System", The Korea Society of Computer and Information, vol. 19, no. 12, pp. 143-152, December 2014.
- [2] J. H. Lim and J. W. Kim, "Privacy-Preserving Data Collection and Aggregation of Health Data with Differential Privacy", The Korea Society of Computer and Information, August 2020. <http://smulib.dcollection.net/common/orgView/200000332991>
- [3] C. Dwork, F. Mcsherry, K. Nissim, and A. Smith, "Calibrating Noise to Sensitivity in Private Data Analysis," Proceedings of the 3rd Theory of Cryptography Conference, pp. 265-284. Mar, 2006.
- [4] L. I. Yu and Y. M. Na, "A Study on Database Encryption Techniques Considering Order Preservation and Differential Privacy", Korean Institute of Information Scientists and Engineers, December 2023. <https://lib.dankook.ac.kr/dcollection/common/orgView/000000199648>
- [5] N. I. Kim, I. C. Baek, H. j. Lee, M. S. Kim and Y. D. Chung, "A Differential-Privacy Technique for Publishing Density-based Clustering Results", KIISE, vol. 51, no. 4, pp. 380-388, 2024
- [6] Malekzadeh M, Hasircioglu B, Mital N, Katarya K, Ozfatura ME, Gündüz D. "Dopamine: Differentially Private Federated Learning on Medical Data.", 2021. <http://arxiv.org/abs/2101.11693>

# 사용자 외출 패턴 기반 선제 발화 및 맞춤형 대화 제공 AI 음성 봇 시스템

송지웅\*, 신지영\*, 황종익\*, 주종화\*\*

## AI voice bot system that pre-emptively speaks and personalizes conversations based on user outing patterns

Ji woong Song\*, Ji young Shin\*, Jong ick Hwang\*, and Jong wha J. Joo\*\*

### 요 약

본 연구는 독거노인의 외출 패턴을 반영하여 최적의 대화 시점을 설정하고, RAG(Retrieval Augmented Generation) 구조를 통해 사용자 발화를 기억함으로써 개인화된 대화를 제공하는 AI 음성 봇 시스템을 개발했다. 본 시스템은 LinUCB 알고리즘과 EWMA를 활용하여 독거노인의 외출 패턴에 맞춘 선제 발화 시각을 동적으로 조정하며, 대화 요약 및 관심사 추출 모듈을 통해 사용자의 취향과 특성에 기반한 맞춤형 상호작용을 실현한다.

### Abstract

This research proposes an AI voice bot system that provides personalized conversations by setting the optimal time to talk to the elderly by reflecting their outing patterns and remembering user utterances through a Retrieval Augmented Generation (RAG) structure. The system utilizes LinUCB algorithm and EWMA to dynamically adjust the preemptive speaking time to suit the outing patterns of the elderly, and realizes personalized interaction based on the user's preferences and characteristics through the conversation summary and interest extraction modules.

### Key words

pre-emptively speaking, personalized conversation, interest extraction, conversation summarize, LLM, RAG, LinUCB, EWMA

## 1. 서 론

한국의 독거노인 비율은 2000년 16.0%에서 2024년 22.1%로 점차 증가하고 있으며, 한국은 OECD 국가 중 독거노인 자살률이 높은 국가이다. 이들은

함께 사는 가족이 없기 때문에 외로움이나 우울감을 더 많이 느끼게 된다[1]. 이를 해결하기 위해 AI 기반의 다양한 돌봄 서비스가 개발되어 독거노인의 외로움을 줄이고자 하는 노력이 지속되고 있다. 기존의 AI 돌봄 시스템은 고정된 시간에 대화를 시도

\* 동국대학교 컴퓨터공학과, {thdwldnd7, sjy140801, paperk}@dgu.ac.kr

\*\* 동국대학교 컴퓨터공학과, jwjoo@dgu.ac.kr (교신저자)

※ 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업(2023-0-00049)과 대학ICT연구센터육성지원사업(IITP-2024-2020-0-01789)의 지원을 받아 수행되었음.

하거나 사용자의 호출에 반응하는 방식으로 작동한다[2][3]. 하지만 이러한 방식은 사용자가 대화를 필요로 하는 순간에 시스템이 반응하지 않거나, 준비되지 않은 상황에서 대화가 시도되는 경우가 있어 자연스러운 상호작용이 어려운 한계가 존재한다. 또한, 사람처럼 느껴지는 대화 모델을 만들기 위해서는 대화 상대가 말한 내용을 기억하는 것은 핵심적인 요소지만[4] 이를 충족하지 못하는 제품들이 많다.

본 연구에서는 사용자의 외출 패턴을 기반으로 최적의 대화 시점을 파악하여, 사용자의 호출 없이도 선제적으로 자연스러운 대화를 제공하는 AI 음성 봇 시스템을 개발했고, RAG(Retrieval Augmented Generation) 구조를 통해 사용자의 이전 발화 내용을 기억하여 더욱 유기적인 대화를 실현하는 시스템을 개발했다.

## II. 사전 지식

### 2.1 RAG

대규모 언어 모델(Large Language Models, LLM)의 발전은 방대한 학습 파라미터와 지식을 기반으로 답변 생성 능력을 크게 향상시켰다[5]. 그러나 LLM은 사실과 다른 결과를 제시하는 환각(hallucination) 현상 등의 문제를 보인다. 이러한 문제를 해결하고자, 검색 증강 생성(Retrieval Augmented Generation, RAG)이 제안되었다.

RAG는 대규모 언어 모델인 LLM의 출력을 최적화하기 위해 사용자의 질의에 대한 답을 생성하는 과정에서 기존에 학습된 자체 지식뿐만 아니라 외부의 신뢰할 수 있는 지식 베이스를 참조하도록 하는 프로세스이다[6]. RAG 구조는 사용자의 질문과 관련된 외부 문서를 검색하고 이를 질의 내용에 추가하여 LLM에 전달함으로써 더욱 정확한 응답을 생성할 수 있도록 한다.

### 2.2 Prompt Engineering

LLM이 발전하면서 프롬프트 엔지니어링이 중요시되고 있다. 프롬프트는 사용자가 제공하는 입력으

로, 프롬프트 입력으로 번역, 소설 작성, 코드 생성 등 LLM이 다양한 작업을 수행하게 할 수 있다[7]. 프롬프트 엔지니어링(Prompt Engineering)이란 LLM의 수행 결과가 사용자가 원하는 결과값에 가까워지도록 프롬프트를 지속해서 수정 및 보완하는 것이다.

## III. 시스템 개발

본 연구에서는 독거노인의 외출 패턴을 반영한 선제 발화와 사용자의 관심사 기반 맞춤형 대화 서비스를 제공하는 AI 음성 봇 시스템을 개발한다. 본 시스템은 선제 발화 모듈, 대화 요약 및 업데이트, 관심사 추출 모듈을 구성 요소로 한다.

### 3.1 System Architecture

먼저, 선제 발화 모듈은 사용자의 외출 패턴을 분석하여 대화의 최적 시점을 설정한다. 설정된 시점에 서버의 대화 시작 API가 호출되면, 사용자 정보 DB에서 사용자의 관심사를 추출하여 RAG를 수행한다. 관심사와 Relevant context를 LLM에 전달하여 대화를 시작할 질문을 생성한다. 생성된 질문은 사용자에게 전달되며, 사용자가 응답을 제공하면 해당 응답에 대해 추가로 RAG를 수행한다. 이후 사용자 응답과 Relevant context가 다시 LLM에 전달되어 후속 질문이 생성되며, 이는 대화가 종료될 때까지 반복된다. 대화가 종료된 후, 대화 요약 및 업데이트와 관심사 추출 과정이 진행된다.

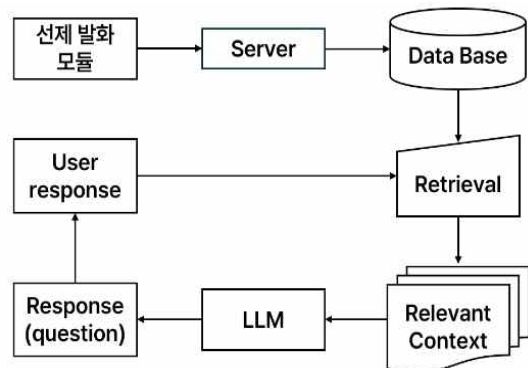


그림 1. System Architecture

### 3.2 선제 발화 모듈

선제 발화 모듈은 매일 유사한 생활 패턴을 보이는 사용자들을 대상으로 사용자가 집에 있고, 평소에 자주 대답했던 시간대에 선제적으로 말을 거는 것을 목표로 한다. 또한 생활 패턴이 바뀌더라도 변화된 패턴에 맞춰 적절하게 발화를 시도하는 것을 목표로 한다.

노인들의 평균적인 취침, 기상 시간을 고려하여 발화 시도 시간대는 07시부터 22시까지로 한정한다 [8]. 전체 시간대를 3시간 간격으로 나누어 총 5개의 section을 구성하며, 각 section 별로 한 번씩 대화를 시도하여 하루에 최소 5회의 발화를 시도한다. 각 section을 15분 간격으로 12개의 slot으로 구분하여 관리한다.

매일 각 slot 별로 사용자가 집에 있는지 여부와 선제 발화에 대한 응답 여부를 기록한다. 집에 있었다면 1, 부재 중이었다면 0으로 기록하고, 선제 발화에 응답했을 경우 1, 응답하지 않았을 경우 0으로 기록한다. 선제 발화 시도가 없었던 slot이라면 하지 않은 것으로 기록한다.

선제 발화 시점은 각 section 별로 LinUCB 알고리즘[9]을 적용하여 결정된 slot 번호를 기준으로 설정한다. LinUCB 알고리즘에서 사용하는 각 slot 별 context 벡터로는 각 slot 별로 최근 30일간 수집한 외출 여부 및 응답 여부 데이터를 EWMA함수[10]를 사용하여 집계한 크기 2인 벡터를 활용한다.

매일 00시 00분에 해당일의 선제 발화 시점이 결정되며, 결정된 시점에 사용자 위치를 확인하여 위치가 집일 경우 선제 발화를 진행한다. 만약 사용자가 집이 아니라면 같은 section 내 다른 slot에 대해서 LinUCB 알고리즘을 재적용하여 slot을 재선정하고, 해당하는 시점에 다시 선제 발화를 시도한다. 이 경우에도 사용자의 위치가 집이 아니라면 더 이상 해당 시간대에 추가적인 시도를 하지 않는다.

선제 발화 시점에 사용자 위치가 집이 아니라면 reward를 -1로 설정한다. 사용자의 위치가 집이고, 사용자가 응답한 경우 reward를 1로 설정한다. 사용자의 위치가 집인데도 응답하지 않은 경우 reward를 -1로 설정한다. 설정된 reward를 기반으로 LinUCB

알고리즘이 해당 slot의 가중치를 갱신하여, 선제 발화 시각의 최적화가 가능하게 한다.

### 3.3 대화 요약 및 업데이트

대화 요약 및 업데이트 모듈은 사용자에게 더욱 개인화된 맞춤 대화를 제공하기 위하여 해당 사용자와 과거에 대화했던 내용을 기억하는 대화 모델을 제공하는 것을 목표로 한다. 이를 위해 대화가 종료될 때마다 해당 대화의 요약문을 생성하고, 이를 관리하기 위한 시스템을 개발했다.

최근 연구[11]에 따르면, LLM이 zero-shot 기준의 요약 생성 작업에서도 기존 요약 모델이나 인간이 작성한 요약보다 우수한 성능을 보인다. 이러한 성능을 바탕으로 본 연구에서는 대화 요약문 생성 및 업데이트에 LLM과 RAG를 활용하였다.

사용자와의 대화 종료 시에, 대화 전체 내용을 프롬프트와 함께 LLM에 전달하여 해당 대화의 요약문을 생성하고, 이를 사용자 개별 DB에 저장된 요약문들과 비교하여 갱신 또는 삭제가 필요한 내용을 식별한다. 이를 위해 RAG의 Ensemble Retriever를 활용하여 기존 요약문 중 새로운 요약문과 유사한 것을 탐색하고, 유지, 갱신, 삭제 등의 작업을 한다.

이와 같은 요약문 생성 및 업데이트 과정을 통해, 본 시스템은 사용자와의 대화 시 이전 대화 내용과 사용자의 특성을 기억하는 대화의 진행이 가능해진다. 이를 통해 사용자는 일반적인 LLM과의 대화보다 더욱 개인화된 대화 경험을 하게 된다.

### 3.4 관심사 추출

관심사 추출 모듈은 사용자의 취향과 흥미를 바탕으로 맞춤형 대화를 제공하는 것을 목표로 한다. 본 연구에서 관심사 추출은 LangChain 모듈과 프롬프트 엔지니어링을 조합하여 다음과 같은 과정을 거친다.

대화 기록 분석 및 관심사 추출 단계에서 사용자의 대화 기록을 기반으로 관심사로 판단할 수 있는 주요 단어를 추출한다. 예를 들어, 사용자가 “취미



는 독서예요"와 같은 문장을 언급할 경우, '독서'라는 단어가 사용자의 관심사로 추출된다. 이러한 단어는 대화 내 단어 빈도와 중요도를 기반으로 추출되어 초기 관심사로 설정된다.

키워드 선정 및 정제 단계에서 추출된 단어들은 의미적 중복성 제거를 위해 필터링 과정을 거친다. 예를 들어 '독서', '읽기' 등의 단어가 추출된 경우, 이를 '독서'로 대표하여 정제한다.

관심사 업데이트 및 저장 단계에서 최종적으로 선정된 관심사는 사용자 정보 DB에 저장되며, 향후 대화에 반영될 수 있도록 업데이트된다.

관심사 추출 모듈은 사용자 맞춤형 대화를 가능하게 하며, 대화의 자연스러움과 연속성을 유지함으로써 사용자와의 정서적 유대감을 강화한다.

#### IV. 실험 결과

실제로 선제 발화 모듈에 적용된 알고리즘이 변화된 외출 패턴에 맞추어 발화 타이밍을 적절하게 업데이트하는지 검증하기 위해, 두 가지 외출 패턴(A, B)에 대해 테스트를 진행하였다. 테스트 시작 후 1일 45일까지는 A 패턴으로, 45일 90일까지는 B 패턴으로 생활하는 사용자를 가정하였다. 왼쪽 그림은 사용자 위치가 집일 확률을, 오른쪽 그림은 사용자가 대답할 확률을 나타낸다.

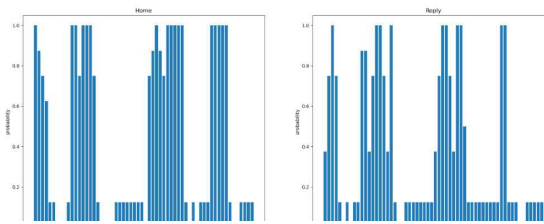


그림 2. 외출 패턴 A

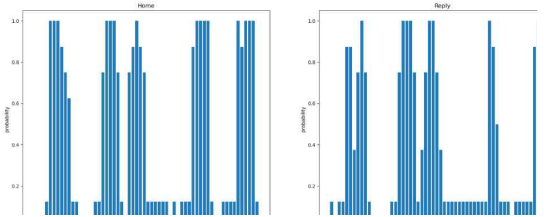


그림 3. 외출 패턴 B

테스트는 EWMA의 alpha 값을 0.8, LinUCB의 alpha 값을 0.1로 설정하여 총 1일부터 90일까지를 대상으로 100회 반복 수행하였다. 그 결과, 성공률은 다음과 같다.

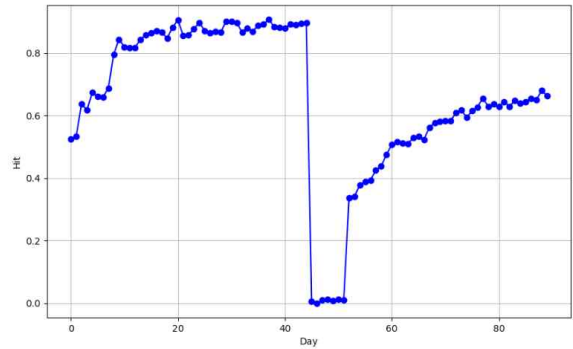


그림 4. 실험 결과

그림 4에서 보듯이, 45일 근방에서 성공률이 일시적으로 감소하였으나, 시간이 지남에 따라 점차 회복되는 양상을 확인할 수 있다.

#### V. 결론

본 연구에서는 사용자 외출 패턴 기반 선제 발화 및 관심사 맞춤형 대화 제공 AI 음성 봇 시스템을 개발하여 독거노인의 정서적 고립을 완화하고자 하였다. 본 시스템은 독거노인의 외출 패턴을 분석하여 최적의 대화 시점을 파악하며, LinUCB 알고리즘과 EWMA 함수를 통해 개인의 생활 변화를 반영한 대화 시점을 지속적으로 업데이트한다. RAG 구조를 활용하여 사용자의 과거 발화를 기억하고, 맞춤형 대화를 제공함으로써 보다 자연스럽게 연속성 있는 상호작용을 가능하게 하였다.

실험 결과, AI 음성 봇 시스템은 사용자가 외출 패턴을 변경하는 경우에도 일정 시간이 경과함에 따라 성공률이 점차 회복되는 것으로 나타났다. 이는 본 시스템이 사용자와의 유의미한 상호작용을 통해 정서적 지원을 강화할 수 있음을 시사한다. 향후 연구에서는 다양한 고령자의 패턴과 관심사를 더 깊이 반영하여 AI 돌봄 시스템의 상호작용 능력을 지속적으로 개선하고자 한다.

## 참 고 문 헌

- [1] 통계청, 「장래가구추계」, 2052, 2024.10.29, 가  
구주의 연령/가구원수별 추계가구-전국
- [2] AI 돌봄 서비스 로봇 효돌[웹사이트].  
(2024.10.29). URL : <https://hyodol.com/>
- [3] KT AI Care 서비스[웹사이트]. (2024.10.29). URL  
: <https://gigagenie.kt.com/m/business/genieAiCare.do>
- [4] NAVER D2. (2023, March 16). Remember Me:  
chatbot for personalized care [Video]. YouTube.  
<https://www.youtube.com/watch?v=d8g3D3Xf1Gw>
- [5] Brown, T., Mann, B., Ryder, N., Subbiah, M.,  
Kaplan, J. D., Dhariwal, P., and Amodei, D.,  
“Language models are few-shot learners,”  
Advances in Neural Information Processing  
Systems, vol. 33, pp. 1877-1901, 2020.
- [6] P. Lewis, E. Perez, et al. "Retrieval-augmented  
generation for knowledge-intensive NLP tasks."  
arXiv:2005.11401, 2020.
- [7] M. Kang and D. Jeon, “Enhancing the  
Performance of Generative AI-Based Educational  
Material Recommendation Functions: Focusing on  
Query-Based Prompt Engineering,” Digital  
Contents Society, Vol. 25, No. 6, pp. 1601-1609,  
2024.
- [8] 통계청. 2019년 생활시간조사 결과, 2020.
- [9] Lihong Li et al. "A Contextual-Bandit Approach to  
Personalized News Article Recommendation."  
Proceedings of the 19th international conference  
on WWW, 661-670, 2010.
- [10] Hunter, J. S. "The Exponentially Weighted  
Moving Average. Journal of Quality Technology"  
cx18(4), 203-210, 1986
- [11] Xiao Pu, Minqi Gao and Xiaojun Wan.  
"Summarization is (Almost) Dead." arXiv preprint  
arXiv:2309.09558, 2023.

# Raspberry Pi 기반 OTP 및 얼굴 인식 도어락 시스템

이용진\*, 이성욱\*, 오건주\*, 이재열\*, 권명준\*, 안영래\*, 지승한\*\*

## Raspberry Pi-based OTP and Face Recognition Door Lock System

Yongjin Lee\*, Seongwook Lee\*, Geonju Oh\*, Jaeyeol Lee\*, Myungjun Kwon\*, Youngra An\*, and Seunghan JI\*\*

### 요 약

1인 가구의 비율이 점점 늘어가는 현대 사회에서 보안은 점차 더 중요하게 자리 잡고 있다. 기존의 도어락은 지문을 통한 패턴 파악이 가능한 등 여러 문제점을 가지고 있는데, 본 연구에서는 이중 보안 시스템으로 전반적인 기능을 Raspberry Pi라는 기기로 구현하여 OTP(One Time Password) 발급을 통한 비밀번호 노출 방지 및 생체인증 시스템(얼굴인식)의 도입을 통해 보안성 강화와 비밀번호 로그 조회, 침입자 조회 등 서버를 통한 웹페이지 상의 관리로 사용자의 접근성 및 편의성을 높이고자 설계했다.

### Abstract

Security is becoming increasingly important in a modern society where the proportion of single-person households is increasing. Existing door locks have the problem that patterns can be identified through fingerprints. In this study, the overall functions of the dual security system were implemented with a device called Raspberry Pi. It was designed to enhance security by preventing password exposure through OTP issuance and introducing a biometric authentication system, and to improve user accessibility and convenience by managing the web page through the server, such as password log inquiry and intruder inquiry.

### Key words

Raspberry Pi, Security, OTP, Face Recognition

## 1. 서 론

현대 사회에서 보안은 스마트 홈 기술의 급속한 발전과 함께 일상생활의 필수 요소로 자리 잡았다. 특히 도어락과 같은 출입 통제 시스템은 개인 안전과 자산 보호의 중요한 역할을 맡고 있다. 비밀번

호 유출, 복제의 위험성이 있으며, 사용자가 자주 비밀번호를 변경하지 않거나 쉽게 추측할 수 있는 비밀번호를 사용할 경우, 불법적인 접근에 노출될 가능성이 크다[1]. 이러한 문제는 비밀번호 기반 시스템의 한계를 보여주며, 더 강력한 보안 시스템의 필요성을 의미한다. 본 연구에서는 Raspberry Pi를

\* 대림대학교 메카트로닉스공학과 gr0721@naver.com

\*\* 대림대학교 메카트로닉스공학과 ish2012@daelim.ac.kr

활용하여 OTP와 얼굴인식을 결합한 이중 인증 보안 도어락 시스템을 제안한다. 이 시스템은 웹 기반 관리 인터페이스를 통해 접근성을 높이는 동시에 높은 보안성을 유지하도록 설계했다. 이를 통해 사용자는 언제 어디서나 시스템을 손쉽게 관리할 수 있으며, 스마트 홈 환경에서 높은 보안성을 제공한다. 따라서 사용자 편의성과 보안성을 동시에 만족시키는 도어락 시스템의 구현 가능성을 제시하며, 스마트 홈 보안 시스템의 발전에 기여할 수 있는 해결책을 모색하고자 설계 연구했다.

## II. 얼굴인식

본 연구에서는 Raspberry Pi 환경에서 효율적인 얼굴인식을 구현하기 위해 Dlib 라이브러리의 Landmark 68을 사용했다[2]. Dlib의 Landmark 68 알고리즘은 회귀 트리, 모양 모델링, Affine 및 Similarity 변환, 비용 함수와 같은 다양한 기법이 복합적으로 사용되어 얼굴의 주요 포인트를 정밀하게 추출하며, 특히 자원이 제한된 환경에서도 높은 성능을 발휘할 수 있는 장점이 있다[2]. Landmark 68은 그림 1과 같이 얼굴의 윤곽선(1~17), 왼쪽 눈썹(18~22), 오른쪽 눈썹(23~27), 코(28~36), 왼쪽 눈(37~42), 오른쪽 눈(43~48), 외부 입술(49~60), 내부 입술(61~68)과 같이 얼굴에서 68개의 고정된 포인트를 회귀 트리(Regression Tree) 알고리즘, 모양 모델링(Shape Modeling), Affine 및 Similarity 변환, 비용 함수(Cost Function)를 사용하여 얼굴의 랜드마크를 추출한다[3].

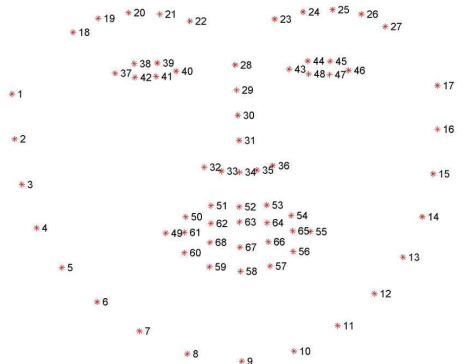


그림 1. Landmark68 얼굴 랜드마크 포인트[3]

Fig 1. Landmark68 facial landmark points[3]

회귀 트리(Regression Tree) 알고리즘은 랜드마크 위치( $L_t$ )와 이미지( $I$ )를 기반으로 회귀 트리 함수( $R_t$ )를 업데이트하여, 다음 단계에서의 새로운 랜드마크 위치( $L_{t+1}$ )를 선정한다. 이는 식 (1)과 같이 표현된다.

$$L_{t+1} = L_t + R_t(I, L_t) \quad (1)$$

또한, 얼굴의 전반적인 형태를 모델링하기 위해 PCA(Principal Component Analysis) 기법을 사용한 모양 모델링(Shape Modeling)을 적용한다. 평균 모양( $L_{mean}$ )과 주성분 벡터( $P$ ), 각 성분의 가중치( $b$ )를 활용하여 얼굴의 전반적인 형태와 랜드마크의 상대적인 위치를 모델링한다. 이는 식 (2)로 나타낼 수 있다.

$$L = L_{mean} + P \cdot b \quad (2)$$

얼굴 정렬(Face Alignment) 과정에서는 Affine 및 Similarity 변환을 통해 원래의 랜드마크 위치를 변환된 위치로 매핑한다. Affine 변환은 얼굴의 회전, 이동, 스케일 변경 등을 처리하여 전체적인 얼굴의 위치와 각도를 맞추는 데 사용되며, Similarity 변환은 얼굴의 비율을 유지하면서 정밀하게 정렬된 이미지를 변환한다. 이러한 변환 과정은 식 (3)과 같이 표현되며, 다양한 촬영 조건에서도 얼굴이 일관되게 정렬될 수 있도록 정규화된다.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} \quad (3)$$

마지막으로, 비용 함수(Cost Function)를 통해 예측된 랜드마크의 위치( $\hat{L}_i$ )와 실제 랜드마크의 위치( $L_i$ ) 사이의 차이를 평균 제곱 오차(MSE)를 기반으로 최소화한다. 식 (4)와 같이 MSE는 예측된 랜드마크와 실제 랜드마크 간의 오차를 간소화하는 데 사용된다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{L}_i - L_i)^2 \quad (4)$$

### III. 도어락 시스템

그림 2는 도어락과 서버의 실시간 상호작용을 나타낸다. 도어락의 작동 방식은 도어락 본체와 웹 페이지가 상호 연동되며, 각 입력된 신호 및 정보를 실시간으로 주고받으며 작동한다.

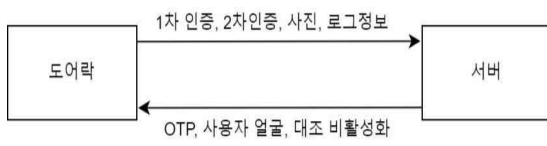


그림 2. 도어락과 서버의 실시간 상호작용  
Fig. 2. Real-time interaction between door lock and server

#### 3.1 웹페이지

웹페이지와 도어락을 서버와 연동하여 데이터를 송수신하는 방식을 활용하였고 각 데이터베이스(Database)의 테이블과 열(column)을 활용하여 OTP 생성, OTP 인증 조회, 사용자 얼굴 등록, 등록된 얼굴 조회, OTP 대조 비활성화 등의 기능을 구현했다. 이러한 기능들로 사용자가 웹페이지에 자신의 계정으로 로그인하여 손쉽게 관리할 수 있도록 했다.

#### 3.2 도어락

도어락은 그림 3의 구성도와 같이 Raspberry Pi를 중심으로 키패드, 카메라, LCD 디스플레이, 잠금장치로 이루어져 있다. 각 센서들은 GPIO핀과 USB 포트를 통해 도어락과 정보를 송수신하며 수신된 정보를 바탕으로 잠금장치의 개폐 여부를 결정한다.

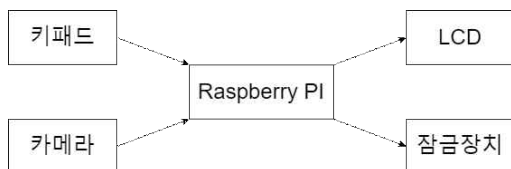


그림 3. 도어락의 구성도  
Fig. 3. Structural diagram of door lock

#### 3.3 도어락 작동 방식

그림 4에 도어락의 작동 과정을 나타내었다. 첫 번째로 OTP 인증 절차가 시작된다. 사용자는 서버로부터 발급된 OTP를 키패드를 통해 입력하고 입력된 번호가 일치하면 얼굴 인식 절차가 이루어진다. 얼굴 인식까지 완료되면 잠금장치가 해제되는 방식으로 OTP 인증 또는 얼굴 인식에 실패할 경우, 시스템은 실패 횟수를 기록하고 정보를 서버에 전송하여 저장한다. OTP 인증이 3회 연속 실패하면 도어락 전방을 촬영하여 서버로 전송하며, 해당 정보를 기반으로 메신저로 경고 알림을 전송한다. 입력을 5회 실패하거나 얼굴 인식에 실패할 경우, 도어락은 자동으로 모든 기능을 잠금 상태로 전환하고 경고 메시지를 전송한다.

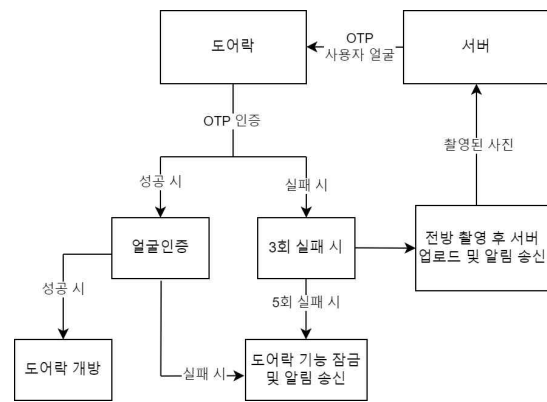


그림 4. 도어락의 작동 방식  
Fig. 4. Workflow of Door lock system

### IV. 실험 및 결과

OTP 인증 방식은 서버에서 0000 - 9999 사이의 난수를 생성해 DB에 저장하고, 지정된 시간 간격이나 사용자가 요청할 때마다 난수를 변경할 수 있도록 설계했다. 3\*4 키패드를 사용하여 사용자가 입력한 비밀번호는 Raspberry Pi를 통해 실시간으로 서버에 전송되고, 서버에 저장된 OTP와 비교한다. 입력된 비밀번호가 서버에 등록된 OTP와 일치하면 그림 5와 같이 'Password Correct'로 일치하지 않는다면 'Password Fail'이 출력된다.



그림 5. 1차 인증  
Fig. 5. First authentication

2차 인증인 얼굴 인식의 성능은 얼굴 인식 성공률과 응답 시간을 측정하여 평가했다. 실험은 카메라를 통해 얼굴을 인식한 후 시스템이 인증을 완료하는 데 걸리는 시간을 응답 시간으로 측정했다. 실험 조건으로는 각 사람의 얼굴을 유사도가 90% 이상일 때 인증이 성공한 것으로 간주하였다. 실험 결과는, 얼굴 인식의 성공률이 평균 93.7%로 나타났으며, 응답 시간은 평균 3.0초 이내였으며, 2.8초에서 3.4초 사이로 측정되었다. 그림 6과 같이, 성공률과 응답 시간에 대한 결과를 그래프로 시각화했으며, 모든 실험에서 성공률이 90% 이상을 기록했다. 이는 저성능 하드웨어 환경에서도 안정적인 얼굴 인식 성능을 제공함을 보여준다.

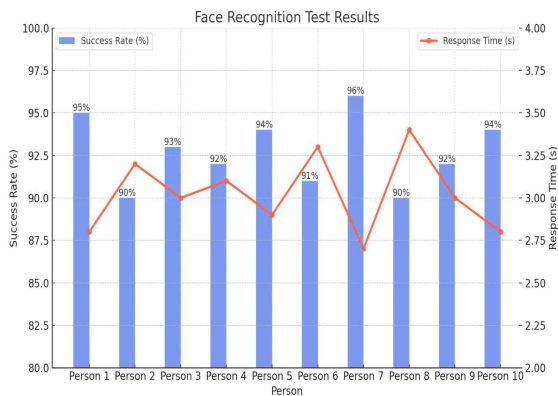


그림 6. 얼굴 인식 실험 결과  
Fig. 6. Face recognition experiment results

그림 7에서 사전에 등록된 사용자 얼굴과 실시간으로 카메라로 들어온 얼굴 이미지를 대조하여 유사도가 지정된 기준치 이상일 경우 인증이 완료되는 방식을 사용했다. 본 실험에서는 얼굴 인증이 성공하면 'Successful', 실패 시에는 'Failed' 라

는 문구를 출력해 실험을 진행하였으며, 실사용을 위해 카메라가 켜진 후 5초 이내에 유사도가 90% 이상일 때 인증이 완료되도록 설정했다.



그림 7. 2차 인증 (데이터, 성공 실패)  
Fig. 7. Second authentication (Data, success, fail)

Landmark68를 활용하여 복잡한 학습 과정 없이 도어락의 카메라로 한 번의 사진 촬영만으로 사용자를 즉시 등록할 수 있는 실시간 사용자 등록 및 해지 기능을 구현했다. 이를 통해 한 번의 사진 촬영만으로 손쉽게 사용자를 등록하고 변경할 수 있으며, 도어락에서 사용자를 초기화하면 서버와 웹페이지에서 기존 사용자의 사진이 자동으로 그림 8의 사용자 관리 인터페이스 기본 이미지로 대체 되어 얼굴 정보 유출을 방지하도록 개발했다.



그림 8. 사용자 관리 인터페이스  
Fig. 8. User management interface

비밀번호 입력이 3회 연속 실패할 경우, 시스템은 도어락 전방을 카메라로 촬영하여 해당 사진을 서버로 전송한다. 촬영된 사진은 FTP(File Transfer Protocol)를 통해 지정된 서버에 업로드되며, 이후 그림 9와 같이 메시지를 통해 사용자에게 경고 메시지가 자동으로 발송된다. 5회 실패하거나 얼굴 인증에 실패할 경우, 시스템은 즉각 사용자에게 경고 알림을 전송하고, 도어락은 웹에서 사용자가 잠금

해제 기능을 수행하기 전까지 잠금 상태로 전환된다. 이 상태에서는 번호를 올바르게 입력하더라도 얼굴 인식 절차는 진행되지 않으며, 도어락은 계속해서 잠금 상태를 유지하여 보안성을 강화한다. 이러한 절차를 통해 무단 접근 시도를 효과적으로 차단하도록 설계했다.



그림 9. 인증 실패별 메시지 알림  
Fig. 9. Messenger notification by authentication failure

그림 10의 작동 내역 기록에서처럼 시스템은 사용자의 모든 시도를 기록하고, 이를 DB에 저장하여 이후 해당 사진과 시간정보를 포함한 작동 내역을 웹페이지에서 확인할 수 있도록 했다.

### ≡ 작동 내역

11.03 04:21:18 비밀번호가 비활성화 되었습니다.  
 11.03 04:21:17 비밀번호가 틀렸습니다  
 11.03 04:21:11 비밀번호가 틀렸습니다  
 11.03 04:21:01 사진을 촬영했습니다.  
 11.03 04:21:00 비밀번호가 틀렸습니다  
 11.03 04:20:54 비밀번호가 틀렸습니다  
 11.03 04:20:48 비밀번호가 틀렸습니다  
 11.03 04:20:33 인증되었습니다.  
 11.03 04:20:17 비밀번호가 틀렸습니다  
 11.03 04:19:43 인증되었습니다.

그림 10. 작동 내역 기록  
Fig. 10. Operation history record

## V. 결과 및 결론

본 연구에서 제안한 시스템은 단순한 접근 제어 장치 이상의 역할을 하며, 실시간으로 사용자와 상호작용하면서 발생할 수 있는 보안 위협에 빠르고 적절하게 대응할 수 있었다. 특히, 얼굴 인식 속도가 평균 3초 이내에 완료되는 신속한 처리 능력과 Raspberry Pi에서도 안정적인 성능을 보임으로써, 다양한 응용 환경에서도 시스템이 적용 가능함을 확인했다. 그러나, 시스템이 사용하는 하드웨어 자원의 제한성으로 인해 대규모 사용자 데이터의 관리에 제약이 있을 수 있다. 이를 극복하기 위해, 향후 연구에서는 시스템의 자원 효율성을 높이고, 확장성 있는 아키텍처를 개발하는 것이 필요하다. 또한, 보안 강화를 위해 추가적인 암호화 및 다중 인증 기법을 도입함으로써, 시스템의 신뢰성과 안전성을 더욱 향상시키는 것이 중요하다.

## 참 고 문 헌

- [1] G. Vardakis, G. Hatzivasilis, E. Koutsaki, and N. Papadakis, "Review of Smart-Home Security Using the Internet of Things," *Electronics*, vol. 13, no. 3343, pp. 1-35, Aug. 2024.
- [2] K. Khabarlak and L. Koriashkina, "Fast Facial Landmark Detection and Applications: A Survey," *arXiv*, vol. 2101.10808v2, pp. 1-18, Apr. 2021.
- [3] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, M. Pantic. 300 Faces in-the-Wild Challenge: The first facial landmark localization Challenge. *Proceedings of IEEE Int'l Conf. on Computer Vision (ICCV-W)*, 300 Faces in-the-Wild Challenge (300-W). Sydney, Australia, December 2013.

# 고독사 방지를 위한 얼굴 인식 기반 출입기록 시스템 설계 및 성능 최적화

김서진\*, 동현진\*, 박주용\*, 배진성\*, 이창민\*, 고석주\*\*, 강성범\*\*\*, 손언호\*\*\*

## Design and Optimization of a Face Recognition-Based Entry Logging System to Prevent Solitary Deaths

Seo-Jin Kim\*, Hyun-Jin Dong\*, Ju-Yong Park\*, Jin-Seong Bae\*, Chang-Min Lee\*,  
Seok-ju Koh\*\*, Seong-Beom Kang\*\*\*, and Eon-Ho Son\*\*\*

### 요약

고령화 사회에서 고독사를 방지하기 위해, 본 연구는 실시간 얼굴 인식 기반 출입기록 시스템을 제안한다. 이 시스템은 Raspberry Pi와 DeepFace 모듈을 사용해 얼굴 임베딩을 생성하고 여러 얼굴 인식 모델과 유사도 메트릭을 비교하여 최적의 조합을 찾아내는 것을 목표로 한다. 연구 결과, 다양한 모델 및 메트릭 조합의 성능을 평가하여 실제 환경에서 가장 효과적인 조합을 도출했으며, 이를 통해 고독사 방지 시스템의 실용적인 가능성을 제시한다.

### Abstract

To prevent loneliness in an aging society, this research proposes a real-time face recognition-based access control system, which uses a Raspberry Pi and DeepFace module to generate face embeddings and aims to find the optimal combination by comparing different face recognition models and similarity metrics. The results of the study evaluate the performance of different combinations of models and metrics to find the most effective combination in real-world environments, which suggests the practical feasibility of such a system.

### Key words

face recognition, elderly care, solitary death prevention, access logging system, deepface

### 1. 서론

현대 사회의 고령화와 1인 가구 증가로 고독사 문제가 심각한 사회적 이슈로 떠오르고 있으며, 이에 대한 효과적인 예방 방안이 요구되고 있다. 기존

의 예방 방법으로는 사회복지사의 정기 방문이나 건강 센서 기반 모니터링이 있지만, 높은 운영 비용과 상시 모니터링의 어려움으로 인해 한계가 있다. 이 문제를 해결하고자 본 연구에서는 비접촉 얼굴 인식 기반 출입기록 시스템을 제안한다. 이 시스

\* 경북대학교 컴퓨터학부, new371@knu.ac.kr, bronzing3412@naver.com, claudeopk@gmail.com · bae1776@gmail.com, rider1106@knu.ac.kr

\*\* 경북대학교 컴퓨터학부, 지도교수, sjkoh@knu.ac.kr

\*\*\* (주)범일정보 기술연구소, ksb0806@bumil.co.kr · kalshade@bumil.co.kr



템은 사용자의 출입 기록을 실시간으로 수집해 비정상적인 생활 패턴을 감지하고, 이상 상황 발생 시 경고를 전송함으로써 고독사를 예방할 수 있도록 설계되었다. 또한 얼굴 인식을 통해 지속적인 인력 투입 없이도 운영 효율성을 높일 수 있다.

본 연구의 목적은 다양한 얼굴 인식 모델과 유사도 메트릭을 비교해 최적의 조합을 도출하고, 고령화 사회에서 고독사 예방에 실질적으로 기여할 수 있는 시스템 개발을 목표로 한다.

## II. 관련 연구

얼굴 인식 모델은 현재 기술적으로는 딥러닝 CNN(Convolutional Neural Network) 구조를 응용한 형태가 대부분이다. 각 모델은 모델마다 고유한 손실 함수와 네트워크 구조 등을 포함하고 있으며 모델 사용 환경(비디오 여부, 밝기 변화 여부, 얼굴 촬영 각도 변화 정도 등)에 따라 얼굴 인식능력에 있어 차별성을 가진다. 본 연구에서 실험 및 비교한 얼굴 인식 모델들 대부분이 딥러닝 CNN 구조를 응용한 것이다.

본 연구에서 활용한 LFW 데이터셋은 얼굴인식 기술 검증에 위한 대표적인 데이터셋 중 하나로 2007년에 발표되었으며 웹에서 수집된 유명한 5,467명에 대한 13,233장의 이미지로 구성되어 있다. 이 데이터셋은 동일 인물에 대하여 자연스러운 환경에서의 밝기, 표정, 얼굴 각도 등이 다양하게 포함되어 있다는 특징이 있다.

## III. 방법론

본 연구에서는 LFW(Labeled Faces in the Wild) 데이터셋에서 1,000개의 이미지 쌍을 사용하여 각기 다른 차원의 벡터를 사용하는 얼굴 인식 모델(Facenet512, Dlib, VGG-Face, ArcFace 등)과 메트릭(cosine, euclidean, euclidean\_l2)에 대해 CPU 환경에서 성능을 평가하였고, 모델별로 정확도, 정밀도, 재현율, 그리고 F1 점수를 측정하여 성능을 비교하였다.

정밀도는 얼굴 인식 모델이 같은 인물로 예측한 경우 중 실제로 같은 인물인 비율을 의미하며 수식은

다음과 같다.

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

정밀도를 측정하는 데 사용된 메트릭은 두 벡터 간의 거리를 측정하는 함수이며 본 연구에서 사용된 메트릭은 다음과 같다.

(가) 코사인 유사도(Cosine Similarity)

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$

두 벡터 A, B에 대해 코사인 각도를 계산한 것.

(나) 유클리드 유사도(Euclidean Distance)

$$\sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

다차원 공간에서 계산한 두 개의 점  $p=(p_1, p_2, \dots, p_n)$ 과  $q=(q_1, q_2, \dots, q_n)$  사이의 거리.

(다) 유클리드 L2 거리(Euclidean L2 Distance)

$$d_{l2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

각 벡터를 L2 norm으로 정규화한 후 두 벡터 사이의 거리를 측정할 것.

## IV. 실험 및 결과

Facenet512 모델은 Euclidean L2 알고리즘을 사용했을 때 가장 높은 정확도와 F1 점수(94.8%, 94.6%)를 기록했지만, 계산 시간은 상대적으로 길어 실시간 처리에 다소 제한이 있을 수 있다. 반면 Dlib 모델은 Euclidean 메트릭에서 성능은 89.4%로 다소 낮지만 Facenet512 모델보다 약 6배 빠른 처리 속도로 실시간 응용에 적합한 가능성을 보였다. VGG-Face

와 ArcFace 모델은 높은 성능을 보였으나, Facenet512 모델과 마찬가지로 계산 시간이 긴 경향이 있었다. 아래 표는 모델별 성능이 가장 높은 메트릭을 요약한 것이다.

Table 1. Experimental data sets

Model	Best Metric	Accuracy	F1 Score
Facenet512	Euclidean L2	94.8	94.6
Facenet	Cosine	86.5	84.4
Dlib	Euclidean	89.4	89.5
VGG-Face	Euclidean L2	92.6	92.1
ArcFace	Cosine	93.2	92.8
GhostFaceNet	Euclidean	88.4	87.1
SFace	Cosine	82.2	78.4
OpenFace	Euclidean	55.5	20.1
DeepID	Euclidean	51.8	7.6

본 실험 결과, Facenet512와 VGG-Face 모델은 정확도와 F1 점수에서 우수한 성능을 보였으나, 비교적 긴 계산 시간이 필요함을 확인할 수 있었다. Dlib 모델은 다소 낮은 성능을 나타냈지만, 약 6배 빠른 처리 속도를 보여 실시간 처리 효율성 측면에서 유리할 수 있음을 시사한다. 이 결과는 응용 환경에 따라 적절한 모델을 선택하는 것이 중요함을 보여준다. 특히, Flask 서버에서 GPU를 사용하여 모델을 처리함으로써 모델의 성능과 처리 속도를 더욱 향상시킬 수 있을 것으로 기대된다.

## V. 결 론

본 연구에서는 고령화 사회에서 고독사를 예방할 수 있는 실시간 얼굴 인식 기반 출입기록 시스템을 개발하고, 그 기술적 가능성과 실용성을 평가했다. 실험 결과, Facenet512 모델과 Euclidean L2 메트릭 조합이 정확도 94.8%, 정밀도 98.48%, 재현율 91%로 가장 우수한 성능을 보였으며, GPU 환경에서 운영할 경우 실제 응용 환경에서 높은 정확도로 출입 기록을 관리할 수 있을 것으로 기대된다. 실시간 성능이 저하될 경우 Dlib 모델을 대안으로 사용할 수 있으며, 정확성은 다소 낮지만 상대적으로 빠른 속도를 제공하여 효율성이 필요한 응용 환경에서 유리할 수 있다.

본 연구는 실시간 출입 관리와 같은 실제 환경에서 고독사 예방에 기여할 수 있는 실질적 대안을 제시하며, 추가적인 실험과 다양한 환경에서의 최적화를 통해 시스템의 성능을 더욱 강화할 수 있을 것이다.

## Acknowledgements

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 통해 수행되었음(2021-0-01082)

이 연구는 과학기술정보통신부 및 정보통신산업진흥원의 “지역산업 SW인재양성 기반조성 사업” 지원으로 수행되었음.

## 참 고 문 헌

- [1] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 1701-1708, 2014.
- [2] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 815-823, 2015.
- [3] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," Proc. Brit. Mach. Vis. Conf. (BMVC), pp. 1-12, 2015.

# 블록체인 기반 분산형 응급의료 지원 시스템 설계

최성\*, 이익호\*, 진기찬\*, 장대규\*, 공태민\*, 안진호\*\*

## Design of Blockchain-Based Decentralized Emergency Medical Support System

Seong Choi\*, Ikho Lee\*, Gichan Jin\*, Daegyung Jang\*, Taemin Kong\*, and Jinho Ahn\*\*

### 요약

본 연구는 응급의료 서비스의 신속성과 효율성을 높이기 위한 블록체인 기반 분산형 응급의료 시스템인 DESK(Distributed Emergency System of Korea)를 제안한다. 기존의 중앙 집중식 구조와 수동식 배정 시스템으로 인한 비효율적인 통신과 의료 자원의 비효율적 배분 문제를 해결하기 위해, 블록체인을 활용한 스마트 컨트랙트와 실시간 데이터 동기화를 도입하여 의료 자원의 신속하고 최적화된 분배와 안전한 데이터 관리 방법을 제시한다. 본 시스템은 환자의 위치와 상태를 기반으로 최적의 병원을 자동 배정하며, 스마트 컨트랙트를 통해 배정 과정의 투명성과 무결성을 보장한다. 또한, 신뢰성 있는 통신으로 병원 간의 정보 교류를 원활히 하여 응급환자에 대한 정확한 진료를 보조한다.

### Abstract

This study proposes a blockchain-based distributed emergency medical system, DESK (Distributed Emergency System of Korea), designed to improve the speed and efficiency of emergency medical services. To address issues such as inefficient communication and resource allocation stemming from a centralized structure and manual assignment systems, DESK introduces smart contracts and real-time data synchronization through blockchain. This enables the swift and optimized distribution of medical resources and secure data management. The system automatically assigns the optimal hospital based on the patient's location and condition, ensuring transparency and integrity in the allocation process via smart contracts. Additionally, reliable communication facilitates information exchange between hospitals, supporting accurate treatment for emergency patients.

### Key words

distributed system, emergency medical service, decentralized network, smart contract, ipfs

---

\* 경기대학교 컴퓨터공학부, {sung000054, dldlrgh28, satba, bigstarchang, taemin2082}@kyonggi.ac.kr

\*\* 경기대학교 AI컴퓨터공학부 교수(교신저자), jhahn@kyonggi.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음 (2021-0-01393).

## 1. 서 론

응급의료 서비스에서 신속한 정보 교환을 통한 골든타임 준수는 매우 중요하다. 그러나 현재 한국의 응급의료 시스템은 자동화되지 않은 중앙 집중형 구조로 인해 병원 간 효율적인 정보 교환과 실시간 응급 대응이 어렵다. 이러한 문제는 응급실 과밀화, 골든타임 확보의 어려움, 의료진의 업무 과중, 응급의료 서비스의 질 저하 등을 초래하고 있다.

반면, 뉴욕의 응급의료 시스템(EMS)은 분산화된 알고리즘과 자동화된 환자 배정 방식을 통해 환자의 상태와 위치를 바탕으로 최적의 병원을 배정함으로써 신속하고 효율적인 응급 대응을 가능하게 하고 있다. 이러한 분산 시스템의 장점을 기반으로 한국형 응급의료체계를 개선할 필요성이 대두되고 있다.

이에 본 연구에서는 블록체인 기술을 활용한 한국형 분산 응급의료 시스템 DESK를 제안하고자 한다. 블록체인 기반의 탈중앙화 네트워크를 통해 병원 간의 신뢰성 있는 정보 공유를 가능하게 하고, 실시간 분산처리를 통한 자동화된 환자 배정을 설계한다. 스마트 컨트랙트(Smart contract)를 사용하여 환자의 상태와 병원 자원 정보를 실시간으로 동기화함으로써, 환자가 잘못된 병원으로 이송되는 일을 방지하고 대규모 재난 상황에서도 살아있는 노드들이 동작하여 신속한 응급 처치를 지원할 수 있도록 한다. 또한, 블록체인의 보안성과 불변성을 통해 데이터의 투명성과 신뢰성을 보장하며, 이를 통해 병원 간의 의료 협력을 촉진하여 신속하고 정확한 응급치료 환경을 조성할 수 있을 것으로 기대된다.

## II. DESK 시스템

### 2.1 전체 시스템 구조

본 시스템은 블록체인 기반 분산 데이터 저장 기술을 사용하여 병원의 병상 정보, 환자의 진료기록, 가용 가능 구급차 등의 정보들을 블록에 저장하여 응급 상황에서 환자가 응급 서비스를 요청하면, 가장 빠르게 도착할 수 있는 구급차를 배정하고 환자

에게 가장 적합한 병원을 자동으로 배정하는 과정을 수행한다.

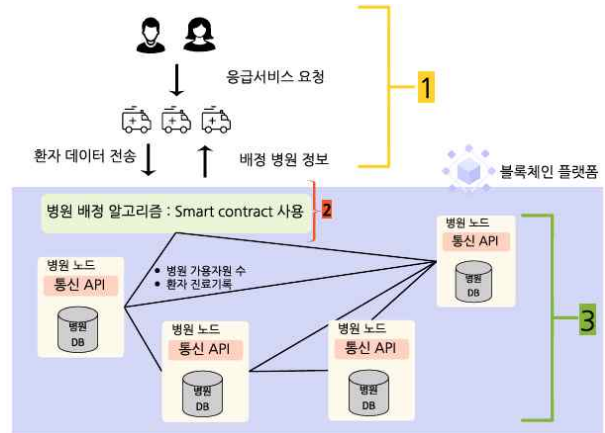


그림 1. 전체 시스템 구조  
Fig 1. Overall System Architecture

응급 서비스를 요청하는 동시에 환자의 정보가 전송되고 참여하고 있는 병원들의 정보를 가지고 블록체인 내에서 스마트 계약을 통해 병원 배정 알고리즘을 수행하게 된다. 이 알고리즘은 환자의 상태와 위치, 병원의 가용 자원 수, 환자의 진료 필요 조건 등을 종합적으로 고려하여 최적의 병원을 선택한다.

각 병원 노드는 통신 API와 분산 파일 저장 구조로 구성되어 네트워크 내에서 상호 연결되어 있으며, 병원의 가용 자원과 배정된 환자의 진료기록 등을 실시간으로 공유하여 빠르게 응급의료 서비스를 제공할 준비를 한다.

### 2.2 병원 배정 알고리즘

병원 배정 알고리즘은 그림 2와 같다. 환자가 응급 의료 시스템을 호출하여 구급차에 탑승하게 되면, 구급대원은 환자의 중증도를 체크한다. 구급대원이 확인한 중증도를 시스템에 설정하면, 중증도가 심각한 순서대로 정렬한다. 만약 중증도가 같게 설정되었다면 병원 배정을 오랫동안 받지 못한 환자가 먼저 정렬된다. 정렬이 완료된 후에는 높은 값에 따라 인덱스를 1, 2, 3 ... 으로 부여한다. 인덱스가 1이 아닌 환자는 시간을 확인하게 된다. 시간을 확인하여 5분 이상 지나있다면 중증도를 올리고 만약 5분이 지나지 않았다면 중증도는 유지한다. 인덱스

가 1인 환자는 병원 배정을 시작하며, 가까운 거리부터 탐색을 시작한다. 환자에게 필요한 진료 과목의 가능 여부를 고려하기 위해 의사 진료 가능 값이 1인 병원이 1개 이상이라면 병상 개수를 확인한다. 의사 진료 가능 값이 1인 병원이 없다면 거리를 늘려 다시 탐색한다. 병상 개수를 확인하여 병상 개수가 하나 이상 존재하는 병원을 먼저 탐색한다. 병원이 존재하지 않는다면, 거리를 증가시켜 탐색하는 과정을 반복한다. 병원 탐색이 완료되고 배정 가능 병상이 있다면 그 병원에 배정하고 알고리즘이 종료된다. 만약 병상 배정이 가능하지 않다면 거리가 같은 값에서 다시 병원 탐색을 진행하게 된다. 탐색 후 병원이 존재하면 해당 병원을 배정하고 존재하지 않는다면 거리를 늘려 다시 탐색한다.

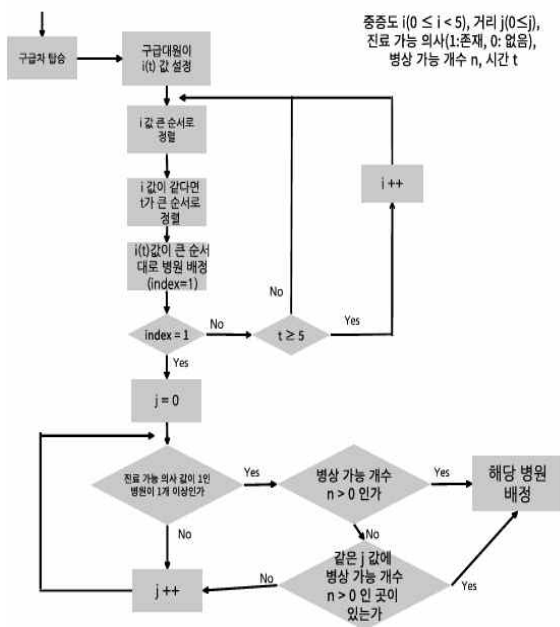


그림 2. 병원 배정 알고리즘  
Fig 2. Hospital Assignment Algorithm

### 2.3 블록체인 시스템

이 논문에서 제안하는 시스템은 비공개형(Private) 블록체인인 Hyperledger Fabric, 병원 배정 스마트 컨트랙트, IPFS(InterPlanetary File System)로 구성된다. Hyperledger Fabric은 허가형 블록체인 네트워크로서, 사용자의 접근권한을 지정할 수 있다. 이를 통해 병원, 구급차 등 상호간의 데이터 접근 권한을 세분화

하여 프라이버시 침해를 방지한다. 또한, Fabric의 채널(Channel)기능을 활용하여 다양한 이해관계자별로 데이터 관리가 가능하여, 응급 상황에서도 민감한 데이터 보호가 가능하다.

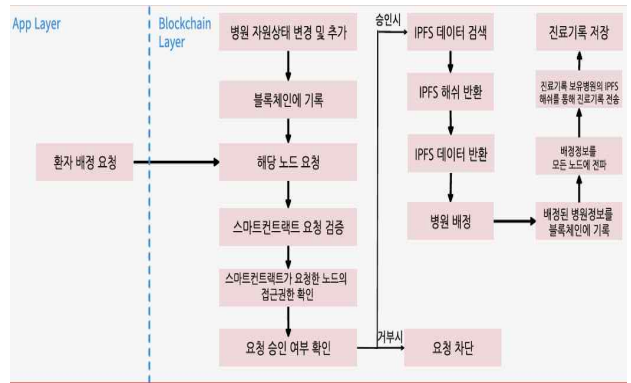


그림 3. 블록체인 서버 순서도  
Fig 3. Blockchain server flow chart

하이퍼레저 페브릭에서는 스마트 컨트랙트를 체인코드(Chaincode)로 구현한다. 체인코드를 통하여 응급 상황의 환자를 적합한 병원에 배정하는 과정을 수행한다. 그림 3은 체인코드로 구현된 DESK시스템이 병원을 배정하는 과정을 서술한 순서도이다. 규격화된 컨트랙트를 통해 신속하게 최적의 병원을 선택하고, 치료지연을 최소화하여 의료 시스템의 효율성을 향상시킨다. 블록체인은 모든 트랜잭션의 내용이 변경하지 못하도록 저장되기 때문에 데이터 위변조의 위험을 원천적으로 차단하며 응급의료 데이터의 투명성과 무결성을 동시에 확보할 수 있다.

IPFS는 데이터 분산 저장을 가능하게 해주는 시스템으로, 방대한 의료 데이터 처리를 가능하게 한다. 대용량 데이터를 전부 블록체인 안에 저장하는 것은 구조상 적합하지 않다. 이러한 문제를 IPFS를 블록체인과 연동하여 해결하였다. 대용량 의료 데이터를 P2P 네트워크 방식의 IPFS에 저장하고 블록체인에 해당 파일의 해시값을 기록하여 데이터를 저장한다. 병원 데이터를 열람하는 과정 또한 IPFS 해시값을 거쳐 데이터를 반환한다.

### III. 결 론

본 연구에서 제안한 한국형 분산 응급의료 시스

템은 기존 중앙집중형 응급의료 시스템의 한계를 개선하고자 블록체인 기반 탈중앙화 네트워크 구조를 도입하였다. 이를 통해 병원 간 실시간 데이터 공유가 원활하게 이루어져 응급 상황 발생 시 환자 상태와 병원 자원 정보를 신속하게 반영할 수 있다. 이러한 실시간 데이터 동기화는 골든타임 확보에 중요한 역할을 하며, 환자를 적절한 병원에 빠르게 배정하여 응급 상황에서의 의사결정을 크게 개선한다.

또한, 스마트 계약을 통해 병원 배정 과정을 자동화하여 인력 개입을 최소화하고 배정 오류를 줄일 수 있다. 탈중앙화된 네트워크 구조를 통해 특정 병원의 문제 발생 시에도 시스템의 가용성을 유지하여 대규모 재난 상황에서도 자원을 효율적으로 분배할 수 있도록 설계되었다. 블록체인의 보안성과 불변성을 통해 의료 데이터가 안전하게 보호되며, 개인정보와 의료 데이터는 고급 암호화 기법을 통해 보호된다. 각 병원은 필수적인 데이터만 선택적으로 공유하여 데이터 유출 위험을 최소화하며, 이를 통해 병원 간 신뢰성 있는 협력 관계가 형성될 수 있다.

향후 발전 방향으로서는 임베디드 헬스케어 기기와 연동하여 실시간으로 환자 상태를 모니터링하고, 고위험군 환자의 위급상황을 자동으로 감지하여 이송하는 서비스를 구축하는 것이다. 이러한 통합적 응급의료 체계는 공공의 안전과 복지 증진에 기여할 것이다.

### 참 고 문 헌

[1] 오영호, “우리나라 응급의료체계의 문제점과 정책방향,” 보건복지 Issue & Focus , vol. 105, pp. 1-8, Sep. 2011.

[2] J. Jeong, S. Kim, M. Kang, W. Cho, Y. Kwon, D. Kim, Y. Yoon, J. Cho, M. Ji, and S. Park, 골든타임 확보를 위한 블록체인의료 데이터 공유 플랫폼 구축, 최종보고서, 충북과학기술혁신원, 발행국가: 대한민국, 발행년월: Dec. 2020. Available: Ministry of Science and ICT, Korea Institute of Science and Technology Evaluation and Planning. [TRKO202100006201]

[3] Hyperledger, <https://hyperledger-fabric.readthedocs.io> [accessed: Oct., 2024]

[4] Seongjun Kang, Seungwon Jeong, Dongwon Jeong, and Hyunjun Jung, "IPFS-based Blockchain storage methods," in Proceedings of KIIT Conference, 2023, pp. 454-458.

# 충돌률과 데이터 길이에 기반한 IEEE 802.11ax 업링크 동적 할당 기법

황영은\*, 김선명\*\*

## Dynamic Allocation of IEEE 802.11ax Uplink Based on Collision Rate and Data Length

Yeongeun Hwang\* and Sunmyeng Kim\*\*

### 요약

본 논문에서는 IEEE 802.11ax 기반의 환경에서 업링크 자원 할당을 최적화하는 방안을 제안한다. [1] Resource Unit (RU) 할당 방식은 사용자당 하나의 RU를 할당할 수 있고 채널 대역폭에 따라 할당 가능한 RU 단위가 정해져있다. 임의로 RU를 배정해주게 되면 패딩 비트로 인해 자원의 낭비가 발생하기 쉽다. 또한, 무선 통신에서는 [2]단말수가 증가하면 높은 충돌 확률로 인해 네트워크 성능이 급격히 저하된다. 이러한 문제를 해결하기 위해 이 논문을 제안한다. 데이터를 대, 중, 소의 세 가지로 그룹화하고, Access Point (AP)가 해당 데이터 길이에 맞는 RU를 할당함으로써 패딩 비트를 줄인다. 대, 중, 소 그룹의 현재 충돌률과 기준 충돌률 값을 비교하여 RU패턴을 조정함으로써 처리율을 향상시켜 준다. 본 연구는 데이터 길이에 따른 RU할당과 충돌률 기반의 동적 자원 관리를 통해 업링크 환경에서 전송의 성능을 향상시키는데 중점을 둔다.

### Abstract

This paper proposes a method for optimizing uplink resource allocation in an IEEE 802.11ax-based environment. [1] The Resource Unit (RU) allocation method allows for one RU to be assigned per user, with the allocable RU units determined by the channel bandwidth. Arbitrarily assigning RUs can lead to resource wastage due to padding bits. Additionally, in wireless communication, [2] an increase in the number of devices can sharply decrease network performance due to a higher collision probability. To address these issues, this paper presents a solution. Data is grouped into large, medium, and small categories, and the Access Point (AP) allocates RUs that match the length of the data to reduce padding bits. By comparing the current collision rates of the large, medium, and small groups to a baseline collision rate, RU patterns are adjusted to improve throughput. This study focuses on enhancing transmission performance in uplink environments through RU allocation based on data length and dynamic resource management based on collision rates.

### Key words

Uplink, RU (Resource Unit) allocation, Dynamic resource management

---

\* 국립금오공과대학교 컴퓨터소프트웨어공학과, 20221392@kumoh.ac.kr

\*\* 국립금오공과대학교 컴퓨터소프트웨어공학과, sunmyeng@kumoh.ac.kr

## 1. 서 론

[ 106, 106, 52, 52, 52, 52, 26 ]

[ 106, 52, 52, 52, 52, 52, 52, 26 ]

IEEE 802.11ax 표준에서는 Orthogonal Frequency Division Multiple Access (OFDMA) 기술을 적용하여 동시에 전송하고자 하는 사용자의 증가에 따른 기존 무선랜 표준의 전송 성능을 최적화하는 것을 목표로 하고 있다. 무선랜에서는 Resource Unit (RU) 할당 개념이 도입되었다. RU 할당 방식의 특징은 사용자당 하나의 RU를 할당할 수 있으며, 채널 대역폭에 따라 수용 가능한 사용자 수와 할당 가능한 RU 단위가 정해져 있다. 할당된 RU에 따라 사용자 간 데이터 전송 시간에 차이가 발생할 경우, 데이터 전송 시간이 짧은 사용자에게 나머지 시간 동안 padding bit를 전송하여 동시에 전송을 마치도록 하고 있다. 그러나 이러한 방식은 동시 전송 사용자 간 데이터 전송 시간의 차이가 클수록 전송 효율성이 떨어지며, 충돌 확률이 증가함에 따라 네트워크 성능이 급격히 저하된다.

따라서 본 논문에서는 IEEE 802.11ax 기반의 고밀도 무선 환경에서 업링크 자원 할당을 효율적으로 최적화하는 방안을 제안한다. 충돌률과 패딩 비트를 고루 고려하여 RU 할당 패턴을 조정함으로써, 네트워크의 전반적인 성능 향상에 기여하고자 한다.

## II. 제안된 방식

본 논문에서는 40MHz 채널 대역폭에서의 OFDMA 기반 업링크 환경을 고려한다. 제안된 방식은 전송할 data의 padding bit를 줄여줌으로써 전송효율의 향상을 목적으로 하고 있다.

AP가 Station (STA) 들의 개수와 데이터 길이를 알고 있다는 가정하에 데이터 길이별로 그룹화를 시킨다. RU는 tone사이즈 별로 구분하여 set를 만들고 충돌률을 고려하여 set 패턴을 동적으로 변경하며 RU를 할당하여 준다. 패턴으로 적용 시킬수 있는 set는 다음과 같다.

[ 242, 106, 52, 52, 26 ]

[ 242, 52, 52, 52, 52, 26 ]

[ 106, 106, 106, 52, 52, 26 ]

균형 있는 데이터를 전송하기 위해 대그룹은 242 tone, 중그룹은 106 tone, 소그룹은 52 tone과 26 tone으로 설정하여 준다. 소그룹의 26 tone은 남은 자원을 채우기 위한 용도로만 사용하여 준다.

전송해야 할 데이터는 전송하기 전 데이터의 길이에 따라 대, 중, 소그룹을 다시 나누어 준다.

처리를 향상을 위한 패턴 변경을 위해 각 그룹별로 충돌률을 계산하여 준다. 그룹별 좋은 효율을 갖는 적정 충돌 확률 범위에서, Clow는 가장 낮은 충돌 확률을 나타내고 Chigh는 가장 높은 충돌 확률을 나타낸다. 그룹별로 충돌률은 다음과 같다.

표 1. 그룹별 충돌률

Table 1. Collision rate by group

	Clow	Chigh
소	0.3	0.5
중	0.25	0.4
대	0.2	0.3

현재 충돌 확률인 Cest가 Chigh을 넘은 그룹이 발생하면 RU의 할당 패턴을 변경하여 데이터를 전송 하여준다. 단, 다른 그룹의 충돌률이 Clow보다 작을 때만 패턴을 변경하여 준다. 각 그룹의 Clow와 Chigh, 우선순위는 다음과 같다.

표 2. 그룹별 우선순위

Table 2. Priority by group

	우선순위
소	0.3
중	0.2
대	0.1

우선순위는 padding bit를 최소화하기 위해 특수한 상황이 아닌 경우, tone의 size가 큰 것에서 작은 것으로 패턴이 바뀌게끔 하기 위함이다.

$$R = M / N * C * P$$

그룹별 데이터의 총 양(M)에 데이터의 개수(N)를



나누고 지난 충돌률(C)을 곱하여 개당 데이터양 대비 충돌률을 계산하여 준다. 여기에 padding bit를 고려하기 위해 우선순위(P)를 곱하여준다. R의 값이 가장 작은 그룹에 할당된 RU를 가지고 와서 현재 충돌률이 Chigh보다 큰 그룹에 추가 할당하여 준다.

만약 같은 값이 나온다면 tone의 size가 더 큰 RU를 이용하여 패턴을 변경하여 주고, 대그룹의 R이 가장 작을 경우에는 중과 소그룹 모두 Cest값이 Clow보다 크면 대그룹의 tone을 중 그룹과 소 그룹에게 모두 나누어 준다.

표 3. 그룹별 데이터 크기와 수의 예시

Table 3. Example data size and number by group

	데이터 총 양	데이터 개수
소	1500	10
중	2500	12
대	2000	2

표 3과 같은 데이터가 있다고 하자. 초기의 패턴은 [ 242, 106, 52, 52, 26 ]으로 전송을 시작한다. 데이터의 개수가 많은 중의 충돌률이 높게 나타나 현재 충돌률의 값이 Chigh을 넘어 표4와 같은 충돌률을 가졌다고 가정하여 보자.

표 4. 그룹별 충돌률 예시

Table 4. Example of collision rate by group

	Cest	Clow	Chigh	R
소	0.4	0.3	0.5	18
중	0.5	0.25	0.4	20
대	0.1	0.2	0.3	10

이때 R의 값은 대그룹이 가장 작으므로 대그룹에 할당된 242tone을 활용하여 패턴을 변경하여 준다. 소그룹의 Cest가 Clow이상 이므로 [ 106, 106, 52, 52, 52, 26 ] 패턴으로 변경하여 준다.

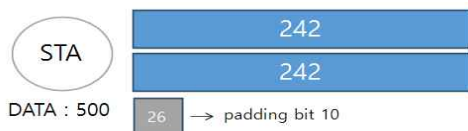


그림 1 패턴 변경 전 STA의 RU할당 예시

Fig. 1. Example of RU allocation for STA before pattern adjustment

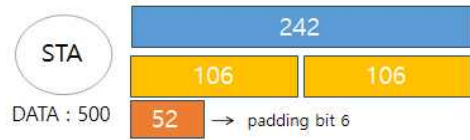


그림 2 패턴 변경 후 STA의 RU할당 예시

Fig. 2. Example of RU allocation for STA after pattern adjustment

그림 1과 그림 2를 보면 패턴 변경을 하더라도 STA의 padding bit도 적절하게 유지되며 할당이 되는 것을 알 수 있다.

### III. 결 론

본 논문에서는 padding bit를 고려한 전송효율 향상을 위한 RU 할당 패턴을 제안하였다. 제안된 방법은 데이터 길이와 충돌률을 고려하여 패턴을 변경시켜 고정된 RU 할당 방법보다 개선된 전송효율을 보여줄 것으로 기대가 된다.

### 참 고 문 헌

- [1] Jeon, Y. (2020). RU-based OFDMA resource allocation for IEEE 802.11ax wireless networks. The Journal of Korean Institute of Communications and Information Sciences, 45(6), 1038-1038.
- [2] Kim, S. (2022). Dynamic uplink channel access protocol based on collision probability in IEEE 802.11ax high density WLANs. Journal of KIIT, 20(2), 71-77.

# 이미지 분류를 위한 전이학습모델 비교

김지원

## A Comparative Study of Transfer Learning Models for Image Classification

Kim Jiwon

### 요약

과도한 사회적 상호작용과 불필요한 콘텐츠가 혼재된 기존 SNS의 한계를 극복하고자 풍경과 동물 사진에 특화된 새로운 플랫폼을 제안하며, 효율적인 콘텐츠 분류를 위해 VGG-16, Inception, ResNet50, EfficientNet-B0 등 CNN기반 전이학습 모델들의 성능을 비교 분석한다.

### Abstract

This study proposes a new platform specialized in landscape and animal photography to overcome the limitations of existing SNS platforms, which are characterized by excessive social interactions and unnecessary mixed content. The performance of CNN-based transfer learning models including VGG-16, Inception, ResNet50 and EfficientNet is compared and analyzed for efficient content classification.

### Key words

Transfer Learning, Convolutional Neural Network(CNN), image Classification, Deep Learning Models

## 1. 서론

현대 사회에서 사진 촬영과 공유는 일상적인 자기표현 및 소통 수단이 되었다. 특히 아름다운 하늘이나 풍경을 마주하였을 때, 길을 걸어가다 귀여운 길고양이를 마주하였을 때, 이를 기록하고 공유하려고 하는 욕구가 보편적이다.

그러나 기존 SNS 플랫폼들은 다음과 같은 문제점을 가지고 있다. 첫째, 댓글, 태그, 공유 등 과도한 사회적 상호작용요구가 존재한다. 둘째, 광고, 텍스트 위주의 불필요한 콘텐츠들이 혼재한다. 셋째,

나와 타인을 계속해서 비교하게 되는 과시 글의 비율이 높다. 넷째, 순수한 사진 공유의 비율이 낮다. 이러한 문제들로 인해 사용자들은 본래의 순수한 공유 목적에서 벗어나 불필요한 심리적 부담과 피로감을 경험하게 된다.

이에 위와 같은 기존 SNS의 문제점은 해결하고 사진 공유의 본질을 회복하여 사진을 사진 그 자체로 집중할 수 있도록 풍경과 동물 사진 콘텐츠만을 올릴 수 있는 애플리케이션을 고안하게 되었다. 이 애플리케이션에서는 사용자가 편하게 사진을 분류하여 검색할 수 있도록 이미지에 태그를 부여하려

고 한다. 본 논문에서는 요구사항에 맞는 태그 부여를 위해 여러 전이 학습 모델을 표1의 클래스 목록과 같이 주변에서 마주할 수 있는 풍경들과 동물들 이미지를 분류하도록 미세조정(fine-tuning)하여 가장 알맞은 모델을 찾고자 한다.

표 1. 데이터셋 구성  
Table 1. Dataset Configuration

클래스 명	train	val
mountain	1,300	510
building	1,300	520
sea	1,300	500
forest	1,200	510
street	1,200	500
dog	1,300	520
cat	1,200	520
squirrel	1,200	500
bird	1,400	510
flower	1,200	490

## II. 전이학습모델

본 논문에서는 대규모 데이터베이스 ImageNet을 데이터셋으로 학습된 CNN(Convolutional neural network) 기반의 전이학습 모델인 ResNet50, VGG16, Inception, EfficientNet-B0 을 활용하여 데이터셋인 표1과 같이 주변에서 자주 마주하는 풍경, 동물을 위주로 하는 이미지분류에 관한 모델들의 성능을 비교, 분석하여 가장 적합한 모델을 찾고자 한다.

### i. VGG16

VGG(Visual Geometry Group)는 CNN의 가장 필수적인 기능을 기반으로 하며, 작은 컨볼루션 필터를 사용하여 낮았던 당시 네트워크의 깊이를 16-19 레이어까지 늘리는 구조를 제안했다. 깊이가 증가함에 따라 분류 정확도가 향상되었으며, Imagenet Challenge 2014에서 localization 1위, classification 2위를 달성한 모델이다. 모델 비교를 위한 학습에서는 VGG-16을 사용하였는데 VGG-16모델은 13개의 컨볼루션 레이어와 3개의 완전 연결 레이어로 구성된 16층 깊이의 네트워크이다. 더 깊은 VGG-19모델에

비해 계산 비용이 적지만 비슷한 성능을 보인다.

### ii. Inception

Inception(GoogleNet)은 VGG와 같은 시기 개발 되었으나, 네트워크 향상을 위한 효율적인 컴퓨팅 자원 활용을 목표로 설계된 모델이며, 모델 비교를 위한 학습에서는 기존 GoogleNet(v1)의 구조를 개선한 모델인 Inception v3를 사용하였는데 이는 큰 필터(5x5, 7x7)를 여러개의 3x3 필터로 분해하여 파라미터 수를 줄이고, RMSProp 옵티마이저와 Batch Normalization을 적용하여 학습 안정성을 높여 v4보다 검증된 안정성을 제공하면서도 v1, v2보다 향상된 성능을 보인다.

### iii. ResNet50

ResNet은 VGG와 같은 기존 CNN 모델들이 층의 깊이가 깊어짐에 따라 발생할 수 있는 기울기 소실(vanishing gradient)와 같은 성능 저하 문제를 해결하기 위해 제안된 모델로 residual learning이라는 개념으로 단순히 층을 쌓는 것이 아닌 Residual Block구조를 사용하여 입출력 사이에 연결을 추가하여 학습과정에 잔차에 대한 학습을 진행할 수 있다. 모델 비교를 위해 학습한 ResNet50 모델은 다른 깊이의 ResNet모델과 비교하여 적절한 특징 추출 능력과 컴퓨팅 자원능력이 더 깊은 모델과 비교해도 비슷한 성능을 가지고 있다.

### iv. EfficientNet

EfficientNet은 CNN 모델의 성능을 효율적으로 향상시키기 위해서 네트워크 깊이, 너비, 해상도를 동시에 균형 있게 조절하는 compound Scaling 방식을 제안한 모델로 NAS(Neural Architecture Search)로 설계한 기본 구조에 최적의 스케일링 계수를 적용하여 B0~B7까지 다양한 크기의 모델을 제공한다. ImageNet에서 기존 CNN모델들보다 적은 연산량으로도 높은 정확도를 달성하였다. 모델 비교를 위한 학습은 가장 기본이 되는 베이스 모델인 EfficientNet-B0을 사용하였는데 더 큰 모델들에 비해 가벼우면서도 효율적인 성능을 제공하여 제한된 컴퓨팅 자원에서도 좋은 결과를 얻을 수 있다.

### III. 모델별 학습결과

데이터 셋은 앞서 보여주었던 표1의 클래스 대로 준비를 하였고, 전처리 과정을 통해 이미지를 각 모델에 맞게 사이즈를 맞춤과 동시에 데이터를 회전시키거나 확대를 시켜 부족한 데이터를 보충하였다. 학습 기본 조건은 Batch size는 64, epochs는 30으로 하여 모델학습을 진행하였고, 그림1, 그림2, 그림3, 그림4에서 보이는 것과 같이 정확도(Accuracy)와 손실도(Loss)값을 측정 하였다.

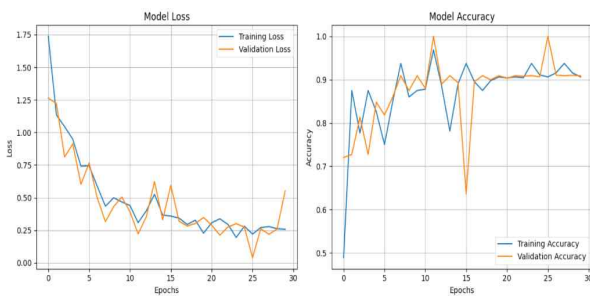


그림 1. VGG-16모델의 훈련 및 검증 손실과 정확도 변화  
Fig. 1. Training and validation loss and accuracy changes of VGG-16.

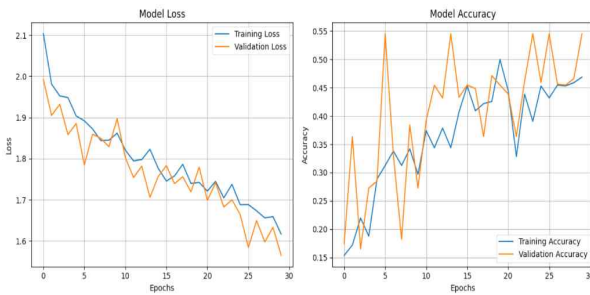


그림 2. Inception(V3)의 훈련 및 검증 손실과 정확도 변화  
Fig 2. Training and validation loss and accuracy changes of Inception(V3)

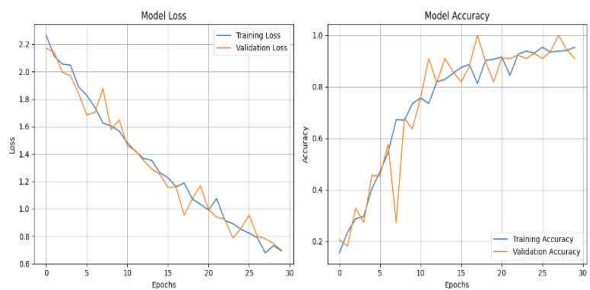


그림 3. ResNet50모델의 훈련 및 검증 손실과 정확도 변화  
Fig. 3. Training and validation loss and accuracy changes of esNet50.

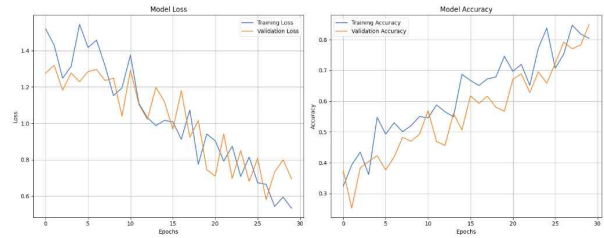


그림 4. EfficientNetB0의 훈련 및 검증 손실과 정확도 변화  
fig4. Training and validation loss and accuracy changes of EfficientNetB0.

학습결과 모든 모델이 높은 정확도(Accuracy)를 보였지만, VGG-16을 제외한 모델들의 경우 손실도(Loss)가 높게 나타나는 것을 확인할 수 있었다. 테스트 결과 정확도(Accuracy)는 높게 수렴하고, 손실도(Loss)값이 낮은 VGG-16모델을 이미지 분류 모델로 사용하고자 한다.

### IV. 결론

기존 SNS 플랫폼의 한계를 해결하고, 사진을 기록하고 공유하는 행위의 본질에 집중할 수 있도록 하는 플랫폼을 고안하였다. 본 논문에서는 앞서 언급한 플랫폼을 위해 이미지 분류 모델설계에 필요한 CNN 기반의 전이학습 모델들을 비교분석 하여 알맞은 모델을 선정하고자 하였다.

전이학습 모델은 VGG-16, Inception(v3), ResNet50, EfficientNetB0을 사용하여 비교하였고, 각 모델별 정확도(Accuracy)와 손실도(Loss)를 비교해본 결과 해당 플랫폼에서는 높은 정확도와 낮은 손실도를 보였던 VGG-16을 사용할 계획이다. 더 정확한 태그 생성을 위해 OCR기능을 함께 이용하여 이 후 실제 서비스에서의 분류 정확도 측정까지 진행할 계획이다.

### 참고 문헌

- [1] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in ICLR, 2014
- [2] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern

Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

- [3] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens 및 Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 미국 네바다주 라스베이거스, 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.eption Architecture for Computer Vision" in

# RAG와 CoVe 기법을 활용한 LLM의 환각 현상 감소 연구: 외국인 노동자 대상 노무 상담 챗봇 사례를 중심으로

우성현\*, 강태현\*, 김건호\*, 이브로힘\*, 김재수\*\*

## Reducing Hallucination in LLMs Using RAG and CoVe Techniques: A Case Study of a Labor Counseling Chatbot for Foreign Workers

Woo Seong Hyun\*, Kang Tae Hyun\*, Kim Geon Ho\*, Ibrahim\*, and Kim Jae Soo\*\*

### 요약

본 연구는 외국인 노동자를 위한 노무사 챗봇을 개발하고, RAG와 CoVe 기법을 활용하여 Hallucination 현상을 감소시키는 전략을 제시한다. GPT-4o-mini 모델을 기반으로 정확한 노동법 정보를 제공함으로써, 외국인 노동자의 권리 보호에 기여하고자 한다.

### Abstract

This study develops a labor attorney chatbot for foreign workers and proposes strategies to reduce hallucination using RAG and CoVe techniques. Based on the GPT-4o-mini model, it aims to provide accurate labor law information, contributing to the protection of foreign workers' rights.

### Key words

Foreign Workers, Labor Attorney Chatbot, RAG, CoVe, Hallucination Reduction, GPT-4o-mini

## 1. 서론

### 1.1 연구 배경 및 목적

외국인 노동자의 증가로 한국 경제에서 그들의 역할이 중요해지고 있으나, 언어 장벽과 복잡한 노동법으로 인해 권리와 의무 이해에 어려움을 겪고 있다. 이는 노동 환경에서의 부당한 대우로 이어질

수 있다. 인공지능 기술, 특히 챗봇은 정보 접근성 향상에 기여하고 있지만, 법률 정보 전달의 과정에서 Hallucination 문제로 인해 정확성과 신뢰성에 한계가 있다.

본 연구는 RAG(Retrieval-Augmented Generation)와 CoVe(Chain-of-Verification) 기법을 활용하여 Hallucination을 줄이고, 외국인 노동자에게 정확한 노동법 정보를 제공하는 신뢰성 높은 노무사 챗봇 개발을 목적으로 한다.

\* 경북대학교 컴퓨터학부 학부생, wshdhkd@naver.com, kang4030@knu.ac.kr, kgh67811@knu.ac.kr, rustamjonov36@gmail.com

\*\* 경북대학교 컴퓨터학부 교수, kis@knu.ac.kr

## II. 선행 연구

### 2.1 RAG(Retrieval-Augmented Generation)

RAG[1]는 대규모 언어 모델과 정보 검색 기술을 결합한 방법론으로, Retriever와 Generator로 구성된다. Retriever는 관련 문서를 검색하고, Generator는 이를 기반으로 답변을 생성한다. RAG는 외부 지식을 활용해 모델의 최신성과 정확성을 향상시키며, Hallucination을 감소시킨다.

### 2.2 CoVe(Chain-of-Verification)

CoVe[2]는 언어 모델이 생성한 답변의 정확성을 자체적으로 검증하는 추가 단계를 도입한 기법이다.

본 연구는 RAG와 CoVe를 통합 활용하여 LLM의 Hallucination 현상을 감소시키는 전략을 수립하고 LLM의 신뢰성과 실용성을 향상시키고자 한다.

## III. 챗봇 LLM 구성

### 3.1 데이터 수집 및 전처리

① 법제처의 국가법령정보센터에서 PDF로 되어 있는 노동법 관련 37개 법률 문서를 저장한다.

② PyMuPDF 라이브러리를 사용하여 PDF 파일에 존재하는 텍스트 추출한다.

③ 정규표현식을 사용하여 각 법률의 목차, 장,

조 항목을 추출한다.

④ 추출된 목차 정보를 기반으로 법률 텍스트를 장과 조 단위로 분할한다.

⑤ tiktoken 라이브러리를 사용하여 텍스트 토큰화 및 최대 토큰 수(8192개)를 초과하지 않도록 청크를 조정한다.

### 3.2 임베딩 처리 및 벡터 데이터베이스 구축

OpenAI의 text-embedding-ada-002 모델로 법률명, 장, 조 제목 등의 청크별 임베딩을 생성하고, FAISS로 인덱싱했다. 임베딩은 law\_embeddings.index 파일에, 메타데이터는 law\_embeddings.json 파일에 저장했다. 일부 데이터를 수동 검증해 법률 조항 간 유사성을 파악하고 검색 효율을 높였다.

### 3.3 파이프라인 구성

RAG와 CoVe 기법을 적용한 LLM의 파이프라인은 그림 1과 같다.

① Baseline Response Generation Stage: RAG 기법 LLM을 활용해 사용자의 질문에 대한 기본적인 답변을 생성한다. 초기 답변은 프로세스의 기초가 되며, 일반적인 정보에 기반한 내용을 제공한다.

② Verification Question Generation Stage: 초기 답변의 정확성을 높이기 위해 Few-shot 학습 기법을 적용해 검증 질문을 생성한다.

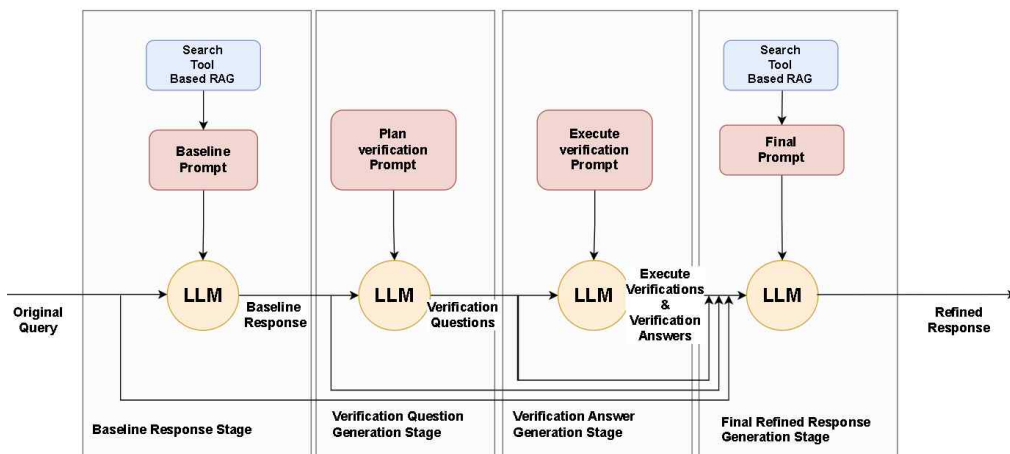


그림 1. LLM 파이프라인 다이어그램

③ Verification Answer Generation Stage: 생성된 각 검증 질문에 대해 RAG를 사용해 답변을 도출한다. 이 단계에서는 추가적인 정보를 바탕으로 초기 답변을 보완하고 더 정밀한 응답을 생성한다.

④ Final Refined Response Generation Stage: ①에서 생성한 초기 답변과 ③에서 생성한 검증 답변을 결합해 최종 개선된 응답을 도출한다. 이때, Few-shot 학습을 다시 적용하여 정제된 정보를 반영한 최종 답변을 생성한다.

### 3.4 사용자 질문에 대한 응답 생성 및 검증

GPT-4o-mini 모델은 RAG를 활용해 지식 기반을 구성하고, 이에 따라 사용자 질문에 대한 답변을 생성하고 할루시네이션을 감소시킨다.

① 질문을 임베딩하여 FAISS 인덱스를 통해 관련 법률 조항을 검색한다.

② 검색된 법률 조항을 바탕으로 GPT-4o-mini 모델이 답변을 생성한다.

RAG를 통한 답변 생성 후, CoVe 기법을 적용하여 문맥과 정보를 다단계로 검증함으로써 할루시네이션을 추가로 감소시킨다. 이 과정에서 모델은 자

체적으로 답변의 정확성을 검증한다. 그 결과는 그림 2와 같다.

## IV. 실험 및 결과

### 4.1 실험 설계 및 평가 방법

본 연구에서는 외국인 근로자들이 자주 하는 35개의 질문을 기반으로 RAG와 CoVe를 적용한 챗봇 모델의 성능을 평가한다. 이 실험은 다음과 같은 방식으로 설계하였다.

RAG 기법을 적용한 모델과 RAG + CoVe 기법을 함께 적용한 모델 두 가지를 비교한다. 아래 3가지 지표를 사용해 모델을 평가한다.

① 정확도(Accuracy): Verification 질문에 대한 답변 중 정답의 비율로 계산한다. 정확도는 생성된 Verification 질문 수를 기준으로 정확한 정보를 제공한 비율을 사용한다.

② 할루시네이션 발생 빈도(Hallucination Rate): 할루시네이션이 발생한 질문 수를 전체 질문 수로 나눈 비율로 측정한다.

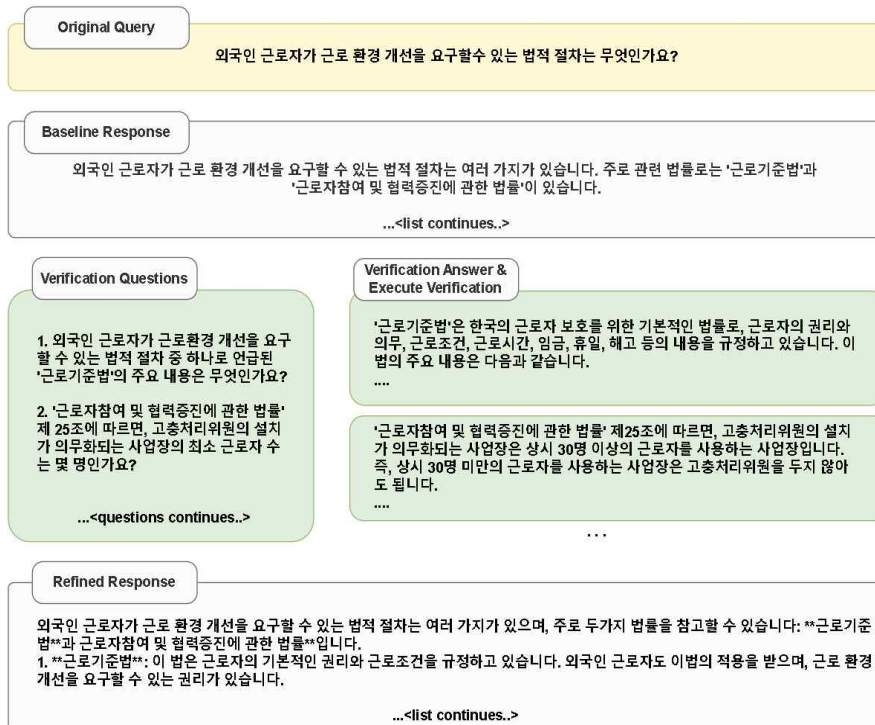


그림 2. 다단계 검증을 통한 응답 예시



③ 응답 속도(Response Time): 사용자 질문이 입력된 시점부터 모델이 답변을 생성한 시점까지의 시간을 측정하여, 각 질문에 대한 평균 응답 시간을 계산하여, 두 모델의 응답 속도를 비교한다.

## 4.2 실험 결과 및 할루시네이션 감소 분석

하나의 질문 당 평균 7.23개의 검증 질문이 생성되었고, 실험 결과는 표 1과 같다.

① 정확도 분석: RAG와 CoVe를 함께 적용한 모델은 RAG만 적용한 모델에 비해 92%의 정확도를 기록하여, 2% 높은 정확도를 보여준다. 이는 CoVe 기법이 추가된 후 검증 절차를 통해 답변의 정확성이 향상되었음을 보여준다.

② 할루시네이션 감소: RAG + CoVe 모델은 할루시네이션 발생 빈도가 51%로, RAG 모델의 57%보다 낮다. 이를 통해 CoVe 기법을 통해 모델이 잘못된 정보를 제공하는 빈도가 감소한 것을 확인할 수 있다.

③ 응답 속도: CoVe 기법은 추가적인 검증 절차로 인해 응답 시간이 RAG에 비해 현저히 길었으며, 이는 실시간 응답이 중요한 서비스에서는 고려해야 할 요소이다.

이 실험을 통해 RAG와 CoVe의 결합이 할루시네이션 빈도를 줄이고 정확도를 높이는 데 긍정적인 효과가 있음을 확인할 수 있다.

## V. 결론 및 향후 연구

### 5.1 연구 요약 및 주요 성과

본 연구에서는 외국인 근로자들을 위한 노무사 챗봇에 RAG와 CoVe 기법을 적용하여 할루시네이션 감소 및 답변 정확도 향상 효과를 분석했다.

실험 결과, RAG + CoVe를 적용한 모델은 RAG

만을 적용한 모델에 비해 더 높은 정확도와 낮은 할루시네이션 발생 빈도를 기록했으며, CoVe 기법을 통한 검증 절차가 답변의 신뢰성을 높이는 데 기여함을 확인했다.

### 5.2 연구 한계 및 향후 개선 방향

연구의 한계로는 CoVe 기법 적용 시 응답 시간이 증가한다는 점이다. 이는 실시간 응답이 필요한 시스템에서 큰 제약이 될 수 있다. 향후 연구에서는 응답 속도를 개선하는 방법과 다양한 도메인에서 RAG + CoVe 모델의 성능을 평가하여 확장 가능성을 확인하는 것이 중요하다. 또한, CoVe 를 적용한 검증 절차가 모든 질문에 대해 동일한 성과를 나타내지 않는 문제를 해결하기 위한 검증 질문 최적화 연구가 필요하다.

## 사사문구

“이 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 통해 수행되었습니다”(2021-0-01082)

## 참 고 문 헌

- [1] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." *Advances in Neural Information Processing Systems*, 33, 9459-9474.
- [2] Li, T., Shuster, K., Weston, J., & Dinan, E. (2023). "Chain-of-Verification Reduces Hallucination in Large Language Models." *arXiv preprint arXiv:2303.05381*.

표 1. 각 기법에 대한 평가 결과

	Accuracy (%)	Hallucination Rate (%)	Response Time (s)
<b>RAG</b>	90	57	7.85
<b>RAG + CoVe</b>	92	51	64.70

# 주파수 해석과 CNN을 이용한 진동 센서에 잡음이 있는 유도전동기 시스템의 고장 진단

김민규\*, 권효선\*, 유지민\*, 이호형\*, 최진우\*, 이인수\*\*

## Fault Diagnosis of an Induction Motor System with Noisy Vibration Sensors Using Frequency Analysis and CNN

Min Kyu Kim\*, Hyo seon Gweon\*, Ji Min Yu\*, Ho Hyeong Lee\*, Jin Woo Choi\*, and In-Soo Lee\*\*

### 요 약

본 논문에서는 시간 도메인 진동 데이터를 주파수 해석하고 CNN을 이용해 진동센서에 잡음이 있는 유도전동기 시스템의 고장 진단 모델을 개발하였다. 정상, 회전자 고장, 베어링 고장 등 세 가지 상태에 대한 진동 데이터를 주파수 해석하여 모델에 학습시켰으며, 모델 평가 시 잡음이 추가된 데이터를 사용하여 실제 진동 센서에 잡음이 존재하는 상황을 가정하였다. 이를 통해 진동 데이터만으로도 빠르고 정확한 고장 진단이 가능하며, 실제 잡음이 있는 산업 현장에서도 우수한 진단 성능을 기대할 수 있음을 제시하였다.

### Abstract

In this study, a fault diagnosis model for an induction motor system with noise in a vibration sensor was developed using frequency analysis of time-domain data and a CNN. Vibration data representing three states—normal, rotor fault, and bearing fault—were frequency-analyzed and used to train the model. To evaluate the model's performance, noise-added data were incorporated to simulate a situation with sensor noise, enabling fault diagnosis under realistic noise conditions.

### Key words

Frequency Analysis, Fault Diagnosis, Vibration Data, CNN, White Gaussian Noise, Bias

## 1. 서 론

유도 전동기의 고장 진단에 시간 도메인 진동 데이터를 사용할 경우, 진동 센서의 잡음이 존재하면 분류 정확도가 크게 저하되는 경향이 있음을 확인할 수 있었다. 산업 현장에서 기계의 고장 진단을 빠르고 정확하게 수행하는 것은 생산성 유지, 안전

성 확보, 설비 수명 연장의 측면에서 매우 중요하므로, 고장 진단의 정확도를 높일 필요가 있다.

본 논문에서는 고장 진단 정확도를 향상시키기 위해 시간 도메인 데이터를 주파수 해석하여 사용하고, 신경망 입력의 형상을 적절히 전처리하였다. 이를 통해 잡음이 포함된 데이터에서도 분류 정확도를 높여 정확하고 안정적인 고장 진단 성능을 구

\* 경북대학교 전자공학부,

\*\* 경북대학교 전기전자공학부 교신저자, insoolee@knu.ac.kr

현하고자 한다.

## II. 주파수 해석과 CNN을 이용한 모델 제작

본 논문에서는 시간 도메인 데이터의 주파수 해석을 통해 얻은 주파수 도메인 데이터와 딥러닝 모델 CNN을 이용하여 정상, 회전자 고장, 베어링 고장 세 가지 상태에 대한 유도전동기의 고장진단을 수행하는 모델을 제작하였다.

### 2.1 주파수 해석

본 논문에서는 주파수 해석된 진동 데이터를 CNN의 입력으로 사용하였다. 시간 도메인 데이터를 입력으로 사용할 경우, 진동 센서의 잡음이 존재할 때 모델의 분류 성능이 크게 저하되는 것을 확인하였다. 이에 따라, 주파수 해석을 통해 얻은 주파수 도메인 데이터를 활용하여 분류 성능을 향상시켰다.

#### 2.1.1 FFT를 이용한 주파수 해석 데이터 획득

본 논문에서는 Python의 Scipy 라이브러리를 이용하여 시간 도메인 진동 데이터를 FFT 변환하였다. 이후 DC 성분을 포함하는 0Hz 값을 제거하고, 양의 주파수 부분만을 사용하였다. 아래 그림 1은 주파수 해석된 데이터의 그래프에서 각 상태에 따른 고유 진동수가 존재함을 보여준다.

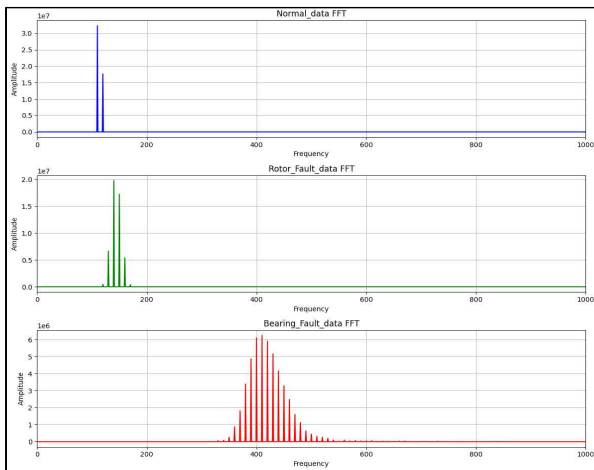


그림 1. 주파수 해석 그래프  
Fig. 1. Frequency Analysis Graph

#### 2.1.2 잡음이 존재하는 데이터 분석

본 논문에서는 진동 센서에 잡음이 있는 경우의 정확한 고장 진단을 목표로, 모델 평가에 사용하는 시험 데이터에 화이트 가우시안 노이즈(WGN)와 Bias를 각각 추가하여 사용하였다. 먼저 모델 입력 형상을 결정하기 위해, 잡음이 존재하는 경우의 주파수 해석 그래프를 분석하였다. 아래 그림 2는 평균 0, 분산 0.1의 WGN을 첨가했을 때의 주파수-진폭 그래프를 보여준다.

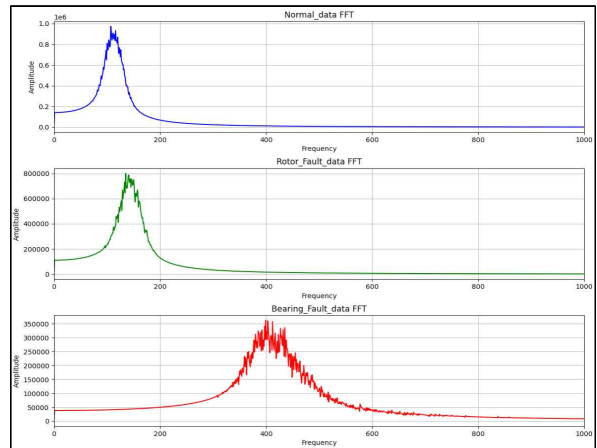


그림 2. 잡음이 첨가된 주파수 해석 그래프  
Fig. 2. Frequency Analysis Graph With Noise

#### 2.1.3 잡음 유무에 따른 피크 주파수 분석

위의 결과를 바탕으로 정확한 분석을 위해 주파수 해석 데이터에서 피크 주파수 값을 출력해 보았다. 그림 3은 각 데이터에서의 잡음 유무에 따른 피크 주파수 출력값을 보여준다.



그림 3. 잡음 유무에 따른 피크 주파수  
Fig. 3. Peak Frequency Based on Noise

## 2.2 CNN 입력을 위한 데이터 전처리

잡음이 첨가된 데이터에서도 유사한 피크 주파수 대역을 유지한다는 점을 활용하여, 주파수 해석 데이터에 대해 적절히 전처리한 후 CNN의 입력으로 사용하였다.

### 2.2.1 진폭 크기 상위 주파수 5개 추출

본 논문에서는 모델 입력 데이터로 진폭 크기 기준 상위 5개의 주파수를 추출하였다. 먼저 획득한 진동수 데이터 100쌍에 대해 주파수 해석을 진행한 후, 주파수-진폭 쌍에서 진폭을 기준으로 내림차순 정렬하여 상위 5개 주파수를 추출하였다. 아래 그림 4는 주파수 해석 데이터에서 상위 주파수 5개를 추출하는 과정을 보여준다.

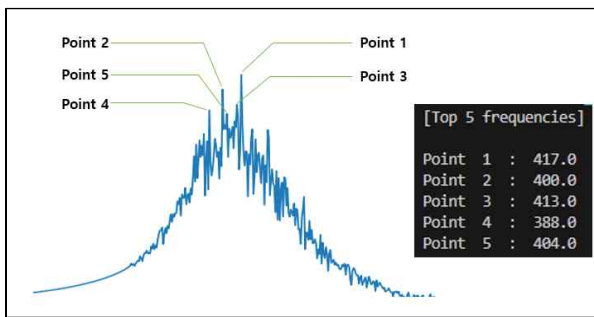


그림 4. 상위 주파수 5개 추출

Fig. 4. Extracting the Top 5 Frequencies

## 2.3 CNN 모델 제작

본 논문에서는 획득한 데이터를 학습하고 고장진단을 위한 분류를 수행하는 모델로 CNN을 선택하였다. CNN의 Conv1D 계층을 이용하여 합성곱층을 구성하였고, 풀링 크기는 입력 데이터의 크기를 유지하기 위하여 1로 설정하였다.

### 2.3.1 1차 합성곱 CNN 모델 제작

본 논문에서는 1차원 데이터를 모델 입력으로 사용하였기에 Conv1D 계층으로 합성곱층을 구성하였다. 활성화 함수는 ReLU를 사용하고, 필터는 16개로 설정하였다. 이후 완전 연결 계층으로 전달하며, 이 계층에서도 활성화 함수로 ReLU를 사용하였다. 마지막 출력층의 활성화 함수는 Softmax를 적용하여 3가지 상태에 대한 예측 정확도를 도출하고, 최

종적으로 원-핫 인코딩된 결과값을 반환받았다.

## III. 유도전동기 고장진단 모델 평가

본 논문에서는 유도전동기의 3가지 상태에 대해 잡음이 없는 데이터를 사용하여 모델을 학습하였다. 모델 평가 과정에서는 실제 환경에서 진동 센서의 잡음 발생 가능성을 고려하여 시험 데이터에 잡음을 추가하여 사용하였다.

### 2.3.2 유도전동기 진동 데이터 학습

정상, 회전자 고장, 베어링 고장 각각의 상태에 대한 시간 도메인 진동 데이터 100쌍씩 총 300쌍을 주파수 해석하였다. 이후 모델 학습에 사용할 훈련 데이터 210쌍과 모델 평가에 사용할 시험 데이터 90쌍으로 분류하여 사용하였다. 학습 Epochs는 20회, 배치 크기는 16으로 설정하였다.

### 2.3.2 모델 성능 분석

본 논문에서는 학습한 모델을 평가하는 과정에서 잡음이 존재하는 진동 센서를 가정하여 시험 데이터에 WGN과 Bias 두 가지 잡음을 추가하여 신경망의 입력으로 사용하였다.

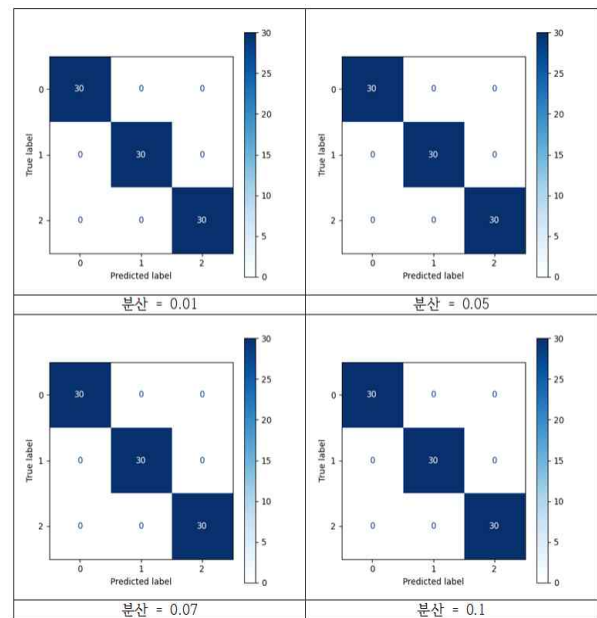


그림 5. WGN가 포함된 데이터 분류 혼동행렬

Fig. 5. Confusion Matrix with WGN

WGN의 경우 평균값은 0이고, 분산 값이 0.01, 0.05, 0.07, 0.1인 4가지 경우에 대해 성능 분석을 진행하였으며, 모든 경우에서 100%의 분류 정확도를 기록하였다. 그림 5는 WGN이 포함된 시험 데이터를 모델 평가에 사용하여 얻은 혼동 행렬을 보여 준다.

Bias의 경우는 0.05, 0.1, 0.5, 1.0인 4가지 경우에 대해 분석하였으며, 모든 Bias에 대해 100%의 분류 정확도를 나타내었다. 아래 그림 6은 Bias가 포함된 시험 데이터를 모델 평가에 사용하여 얻은 혼동 행렬을 보여준다.

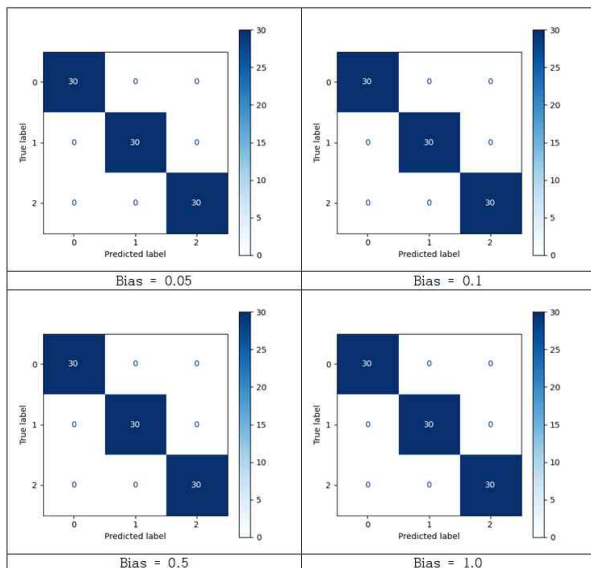


그림 6. Bias가 포함된 데이터 분류 혼동행렬  
Fig. 6. Confusion Matrix with Bias

#### IV. 결 론

본 논문에서는 진동 센서에 잡음이 있는 유도전동기 시스템의 고장 진단을 목표로 CNN 분류 모델을 제작하였으며, 시험 데이터에 잡음을 추가하여 모델 평가를 진행하였다. 분류 성능을 향상시키기 위해 주파수 해석 데이터를 활용하였고, 진폭 기준 상위 주파수 5개를 모델 입력으로 사용하여 피크 주파수 대역을 전달하는 효과를 얻음과 동시에 메모리 사용량과 자원 소비를 줄일 수 있었다. 그 결과, 잡음이 첨가된 시험 데이터에서도 우수한 분류

정확도를 기록하며 모델의 성능을 크게 향상시킬 수 있었다.

이번 연구를 통해 진동 데이터의 주파수 특징을 활용해 잡음을 효과적으로 제거함으로써, 주파수 해석 기반의 고장 진단 효과성을 확인할 수 있었다. 특히, 시간 도메인 해석에서는 Bias가 존재하는 경우 분류 정확도가 크게 저하된 반면, 주파수 해석의 경우 DC 성분을 제거하는 방법으로 분류 정확도를 크게 향상시킬 수 있었다. 이는 산업 현장에서 진동 센서의 잡음 유무와 관계없이 정확한 진단을 가능하게 하는 실시간 유도전동기 고장 진단 시스템 개발에 기여할 수 있을 것으로 기대된다. 향후 연구에서는 다양한 환경에서 발생하는 잡음 유형을 반영한 모델의 성능 개선을 목표로, 시간 및 주파수 도메인 정보를 함께 활용한 고장 진단 기법의 적용 가능성을 탐구할 예정이다.

#### 참 고 문 헌

- [1] Jarmo Ilonen, Joni-Kristian Kamarainen, "Diagnosis Tool for Motor Condition Monitoring", IEEE Trans. Ind. Appl. vol. 41, no. 4, July/August, 2005.
- [2] M. Benbouzid, M. Vieira, and C. Theys, "Induction motors' faults detection and localization using stator current advanced signal processing techniques," IEEE Trans, Power Electron., vol. 14, no. 1, pp. 14-22, Jan. 1999.

# 소규모 개발 환경에서 효율적인 딥러닝 모델 선택 방안

박상우\*, 한동균\*, 서동섭, 손경락, 박세진\*

## Efficient Deep Learning Model Selection in Small Development Environments

Sang-Woo Park\*, Dong-Gyun Han\*, Dong-Seop Seo, Kyeong-Rak Son and Se-Jin Park\*\*

### 요약

인공지능의 일상적 사용이 증가하면서 다양한 분야에서 혁신적인 성과가 나타나고 있다. 하지만 대부분의 딥러닝(Deep Learning) 기술은 대규모 자원을 요구하는 환경에서 개발된다. 반면, 소규모 개발팀이나 개인 개발자는 제한된 자원으로 인해 딥러닝 기술을 효율적으로 활용하기 어려운 상황이다. 본 논문에서는 소규모 개발팀과 개발자를 위한 제한된 환경에서 효율적인 딥러닝 사용 방안을 분석하고, 컴퓨팅 자원을 고려한 작은 데이터 세트에 적합한 모델을 제안한다. 이를 통해 소규모 개발팀과 개발자들이 제한된 자원에서도 최적의 모델을 선택하고, 시간 대비 효율적인 결과를 얻을 수 있도록 돕는다.

### Abstract

As the daily use of artificial intelligence increases, innovative achievements are being made in various fields. However, most deep learning technologies are developed in environments that require large-scale resources. On the other hand, it is difficult for small development teams or individual developers to efficiently use deep learning technologies due to limited resources. In this paper, we analyze efficient use of deep learning in a limited environment for small development teams and developers, and propose a model suitable for a small dataset considering computing resources. Through this, small development teams and developers can select the optimal model even with limited resources and obtain time-efficient results.

### Key words

Deep learning, optimal model, limited resources, obtain time-efficient results

### 1. 서론

딥러닝 기술의 발전은 다양한 산업에서 큰 성과

를 낳았지만, 대부분의 딥러닝 연구는 대규모 자원을 전제로 한 환경에서 이루어진다. 반면, 소규모 개발팀이나 개인 개발자들은 제한된 자원과 환경에

\*계명대학교 컴퓨터공학과, ja513698@naver.com, hdg5639@naver.com, ehdtjq1278@naver.com, rudfkrlife@naver.com, \*\*계명대학교 컴퓨터공학과, baksejin@kmu.ac.kr

서 작업해야 하므로, 이들을 위한 연구는 상대적으로 부족한 실정이다. 대규모 자원을 보유한 기업들이 딥러닝 기술을 활용할 때 시간 대비 성과가 훨씬 더 뛰어나지만, 소규모 개발팀과 개인 개발자들의 시간 대비 성과는 크게 떨어진다. 이는 자원 대비 효율적인 모델을 사용하지 못한 결과이며, 딥러닝 기술의 한계를 드러내고 있다. 소규모 팀들이 이러한 한계를 극복하기 위해서는 자원의 제약을 고려한 경량화된 모델을 사용한 효율적인 학습 전략이 필요하다. 이러한 모델들은 적은 데이터 세트와 제한된 컴퓨팅 자원에서도 높은 성능을 발휘할 수 있어야 한다.

본 논문에서는 소규모 개발팀과 개인 개발자를 위한 제한된 자원 환경에서 효율적으로 사용할 수 있는 인공지능 모델을 제안하고자 한다. 이를 위해, 다양한 모델을 데이터 세트 크기와 복잡도에 따라 비교하며 딥러닝 모델의 성능 차이를 분석한다. 이미지 분류, 객체 탐지, 장면 분류 등의 작업을 통해 각 모델이 어떻게 성능을 발휘하는지 구체적으로 보여준다. 이를 바탕으로, 소규모 개발팀은 상황에 맞춰 가장 효율적인 모델을 선택할 수 있게 된다.

## II. 딥러닝 모델 성능 비교

### 2.1 실험 환경

#### 2.1.1 하드웨어 구성

본 실험에서 사용되는 컴퓨터 사양이다. CPU는 Intel i9-11900KF, GPU는 NVIDIA GeForce RTX 3080, RAM은 64GB의 사양으로 실험이 진행되었다.

#### 2.1.2 개발 환경 및 목적

본 실험은 비교적 고사양의 컴퓨팅 환경에서 수행되었으나, 소규모 팀 및 개인 개발자들이 주로 사용하는 중간 혹은 저 사양의 환경에서도 충분히 재현 가능하고, 성능 저하 없이 유사한 결과를 제공할 수 있는 모델 선택과 권장 가이드 사용을 목표로 진행되었다.

### 2.2 이미지 분류 모델별 테스트

#### 2.2.1 이미지 분류 평균 정확도 테스트

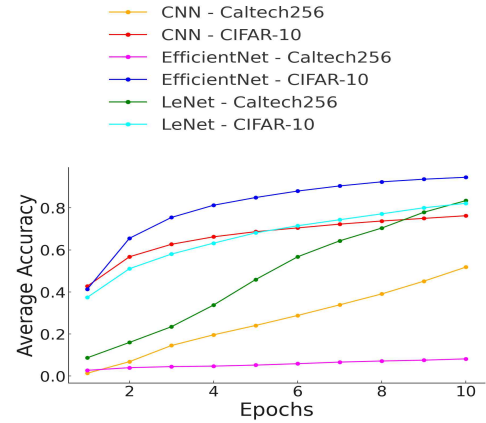


그림 1. 이미지 분류 평균 정확도  
Fig. 1. Image Classification Average Accuracy

본 실험에서는 CIFAR-10과 Caltech256 데이터 세트에서 CNN, EfficientNet, LeNet 모델의 이미지 분류 평균 정확도, 이미지 분류 평균 로스를 Epoch 수에 따라 비교한 결과를 보여준다. <그림 1>에 따르면 CIFAR-10 데이터 세트를 사용한 결과 CNN은 Epoch 2에서 이미 0.6 이상의 정확도를 기록하였다. 이후 Epoch 10에 가까워질수록 최종적으로 0.8에 가까운 정확도를 보였다. EfficientNet은 Epoch 2에서 이미 0.6 이상의 정확도를 기록하며, Epoch 10에 가까워질수록 0.9에 근접한 성능을 보여준다. LeNet은 Epoch 10에 가까워질수록 0.8에 가까운 높은 정확도를 보였다.

반면에, Caltech256 데이터 세트에서는 CNN은 Epoch 4 이전에는 0.2의 낮은 정확도를 보이다가 점차 상승하여 Epoch 10에서 0.5의 정확도를 보였다. EfficientNet은 Epoch 10까지 0.1에 가까운 매우 낮은 정확도를 기록하였다. LeNet은 Epoch 4 이전에는 0.4에 근접하는 낮은 정확도를 보였지만 Epoch 10에 가까워질수록 0.8의 정확도를 보였다.

#### 2.2.2 이미지 분류 평균 로스 테스트

<그림 2>에 따르면 EfficientNet은 CIFAR-10에서 가장 빠르게 손실이 감소하는 모습을 보였다. Epoch 초반부터 손실이 급격히 줄어들며, Epoch 10에 가까워질수록 약 0.5에 가까운 평균 손실을 기록하여

가장 뛰어난 성능을 나타냈다. CNN과 LeNet 역시 손실이 꾸준히 감소하며, Epoch 10에서는 각각 1.0 과 0.8에 가까운 손실 값을 기록했다.

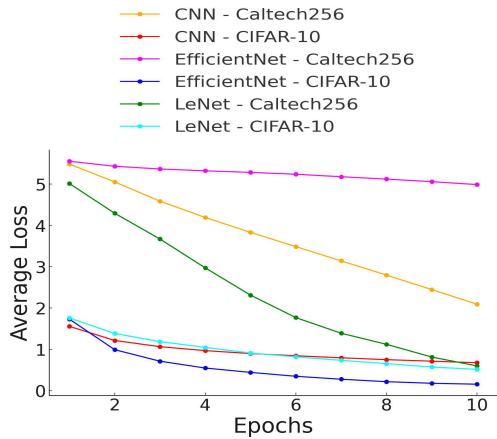


그림 2. 이미지 분류 평균 로스  
Fig. 2. Image Classification Average Loss

반면에, Caltech256 데이터 세트를 사용 한 결과 EfficientNet은 Epoch 10에 도달할 때까지 평균 손실 값이 5.0에 가까워 성능 향상이 거의 없었다. CNN은 평균 손실이 5.2에서 시작해 Epoch가 진행되면서 급격히 감소하는 경향을 보였다. LeNet은 Epoch 10에서는 약 0.5의 평균 손실을 기록하였다.

## 2.3 장면 분류 모델별 테스트

### 2.3.1 장면 분류 평균 정확도 테스트

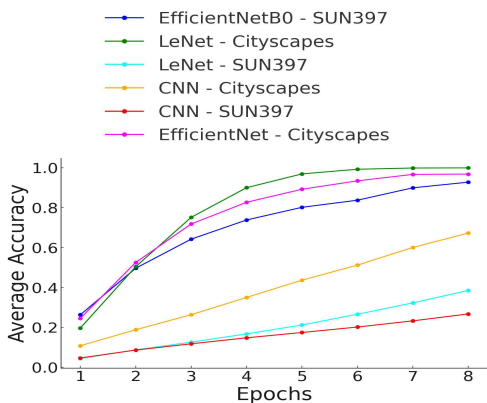


그림 3. 장면 분류 평균 정확도  
Fig. 3. Scene Classification Average Accuracy

본 실험에서는 Cityscapes와 SUN397 데이터 세트에서 CNN, EfficientNet, LeNet 모델의 장면 분류 평균 정확도, 장면 분류 평균 로스 변화를 Epoch 수에 따라 비교한 결과를 보여준다. <그림 3>에 따르면 Cityscapes 데이터 세트를 사용한 결과, EfficientNet은 Epoch가 진행될수록 빠르게 증가하여 4 Epoch에서는 약 0.8에 도달했다. 이후 지속해서 상승하여 8 Epoch 시점에서는 약 0.97의 정확도를 기록하였다. CNN은 초기 Epoch에서 약 0.1의 낮은 정확도로 시작했으며, 8 Epoch 시점에서 약 0.65의 정확도를 기록했다. LeNet은 8 Epoch 시점에서는 약 1.0의 높은 정확도를 기록하였다.

SUN397 데이터 세트를 사용한 결과, EfficientNet은 초기 Epoch에서 약 0.3의 정확도로 시작해, Epoch이 진행됨에 따라 꾸준히 상승하여 8 Epoch에서는 약 0.75의 정확도를 기록했다. CNN은 Epoch가 진행될수록 느리게 증가하여 8 Epoch 시점에서 약 0.4에 도달했다. LeNet은 8 Epoch 시점에서는 약 0.5의 정확도를 기록하며, EfficientNet보다 낮지만, CNN보다는 높은 성능을 보였다.

### 2.3.2 장면 분류 평균 로스 테스트

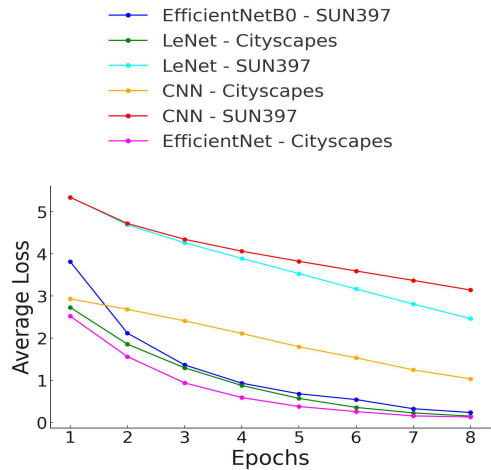


그림 4. 장면 분류 평균 로스  
Fig. 4. Scene Classification Average Loss

<그림 4>에 따르면, Cityscapes 데이터 세트를 사용한 결과 EfficientNet은 초기 Epoch에서 약 2.0의 손실 값을 기록하며 시작했다. Epoch가 진행됨에 따라 빠르게 손실이 감소하여 4 Epoch에서는 약 1.0



이하로 떨어졌고, 8 Epoch 시점에서는 약 0.5로 안정화되었다. CNN은 초기 손실 값이 약 3.5에서 시작했으며, Epoch가 진행될수록 꾸준히 감소하여 4 Epoch에서는 약 2.0, 8 Epoch에서는 약 1.5를 기록했다. LeNet은 초기 손실 값이 약 3.0에서 시작해, Epoch이 증가할수록 점진적으로 감소하여 4 Epoch에서는 약 2.0, 8 Epoch 시점에서는 약 1.0 수준을 기록했다.

SUN397 데이터 세트를 사용한 결과, EfficientNet은 초기 손실 값이 약 5.0으로 가장 높은 수준에서 시작했다. 이후 빠르게 감소하여 8 Epoch에서는 약 1.0 이하로 감소하였다. CNN은 초기 손실 값이 약 5.0으로 시작하였고, 8 Epoch에서는 약 2.0 수준의 손실 값을 기록하였다. LeNet은 초기 손실 값이 약 5.5로 시작하였고, 8 Epoch에서는 약 2.5를 기록하며 가장 높은 손실 값을 유지하였다.

## 2.4 객체 탐지 모델별 테스트

### 2.4.1 객체 탐지 평균 정확도 테스트

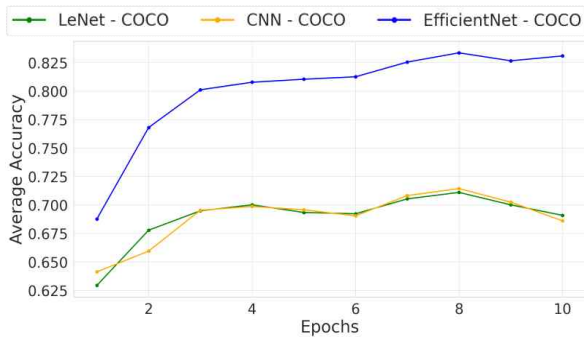


그림 5. 객체 탐지 평균 정확도  
Fig. 5. Object Detection Average Accuracy

본 실험에서는 COCO 데이터 세트에서 CNN, EfficientNet, LeNet 모델의 객체 탐지 평균 정확도, 객체 탐지 평균 로스 변화를 Epoch 수에 따라 비교한 결과를 보여준다. <그림 5>에 따르면 EfficientNet은 Epoch가 증가함에 따라 빠르게 정확도가 향상되어 3 Epoch에서 약 0.81의 정확도를 달성했다. 이후로도 정확도는 완만하게 증가하여 10 Epoch에서는 약 0.83에 도달했다. LeNet은 Epoch가 증가함에 따라 정확도가 서서히 증가했으며, 4 Epoch 이후로는 비교적 안정적인 성능을 유지했다. 10 Epoch에서는

약 0.67에 도달했다. CNN은 4 Epoch 이후에는 LeNet과 비슷한 정확도 수준을 유지하였고, 10 Epoch에서는 약 0.66의 정확도에 도달했다.

### 2.4.2 객체 탐지 평균 로스 테스트

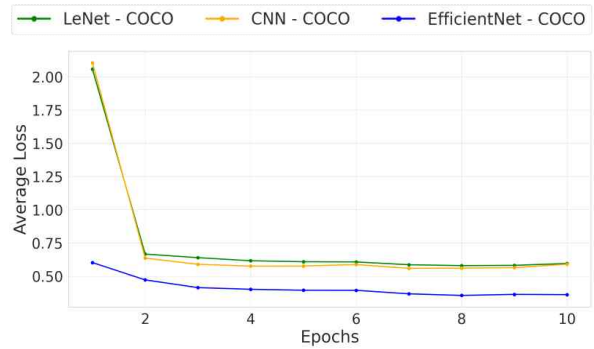


그림 6. 객체 탐지 평균 로스  
Fig. 6. Object Detection Average Loss

COCO 데이터 세트를 사용한 결과, EfficientNet은 초기부터 손실 값이 낮은 상태로 시작하여 Epoch가 증가함에 따라 손실 값이 빠르게 줄어들어, 2 Epoch 이후에는 약 0.5 근처에서 안정화되었다. LeNet은 초기 Epoch에서 손실 값이 급격히 감소하여 2 Epoch 이후 약 0.7 부근에서 안정화되는 경향을 보였다. CNN 역시 초기 Epoch에서 손실 값이 빠르게 감소하여 2 Epoch 이후 약 0.7 근처에서 안정화되었다. 이후 Epoch에서도 안정적인 손실 값을 유지하며 LeNet과 비슷한 수준의 성능을 보였다. 다만, EfficientNet에 비해서는 상대적으로 높은 손실 값을 보였다.

## 2.5 모델별 분석 결과

### 2.5.1 모델별 평균 정확도 증가율

<표 1>에 따르면 장면 분류, 객체 탐지, 이미지 분류에서 CNN, EfficientNet, LeNet 모델의 정확도와 그 차이를 비교하고 있다. EfficientNet은 세 작업 모두에서 가장 높은 평균 정확도를 기록했으며, 장면 분류에서 26.92%, 객체 탐지에서 23.58%, 이미지 분류에서 21.43%로 LeNet 대비 가장 큰 성능 향상을 나타냈다. LeNet은 모든 작업에서 가장 낮은 평균

정확도를 보였다. CNN 모델은 EfficientNet보다 낮은 정확도를 기록했지만, LeNet보다 높은 정확도를 보여주며 중간 수준의 성능을 나타냈다.

표 1. 모델별 평균 정확도 차이

Table 1. Differences in average accuracy by model

작업	모델	정확도 (%)	차이 (%)
장면 분류	CNN	68.0	4.62
장면 분류	EfficientNet	82.5	26.92
장면 분류	LeNet	65.0	0.0
객체 탐지	CNN	70.0	4.48
객체 탐지	EfficientNet	82.8	23.58
객체 탐지	LeNet	67.0	0.0
이미지 분류	CNN	72.0	2.86
이미지 분류	EfficientNet	85.0	21.43
이미지 분류	LeNet	70.0	0.0

### 2.5.2 모델별 테스트 시간 및 학습 시간 차이

표 2. 모델별 테스트 시간과 학습 시간

Table 2. Test time and learning time by model

작업	모델	학습 시간 (sec)	테스트 시간 (sec)
장면 분류	CNN	1200	300
장면 분류	EfficientNet	1500	400
장면 분류	LeNet	800	200
객체 탐지	CNN	1800	450
객체 탐지	EfficientNet	2200	600
객체 탐지	LeNet	1400	350
이미지 분류	CNN	1000	250
이미지 분류	EfficientNet	1300	300
이미지 분류	LeNet	700	150

표 3. 모델별 테스트 시간 및 학습 시간 차이

Table 3. Differences between test and training times by model

작업	모델	학습 시간 차이 (%)	테스트 시간 차이 (%)
장면 분류	CNN	50.0	50.0
장면 분류	EfficientNet	87.5	100.0
장면 분류	LeNet	0.0	0.0
객체 탐지	CNN	28.57	28.57
객체 탐지	EfficientNet	57.14	71.43
객체 탐지	LeNet	0.0	0.0
이미지 분류	CNN	42.86	66.67
이미지 분류	EfficientNet	85.71	100.0
이미지 분류	LeNet	0.0	0.0

<표 2>와 <표 3>에 따르면 장면 분류, 객체 탐지, 이미지 분류에서 CNN, EfficientNet, LeNet 모델의 훈련 시간과 테스트 시간을 비교한 것이다. EfficientNet은 가장 높은 훈련 및 테스트 시간을 기록했으며, 특히 장면 분류에서 훈련 시간이 87.5%, 테스트 시간이 100% 증가했다. CNN은 LeNet보다 긴 훈련 및 테스트 시간을 보였지만 EfficientNet보다는 짧은 시간을 기록했다. LeNet은 가장 빠른 훈련과 테스트 시간을 보였으며, 기준으로 사용되어 증가율이 0%로 표시되었다.

### 2.5.3 모델별 파라미터 수에 따른 메모리 사용량 차이

표 4. 모델별 파라미터 수에 따른 메모리 사용량

Table 4. Memory usage by number of parameters per model

작업	모델	파라미터 수 (백만 개)	메모리 사용량 (MB)
장면 분류	CNN	3.2	1500
장면 분류	EfficientNet	5.4	3000
장면 분류	LeNet	0.6	800
객체 탐지	CNN	4.0	2000
객체 탐지	EfficientNet	6.5	4000
객체 탐지	LeNet	0.8	1200
이미지 분류	CNN	3.0	1300
이미지 분류	EfficientNet	5.0	2500
이미지 분류	LeNet	0.5	700

표 5. 모델별 파라미터 수에 따른 메모리 사용량 차이

Table 5. Differences in memory usage by number of parameters by model

작업	모델	파라미터 증가율 (%)	메모리 사용량 증가율 (%)
장면 분류	CNN	433.33	87.5
장면 분류	EfficientNet	800.0	275.0
장면 분류	LeNet	0.0	0.0
객체 탐지	CNN	400.0	66.67
객체 탐지	EfficientNet	712.5	233.33
객체 탐지	LeNet	0.0	0.0
이미지 분류	CNN	500.0	85.71
이미지 분류	EfficientNet	900.0	257.14
이미지 분류	LeNet	0.0	0.0

<표 4>와 <표 5>에 따르면 장면 분류, 객체 탐지, 이미지 분류 작업에서 CNN, EfficientNet, LeNet 모델의 파라미터 수와 메모리 사용량을 비교한 것

이다. EfficientNet은 세 작업 모두에서 가장 많은 파라미터 수와 높은 메모리 사용량을 기록했으며, 장면 분류에서는 800%의 파라미터 증가율을 보였다. CNN은 EfficientNet보다는 적지만 LeNet보다는 많은 파라미터와 메모리 사용량을 보였으며, 이미지 분류 작업에서 500%의 파라미터 증가율을 기록했다. LeNet은 가장 적은 파라미터 수와 메모리 사용량을 보이며 기준 모델로 설정되어 증가율이 0%로 표시되었다.

## 2.6 딥러닝 경량화 관련 유사 연구

### 2.6.1 ACM Symposium on Neural Gaze Detection 연구

실제 ACM Symposium on Neural Gaze Detection 연구[15]에 따르면, 딥러닝 모델의 경량화 아키텍처와 모델압축에 중점을 두고 있다. 이 연구에서는 주로 자원이 제한된 환경에서도 높은 성능을 유지하면서 딥러닝 모델을 실행할 방법을 탐구한다. 예를 들어, MobileNet과 같은 경량화된 CNN 모델은 메모리 사용량과 계산 비용을 최소화하도록 설계되어, 중소규모의 개발 환경에서도 원활하게 실행될 수 있도록 최적화된 구조를 가진다는 결과를 보여주었다.

### 2.6.2 유사 연구와의 차이점

ACM Symposium on Neural Gaze Detection 연구[15]에서는 MobileNet과 같은 모델을 사용하여, 저전력 디바이스나 제한된 자원에서의 실행 결과에 대해서만 언급하였다. 반면에, 우리의 연구에서는 다양한 컴퓨팅 환경에서 적용할 수 있는 모델을 선택하는 데 집중하며, 모델 성능 외에도 파라미터 수, 메모리 사용량, 훈련 및 테스트 시간 등의 지표를 종합적으로 비교하였다. 따라서 정확도뿐만 아니라, 소규모 개발팀 혹은 개인 개발자들이 실제로 사용할 때 효율적인 모델을 선정할 수 있게 된다.

## III. 결 론

본 논문은 소규모 팀이나 개인 개발자가 제한된

자원에서 효과적으로 활용할 수 있는 딥러닝 모델의 선택을 돕기 위해 이미지 분류와 객체 탐지 작업에서 다양한 모델의 성능과 효율성을 비교하였다. 실험 결과, EfficientNet은 이미지 분류와 객체 탐지와 같은 복잡한 작업에서 가장 높은 정확도를 기록하며, 특히 성능이 중요한 작업에서 적합한 선택임을 보여주었다. 그러나 이 모델은 높은 정확도를 달성하는 대신 상대적으로 많은 파라미터와 메모리 사용량을 요구하여 자원 소모가 큰 것으로 나타났다. 반면에, CNN과 LeNet과 같은 모델은 EfficientNet보다 정확도는 낮았으나, 상대적으로 적은 메모리와 파라미터 수를 통해 경량화된 구조를 보여주었다. 이러한 모델들은 높은 성능보다는 자원 효율성을 우선시하는 작업에 적합하며, 특히 소규모 팀이나 개인 개발자가 제한된 자원으로 높은 성능을 달성해야 할 때 유용할 수 있다.

결론적으로, 중소규모의 개발 환경에서 사용하기 적합한 모델은 CNN이다. CNN은 EfficientNet보다 자원 소모가 적으면서도, LeNet보다 더 높은 정확도를 제공하여 고사양, 중간사양, 저사양 다양한 컴퓨팅 환경에서 안정적으로 우수한 성능을 발휘할 수 있다.

## 참 고 문 헌

- [1] Bong, Kang Ho, "Determinants of artificial intelligence adoption in firms: Evidence from Korean firm-level data", *Journal of Computer Information Systems*, in-<https://doi.org/10.22693/NIAIP.2024.31.3.034>, vol. 31, pp. 34-47, September 2024.
- [2] Chung In Kwan, "Exploratory Research on Awareness, Use, and Future Intention for A.I. Services among Koreans", *Journal of Computer Information Systems*, in-DOI : 10.46349/kjhss.2024.09.48.3.133, pp. 133-158, September 2024.
- [3] Oh Kyu-seol, "Strategies for an Autonomous Literacy Network", *Journal of Computer Information Systems*,

- in-<https://doi.org/10.22557/HG.2024.9.85.3.717>, pp. 717-751, September 2024.
- [4] Ji-Yeon Lee, "Exploring Art Educational Use of Image Generation AI as an Art Making To". ISSN (Print) 1229-747X Journal of Research in ArtEducation <http://dx.doi.org/10.20977/kkosea.2024.25.3.65> 2024 Vol. 25-3 pp. 65 ~ 79, July 2024
- [5] Changhyun Park, "KISTEP 10 Emerging Technologies in the Era of Generative Artificial Intelligenc", pp. 1-43, February 2024.
- [6] Youngkak Chol, Hyun Ju Kim, "Generative AI Art on the View of the Philosophy of Technology", Journal of Digital Contents Society Vol. 25, No. 9, pp. 2525-2541, Sep. 2024
- [7] Jeong-Hoon Jang, Jang-Mook Kang, Yong-Joon Lee, "A Study on Moral Machine and Laws in the Advancement of Artificial Intelligence", pJournal of the Korea Academia-Industrial cooperation Society 2024, <https://doi.org/10.5762/KAIS.2024.25.9.360> ISSN 1975-4701 / eISSN 2288-468 Vol. 25, No. 9 pp. 360-365, September 2024
- [8] Park Yejong, Lee Seulgi, "An Analysis of Perceptions Toward Artificial Intelligence: Focusing on AI Regulation, Technology, and Digital Literacy", Journal of Computer Information Systems, in- DOI <http://dx.doi.org/10.21484/kros.2024.21.2.161>, pp. 161-200, August 2024.
- [9] A. Krizhevsky, "Learning multiple layers of features from tiny images", Technical Report, University of Toronto, 2009.
- [10] G. Griffin, A. Holub, and P. Perona, "The Caltech-256 object category dataset", Technical Report, California Institute of Technology, 2007.
- [11] COCO Consortium, "COCO: Common Objects in Context". [Online]. Available: <https://cocodataset.org/>
- [12] Cityscapes Dataset, "Cityscapes dataset for semantic urban scene understanding". [Online]. Available: <https://www.cityscapes-dataset.com/>
- [13] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba, "SUN database: Large-scale scene recognition from abbey to zoo", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3485-3492, 2010. <https://doi.org/10.1109/CVPR.2010.5539970>.
- [14] TensorFlow Datasets, "SUN397 dataset". [Online]. Available: <https://www.tensorflow.org/datasets>.
- [15] Ching-Hao Wang, Kang-Yang Huang, Yi Yao, Jun-Cheng Chen, Hong-Han Shuai, Wen-Huang Cheng, "Lightweight Deep Learning: An Overview" ,IEEE Consumer Electronics Magazine Volume: 13, Issue: 4, pp. 51-64, July 2024

# 유니티에서 SortedSet 다익스트라 알고리즘과 DFS를 이용한 탐색시간의 비교분석

한승주\*, 이승민\*\*, 정영훈\*\*\*,이부형\*\*\*\*

## Comparison of Search Time Between SortedSet Dykstra Algorithm and DFS in Unity

SeungJu Han\*, SeungMin Lee\*, YoungHun Jeong\*, Boo-Hyung Lee\*\*

### 요 약

게임 개발 프로그램인 유니티에서 지도나 네이게이션 알고리즘을 만들 때 기본적으로 하나의 프로젝트에서 독립적인 오브젝트들이 병렬적으로 처리되는데 이에 따라 하나의 오브젝트에서 너무 많은 공간복잡도와 시간 복잡도를 가지게 되어 프로젝트의 복잡도를 높이고 많은 메모리 누수가 발생한다. 이와 같은 문제점을 해결하기 위해 BFS(Breadth First Search)개념인 우선순위 큐를 사용한 다익스트라 알고리즘을 사용하는데 .NET 낮은 버전의 유니티에서는 사용하기 어렵다. 본 논문에서는 우선순위큐를 사용한 다익스트라 알고리즘 대신 SortedSet이라는 자료구조를 사용한 다익스트라 알고리즘을 이용하여 위의 문제를 해결하고 DFS(Depth First Search:깊이우선탐색)알고리즘과 탐색 성능을 비교하였다. 구현결과 SortedSet 자료구조를 사용한 다익스트라 알고리즘이 DFS알고리즘에 비해 100배 이상의 탐색시간 향상을 보였으며 낮은버전의 유니티에서도 충분히 사용가능함을 보였다.

### Key words

BFS(Breadth First Search), Unity, Dijkstra, DFS(Depth First Search) SortedSet

### I. 서 론

오늘날에 게임 개발과 메타버스에서 자주 사용되는 게임 개발 프로그램인 유니티는 기본적으로 하나의 프로젝트에서 독립적인 오브젝트들이 병렬적으로 처리된다. 그렇기 때문에 하나의 오브젝트에서 너무 많은 공간복잡도와 시간복잡도를 가지는 것은 많은 오브젝트들이 병렬적으로 처리되기 때문에 결과적으로 프로젝트의 복잡도를 매우 높이게 되고, 이는 많은 메모리 누수를 가져온다[1][2]. 특히 이러한 문제는 탐색이나 정렬 알고리즘에서 두드러지게

나타난다.

그렇기 때문에 유니티 환경에서 지도나 네비게이션 알고리즘을 만들 때, 더 시간복잡도와 공간복잡도를 줄이기 위해서 BFS(Breadth First Search)개념인 우선순위 큐를 사용한 다익스트라 알고리즘을 사용하는데 낮은 .NET버전의 유니티에서는 추가된 자료구조나 함수들을 사용하는 것이 불가능하다[3]. 이와 같은 문제점을 해결하기 위해 본 논문에서는 우선순위큐를 사용한 다익스트라 알고리즘 대신 SortedSet 자료구조(이후 SortedSet 다익스트라)를 사용한 후 DFS알고리즘과 비교한다.

\*공주대학교 소프트웨어학과, hansengjuu@naver.com, \*\*공주대학교 소프트웨어학과교수(교신저자)bhl1998@kongju.ac.kr

## II. SortedSet 다익스트라의 구현

### 2.1. 초기화

SortedSet 다익스트라를 구현하기 전에 먼저 각 노드에 대해서 저장할 데이터를 정의한다. 단방향으로 연결된 노드의 번호를 저장하고, 경로 추적을 하기 위해서 이전 노드와 지금까지의 최소 코스트를 저장해 주었다. 정의한 데이터는 그림 1과 같다.

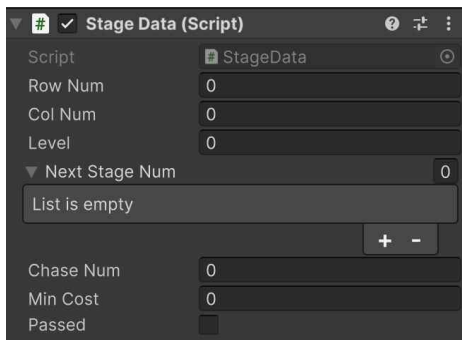


그림 1. 노드 데이터

그림 1에서의 데이터는 행에 위 데이터를 정해진 수만큼 추가하고 각 행의 데이터에서 다음 행의 데이터로 넘어가도록 구현하였다. 각 행을 관리해 주기 위해서 List를 이용해서 데이터를 관리해준다. 또한 행데이터를 모아놓은 List를 만들어 행데이터를 관리한다. 먼저 각 노드에 가중치를 무작위로 나누어 주기 위해서 Fisher-Yates Shuffle를 이용하는데 이는 유니티 내부의 Random을 이용해서 구현 해 주었다. 그림 2는 구현된 노드를 시각적으로 나타낸 것이다.

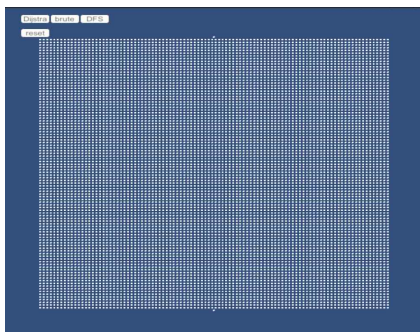


그림 2. 시각화된 구현 노드

### 2.2. SortedSet 다익스트라 구현

BFS 개념인 우선순위큐를 이용한 다익스트라는 유니티가 사용하는 .NET에는 라이브러리에서 우선순위큐를 지원하지 않기 때문에 낮은 버전의 유니티에서 BFS개념을 이용하여 탐색시간을 줄이기 위해서는 SortedSet을 이용해서 쉽게 구현할 수 있다.

SortedSet은 이름 그대로 정해진 비교자를 이용해서 자동정렬되는 Set이다. 비교자를 노드의 지금까지의 최소코스트 mincost 변수를 기준으로 오름차순으로 정렬한다. 노드 탐색 횟수를 줄여주기 위해서 조건은 다음 mincost가 현재노드의 mincost + 다음노드의 level보다 클 때 탐색하도록 제한해주고 탐색 시에 mincost를 갱신해서 횟수를 줄여주었다.

## III. 실험

실험에서는 SortedSet다익스트라와 DFS의 탐색시간을 비교하였다. 실험은 그림 2의 노드를 사용하였는데 100줄의 행에 각 행마다 100개의 노드를 부여하였고 모든 노드를 모르는 상태 확인하지 않고 DFS와 SortedSet 다익스트라를 각각 100회 시행한 후 결과를 평균하였다. 두 알고리즘 모두 다음의 mincost의 크기가 현재 계산보다 작을 때만 수행하도록 조건을 한정했다. 실험결과 Sorted 다익스트라의 탐색시간은 9ms, DFS의 탐색시간은 1363ms이었다.

표 1에서와 같이 제안된 SortedSet 다익스트라가 DFS탐색에 비해서 100배이상의 빠른 수행 시간을 나타냄을 알 수 있다.

표 1. 탐색시간 비교

알고리즘	탐색시간(ms)
SortedSet 다익스트라	9
DFS	1363

## IV. 결론

본 논문에서는 SortedSet다익스트라를 구현한 후와 DFS 탐색알고리즘과의 탐색 시간 차이를 비교

하였다. SortedSet다익스트라는 노드데이터를 먼저 정의한 후 각 노드에 가중치를 무작위로 부여하여 노드를 구현하였다. 결과적으로 본 논문에서 제안한

SortedSet 다익스트라가 DFS탐색에 비해서 100배 이상의 빠른 수행 시간을 나타냄을 알 수 있다. 제안된 알고리즘은 모든 노드를 무작위로 만든 상황에서 노드들을 이용해서 지도를 만들거나 만약 게임이라면 맵을 구성하는데에 응용할 수 있을 것으로 본다.

### 참 고 문 헌

- [1] 안성진, 조별희, 박규철, 이부형, “다양한 지형을 가진 다각형 랜덤 자연맵 생성 방법” 2022한국정보기술학회 추계학술대회, pp.677-679, 2022
- [2] 이용후, 김상운 “내비게이션 경로설정에서 최단 거리경로 탐색을 위한 A\* 와 Dijkstra 알고리즘의 하이브리드 검색법” 전자공학회논문지, 2014, 51.10: 109-117.
- [3] 이종원, 정종필, 「게임제작으로 배우는 유니티」, 한빛미디어

# FSM을 이용한 NPC와 플레이어 행동(상태) 추천 알고리즘

김기웅\*, 강상현\*, 최진영\*, 손우성\*\*, 이부형\*\*\*

## Behavior(State) Recommendation Algorithm of NPC and Player Using FSM

Kim Gi Woong\*, Kang Sang Hyeon\*, Choi Jin Young\*, Son Woo Seong\*,  
and Boo-Hyung Lee\*\*\*

### 요 약

플레이어들은 게임을 즐길 때 게임 내에 등장하는 NPC(Non Player Character)와 다양한 상호작용을 하게 되는데 NPC의 행동에 따라 플레이어의 행동이 결정되기 때문에 많은 연구에서 NPC의 상태를 다양하게 변화 시키기 위한 방법을 제시하고 있다. 본 논문에서는 기존 연구들이 NPC의 행동 변화를 유연하게 제어하는 데 초점을 맞춘 것과 달리, 플레이어 및 NPC의 상호작용 데이터를 실시간으로 반영하여 맞춤형 추천을 제공하는 알고리즘을 제안한다. 이를 위해 게임 환경에서 많이 사용되는 FSM 알고리즘을 기반으로 NPC와 플레이어 행동 또는 상태를 추천한다. 게임 플레이 시 NPC의 행동에 대응하는 플레이어의 대응을 바탕으로 베이지안 정리에 따른 추천 값을 조정하여 사용자에게 가장 높은 값을 지닌 결과를 제공하여 안정된 상태또는 행동이 선택 될 수 있도록 하며 게임제작을 통해 NPC와 플레이어사이의 많은 상호작용 후에 안정된 행동(상태)이 선택됨을 보였다. ,

### Key words

NPC(Non-Player Character), Player, FSM(Finite State Machine), Bayesian Optimization(베이지안 최적화)

### I. 서 론

게임에서는 유저가 제어하는 플레이어 캐릭터가 있고, 컴퓨터가 제어하는 NPC (Non-Player Character)가 있다. 이중 NPC는 유저가 게임을 진행하는데 도움을 주거나 유저 캐릭터와 맞서 전투를 진행하는 등의 역할이 주어진다[1].

플레이어들은 게임을 즐길 때 게임 내에 등장하는 NPC(Non Player Character)와 다양한 상호작용을 하게 되는데 NPC의 행동에 따라 플레이어의 행

동이 결정되기 때문에 많은 연구에서 NPC의 상태를 다양하게 변화시키기 위한 방법을 제시하고 있다[2]. 그러나 이와 같은 방법은 플레이어의 행동 데이터를 바탕으로 최적의 대응을 추천하기보다는 NPC의 상태 변화 자체에만 집중하는 경향이 있다.

본 논문에서는 기존 연구들이 NPC의 행동 변화를 유연하게 제어하는 데 초점을 맞춘 것과 달리, 플레이어 및 NPC의 상호작용 데이터를 실시간으로 반영하여 맞춤형 추천을 제공하는 알고리즘을 제안한다.

\*공주대학교 소프트웨어학과kimgiwoong127@smail.kongju.ac.kr ,\*\*공주대학교 컴퓨터공학과,

\*\*\*공주대학교 소프트웨어학과교수(교신저자)bhl1998@kongju.ac.kr



## II. FSM을 이용한 행동 추천 시스템

그림 1은 FSM을 이용한 행동추천 시스템을 나타낸다. 그림 1에서 NPC는 FSM과 데이터베이스 생성 모듈 등으로 구성되어 있다. FSM은 NPC의 상태를 제어하고 상태의 변화를 구현한다. FSM을 통해 NPC에서 상태를 얻은 뒤 플레이어와 상호작용을 하고, 결과를 NPC의 데이터베이스 생성 모듈을 거쳐 데이터베이스에 갱신한다. 갱신된 정보는 추천 알고리즘을 거쳐 새로운 상태를 추천하고 기존 추천 정보를 삭제 및 수정한다. 이와 동시에 데이터베이스에서 추천 정보를 플레이어에게 전달한다.

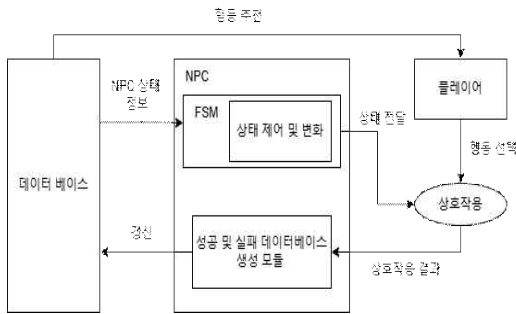


그림 1. FSM을 활용한 추천 시스템

### 2.1 데이터베이스 형식

FSM을 위한 요소는 NPC의 행동 데이터베이스에 저장하며 게임에서 발생하는 상태를 S, 각 상태를 선택한 횟수를 List, 상태의 성공/실패 횟수를 각각 W와 L, 해당 상태를 선택했을 때의 효율 값 E를 포함한다.

표 1. 데이터베이스 저장 형태

S	List	W	L	E
게임 내 NPC 및 플레이어가 선택하는 상태	S 내 전이가 일어난 횟수	상호 작용 이후 유효한 경우의 횟수	상호 작용 이후 불리한 경우의 횟수	S에 따른 상태별 효율 값

$$E = \frac{(W-L)}{U}, \quad U: List \text{ 전체}$$

### 2.2 추천 알고리즘

본 논문에서 사용되는 추천 알고리즘을 구현하기 위해서 베이저 정리를 활용한다. 베이저 정리는 다음과 같이 사전 확률을 통하여 사후 확률을 구하는 방법이다.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

본 연구에서는 표 1내 데이터베이스에 저장된 요소들을 이용한 후 사후 확률을 다음과 같이 계산한다.

$$P(W) = \frac{\left(\frac{W}{(W+L)}\right) * \frac{W}{List(W)}}{List(W) + List(L)}$$

위 공식은 전체 정보가 저장된 List 값과 W, L에 저장된 값을 통해 해당 S로 진입할 시에 성공할 확률을 도출하며, 데이터베이스에 확률을 저장한다.

## III. 실험 및 분석

### 3.1 실험환경 구축

NPC의 상태 데이터를 기반으로 플레이어에게 효과적인 행동을 제시하는지 실험한다. 이를 위해 먼저 그림 2와 같은 게임환경, NPC와 플레이어를 제작하였으며 실험용 게임은 플레이어(왼쪽)와 NPC(오른쪽)가 각자 자신의 속도에 따라 공격과 방어를 주고받는 턴제 전략 게임으로 Unity 환경에서 제작하였다.

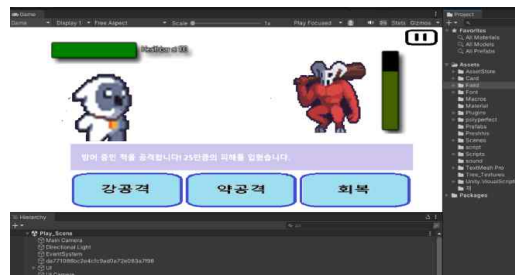


그림 2. 실험 게임 화면

본 실험은 2가지 형태로 진행한다. 먼저 FSM 알고리즘이 NPC에 정상적으로 적용되는지 확인하기 위해 플레이어의 행동 변화 수와 NPC의 상태 변화를 비교하는 실험을 진행한다. 다음 실험은 FSM이 정상적으로 작동하는 경우 NPC의 변화량에 영향을 주는 요소들을 기록하며 효율 값과 비교하여 플레이어에게 성공률이 가장 높은 행동을 추천하는 것을 확인하는 실험을 진행한다. 표 2는 게임 실행 중 플레이어가 선택할 수 있는 행동을 각각 정리한 표이다.

표 2. 게임 내 사용하는 기술

종류	기술명	데미지	효과
공격	일반 공격	15	특수 효과 없음
	강 공격	30	일정 확률로 공격
수비	방어	-	입는 데미지를 경감.
	회피	-	일정 확률로 데미지를 무효
	회복	-	체력 회복 (최대 HP를 초과하지 않음)

성공과 실패를 판정하는 결과는 현재 체력의 변화를 공격 턴이 끝날 때마다 확인하여 효율 값을 구한다. 상호작용 이후 체력(hitpoint:HP)의 감소가 10%에서 W의 값이 증가하고 반대라면 L의 값이 증가하도록 설정했다.

$$W : (HP \times 0.1) \geq damage$$

$$L : (HP \times 0.1) < damage$$

### 3.2 실험 결과

그림 3은 플레이어 행동변화와 NPC의 상태변화를 비교분석한 것으로 x축은 NPC와 플레이어와 상호작용 횟수를 나타내고 y축은 행동추천 확률값을 나타낸다. NPC와 플레이어의 상호작용 횟수가 많아질수록 NPC의 행동 추천 확률값의 변화율이 적은 반면 플레이어의 추천하는 행동의 상태는 변화율이 증가함을 볼 수 있다. 그림 4는 NPC상태변화에 따른 행동추천 확률을 비교한 그림으로 NPC의 행동 변화가 증가할수록 추천 확률의 변화량이 점점 줄어드는 결과를 보인다.

## V. 결론

본 연구는 FSM(유한상태기계)과 베イズ 정리를 결합하여, NPC와의 상호작용에서 사용자에게 최적의 대응을 추천하는 시스템을 제안하였다. 이는 플레이어의 행동 데이터를 실시간으로 반영하여 능동적인 추천을 제공한다는 점에서 기존 방식과 차별화된다.

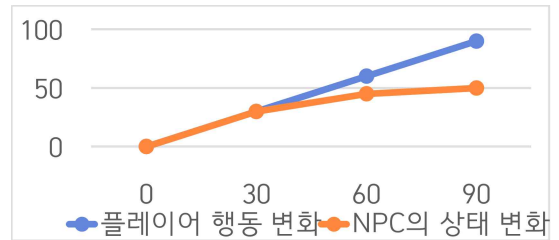


그림 3. 플레이어/NPC의 상호작용에 따른 행동추천 확률값의 비교

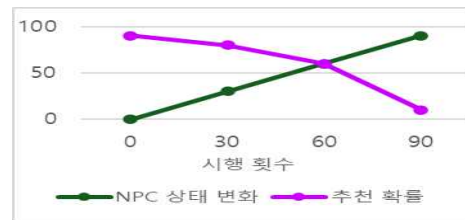


그림 4. NPC의 상태변화에 따른 행동추천 확률 변화량

기존 연구들이 NPC의 고정적 행동 패턴에 집중한 것과 달리, 본 연구는 실험을 통해 FSM을 활용하여 상태를 관리하고, 데이터에 기반한 효율적 상태 전이를 구현하였다. 이를 통해 FSM 전이 과정에서 상호작용 데이터를 기반으로 효율적인 상태를 실시간으로 반영하며, NPC의 행동 패턴을 정교하게 제어할 수 있다.

## 참고 문헌

- [1] Ho Suk Lee, "On-line Bayesian Learning based on Wireless Sensor Network", 한국정보과학회 학술 발표논문집, pp. 105-108, June 2007
- [2] 임경원, 이부형 "Raycast와 베지어곡선을 이용한 NPC의 부드러운 이동경로 탐색 방법", 한국정보기술학회 하계 종합학술대회 논문집, pp. 69-73, 2023

# 3D 프린팅에서 PLA+ 재료의 출력 조건에 따른 인장강도와 전단강도에 관한 연구

김호준\*, 전진표\*, 정진\*, 이치원\*, 나두현\*\*\*, 지승한\*\*

## Study of Tensile Strength and Shear Strength According to Printing Conditions of PLA+ Material in 3D Printing

HoJun Kim\*, JinPyo Jeon\*, Jin Jung\*, ChiWon Lee\*, DuHyeon Na\*\*\*, and SeungHan Ji\*\*

### 요약

FDM방식의 3D 프린팅은 사용자 친화적인 출력 방식으로 접근성이 좋아 다양한 산업 분야에서 활용되고 있다. 하지만 출력에 사용되는 플라스틱 재료의 특성상 철재 재료에 비해 강도가 낮다는 단점이 있다. 본 연구에서는 3D 프린팅 출력물의 인장강도와 전단강도 향상을 위해 내부 채움 패턴, 내부 채움 밀도를 변수로 하여 인장시험, 전단시험을 진행하였다. 인장강도는 동심원 패턴이 높았지만 전단강도는 지그재그 패턴이 높은 결과를 나타내었고 내부 채움 밀도가 증가할수록 인장강도와 전단강도가 선형적으로 증가함을 보였다.

### Abstract

FDM-based 3D printing is a user-friendly printing method with high accessibility, making it widely applicable across various industries. However, plastic materials used in printing generally offer lower strength than metal materials. In this study, tensile and shear tests were conducted with infill patterns and densities as variables to improve the tensile and shear strength of 3D-printed outputs. The tensile strength was higher with the concentric pattern, while the shear strength was higher with the zigzag pattern. Furthermore, both tensile and shear strength increased linearly as the infill density increased.

### Key words

3D printing, FDM, tensile strength, shear strength, PLA+, infill pattern, infill density

## 1. 서론

FDM(Fused Deposition Modeling) 방식의 3D 프린팅은 가열된 플라스틱 재료를 노즐로 압출하여 원하는 위치에 적층하며 제품을 제조한다. 기존 절삭

가공보다 재료의 낭비가 적고 제품 생산이 비교적 쉬워 누구나 쉽게 사용할 수 있다는 장점이 있다. 하지만 플라스틱 재료의 낮은 강도로 인해 복합적인 응력이 작용하는 기계 부품에 적용하기 어렵다는 단점이 있다. 따라서 플라스틱 재료를 사용한

\* 대림대학교 메카트로닉스공학과, khj1025119@naver.com

\*\* 대림대학교 메카트로닉스공학과, jsh2012@daelim.ac.kr

\*\*\* 대림대학교 기계공학과 부교수, dhna@daelim.ac.kr

3D 프린팅 출력물의 강도 향상을 위한 연구가 지속적으로 수행되고 있다.

나두현 등[1][2]은 출력 옵션 중 내부 채움 패턴, 노즐 온도, 베드 온도와 출력 속도를 변수로 하여 각 출력 옵션에 따른 인장강도에 대해 연구하였다. 저자는 노즐 온도, 베드 온도가 높을수록 혹은 출력 속도가 느릴수록 인장강도는 선형적으로 증가함을 확인하였다. 또한 소재, 채움 밀도, 레이어 방향, 채움 패턴을 변수로 하여 인장강도를 비교 분석하였다. PLA+소재가 ABS소재보다 인장강도와 연신율이 높았다. 연구자는 각 변수를 최적의 조건으로 출력했을 때 채움 밀도가 60%인 시편보다 인장강도를 2.07배 높일 수 있다고 언급하였다.

김기대[3]는 3D 프린팅 출력물의 기계적 물성치 (mechanical property)와 형상 정밀도(form accuracy) 개선을 위해 출력 방향, 적층 방향과 적층 두께를 변수로 한 인장, 충격시험을 진행하였다. 단면 층간 접합 강도가 수지 결합 강도보다 약하기 때문에 적층 방향이 단면 방향일 때 인장, 충격강도가 가장 낮은 것을 언급하였다. 또한 적층 두께가 얇을수록 주사 간격은 줄어들고 주사 횟수가 증가되어 인장 강도가 증가하였다.

Liviu Marşavina 등[4]은 시편 출력 방향, 3D 프린터 종류, 필라멘트 색을 변수로 한 인장, 파괴인성 시험을 진행하였다. 시편의 출력 방향이 45°일 때 가장 낮은 인장강도와 파괴인성을 보였는데 이는 응력에 저항하는 레이어가 0, 90°의 시편보다 적기 때문이라고 언급하였다.

기존 연구자들은 FDM방식의 3D 프린팅 출력물의 인장, 파괴인성에 대한 연구는 수행하였지만 전단강도에 대한 연구는 수행하지 않았다. 전단응력은 부품에 미세 균열의 발생과 성장에 영향을 미친다. 미세 균열에 응력이 누적될 경우 피로 파괴로 이어질 수 있다. 3D 프린팅 출력물이 다양한 산업 분야에 사용되기 위해서는 인장강도와 전단강도 향상에 대한 연구가 필수적이다. 따라서 본 연구에서는 내부 채움 패턴(infill pattern)과 내부 채움 밀도(infill density)를 변수로 하여 인장, 전단시험을 진행하였다. 또한 인장, 전단시험을 진행하였을 때 인장강도(tensile strength)와 전단강도(shear strength)가 어떤

차이가 있는지에 대한 비교, 분석도 진행하였다.

## II. 실험 방법

### 2.1 시편 제작

본 연구에서는 인장 및 전단시험을 위해 인장, 전단시편을 출력하였다. 인장시편은 ASTM D638 Type I 규격으로 제작하였다(그림 1(a)). 전단시편은 그림 1(b)과 같이 빨간색 구역에서 파단이 날 수 있도록 제작하였다. 제작한 전단시편을 인장시험기에 장착하여 전단시험을 진행하였다[5]. 출력 재료는 PLA+(Poly Lactic Acid-Plus)를 이용하였고 Cubicon사의 슬라이싱 프로그램 Cubicreator4 V4.4.0을 이용하여 내부 채움 패턴, 내부 채움 밀도와 같은 출력 조건을 변경하고 G코드를 생성하였다.

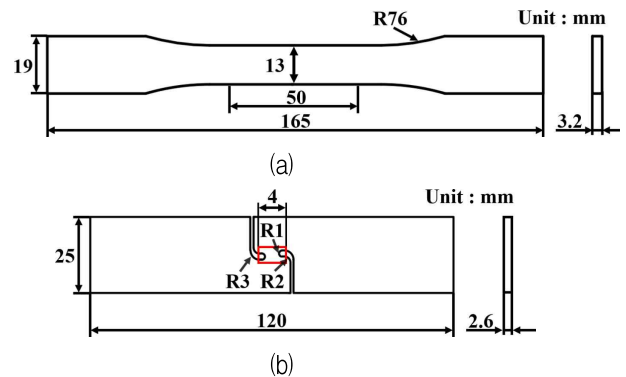


그림 1. 시험 시편

Fig. 1. Test specimen

(a) Tensile specimen(ASTM D638)

(b) Shear specimen

### 2.2 시험 장비 및 조건

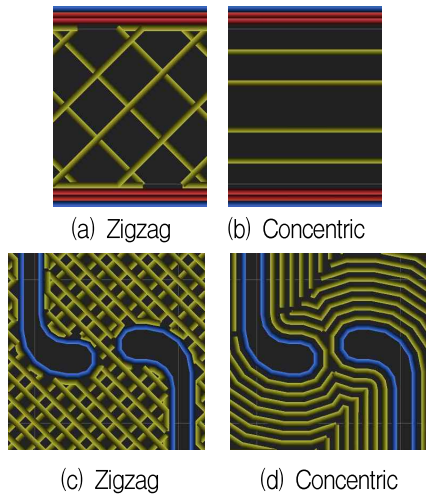
인장, 전단 시편 출력을 위해 Cubicon사의 Style NEO-A31C를 사용하여 시편을 출력하였다.(그림 2 (a)) 이 프린터는 FDM방식으로 출력물을 제작한다 [6]. 인장, 전단시험은 Testone사의 만능재료시험기 (Universal Testing Machine) TO-102를 이용하였다(그림 2(b)).

그림 3은 내부 채움 패턴에 따른 내부 형상을 나타낸다. (a)와 (b)는 인장시편, (c)와 (d)는 전단시편의 내부 형상이다. 본 연구에서는 내부 채움 패턴과



(a) (b)  
 그림 2. 3D 프린터와 시험 장비  
 Fig. 2. 3D printer and testing machine  
 (a) 3D printer(Style NEO-A31C)  
 (b) Tensile, shear testing machine(TO-102)

내부 채움 밀도에 변수를 두어 인장, 전단시험을 진행하였다. 내부 채움 패턴은 zigzag, concentric 패턴을 사용했다. 내부 채움 밀도는 20, 40, 60, 80, 100%로 변경하며 시험을 진행하였다. 전단시험에서는 내부 채움의 경향을 확인하기 위해 외벽의 개수를 1개로 하여 시험을 진행하였다. 인장시험과 전단시험은 각 변수 별로 2번씩 진행하였다.

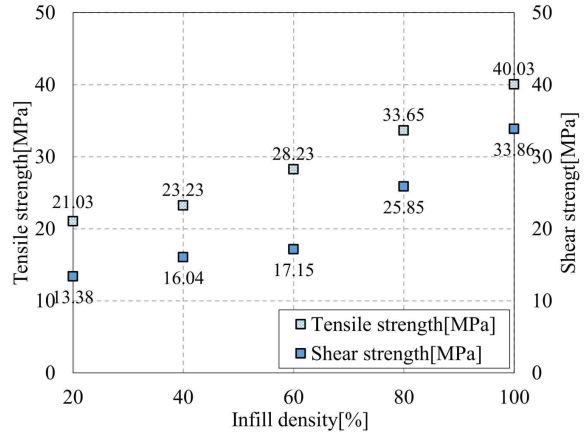


(a) Zigzag (b) Concentric  
 (c) Zigzag (d) Concentric  
 그림 3. 내부 채움 패턴에 따른 내부 형상  
 Fig. 3. Infill shape of infill patterns

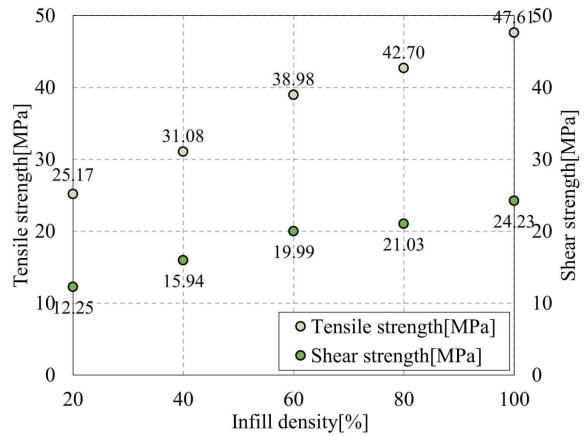
### III. 결과 및 분석

그림 4 (a)는 zigzag패턴의 인장, 전단시험 결과를 나타낸다. 내부 채움 밀도가 높아질수록 인장, 전단강도 또한 높아지는 경향을 보였다. Zigzag패턴의 전단강도도 내부 채움 밀도가 증가할수록 전단강도가 선형적으로 증가하는 경향을 보였다.

그림 4(b)는 concentric패턴의 인장, 전단시험 결과로 zigzag패턴과 마찬가지로 내부 채움 밀도가 증가하며 인장, 전단강도가 증가하였다.



(a)



(b)

그림 4. 인장, 전단시험 결과  
 Fig. 4. The results of tensile and shear test  
 (a) Zigzag pattern  
 (b) Concentric pattern

표 1은 내부 채움 밀도가 증가할수록 인장, 전단강도가 상승한 비율을 나타낸다. 가장 높은 강도 상승률은 zigzag패턴의 내부 채움 밀도가 60%에서 80%로 높아질 때 전단강도는 50.73% 상승하였다.

Zigzag와 concentric패턴 모두 전단강도가 인장강도에 비해 낮은 값을 나타냈다. 내부 채움 밀도 80%를 기준으로 zigzag패턴은 인장강도가 전단강도보다 30.17%(25.85 vs 33.65MPa) 높았고 concentric패턴은 103.00%(21.03 vs 42.70MPa)의 차이로 인장강도가 더 높은 값을 보였다. 그림 1(a)의 인장시험편은 표점거리 내에 내부 채움 레이어가 배치될 공간이

충분하여 더 많은 응력을 버틸 수 있지만 그림 1(b)의 전단시편은 표점거리 내에 내부 채움 레이어가 배치될 공간이 인장시편보다 상대적으로 작기 때문에 전단강도가 인장강도보다 낮게 나타난 것으로 판단된다.

또한 zigzag패턴의 경우 인장강도 평균값이 28.09MPa, concentric패턴은 37.11MPa로 concentric패턴이 32.11% 높았다. 전단강도 평균값은 zigzag패턴이 21.26MPa, concentric패턴은 18.69MPa으로 전단강도는 zigzag패턴이 13.75% 더 높은 것을 확인하였다. 그림 3(b)와 같이 concentric패턴은 벽을 따라 내부 채움 레이어가 생성되어 zigzag패턴보다 인장응력 방향에 대한 큰 저항을 나타내어 인장강도 값이 더 높은 것으로 판단된다. 전단강도의 경우 zigzag패턴이 더 높았는데 이는 그림 3(c), (d)와 같이 concentric패턴에서 전단응력이 가해지는 방향에 대해 저항하는 레이어가 zigzag패턴에 비해 적기 때문이라고 판단된다.

표 1. 내부 채움 밀도 변화에 따른 강도 상승률  
Table 1. Strength Improvement Rate with Increasing Infill Density

내부 채움 밀도 변화에 따른 강도 상승률 [%]				
내부 채움 밀도 변화	Zigzag패턴		Concentric패턴	
	인장	전단	인장	전단
20 → 40%	10.46	19.88	23.48	30.12
40 → 60%	21.52	6.92	25.42	25.41
60 → 80%	19.20	50.73	9.54	5.20
80 → 100%	18.96	30.99	11.50	15.22

#### IV. 결론

본 연구는 FDM방식의 3D 프린팅에서 내부 채움 패턴, 내부 채움 밀도에 따른 인장시험과 전단시험을 진행하였으며 결론은 다음과 같다.

(1) Zigzag패턴과 concentric패턴 모두 내부 채움 밀도가 올라갈수록 인장강도가 증가하는 경향을 나타냈다. 가장 높은 강도 상승률은 zigzag패턴의 내부 채움 밀도가 60%에서 80% 구간에서 전단강도가 50.73% 증가하였다.

(2) Zigzag, concentric에서 전단강도 또한 인장강

도와 동일하게 내부 채움 밀도가 증가할수록 전단강도가 선형적으로 증가하였으며 전단강도는 인장강도에 비해 낮은 결과 값을 보였다.

(3) 인장강도는 concentric패턴이 더 높았으며 전단강도는 zigzag패턴이 더 높았다. 이는 내부 채움 패턴에 따른 내부 형상 차이 때문인 것으로 판단된다. 따라서 응력이 가해지는 방향에 저항을 나타내는 내부 채움 패턴을 적절히 선택해야 한다.

#### 참 고 문 헌

- [1] D. H. Na and S. G. Kim, "A Study on Tensile Strength According to Various Output Conditions of PLA+ Materials Using 3D Printing", *Trans. Mater. Process.*, vol. 31, no. 2, pp. 89-95, March 2022.
- [2] D. H. Na and S. G. Kim, "A Study on Tensile Strength of PLA+ and ABS Materials by 3D Printing Output Conditions", *Trans. Mater. Process.*, vol. 30, no. 6, pp. 284-290, November 2021.
- [3] G. D. Kim, "Comparison of Mechanical Properties and Form Accuracy in FDM 3D Printing Based on Building Conditions", *J. Kor. Soc. Manufacturing Process Eng.*, vol. 20, no. 8, pp. 52-59, August 2021.
- [4] L. Marşavina, C. Vălean, M. Mărghitaş, E. Linul, N. Razavi, F. Berto and R. Brighenti, "Effect of the manufacturing parameters on the tensile and fracture properties of FDM 3D-printed PLA specimens", *Eng. Fract. Mech.*, vol. 274, pp. 108766, August 2022.
- [5] F. Khameneh, A. Abedini and C. Butcher, "Lengthscale effects in optical strain measurement for fracture characterization in simple shear", *Int. J. Fract.*, vol. 232, pp. 153-180, October 2021
- [6] [http://www.3dcubicon.com/shop/item.php?it\\_id=1631775559](http://www.3dcubicon.com/shop/item.php?it_id=1631775559)

# 병렬처리 도구를 이용한 한국어 문법 검사 시스템 구조

전민주\*, 이윤종\*, 김하민\*, 노태윤\*, 박영호\*, 안진호\*\*

## Korean Grammar Check System Architecture Using Parallel Processing Tools

MinJu Jeon\*, YoonJong Lee\*, HaMin Kim\*, TaeYoon Roh\*, YoungHo Park\*, and JinHo Ahn\*\*

### 요 약

이 연구는 맞춤법 검사기의 용도가 일상 속 짧은 글에 국한되는 현황을 넘어, 과제, 자기소개서, 도서 출판 등 대용량 텍스트의 문법 검사를 위한 시스템을 설계하는 데 중점을 두고 있다. 대학생의 과제, 취업 준비생의 자기소개서, 그리고 일반 대중의 일상 대화 속에서는 짧은 글의 문법 검사가 자주 사용되지만, 출판사와 개인 작가에게는 도서 한 권 분량 이상의 긴 텍스트를 검토할 수 있는 시스템이 요구된다. 이를 위해 연구는 대용량 텍스트를 효과적으로 검사할 수 있는 맞춤법 검사기를 기획 및 초고 단계부터 편집자의 검토 단계 전까지 실질적 사용 사례에 맞춰 설계하고자 한다. 백엔드 시스템과 웹 기반 인터페이스를 함께 구축하여 편집자와 작가가 초고 검토, 피드백 반영, 교정 및 운문 작업을 더욱 효율적으로 수행할 수 있도록 지원한다. 이러한 접근 방식은 출판사 및 개인 작가를 위한 대용량 맞춤법 검사의 필요성을 충족하며, 다양한 사용자층이 포괄적으로 문법 검사를 활용할 수 있도록 돕는 것을 목표로 한다.

### Abstract

This study focuses on extending the functionality of spell-checking systems, which are currently limited to short texts, to accommodate large volumes of text, such as academic assignments, resumes, and book publishing. While grammar checkers are commonly used for short texts in everyday contexts, a system that can review long texts—such as entire books—is in demand by publishers and individual authors. To address this, the study aims to design a spell-checking system capable of handling large texts and covering all stages, from drafting and initial review to pre-editorial inspection. A backend system and web-based interface will be built, supporting editors and authors through preliminary review, feedback incorporation, editing, and final proofreading. This approach meets the needs of publishers and independent authors for comprehensive spell-checking of large volumes, creating a tool that a wide range of users can utilize effectively.

### Key words

parallel processing, grammar check, apach spark, mapreduce, redis

---

\* 경기대학교 AI컴퓨터공학부 컴퓨터공학과 {mingmingmon,zoniblack,nugs200012,iantyrho13,youngho}@kgu.ac.kr

\*\* 경기대학교 AI컴퓨터공학부 교수 jhahn@kyonggi.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음 (2021-0-01393).

## 1. 서 론

최근 개인 출판의 증가와 함께 대용량 텍스트의 품질을 관리하기 위한 맞춤법 검사기 활용이 급증하고 있다. 디지털 출판과 블로그, 독립 출판 활동이 활성화되면서 맞춤법 검사기는 개인 출판자와 작가들이 텍스트 품질을 효율적으로 관리하는 필수 도구로 자리 잡았다. 기존의 맞춤법 검사기는 주로 짧은 텍스트에 한정된 기능을 제공하여, 긴 원고를 검사하는 데 한계가 있었다. 이 때문에 개인 출판 활동의 활성화와 더불어, 한 권의 책에 이르는 대용량 텍스트도 빠르게 검사할 수 있는 시스템의 수요가 증가할 것이라고 보인다.

맞춤법 검사기에서 사용하는 기술로는 문장 내 문법 오류를 자동으로 분석하고 수정하는 알고리즘이 있다. 예를 들어, 텍스트를 형태소 단위로 분해해 분석하거나, 문법 규칙에 따라 오류를 탐지하는 방식이다. 한국어 맞춤법 검사기는 단어의 형태와 문맥에서의 올바른 용법을 확인하는 작업이 필수적이다. 이를 위해 음절이나 어미의 규칙에 따라 단어를 비교하고, 사전에 없는 단어를 탐지해 오류를 검사한다. 그러나 이러한 방식은 텍스트 길이에 제한을 받을 수밖에 없고, 대규모 텍스트를 한꺼번에 처리하기에는 처리 속도가 크게 저하되는 문제가 있다.

이와 같은 문제를 해결하기 위해 분산 및 병렬 처리 기술을 적용한 대용량 맞춤법 검사 시스템을 설계하고 있다. 분산 및 병렬 처리 기술을 도입하면, 긴 텍스트를 여러 노드에 나누어 병렬로 검사할 수 있기 때문에 더 많은 양의 텍스트에서 빠르게 오류를 탐지할 수 있다. 예를 들어, 한 권의 도서에 이르는 긴 원고를 300페이지 단위로 나누어 여러 서버에서 동시에 검사하고, 그 결과를 종합하는 방식을 적용할 수 있다. 이는 개인 출판자에게 실질적인 시간을 절감시켜주며, 작업 효율을 크게 향상시킨다.

## II. 대용량 맞춤법 검사기를 위한 기술 구현

### 2.1 분산병렬처리 설계

맞춤법 검사기 프로젝트에서는 Data-Centered 아키텍처와 Apache Spark[1] 기반의 MapReduce 논리, Vertical Distribution 시스템 아키텍처를 채택하였다. 이를 통해 대용량 텍스트 데이터를 효율적으로 처리하고, 병렬 작업의 성능을 극대화하며 데이터의 일관성을 보장하고자 한다. 특히, 본 프로젝트는 기존의 Hunspell 한국어 사전(hunspell-dict-ko) [2] 를 Spark 기반 아키텍처에 통합함으로써 개발 효율성을 높이고, 대규모 텍스트 데이터를 분산 환경에서 효과적으로 처리할 수 있다.

### 2.2 Hunspell과 Data-Centered 아키텍처의 활용

Hunspell 한국어 사전을 사용함으로써 맞춤법 검사에 필요한 주요 데이터(사전, 문법 규칙 등)는 사전 내에 내장된 자료에 의존하게 된다. Hunspell 한국어 사전은 자체적인 사전과 규칙을 포함하고 있으며, 이 데이터는 별도의 관리 없이 각 노드에서 참조하여 사용할 수 있다. 그러나 추가적인 예외 규칙, 수정 제안과 같은 데이터를 프로젝트의 특성에 맞게 Data-Centered 아키텍처를 통해 중앙 관리하여 Spark 노드에서 일관되게 사용할 수 있다.

### 2.3 Apache Spark와 MapReduce의 활용

Apache Spark와 MapReduce 기반의 병렬 처리 방식은 본 프로젝트의 핵심 성능 요소 중 하나다. Spark는 전통적인 Hadoop MapReduce 모델[3]과 달리, 메모리 기반의 고속 데이터 처리를 통해 더 빠르게 작업을 수행할 수 있다. 이를 통해 맞춤법 검사기 프로젝트에서는 MapReduce의 논리를 Spark의 RDD(Resilient Distributed Dataset) API[4]를 통해 구현하여, 각 텍스트 단위(문장, 문단 등)를 개별적으로 병렬 처리할 수 있게 된다. 구체적으로, MapReduce 논리는 다음과 같이 적용된다. map() 연산을 통해 각 문장을 독립적인 맞춤법 검사 단위로 나누고, 이를 분산 노드에서 병렬로 처리하여 각 텍스트 조각에 대해 개별 맞춤법 검사를 수행한다. 이후 reduce() 연산을 통해 각 맞춤법 검사 결과를 집계하고, 이를 바탕으로 최종적인 오류 보고서와 수정 제안을 사용자에게 제공한다.

Spark의 MapReduce 모델은 이와 같은 맞춤법 검



사 프로세스에서 데이터 병렬 처리를 극대화하며, 데이터 일관성 및 응답성을 크게 향상시킨다. Map 단계에서는 맞춤법 검사를 병렬로 수행하며, 각 문장의 검사 결과는 Reduce 단계를 통해 중앙에서 집계되어 최종 결과로 출력된다. Spark의 RDD는 이와 같은 MapReduce 작업을 효율적으로 관리하며, 메모리 내에서 데이터를 다루기 때문에 기존 MapReduce 모델보다 훨씬 빠르게 작동한다.

이해할 수 있다.

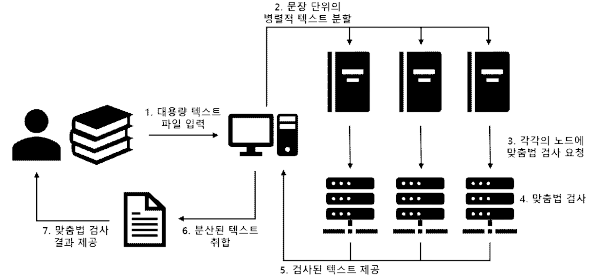


그림 1. 데이터 흐름 개요

## 2.4 Vertical Distribution 아키텍처의 활용

본 프로젝트에서는 Vertical Distribution 아키텍처를 적용하여 Data-Centered 아키텍처의 컴포넌트를 수직적으로 분산하는 것을 제안한다. 각 컴포넌트는 특정 기능을 전담하는 서버 또는 서비스로 독립적으로 배치되며, 기능별로 역할을 분담하여 작업을 효율적으로 수행한다. 예를 들어, 데이터 관리 컴포넌트는 중앙의 데이터베이스 서버에 배치되어 모든 노드에 동일한 사전 및 맞춤법 규칙을 제공한다. Spark를 활용한 맞춤법 검사 엔진은 개별 노드로 배치되어 텍스트 단위별 맞춤법 검사를 독립적으로 수행한다. 더 나아가 캐싱 및 수정 제안 컴포넌트는 Redis[5]와 같은 분산 캐시 시스템을 통해 자주 사용되는 검사 결과를 빠르게 제공하여 성능을 최적화할 수 있다.

이와 같이 기능별로 수직적으로 분산된 구조를 통해 특정 컴포넌트에 대한 요청이 증가할 때 해당 컴포넌트만 독립적으로 확장할 수 있으며, 시스템의 유연성과 확장성을 높일 수 있다. 예를 들어, 맞춤법 검사 요청이 급증할 경우 맞춤법 검사 엔진 노드만 추가하여 대응할 수 있어 전체 시스템의 효율성을 극대화할 수 있다.

## 2.5 데이터 흐름 개요

데이터 흐름 개요는 사용자가 맞춤법 검사 시스템을 통해 서비스를 요청하고 결과를 반환받는 과정에서의 상호작용을 개략적으로 나타낸다. 이 흐름도는 맞춤법 검사 요청, 시스템 응답, 검사 결과의 확인 등 사용자가 체감하는 주요 단계를 시각화하여, 사용자가 시스템을 어떻게 경험하고 활용하는지 직관적으로 보여준다. 이를 통해, 사용자 관점에서 바라본 맞춤법 검사 시스템의 서비스 흐름을 쉽게

## 2.6 데이터 처리 프로세스 흐름도

데이터 처리 프로세스 흐름도는 맞춤법 검사 요청이 시스템 내부에서 처리되는 과정을 상세히 나타내며, Apache Spark 클러스터, Data-Centered 아키텍처의 컴포넌트, MapReduce 연산 등의 주요 컴포넌트 간 데이터 흐름과 로직 처리를 시각화한다. 사용자의 요청이 맞춤법 검사 엔진으로 전달되면, Spark 클러스터가 맞춤법 검사 요청을 병렬 처리하며, Hunspell 한국어 사전(hunspell-dict-ko)을 참조해 텍스트의 맞춤법과 문법 오류를 검사한다. 필요한 경우 Data-Centered 아키텍처에서 사용자 정의 단어 나 예외 규칙 같은 추가 데이터를 조회한다.

각 텍스트 단위에 대한 검사 결과는 Spark의 MapReduce 연산을 통해 집계되며, 자주 사용되는 검사 결과와 수정 제안은 Redis 캐시에 저장되어 후속 요청 시 성능을 향상시킨다. 최종 맞춤법 검사 결과는 사용자에게 반환되어 맞춤법 검사 요청에 대한 응답으로 제공된다.

## III. 결 론

본 연구는 개인 출판자들이 대용량 텍스트를 효과적으로 관리할 수 있도록 돕는 맞춤법 검사 시스템의 필요성을 기반으로 한다. 분산 및 병렬 처리 기술을 활용해 기존의 짧은 텍스트 위주 검사 방식을 넘어서, 장편 원고의 전체를 동시에 분석하는 방법론을 제시하고자 한다. 이를 통해 출판자와 작가들은 보다 빠르고 효율적으로 맞춤법을 검사하여 고품질의 콘텐츠를 제작할 수 있으며, 나아가 개인 출판의 접근성과 품질을 높이는 데 기여할 수 있을 것이다.

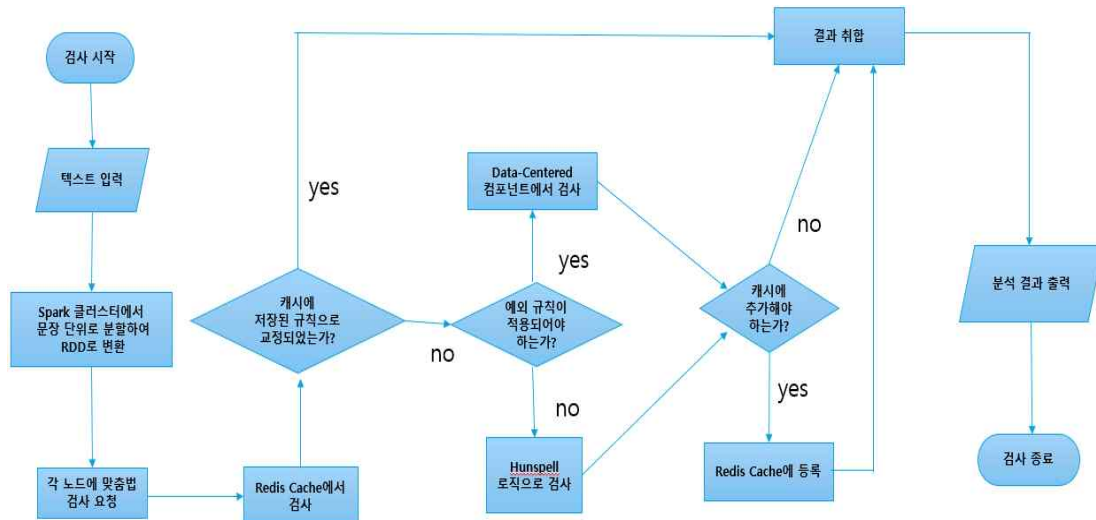


그림 2. 데이터 처리 프로세스 흐름도

## 참 고 문 헌

- [1] Seyoon Ko, & Joong-ho Won (2016). Processing large-scale data with Apache Spark. 응용통계연구, 29(6), 1077-1094.
- [2] hunspell-dict-ko. (n.d.). A Hunspell-based spell-checking dictionary for Korean, GitHub repository. Available: <https://github.com/spellcheck-ko/hunspell-dict-ko>
- [3] T. Tekdogan and A. Cakmak, "Benchmarking Apache Spark and Hadoop MapReduce on Big Data Classification," in 2021 5th International Conference on Cloud and Big Data Computing (ICCBDC), Liverpool, United Kingdom, Aug. 13-15, 2021, pp. 10. Available: <https://doi.org/10.1145/3481646.3481649>
- [4] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauly, M. J. Franklin, S. Shenker, and I. Stoica, "Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing," in Proc. 9th USENIX Symp. Networked Systems Design and Implementation (NSDI), 2012, pp. 15-28.
- [5] "Redis Documentation," Redis.io. [Online]. Available: <https://redis.io/>. Accessed: Oct. 30, 2024.

# 국내 산림 산불감시드론의 충전방식제안

정유영\*, 이서윤\*, 진정민\*, 김지원\*, 지승한\*\*

## Proposed Charging Methods for Forest Wildfire Monitoring Drones in South Korea

Jung youyoung\*, Lee seoyun\*, Jin jungmin\*, Kim jiwon\*, and Ji seung hwan\*\*

### 요 약

본 연구는 산불 감시 및 초기 진압에서 회전익 드론의 효율적 활용과 감시 지속 시간을 늘리기 위한 보조 충전 방법을 제안한다. 회전익 드론의 초기 화재 대응에 유리한 특성을 바탕으로, 국내 주요 산림의 지형과 드론 성능을 분석하여 태양광 충전, 정류장식 충전스테이션, 교체형 배터리, 송전탑 직접 충전 방식의 장단점을 평가하였다. 이를 통해 산불 예방 및 감시 시스템에서 드론의 지속 감시와 신속 대응을 위한 충전 방안의 필요성을 강조한다.

### Abstract

This study proposes auxiliary charging methods to enhance the efficiency and monitoring duration of rotary-wing drones for wildfire detection and initial suppression. Based on the advantages of rotary-wing drones in early fire response, we analyze the geographical features of major forests in South Korea and evaluate the strengths and weaknesses of solar charging, docking station charging, replaceable batteries, and direct transmission tower charging. This research underscores the need for reliable charging solutions to support continuous monitoring and rapid response in wildfire prevention and management systems.

### Key words

wildfire monitoring, solar charging, ground charging stations, battery swapping systems, power line charging

## I. 서 론

산불 재난 발생 시 드론 활용 방안에 관한 연구에 따르면, 산불 발생 시 효과적인 화재 감시와 진압을 위해 회전익 드론이 활용되는 것이 적합하다고 보고된다.[1]

이 연구의 목적은 회전익 드론의 배터리 사용 시간을 연장하기 위한 보조 충전 방식을 제안하여, 산

불과 같은 장시간의 감시가 필요한 상황에서 회전익 드론의 효과적인 활용을 가능하게 하는 것이다.

## II. 산불 감시 드론 지역별 활용

### 2.1 고정익 드론과 회전익 드론의 비교

회전익 드론은 고정익 드론에 비해 기동성과 정

\*대림대학교 메카트로닉스공학과 sjadw1004@naver.com

\*\*대림대학교 메카트로닉스공학과, jsh2012@daelim.ac.kr

밀한 감시 능력을 갖추고 있어 실시간으로 화재 현장에 대응할 수 있다.[1] 특히, 고정익 드론은 넓은 범위의 감시에는 적합하지만, 정지 비행이 어려워 초기 화재 진압에는 한계가 있다. 반면, 회전익 드론은 정밀 제어가 가능하여 초기 진압에 유리한 것으로 평가된다.

인공지능 및 드론을 이용한 화재 초기 진압 효율 향상에 관한 연구에서는 다양한 센서가 장착된 드론이 초기 화재 대응에 중요한 역할을 할 수 있음을 강조하고 있다. 특히, 다중 센서 기술이 적용된 회전날개회전익 드론은 화재 현장의 발화 원과 확산 경로를 정확하게 파악하여 AI 시스템에 데이터를 제공함으로써 신속한 대응이 가능하다.[2] 또한, 발전 과정은 회전날개회전익 드론에 적용할 수 있는 열화상 카메라, RGB 카메라, LIDAR 등 다양한 센서로 설명되며, 이러한 센서들이 초기 화재 대응에 중요한 역할을 한다는 점을 강조하고 있다.

고층 건물의 초기 화재 진압을 위한 소화탄 발사기와 분말 소화기를 장착한 소방 드론 개발 연구에서는[3] 회전익 드론의 경우 소화탄 발사 및 분말 분사 장비를 탑재하여 초기 화재 진압에 효과적으로 사용될 수 있음을 보여주고 있다.

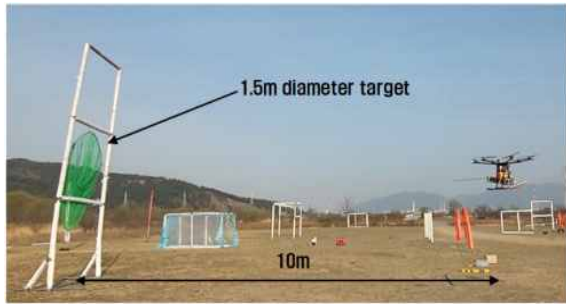


그림 1. 목표 지점에서의 사격 정확도 테스트 장면 (출처: 안성수, 2024, p390)

Fig. 1 The scene of the shooting accuracy test at the target point(source: An, Seong-su, 2024, p390)

그림 1에서 확인할 수 있듯이, 이 연구에서는 회전익 드론의 호버링 기능을 활용해 소화탄을 정확한 목표 지점에 투척하는 실험을 나타내었다. 실험 결과 80%의 명중률을 보였으며, 이는 고정익 드론에서는 불가능한 정밀한 화재 진압 능력을 보여주는 사례라 할 수 있다.[3]

## 2.2 국내의 산불 감시 드론 현황

최근 국내에서 산불 발생이 증가하면서 이에 대한 효과적인 감시와 대응 시스템의 필요성이 커지고 있다. 이러한 상황에서 산불 감시 드론은 예방과 조기 발견을 위한 중요한 도구로 자리 잡고 있다. 현재 한국에서 운영되고 있는 드론은 고해상도 카메라와 열화상 카메라를 장착하여, 실시간으로 산불의 징후를 감시하는 데 활용되고 있다.

특히 경상북도와 강원도에서는 드론을 통해 산불 예방을 위한 시스템을 구축하고 있다. 경상북도는 드론을 활용하여 3차원 모델링을 통해 산불 발생 위험 지역을 점검하고 있으며, 강원도는 드론을 이용한 감시와 더불어 항공 사진과 위성 이미지를 결합하여 보다 정확한 모니터링을 하고 있다. 이러한 다양한 방법으로 드론을 활용하는 사례는 계속해서 늘어나고 있다.

이렇듯, 국내에서의 드론 활용은 산불 감시에 있어 점차 중요한 역할을 하고 있으며, 지속적인 기술 발전과 인프라 구축이 이루어지고 있다.[4][5][6]

## 2.3 드론 기기별 성능 비교

그림 2 와 같이 드론의 카메라 탐지 거리와 관련한 연구 자료나 문헌을 통해[7][8][9] 일반적으로 사용되는 드론 감시 장비와 그 탐지 거리를 확인할 수 있다.

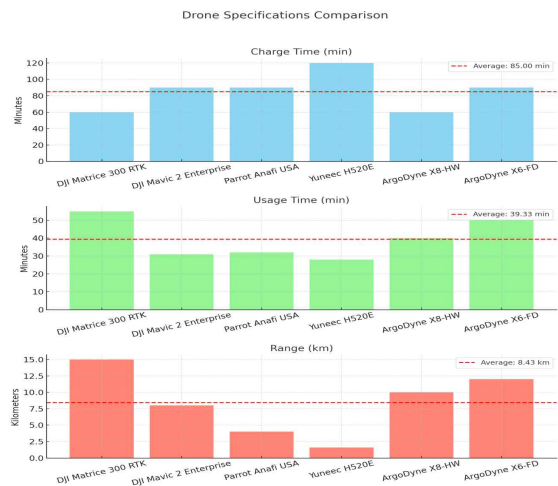


그림 2. 상용 드론별 성능과 평균

Fig. 2. Average Performance of Commercial Drones

산불 감시에 활용되는 드론 카메라인 FLIR Systems, Inc는 열화상 카메라를 제공하며, 산불 감시용 드론에 자주 사용되는 FLIR Vue 및 Duo 시리즈는 약 500m에서 1km까지 탐지가 가능하다고 알려져 있다. 이 카메라들은 열을 감지하여 화재와 같은 열원에서 발생하는 변화를 감지하는 데 효과적이다.[7]

DJI의 Zenmuse XT2는 FLIR 열화상 기술과 결합한 카메라로, 산불 감시와 같은 열 감지 작업에 사용된다. 이 카메라는 대략 700m에서 1km까지의 열원을 탐지할 수 있어 산불 감시용으로 적합하다.[8]

O. M. Bushnaq, A. Chaaban, T. Y. Al-Naffouri는 드론에 장착된 열화상 카메라의 탐지 거리가 500m에서 1km 사이일 때 산불 감시에 적합하다고 하였다.[9]

이 외에도 드론 제조사들의 기술 문서나 산불 감시 관련 논문에서 카메라의 탐지 거리 및 성능을 확인할 수 있다. 드론의 평균 사용 가능 범위가 9.33km일 때, 이를 기반으로 산림의 평균 면적 87.33km<sup>2</sup>에 대해 커버할 수 있는 비율을 계산하면 드론이 이동가능한 범위를 9.33 km 직선거리로만 생각하지 않고, 원형으로 생각할 때 커버할 수 있는 면적은 다음과 같다.

$$\text{면적} = \pi \times (9.33)^2 \approx 273.55 \text{Km}^2 \quad (1)$$

식 (1)은 드론이 한 번 충전으로 이론상 커버할 수 있는 범위이며 이를 산림의 평균 면적인 87.33km<sup>2</sup>와 비교하여 보면,

$$\text{커버 비율} = \left( \frac{87.33}{273.55} \right) \times 100 \approx 31.92\% \quad (2)$$

식 (2)의 결과로 드론의 평균 감시 범위는 산림 평균 면적의 약 31.92%를 커버할 수 있다.

이를 바탕으로 단일 충전 스테이션만으로 감시 드론이 산림 평균 면적을 완전히 커버하기에 어려움이 있다고 판단하였고, 드론의 사용 범위를 늘릴 방안을 모색하였다.

## 2.4 지역별 드론 보조 충전방식 제안

울진군은 전국적으로 일조량이 높아 태양광 충전에 유리한 환경을 제공하며, 산악지대에서 태양광 패널을 통한 지속적인 에너지 공급이 가능하다. 따라서 태양광 충전방식이 적합한데 태양광 충전 방식은 정찰 범위와 에너지 효율성에서 높은 평가를 받으며, 인프라 요구사항이 적어 경제적이다. 태양 드론 표면의 태양광 패널을 통해 생성된 전력을 배터리에 저장하거나 시스템에 직접 사용함으로써, 자율 충전과 연속 감시가 가능하다.

이우람은[10] 태양광 패널을 추가한 드론의 성능이 약 2~3% 정도 향상된다고 발표하였다. 이는 태양광 패널이 비행 중 추가 전력을 공급함으로써 배터리 소모를 약간 줄여 비행시간을 늘리는데 이바지했음을 의미한다. 높지 않은 증가 폭이라 볼 수 있지만, 연속적으로 감시를 이어가야 하는 감시 드론의 특성상, 장기적으로 보았을 때는 유의미한 차이를 보인다고 볼 수 있다. 이는 성능 향상의 관점에서는 2~3%이지만, 우리가 중점으로 제안하는 배터리 사용 시간 증대의 관점에서는 더욱 높은 효율을 가진다고 판단할 수 있다.

계룡산은 약 65.335km<sup>2</sup>의 넓은 면적을 지니고 있고, 복잡한 지형으로 인해 사람이 직접 감시하기 어려운 구역이 많아 드론을 활용한 감시의 필요성이 강조된다. 그러나 드론이 한 번의 충전으로 감시하기에는 면적이 너무 넓어 지속적인 감시를 위해 다양한 충전 방식과 경로 최적화가 요구된다. 정류장식 충전스테이션 방식은[11] 넓은 산림 지역 감시에 적합하기 때문에 이러한 지역에서는 드론이 정류장식 충전 스테이션을 통해 비행을 지속적으로 지원 받을 수 있어 넓은 구역을 안정적으로 감시하는 것이 이상적이다.



그림 3. 아르고스다인 Aquila-3F (출처: 아르고스다인)  
Fig. 3. Argosdyne Aquila-3F (source: Argosdyne, 2022)

그림 3의 아르고스다인의 Aquila-3F 드론은 비행 시간과 평균 속도 10m/s를 기준으로 30분 기동이 가능한데, 평균적인 작동 시속인 12m/s 속도로 30분간 비행한다고 하였을 때, 21.6km 비행이 가능하다. 이 드론은 20km 구간마다 설치된 충전스테이션에 착륙하여 배터리를 충전함으로써 연속적인 감시를 가능하게 한다.[11] 하지만 정류장식 충전스테이션의 문제점도 있는데, 강원도의 산악지대와 같이 험준한 산악지형일 경우 설치지역이 유지보수의 접근성이 낮아 경제성 또한 낮아진다는 문제점이 있다.

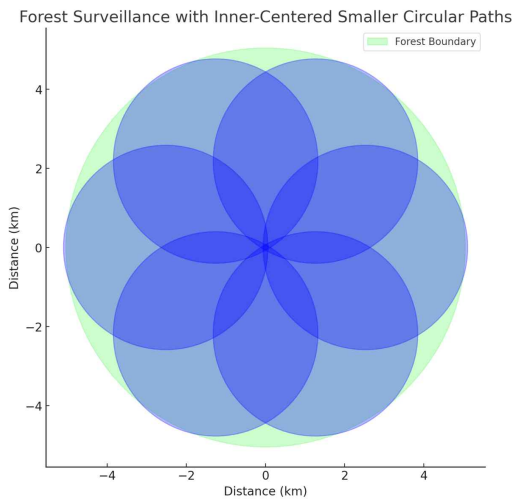


그림 4. 산림감시 예상 매커니즘  
Fig. 4. Expected Mechanisms for Forest Monitoring

그림 4는 이를 위한 해결 방안으로 예상하여 우리는 이 알고리즘을 제시한다. 초록색 원이 산림, 파란색 원이 드론이 감시하는 구역이다. 원형 산림의 면적을 80km<sup>2</sup> 드론 1대의 감시 면적을 21km<sup>2</sup>라고 가정한다. 이때 산림과 드론의 감시 면적을 원형으로 가정하면 식 (3), (4)를 적용할 수 있다.

$$r_{forest} = \sqrt{\frac{A_{forest}}{\pi}} = \sqrt{\frac{80}{\pi}} \approx 5.05Km \quad (3)$$

$$r_{drone} = \sqrt{\frac{A_{drone}}{\pi}} = \sqrt{\frac{21}{\pi}} \approx 2.59Km \quad (4)$$

총 6개의 스테이션을 산림 외곽에 원형으로 배치해 산림을 완전히 커버한다. 이때, 각 드론의 감시 반경이 서로 중첩되며 중앙까지 포함할 수 있다. 이러한 알고리즘을 사용한다면, 산림 중심지에서 멀어

져 접근성이 용이해지고 유지관리에 이점을 가질 수 있다. 각 드론은 일정 거리 내에서 중앙을 향하며 순환 비행하도록 배치되어, 드론 간 감시 영역이 서로 적절히 중첩되면서 중복을 최소화한다. 이 방식은 전체 산림을 균형 있게 커버하며, 모든 구역이 빠짐없이 감시되도록 한다. 따라서, 드론의 효율적 배치와 감시 성능을 극대화하여 산림 지역의 주요 구역을 효과적으로 감시할 수 있는 구조를 구성한다.

남산은 약 3.3 제곱킬로미터(330헥타르)의 비교적 작은 면적을 가지고 있다. 도심과 인접해 있어 많은 방문객이 찾으며, 산불이나 재난 발생 시 피해가 주저지로 빠르게 확산될 위험이 있다. 이러한 위치적 특성은 드론을 활용한 실시간 감시가 특히 효과적일 수 있는 환경을 제공한다. 따라서, 한 번의 충전으로도 주요 구역을 효율적으로 순찰하고 감시할 수 있는 교체형 배터리 방식을[12] 제안한다. 이는 빠른 응답 속도와 확장성을 제공해 산불 감시에서 효율적이고, 지속적인 모니터링이 가능하다. 하지만 초기 설치와 운영 비용이 높다는 단점이 있다.

지속적인 모니터링을 위해 특화된 배터리 교체 드론에는, PRENEU의 PANDION S2 드론이 소형이면서도 감시 및 모니터링에 최적화된 설계로 제작되었다.

그림 5의 프리뉴의 판디온 S2 드론은[12] 최대 비행시간 30분, 최고 속도 45km/h로 최대 22.5km 범위를 커버할 수 있다.



그림 5. PRENEU PANDION S2 (출처: PRENEU, 2023)  
Fig. 5. PRENEU PANDION S2 (source: PRENEU, 2023)

그림 6의 드론 스테이션은[12] 최대 5개의 배터리를 저장하며, 드론이 배터리를 소모할 때마다 착륙해 약 90초 이내에 자동으로 배터리를 교체해 준다. 이 방식은 연속적인 감시를 가능하게 하지만, 여분의 배터리 구비와 배터리 수명 관리가 필요하다. 따라서 인프라 구비에 있어 경제적인 측면도 충분히 고려해야 할 것이다.



그림 6. PRENEU 드론 스테이션 (출처: PRENEU, 2023)  
Fig. 6 PRENEU Drone Station (source: PRENEU, 2023)

태백전력 지사에는 그림 7과 같이 861개의 송전탑과 610.4c-km의 송전선로가 설치되어 있다.[13] 송전탑은 154kV의 초고전압 송전선로로 구성되어 있는데, 태백산과 같은 송전탑 밀집 산림 지역에서의 드론 충전 방식으로 송전탑 직접 충전 방식을 제안한다. 이미 존재하는 송전탑과 전력망을 그대로 사용하여 추가적인 인프라를 설치할 필요가 없어 경제성 면에서 유리하고 드론의 배터리가 부족할 시 특정 지점으로 돌아갈 필요 없이 근처 송전탑에 접근해 충전할 수 있다.

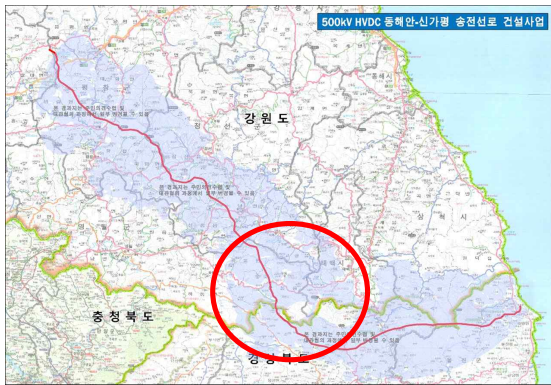


그림 7. 태백산 산림의 송전탑 구간  
Fig. 7. The section of the transmission tower in the Taebaeksan Forest(source: Park, Cham-beom, 2020)

Boaz Ben-Moshe는[14] 특정 조건에서 100~250V AC의 송전선 전압을 이용하여 드론을 충전할 수 있으며, 온보드 충전기를 통해 드론이 필요한 전력을 지속적으로 보충할 수 있도록 설계하였다. 이 온보드 충전 시스템은 배터리를 10%에서 70%까지 45

분 동안 충전하여 추가 비행시간 30분 이상을 확보할 수 있게 설계하였고, 무게에 따른 전력 소비량이 선형적으로 증가한다.



그림 8. 온보드 충전형 드론의 충전 시연  
Fig. 8. Charging Demonstration of Onboard Charging Drones(source: Ben-Moshe B, 2021, p11)

송전탑 직접 충전 방식은 산불 감시 드론의 운영 효율을 크게 높일수 있으며, 특히 장시간 비행과 주기적인 충전이 필요한 산림 지역 감시에서 매우 효과적인 방식이다.

### III. 결 론

본 연구는 산불 감시 및 초기 진압에서 회전익 드론의 활용 가능성과 이를 지원하는 충전 방식에 대해 논의했다. 회전익 드론은 고정익 드론에 비해 정밀 제어가 가능하여 초기 화재 진압과 같은 상황에 더 유리하며, 다양한 센서를 통해 실시간 모니터링이 가능하다. 그러나 장시간의 감시가 필요한 상황에서는 배터리 사용 시간이 제한적이라는 문제점이 있다.

이를 해결하기 위해 본 연구는 국내 산림의 지형적 특성에 따라 네 가지 보조 충전 방식을 제안했다. 각 방식은 특정 지역의 특성과 활용 목적에 따라 일조량이 높은 지역에는 태양광 충전이, 넓은 산림 지역에는 정류장 식 충전스테이션 방식이 적합하다. 교체형 배터리 방식은 빠른 응답과 확장성에 강점을 가지며, 송전탑 직접 충전 방식은 인프라 비용을 절감할 수 있는 경제적인 대안으로 제시된다.

이를 통해 산불 감시 드론의 운영 효율성을 높이는 데에 중요한 기초 자료를 제공하며, 향후 드론의 충전 인프라 구축과 충전 방식의 발전에 기여할 것으로 사료된다. 각 충전 방식의 장단점을 고려하여 드론 운영 계획을 수립함으로써, 산불 감시 및 초기 진압의 효율성을 한층 더 높일 수 있을 것이다.

### 참 고 문 헌

- [1] Lee, Sungjae, & Choi, Yosoon. (2016). "Comparison of Topographic Surveying Results using a Fixed-wing and a Popular Rotary-wing Unmanned Aerial Vehicle (Drone)." *Tunnel and Underground Space*, 26(1), 24-31.
- [2] Kim, Jang-Ook (2023-04-26). "A study on improving the efficiency of early fire suppression using artificial intelligence and drones."
- [3] An, Seong-su. (2024). "Development of a firefighting drone equipped with a fire extinguisher launcher and a powder extinguisher for initial fire suppression in high-rise buildings." *Journal of the Korean Society for Industry-Academic Technology*, 25(3), 385-392.
- [4] Tongyeong News. (2024). "Drones Now Take Flight First in Wildfire Monitoring in Tongyeong City."
- [5] Hanok News. (2023). "The role of drones in wildfire monitoring and response."
- [6] Korea Herald. (2023). "Wildfire monitoring drones deployed in Gyeongbuk region."
- [7] FLIR Korea. (2023). "How far can you measure with a thermal camera?"
- [8] IT Chosun. (2018). "Thermal imaging camera applications expand in drones and security industries."
- [9] O. M. Bushnaq, A. Chaaban and T. Y. Al-Naffouri, "The Role of UAV-IoT Networks in Future Wildfire Detection," in *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16984-16999, 1 Dec. 2021, doi: 10.1109/JIOT.2021.3077593.
- [10] Wooram Lee (2023). "Experimental Verification on the Extending Flight Time of Solar Paper for Drone using Battery for Electric Vehicles," 24(4), 229-235.
- [11] Argosdyne. (2022). "Aquila-3F."
- [12] PRENEU. (2023). "PANDION S2."
- [13] Ahn Kwang-hoon (2006). "154kV Taeback~Gohan Transmission Line Expansion Project 'On Track.'" *Electric News*.
- [14] Ben-Moshe B. "Power Line Charging Mechanism for Drones." *Drones*. 2021; 5(4):108.