

2022 한국정보과학회  
학술심포지움 논문집

유비쿼터스 컴퓨팅과 웹 정보기술

Proceedings of UCWIT 2022

Ubiquitous Computing and Web Information Technology  
Kyungpook National University, Dec. 2-3, 2022

제15권 제2호

일시: 2022. 12. 2. (금) ~ 2022. 12. 3.(토)  
장소: 경북대학교  
주최: 사단법인 한국정보과학회  
주관: 한국정보과학회 영남지부, 경북대학교 지능융합 소프트웨어 교육연구단,  
경북대학교 소프트웨어기술연구소



경북대학교  
KYUNGPOOK NATIONAL UNIVERSITY



# 대회장 인사말

어느덧 초겨울이 오고 2022년도를 마무리하는 이 시점에 올 한해에도 많은 성과가 있으셨기를 바랍니다. 명실상부 IT 분야 전국 최대 규모의 교육, 연구 및 산학협력 인프라를 보유하고 있어 국내외로 우수한 IT 전문인력을 배출하고 여러 연구성과를 자랑하는 경북대학교에서 2022년 한국정보과학회 영남지부 학술대회를 개최하게 되어 매우 기쁘게 생각합니다. 본 행사에 참석해 주신 많은 회원 여러분과 본 행사를 위해 준비해 주신 관계자분들에게 깊은 감사의 말씀드립니다.

한국정보과학회는 컴퓨터와 소프트웨어 및 정보통신 분야에서 우리나라 최고의 학회로 발전을 거듭하고 있으며 국내외 IT 산업에 크게 기여하고 있습니다. 본 학술대회 행사는 한국정보과학회가 주최하고 한국정보과학회 영남지부, 경북대학교 컴퓨터학부/BK21 지능융합 소프트웨어 교육연구단이 공동주관하고 있습니다. 한국정보과학회는 기술 분야별 소사이어티 활성화와 지부 운영을 적극 지원하고 있으며 영남지부의 본 행사와 같은 활발한 연구 활동은 한국정보과학회 발전을 비롯하여 국가 IT 발전에 지대한 공헌을 할 것이라 기대됩니다.

한국정보과학회 영남지부는 “유비쿼터스 컴퓨팅과 웹 정보기술 (Ubiquitous Computing and Web Information Technology, UCWIT)” 이란 명칭의 학술대회를 지금까지 운영해 오고 있습니다. 본 행사는 유비쿼터스와 웹 정보기술을 중심으로 인공지능과 빅데이터를 비롯한 정보 통신 및 전통 산업과의 IT 융합 기술을 주제로 분야별 전문가들이 모여 최신 주제를 토론하는 학술대회입니다. 특히나, 요즘 인공지능, 빅데이터, 블록체인, 메타버스, 6G와 같은 4차 산업혁명을 주도하는 요소 기술들이 급속도로 발전하는 가운데 본 행사를 통해 해당분야의 다양한 연구 성과들을 공유하며 토론할 수 있는 좋은 기회가 될 것이라 기대됩니다.

끝으로, 그동안 본 학술대회 개최를 위하여 애써온 조직위원장, 학술위원장, 집행위원장 및 위원님들의 노고에 심심한 감사의 마음을 전합니다. 또한 본 대회 개최를 위해 장소를 제공하고 여러 가지 지원을 해주신 경북대학교 관계자 여러분과 후원 기관 여러분에게 고마움을 전합니다. 이번 학술회의에 참석하신 참석자 여러분의 관심과 협력에 깊은 감사의 말씀을 드리며 항상 건강하시길 바랍니다.

대회장 한국정보과학회 회장 심규석

# 조직/학술 위원장 인사말

열역학 제2법칙에 따르면 닫힌계에서 엔트로피는 점점 증가한다고 합니다. 우리가 사는 세상도 이 법칙을 따르는 것 같습니다. 새로운 질병이 나타나기도 하고, 새로운 전쟁이 일어나기도 합니다. 세상이 점점 어지러워질 때 등장하는 구세주는 이젠 당연히 컴퓨터와 IT 기술입니다. 이러한 상황에서 개최되는 한국정보과학회 학술심포지움 유틸리티 컴퓨팅과 웹 정보기술(Ubiquitous Computing and Web Information Technology, UCWIT)는 그 의미가 남다르다 하겠습니다. 한국정보과학회 영남지부 주관의 UCWIT 2022 행사는 산업체, 연구소 및 학계에서 활동하시는 지역 IT 전문가들을 위한 교류의 장입니다.

우리 지역에서 국제적 경쟁력을 갖춘 지역의 주력산업인 반도체, 기계, 자동차, 조선, 에너지, 철강, 물류 등 전통산업은 많은 변화를 요구하고 있으며, 그 변화의 중심에는 제4차 산업혁명이라고 일컬어지는 ICT와의 융합이 차지하고 있습니다. 이번 학술심포지움을 통하여 지역의 산업이 더욱 업그레이드되어 경쟁력을 강화하는 기회가 될 수 있기를 기대합니다.

조직위원회에서는 제4차산업혁명의 핵심 이슈들을 중심으로 금년 행사를 준비하였습니다. 본 행사는 논문발표와 특별강연으로 구성됩니다. 논문발표는 총 4개의 세션으로 이루어집니다. 산학연 각계에서 제출된 논문은 메타버스/콘텐츠, 인공지능/딥러닝, IOT/클라우드/빅데이터 및 학생들의 참신한 아이디어를 구현한 프로그램들에 대한 다양하고 흥미로운 주제를 다루고 있습니다.

본 행사가 활발한 교류의 장, 그리고 격의 없는 토론의 장이 될 수 있도록 여러분의 많은 참여를 바랍니다. 그리하여 본 행사가 지역 IT 분야의 학문적 발전과 관련 산업의 성장에 이바지할 것을 기대합니다.

행사 준비를 위해 수고하신 심규석 한국정보과학회 회장님을 비롯하여, 영남지부 운영위원님들, 대회 조직위원님들과 프로그램위원님들께 감사드립니다. 아울러 물심양면으로 지원해주신 하이브레인넷, (주)큐텍코리아 및 행사 장소를 기꺼이 제공해 주신 경북대학교 홍원화 총장님께 감사드리며 그리고 학문적 성과의 공유를 위해 기꺼이 논문을 투고해주신 저자님들께 진심으로 감사합니다.

조직위원장 **이 수 현** (창원대, 한국정보과학회 영남지부 지 부 장)  
학술위원장 **김 일 곤** (경북대)

# 대회 조직

대회장 한국정보과학회장 심규석 (서울대학교)

조직위원장 정보과학회 영남지부장 이수현 (창원대학교)

학술위원장 김일곤 (경북대학교)

학술위원 박혜영, 탁병철, 김용태 (경북대학교)

집행위원장 김동균 (경북대학교)

집행위원 정인욱, 김정근 (경북대학교)

고문 강순주 (경북대학교)  
김석규 (안동대학교)  
김순철 (대구대학교)  
김영학 (금오공과대학교)  
김정구 (부산대학교)  
변정용 (동국대학교)  
우용태 (창원대학교)  
송하주 (부경대학교)  
정민수 (경남대학교)  
정중수 (안동대학교)

# 행사 일정

12월 2일(금)

시간	행사		
	Room #1 (융복합관 345호)	Room #2 (융복합관 245호)	Room #3 (융복합관 248호)
13:30 - 14:00	등록 및 접수 (융복합관 1층)		
14:00 - 14:30 (개회식)	개회식 (장소: IT융복합관 345호, 좌장: 경북대학교 김동균 교수)  개회사: 이수현 한국정보과학회 영남지부장 축사: 한동석 경북대학교 IT대학장 환영사: 심규석 한국정보과학회회장		
14:30 - 15:10 (특강)	특강 1: "Powerful ABB (AI·Bigdata·Blockchain)" (장소: IT융복합관 345호, 좌장: 경북대학교 김동균 교수)  대구광역시 미래ICT국 AI블록체인과장 류동현 박사		
15:20 - 16:40 (논문 발표)	Session #1: Metaverse, Contents & Network Technology  좌장: 이종택 (경북대학교)	Session #2: Recent Advances in Deep Learning  좌장: 정인욱 (경북대학교)	Session #3: Applications of Artificial Intelligence  좌장: 탁병철 (경북대학교)
16:40 - 16:50	휴식		
16:50 - 18:00 (논문 발표)	Session #4: ICBMS (IoT, Cloud, Bigdata, Mobile, Security) Technology [Online & Offline]  좌장: 김정근 (경북대학교)	포스터 세션  좌장: 남덕윤 (경북대학교)	
18:00 - 18:30	학술대회 폐회		

12월 3일(토)

시간	행사
10:00 - 12:00	한국정보과학회 영남지부 총회

## Session #1: Metaverse, Contents & Network Technology

좌장: 이종택 (경북대학교)

번호	제목 저자 및 소속
1-1	<b>Cell-Free mMIMO: A Next Generation Paradigm</b> Mahnoor Ajmal, Ayesha Siddiqa, 김동균 (경북대학교)
1-2	<b>골목길에서 노드 이동성을 고려한 5G 기반 사고 위험 예측 모델</b> 전혜림, 서경희, 최희정, 홍승은, 백호기* (경북대학교)
1-3	<b>실시간 주문형 디지털 콘텐츠의 DRM 패키징 방법</b> 신서윤, 김순철 (대구대학교)
1-4	<b>대규모 미팅 행사를 위한 어플리케이션 구조 제안</b> 이승민, 노준수, 신지한, 이수인, 최혜림 (경북대학교)
1-5	<b>판매데이터와 공공데이터를 활용한 상권분석 시스템 구축</b> 김건준, 이상협, 조원영, 신지송, 박찬식 (경북대학교)

## Session #2: Recent Advances in Deep Learning

좌장: 정인욱 (경북대학교)

번호	제목 저자 및 소속
2-1	<b>Theoretical analysis of forward and backward propagation in Neural Networks</b> Musulmon Lolaev, Anand Paul (경북대학교)
2-2	<b>A study of Medical prior Knowledge Automatic CAC Voxel Classification in CT Using Unsupervised Deep learning Architecture</b> Jabeen Farah, 김재일 (경북대학교)
2-3	<b>YOLOv5를 이용한 플라스틱 재활용 폐기물 선별 정확도 향상에 대한 연구</b> 유준호, 안현아, 이윤서, 이인희, 이지현, 김한영, 정설영 (주식회사 워, 경북대학교)
2-4	<b>트랜스포머 기반의 BERT 모델과 온라인 포럼 데이터를 활용한 시스템 근본 원인 종류 탐색</b> 김유양, 송소민, 탁병철 (경북대학교)
2-5	<b>Falling Detection of Workers using Deep Learning with Sensing Data</b> Jabeen Shaisata, 홍봉희 (부산대학교)

### Session #3: Applications of Artificial Intelligence

좌장: 탁병철 (경북대학교)

번호	제목 저자 및 소속
3-1	시계열 생체 데이터를 RP신호로 변환하기 위한 효과적인 방법 전영준, 강순주 (경북대학교)
3-2	염색 공정 데이터를 이용한 친환경 섬유 소재 내광성 등급 예측 모델 개발 서강복, 김덕엽, 이우진 (경북대학교)
3-3	블루투스 센서와 자기장 센서를 혼합한 실내 측위 방법 한승준, 권영우 (경북대학교)
3-4	프로그래밍 교육 피드백을 위한 누적 학습 데이터 기반 학습자 수준 예측 오승민, 김덕엽, 이우진 (경북대학교)
3-5	한글 손글씨 문장 인식에 대한 OCR 오픈 소스 성능 비교 분석 신찬규, 김세연, Barathi Subramanian, 김정홍 (경북대학교)

### Session #4: ICBMS (IoT, Cloud, Bigdata, Mobile, Security) Technology [Online & Offline]

좌장: 김정근 (경북대학교)

번호	제목 저자 및 소속
4-1	가상 환경을 통한 커스텀 학습 데이터 생성 및 기계 학습의 유효성 검증 (온라인) 배진호, 구은성, 김근찬, 김영호, 최희정, 이효경 (경북대학교)
4-2	입원 환자 도움 프로그램 (온라인) 김재우, 전찬용, 안동혁, 최도진 (창원대학교)
4-3	쿠버네티스를 활용한 MLOps 플랫폼 'MeLonKube' 개발 (온라인) 김석희, 박보경, 박종민, 신시온, 이에림, 이동욱, 고석주 (경북대학교)
4-4	효과적인 기상 정보 안내를 위한 딥러닝 기반 수어 교육 시스템 (온라인) 손혜정, 윤선아, 임가은, 조건형, 손윤식 (동국대학교)
4-5	IoT Hub 보안성 향상을 위한 취약점 분석 프로세스에 관한 연구 (온라인) 장지원, 윤원희, 김성은, 김지연, 안영준, 강준혁, 강인욱, 김홍진 (한국정보기술연구원 차세대 보안 리더 양성 프로그램)
4-6	유니티를 이용한 학습용 콘텐츠 기반 게임 설계 (온라인) 강지윤, 안진수, 이수민, 김재우, 최도진 (창원대학교)
4-7	생산라인의 이상유무와 예지보전 판단을 위한 인공지능 기반 모바일 애플리케이션 (온라인) 김명균, 전왕수, 이상용 (경남대학교)
4-8	저장 매체에 따른 카프카 처리량 분석 (온라인) 이동영, 강동현 (창원대학교)

## 포스터 세션

번호	제목 저자 및 소속
5-1	<b>바코드를 이용한 사물인식 시스템 설계</b> 권순성, 이 셋, 강호윤, 김우현, 정중수 (안동대학교)
5-2	<b>스마트밴드를 이용한 위험 심박수 감지 모바일앱</b> 강동현, 정근모, 송하주 (부경대학교)
5-3	<b>양상블 기법을 이용한 PCB 불량 검출 방법에 대한 연구</b> 최성혁, 배완기, 유병주, 김동현, 정설영, 이문용 (경북대학교, JS SYSTEM Co., Ltd)
5-4	<b>IoT 기반 화재 정보 감지 시스템의 설계</b> 이셋, 정중수, 유은상 (안동대학교, (주) 라성에너지)
5-5	<b>Streamlit 프레임워크를 이용한 TTS 및 SVS 기반 웹 서비스 개발</b> 김예원, 이충현, 김성재, 김인중 (한동대학교)
5-6	<b>반려동물 돌봄을 위한 모바일 앱 설계 및 구현</b> 김도선, 정민욱, 박희원, 정원준, 김동균*, 임호만 (경북대학교, (주)에스엠에스)
5-7	<b>머신러닝을 활용한 인지기능 상태 예측</b> 구동준, 이아라, 원경재, 김일곤 (경북대학교)
5-8	<b>아두이노를 이용한 무호흡 감지 및 증상 완화 모듈 설계 및 구현</b> 이현로, 김대영, 서민석, 김동균, 이유락 (경북대학교, 조은캠프)
5-9	<b>WebRTC 기반 실시간 영상공유 및 모니터링 시스템의 설계 및 구현</b> 이원근, 이성원, 천용태, 최명빈 (경북대학교)
5-10	<b>아두이노 압력센서 데이터 기반 딥러닝 자세 예측 의자</b> 김주형, 심규혁, 우용태 (창원대학교)
5-11	<b>게임에서 NFT의 활용</b> 김정훈, 이수현 (창원대학교)
5-12	<b>Social IoT traffic prediction</b> Dong Hu, Limei Peng(경북대학교)



# Cell-Free mMIMO: A Next Generation Paradigm

Mahnoor Ajmal<sup>o</sup>, Ayesha Siddiqa, Dongkyun Kim

mahnoor.ajmal@knu.ac.kr, asiddiqa@knu.ac.kr, dongkyun@knu.ac.kr

School of Computer Science and Engineering, Kyungpook National University

## 요약

Cell-Free (CF) massive MIMO (mMIMO) system features dense deployment of access points (APs), distributed over a large area to serve a small number of users simultaneously. This approach removes cell edges and introduces overlapping regions for serving users hence providing a uniform quality of services. CF systems have been regarded as a potential next-generation technology for 5G networks. In this paper, detailed system architecture, principle methods, and key challenges for CF mMIMO deployment are discussed. Furthermore, a succinct overview of the recent research contributions and future considerations for the aforementioned scheme is investigated.

## 1. INTRODUCTION

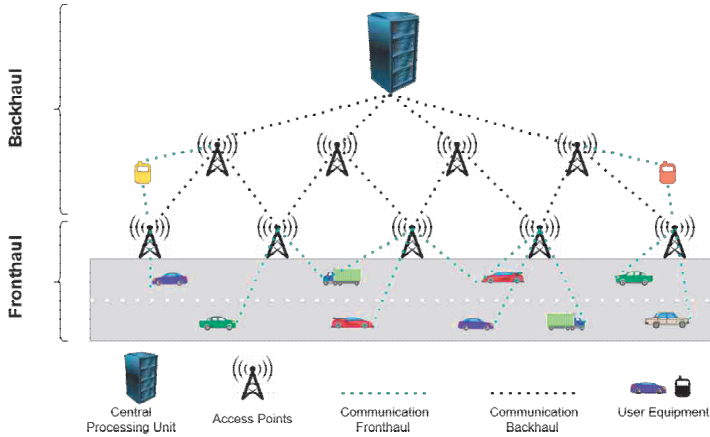
International Telecommunication Union (ITU) issued specifications for the 5G (MT-2020) in 2015 [1]. According to these specifications, 5G services must deliver enhanced mobile broadband (eMBB) with immense downlink and uplink speed. For eMBB, ITU has recommended 20 Gbps downlink (DL) speed and 1 Gbps uplink (UL) speed so it can support applications like smart health services, drone surveillance, and IoT. 5G is also required to have one million machines connected per square km with massive machine-type communication (mMTC). This specification will support various applications like smart cars, smart cities, robotics, and smart machines. Ultra-high reliability and low latency (URLC) is another important specification for 5G networks set by ITU. The aforementioned specifications motivate researchers to investigate efficient techniques that have the capability to incorporate 5G communication network demands. Cell-Free (CF) massive multiple input multiple outputs (mMIMO) is a new technology that has the potential to implement these new specifications, specifically mMTC and IoT in which devices are sparse in large geographical areas. Over the recent years, researchers came up with multiple technologies to meet the ever-increasing demands of low latency, efficient and minimal bandwidth consumption, high throughput, reliable connection, and much more. Multiple input multiple output (MIMO) is one of the promising technologies that accommodate most of the aforementioned requirements. In MIMO, multiple antennas are mounted on the base station to serve multiple users in a cell. MIMO transmitter has the ability to deliver multiple signals simultaneously while utilizing the installed multiple antennas at the base station. MIMO enhances the quality of communication by providing antenna diversity, beamforming capabilities, and spatial multiplexing [2]. Massive MIMO (mMIMO) is an extension of MIMO where a massive number

of antennas are installed at base stations that not only improve signal quality but also accommodate more users. mMIMO improves data rates while utilizing the same frequency-time resources [2]. mMIMO can be categorized into two broad categories depending on the deployment of antennae.

- 1) If all the antennas are mounted on the base station, then we term it as centralized mMIMO
- 2) If antennas are geographically distributed in a large area, on different access points (APs), then we term it as distributed mMIMO

Centralized mMIMO enables spatial multiplexing in which different data streams are sent using the same time-frequency resources but there is a considerable difference between the signal quality of the user placed near the base station and the one placed at the edge of the cell. In this regard, distributed mMIMO gives better performance as it ensures the same QoS to all the users regardless of their position in a cell. Furthermore, it also removes inter-cell interference and reduces the required transmission energy. mMIMO technology has been one of the most significant developments in wireless communication but still it has some issues. In distributed mMIMO, interference from neighboring base stations is a crucial issue because of the disjoint region. To overcome this, an incarnation of distributed massive MIMO system termed as Cell-free (CF) massive MIMO (mMIMO) [3] system was introduced. In CF- mMIMO, large numbers of Access Points (APs) are distributed in a wide area and they are not bounded by cellular regions. It not only ensures the uniform quality of service but also provides reduced multi-user interference [4]. The outline of the rest of the paper is as follows. The working of CF- mMIMO is discussed in Section-II. We start with the brief working

model and narrow it down to fronthaul working is subsection-A. Channel estimation, data transmission from users to APs is discussed. In subsection B, backhaul working is explained. Section III examines the literature for both fronthaul and backhaul. This section also investigates the issues faced while deployment of CF-mMIMO networks and provide future recommendation on the aforementioned topic. Finally, in section IV, we conclude the paper.



1. Fig. 1. Cell Free massive MIMO Architecture

## 2. SYSTEM MODEL

Cell-Free massive MIMO ensures macro-diversity by deploying access points in a wide range of areas. If we look into the architecture of CF-mMIMO depicted in “Fig.1 refs fig”, we have two links in which it operates: fronthaul and backhaul. In fronthaul, a large number of APs are serving users via wireless connection while in backhaul these access points are connected to a central processing unit (CPU) where most of the processing takes place. In a typical scenario, data transmission in a CF-mMIMO system operates in three stages

- Users generate their backhaul signal and transmit it to numerous access points in their vicinity
- APs detect receive user signals by using detection algorithms like Maximum ratio combining (MRC) and then transmit the detected user signal to the CPU via a backhaul link.
- Finally, User source signals are recovered at the CPU by aggregation of received backhaul signals.

### A. Front-haul (From Users to Access Points)

Fronthaul refers to the link from Users to APs. In a CF-mMIMO network, if a user is in the vicinity of more than one AP at a time, then it connects with all of them simultaneously. For working, initially, APs periodically broadcast reference signals (RF) in the networks. Users

receive the reference signal (RF) signal and send it back to the APs along with channel noise. This transmission path from the user to APs is termed an uplink (UL). APs estimate the channel on observed signals power and channel noise. After channel estimation, APs assign network resources to users for uplink (UL) and downlink (DL) communication. After threshold time APs send orthogonal pilot information to users via downlink (DL). Finally, a communication channel has been established between APs and users for UL/DL data communication. The frequent channel estimation and resource allocation in the case of highly mobile users is still an open issue.

### B. Back-haul (From Access Points to CPU)

The backhaul refers to a link from the access points to the CPU. Access points have to gather user data and forward it to the CPU via wired or wireless link. The amount of data that needs to be forwarded to the CPU varies from case to case. Generally, the data that APs send to the CPU consist of two types, one containing the channel information from each user and the other with the received user signal [5]. Due to bandwidth constraints, AP sends a quantized version of signals only. Each AP multiplies the received signal by the conjugate of the estimated channel from each user and sends back a quantized version of this weighted signal to the CPU.

## 3. SCOPE, CONTRIBUTION, AND FUTURE RECOMMENDATIONS

CF- mMIMO networks is a potential 5G New Radio (NR) topic currently being investigated by numerous researchers. For fronthaul, the pilot assignment on an increasing number of users is a research perspective that is being studied. To overcome the pilot contamination issue in the network various techniques have been discussed in [6]. The proposed algo- rithm iteratively estimates the LSF coefficient and assigns a pilot sequence with a minimum data rate. The author in [7] proposed the search-based approach to select the sub-optimal pilot assignment. They have suggested that If the user is served by the limited APs in the network, then the probability of pilot contamination can be minimized. Various frequency division duplexing (FDD) and time division duplexing (TDD) protocols have been proposed to overcome channel estimation and communication problems on fronthaul. The authors in [8] have proposed the system model to utilize frequency division du- plexing (FDD) to estimate the channel. Moreover, it minimizes the angle reciprocity property of the FDD system by assuming that uplink (UL) and downlink (DL) channels are similar. The author in [9], [10] compared the performance of LS and MMSE for channel estimation in CF networks. The results depicted that LS has a lower data rate as compared to MMSE. This is because it has considered a Rician fading channel for information

collection. The proposed method is based on line of sight (LoS) components. The authors in [11], [12] estimated the performance of CF-mMIMO system in the non-phase aware channel estimation. The result shows significant degradation in the system performance compared to the phase-aware method. Most of the research work targeting backhaul considers wired backhaul that uses cables or optical fiber for data transmission but in the future, this architecture will not be suitable due to its low flexibility, deployment cost, and above all the fact that wired connection cannot be implemented everywhere [13]. Wireless Backhaul is an open research area that needs to be explored for Cell-Free mMIMO systems due to its high flexibility, low cost, and low implementation cost. The dedicated international telecommunications union (ITU) frequency bands for 5G; FR1, Sub 6-GHz, operates below 6-GHz, and FR2, mmWave, operates above 24-GHz requires the use of advanced technologies for signal transmission like millimeter wave (mmWave), Non-Orthogonal Multiple access (NOMA), Orthogonal Frequency Domain Multiple Access (OFDMA). In [14], the authors have proposed the use of NOMA for backhaul signal processing along with DNN-based encoders and decoders. The use of NOMA results in system bandwidth reduction while DNN-based encoders and decoders ensure computation mapping optimization by removing redundancy and ensuring transmission accuracy. Note that we can even improve the performance at backhaul by integrating data mining techniques. For example, in [5], authors have proposed to implement NOMA by using the DNN model. DNN model performs efficient user clustering by optimizing hyperparameters that are power and channel diversity in [5]. A hybrid scheme merging approaches used in [14] and [5] has the potential to drastically improve the performance and reduce network overhead. Another relatively new concept for 5G New Radio (NR) is Integrated Access and backhaul (IAB) [13]. In this technique, the same wireless channel is used for coverage and backhaul connectivity among Access points. In 5G, ITU dedicated frequency bands are high as discussed before so this technique makes sure of efficient use of spectrum, improved performance along with the lower cost of equipment. IAB is essentially a multi-hop approach that functions by having a fraction of the deployed BSs act as donor nodes, using a wired/fiber connection. The remaining Base stations that do not have a wired connection are termed IAB nodes. Both IAB nodes and IAB donors foster the same cellular coverage area and appear identical to Users. Another advantage for operators of integrated Access and Backhaul operators is that they can make use of their current mmWave spectrum licenses and use separate backhaul as needed without having to conduct additional interference studies that may be subject to regulatory approval.

#### 4. CONCLUSION

In this paper, we have presented a detailed discussion of the Cell-Free (CF) massive MIMO (mMIMO) systems model, its scope, contributions, and future recommendations. The CF system model separately discussed the fronthaul and backhaul working. Moreover, it also specifies the communication process for UL/DL data transmission. The paper has also specified the scope of CF networks, discussed the recent research contributions, and recommends future research directions. In summary, this paper gives a comprehensive overview of CF systems with state-of-the-art contributions.

#### ACKNOWLEDGMENT

This research has been supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT), (NRF-2022R1A2C1003620). In addition, the research has also been funded by the BK21 FOUR project (AI-driven Convergence Software Education 220 Research Program) funded by the Ministry of Education, School of Computer Science and Engineering, 221 Kyungpook National University, Korea (4199990214394).

#### REFERENCES

- [1] Erik Dahlman, Stefan Parkvall, and Johan Skold. 5G NR: The next generation wireless access technology. Academic Press, 2020.
- [2] Jiayi Zhang, Shuaifei Chen, Yan Lin, Jiakang Zheng, Bo Ai, and Lajos Hanzo. Cell-free massive mimo: A new next-generation paradigm. *IEEE Access*, 7:99878–99888, 2019.
- [3] Salah Elhoushy, Mohamed Ibrahim, and Walaa Hamouda. Cell-free massive mimo: A survey. *IEEE Communications Surveys & Tutorials*, 24(1):492–523, 2021.
- [4] Hussein A Ammar, Raviraj Adve, Shahram Shahbazpanahi, Gary Boudreau, and Kothapalli Venkata Srinivas. User-centric cell-free massive mimo networks: A survey of opportunities, challenges and solutions. *IEEE Communications Surveys & Tutorials*, 24(1):611–652, 2021.
- [5] Manijeh Bashar, Kanapathippillai Cumanan, Alister G Burr, Hien Quoc Ngo, and Merouane Debbah. Cell-free massive mimo with limited backhaul. In 2018 IEEE International Conference on Communications (ICC), pages 1–7. IEEE, 2018.
- [6] Giovanni Interdonato, Hien Quoc Ngo, Pal Frenger, and Erik G Larsson. Downlink training in cell-free massive mimo: A blessing in disguise. *IEEE Transactions on Wireless Communications*, 18(11):5153–5169, 2019.

- [7] Heng Liu, Jiayi Zhang, Xiaodan Zhang, Adit Kurniawan, Tutun Juhana, and Bo Ai. Tabu-search-based pilot assignment for cell-free massive mimo systems. *IEEE Transactions on Vehicular Technology*, 69(2):2286–2290, 2019.
- [8] Asmaa Abdallah and Mohammad M Mansour. Angle-based multipath estimation and beamforming for fdd cell-free massive mimo. In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2019.
- [9] Stefano Buzzi, Carmen D’ Andrea, and Ciro D’ Elia. User-centric cell free massive mimo with interference cancellation and local zf downlink precoding. In *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–5. IEEE, 2018.
- [10] Stefano Buzzi, Carmen D’ Andrea, Alessio Zappone, and Ciro D’ Elia. User-centric 5g cellular networks: Resource allocation and comparison with the cell-free massive mimo approach. *IEEE Transactions on Wireless Communications*, 19(2):1250–1264, 2019. 4
- [11] Zhe Wang, Jiayi Zhang, Emil Bjornson, and Bo Ai. Uplink performance of cell-free massive mimo over spatially correlated rician fading channels. *IEEE Communications Letters*, 25(4):1348–1352, 2020. [12] Ozgecan Ozdogan, Emil Bjornson, and Jiayi Zhang. Downlink performance of cell-free massive mimo with rician fading and phase shifts. In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2019.
- [13] Ali Hosseinalipour Jazi, S Mohammad Razavizadeh, and Tommy Svensson. Integrated access and backhaul in cell-free massive mimo systems. *arXiv preprint arXiv:2210.12633*, 2022.
- [14] Hanxiao Yu, Neng Ye, and Aihua Wang. Non-orthogonal wireless backhaul design for cell-free massive mimo: An integrated computation and communication approach. *IEEE Wireless Communications Letters*, 10(2):281–285, 2020.

# 골목길에서 노드 이동성을 고려한 5G 기반 사고 위험 예측 모델 성능 평가

전혜림<sup>○</sup>, 서경희, 최희정, 홍승은, 백호기(교신저자)

경북대학교 컴퓨터학부

{jhr020528, k0\_h2\_218, chj00, y.hong, neloyou}@knu.ac.kr

## Performance Evaluation of 5G-Based Accident Risk Prediction Model Considering Node Mobility in Alleyway

Hyerim Jeon<sup>○</sup>, Kyunghee Seo, HeeJung Choi, Songeun Hong, Hoki Baek(Corresponding author)

School of Computer Science and Engineering, Kyungpook National University

### 요약

최근 자율주행 기술이 발전함에 따라 골목길 및 이면도로에서의 교통사고를 줄이기 위해서 사고 위험 예측 모델의 중요성이 대두되고 있다. 본 논문에서는 골목길 및 이면 도로에서의 5G 기반 사고 위험 예측 모델을 제안하였다. 기존 모델과는 달리 기지국으로부터 매 초마다 전달받는 user equipment(UE)들의 위치, 속도, 가속도, 방향을 고려하여 동적으로 위험 반경을 정의하기 때문에 사고 전에 신속하고 정확한 알림을 사용자에게 전달할 수 있다. 성능 평가에서는 제안한 모델이 사고가 나는 시나리오에서 미리 사고 위험을 예측하여 사용자에게 알림을 보낸 비율이 다른 모델들보다 월등히 높다는 것을 알 수 있었다. 특히 사고가 나지 않았는데도 알림이 잘못 울리는 비율 또한 평균 30%이상으로 낮아 기존의 모델들에게서 잦게 발생했던 잘못된 알림으로 인해 발생하는 사용자의 경각심 저하 문제 또한 해결할 수 있었다.

### 1. 서론

교통 환경이 발전함에 따라 삶의 질이 향상되고 있으나 끊임 없이 발생하는 교통사고는 인명 피해뿐만 아니라 경제적 손실까지 초래하며 심각한 사회 문제로 대두된다. 삼성교통안전문화연구소에 따르면, 도시부 발생 교통사고의 64.9%, 교통사고 사망자의 58.8%가 도로 폭 13m 미만 도로에서 발생한 것으로 나타났다[1]. 도로의 폭이 좁으면 운전자의 시야 역시 좁아지므로 상대 차량의 위치를 빠르게 파악하기 어려워 사고 발생 가능성이 커진다. 또한 차도와 보도의 구분이 없는 골목길과 이면 도로의 특성상 갑자기 튀어나오는 보행자에 대한 대비가 어려워 차량과 보행자 사이의 교통사고인 차 대 사람 사고 역시 빈번하게 발생한다.

자율주행 자동차의 위험 예측 기반 기술은 사전에 주변 User Equipment(UE)와의 충돌을 예측함으로써 이러한 사고 발생률을 감소시킨다. UE와의 충돌이 예상될 때 발생하는 알림은 3가지 평가지표를 통해 정확도를 검증받는다. 평가 지표는 correct(실제 충돌이 일어나기 전 알림을 준 경우), false alarm(FA: 실제 충돌이 일어나지 않았지만 알림을 준 경우, 충돌이 일어났지만 알림을 충돌 시점에 비해 너무 일찍 준 경우) 그리고 miss(충돌이 일어났지만 알림이 울리지 않거나 늦게 울린 경우)이다.

기존 연구는 차량과 보행자의 속도 및 가속도를 고정된 값으로 제한하여 이동 반경을 파악하고 충돌을 감지하였

다. 그러나, 특정 시간 동안 위험 반경이 고정된다면 실시간 특성이 고려되지 않아 FA가 높아져 운전자의 경각심을 낮춘다는 문제가 발생한다. 본 연구에서는 속도, 가속도, 두 UE 간의 상대 속도까지 매초 고려하는 속도·가속도 고려 동적 알고리즘

을 제안하며 FA와 miss를 최소화하는 것을 목표로 한다. 이를 통해 충돌 예측의 정확도를 높임으로써 사고를 예방함과 FA로 인한 사용자의 경각심 문제도 해결하고자 한다.

### 2. 관련 연구

기존의 선행 연구에서는 차량과 보행자 간 충돌 위험보다 차량 간 충돌 방지를 위해 위험 반경을 설정하는 경우가 대부분이다. 그러므로 차도와 보도의 구분이 명확하지 않고 차량의 진행 방향이 확실하지 않은 이면 도로에서는 성능을 보장하기 어렵다.

최근 주행 중 충돌을 감지하기 위한 기술이 꾸준히 연구되고 있으며, 시중에 출시되고 있는 서비스들도 적지 않다. 예를 들어 충돌을 감지하는 차량과 전방 차량의 위치와 속도를 받아 계산하는 [2], 위험 영역을 설정하고, 영역 내 위치한 차량 중 주변 차량과의 충돌까지의 시간이 1.6초 미만일 경우 범위 내 차량에만 경고 알림이 울리도록 하는 [3], 충돌 발생 예상 구역을 사전에 정의해두고, 해당 구역에 진입하는 차량에만 경고 및 제어를 실행하는 [4], 그리고 보행자의 진행 방향 및 주변 보행자 거동을 고려하여 현재 위치, 예측 위치로 보행자의 진행 방향을 예측하는 [5]가 있다.

### 3. 가속도·속도 고려 동적 알고리즘

#### 3.1 가정 사항

Sub-carrier Spacing(SCS)를 키워 symbol 길이를 줄임으로써 마이크로초 단위로 시간 할당 및 스케줄링을 할 수 있다는 것과 가변 슬롯을 적용하여 대역폭을 설정할 수 있다는 5G의 특징을 이용하기 위해 본 연구는 5G 표준을 기반으로 진행됨을 가정한다. 30kHz 대역폭을 이용하며, downlink, uplink 시 자원 할당 주기가 초 단위로 진행됨을 보장하고, 각 UE에는 5G 통신 단말기가 부착되어 있음을 가정한다.

UE에서 기지국으로 보내는 uplink 메시지에 UE의 종류, UE의 속도 및 가속도, UE 위치, 이동 각도를 포함한다. UE의 종류에는 차량, 전동 키보드, 오토바이, 사람이 포함되며, 사람을 제외한 vehicles는 모두 용어를 차량이라고 통칭한다. 지상에 존재하는 모든 UE의 고도는 동일하다고 가정하며, UE의 위치는 xy 평면상의 좌표를 이용하여 표현한다. UE의 이동 방향은 각도로 표현한다. UE의 이동 각도는 가장 처음 이동하기 시작할 때의 각도를 일정하게 유지한다. Downlink로 보내는 메시지는 충돌 위험이 감지되었다는 경고 메시지만을 포함하고 경고 알림은 각 차의 상태를 고려한 값으로 생성된 두 개의 반원이 겹친 경우에만 전송된다.

차량의 속도는 차량에 탑재된 휠 속도 센서로 측정되며 휠 속도 센서를 이용해 측정된 속도는 CAN 통신선을 따라 On Board Unit(OBU)로 전달 및 기지국으로 전송된다. CAN 통신선을 이용할 때의 통신 속도는 승용차의 평균 통신 속도인 500Kbps 임을 가정한다.

### 3.2 전체 시스템 개요

표 1. Parameters 정의

Parameters	
차량 초기 속도(m/s)	$v_0$
보행자 초기 속도(m/s)	$v_0^p$
차량 초기 가속도(m/s)	$a_v$
보행자 초기 가속도(m/s)	$a_p$
가속도 고려 속도(m/s)	$v$
UEs 위험 반경 반지름끼리의 합	$r_{sum}$
UEs 사이의 거리	$D$
업데이트 주기(s)	$T$

본 연구에서 제안하는 속도·가속도 고려 동적 알고리즘의 전체 개요도는 그림 1과 같다.

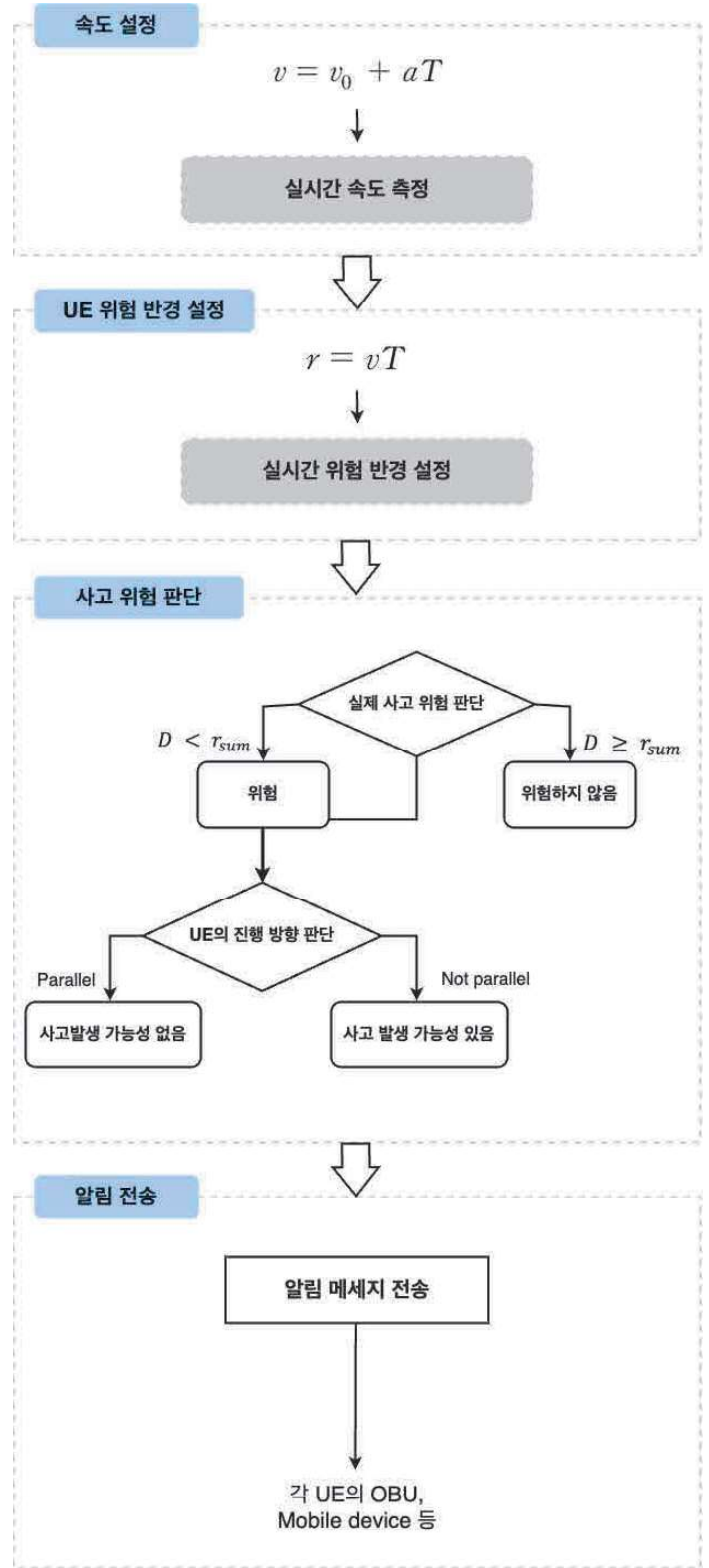


그림 1. 속도·가속도 고려 동적 알고리즘 전체 개요도

그림 1에서 속도 설정의 경우, 수식 (1)과 같이 UE의 가속도  $a$ 를 고려해 등가속도 직선 운동 공식을 사용하여  $T$ 초마다 바뀌는 속도  $v$ 를 구한다. UE는 3.1 가정 사항에서 구분한 유형에 따라 설정한 범위 내의 가속도를 가지며  $T$ 마다 이전 속도  $v_0$ 에  $a$ 와  $T$ 를 곱하여 변화된 속도를 구한다.

$$v = v_0 + aT \quad (1)$$

차량의 위험 반경  $r$ 의 경우, 고정된 값을 사용했던 이전 모델들과는 달리  $T$ 초 동안 이동하는 거리를 반영하여  $v$ 와  $T$ 의 곱으로 계산한다.

$$r = vT \quad (2)$$

실제 사고 위험을 판단하기 위해서는 두 UE 사이의 거리  $D$ 와 각 UE의  $r$ 값을 더한 값의 크기를 비교해야 한다.  $D > r_{sum}$ 인 경우, UE 간의 거리가 멀기 때문에 위험하지 않다고 판단하여 경고 알림 대상에서 제외한다.  $D < r_{sum}$ 인 경우 UE의 진행 방향에 따라 사고 발생의 가능성을 추가로 판단하여 경고 알림 대상으로 선정한다. 본 연구에서는 2차원 좌표의 중심에 교차로가 존재하는 골목길에서의 측면 직각 충돌을 고려하므로, UE가 접근할 수 있는 위치는  $xy$  평면의 1-4사분면이며, UE가 교차로에 접근하는 이동 각도는 오른쪽( $0^\circ$ )부터 반시계 방향으로 각각  $90^\circ, 180^\circ, 270^\circ$ 라고 설정하였다. 또한, 측면 충돌하는 UE의 사고 발생 가능성만을 고려하므로, 두 UE의 이동 각도가 평행한 경우는 고려하지 않는다. 따라서 UE가 원점에 가까워지는 형태로 이동하며, 두 UE가 이동 각도가 평행하지 않은 경우에는 사고 발생 가능성이 있다고 판단한다. 예를 들어, UE1이 1사분면 쪽 도로에서  $270^\circ$ 의 방향으로 교차로에 접근할 경우,  $x$  좌표가 음수인 위치에 존재하는 UE2가  $180^\circ$ 의 방향으로 교차로에 접근할 때에는 해당 UE의 사고 발생 가능성을 고려할 필요가 없다.

기지국은 사고 발생 가능성이 있는 UE에 대해서만 경고 알림을 전송하며 OBU가 내장된 UE의 경우와 그렇지 않은 경우를 구분하여 UE나 mobile device에 경고 알림을 전송한다.

## 4. 시뮬레이션 및 평가

### 4.1.1 시뮬레이션 개요

본 연구에서는 사고가 자주 발생하는 십자 형태의 이면 도로에서의 상황을 가정해 측면 충돌이 일어나는 시나리오를 생성하였다. 충돌 기준이 되는 차량은 경로를 고정

시켜 수직 방향으로 주행하였고, 기준 차량 외의 UE는 가능한 도로 범위 내에서 4개의 방향 중 어느 방향으로든 주행할 수 있도록 초기값을 설정하였다. 속도와 가속도는 매초 범위 내에서 랜덤하게 생성하여 다른 UE들의 경로를 설정한 후 기준 차량과의 충돌을 판단하였고 해당 시나리오에서의 실제 충돌 여부와 본 예측 시스템(속도·가속도 고려 동적 알고리즘)이 위험을 감지하여 보낸 알림의 정확도를 비교하였다. 또한 같은 시나리오 내에서, 기존 선행되었던 연구의 알고리즘 정확도를 계산해 비교해 보았다.

### 4.1.2 비교 기법

속도·가속도를 고려한 동적 알림 알고리즘을 제안하면서, 기존의 두 알고리즘인 위험 반경 고정 방식과 사용자 타입 구분

방식을 같이 비교 분석한다.

#### 1) 위험 반경 고정 알고리즘

UE의 종류와 관계없이, 모든 위험 반경을 10m로 설정한다. 범위를 크게 설정하여 miss를 줄이기 위해서 흔히 고안되는 방법이다.

#### 2) UE 타입 구분 알고리즘

보행자와 차량의 위험 반경을 각각 4m, 8m로 고정해서 경고 알림을 주는 알고리즘이다.

## 4.2 시뮬레이션 환경 설정

표 2. 시뮬레이션 설정 값

	고정값	범위
$T$	0.5, 1, 2	-
$a_v$	-	-8 ~ 3
$a_p$	-	-1 ~ 1
$v_0$	8	5 ~ 16
$v_0^p$	1	0.8 ~ 1.3
도로 길이(m)	60	-
폭(m)	10	-

차량과 보행자의 속도는 각각 평균 8m/s와 1m/s이며 표준편차가 1m/s와 0.1m/s인 정규 분포의 확률을 따른다. 나머지 설정값들은 표 1을 따른다.

시나리오 당 정확한 결과값을 산출하기 위하여 1,000번의 시뮬레이션을 반복 진행하였고, 1,000개의 산출된 결과값을 평균하여 값을 나타내었다.

## 4.3 시뮬레이션 평가

사고 위험 예측 모델들을 평가하기 위해 실제로 시뮬레이션 상에서 사고가 일어나기 전까지의 위험 반경과 알림의 유무를 지속적으로 탐지한다. 사고가 실제로 일어났지만 최근  $T \times 2$ 의 주기 동안 아무런 알림이 발생하지 않았다면 miss로, 알림이 울렸지만  $T \times 2$ 의 주기 이전부터 알림이 울린 경우를 포함하여 사고가 실제로 일어나지 않았지만 알림이 울린 경우를 false로 설정하였다.

## 4.4 시뮬레이션 결과

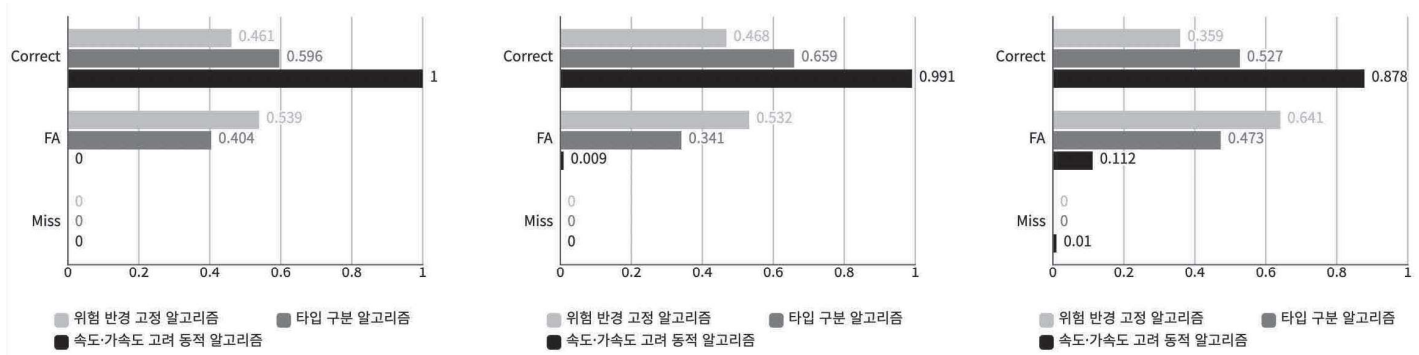


그림 2. 시뮬레이션 측정값

(1) T = 1일 때

그림 2를 보면, 다른 알고리즘들에 비해 속도·가속도 고려 동적 알고리즘이 99.1%의 높은 정확도를 가진다. FA 비율도 0.9%로 거의 나오지 않는 것으로 확인된다. FA 비율이 낮다는 것은 위험 반경이 최소한으로 설정된 것으로 FA 지표가 정확하지만 miss 값이 발생할 수 있다. 하지만 본 시뮬레이션에서는 miss 값 또한 발생하지 않아 다른 알고리즘들에 비해 우수하다고 할 수 있다.

(2) T = 2일 때

자원 할당 주기가 길어짐에 따라, T = 1일 때에 비해 속도·가속도 고려 동적 알고리즘의 정확도가 87.8%로 감소하게 된다. miss 값도 1%로 미비하지만, 처음으로 발생하게 되는 구간이다. FA 측면에선 기존 알고리즘들보다 우수한 결과를 보이지만, miss 값은 사고로 직결되기에 FA 값보다 민감한 판단이 필요하다.

(2) T = 0.5일 때

빠르게 차량의 현재 상태가 업데이트되기 때문에 동적으로 위험을 예측하는 속도·가속도 고려 동적 알고리즘이 100%의 확률로 우수한 정확도를 가진다. 위험 반경 고정 알고리즘과 UE 타입 구분 알고리즘은 실시간 정보에 따라 위험 반경이 설정되는 것이 아니므로 T 값이 감소했다고 해서 성능이 더 개선되지 않았다. 오히려 긴급하게 속도와 가속도가 바뀌는 시나리오에서는 더 성능이 떨어짐을 볼 수 있었다.

5. 결론

본 연구에서는 실시간으로 1초간 자원 할당이 보장되는 5G 환경에서 사용자의 위치, 속도, 가속도 값을 모니터링하며 동적으로 바뀌는 상황에서의 충돌 알림 방식을 고안하였다. 다른 알고리즘과 비교하여 시뮬레이션한 결과 FA의 비율이 다른 모델들에 비해 현저히 낮았고, miss는 거의 발생하지 않았다. 따라서 FA로 인한 사용자의 경각심 저하에 대한 문제도 해결될 것으로 생각한다.

사 사 문 구

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음”(2021-0-01082)

참 고 문 헌

[1] 김상욱, 전제호, 김인석, 임채홍, “국내 도시부 교통사고 특성 분석”, 교통기술과정책, pp. 53-59, 2017

[2] 김범성, 최배훈, 김은태, 고창성, 황재호, “전방 차량 정보를 이용한 충돌 위험도 판단 알고리즘”, 한국자동차공학회 춘계학술대회, pp. 925-927, 2013

[3] 정수현, 김주희, 안창선, “V2V기반 위험 차량 판별 및 충돌 방지 알고리즘”, 한국자동차공학회논문집, 26(6), pp. 773-782, 2018

[4] 공규식, 박다연, 정재일, “신호등이 없는 교차로 환경에서 차량용 에플레이터를 활용한 유럽 V2V 기반의 차량 충돌 방지 Safety Application에 관한 연구” 한국통신학회 학술대회는문집, pp. 1380-1381, 2017

[5] 황준연, 이상엽, 김형기, 이형직, 홍대건, “보행자 보호를 위한 충돌 예측 모델 개발”, 한국자동차공학회 춘계학술대회, pp. 745-746, 2014.



# 실시간 주문형 디지털 콘텐츠의 DRM 패키징 방법

신 서 윤°, 김 순 철

대구대학교 컴퓨터정보공학부  
keffred@daegu.ac.kr, kimsc@daegu.ac.kr

## DRM Packaging For Real-time On-demand Digital Contents

Seoyun Shin°, SoonCheol Kim

Dept. of Computer Engineering, Daegu University

### 요 약

본 논문은 실시간 주문형 디지털 콘텐츠의 DRM 패키징 방법과 이를 실행하기 위한 컴퓨팅 장치를 설명한다. DRM 서버에서 수행되는 실시간 주문형 디지털 콘텐츠의 DRM 패키징 방식은 총 3단계로 이루어지며 라이선스 서버로부터 사용자 단말에 대한 라이선스 유효성 검증 결과를 수신하는 단계, 해당 사용자 단말에 대한 라이선스가 유효한 것으로 판단되는 경우, 해당 DRM 서버의 비밀키에 기초하여 콘텐츠 복호화 키를 암호화하는 단계, 암호화된 콘텐츠 복호화 키를 해당 라이선스 서버로 송신하는 단계로 나누어진다.

### I. 서 론

최근 IT의 패러다임이 인터넷 중심의 e-society 사회에서 인간, 사물, 컴퓨터가 융합되는 u-society 사회로 급격하게 변해가고 있으며 그 중심에 실시간 주문형 디지털 콘텐츠 서비스 기술이 있다. 실시간 주문형 디지털 콘텐츠 서비스 기술은 디지털 융합(convergence) 기술 추세에 따라 장치, 네트워크, 서비스의 통합을 가속화시켜, 음악, 영화, AR/VR 등 다양한 콘텐츠를 생산할 수 있는 핵심 기술로서 빠르게 성장 중인 국내 온라인 콘텐츠 시장에서 다양한 콘텐츠 분야에 활용될 수 있는 기술이다. 실시간 주문형 디지털 콘텐츠는 시청자가 미처 제시간에 보지 못한 TV 프로그램이나 영화 등을 원하는 때에 볼 수 있도록, 서버에 저장되었다가 시청자의 주문에 따라 실시간으로 사용자 단말로 제공되는 콘텐츠를 의미한다. 실시간 주문형 디지털 콘텐츠가 보편적인 시청각 콘텐츠의 일반적인 형태로 자리 잡음에 따라, 시청자는 콘텐츠를 일방적으로 받아들이는 수용자의 위치에서 능동적으로 변한 소비자가 되어, 시청각 콘텐츠 분야에서 전환이 일어났다.

#### 1.1. 연구의 필요성

이러한 실시간 주문형 디지털 콘텐츠는 각 콘텐츠당 비용을 내거나, 소정 동안의 이용료를 내는 방식으로 누릴 수 있는데, 이때 콘텐츠 생산자와 유통사에서 정당한 대가가 돌아가게 하도록 이용 기간에 제한을 두고, 다운로드 및 불법복제를 방지하는 DRM 패키징 기술이 적용되는 것이 일반적이다. 그런데, 기존에 적용되는 DRM 패키징 기술은 디지털 콘텐츠에 대한 메타데이터를 암호화하지 않거나, 단순한 암호화 방식을 취하고 있어, 웹 페이지에서 곧바로 다운로드가 가능하거나, 크로미엄 기반의 다른 브라우저에서 확장 프로그램을 사용하여 손쉽게 다운로드 가능한 문제점이 있다.

#### 1.2. 연구의 목표 및 내용

본 논문에서 설명하는 DRM 패키징 방법에 따르면, 사용자 단말에서 직접 라이선스를 요청하기 위한 서명을 생성함으로써, 단순히 디지털 콘텐츠를 암호화하는 종래의 디지털 콘텐츠 보안 방식에 비해 디지털 콘텐츠의 무분별한 다운로드를 방지할 수 있다. 또한, 암호화된 디지털 콘텐츠의 재생에 필요한 복호화 키를 DRM 서버의 비밀키로 한 번 더 암호화함으로써, 종래의 암호화 키를 이용한 디지털 콘텐츠 보안 방식보다 효과적으로 디지털 콘텐츠의 불법 추출을 방지할 수 있다.

추가로 자칫하면 무거워질 수 있는 DRM 기능의 성능 이슈를 경량화된 패킷 캡슐화 및 인증서를 통해 해결하고자 한다.

### II. 선행 연구

선행 연구[1]의 DRM(Digital Rights Management) 기술 요약에 의하면 CP(콘텐츠 제작자) 서버에서 DRM 서버의 패키징 신호와 단말에서의 면허 열람 허가 신청을 모두 받아 비교하여, 양자 간의 콘텐츠가 일치되는 경우에만 열람 허가 신호를 사용자 단말에 전달하기 때문에, CP 서버에서 실시간으로 콘텐츠의 판매정보를 확인할 수 있게 되는 효과가 있다.

해당 개발된 프로그램의 선행문헌[1-청구항 6]에서는 해당 DRM 서버와 CP 서버에 소비자가 요청한 콘텐츠에 대한 패키징 신호를 전송하고, 해당 CP 서버는 라이선스가 요청된 콘텐츠와 해당 패키징 신호가 전송된 콘텐츠와 일치할 때 해당 콘텐츠에 대한 열람 허가 신호를 사용자의 단말에 전송하는 것을 특징으로 하는 DRM 시스템이다.

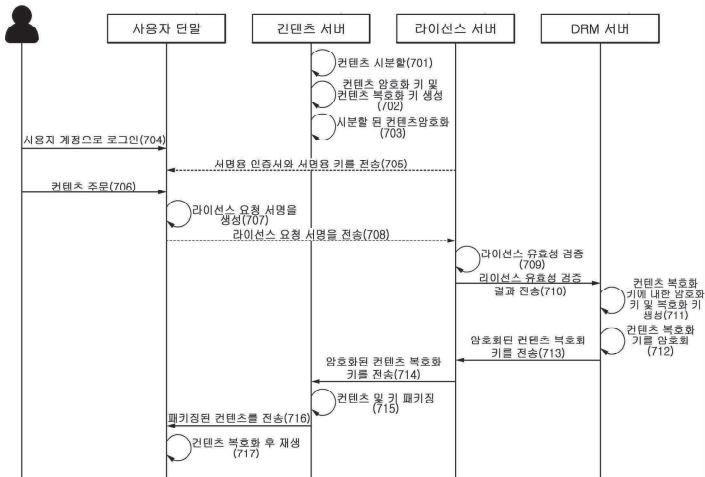
선행연구[1]은 DRM 콘텐츠를 제공하는 방법 내지는 시스템에 관한 것으로, 콘텐츠에 대한 라이선스를 요청받는 서버(CP

서버)와 콘텐츠에 대한 DRM 패키징을 수행하는 서버(DRM 서버)가 분리되어 있고, 패키징은 라이선스 정보, 사용자 단말의 기기 정보, 콘텐츠가 요청된 시간 정보 등을 이용하여 수행된다는 점 (선행문헌[1-청구항 7, 9] 참조)에서, 본 연구의 라이선스 서버와 DRM 서버가 분리된 구성, DRM 서버가 사용자 단말의 기기 정보를 고려하여 분할된 데이터에 접근제한을 설정하는 구성과 일부 유사하다.

그러나, 선행연구[1]은 데이터 복호화 키를 별도의 비밀키로 암호화하지 않고 있으며, 원본 데이터(원본 콘텐츠)를 나누어 각 분할된 데이터마다 다른 암호화 키로 암호화하는 구성도 전혀 개시하고 있지 않다. 또한, 선행연구[1]은 본 연구의 핵심적인 기술적 사상인 원본 데이터가 분할되어 각기 다른 암호화 키로 암호화되는 구성, 각 분할된 데이터의 데이터 복호화 키가 DRM 서버의 비밀키로 암호화되는 구성에 관해서 개시하거나, 시사, 암시하지 않고 있다.

### III. 제안 기법

[그림 1]은 본 논문에서 제안하는 DRM 패키징 시스템의 전체적인 흐름을 보여주고 있다.



[그림 1 : DRM 패키징 시스템 흐름도]

콘텐츠 제공 서비스를 이용하고자 하는 사용자는 사용자 단말을 통해 자신의 사용자 계정으로 콘텐츠 제공 서비스에 로그인한다. 사용자 계정에 대한 로그인이 수행되면, 라이선스 서버는 사용자 계정이 로그인 된 사용자 단말로 서명용 인증서 및 서명용 키를 전송한다. 그 후 사용자는 디지털 콘텐츠를 주문하고, 이에 따라 사용자 단말은 수신한 서명용 인증서 및 서명용 키에 기초하여 라이선스 요청 서명을 생성하고, 생성된 라이선스 요청 서명을 라이선스 서버로 전송한다.

이어서, 라이선스 서버는 수신한 라이선스 요청 서명에 기초하여 사용자 단말에 대한 라이선스의 유효성을 검증한다. 라이선스 서버는 DRM 서버로 사용자 단말에 대한 라이선스 유효성 검증 결과를 전송한다.

DRM 서버는 사용자 단말에 대한 라이선스가 유효한 것으로 판단되는 경우, 주문된 디지털 콘텐츠의 콘텐츠 복호화 키에 대한 DRM 서버의 공개키 및 비밀키를 생성한다. DRM 서버는 콘텐츠 복호화 키에 대한 암호화 키 및 복호화 키 생성단계에서 생성된 비밀키에 기초하여 콘텐츠 복호화 키를 암호화한

다.

DRM 서버는 암호화된 콘텐츠 복호화 키를 라이선스 서버로 전송하고, 라이선스 서버는 이를 수신하여 콘텐츠 서버로 전송한다. 콘텐츠 서버는 시분할 된 디지털 콘텐츠 중 사용자가 주문한 디지털 콘텐츠와, 콘텐츠 복호화 키에 대한 암호화 키 및 복호화 키 생성단계에서 생성된 공개키 및 암호화된 콘텐츠 복호화 키를 패키징하고, 패키징 된 콘텐츠를 사용자 단말로 전송한다.

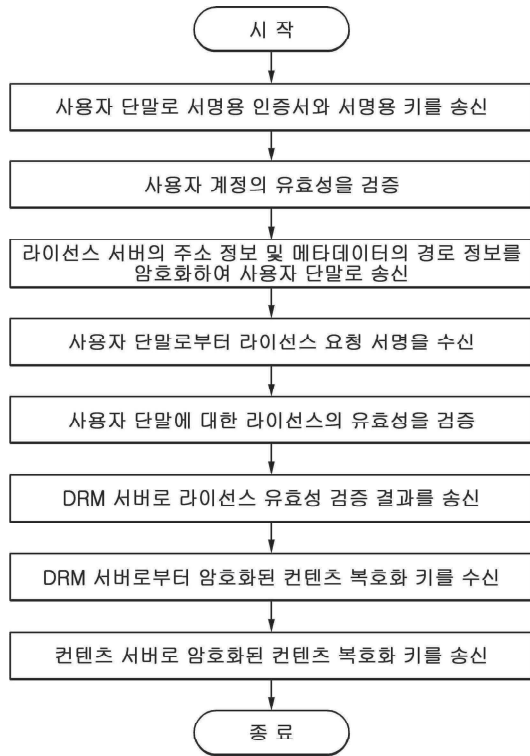
사용자 단말은 콘텐츠 복호화 키에 대한 암호화 키 및 복호화 키 생성단계에서 생성된 공개키로 암호화된 콘텐츠 복호화 키를 복호화하고, 이를 통해 획득한 콘텐츠 복호화 키를 이용하여 암호화된 디지털 콘텐츠를 복호화한 이후, 이를 재생한다.

사용자 단말은 디지털 콘텐츠에 적용된 암호화를 해제(복호화)하는 데 실패한 경우, DRM 조건을 만족하지 못하여 디지털 콘텐츠로의 접근이 만료된 것으로 판단할 수 있다. 이 경우, 사용자 단말은 사용자에게 다시 콘텐츠를 주문해 줄 것을 요청하는 알림 또는 메시지를 제공할 수도 있고, 사용자의 주문을 별도로 요청함이 없이 자동으로 라이선스 요청 서명을 생성하는 단계부터 반복할 수 있다.

[그림 2]는 추가적인 실시 예에 따라 라이선스 서버에서 수행되는 실시간 주문형 디지털 콘텐츠의 DRM 패키징 방법을 설명하기 위한 순서도이다.

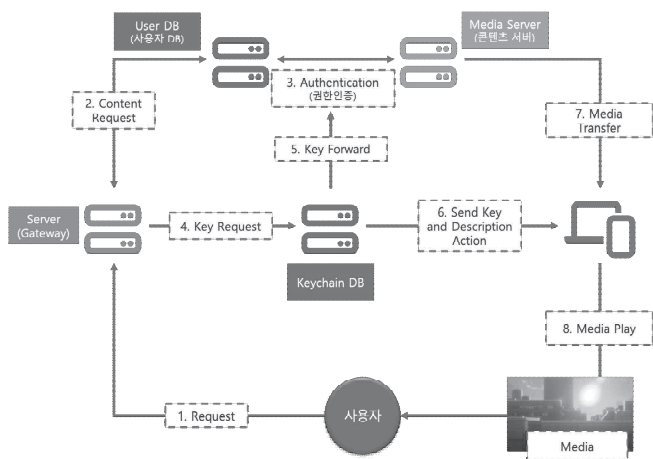
라이선스 서버는 콘텐츠 제공 서비스 플랫폼에 로그인한 사용자 단말로 서명용 인증서와 서명용 키를 송신한다. 라이선스 서버는 콘텐츠 제공 서비스 플랫폼에 로그인한 사용자 계정의 유효성을 검증한다. 이후, 라이선스 서버는 사용자 계정의 유효성이 검증된 경우, 라이선스 서버의 주소 정보 및 사용자 단말이 요청하는 콘텐츠에 대한 메타데이터(metadata)의 경로 정보를 암호화하여 사용자 단말로 송신한다.

라이선스 서버의 주소 정보 및 메타데이터의 경로 정보는 AES-256 방식에 기초하여 암호화될 수 있다. 암호화된 라이선스 서버의 주소 정보 및 암호화된 메타데이터의 경로 정보는 json 형식의 인코딩 파일로 사용자 단말에 제공될 수 있다. 라이선스 서버는 사용자 단말로부터 서명용 인증서와 서명용 키에 기초하여 생성된 라이선스 요청 서명을 수신하며 수신한 라이선스 요청 서명에 기초하여 사용자 단말에 대한 라이선스의 유효성을 검증한다. 또한, 라이선스 서버는 DRM 서버로 사용자 단말에 대한 라이선스 유효성 검증 결과를 송신하며 DRM 서버로부터 DRM 서버의 비밀키에 기초하여 암호화된 콘텐츠 복호화 키를 수신한다. 라이선스 서버는 콘텐츠 서버로 암호화된 콘텐츠 복호화 키를 송신한다. 이후, 라이선스 서버는 암호화된 콘텐츠 복호화 키와 함께, 해당 암호화된 콘텐츠 복호화 키를 복호화하기 위한 DRM 서버의 공개키의 유효기간에 대한 정보를 송신할 수 있다. 이 경우, 콘텐츠 서버는 암호화된 콘텐츠 복호화 키 및 DRM 서버의 공개키의 유효기간에 대한 정보를 사용자 단말로 전달(전송)할 수 있다.



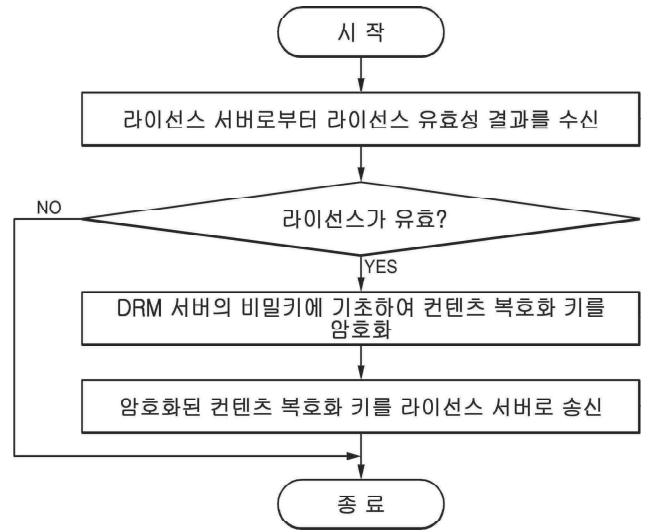
[그림2 : 라이선스 서버에서 수행되는 시스템 흐름도]

[그림3]에서 콘텐츠 서버는 콘텐츠 서버 내 저장 공간 또는 콘텐츠 서버와 연동된 별도의 데이터베이스에 저장된 디지털 콘텐츠들을 기 설정된 시간 간격으로 시분할하고, 시분할 된 디지털 콘텐츠들에 대한 콘텐츠 암호화 키 및 콘텐츠 복호화 키를 생성하고, 생성된 콘텐츠 암호화 키에 기초하여 시분할 된 디지털 콘텐츠들을 암호화한다. 이후, 라이선스 서버에서 암호화된 콘텐츠 복호화 키를 전송받으면, 콘텐츠 및 키를 패키징하여 사용자 단말로 전송한다.



[그림3 : 콘텐츠 서버 관점의 시스템 흐름도]

[그림 4]은 DRM 서버에서 수행되는 실시간 주문형 디지털 콘텐츠의 DRM 패키징 방법을 설명하기 위한 순서도이다.



[그림4 : DRM 서버에서 수행되는 시스템 흐름도]

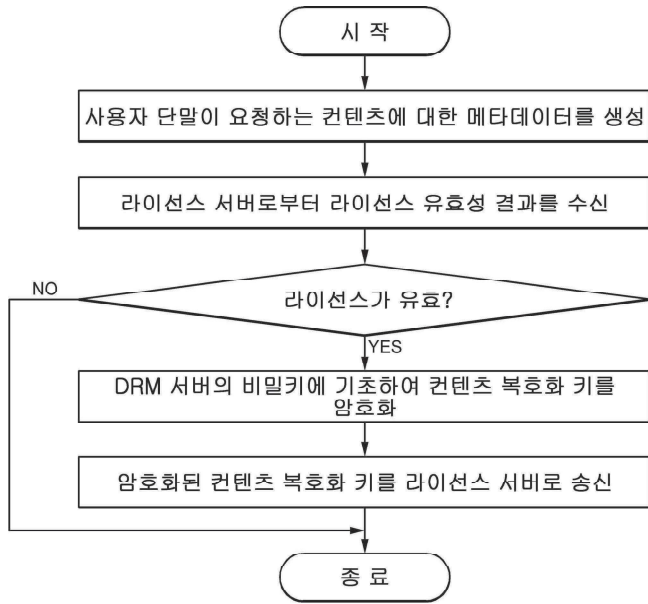
DRM 서버는 라이선스 서버로부터 사용자 단말에 대한 라이선스 유효성 검증 결과를 수신하며 라이선스 서버로부터 라이선스 유효성 검증 결과뿐만 아니라, 사용자 단말의 기기 정보, 사용자 계정의 식별 정보 및 사용자 단말이 요청하는 콘텐츠의 식별 정보 중 적어도 하나를 추가로 수신할 수 있다.

이후, 사용자 단말에 대한 라이선스가 유효한 것으로 판단되는 경우, DRM 서버는 DRM 서버의 비밀키에 기초하여 콘텐츠 복호화 키를 암호화한다. DRM 서버는 라이선스 서버로부터 라이선스 유효성 결과를 수신함에 앞서서 사전에, 콘텐츠 복호화 키의 암호화 및 복호화를 위한 공개키 및 비밀키 쌍을 생성할 수 있다. DRM 서버는 라이선스가 유효한지 확인하는 과정에서 라이선스가 유효한 것으로 판단되는 경우, 비로소 콘텐츠 복호화 키의 암호화 및 복호화를 위한 공개키 및 비밀키 쌍을 생성할 수 있다.

DRM 서버는 콘텐츠 복호화 키의 복호화를 위한 공개키를 콘텐츠 서버 및 라이선스 서버에 바이너리 데이터(binary data) 형식으로 제공할 수 있다. 이를 통해, 공개키가 공유되는 과정에서 DRM 패키징 시스템 외부로 유출되는 것을 방지할 수 있다. DRM 서버는 콘텐츠 복호화 키를 암호화하되, 사용자 단말의 기기 정보, 사용자 계정의 식별 정보 및 사용자 단말이 요청하는 콘텐츠의 식별 정보 중 적어도 하나에 기초하여 콘텐츠 복호화 키의 복호화에 대한 접근 제약을 설정할 수 있다.

DRM 서버는 해당 주문된 콘텐츠에 있어서, 해당 사용자 단말의 기종과 상이한 기종의 단말을 통해 콘텐츠 복호화 키의 복호화를 할 수 없도록 설정할 수 있다. DRM 서버는 사용자 계정의 식별 정보를 이용하여 사용자 계정을 식별하고, 콘텐츠의 식별 정보에 기초하여 주문된 콘텐츠를 식별할 수 있다. DRM 서버는 해당 주문된 콘텐츠에 있어서, 해당 사용자 계정을 통해 콘텐츠 복호화 키의 복호화가 가능한 기간의 제약을 설정할 수 있다. 이후, DRM 서버는 암호화된 콘텐츠 복호화 키를 라이선스 서버로 송신한다.

[그림5]는 DRM 서버에서의 추가적인 보안 시스템 흐름도를 보여주고 있다.



[그림 5 : DRM 서버에서의 추가적 보안 시스템 흐름도]

추가적인 보안을 위해서 라이선스 유효성 결과를 수신하기 전 단계에서 DRM 서버는 HLS 포맷 또는 DASH 포맷에 기초하여 사용자 단말이 요청하는 콘텐츠에 대한 메타데이터를 생성한다. 그후, DRM 서버는 라이선스 서버로부터 사용자 단말이 요청한 콘텐츠의 복호화를 위한 콘텐츠 복호화 키를 수신하여 사용자 단말에 대한 라이선스가 유효한 것으로 판단되는 경우, DRM 서버는 DRM 서버의 비밀키에 기초하여 콘텐츠 복호화 키를 암호화한다. 이후, DRM 서버는 암호화된 콘텐츠 복호화 키를 라이선스 서버로 송신한다.

#### IV. 결 론

본 실시간 주문형 디지털 콘텐츠의 DRM 패키징 방법과 이를 실행하기 위한 컴퓨팅 장치를 연구 및 개발하여 얻은 결과는 다음과 같다.

1) 기 DRM 시스템들은 사용자 단말이 사용자 데이터에 직접 서명하여 라이선스 검증을 요청하지 않고, 데이터 복호화 키를 별도의 비밀키로 암호화하지 않고 있다.

2) 해당 보안 문제를 해결하기 위해 해당 라이선스 서버의 주소 정보 및 해당 사용자 단말이 요청하는 콘텐츠에 대한 메타데이터(metadata)의 경로 정보를 암호화하여 사용자 단말로 송신하는 단계를 더 포함한다.

3) 해당 프로그램은, HLS 포맷 또는 DASH 포맷에 기초하여 해당 사용자 단말이 요청하는 콘텐츠에 대한 메타데이터(metadata)를 생성하는 단계를 실행하기 위한 명령어들을 더 포함할 수 있다.

결론적으로, 이러한 저작물 불법 유통사태를 막기 위해서는 저작권보호 기술들의 균형 있는 활용 및 발전이 필요하다. 디지털 콘텐츠 시장은 방송과 통신의 융합, 무선통신 환경의 발달, 휴대용 멀티미디어 기기의 융합 가속화로 인해 빠르게 성장하고 있으며, 소비자들이 디지털 콘텐츠를 구매할 수 있는 채널들이 지속해서 확대되고 있어 미래 디지털 콘텐츠 산업의 지식재산권 보호를 위해 어떠한 저작권보호 기술들이 사용되고, 표준화되어 뿌리를 내릴지 모르기 때문이다[2]. 이러한 변

화된 환경에 대한 대처 차원에서라도 디지털 콘텐츠 시장의 성장 단계에 따른 DRM 기술의 개발 및 표준화 작업에 대한 정부의 지원 및 관련 업계의 관심이 필요하다.

#### 참 고 문 헌

- [1] 대한민국 공개특허공보 제 2010-0011223호 (2011.08.11. 공개)
- [2] DRM 시장의 현황 및 문제점 - 한국저작권위원회 기술연구소 김현미 선임
- [3] 저작권보호센터, 2010 저작권 보호 연차보고서, 2010
- [4] 한국정보통신기술협회, 정보통신 중점기술 표준화로드맵 Ver.2010, 2010.2
- [5] 한국콘텐츠진흥원, 2009 해외콘텐츠시장조사, 2010
- [6] LG경제연구원, 공짜 늪에 빠진 디지털콘텐츠, 2010
- [7] 한국저작권위원회, DRM 호환성 해결을 위한 표준화 연구, 2009.12
- [8] 한국인터넷진흥원, 2009 국내 지식정보보안산업 시장 및 동향 조사, 2009.12
- [9] 황성운, 디지털 저작권 관리(DRM), 2008

# 대규모 미팅 행사를 위한 어플리케이션 구조 제안

이승민<sup>o</sup> 노준수 신지한 이수인 최혜림  
경북대학교 IT 대학 컴퓨터학부

milouw56@gmail.com, [wnstn82612@naver.com](mailto:wnstn82612@naver.com), [deliciousqoo@naver.com](mailto:deliciousqoo@naver.com), [chr121910@gmail.com](mailto:chr121910@gmail.com)

## Propose an application structure for large meeting events

Seung MinLee<sup>o</sup>, Joon Soo Roh, Jihan Shin, Lee, Hyerim Choi  
Department of Computer Science, College of IT Engineering, Kyungpook National University

### 요 약

조직의 이사회, 주주 총회 등 다양한 정보가 오가고, 의사결정 등이 필요한 대형 행사에서의 접수대 기능을 개선하여 감염과 혼잡을 개선하여 포스트 코로나 시대를 대비하고, 모바일을 통한 의사결정 및 자료 공유를 통해 정보화 시대에서의 더욱 효율적인 행사진행을 돕는 모바일 회원수첩을 개발한다.

으로 인한 안전사고 및 코로나 감염을 예방할 수 있을 것이다.

## 1. 서 론

### 1.1 연구 배경 및 목적

사회적 거리두기가 약화되며 최근 대규모 미팅 행사가 늘고 있다. 이러한 상황에서 대규모 행사의 질서 유지 및 통제, 의견 공유의 어려움에 관해 이야기하는 목소리가 점점 커지고 있다. 이에 요즈음 디지털로 변화되는 여러 문서 작업과 이제는 필수가 된 스마트폰 어플리케이션을 이용해 본 문제점을 해결하고자 한다.

대규모 행사는 여러 사람이 참여하는 행사인 만큼 질서 유지 및 통제가 어려운 것이 현실이다. 특히 접수대의 경우, 참석자 접수를 위해 사람들이 줄지어 기다리고, 행사장 안내 등 진행 과정에서 많은 사람이 오고 가는 곳이기 때문에 혼잡스러운 편이다. 마스크 착용을 의무화하여 감염을 예방하고는 있지만, 여러 사람이 모이게 되는 행사에서 감염을 온전히 피하기는 어려운 실정이다. 따라서 사람들이 한 곳에 집중적으로 모이는 상황을 최소화하고 혼잡을 줄이기 위해 접수대의 기능을 분산시킬 필요가 있다. 또한, 대규모 행사의 경우 의사 결정 및 투표 과정이 수기로 진행됨에 따라 단순 작업에 투입되는 시간과 인력이 많다. 이에 본 연구는 각자의 자리에서 의견을 손쉽게 공유하고 취합할 수 있는 모바일 작업(어플리케이션)을 활용하여 이러한 문제점을 해결하고자 한다.

따라서 본 연구에서는 대규모 미팅 행사를 위한 어플리케이션을 제안한다. 이는 대규모 미팅 행사에서 이루어지는 업무를 파악하고 해당 업무를 효율적으로 해결하기 위한 모바일 작업을 연구하여 대규모 미팅 행사를 위한 어플리케이션의 구조를 제시할 것이다. 이를 통해 행사 시 정보 전달 및 접수 업무를 스마트폰 작업으로 변경하여 사람이 한곳에 모이는 일을 방지하고, 개인의 의견을 손쉽게 얻고자 한다.

본 연구를 통해 기존에 대규모 미팅 행사에 들어가는 비효율적인 인력과 시간 소모를 줄이고 원활한 행사 진행을 이룰 수 있을 것이다. 나아가 진행 및 관리 비용을 줄이고, 대규모 인원

### 1.2 연구 목표

모바일 회원 수첩 “moaDo”를 서비스 중인 ‘플래시 21’과 협력하여 기존에 제공 중인 회원 정보, 공지사항 알림, 일정 관리 등의 기능에 행사 안내 기능을 더하여 조직의 행사 진행과 인력 및 시간 관리의 효율성을 높인다. 행사 안내 기능은 참석자 등록 및 확인 기능, 의사 결정 시스템, 자료공유 기능 등을 포함한다. 이러한 차원에서 본 연구가 가지는 목적은 다음과 같다.

첫째, 대규모 행사 진행 과정 중 디지털화가 가능한 작업을 분류하고 각 작업의 요구사항을 분석해 최적의 디지털화를 파악하고자 한다.

둘째, 위 과정을 바탕으로 도출된 최적의 디지털화에 대해 제시하고 이를 어플리케이션과 결합하고자 한다.

## 2. 본 론

### 2.1. 대규모 미팅 행사 작업 분석

대규모 미팅 행사 진행 시 필요한 업무는 다음과 같다. 우선 행사에 참석한 인원을 파악하기 위해 참석자 접수 업무를 필요로 한다. 다음으로, 행사를 진행하는데 필요한 자료를 배포하는 업무가 있다. 마지막으로, 안전에 대한 투표 등의 의사결정 업무가 있다. 본 과제에서는 기존에 아날로그 방식으로 이루어지던 행사 접수, 의사결정 방식, 자료 공유 업무를 디지털화하여 행사 진행의 편의성을 개선하고자 한다.

### 2.2. 작업 디지털화

#### 2.2.1. 접수 디지털화

첫째로 행사 참석자 접수 과정을 디지털화할 필요가 있다. 기존에 사용해오고 있는 참석자 접수 방식으로는 수기 명단 작

성, 진행요원이 호명하는 방식 등이 있다. 먼저 행사장 내에 마련된 접수대에서 수기 명단을 작성할 경우, 접수대에 많은 인원이 몰려 혼선이 생길 수 있다. 진행요원을 배치하여 혼선을 방지하여도 참석자들이 기다려야 한다는 불편함이 존재한다. 두 번째로 진행요원이 한 사람씩 호명하여 참석을 확인할 경우, 참석자는 자신의 이름이 호명될 때까지 기다려야 하는 번거로움이 존재한다. 이 과정에서 인원수에 비례하는 시간이 소모된다. 또한 참석자가 일시적으로 자리를 비울 경우 접수 정확도가 떨어진다. 따라서 기존의 접수 방식에서 참석자의 편의성, 접수현황의 정확성, 접수 과정에서 소모되는 인력을 고려하여 디지털화하고자 한다.

디지털 접수 방법으로 세 가지 방법의 접수 방법을 고안하였다. 첫 번째는 행사장의 입구에 QR코드를 두고, 입장하는 참석자가 개인 모바일 앱으로 해당 QR코드를 스캔하도록 하는 방법이다. 두 번째는 애플리케이션에 숫자 또는 문자 코드를 입력하는 방법이다. 마지막으로 행사 진행 요원이 각 참석자를 확인하고 애플리케이션의 참석 버튼을 눌러 접수해주는 방법이 있다.

세 방법은 서로 다른 장단점을 가진다. 첫 번째 방법은 참석자가 행사장에 입장하며 접수하기 때문에 다른 방법에 비해 정확한 현황 파악이 가능하다는 장점이 있다. 하지만 한정된 QR코드를 스캔하기 위해 참석자들이 특정 장소에 몰리면서 입장에 혼선이 생길 수 있다는 단점이 존재한다.

두 번째 방법은 진행요원이 앱을 통해 생성한 코드를 개인이 휴대한 기기로 애플리케이션의 코드 입력란에 입력하여 접수하기 때문에 행사장 내에서 혼란스러운 상황이 발생할 가능성이 매우 낮아진다. 하지만 코드만 알고 있다면 행사에 참여하지 않고도 접수할 수 있으며, 입력 시 개인의 실수 등으로 현황 파악의 정확도가 떨어질 수 있다.

세 번째 방법은 애플리케이션 사용이 어려운 참석자에게 부담이 적으며, 진행 요원이 확인한 후 접수하기 때문에 접수현황 파악의 정확도 또한 높다. 하지만 앞의 두 방법에 비해 추가 인력이 필요하며, 인력이 부족할 경우 시간이 소모된다는 단점이 있다.

행사 참석 인원수, 행사 장소의 규모 등에 따라 적절한 접수 방법은 다를 것이다. 따라서 본 과제에서는 세 가지 방법을 모두 적용한 애플리케이션을 통해 행사를 시행하는 집단에서 선호하는 접수 방법을 선택할 수 있도록 하고자 한다.

## 2.2.2. 투표 디지털화

마지막으로 행사/회의를 진행하는 데 있어서 필요에 의해 시행되는 안전에 대한 투표를 디지털화하는 과정이 필요하다. 현재 국회나 이사회, 모임 등에서 일부 안전에 대한 투표를 진행할 때 투표 결과 수합을 거수하는 방식으로 많이들 진행한다. 이는 일정 인원 이상이 되었을 경우 의결 결과를 수합하는데 시간적 손해와 번거로움을 유발한다. 본 연구자는 이를 디지털화하여 조직의 비효율적인 인적 관리 해결과 불필요한 시간 소모를 없앨 수 있도록 할 것이다. 우선 투표가 진행되는 방식은 단체 대화방을 통해 이루어지는 투표의 구조와 유사하다. 회원 수첩 앱에서 투표를 등록하고, 참여자들은 자신의 모바일 기기를 활용하여 투표를 진행하게 된다. 그리고 해당 투표의 결과는 진행자와 참여자가 모두 확인할 수 있다.

투표의 디지털화가 실현되었을 경우 얻을 수 있는 이점은 다음과 같다[1]. 첫째로 투표를 진행하고 결과를 수합하는 과정에서

최소의 인력과 비용으로 개표가 가능하다. 두 번째로 디지털화된 투표의 경우 장소에 구애받지 않고도 투표를 진행할 수 있으므로 의결을 진행할 때 꼭 대면이 아니더라도 비대면으로도 의결을 진행할 수 있다. 마지막으로 디지털 투표의 특성상 무효표가 없으므로 투표율을 높일 수 있으며 회원들의 뜻을 정확히 반영하는 효과를 얻을 수 있다.

하지만 기존 거수 같은 아날로그 방식에서 디지털 투표로 바뀔 시에 고려해야 할 사항들이 있다. 우선 디지털 투표를 진행하는 해당 시스템에 문제가 생겼을 시엔 투표 결과에 혼란을 초래할 수 있으며 해당 디지털 서버를 구축하는 데에 초기 비용이 발생하고 추후 관리하는데도 비용이 소모된다. 또한 디지털 특성상 투표 결과가 조작되거나 해킹 등의 이슈가 발생할 수 있다. 또 노년층이나 정보화 소외계층에게는 디지털화 투표 방식이 익숙하지 않아 일부 사용자들에게 불편함을 유발할 수 있다.

이를 해결하기 위해선 디지털 투표의 신뢰성과 사용자들에게 친숙함을 제공할 수 있어야 할 것이다. 투표의 신뢰성은 투표자 개개인이 자기 뜻이 투표에 제대로 반영되었는지 알아야 하며 그 결과에 조작이나 해킹 등의 위험이 없음을 인지할 수 있어야 한다. 사용자들에게 친숙함은 투표하는 과정에서 그 과정이 복잡하지 않고 직관적이어야 하며 편리해야 할 것이다. 우리는 이를 고려하여 투표 작업을 디지털화할 것이다.

## 2.2.3. 자료 공유 디지털화

다음으로, 행사/회의를 진행하는데 필요한 자료 배포 업무를 디지털화하는 과정이 필요하다. 기존에 행사/회의를 진행할 때 행사자료를 대부분 주최 측에서 종이에 출력하여 준비해주는 방식으로 진행되는데 준비하는데 드는 경제적 비용이나 행사/회의를 진행하면서 종이에 출력된 자료의 효율성을 생각해보면 아날로그 방식으로 자료를 배포하는 방식이 그렇게 좋은 방식은 아님을 알 수 있다. 그래서 우리는 이를 디지털화하여 자료 배포에 드는 비용과 행사/회의를 진행하면서 자료를 활용하는 효율성을 높이고자 한다.

자료 배포 업무를 디지털화하는 것은 대화 앱의 단체 대화방 기능이나 현재 여러 학교에서 진행하고 있는 학습관리시스템을 예로 들 수 있다. 대화 앱의 단체 대화방에서 누군가가 자료를 게시하면 사람들은 터치 한 번으로 개인의 기기에 그 자료를 내려받을 수 있다. 또 학습관리시스템의 경우 해당 강의의 관리자가 자료실에 강의자료를 업로드하면 강의를 수강하는 학생들은 자료를 다운받아서 개인 기기에 자료를 확인하거나 웹페이지 상에서 바로 강의자료를 확인할 수 있다.

이처럼 자료를 제공할 참석자 또는 주최자가 회원 수첩 앱에 미리 자료를 업로드할 수 있고 참여자가 그 자료를 내려받을 수 있도록 하는 구조는 앞서 말했던 자료를 준비하는데 드는 경제적 비용과 행사/회의를 진행하는데 드는 효율성을 높일 수 있으므로 해당 업무를 디지털화하려고 한다.

## 2.3. 예상 UI

모바일 회원 수첩 앱을 실행하고 행사/회의 참여를 하기 위해선 [그림1]과 같은 시작 화면을 거치게 된다. 왼쪽 화면에서 먼저 개설된 행사/회의 항목을 선택하게 되면 오른쪽 화면이 표시된다.



[그림1] 앱에서 해당 회의 접속 화면

해당 행사/회의를 클릭 후 사용자는 회의에 관련된 다양한 기능을 사용할 수 있다. 구상했던 회의에 필요한 기능은 사용자가 해당 행사/회의의 개요를 알 수 있고 참석자 본인이 자신이 참석 여부를 체크할 수 있으며 자신 이외에 어떤 참석자가 참석했는지 확인할 수 있고 안전투표 및 행사/회의에 필요한 자료는 앱을 통해 공유받을 수 있다.

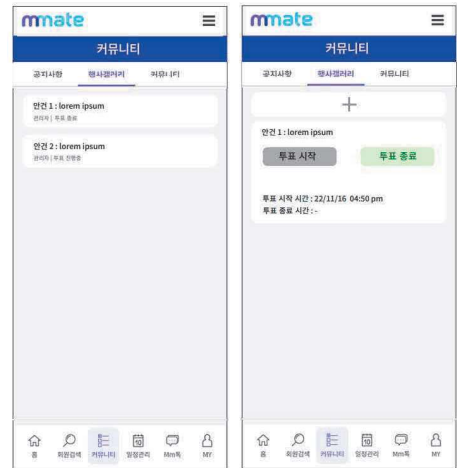
먼저 사용자가 행사/회의에 참석했을 때 참석 여부를 체크하기 위해 [그림1] 오른쪽 화면의 참석 체크 버튼을 클릭하였을시 [그림2] 왼쪽 화면으로 넘어가게 된다. 해당 화면에서는 QR코드를 찍거나 행사/회의에서 생성된 출석 코드를 입력하는 방식으로 사용자가 직접 자신의 참석 여부를 결정할 수 있도록 한다.

또 [그림1] 오른쪽 화면의 참석자 확인 버튼을 클릭하였을시 [그림2]의 오른쪽 화면으로 넘어가게 되어 자신 이외에 어떤 참석자가 참석했는지 확인할 수 있으며 회의를 주최한 관리자는 해당 페이지에서 다른 참석자들의 참석 여부를 직접 변경할 수 있도록 하여 앞서 2.2.1에서 말했던 접수 디지털화를 모두 가능하게 하였다.



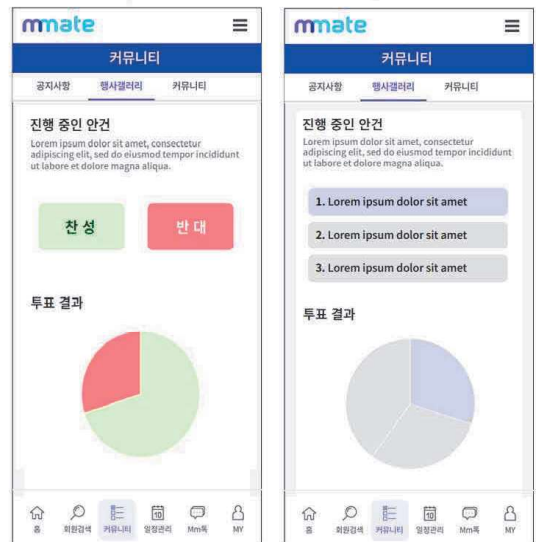
[그림2] 앱에서 디지털화 된 접수 기능 화면

안전투표 기능은 [그림1] 오른쪽 화면에서 안전투표 버튼을 클릭시 [그림3]의 화면으로 넘어가게 된다. 사용자는 어떤 안전이 진행중이고 완료되었는지 확인할 수 있으며 안전 제안자나 주최자는 [그림3] 오른쪽 화면과 같이 안전을 생성하여 투표를 진행하거나 투표를 종료시킬 수 있다.



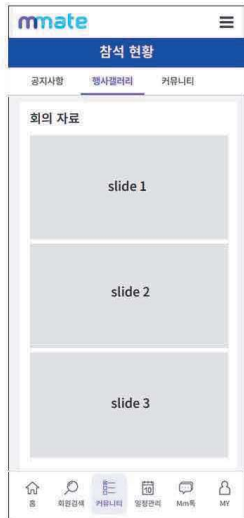
[그림3] 안전 확인 및 안전 생성 화면

해당 안전을 클릭시 [그림4]에서 보이는것 처럼 진행중인 안전에 대해서 사용자는 찬성/반대 또는 제시되어 있는 선택지 중에서 선택할 수 있으며 투표결과를 바로 직관적으로 확인할 수 있어 앞서 2.2.2에서 말했던 투표의 디지털화를 이룰 수 있다



[그림4] 안전 투표 및 결과 확인 화면

마지막으로 2.2.3의 자료 공유의 디지털화는 [그림1] 오른쪽 화면의 행사 자료를 클릭시 해당 행사/회의 에서 진행중인 사안에 대한 자료가 업로드 되어있을 것이며 사용자는 해당 자료를 다운받아 [그림5]에서처럼 앱에서 바로 자료를 확인할 수 있을 것이다.



[그림5] 앱에서 디지털화된 자료공유 기능 화면

### 3. 결 론

#### 3.1 기대 효과 및 활용 방안

본 과제에서는 다양한 행사 현장에서 대규모 인원을 통제하는데 질서 유지 및 통제가 힘들고 주최측이 진행하는데 있어서 반복적인 업무와 인적오류들로 인한 번거로움을 모바일 수첩 앱을 통해 주최자 뿐만 아니라 행사/회의 참여자들에게도 기존의 방식을 디지털화하고 새롭게 추가된 기능을 활용하여 현장에서 이루어지는 행사/회의를 원활하게 진행할 수 있도록 하는 것에 초점을 맞추고 있다.

또한 행사 안내 및 정보 전달의 역할도 어플리케이션에서 제공함으로써 참석자는 해당 행사에 대해서 좀더 쉽게 파악할 수 있고 행사에 대한 정보에 보다 쉽게 접근할 수 있으며 주최 측에서도 접수대에서의 혼잡도를 개선하고, 현장 인력을 좀 더 효과적으로 운용함으로써 경제적인 부분도 개선될 수 있을 것이다.

기존에 모바일 회원 수첩에서 제공하고있는 기능을 기반으로 본과제에서는 디지털 접수방식, 디지털 의결 투표기능, 디지털 자료공유 등의 디지털화 된 기능등을 추가하였다.

기존의 아날로그 방식으로 진행되던 접수기능은 사용자들은 자신이 행사/회의에 참석했는지에 대해 명확하게 기록을 남길 수 있으며 관리자도 참석자의 참석여부를 결정할 수 있게 하여 사용자와 관리자 둘다의 편리성을 높였다. 또한 의결상황에서도 좀더 명확하게 자신의 의견에 따라 결정할 수 있고 직관적으로 투표 결과를 바로 알 수 있게 되었으며 좀더 편리하게 행사/회의 때 사용되는 자료에 접근할 수 있게 될 것이다.

이는 현대사회에서 모두가 각자의 모바일 기기를 소지하고 있고 경제성이 중요시 되는 사회인만큼 더 필요한 기능이라고 생각한다. 모든 과정에서 아날로그보다 더 빠르고 편리한 디지털화를 이루어내어 행사/회의를 진행함에 있어서 행사에 필요한 인적 자원을 줄이고, 진행 및 관리 비용을 줄이고, 행사의 빠르

고 정확한 진행을 도움으로써 경제성을 높일 수 있을거라고 기대한다.

이를 확실히 이루어내기 위해선 이를 뒷받침할 앱의 신뢰성과 친숙함을 만들어나가고 유지하는 것이 앞으로의 숙제일 것이다.

#### 사 사 문 구

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음”(2021-0-01082)

#### 참 고 문 헌

[1]전웅렬, 이운호, 원동호 “전자투표 시스템 실용화 현황과 전망” 정보보호학회지 제21권 제2호 제2호  
<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE01635202>



# 판매데이터와 공공데이터를 활용한 카페 상권분석 시스템

## 구축

김건준\*, 이상협, 조원영, 신지송, 박찬식

경북대학교 컴퓨터학부

[kgjsmart1559@gmail.com](mailto:kgjsmart1559@gmail.com), [leetkdguq73@naver.com](mailto:leetkdguq73@naver.com), [111\\_yeong@naver.com](mailto:111_yeong@naver.com), [songjin1127@naver.com](mailto:songjin1127@naver.com),

[cspmobilelab@naver.com](mailto:cspmobilelab@naver.com)

## Establishment of cafe commercial analysis system using sales data and public data

GeonJun Kim\*, Sanghyeop Lee, Wonyeong Jo, Jisong Shin, Chansik Park

Kyungpook National University School of Computer Science and Technology

### 요 약

코로나 19의 영향으로 많은 소상공인들의 피해가 커지고 있다. 본 연구에서는 대구 수성구의 여러 공공 데이터를 수집하고 가공하여 상권분석을 한 뒤, 이것을 시각화하여 제공한다. 더 나아가 ‘CSP 모바일 앱’의 카페 판매데이터를 활용하여 딥러닝을 통해 여러 정보를 제공하여 소상공인들에게 도움이 되는 플랫폼을 제시한다.

### 1. 서 론

코로나 19 사태가 확산 및 장기화되면서 국가 경제의 침체 또한 깊어지고 있으며, 그에 따라 국가 경제의 25%를 차지하는 소상공인들의 피해도 커지고 있다[1]. 중앙로역 상권은 대구에서 가장 큰 상권으로 ‘동성로’라는 이름으로 많이 알려져 있다. 중앙로역 음식점종의 상권을 분석해보면, 2019년에 운영되던 점포 수 평균은 710개였지만 2020년에 652개, 2021년에 623개로 줄어들어 운영 점포 수가 연간 8%, 4%씩 감소했고, 전체 매출 규모는 2020년 33%, 2021년 5%가 감소했다[2][표 1].

[단위 : 억 원, 개소, %]

구분	2019년	2020년	2021년(추정)	2019년 대비 2020년	2020년 대비 2021년(환산)
매출액 (기간 합산)	2,770	1,861	1,776	-32.8%	-4.6%
점포 수 (기간 평균)	710	652	623	-8.2%	-4.4%

<표 1. 중앙로역 상권 전체 음식점 매출/점포 수 변화 (출처:[2])>

이렇게 대구 상권의 중심지 또한 어려움을 겪는 상황에서 큰 타격을 입었을 대구의 소상공인들에게 도움이 되는 정보를 제공하고자 한다.

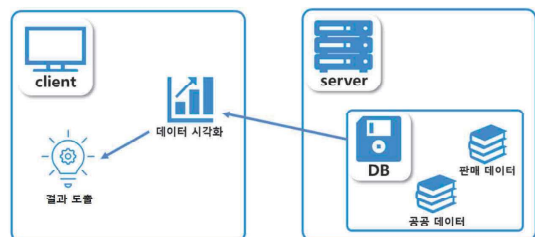
상권분석이란 ‘상업이 영향을 미치는 구역의 점포 수와 인구 분포, 교통량 등을 분석하는 일’로 이를 바탕으로 점포가 들

어설 곳을 결정한다[네이버 국어사전]. 본 연구는 대구 안에서 가장 많고 다양한 공공데이터를 제공하는 수성구를 선택해 상권분석에 도움이 되는 데이터들을 수집하고 분석한 정보를 시각화하여 제공한다.

더 나아가 ‘CSP 모바일 앱’의 카페 판매데이터를 활용하여 일별, 월별 카페의 추천 메뉴와 판매량을 예상하여 적절한 원재료 수급에 도움을 주고자 한다.

### 2. 본 론

#### 2.1. 시스템의 전체 구조



<그림 1. 상권분석 시스템 전체 구조 >

다양한 방법으로 공공데이터들을 수집하여 데이터들을 분석한

후 분석한 결과물을 데이터베이스에 저장한 뒤 저장된 데이터들을 서버를 통해 클라이언트로 전달한다. 차트와 같은 시각자료를 이용해 사용자들에게 정보를 더 효율적으로 전달한다[그림 1].

주 언어는 javascript 위주로, 서버는 node.js(express.js)와 mongoDB를 사용하여 구축하고, UI는 react.js를 활용할 예정이다. 추가적으로 공공데이터를 통하여 해당 가게의 매출을 분석하고 예측하는 모델을 구현한다. 이는 tensorflow를 이용해 딥러닝 기법으로 구현한다.

## 2.2. 수집한 공공데이터

수정구청에서 제공하는 공공데이터를 통해 동별 최근 3년간 개폐업 현황과 연간 평균 유동인구, 연령별, 성별별 인구현황 데이터를 수집하고 가공한다.

카카오맵 api와 python 언어, 셀레니움 프레임워크를 활용한 크롤링을 통해 현재 수정구에 위치한 가게 정보를 업종별로 수집하고 csv파일로 정리한다. 해당 파일에는 가게명, 도로명 주소, 지번 주소, 위도, 경도, 행정동 정보가 포함되어 있고, 이는 추후 지도에 표시할 때 사용된다.

	A	B	C	D	E	F	G	H	I	J	K
1 행정동	남대 바육(남/여 평균연령	연간 평균유동인구	평균영업기간(년)	평균영업기간(수)	평균영업기간(목)	평균영업기간(금)	평균영업기간(토)	평균영업기간(일)	평균영업기간(휴)	영업기간 평균 이상 업종	
2 고산1동	0.94	39.8	424.78	3.88	1.5	0.92	1.24	1.04	1.04	식육, 한식	
3 고산2동	0.99	41.6	145.38	1.41	5.21	0.5	1.13	1.26	일식, 유계		
4 고산3동	0.94	40.5	219.8	1.08	1.36	17.58	1.31	1.22	중식, 한식, 유계		
5 무산1동	0.95	42.2	259.21	0.9	1.37	1.28	1.21	1.2	유계		
6 만촌1동	0.96	42.5	108.46	3.46	0.81	2	0.98	1.35	식육, 유계		
7 만촌2동	1.02	43.4	87.61	1.67	1.08	2.17	0.94	1.21	유계		
8 만촌3동	1.09	39.2	262.57	3.17	0.75	1.5	1.21	1.13	음료		
9 평암1동	0.94	50.4	311.81	3.83	2.17	14.92	1.72	1.15	식육, 한식, 중식, 한식		
10 평암2동	0.95	42.5	213.06	1.58	1.75	21.54	1.29	1.34	중식, 한식, 유계		
11 병어1동	0.92	41.2	247.47	1.17	1.42	4.17	1.15	1.36	중식, 한식		
12 병어2동	0.93	41.4	114.44	1.23	1.38	3.02	1.19	1.08	유계		
13 병어3동	0.92	37.7	294.66	3.08	0.84	2	1.23	1.21	유계		
14 병어4동	0.99	57.5	103.26	3.73	0.99	1.1	1.10	1.21	식육, 유계		
15 상동	0.94	45.2	189.39	0.46	2.54	22.17	1.15	1.07	일식, 중식		
16 수선1가동	0.87	43.7	269.8	16.65	2.08	1.5	1.69	1.32	식육, 일식, 한식, 유계		
17 수선2가동	0.94	40.6	201.4	3.08	2.33	1.58	1.28	1.2	일식, 한식, 유계		
18 수선4가동	0.89	44.5	183.6	1.61	2.72	2.83	1.15	1.11	일식		
19 중동	0.98	47	307.8	3.71	1.17	2.08	1.04	0.88	식육		
20 지산1동	0.91	45.4	261.05	0.9	1.21	1.92	1.3	1.25	한식, 유계		
21 지산2동	0.91	44.4	336.63	3	6	3.5	1.03	1.15	식육, 중식		
22 평동	0.95	45.7	94.07	8.12	1.83	21.83	0.93	0.82	식육, 중식		
23 평간1동	0.93	39.2	154.38	4.92	0.5	21.17	1.16	1.39	식육, 중식, 유계		
24 평간2동	0.92	44.3	198.43	1.1	0.83	1.5	1.23	1.22	유계		

<그림 2. 최종 정리한 데이터>

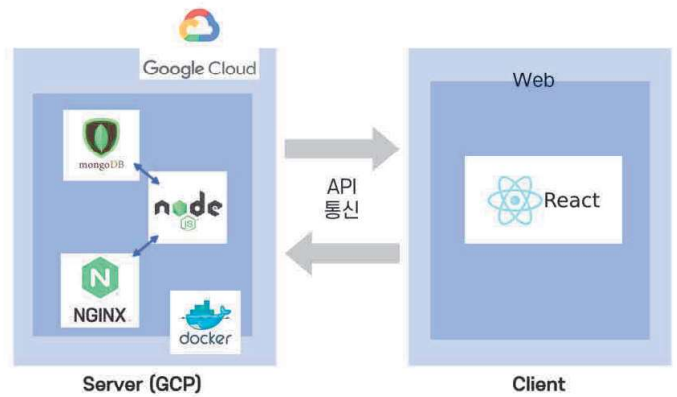
[그림 2]는 수집한 여러 데이터들을 종합하여 최종적으로 정리한 데이터이다. 최종 데이터는 수집한 공공데이터를 평균 영업기간을 중심으로 정리하였는데, 이는 점포가 유지될 수 있는 것과 이익이 큰 관련이 있다고 판단했기 때문이다.

## 2.3. 판매데이터 분석 및 가공

‘CSP 모바일 웹’의 판매데이터를 제공받아 이 판매데이터를 기준으로 기상 데이터(일별 최저기온, 일별 최고기온, 전운량, 강수량, 풍속 등)를 활용해 일별 날씨별 메뉴 판매량과 일별 메뉴 판매량을 예상하여 추천 메뉴와 원재료 수급 적정량을 제공해 줄 것이다.

그 방법으로 tensorflow 딥러닝 기술을 사용하여 해당 일의 기상 데이터를 입력 값으로 넣고 메뉴별 매출을 결과 값으로 얻는 것을 목표로 한다.

## 2.4. 상권분석 시스템 설계



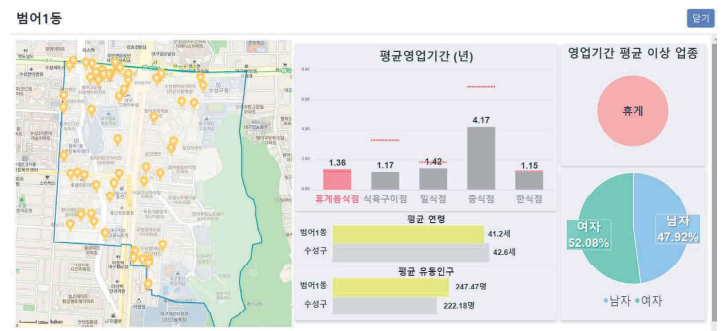
<그림 3. 서비스 구조도 및 활용 라이브러리>

주 개발도구는 VSCode를, 개발언어는 Javascript를 채택하였고, 서버는 구글 클라우드를 통해 도커를 활용한 mongoDB, node.js, Nginx를 연동하여 배포한다. 클라이언트는 React.js를 활용해 서버에 페이지 또는 데이터 통신을 요청할 수 있다.

수집한 데이터를 효율적으로 나타내기 위해 카카오맵 api와 apexcharts 라이브러리를 활용한다[그림 3].



<그림 4. 웹 페이지 접속 시 보여주는 화면>



<그림 5. 행정동 클릭 시 보여주는 화면>

먼저, 웹 페이지에 접속하면 왼쪽에는 수정구의 오늘, 내일 날씨와 예상 최고매출 메뉴가 뜨고, 오른쪽에는 수정구 행정동이 표시된 지도가 뜨고 마우스를 올리면 해당 행정동명을 마우스 포인터 옆에 띄워주며 해당 행정동 색깔을 바꿔준다[그림 4]. 행정동을 클릭하면 해당 동 명을 서버에게 보내주고 그 동에 대한 정보들을 받아오며 창을 띄운다. 창 의 왼쪽에는 해당 동 의 카페 매장 위치가 마커로 표시된 지도를 보여준다. 오른쪽에는 apexcharts를 활용해 받아온 정보들을 시각화하였다. 평균 영업기간과 평균연령은 막대그래프로, 남녀 비율은 원그래프로 나타낸다. 평균영업기간 그래프에는 빨간 점선으로 수정구의

업종별 평균영업기간을 표시했는데, 이 점선보다 높은 영업기간을 가지는 업종만 따로 표시해 준다[그림 5].

### 3. 결 론

본 고에서는 소상공인의 카페 창업과 카페 유지를 위하여, 공공데이터를 수집하고 분석하여 이를 웹 페이지로 구현하는 방법에 대해 상세하게 소개한다.

대구광역시 수성구의 인구, 성별 비율, 연령대, 유동인구, 평균 영업 지속 기간 등을 분석하고 이를 시각화함으로써 창업을 하려는 소상공인들은 관련 정보를 한눈에 파악할 수 있다.

딥러닝을 사용한다면 대구광역시 수성구가 아닌 다른 지역의 데이터도 예측이 가능하다. 대부분 제공하는 인구, 성별 비율, 연령대 등을 활용해 만든 딥러닝 모델을 통해 해당 동의 인구 수, 유동인구, 연령만 있다면 해당 동에서 유리한 업종에 대한 정보를 알 수 있을 것이다.

이를 통해 소상공인들은 희망하는 동에서 무엇을 창업하는 것이 유리할지를 비교할 수 있을 것이고 이는 소자본 창업에 도움을 줄 것이라 기대한다.

나아가 데이터의 양이 더 축적되고 적용할 수 있다면 해당 모델을 더욱 발전시킬 수 있을 것이다. 또한 적용한 데이터 외에도 다른 분야의 데이터를 제공받아 적용하고 상관 관계를 파악한다면 모델의 신뢰성과 정확성은 더욱 증가할 것이다. 이를 통해 최종적으로 전국에 대한 상권분석 시스템을 제공하는 플랫폼으로 발전할 수 있을 것이다.

### 사 사 문 구

“본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 sw 중심대학 사업의 연구결과로 수행되었음” (2021-0-01082)

### 참 고 문 헌

- [1] 이현승, 코로나19로 인한 피해 정도에 따른 매출액 영향 요인 분석, 2021
- [2] 한국농수산물유통공사의 전국 상권 트렌드 조사 보고서

### 참고 인터넷 사이트

- [1] 카카오맵 API  
<https://apis.map.kakao.com/web/>
- [2] 수성구청 공공데이터 포털  
[https://www.suseong.kr/index.do?menu\\_id=00001145&servletPath=%2Findex.do](https://www.suseong.kr/index.do?menu_id=00001145&servletPath=%2Findex.do)

# 신경망의 순방향 및 역방향 전파에 대한 이론적 분석

무술몬 롤라에프<sup>o</sup> 아난드 폴

경북대학교 CCMP 연구실

pauleditor@gmail.com, musulmon@knu.ac.kr

## Theoretical analysis of forward and backward propagation in Neural Networks

Musulmon Lolaev<sup>o</sup> Anand Paul

CCMP lab , The School of CSE, Kyungpook National University

### Abstract

Artificial Neural Networks(NN) is one of the widely used method of Machine Learning(ML) for solving various types of artificial learning tasks such as Supervised, Unsupervised, Reinforcement learning, etc. The real-life issues require more complicated algorithms to achieve expected results, especially while modeling unstructured datasets like images, text, and signals in real time. Many models have been developed based on NN including Deep learning. NN is a “black box” algorithm, therefore, in this paper, we will clarify its theoretical background by providing more sophisticated examples intending to compare its distinctive properties

### 1. Introduction

Over the last two decades, the computational performance of computers has increased significantly, and as result, large numbers of ML algorithms are available in order to model various ML tasks along with object detection, machine translations, and signal processing. Additionally, efficiently storing huge amounts of data from social site users and industries with the fast speed of the Internet provides data for sake of training, testing ML models, and applying them to industry problems. On the other hand, giant data warehouses from various sources often affect negatively the performance and make rising up the problem of Big Data [1].

There are many ML models that have been proposed for developing technologies without human intervention in individual fields including medicine, and industries after the invention of computing machines. Even history of NN back to the middle of the last century, its models are utilized to achieve expected outcomes for the most problematic fields such as machine translation and computer vision, and powerful improvements are regularly proposed.

The idea of NN was named perceptron and the first artificial NN was invented by psychologist Frank Rosenblatt in 1958 [2], and the first many-layer NN which is called nowadays Deep Neural Network was created in 1965 [3].

This paper emphasizes how NN and its forward and backward propagations work with various properties in classification tasks by supporting graphical examples.

### 2. Neural network architecture

Let's denote dataset objects  $(x^{(i)}, y^{(i)})_{i=1}^N$  where  $N$  - the number of objects in dataset,  $x^{(i)} \in \mathbb{R}^n$  - features and  $y^{(i)} \in \mathbb{R}$  - target (observation) value of object  $i$ . The following formula (equation 1) illustrated an one-layer linear NN model for linear regression task:

$$f(x^{(i)}, w) = \sum_{j=1}^n x_j^{(i)} w_j \quad (1)$$

where  $w \in \mathbb{R}^n$  - learnable parameters of NN, usually called weights of features.

Parameters  $w$  are learned by minimizing cost function using gradient descent optimization (GDO) for any given datasets  $(x^{(i)}, y^{(i)})_{i=1}^N$  separately. There are several cost functions proposed over the years for modeling numerous problems, the

least squares error (LSE) cost function is widely used for linear regression tasks [4]. Cost function is shown in equation (2) below.

$$J = \frac{1}{N} \sum_{i=1}^N (f(x^{(i)}, w) - y^{(i)})^2 \quad (2)$$

The analytical derivative of LSE for moving next step (updating parameters  $w$ ) in GDO

$$\frac{\partial J}{\partial w} = \frac{2}{N} \sum_{i=1}^N x^{(i)} (f(x^{(i)}, w) - y^{(i)}) \quad (3)$$

In general, finding global optimal parameters  $w$  using equation (1) for linear regression can be solved analytically by applying a normal equation (usually named Ordinary Least Squares), but it has many limitations in modeling real-life problems. In contrast,

GDO finds the local optimal parameters  $w$ .

### 3. Forward and backward propagation

The key idea of GDO is feeding NN forward and backward several times and updating parameters  $w$  slightly in each iteration to get the local optima based on minimizing the cost function value. The following Fig.1. shows how it works in the one forward and backward step using more accessible NN architecture. After computing forward step, the derivative of the cost function with respect to parameters  $w$ , is computed and then the backward propagation can be computed topologically reserved order of NN uses the cost value and chain rule

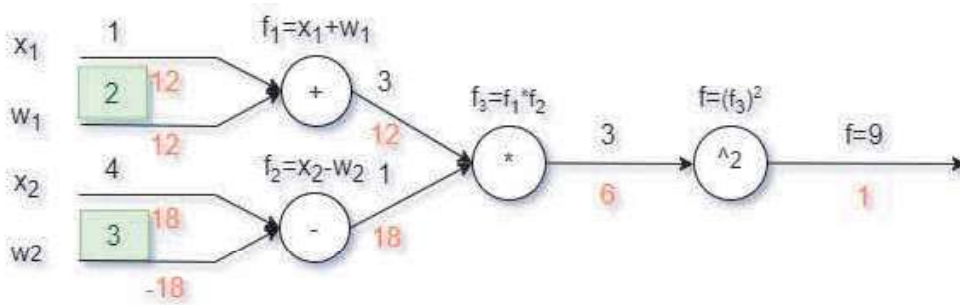


Figure 1. The illustration of forward and backward propagations. Initially, there are 2 feature value in black and 2 parameters value in green which are used for computing one step forward, and backward propagation.

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial f} \frac{\partial f}{\partial x} \quad (4)$$

where  $f(x, w)$  - function in each gate. When moving to back, it requires to find the partial derivatives to compute for each step. For first step of back propagation in the Fig.1., the derivative calculated following using equation (4)

$$\frac{dL}{df_3} = \frac{dL}{df} \frac{df}{df_3} \quad (5)$$

where  $\frac{dL}{df} = 1$ ,  $\frac{df}{df_3} = 2f_3$  and  $f_3 = 3$ . After applying

them for equation (5), the derivative will be  $\frac{dL}{df_3} = 6$ .

### 4. Activation functions

In classification tasks, NN attempts to build hyper planes which separate space into several subspaces, and the number of subspaces in each layer is independent of the number of neurons. Fig. 2 shows the multi-layer NN architecture

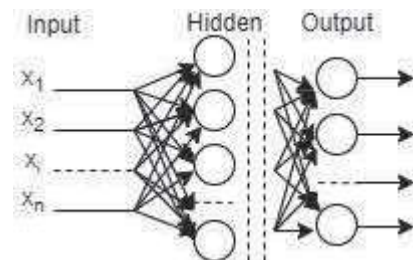


Figure 2. Multi-layer NN architecture.

Each neuron has its own activation function which computes input values as one value in fig. 3.

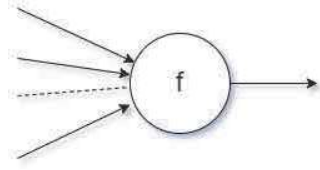


Figure 3. One neuron with an activation function.

There are several activation functions for developing NN model, and almost all of them build decision boundary differently. Therefore, choosing the best-suited activation function for each neuron affects significantly the performance of the model. We will provide visual illustrations for understanding how they affect in the task of binary classification in 2-dimensional space. The fig. 4 presents the decision boundary when applying the ReLU and Sigmoid activations.

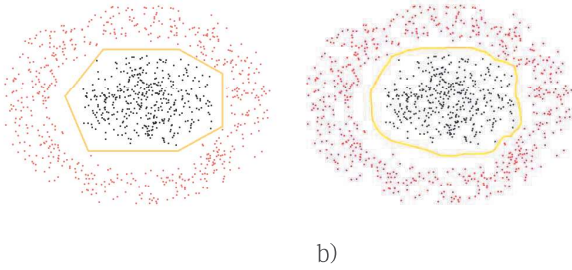


Figure 4. a)  $\max(0, f)$  - ReLU activation functions decision boundary. b)  $f(x, w) = \frac{1}{1 + e^{-wx}}$  - sigmoid activation functions decision boundary..

## 5. Loss functions

Estimating overall model error, and updating gradient descent parameters based on the cost is one the crucial indicator which affects reasonably the performance of the model. Softmax is general equation of Logistic regression for multi-class classification task and given in the equation (6).

$$J = -\sum_{i=1}^n \log \left( \frac{e^{f_{y^{(i)}}}}{\sum_{j=1}^c e^{f_j}} \right) \quad (6)$$

where  $f_i(x^{(i)}, w)$  - out values of last layer,  $c$  - number of classes in dataset.

The next loss function is SVMs(hinge loss) in equation (7).

$$L = \sum_{i=1}^n \max(0, y_o^{(i)} - y_c^{(i)}) \quad (7)$$

where  $y_o^{(i)}$  - other class scores,  $y_c^{(i)}$  - correct class scores. Briefly, the two loss functions compute the overall cost differently, the first function sums correct class scores while the second function computes it subtraction between other classes and correct class scores.

## References

1. Jianqing Fan., Fang Han and Han Liu. "Challenges of Big Data analysis" National Science Review. 2014, Vol. 1, No. 2.
2. Rosenblatt, F. "The Perceptron: A Probabilistic Model For Information Storage And Organization in the Brain". Psychological Review. 65 (6): 386-408. CiteSeerX 10.1.1.588.3775. 1958. doi:10.1037/h0042519. PMID 13602029.
3. Schmidhuber, J. "Deep Learning in Neural Networks: An Overview". Neural Networks. 61: 85-117. arXiv:1404.7828. 2015 doi:10.1016/j.neunet.2014.09.003. PMID 25462637. S2CID 11715509.
4. Goldberger, Arthur S. "Classical Linear Regression". Econometric Theory. New York: John Wiley & Sons. pp. 158. ISBN 0-471-31101-4. (1964).

# 분류에 관한 연구 무감독 딥러닝

자빈 파라<sup>o</sup>, 김재일

경북대학교 IT 대학 컴퓨터학부 (저자 소속명)

[farah\\_jabeen@knu.ac.kr](mailto:farah_jabeen@knu.ac.kr), [jaeilkim@knu.ac.kr](mailto:jaeilkim@knu.ac.kr)

## A Study of voxel classification using unsupervised deep learning

Jabeen Farah<sup>o</sup>, Jaeil Kim

Department of Computer Science, College of IT, Kyungpook National University

### 요약

The development of deep learning technology in the field of image processing has made it possible to apply deep learning to various fields based on images. In addition, medical image-based deep learning technologies are being applied to diseases in various areas such as the chest, brain, and other body parts [1]. Among the diverse reason for death, like heart attack, cancer, and another disease that the human eye is not able to predict at the initial level coronary artery disease (CAD) are the main worldwide reason for death, According to World Health Organization (WHO) [2] statistics about 17.9 million humans died from cardiovascular disease in 2019. The cause of the Coronary artery is wasted, or calcium accumulated in the walls of the coronary artery. Cardiovascular disease can be predicted by using Coronary Artery Calcification which measures the calcium in CT images.

Where deep learning indicates CAC and radiologist determine the accuracy take a lot of time. In recent studies, automatic CAC scoring methods have been introduced. In this paper, to classify voxel our model used the Encoder[3]to extract sharp and classify them to achieve better accuracy as compared to previous methods.

## 1. Introduction and Background

Depending upon the occupancy and non-existence of coronary artery segmentation CAC scoring method is evaluated. That consists of a two-stage process. In the first stage giving cardiac CT images, coronary arteries are segmented [4]-[5]. By excluding the area of the coronary artery, it helps us to easily evaluate the CAC. However, if you want to compute the CAC score is only detectable in the coronary arteries. Lessman et al. proposed a method that can calculate CAC without coronary artery segmentation thus we can skip the first step. In contrast to that, it proposed an automatic CAC scoring method that consists of two CNN models sequentially [6]. The purpose of the first CNN model is to detect global characteristics like structural position and leads to the second CNN model that helps to obtain spatial information most likely size, shape, etc. This model gives a new idea where we can skip coronary artery segmentation but follows one drawback. The performance of the model is heavily dependent on the first CNN if it does not detect the calcified region of the voxel second CNN model will not be executed

requires a lot of data, which uses a lot of bandwidth of network in IoT system.

## 2. Methodology

This section gives a brief and detail about datasets and also describe the method used for classification purpose.

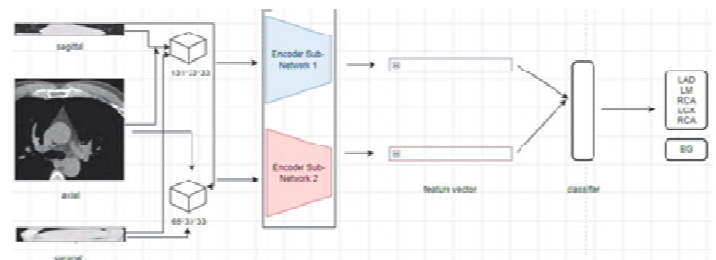


Figure 1. Proposed Model

### 2.1 Dataset

The above Figure 4 Input images consist of 2-D voxel patches from Axial, sagittal and Coronal images that fed into

the encoder and by using feature extractors we can classify voxels into different categories.

## 2.2 Data Preparation

We have 438 CT images that consist of 2.5 mm slice thickness and data has been classified into train, valid and test data. Test ratio divide by 0.6:0.2:0.2 that randomly selected 264:87:87 ct. Normalize pixel value using mean and stddev. Minimum and maximum values clipped between 2000 and -1000 respectively. Intercept value and slope value are required to convert the value loaded from the dicom file into HU value by using Equation1

$$HU = (Value * slope) + intercept \quad (1)$$

Histogram Equalization has been implemented that helps to increase the brightness and contrast the images

## 2.3 Encoder\_EMR Model

### Encoder

The encoder plays a role in learning information such as the size and shapes of the patches are different depending on the converted size. We have used two sub-networks and sub-networks are sharing weights. Sub-network structures are the same, however, generate feature vectors of the same output.

### Classifier

The classifier is implemented in that helps to classify the voxels into different categories. The classifier consists of 1 fully connected layer that takes the concatenated vector as input.

### EMR-CT Model

EMR data has been classified and by using the late fusion model it has been concatenated with the CT classifier and with the help of sigmoid classification we can predict voxel calcification.

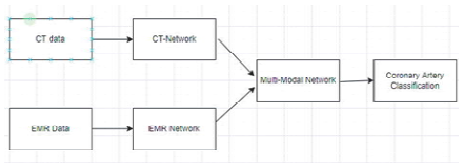


Fig. 5: EMR-CT model

## 3. Implementation Details

In this section what parameters and framework has been used will be describe.

### 3.1 Setup of parameters

The entire model implemented is based on the deep learning framework TensorFlow 2.4.2 and the GPU used for training is Quadro RTX 8000. The training of the model was conducted using Adam optimizer [7] with a batch size of 50 for 1000 epochs. The initial learning rate was set to 0.0001

## 4. Result and Discussion

In this paper, the prediction performance of the proposed method was evaluated using Accuracy 2 and F-13 Score

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (2)$$

$$F1Score = \frac{2 * precision * recall}{precision + recall} \quad (3)$$

The encoder which extracts features from the input image plays an important role in voxel classification. The performance has been measured by using a neural network of Encoder based on Resnet, Densenet. The results have been classified and compared using 5-cross validation

### Encoder Based CT classification

Fold	Encoder		Encoder-Resnet		Encoder-DenseNet	
	Acc%	F-1 score	Acc%	F-1 score	Acc%	F-1 score
Fold-1	90.1%	0.8948	94.18%	0.9052	94.96%	0.8705
Fold-2	92.6%	0.9173	93.90%	0.9319	96.21%	0.9286
Fold-3	91.7%	0.9312	91.64%	0.9277	95.77%	0.9443
Fold-4	94.4%	0.9166	91.92%	0.9164	93.06%	0.9349
Fold-5	96.9%	0.9108	89.20%	0.8759	95.50%	0.8749
Average	93.3%	0.9123	91.14%	0.9214	94.08%	0.9108

### Encoder Based CT-EMR classification

Fold	Encoder		Encoder-Resnet		Encoder-DenseNet	
	Acc%	F-1 score	Acc%	F-1 score	Acc%	F-1 score
Fold-1	93.40%	0.9018	96.94%	0.9125	95.45%	0.8710
Fold-2	93.57%	0.9391	93.46%	0.9444	94.78%	0.9351
Fold-3	93.07%	0.9275	93.41%	0.9390	93.81%	0.9299
Fold-4	93.40%	0.9344	91.92%	0.9194	94.06%	0.9247
Fold-5	94.53%	0.8876	95.20%	0.8649	95.89%	0.9249
Average	92.99%	0.9246	93.74%	0.9419	94.98%	0.9218



## 5. Conclusion

Heavy smokers undergoing screening with low-dose chest CT are affected by cardiovascular disease which is the global leading cause of death. Coronary artery calcium scoring is one of the most common methods to detect calcified plaque in coronary arteries one of the major future causes of coronary artery disease risk. Recent development in deep learning is exceedingly popular in medical images, however, very few models have been developed to integrate both clinical and imaging data. In this study, we developed a multi-model architecture that classifies the low-dose CT images with EMR data by using multi-scale encoders that classifies the coronary artery voxels. Recent methods in deep learning classify calcified pixels(2D) or voxels(2.5D or 3D) and calculate the CAC score. Dataset has been provided by Asan Medical Center in South Korea. By using the dataset in this study we have used 438 patient CT-EMR data. Our Encoder base multi-scale achieved 94% accuracy as compared to other models that have been used for similar classification purposes.

## References

1. Y. C. K. S. R. J. H.-j. B. N. K. Mingyu Kim Jihye Yun, in Deep Learning in Medical Imaging, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., Neurospine, 2019, 579–596. [Online]. Available: <https://www.e-neurospine.org/journal/view.php?doi=10.14245/ns.1938396.198>
2. World Health Organization. Cardiovascular Diseases. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
3. —, “Cardiovascular Imaging, 2019, 383–384. [Online]. Available: <https://www.e-neurospine.org/journal/view.php?doi=10.14245/ns.1938396.198>
4. M. A. V. T. L. J. M. Wolterink R. W. van Hamersvelt and I. I. S. G. M., “Coronary artery centerline extraction in cardiac ct angiography using a cnn-based orientation classifier,” 2, vol. 51, Medical Image Analysis, 2018, pp. 46–60.
5. —, “An automatic machine learning system for coronary calcium scoring in clinical non-contrast enhanced, ecg-triggered cardiac ct,” 2, vol. 9035, Presented at the SPIE Medical Imaging, SPIE, 2014, 90350E–8.
6. M. Z. P. A. d. J. B. d. V. M. V. N. Lessmann B. van Ginneken and I. I. S. G. M., “Automatic calcium scoring in low-dose chest ct using deep neural networks with dilated convolutions,” vol. 37, IEEE Transactions on Medical Imaging, 2018, pp. 615–625.
7. J. B. Diederik P. Kingma, “Adam: A method for stochastic optimization,” ICLR, 2014.

# YOLOv5를 이용한 플라스틱 재활용 폐기물 선별 정확도 향상에 대한 연구

유준호<sup>1</sup> · 안현아<sup>1</sup> · 이윤서<sup>1</sup> · 이인희<sup>1</sup> · 이지현<sup>1</sup> · 김한영<sup>2</sup> · 정설영<sup>1</sup>

경북대학교 IT 대학 컴퓨터학부<sup>1</sup>, 주식회사 윈<sup>2</sup>

niceyoo0221@knu.ac.kr, aha7969@knu.ac.kr, roiyoon@knu.ac.kr, dlsgudrpdla@knu.ac.kr,

gee9328@knu.ac.kr, henry@wimcorp.co.kr, snowflower@knu.ac.kr

## The Improvement of Plastic Recycling Waste Detection Accuracy Using YOLOv5

Junho Yu, Hyeona An, Yoonseo Lee, Inhee Lee, Jihyeon Lee, Hanyoung Kim, Seolyoung Jeong

Department of Computer Science, College of IT, Kyungpook National University<sup>1</sup>, WIM Inc.<sup>2</sup>

### 요 약

지속적으로 증가하는 재활용 폐기물의 양과 대조적으로 수작업을 통한 분류 작업을 수행하는 폐기물 선별장에서는 근무 기피 및 근로자 고령화, 인건비 상승 등의 이유로 분류 처리량 감소 문제를 겪고 있다. 본 연구에서는 재활용 폐기물의 가장 큰 비중을 차지하는 PE 플라스틱의 선별 자동화 시스템을 위한 YOLOv5 모델에 가장 적합한 학습 방법을 찾고자 보조 class를 추가하는 방안을 제안하고, 제안한 학습방법을 사용한 모델과 일반적인 모델의 AP를 비교함으로써 정확도의 향상을 확인하였다.

## I. 서 론

COVID-19 발생 이후 비대면·비접촉 문화의 빠른 확산과 함께 각종 폐기물 배출량이 급증하였다. 특히 플라스틱의 경우 배달 및 배송의 증가와 위생용품의 사용으로 그 양이 폭발적으로 증가하였으며, 지속적인 배출량 증가가 예상된다.

이러한 폐기물 배출량의 증가와 대조적으로, 각 지역 폐기물 선별장은 인력이 매우 부족한 상황이다. 대부분의 작업을 수작업에 의존하는 폐기물 선별장의 분류 시스템을 개선하기 위하여 국내외 여러 기업에서 재활용 폐기물의 분류 자동화 시스템 도입을 위한 연구를 수행하고 있다.

실제 환경의 생활폐기물들은 끊임없이 새로운 상품이 생산되고, 재활용을 위한 우선순위 또한 다양하다. 이에 본 연구에서는 특정 종류의 재활용 플라스틱 선별 정확도를 향상시키기 위한 YOLOv5 모델을 활용하며 보조 class를 추가하는 학습 방안을 제안한다.

## II. 관련 연구

R-CNN은 Selective Search를 통해 각 이미지의 Region Proposal을 생성한다. 이를 wrapping함으로써 CNN의 input으로 설정하며, wrapping 과정에서 이미지 손실이 발생하고 후보영역마다 CNN 알고리즘을 적용해야 하기 때문에 시간복잡도와 공간복잡도가 크다는 단점이 있어 본 연구에 활용하기에 부적합하다[1].

앞서 언급한 R-CNN과 비교하여 Fast R-CNN은 wrapping 과정이 없으며, 원본 이미지 전체에 CNN을 적용하여 Region Proposal을 생성하고 RoI Pooling을 통해 고정 크기를 가진 Feature vector로 추출한다. R-CNN 대비 정확도와 속도가 상승하였으나, 여전히 Selective search를 통해 Region Proposal을 진행하기 때문에 속도가 느려 본 연구에 활용하기 어렵다[2].

YOLO는 Bounding Box와 Class Probability를 동시에 고려하여 객체의 종류와 위치를 한 번에 예측한다. 이미지를 고정 크기의 그리드로 쪼개고 각 그리드의 Bounding Box를 예측한다. 처리과정이 간단하고 속도가 매우 빨라 실시간 영상 처리에 강한 모습을 보인다[3]. 실시간 재활용 플라스틱 선별 딥러닝 모델로 가장 적합

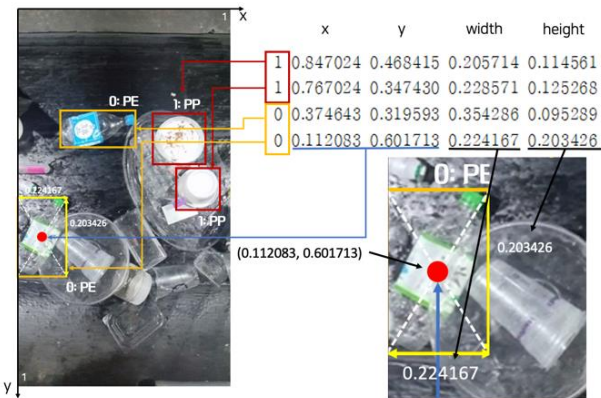
하여, 본 연구에서는 YOLOv5 모델을 활용하였다.

### III. 설계 및 구현

#### 3.1 데이터 가공

YOLOv5 모델 학습을 위한 학습 데이터셋을 구축하기 위하여 데이터 가공이 필요하다. 폐기물 선별장으로부터 수집한 원천 데이터 영상을 이미지 프레임으로 나누고, 딥러닝 학습을 위한 태그를 생성하여 재활용 플라스틱과 매칭하는 데이터 라벨링을 수행한다. 큰 범위의 class를 PE와 PP 두 가지로 정한 후 class 내부에서 딥러닝 모델이 특징을 잘 잡아 선별할 수 있도록 각 폐기물의 형태를 기준으로 태그를 구분하여 총 8,896장의 가공 데이터를 생성하였다.

원천 데이터 영상의 프레임 하나당 원본 이미지 파일 하나와 바운딩 박스 라벨링 정보 파일 하나가 대응된다. 바운딩 박스 라벨링 정보는 <그림 1>와 같이 이미지의 Bounding Box에 대한 속성값이 txt 파일로 저장되며, Bounding Box 속성값은 class 인덱스, 박스중심의 x좌표, 박스중심의 y좌표, 박스의 너비, 박스의 높이로 구성된다. class 인덱스 0은 PE를, 1은 PP를 의미하므로, <그림 1>의 경우 학습을 위한 PE 2개와 PP 2개 데이터를 가공한 파일에 해당한다.



<그림 1> 데이터 가공을 통한 txt 파일 생성

#### 3.2 데이터 학습

학습의 효율성과 다양한 활용을 위하여 본 연구에서는 데이터를 여러 방식으로 증강시키는 방법을 채택하

였다. 데이터 증강에 사용된 방식과 하이퍼 파라미터 값은 다음과 같다.

사진의 H(색조), S(채도), V(명도) 값을  $hsv\_h = 0.015$ ,  $hsv\_s = 0.7$ ,  $hsv\_v = 0.4$ 로 조정하여 데이터를 증강한다. 또한,  $translate$  값을 0.1로 설정하여 이미지를  $\pm 0.1$  이동시키거나  $scale$  값을  $\pm 0.1$ 의 범위로 부여하여 이미지 크기를 확대, 축소함으로써 데이터 증강이 가능하다.  $flipplr$  파라미터는 50% 확률로 이미지 좌우반전을 수행한다.  $Mosaic$  파라미터는 이미지들을 조각조각 붙여 데이터를 증강하며, 확률값에 1.0을 부여하여 모든 데이터에 적용되도록 설정하였다.



<그림 2> (좌)원천 데이터, (우)파라미터 조정 데이터

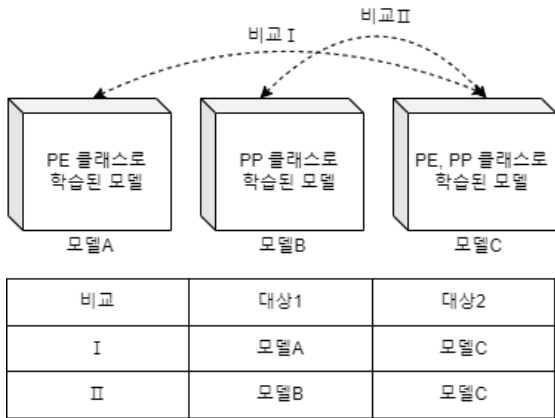
#### 3.3 Assistance class learning

본 연구에서는 YOLOv5 모델의 Object Detection 정확도 향상을 위하여 보조적인 학습 방법을 제안한다. 하나의 특정 class를 detection하는 경우에 대하여 보조 class를 활용할 경우 대상 class의 정확성을 향상시킬 수 있다.

실제 폐기물 선별장에 들어오는 재활용 폐기물 중 가장 많은 양을 차지하는 PE의 효율적인 선별 작업을 위하여, 선별장 내부에서는 오직 PE에 대한 정확도에만 집중하였다. 그러나 YOLOv5 모델을 활용한 재활용 플라스틱 선별 자동화 딥러닝 모델의 경우 단순 PE의 학습에만 집중하기보다는 PP 또한 학습하는 것이 정확도를 포함한 전체 성능 향상에 도움될 것으로 예측된다.

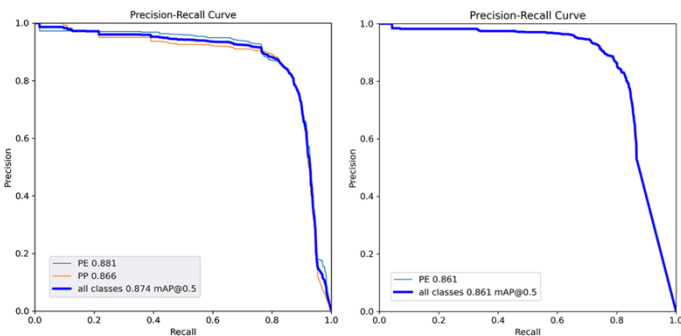
#### IV. 연구 결과

해당 방안을 증명하기 위하여 두가지 비교연구를 실행하였다. <그림 3>은 보조 class를 추가하여 학습한 모델의 성능 변화 양상을 알아보기 위한 비교 연구를 도식화한 것이다. 각 모델 A, B, C의 AP를 비교함으로써 class 추가에 따른 정확도 차이를 살펴보았다.



<그림 3> 비교 연구 도식화

확보한 데이터셋을 Validation set에 PE와 PP의 수가 균일하도록 Train set 7804개, Validation set 1219개로 분할했다. YOLOv5 small, medium, large model로 학습 진행 후 테스트를 수행하였다.



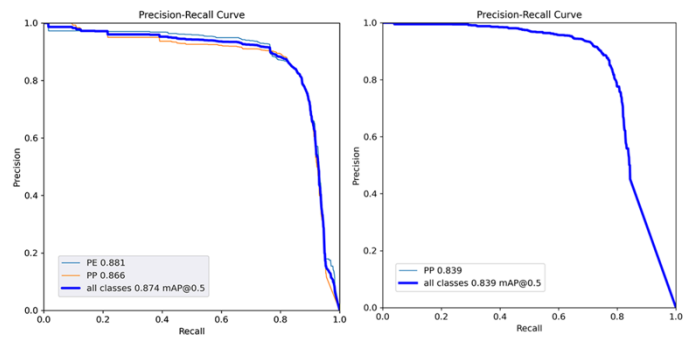
<그림 4> 비교연구 I Precision-Recall Curve

<그림 4> 과 <그림 5> 는 학습한 YOLOv5 model의 Precision-Recall 그래프이다. 각 그림의 왼쪽은 모델 C, 오른쪽은 각 비교실험에 대한 대상1의 모델을 사용

했다. Precision-Recall 그래프 아래 면적은 AP를 나타낸다.

<그림 4>은 비교연구 I에 대한 결과 그래프로 보조 class가 추가된 모델 C가 모델 A에 비해 AP가 높은 것을 시각적으로 확인할 수 있다.

<그림 5>은 비교연구 II에 대한 결과 그래프로 보조 class가 추가된 모델 C가 모델 B에 비해 AP가 높은 것을 시각적으로 확인할 수 있다.



<그림 5> 비교연구 II Precision-Recall Curve

<표 1>은 보조 class를 추가하여 학습한 모델의 성능 변화 양상이다. 하나의 class를 가지고 있는 모델 A와 모델 B의 AP는 각각 0.861, 0.839이고, class가 추가된 모델 C의 PE AP는 0.881, PP AP는 0.865로 기존의 모델 A, B보다 높다.

	모델 A (PE)	모델 B (PP)	모델 C (PE+PP)
PE(AP)	0.861	-	0.881
PP(AP)	-	0.839	0.865

<표 1> class 추가에 따른 성능 변화 양상

<표 2>와 <표 3>은 보조 class를 추가하여 학습한 모델의 성능 변화 양상을 다양한 환경에서 확인하고

자 테스트한 결과이다. model depth multiple과 layer width multiple에서 차이를 보이는 medium model과 large model에서도 동일한 테스트를 수행하였으며, 앞선 small model과 마찬가지로 PE와 PP를 모두 학습한 모델에서 더 높은 성능을 보임을 확인하였다.

	모델 A (PE)	모델 B (PP)	모델 C (PE+PP)
A11(AP)	-	-	0.886
PE(AP)	0.873	-	0.898
PP(AP)	-	0.852	0.873

〈표 2〉 YOLOv5m 모델을 활용한 성능 변화 양상

	모델 A (PE)	모델 B (PP)	모델 C (PE+PP)
A11(AP)	-	-	0.899
PE(AP)	0.872	-	0.907
PP(AP)	-	0.860	0.891

〈표 3〉 YOLOv5l 모델을 활용한 성능 변화 양상

## V. 결 론

본 논문에서는 실시간 객체 탐지 알고리즘 YOLOv5를 이용한 재활용 플라스틱 모델의 성능을 향상시키기 위하여 보조 class를 추가하는 방안을 제안하고, 모델 별 AP를 비교함으로써 정확도의 향상을 확인하였다.

해당 연구를 통해 개발한 재활용 플라스틱 폐기물 선별 딥러닝 모델은 지속적인 고도화를 통해 폐기물 선별장의 플라스틱 폐기물 분류 자동화 시스템 구축을 위하

여 사용될 수 있다. 학습 데이터셋의 구체화를 통해 PE와 PP 외 PS, PVC 등 선별 가능한 플라스틱의 종류를 다양화하거나 알루미늄, 유리 등 선별 폐기물 범위를 확장하여 응용할 수 있을 것이다. 나아가 선별 작업을 필요로 하는 여러 사업에 적용하여 작업의 자동화를 꾀하는 하나의 방안이 될 것이다.

## 사 사 문 구

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음. (2021-0-01082)

본 연구를 위한 원천 데이터는 주식회사 워으로부터 제공받았음.

## 참 고 문 헌

- [1] Girshick et al, "Rich feature hierarchies for accurate object detection and semantic segmentation", 2014
- [2] Girshick, "Fast R-CNN", ICCV 2015, 2015.
- [3] Redmon et al, "You Only Look Once: Unified, Real-Time Object Detection", 2016

# 트랜스포머 기반의 BERT 모델과 온라인 포럼 데이터를 활용한 시스템 근본 원인 종류 탐색

김유양<sup>0</sup> 송소민 탁병철

경북대학교

[youyangkim@knu.ac.kr](mailto:youyangkim@knu.ac.kr), [sominsong@knu.ac.kr](mailto:sominsong@knu.ac.kr), [bctak@knu.ac.kr](mailto:bctak@knu.ac.kr)

## Root Cause Categorization Using Online Forum Data and Transformer-Based BERT Model

Youyang Kim<sup>0</sup> Somin Song Byungchul Tak  
Kyungpook National University

### 요약

시스템에서 발생하는 여러 종류의 문제를 해결하기 위해서는 사전에 어떤 종류의 근본 원인이 문제를 발생시키는지 인지하는 것이 중요하다. 본 논문에서는 온라인 포럼 중 하나인 스택 오버플로우 데이터를 활용하여 시스템에서 발생하는 문제의 근본 원인의 종류를 찾는 방법에 대해 제시하고, 오픈소스 클라우드 플랫폼인 오픈 스택에서 발생하는 문제 근본 원인에 대해 탐색한다.

### 1. 서론

현대 시스템은 계속해서 구조가 복잡해지고 이에 따라 시스템 내부에서 발생하는 문제의 종류도 다양해지고 있다. 시스템에서 생긴 문제를 해결하고 복구하여 다시 시스템이 정상적으로 수행되게 하는 트러블슈팅은 현재까지 IT 운영 팀 (ITops)의 가장 큰 목표이다. 또한 기존의 대부분 기술들은 문제가 발생함을 인지하는 데에 그쳤고, 어떠한 근본 원인 때문에 문제가 발생했는지에 대해서는 제시하지 않고 있다. 문제의 원인을 파악하는 것은 문제를 해결하고 복구되는데 걸리는 시간을 감소시킨다. 그렇기 때문에 사용자가 시스템에서 발생할 수 있는 문제의 근본 원인의 종류에 대해 사전에 인지하고 있는 것은 문제를 해결하는 시간을 크게 감소시킬 수 있다. 본 논문에서는 특정 시스템에서 주로 발생하는 근본 원인의 종류를 찾는 방법에 대해 제시하고, 오픈소스 클라우드 플랫폼인 오픈 스택[4]에 해당 방법을 적용하여 어떤 종류의 근본 원인이 존재하는지를 찾는다.

현재 트러블슈팅을 위한 많은 모니터링 기술과 분석 기술이 있지만, 대부분은 숙련된 사람들의 지식이나 경험에 의존한다. 최근에는 온라인 포럼과 같은 지식을 공유하고 토론하는 공간이 트러블 슈팅과 문제 해결을 위한 필수 장소가 되고 있다. 온라인 포럼 (예를 들어, stackoverflow[3])에서 사용자는 시스템을 사용하다 생긴 문제에 대해 글을 작성하고, 또 다른 사용자는 해당 글의 답변에 문제의 원인 또는 해결책을 제시한다. 우리는 이러한 답변 글에 존재하는 문제의 원인을 나타내는 문장들을 사용하여 시스템의 근본 원인 종류에 대해

탐색한다.

문장의 내용을 분석하여 근본 원인의 종류를 찾기 위해 자연어처리 기법 (NLP)을 적용한다. 먼저 많은 답변 글 중에서 문제 원인을 나타내는 문장만 분류한다. 대부분의 답변 글은 문제 원인을 나타내는 문장, 해결책을 제시하는 문장, 어디에도 속하지 않는 문장 과 같이 여러 종류의 문장으로 구성되어 있기 때문에 근본 원인을 나타내는 문장만 분류해 내는 작업이 필요하다. 또한, 유사한 근본 원인을 내포하고 있는 문장들은 비슷한 언어 패턴을 나타낸다. 예를 들어, “The problem is configuration file”, “The configuration is the root cause.” 를 통해 우리는 시스템의 잘못된 설정 파일이 문제의 근본 원인임을 알 수 있다. 유사한 의미의 문장은 비슷한 패턴을 가진다는 점을 이용하여 비슷한 의미의 문장을 벡터공간에서 가깝게 위치시키는 문장 임베딩 기법을 적용한다. 이후 벡터 공간에서 가깝게 위치하는 데이터들끼리 그룹화 하는 과정을 거쳐 문장 그룹을 생성하고 이를 근본 원인의 종류로 정의한다.

본 논문에서는 대표적인 온라인 포럼 중 하나인 스택 오버플로우(stackoverflow)사이트의 데이터를 이용하여 오픈스택 시스템에서 어떤 종류의 근본원인이 주로 문제를 발생시키는지 찾아본다.

### 2. 관련 연구

시스템에서 발생하는 문제의 근본 원인을 찾는 연구는 많이 진행되었지만, 여전히 문제점 들이 존재한다. 연구자들은 시스템에서 발생한 문제를 해결하기 위해 이전의 경험에

의존하거나[5], 시간에 따라 이상현상이 전파되는 그래프를 구현하는 방식[6]을 사용하였다. 다른 연구에서는 시스템에 존재하는 센서의 시간에 따른 변화를 분석하거나[7], 시스템 콜을 분석하는 방법[8]을 통해 문제의 원인을 찾으려 한다. 하지만 해당 연구들은 대부분 시스템에서 문제가 발생했을 때, 이를 탐지하는 역할에서 그치고, 문제의 근본 원인은 찾지 못한다.

자연어 분석을 위해서는 먼저 텍스트를 실제 값 벡터로 변환하는 임베딩 작업이 필요하다. BERT[1]는 최근 자연어 처리 모델 중 가장 뛰어난 성능을 보이는 Pre-trained 모델이다. BERT는 같은 단어라도 문맥에 따라 다른 임베딩 값을 만들어내는 모델로 문맥을 반영하여 임베딩 하는 모델이다. 본 연구에서는 문장 데이터로 학습된 SBERT(Sentence-BERT)[2]를 사용하여 문장 임베딩을 진행한다.

### 3. 실험 데이터

본 논문에서는 스택 오버플로우 데이터를 사용하여 오픈 스택에서 발생하는 문제의 근본 원인을 탐색한다. 스택 오버플로우 사이트에서 사용자들은 질문을 통해 시스템 운영 중 생긴 문제에 대해 공유하고, 답변을 통해 문제의 원인이나 해결책에 대한 정보를 얻는다. 우리는 이 질문과 답변, 2개의 부분 중에 답변 부분만을 사용해 문제 원인의 종류들을 파악하고자 한다. 또한 오픈 스택에 관련된 글만 사용하기 위해 오픈스택 관련 단어들(openstack, nova, neutron, cinder, etc.) 로 태그 된 질문 글만 사용하고, 각 질문 글에서 질문자가 채택한 답변 글만 사용한다. 스택 오버플로우에서 오픈스택 관련된 단어들로 태그 된 글의 수는 3058개이며, 그 중 질문자가 채택한 답변 글이 존재하는 글은 930개이다. 본 논문에서는 930개의 글을 사용하여 오픈 스택에 존재하는 근본 원인의 종류에 대해 탐색한다.

온라인 포럼 답변 글을 구성하는 문장들은 문제 원인, 해결책, 해당 사항 없음 과 같이 3가지의 문장 종류로 구분할 수 있다. 표 1. 에서는 각 문장 종류의 예시를 나타내며, 각 종류별 특징을 보여준다. 우리는 3가지 문장 종류 중 문제 원인을 나타내는 근본 원인 문장만 사용한다. 해당 문장을 전체 데이터에서 분류하기 위해 트랜스포머 기반의 BERT 모델을 문장 분류 작업으로 fine-tuning 하여 사용한다. 문장 분류 모델을 생성하기에 앞서 먼저 불필요한 데이터를 제거하는 데이터 정제 과정을 진행한다. 이후 정제된 문장 데이터를 사용하여 BERT fine-tuning을 진행한다.

표 1. 근본 원인, 해결책, 해당 사항 없음에 해당하는 문장의 예시.

종류	문장
근본 원인 (root cause)	The root cause is that flavor with which instance was created has capacity 80GB, but only get free_disk_gb is 47 by hypervisor_stats.
해결책 (solution)	The solution is to modify the source.rc file the devstack installation
해당 사항 없음	Update for other reference.

### 3.1 데이터 정제

온라인 포럼 데이터의 특성은 코드, 콘솔 출력, 로그와 같은

데이터가 포함되어 있다는 점이다. 해당 데이터들은 자연어가 아니기 때문에 일반적인 단어 임베딩 기법을 적용할 수 없다. 스택 오버플로우에서는 <code> HTML 태그를 사용하여 해당 데이터들을 표현하고, 우리는 <code> 태그를 사용하여 해당 데이터들을 모두 제거하였다.

### 3.2. BERT fine-tuning을 통한 근본 원인 문장 분류

근본 원인을 나타내는 문장은 다른 문장들과 비교했을 때, 뚜렷한 언어 패턴이 존재한다. 예를 들어, “The problem is ~”, “It looks like the ~ not due to ~”, 와 같은 문장은 근본 원인을 표현하는 문장의 특징이다. 우리는 근본 원인 문장을 분류하는 작업이 감정 분석과 매우 유사하다는 점을 반영하여 NLP 분류기를 사용한다. 본 연구에서는 트랜스포머 기반의 BERT 모델을 사용하여 근본 원인을 나타내는 문장과 아닌 문장을 분류한다.

표 2. 에서는 분류기로 분류한 근본 원인 문장과 근본 원인 문장이 아닌 것의 예시이다. 근본 원인 (root cause)문장은 “The main reason for this ~” 를 통해 해당 문장이 문제의 근본 원인을 나타내고 있음을 알 수 있으며, 아닌 문장 (N/A)로 분류된 문장에서는 근본 원인에 대한 설명이 없음을 파악할 수 있다. 이 분류기는 0.8946의 정확도를 가진다.

표 2. 학습한 분류기를 사용하여 근본 원인 문장과 아닌 문장을 분류한 결과의 예시

종류	문장
근본 원인 (Root cause)	The main reason for this configuration issue was while creating the subnet I used auto for -gateway option and obviously it didn't pick the correct gateway.
근본 원인 아님 (N / A)	The following method worked for me.

### 4. 문장의 형상 추출 (Feature Extraction)

자연어 처리(NLP)에서는 텍스트 분석을 위해 각 단어, 또는 문장을 실제 값 벡터로 변환하는 작업이 필요하다. 이때, 유사한 단어나 문장은 벡터 공간에서 가까운 공간에 위치할 것이라 가정하는데 이를 단어 임베딩 또는 문장 임베딩이라 한다. 본 연구에서는 근본 원인을 나타내는 문장을 SBERT (sentence-BERT)를 사용하여 문장 임베딩 한다.

#### 4.1. 문장 전처리

텍스트를 실제 값 형태의 벡터로 변환하기 전에 불필요한 값들을 제거하는 전처리 작업이 필요하다. 우리는 이 단계에서 불필요한 문자 제거 (“ ”, “\*” 와 같은 특수문자), 소문자로의 변환을 진행하였다.

#### 4.2. SBERT를 사용한 문장 임베딩

유사한 근본 원인을 내포한 문장은 유사한 언어 패턴을 지닌다. 표 3. 은 유사한 근본 원인을 가진 문장과 유사하지 않은 근본 원인을 가진 문장으로 선정한 ground-truth 예시를 나타내고, 해당 표를 통해 같은 근본 원인을 나타내는 문장의 공통점을 파악할 수 있다. 또한 표 3. 에서 유사하지 않은 근본 원인 예시를 통해 같은 단어이더라도 사용되는 문맥에

따라 문장의 의미가 달라짐으로, 문맥을 반영한 문장 임베딩이 필요하다는 것을 알 수 있다.

우리는 문맥을 반영한 문장 임베딩을 하는 모델인 SBERT를 사용하여 전처리 과정을 거친 각 문장들을 실제 값 벡터로 변환한다.

표 3. 유사한 근본 원인을 가진 문장과 유사하지 않은 근본 원인을 가진 문장 예시

종류	문장
유사한 근본 원인	Seems auth_url is unavailable
	The auth_url variable is wrong.
유사하지 않은 근본 원인	Seems auth_url is unavailable.
	Auth_url is not the problem, but your password is incorrect.

### 5. K-Means 알고리즘을 사용한 문장 임베딩 그룹화

임베딩을 통해 각 문장들이 실제 값 벡터로 변환된 후, 비슷한 벡터를 가지는 문장을 하나의 그룹으로 묶는 그룹화를 실행한다.

이 그룹화 과정에서는 K-Means 클러스터링을 사용한다. K-Means 클러스터링은 유클리디안 거리를 측정을 통해 그룹의 중심과 각 데이터 간의 거리를 최소화하는 작업으로 데이터를 K개의 클러스터로 묶는 머신 러닝의 비지도 학습 기법이다.

K-Means 클러스터링 기법은 데이터들 간의 거리 계산이 주를 이뤄 속도가 빠르기 때문에 대용량 데이터 처리에 적합하므로 우리는 해당 기법을 사용한다.

#### 5.1. 실루엣 값을 통한 그룹 수 결정

K-Means 클러스터링 기법에서는 클러스터의 개수 K를 정해야 한다. 우리는 이를 실루엣 값을 통해 K를 정한다. 실루엣 값은 한 데이터가 다른 클러스터에 비해 현재 속한 클러스터와 얼마나 유사한지를 측정하는 값이며 해당 값이 1에 가까울수록 각 데이터가 속한 그룹과의 유사도가 높음을 의미한다. 본 연구에서는 K를 120 부터 150까지 변경시켜가며 실루엣 값을 측정하였고, 실루엣 값이 가장 높게 나타난 값 148을 그룹의 수 K로 지정한다.

그림 1.에서는 스택 오버플로우 데이터를 사용한 근본 원인 탐색의 전체 과정을 나타낸다.

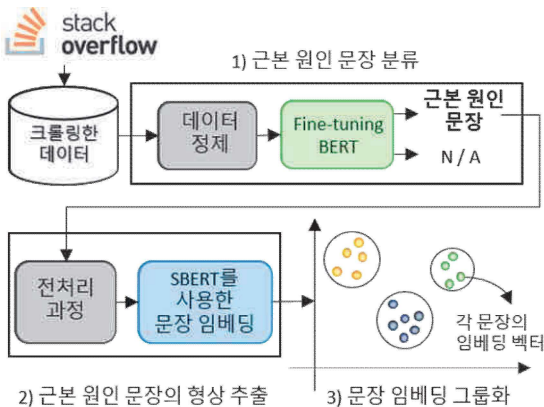


그림 1. 근본 원인의 종류를 분석하기 위한 전체 과정

## 6. 오픈 스택에서 발생하는 근본 원인의 종류

근본 원인 종류 탐색 방법을 통해 오픈 스택에서 어떤 종류의 근본 원인이 존재하는지를 분석한다.

본 연구에서 사용한 오픈 스택 관련 스택 오버플로우 질문 글의 수는 930개이며, 해당 질문 글에서 질문자가 선택한 답변 글만 사용한다. 답변 글을 구성하는 문장은 총 3633개이다.

표 4.는 오픈 스택과 관련된 스택 오버플로우 데이터 중 근본 원인을 나타내는 문장을 사용하여 문장 임베딩 후 그룹화 한 결과로, 많이 발생하는 문제의 근본 원인을 순서대로 정렬한 결과이다. 또한 각 근본 원인은 해당 그룹에 속한 문장들의 공통적인 내용을 바탕으로 정의한 것이다. 해당 표에서는 가장 많이 발생하는 근본 원인 top 4 를 나타낸다. 1순위로 많이 발생하는 문제 상황 및 근본 원인은 설정 오류로 인해 오픈 스택의 컴포넌트의 한 종류인 nova-compute가 존재하지 않는 것으로 나타났으며, 2위로 많이 발생한 근본 원인은 openstack 컴포넌트 중 하나인 Swift가 불가능 (disable)상태 이거나 설치되지 않았을 때 발생하는 문제이다. 3위에는 각각 오픈 스택 네트워크 중 provider 네트워크가 존재하지 않는 것이 근본 원인이며, 4위로는 오픈 스택 버전 차이가 근본 원인으로 나타났다. 표 5.에서는 각 근본 원인 그룹에 속하는 질문 글의 수를 나타낸다. 1순위의 근본 원인 그룹에는 총 32개의 질문 글이 속하며, 전체 질문 글 중 3.44%를 차지한다. 2, 3, 4 순위 근본 원인 그룹에는 각각 28, 27, 27개의 질문 글이 해당 그룹에 속하며, 전체 질문 글 중 3.01%, 2.90%, 2.90%를 차지한다.

표 4. 오픈 스택에서 많이 발생하는 근본 원인의 종류 top4

순위	근본 원인
1	설정 오류로 인해 Nova-compute가 존재하지 않음
2	Swift가 불가능 상태 (disable) 또는 설치되지 않음
3	네트워크 중 provider(external) network가 존재하지 않음
4	오픈스택 버전 차이로 인해 문제 발생

표 5. 각 그룹별 질문 글 수 (총 930개)

순위	1	2	3	4
개수	32 (3.44%)	28 (3.01%)	27 (2.90%)	27 (2.90%)

### 6.1 각 그룹별 높은 빈도수의 단어 정렬

각 근본 원인 그룹별 공통적인 특성을 파악하기 위해 각 그룹에 속한 문장을 구성하는 단어를 빈도수에 따라 정렬한다. 표 6.은 각 그룹 별 높은 빈도수를 나타내는 단어를 내림차순 정렬하고 그 중 상위 6-7개의 단어를 나타낸다. 1순위로 많이 발생하는 근본 원인을 표현하는 문장에는 “nova, deleted, nova.conf, node, restart” 와 같이 nova(nova-compute)를 표현하는 단어들이 많이 사용됨을 보이며, 이는 그룹에 존재하는 문장들이 공통적으로 nova 컴포넌트에 대해 말함을 알 수 있고, nova.conf를 통해 nova 컴포넌트의 설정 파일에 관련된 원인임을 알 수 있다. 또한 2순위 그룹에서는 swift, storage, installation, container와 같은 단어의 빈도수가 높음을 통해 해당 그룹에 속한 문장들이 공통적으로 swift 컴포넌트에 대해 설명하고, storage 서비스를 지원하는 swift 컴포넌트가 설치되어야 함을 나타낸다. 3



순위의 그룹에서는 network, external, provider의 단어를 통해 vm 인스턴스에 네트워크 서비스를 제공하는 provider(external) 네트워크에 대해 말하고 있음을 알 수 있으며, 4순위의 version, not, running, latest, release를 통해 오픈 스택 버전이 맞지 않고, 최신 버전이 필요함을 알 수 있다.

표 6. 오픈스택 근본 원인의 그룹별 문장을 구성하는 단어의 빈도수 내림차순 정렬 중 상위 7개

순위	최고 빈도 단어 정렬
1	Nova, deleted, nova.conf, node, instance, service, restart
2	Swift, support, storage, installation, repositories, container
3	Network, external, instance, create, router, provider, ip
4	Version, not, running, latest, release, command, client

## 7. 결론 및 향후 연구

본 연구에서는 대표적인 온라인 포럼인 스택 오버플로우 데이터를 사용하여 시스템에 존재하는 근본 원인의 종류에 대해 분석하는 방법을 제시하고, 오픈소스 기반의 플랫폼인 오픈 스택에 존재하는 근본 원인의 종류에 대해 탐색하였다. 스택 오버플로우의 글 중 근본 원인을 나타내는 문장을 분류하기 위해 BERT 모델을 사용하였으며, 문장 임베딩과 클러스터링을 통해 공통적인 특징을 보이는 문장들을 그룹화하고 근본 원인의 종류를 정의하였다. 그 결과, 오픈 스택에서 발생하는 문제는 설정 오류로 인해 컴포넌트가 존재하지 않는 원인, 오픈스택 컴포넌트 중 하나인 swift가 동작하지 않는 원인, 인스턴스에 연결할 네트워크가 존재하지 않는 원인 그리고 오픈스택 버전의 차이가 문제를 발생시키는 주 근본 원인임을 발견하였다. 향후 연구 계획으로는 다른 온라인 포럼 데이터를 활용한 근본 원인 종류 탐색, 코드, 로그와 같은 데이터를 활용한 근본 원인 종류 탐색 등이 있다.

## 사 사 문 구

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 한국 연구재단의 지원을 받아 수행된 연구임(No. 2021R1A5A1021944)

## 참고 문헌

[1] Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)

[2] Reimers, Nils, and Iryna Gurevych. "Sentence-bert: Sentence embeddings using siamese bert-networks." arXiv preprint arXiv:1908.10084 (2019).

[3] stackoverflow. <https://stackoverflow.com/>, online

[4] openstack. <https://openstack.org/>, online; accessed 28 March 2022

[5] Melchior, C., Tarouco, L.M.R.: Troubleshooting network faults using past experience. In: NOMS 2000. 2000 IEEE/IFIP Network Operations and Management Symposium 'The Networked Planet: Management Beyond 2000' (Cat. No. 00CB37074). pp. 549-562. IEEE (2000)

[6] Lin, J., Zhang, Q., Bannazadeh, H., Leon-Garcia, A.: Automated anomaly detection and root cause analysis in virtualized cloud infrastructures. In: NOMS 2016-2016 IEEE/IFIP Network Operations and Management Symposium. pp. 550-556. IEEE (2016)

[7] Kim, M., Sumbaly, R., Shah, S.: Root cause detection in a service oriented architecture. ACM SIGMETRICS Performance Evaluation Review 41(1), 93-104 (2013)

[8] Ren, Z., Liu, C., Xiao, X., Jiang, H., Xie, T.: Root cause localization for unreproducible builds via causality analysis over system call tracing. In: 2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE). pp. 527-538. IEEE (2019)

# 감지 데이터가 포함된 딥 러닝을 사용하는

## 작업자의 추락 감지

자빈 샤이사타 홍봉희

[shaistaj8@gmail.com](mailto:shaistaj8@gmail.com), [vbhhong@gmail.com](mailto:vbhhong@gmail.com)

## Falling Detection of Workers using Deep Learning with Sensing Data

Jabeen Shaisata, Hong Bong Hee

### Abstract

Any nation's development is greatly influenced by the construction industry. The workers at construction area have high risks of danger to their lives and needs to be fully alert and active as they are working with heavy machinery. Wearable fall detection systems that have sensors are considered more accurate and efficient as they are easy to carry with workers and does not become a hurdle to their work. As workers have to work with heavy machinery for a long time and other environmental changes can cause accidental fall which is great damage to the health of workers, as it can cause injury and can be the reason of death as well. In this study an open-source dataset is collected using wearable sensors while performing some regular tasks. The data collected from gyro and accelerometer sensor is processed and clean and then the learning model is applied to it. CNN-LSTM based model for detection of fall is introduced and applied together because of CNN feature extraction capability together with LSTM as its popular for its time series ability.

### 1. Introduction and Background

Today, fall detection is one of the major concerns of construction industry. Fall accidents are one of the major causes of serious fatalities and injuries at the construction sector. Falls are a leading cause of injuries and can lead to devastating disabilities.

As there are many methods and systems introduced to detect Fall. One of the systems uses the sensors as accelerometer and gyroscope, as these sensors are attached to person whose movements are to be observed, these sensors generate radio waves which are examine on a screen. The fall detection devices use three-axis data.

One of the leading causes of fatal injuries in older individuals is unintentional falls. According to [1], more over 25% of those over 65-years of age experience fall and the ratio increases to 32%-42% for those over 70. Additionally, between 30% and 50% of residents of long-term care facilities experience falls each year, with nearly half of them having repeated falls. 40% of injury deaths and 20% to 30% of mild to severe injuries are caused by falls. In the US, the average cost of a single hospitalization for fall-related injuries among patients 65 years of age was \$17483 in 2004,

with a prediction of \$240 billion in total expenses by 2040. As per alarming situation a reliable and accurate model for the detection of fall is therefore needed for society.

Deep learning has been used more and more in fall detection in recent years. In order to streamline the data preprocessing procedure, Fakhrulddin et al. (2017) [2] introduced an algorithm with a detection of fall accuracy of 92.3% as it transforms accelerometer data into RGB images and categorizes them using a convolutional neural network. Deep learning-based techniques can't be installed directly on embedded device; thus, the data must be sent over the network to the server for processing. Although this method can cause some delays, it makes it easier to monitor the movement status of the elderly in the park in a unified manner to provide prompt rescue in the event of falls in situations like nursing homes. To detect a fall, the method in [2] still requires a lot of data, which uses a lot of bandwidth of network in IoT system.

We suggested a fall detection model which is based on CNN-LSTM combine together to address the aforementioned issues. The model uses neural networks to extract the information automatically and classify them without the need for a laborious preparation step. In order to save data, the three-axis accelerometer's raw data is first converted to the resultant acceleration value before being transferred to the

CNN-LSTM network for categorization. It is ideal for IoT applications since it just requires a fewest possible data to complete the fall detection method. The model applied to dataset [3] shows the accuracy of 98.97%.

The remainder of the paper is presented as follows. Firstly, scheme and model utilized are presented in Section 2. The dataset and experimental parameters are described in Section 3., while Section 4 shows the experimental result achieved on a selected dataset and its comparison with other works. Section 5 includes the conclusion of paper.

## 2. Methodology

This section gives a brief and detail about datasets and also describe the method used for the detection of fall.

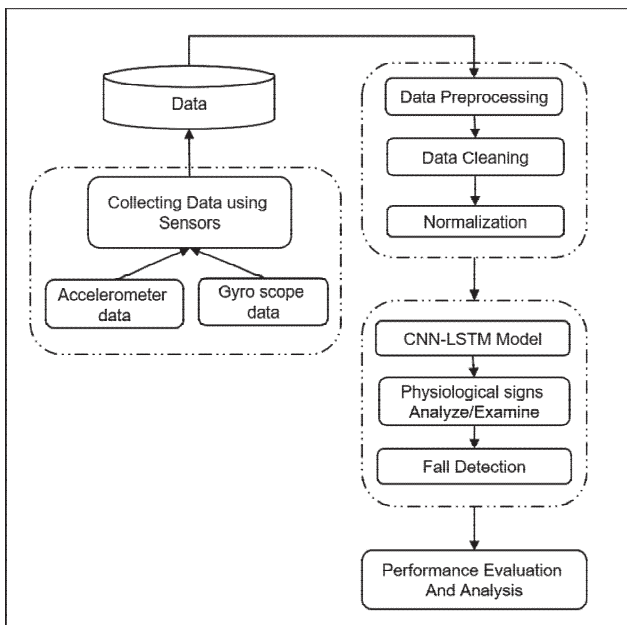


Figure 1. Proposed Model

### 2.1 Dataset

The data is collected using sensors' gyroscope and accelerometer. The motion data was specifically recorded using a mobile device equipped with the LSM330DLC inertial sensor (3D accelerometer and gyroscope) [4]. The dataset [5] acquired with sensors includes a daily living scenario, which is a combination of seven different Activities of Daily Living (ADL), and four different types of falls from a total of 66 individuals with more than 3200 trials. The dataset makes use of the Samsung smartphone which is integrated with sensor. A frequency of 200 Hz is used by the phone's sensors, gyroscope, and three-axis accelerometer to collect data [4]. When performing the fall, it was observed that the falling takes between two to four seconds, So the data is selected on this parameter as 4 sec at a frequency of 200 Hz.

Both sensors together with device sensor makes the recordings of data generated. For the purpose of data capture an Android application was developed that records

raw data for acceleration, angular velocity. This provides the highest possible sampling rate. The signals can be subsampled at any time if lower sampling rates are desired. For the fall simulation a relative hard mattress of 5 cm in thickness (as used in martial arts) was utilized to dampen the fall. To ensure that the individuals completed the proper fall in a realistic manner, the authors pay special attention on the participants on how to do each fall.

### 2.2 Data Preparation

Data preprocessing is essential in any machine learning application for increasing the model's effectiveness. When analyzing data, the first step is to ensure that the data is correct and cleaned it is the part of Cleaning. Four main cleaning processes for data from wearable sensors are suggested. The acceleration data should first have a low-pass filter applied to it to reduce noise. In order to find any extra incorrect data, or data that were not fixed by the automated filtering in step 1, the collected data should be visualized in order to be checked. Possible examples of inaccurate data include activities that are fall-like were listed first, which is based on certain requirements. In certain sequences, the subject is seen in a variety of positions, such as sitting on a chair or getting in and out of a car, and is often motionless at the end. Examples of abrupt or rapid movements that are analogous to falls include jumping and sprinting. Regular daily activities include walking, standing up, and ascending and descending stairs (sometimes known as "stairs up" and "stairs down"). Third, any observations that were made outside the experimental window should be removed and the data from the various sensors should be synced. The information is kept in ns along with its timestamp. The actions and subject data are stored in the dataset using an SQLite database.

As a result, the best models can be selected by sensing a greater grasp of the data. Once a model has been selected, the data can be pre-processed to enhance learning or the best understanding the model can have. An CNN-LSTM network that utilized floating and integers was trained using dataset. All of the data in the features were subjected to normalization in order to increase the machine learning model's accuracy. Before being used by the LSTM model, the data for each feature is modified. We employ the normalization procedure, which involves subtracting the minimum value from the data set and then dividing by the data set's range, to construct the recurrent neural network. Additionally, the preprocessing module contains the "MinMaxScaler" class, which is used to scale the data's range.

### 2.3 CNN-LSTM Model

After the preprocessing is done the data is ready to apply the model. The sensors send the sequence of the resulting acceleration data to the input layer. If and are the values from the 3-axis accelerometer, the resulting acceleration can be determined using the following equation:

$$a_r = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

Each convolution layer in the CNN section of the network is made up of a 1D convolution, a layer of linear activation, and a layer that maximizes pooling. Additionally, batch normalization layer [5] is used on the network to enhance model correctness and convergence speed. The goal of batch normalization is to ensure that the model's nonlinear expression capability while changing the variance's magnitude and mean position to make the new distribution more consistent with the data's actual distribution.

To make sure the model can learn enough features, two LSTM layers employed. Highly abstract features are generated by the CNN and LSTM layers, and these features are subsequently translated to the sample marker space by two completely linked layers and a Softmax layer. This network's loss function is provided as follows.

$$Loss = -\frac{1}{N} \sum_{i=0}^{N-1} [y_i \ln a_i + (1 - y_i) \ln(1 - a_i)] + \lambda \omega^2$$

where N represents number of categories, shows actual output of neurons, displays the expected output. To increase the network's capacity for generalization, an L2 regularization term is applied, is the regularization term's proportional parameter, and represents the network's weight.

### 3. Experimental Analysis

The suggested model will be tested against the dataset in this section to determine its effectiveness and viability.

#### 3.1 Setting of Dataset

The experiment includes, the data on 3-axis acceleration to determine the resulting acceleration after the dataset is decreased to create the signal with a 50 Hz sampling rate. In order to verify that the model uses the fewest possible data, additional information from the gyroscope and accelerometer is disregarded. Seven different types of Activities of Daily Living (ADL)s are chosen for the training of model, along with fall behavior. Figure 2 displays the acceleration curve of some activities from daily task.

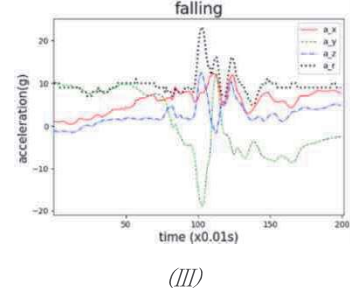
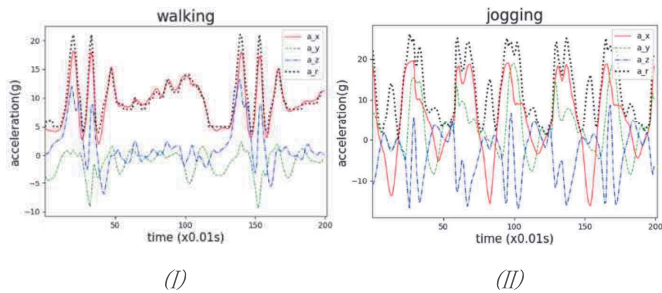


Figure 2. Curves of acceleration for various Activities of Daily Living (ADL): (I) walking acceleration curve, (II) jogging acceleration curve, (III) falling acceleration curve

Given that falling takes between two to four seconds, we take 200 samples starting 4 seconds the moment fall occurs. 5,890 train samples and 1474 test samples, with a total of 80% and 20% of the experimental data set, respectively, are acquired. 15% of the train set is utilized as a validation set during training in order to ensure the generalization ability of model.

#### 3.2 Setup of parameters

For the following experiment, the CNN-LSTM model is constructed using the deep learning framework, Keras. In Table 1, the parameters are displayed. In order to compare results, Additionally, we train a 1-D CNN and a pure LSTM with the same parameters.

Table 1: Setting up the training parameters.

Parameter	Value
Epoch	300
Optimizer	Adam
Batch size	32
Learning Rate	0.001

### 4. Result and Discussion

On the dataset, we tried a variety of models. Figure 3 depicts the training procedure's accuracy curve. The experimental findings indicate that the CNN-LSTM network has a higher accuracy than the 1D CNN and LSTM network. Additionally, as evidenced by the data, the CNN-based network converges significantly more quickly than the LSTM network.

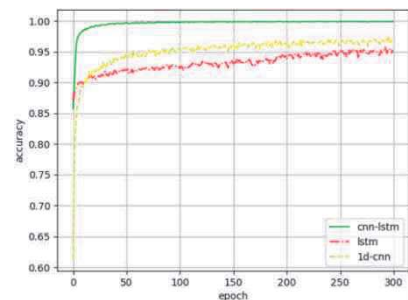


Figure 3. The training process's accuracy curve.

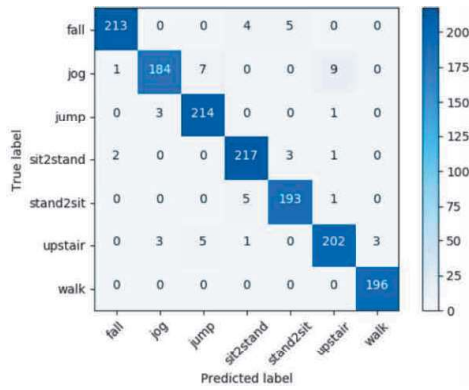


Figure 4. Confusion matrix.

Figure 4 displays a confusion matrix representing the test outcomes on the CNN-LSTM model. In seven different categories of daily behaviors, it is demonstrated that the classification accuracy for the CNN-LSTM model is 98.97%. The fall behavior has a detection precision and specificity of 98.61% and 99.77%, respectively. In order to support our study, we check the sensitivity, G-mean, and consistency to support our claim. With the sensitivity of 94.08% the G-mean is 94.08% while the consistency lies at 19.8%. In the meanwhile, we contrast the outcomes with a number of different algorithms tested on the dataset. We contrast the number of categories, sample frequency, and classification precision of ADL. It should be mentioned that some algorithms use binary classification methods, which means that the classification accuracy and fall detection precision are comparable.

In order to determine whether the model is appropriate for IoT applications, factors such as the number of categories used per second and as part of the comparison set, the relevant sensors are also displayed in the table. Table 2 displays each method's outcomes.

Table 2: Comparing various techniques for the detection of fall

Study	Sample frequency	No. of categories	Method	Accuracy in classifying	Sensors
(AkramBayat, 2014)	100 Hz	4	Support Vector Machine	91.20%	Accelerometer
(Ivan Miguel Pires, October 2017)	100 Hz	5	Deep Neural Network	89.51%	magnetometer gyroscope
(Fakhrulddin A H, November 2017)	50 Hz	2	Convolutional Neural Network	92.30%	Accelerometer
Ours	50 Hz	7	CNN-LSTM	98.97%	Gyro sensor Accelerometer

Table 3: Fall Detection Model Performance

Method	Accuracy	Sensitivity	Specificity	G-mean	Consistency
CNN-LSTM	98.97%	94.08%	94.02%	94.08%	19.8%

## 5. Conclusion

In this paper, we suggested a CNN-LSTM network-based fall detection algorithm. The dataset's average classification accuracy for seven different types of ADLs, including fall, is 98.97%. Additionally, the fall detection accuracy and specificity are 98.61% and 94.02%, respectively. Instead of the 3 axis acceleration data, the model requires the resulting acceleration, which results in a relatively small data volume. This study of creating a model for fall detection system can be shown to be an inventive technique to control injury and mortality rates by offering proper medical care in the event of an early fall detection. If used in the works below, the CNN-LSTM model can make building sites more convenient by spotting falls earlier.

## References

1. AkramBayat, M. D. (2014). A Study on Human Activity Recognition Using Accelerometer Data from Smartphones. *The 9th International Conference on Future Networks and Communications*, 450-457.
2. Fakhrulddin A H, F. X. (November 2017). Convolutional neural networks (CNN) based human fall detection on Body Sensor Networks (BSN) sensor data. *2017 4th International Conference on Systems and Informatics (ICSAI)*.
3. George Vavoulas, C. C. (2016). The MobiAct Dataset: Recognition of Activities of Daily Living using. *International Conference on Information and Communication Technologies for Ageing Well and e-Health*, (p. 9).
4. Ivan Miguel Pires, N. G. (October 2017). Identification of Activities of Daily Living through Data Fusion on Motion and Magnetic Sensors embedded on Mobile Devices. 19.
5. Lab, B. (2016). *Biomedical Informatics and eHealth Laboratory*. Retrieved from <https://bmi.hmu.gr/the-mobifall-and-mobiact-datasets-2/>
6. Sergey Ioffe, C. S. ( 2 Mar 2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.
7. Unit, W. A. (Geneva : WHO, 2008.). *WHO global report on falls prevention in older age / Ageing and Life Course, Family and Community Health*.

# 시계열 생체 데이터를 RP신호로 변환하기 위한 효과적인 방법

전영준<sup>o</sup>, 강순주

경북대학교 전자전기공학부

yjjeon@knu.ac.kr, sjkang@ee.knu.ac.kr

## Effective method for converting time-series bio-signal into RP images

YeongJun Jeon<sup>o</sup>, SoonJu Kang

Electronic and Electrical Engineering at Kyungpook National University

### 요 약

본 논문은 시계열 데이터로 측정되는 생체신호들을 2D CNN 모델에서 활용하기 위해서 Recurrence Plot으로 변환할 수 있는 효과적인 방법을 보인다. 대부분의 bio-signal을 분석하는 연구들은 가속도 신호 및 ECG 데이터를 기반으로 하고 있다. 하지만, 이러한 신호에서 우리가 원하는 anomaly detection, person identification, behavior classification를 수행하기 위해서 time series 데이터를 1D CNN에서 학습하는 방식은 한계가 존재한다. 데이터 자체가 간단하기 때문에 특징을 다양하게 추출하기 힘들며, 에러 데이터가 포함될 때 더욱 큰 문제를 야기할 수 있다. 본 연구에서는 이를 개선하기 위해서 time series 데이터를 2D image로 변환할 수 있는 Recurrence Plot의 효과적인 적용 방법을 제안한다. 변환 후의 데이터는 에러 데이터가 포함되었을 때도 일관성이 있는 결과를 보임을 확인하였으며, 주기성이 있는 생체 신호에서 효과적으로 쓰일 수 있는 방법임을 보인다.

### 1. 서 론

Big data 시스템이 전세계적으로 구축되고 임베디드 시스템이 발전함에 따라 이를 활용한 다양한 분석방법이 연구되고 있다. 이 중 생체신호 데이터를 이용한 다양한 연구들이 수행되고 있으며, 이 중에서 ECG를 활용한 연구들은 매우 활발히 이루어지고 있고, 건강 상태를 분석하거나 사람을 식별하고, 이상 신호를 감지하고 이를 해결하는 등 다양하게 이루어지고 있다 [1][2]. 그리고 이와 관련된 대부분의 연구들은 공개데이터셋을 이용하고 있으며 여기에 사용되는 대다수의 데이터는 실험 환경에서 측정된 이상적인 데이터만으로 이루어져 있다. 하지만 이상적인 데이터만을 활용한 연구로 분석된 결과는 일상환경에서 실시간으로 측정된 데이터를 상대로는 정확도를 보장하기 어렵다. 특히, 측정 대상은 주로 노년층으로 대부분 스마트기기에 익숙하지 못하다. 즉, 일상환경에서 측정된 데이터에서는 센서의 착용 상태, 움직임 등으로 인해 데이터에 왜곡된 형태가 많이 발생하게 된다. 본 연구에서는 이처럼 왜곡된 형태가 있는 데이터를 포함한 데이터에서도 효과적으로 머신러닝 및 딥러닝 기법을 적용할 수 있도록 하는 전처리 방법인 Recurrence Plot의 효과적인 적용방법을 소개한다. 본 연구에서는 PPG 신호를 주로 다루었으며, 광적용맥파(photoplethysmogram, PPG) 신호는 심장의 수축과 이완으로 변화하는 혈액용적(blood volume)과 혈액내의 헤모글로빈에 흡수되는 빛의 양의 선형적 관계를 이용하여 신호를 획득하는 방법으로, 맥박 또는 혈중 산소포화도를 측정할 수 있어 많은 분야에서 활용되고 있다 [3][4]. 최근에는 ECG뿐 아니라 PPG도 건강에 관련한 요소와 개인별 특성을 파악할 수 있는 신호

로 활발히 연구되고 있다 [5]. 우리는 자체 개발한 웨어러블 헬스케어 기기인 Sleepcare Kit로 다수의 사람의 일상생활 중의 PPG신호를 포함한 다양한 생체신호를 측정하였으며, 이를 이용한 사람을 식별할 수 있는 연구를 진행했다 [6]. 본 논문에서는 이 연구를 진행하는 과정 중 일상환경에서 측정된 PPG 신호를 활용하기 위한 전처리 방법으로 Recurrence Plot을 효과적으로 사용할 수 있는 방법을 제안한다.

### 2. 연구 내용

본 연구에서 제안하는 전처리 방법인 Recurrence Plot은 정형(Tabular) 데이터를 이미지화 혹은 텍스트화 할 때 더 많은 종류의 딥러닝 알고리즘들에 적용할 수 있어 더 좋은 성과를 거둘 수 있다는 취지로 개발된 것이다 [7]. 즉, 정형 데이터를 비정형(이미지, 텍스트, 음성...)처럼 학습하기 위한 방법으로 제안된 것이다. Recurrence Plot에서 주로 사용되는 것은 Cross-Joint Recurrence Plot(CRP)로 수식은 Eq(1)과 같다 [8].

$$CRP_j = \theta(\vec{a} - \|\vec{x}(i) - \vec{x}(j)\|).....(1)$$

여기서  $\vec{x}(i)$ 는 시간  $t$ 에 대해 위상 공간으로 재구성된 벡터이며,  $\vec{x}(i)$ 와  $\vec{x}(j)$ 사이의 거리를 계산한 후 지정된  $\epsilon$ 와 비교한다. 이 때 Heaviside 함수( $\theta$ )를 통해 비교한 값이 임계 값보다 크면 1, 그렇지 않으면 0을 출력하게 된다. 이는 간단하게 구현될 수 있는 방법이지만 실제 적용한 결과를 학습할 때는 좋은 성능을 이끌어내기 힘들다. 이는 0과 1만으로 표현됨으로 인

해 RP 이미지가 단순하게 표현되므로 다양한 특징을 추출하기 힘들기 때문이다. 특히, 이 수식에서는  $\epsilon$ , 임계 값을 설정하는 것이 매우 중요하다.

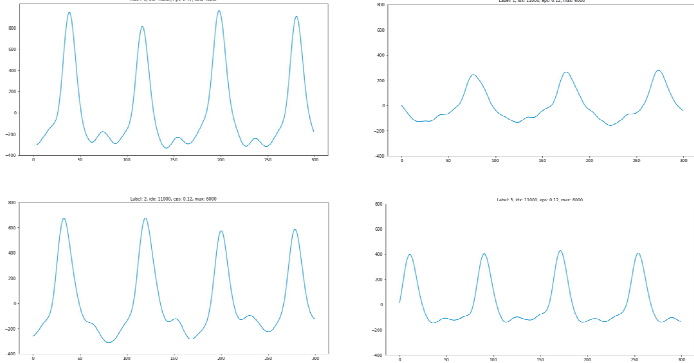


그림 1. Butterworth 필터를 거친 후 개인 별 PPG 신호

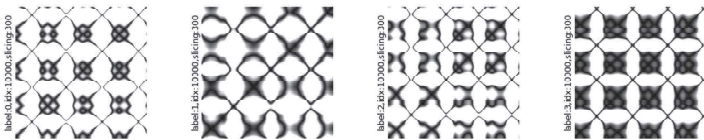


그림 2. 그림 1의 PPG 신호를 RP 변환한 모습

본 연구에서는 이를 PPG신호에 적용하기 위한 최적의  $\epsilon$  값을 찾는 과정과 적용한 결과를 분석한다. 일반적으로 PPG 신호는 측정된 결과를 Butterworth 필터를 거친 후의 결과를 사용하며, 이 모습은 그림 1의 그림에서 확인할 수 있다 [9]. 그림 1은 y축을 고정한 후 출력한 그림이며, 사람에 따라서 다른 형태를 가짐을 확인할 수 있다. PPG 신호는 Peak 파형 이후에 Notch 구간이 일반적으로 발생하게 되는데 이 형태가 올바르게 나타남을 볼 수 있으며, 모든 신호에서 높낮이의 차이는 있지만 특성이 각각 드러남을 확인할 수 있다. 이러한 이상적인 신호를 RP로 변환한 모습은 그림 2에서 보인다. 주기성이 있는 신호이기 때문에 타일의 형태처럼 나타나는 것을 확인할 수 있다. 각 RP는 time series일 때보다 훨씬 각자의 특징이 드러남도 보인다. 특히 그림 1에서 왼쪽 위 아래의 그래프의 차이를 육안으로는 확인하기 쉽지 않은데 RP로 변환한 결과에서는 무늬 사이의 표식모양에서도 차이가 두드러짐을 보인다. 그리고, 이러한 변환이 효과적인 가장 큰 이유는 왜곡된 신호가 포함된 신호에서도 견고한 결과를 출력한다는 점이다. 그림 3은 이를 보이며, 3-t-(a)는 200 지점에서 한번 큰 왜곡 신호가 발생함을 볼 수 있다. 이러한 상태는 일상생활에서 측정 시 순간적인 움직임 또는 센서의 착용 상태에 따라서 흔하게 발견되는 데이터이다. 하지만 이러한 데이터가 포함된 time series 데이터 셋을 1D CNN 모델에서 학습하면 학습 성능이 매우 떨어지게 된다. 하지만 위 신호를 RP로 변환하면 3-rp-(a)에서 보이듯이 왜곡 신호가 발생한 부분이 있더라도 3-rp-(b)와 비교하여 특징을 올바르게 추출할 수 있음을 확인할 수 있다.

그럼, RP로 변환하기 위한 최적의 hyper parameter를 설정하는 방법은 분석할 데이터의 특성을 파악한 상태에서

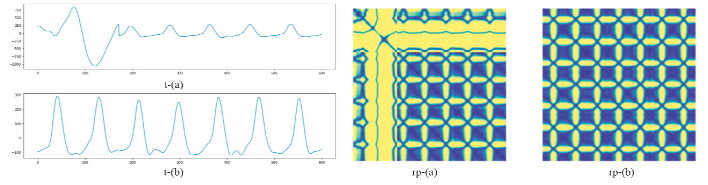


그림 3. 데이터 비교; t: time series, rp: recurrence plot;(a) 왜곡된 데이터가 포함된 RP; (b)이상적인 데이터의 RP

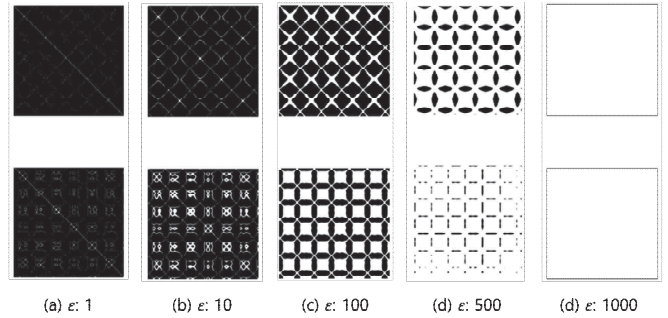


그림 4.  $\epsilon$  에 따른 RP 변환 후의 모습 비교

값을 세팅하는 것이다. 본 연구에서는 PPG 신호가 filter를 거친 후의 신호의 값을 기반으로  $\epsilon$  및 threshold 값을 적용한 결과를 보인다. 그림 4는 이를 보인 그림으로  $\epsilon$  값에 따라 정보의 소실이 얼마나 발생하는지를 파악할 수 있다. 당연하게도 그림 4-(a)와 4-(b)처럼 작은  $\epsilon$  값에서는 제대로 된 특징을 추출하기 힘든 모습을 보인다. 그리고 그림 4-(c)에서 보듯이  $\epsilon$  값이 신호에 비해 커지게 되면 특성이 점점 사라짐을 확인할 수 있고  $\epsilon$  값이 1000이 되었을 때는 모두 1로 표현이 되기 때문에 특성을 뽑아낼 수 없다. 즉, 본 연구에서 사용된 PPG 센서의 데이터에서의 최적의  $\epsilon$  값은 100전후가 될 것이다. 이를 판단하기 위한 가장 쉬운 기준은 1과 0의 비율이 1:1에 가장 가까운  $\epsilon$  값을 찾는 것이다. 실제 변환한 데이터에서도  $\epsilon$ 가 100인 그림 4-(c)에서의 1과 0의 비율은 0.45~0.58정도 사이인 것으로 나타났다. 만약 다른 생체신호 및 time series 데이터를 가공할 때에도 이러한 기준을 토대로  $\epsilon$  값을 설정한다면 최적의 값을 찾을 수 있을 것이다. 또한, 여러 label의 데이터를 변환한다고 할 때도 모두 같은  $\epsilon$ 를 사용해야 하는 것은 아니다. 만약, 데이터에 따라  $\epsilon$  값을 조금 더 미세하게 조정하는 과정은 추후 학습과정에도 용이하게 작용할 수 있다.

### 3. 결 론

본 연구에서는 time series 생체신호를 RP변환 했을 때의 이점과 어떻게 변환하는 것이 효과적인지에 관한 내용을 주로 다루고 있다. 본 논문에서는 다루지 않았지만 time series 데이터를 1D CNN 모델에 학습한 것과 RP로 변환한 데이터를 2D CNN 모델에서 학습한 결과를 비교한 결과 RP로 변환한 것이 10~15% 정확도가 향상됨을 확인했다. 또한 RP 데이터를 사용하였을 때는 데이터에 왜곡이 포함된 상황이 많은 데이터셋에서는 더 큰 강점을 보인다. 본 논문에서는 CRP를 기반으로 한 hyper parameter를 설정하는 방법을 간단히 다루었다. 하지만 상황에 따라 joint recurrence plot (JRP) 방법을 적용하거나 또는 RP 수식을 본인의 연구에 맞게 변형하여 적용하면 보다 효

과적인 RP 이미지를 얻을 수 있다.

### 참 고 문 헌

- [1] 이은혜. “다양한 상황에 대한 심전도 신호를 이용한 개인 식별 및 인증.” Diss. 한양대학교, 2017.
- [2] Li, Hongzu, and Pierre Boulanger. “A survey of heart anomaly detection using ambulatory Electrocardiogram (ECG).” *Sensors* 20(5), 1461, 2020
- [3] 정인철. “PPG 신호를 이용한 심혈관 기능 및 혈압의 추정에 관한 연구.” Diss. 연세대학교 대학원, 2005.
- [4] Song, Je-Min, et al. “Design and implementation of a prediction system for cardiovascular diseases using PPG.” *Journal of the Korean Society of Radiology* 5(1), 19-25, 2011
- [5] Everson, Luke, et al. “BiometricNet: Deep learning based biometric identification using wrist-worn PPG.” 2018 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2018.
- [6] Jeon, Y. J., & Kang, S. J. “Wearable sleepcare kit: Analysis and prevention of sleep apnea symptoms in real-time.” *IEEE Access*, 7, 60634-60649, 2019
- [7] Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. “Recurrence-plot-based measures of complexity and their application to heart-rate-variability data.” *Physical review E*, 66(2), 026702, 2002
- [8] Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. “Recurrence plots for the analysis of complex systems.” *Physics reports*, 438(5-6), 237-329, 2007
- [9] Chatterjee, Ayan. “Calculate Pulse from Touch Error Free PPG Signal with 2 nd Order Butterworth Filter.” 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). IEEE, 2021.



# 염색 공정 데이터를 이용한 친환경 섬유 소재

## 내광성 등급 예측 모델 개발

서강복<sup>○</sup> 김덕엽 이우진

경북대학교 IT대학 컴퓨터학부

dating1227@gmail.com, ejrduq77@naver.com, woojin@knu.ac.kr

### Development of light resistance class prediction model

### for eco-friendly textile materials using dyeing process data

Kang Bok Seo<sup>○</sup> Deok Yeop Kim Woo Jin Lee

School of Computer Science and Engineering, Kyungpook National University

#### 요 약

최근 노동 집약적인 산업 분야인 섬유 산업에서는 AI를 통해 공정 과정을 더욱 효율적으로 개선하기 위한 시도를 하고 있다. 그러나 섬유 산업 분야의 경우 데이터를 수집하기 위해 소요되는 시간과 비용이 크기 때문에 다양한 조합의 데이터를 수집하기 어렵다. 특히, 염색 공정의 경우 염색 초기 과정부터 친환경 섬유 소재의 내광성을 확인하는 과정까지 오랜 시간이 소요된다. 내광성은 염색 공정에 존재하는 다양한 특성에 영향을 받으며 특정 데이터만 참고해서는 그 결과를 예측하기가 쉽지 않다. 본 논문에서는 염료 공정에 존재하는 데이터를 Feature Selection Method를 적용하여 친환경 섬유 소재의 내광성 등급과 관련이 있다고 판단되는 특성들을 추출하고 MLP 모델에 학습시켜 친환경 섬유 소재의 내광성 예측 모델을 개발한다. 학습된 모델로 내광성 등급을 예측하고 예측 결과와 데이터 분석 과정을 통해 향후 연구 방향을 제안한다.

#### 1. 서 론

최근 노동 집약적인 산업 분야 중 하나인 섬유 분야에 AI를 적용하기 위한 시도[1]가 이루어지고 있다. 노동 집약적인 성격이 강한 분야에 AI를 적용하면 생산 시스템을 최적화할 수 있어 제품 생산이 더욱 효율적으로 이루어질 수 있다. 특정 분야에 AI 모델을 적용하기 위해서는 모델 학습에 사용되는 데이터의 양과 품질을 고려해야 한다. 섬유 산업 분야의 경우 데이터를 수집하기 위해 소요되는 시간과 비용이 크기 때문에 "다양한 조합의 데이터를 수집하기 어렵다. 특히 염색 공정의 경우 염색 초기 과정부터 친환경 섬유 소재의 내광성을 확인하는 과정까지 시간이 오래 요구되어 더욱 데이터를 수집하기 어렵다. 내광성은 자외선에 노출되었을 때 재료의 성능 저하에 대한 저항성을 나타내는 수치로 9개의 등급으로 이루어져 있다. 내광성 등급은 염색 공정에 존재하는 다양한 특성에 영향을 받으며 특정 데이터만을 참고해서는 등급을 예측하기가 쉽지 않다.

이에 본 논문에서는 염료 공정에 존재하는 다양한 데이터를 분석하고 친환경 섬유 소재의 내광성 등급과 관련이 있는 데

이터들을 추출하여 내광성 등급을 예측하고자 한다. 예측 성능 향상을 위해 다양한 Feature Selection Method를 적용하여 학습에 사용될 주요 특성을 선정하여 학습을 진행하였다. 내광성 등급의 예측은 MLP(Multi-Layer Perceptron)기법을 적용하였다. 데이터 분석과정과 예측 결과를 바탕으로 내광성 등급 예측에 대한 향후 연구 방향을 제안한다.

#### 2. 관련 연구

기존 섬유 산업 연구에서는 품질 개선을 위해 통계적으로 데이터를 분석하거나 수학적 모델링으로 데이터를 분석하였다. 하지만 통계적 분석 방식[2]은 단순한 선형 모델의 경우에는 유용하지만 복잡한 비선형 모델의 경우 적용하기 어렵다는 단점이 있다. 또 수학적 모델링[3]의 경우 관련 수식이 있거나 구조화가 용이할 정도로 모델이 단순한 경우는 유용하지만 그렇지 않은 경우 수학적 모델링 자체를 할 수 없다는 단점이 있다. 이 외에도 30개 정도의 매우 적은 데이터를 AI에 적용시켜 면방직사의 강도를 예측한 연구[4]도 있다. 이는 이른 시기에 AI를 섬유 산업에 적용시켰다는 의의가 있으나 30개 데이터가 모집단을 대표한다고 보기 어렵기 때문에 이를 통해 만들어진 모델의 신뢰도가 떨어진다. 기존 섬유 산업 관련 연구들 중 AI 적용 사례가 많지 않은 것은 학습에 사용될 충분한 데이터를 확보하기 어렵고, 축적된 데이터의 양이 많지 않기 때문이다. 또한 데이터를 수집하더라도 수집하는 데이터에 대

※ 본 연구는 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0022335).

한 분석이 충분히 이루어지지 않아 활용하기가 어렵기 때문이다.

### 3. 염색 공정 데이터

본 논문에서 활용하는 데이터는 1043개의 염색 공정 데이터이다. 이 데이터는 231개의 염색 공정 변수와 1개의 예측 대상인 내광성 등급으로 이루어지며 데이터의 구성은 표 1과 같다.

표 1. 염색 공정 데이터 구성

구분	공정명	변수 개수(개)
소재 정보	섬유 소재 정보	21
가공 공정	전처리(Pilot)	87
	배합(Lab)	39
	염색(Lab)	24
	후처리(Lab)	45
	후가공(Lab)	15

염색 공정 데이터는 크게 소재 정보, 가공 공정(Pilot), 가공 공정(Lab) 3부분으로 구성된다. 이 중 섬유 소재 정보 20개, 가공 공정(Pilot) 85개, 가공 공정(Lab) 37개 변수가 단일 값 변수로 수집되었다. 하지만 데이터가 추가로 수집되면서 많은 변수가 이전과는 다른 형태로 수집되었다. 데이터가 추가될 때마다 이전에 수행했던 데이터 분석과는 다른 분석 결과가 나타났다. 이에 현재 수집된 데이터에서 단일 값 및 올바르게 수집되지 않은 데이터를 제외하고 41개의 특성으로 데이터 분석에 적용했다.

### 4. Feature Selection Method[5]를 주요 특성 추출

내광성 예측에 활용될 데이터를 학습에 활용하기 위해 Feature Selection Method에 존재하는 Filter, Wrapper, Embedded Method를 적용하여 데이터 분석을 진행하였다. 표 2는 Filter Method에서 주로 사용되는 피어슨 상관 계수를 적용하여 절대 값 기준으로 상위 10개 특성을 나타낸 것이다.

표 2. 피어슨 상관 계수 Top 10

순위	상관 계수	특성 명칭
1	0.429019	Lab 염색 하강속도 #1
2	0.298021	배합_Fadex® TS liq
3	0.280198	Dorospense Yellow KKL_prop
4	0.276451	배합_UVK-200
5	0.26677	Dorospense Yellow KKL
6	0.262747	배합_Sera Fast P-UVC
7	-0.223441	Dorospense Blue KKL_prop
8	-0.212875	Dorospense Red KKL_prop
9	-0.211858	Dorospense Red KKL
10	0.203054	배합_UVK-400 H/C

표 2에서 나타나듯이 내광성 등급과 상관 계수가 높은 데이터의 경우 ‘Lab 염색 하강속도 #1’을 제외하고는 모두 ‘배합(Lab)’ 공정에 존재하는 변수이다. 주로 분석해야 될 특성들이 ‘배합(Lab)’ 공정에 속해있는 것으로 나타난다. 표 3은 Embedded Method 중 하나인 랜덤 포레스트 모델을 활용한 특성의 중요도를 기준으로 상위 10개의 특성을 나타낸 것이다.

표 3. 랜덤 포레스트 특성 중요도 Top 10

순위	특성 중요도	특성 명칭
1	0.063784	Lab 염색 상승온도 #3 유지시간
2	0.042533	Dorospense Yellow KKL
3	0.042154	Lab 염색 상승속도 #3
4	0.042049	Dorospense Dark Grey KKL
5	0.038195	Dorospense B/K KKL
6	0.036994	Dorospense Brown K-3LR
7	0.034325	Dorospense Red KKL
8	0.034169	Dorospense Blue KKL
9	0.027517	제직_중량_생지
10	0.027392	Lab 염색 상승속도 #2

표 3에서 나타나듯이 ‘배합(Lab)’ 공정에 속한 특성들이 가장 높은 비율을 가지고 있으며 피어슨 상관 계수와 달리 ‘염색(Lab)’ 공정에 속한 변수가 추가되었다. 마지막으로 표 4는 Wrapper Method 방법 중 하나인 RFE(Recursive Feature Elimination)를 적용하여 상위 10개의 특성을 나타낸 것이다.

표 4. RFE Top 10

순위	특성 명칭
1	Dorospense Yellow KKL
2	Dorospense Red KKL
3	Dorospense Blue KKL
4	Dianix Yellow AM-2R
5	Dorospense B/K KKL
6	Dorospense Dark Grey KKL
7	Dorospense Brown K-3LR
8	Lab 염색상승속도#3
9	Lab 염색상승온도#3유지시간
10	Lab 후처리환원제정용수량

표 4에 나타난 RFE의 특정한 수치 기준이 아닌 중요도 높은 특성들을 부분집합으로 나타내며 해당 분석에서도 ‘배합(Lab)’ 공정에 속한 특성들이 가장 높은 비율로 나타났다. ‘배합(Lab)’ 공정에 속한 변수들은 대부분 염색에 사용되는 염료의 종류와 투여량이며 해당 특성들이 3개의 분석 결과에 모두 공통적으로 포함되어 있으며 이는 수집된 데이터에서 내광성 등급에 가장 많은 영향을 주는 주요 특성으로 판단된다.

### 5. MLP 기법을 이용한 모델 학습과 예측

내광성 예측하는 모델의 학습은 수집된 데이터 1043개의 데이터를 사용하여 진행했다. 해당 데이터 중 무작위 102개는 예측용 테스트 데이터로, 나머지를 모두 학습용 데이터로 사

용했다. 학습 모델은 MLP(Multi-layer Perceptron)를 사용하였으며 사용 옵션은 표 4와 같다.

표 5. 학습 및 예측 모델 실험 조건

구분	범위	최적화 값
Layer	36 ~ 512	144
Dropout	0 ~ 0.8	0.5
Optimizer	Adam, RMSprop, Nadam	Adam
regularizer	L1, L2	L2
Activation Function	selu, relu	relu
Learning rate	0.0001 ~ 0.001	0.0035

예측 성능 향상을 위해 다양한 조건에서 학습을 수행하였으며 표 4에서 나타난 최적화 값을 사용하였을 때 가장 높은 성능을 보였으며 이때 Accuracy\_Score가 71%로 비교적 낮은 성능을 보여준다. 이는 수집된 데이터가 편향되어 있고, 수집된 데이터의 양이 부족하여 나타난 것으로 판단된다. 표 5는 학습한 모델로 테스트 데이터를 이용하여 예측한 Confusion matrix 결과이다.

표 6. 내광성 등급 예측 Confusion Matrix

등급	1	1-2	2	2-3	3	3-4	4	4-5	5
1	13	1	0	0	0	1	0	0	0
1-2	1	11	2	0	1	0	0	0	0
2	1	2	9	2	0	0	0	0	0
2-3	0	2	2	11	0	0	0	0	0
3	0	1	0	4	12	1	0	0	0
3-4	0	1	0	2	0	8	3	0	0
4	0	0	0	0	0	0	7	1	0
4-5	0	0	0	0	0	0	1	1	0
5	0	0	0	0	0	0	0	1	0

표 5에서 나타나듯이 각 등급에 따른 성능 차이가 많이 나타난다. 이는 수집된 데이터가 편향되어 있어 4-5, 5등급의 경우 매우 저조한 성능을 보여준다. 이 두 등급은 전체 데이터에서 차지하는 비율이 1%미만으로 해당 등급들은 올바른 학습이 이루어지지 않고 있으며 해당 등급의 데이터 수가 매우 부족하여 발생한 것으로 데이터가 추가적으로 수집되면 해결 될 것으로 판단된다. 위 두 등급을 제외한 데이터의 경우 비교적 많은 양의 데이터가 수집되어 좀 더 나은 성능을 보여준다. 특히 1~3등급 사이의 데이터는 전체 데이터의 80%이상을 차지하고 있어 데이터가 추가적으로 수집되어 추가적인 분석을 통해 더욱 성능을 높일 수 있을 것으로 판단된다.

## 6. 결론 및 향후 연구

본 논문에서는 친환경 섬유 소재의 염색 공정 데이터 분석

을 통해 내광성 등급과 관련 있는 주요 특성들을 선정하고 그 데이터를 MLP 모델에 학습시켜 내광성 예측 모델을 개발하였다. 데이터 분석은 Feature Selection Method에 존재하는 3가지 분석 방법(Filter, Wrapper, Embedded)을 적용하였으며, 이를 통해 선택된 주요 특성들을 중점으로 학습을 진행하였다. 학습에는 MLP(Multi-Layer Perceptron) 모델을 사용하였으며 다양한 조건에서 학습을 수행한 결과 약 71%의 정확도를 보여 주었다. 수집된 내광성 등급 데이터는 9개 등급 중 4-5, 5 등급의 데이터가 매우 부족하여 편향된 데이터였으며 이 두 등급에 대한 예측 성능은 매우 낮게 나타났다. 하지만 나머지 등급의 경우 비교적 균등하게 데이터가 수집되어 있어 성능도 높게 나타났다. 따라서 부족한 등급의 데이터를 추가적으로 확보하고 각 데이터에 대한 분석이 더욱 자세히 이루어지면 내광성 등급 예측 모델 성능을 더욱 개선시킬 수 있을 것으로 판단된다.

## 참고 문헌

- [1] Zhou, Z. et al, "Fabric defect detection based on feature fusion of a convolutional neural network and optimized extreme learning machine", Textile Research Journal, Vol. 92, No. 7-8, pp. 1161-1182, 2022.
- [2] 조기환, 정성훈, "P-71 : PET single filament 데이터 로부터의 번들강도 결정을 위한 통계적 접근", 한국섬유공학회 학술발표논문집, Vol. 2003, pp. 291-292, 2006.
- [3] 허유, 김종성, "콤팩 형성에 따른 링 정방장력 거동의 실험적 모델링", 한국섬유공학회지, Vol. 39, pp. 209-216, 2002.
- [4] 전봉수, 양철곤, "회귀분석과 신경망을 이용한 면방적사의 강도 예측", 한국섬유공학회지, Vol. 34, pp. 731-738, 1997.
- [5] Tang, Jiliang, Salem Alelyani, and Huan Liu. "Feature selection for classification: A review." Data classification: Algorithms and applications (2014): 37.

# 블루투스 센서와 자기장 센서를 혼합한 실내 측위 방법

한승준<sup>o</sup>, 권영우

경북대학교 컴퓨터학부

dk07564@knu.ac.kr, ywkwon@knu.ac.kr

## Indoor Positioning Method of Combining Bluetooth Sensor and Magnetometer

Seungjun Han<sup>o</sup>, Youngwoo Kwon

School of Computer Science and Engineering, Kyungpook National University

### 요약

실내 위치 기반 서비스를 제공하기 위해 GPS를 대체할 수 있는 실내 측위 기법이 연구되고 있다. 스마트폰의 보급으로 블루투스 센서, 지자기 센서와 같은 스마트폰의 센서를 측위에 활용할 수 있다. 그러나 센서 하나를 이용하는 단일 측위 방법은 한계점이 존재하며 이를 해결하기 위해서 다른 센서와 혼합하여 사용할 필요가 있다. 블루투스 센서와 자기장 센서를 혼합하여 블루투스 비콘의 RSSI 값을 이용해 자기장 세기 참조 범위를 줄여 실내에서 위치를 추측하는 방법을 제안하고자 한다.

### 1. 서론

실외에서 사용할 수 있는 위치 기반 서비스는 GPS와 같은 시스템을 이용하여 높은 정확도에 기반한 서비스가 가능하지만, 실내에서는 GPS 신호가 왜곡되어 사용하기 힘들다는 문제점이 있다. 따라서 실내에서 위치 기반 서비스를 제공하기 위해서 다른 측위 기술이 필요하다.

최근에는 스마트폰의 보급으로 스마트폰 센서를 이용하면 추가적인 장비를 설치할 필요 없이 측위 기술을 사용할 수 있게 되었다. 대표적으로 블루투스 센서를 이용하여 블루투스 신호 세기 RSSI(Received Signal Strength Indication)[1]를 사용하는 방법이나, 지자기 센서를 이용하여 철골 구조나 엘리베이터와 같은 구조물로 인해 지역적으로 고유하게 생성되는 자기장을 이용하여 핑거프린팅을 사용하는 방법이 있다[2]. 블루투스 신호와 자기장 세기는 서로 간섭을 일으키지 않고, 스마트폰에 블루투스 센서와 지자기 센서가 설치되어 있어 실내 측위에서 함께 사용할 수 있다[3].

본 논문에서는 스마트폰의 블루투스 센서와 자기장 센서를 결합하여 블루투스 센서를 이용한 방법에서 발생하는 신호 불안정, 기기 이상 문제와 자기장 핑거프린팅에서 참조 지점이 다수 개 존재하는 문제를 해결하는 방법을 제안한다.

### 2. 관련 연구

#### 2.1 블루투스 RSSI

블루투스 RSSI는 -99dBm에서 35dBm까지의 값을 가진다. 값이 높을수록 스마트폰에서 가깝다는 의미이다. 블루투스 RSSI 값을 이용하면 블루투스 비콘과 신호를 수신하는 스마트폰 사이의 거리를 측정할 수 있다. 식 1은 스마트폰에서 비콘이 송출하는 전파 신호 세기를 수신할 때의 값이고, 식 2는 식 1을 이용하여 거리  $d$ 를 측정하는 식이다.

$$RSSI = -10n \log(d) + txPower \quad (1)$$

$$d = 10^{\frac{txPower - RSSI}{10n}} \quad (2)$$

$n$ 은 경로 손실을 의미하며 2에서 4 사이의 값을 가지는 상수이다.  $txPower$ 는 비콘에서 송출하는 신호의 세기이며 스마트폰이 신호를 수신하면 RSSI 값으로 환산된다.

하나의 비콘으로부터 얻은 거리 값은 정확한 측위 결과를 제공하지 못하므로 최소 3개를 이용하는데 실내 공간에서는 일부 블루투스 신호가 벽이나 장애물에 의해 왜곡되거나 기기 손상되었을 경우 측위 결과에 오차를 발생시킬 수 있다.

#### 2.2 자기장 핑거프린팅

실내에는 건축자재, 장애물 등에 의해 고유한 자기장 영역이 생성된다. 이 점을 이용하면 자기장 핑거프린팅을 이용할 수 있다. 핑거프린팅은 크게 오프라인 단계, 온라인 단계 2개의 단계로 나뉜다. 오프라인 단계에서는 전체 공간을 일정한 단위로 나누어 작은 공간 단위인 셀(cell)로 분할한다. 각 셀에서 자기장 세기를 수집하여 위치좌표와 함께 데이터베이스에

저장하여 자기장 맵을 구성한다. 온라인 단계에서는 측정되는 자기장 세기를 오프라인 단계에서 구성한 자기장 맵과 비교하여 가장 유사한 지점을 찾아 해당 좌표로 위치를 추측한다.

자기장 핑거프린팅에서는 실내 공간에 유사한 자기장 세기 값을 가지는 셀이 다수 존재할 수 있다. 따라서 측위 오차를 줄이기 위해 수집되는 자기장 세기를 바탕으로 셀을 추측할 때 자기장 맵에서 참조할 영역을 최소화할 필요가 있다. 또한 자기장 세기는 스마트폰의 지자기 센서에 따라 값이 다르게 측정될 수 있다. 그림 2는 총 3개의 스마트폰을 이용해 전체 88개의 셀에서 자기장 세기를 수집한 후 각 셀에서 수집한 자기장 세기의 평균을 그래프로 나타낸 것이다. 각 셀에서 자기장의 세기는 다르지만 전체 자기장 세기의 분포는 유사하였다.

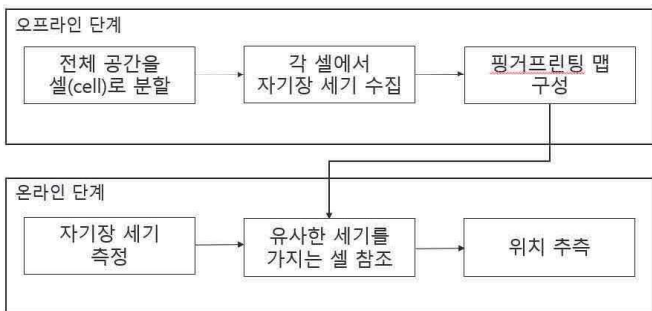


그림 1. 자기장 핑거프린팅

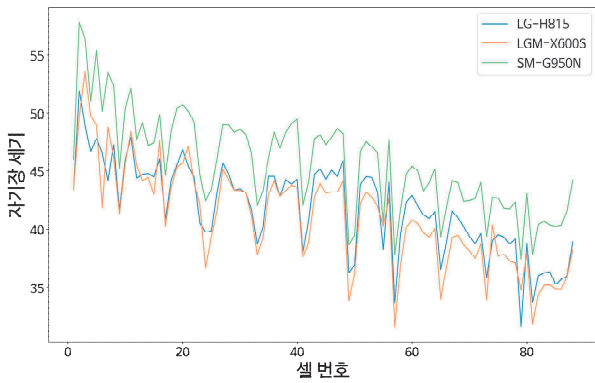


그림 2 (a). 보정 전 셀 별 자기장 세기

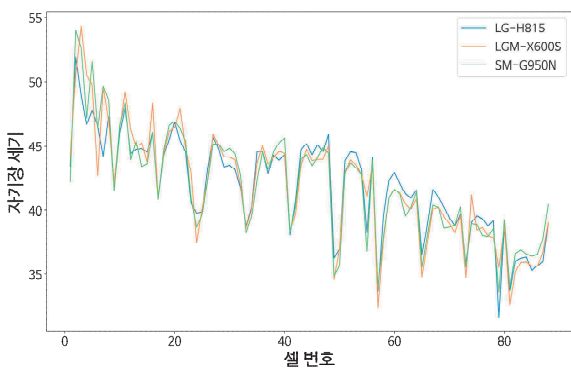


그림 2 (b). 보정 후 셀 별 자기장 세기

### 3. 제안 방법

제안하는 알고리즘은 블루투스 RSSI와 자기장 핑거프린팅을 혼합한 알고리즘이다. 블루투스 RSSI를 이용해 비콘으로부터 거리를 측정하고, 해당 거리만큼 떨어진 셀들 중 측정되는 자기장 세기와 가장 유사한 값을 가지는 지점으로 위치를 추측하는 방법이다. 그림 3은 제안하는 알고리즘의 개요이다.

자기장 핑거프린팅의 오프라인 단계와 같은 방식으로 전체 실내 공간을 일정한 셀로 분할하고 각 셀에서 자기장 세기를 측정하여 자기장 맵을 구성한다. 이후 측위를 할 때 0.5초 동안 받아들인 비콘 신호 중 평균 세기가 가장 강한 비콘을 선택하고, 그 비콘이 가장 신호를 적게 보낸 비콘이 아니면 해당 비콘을 이용하여 거리를 계산한다. 가장 신호를 적게 보낸 비콘이라면 신호를 가장 많이 보낸 비콘을 이용하여 거리를 계산한다. 이후, 그 거리만큼 떨어진 셀들을 후보군으로 둔다. 자기장 세기를 측정하여 후보군의 자기장 셀과 비교하여 가장 유사한 세기 값을 가지는 셀로 사용자의 위치를 추측한다. 자기장 매칭에서는 작성한 자기장 맵의 (0, 0) 셀과 측위 단계에서 (0, 0) 셀의 자기장 세기를 비교하여 차이 값만큼 보정하여 사용한다.

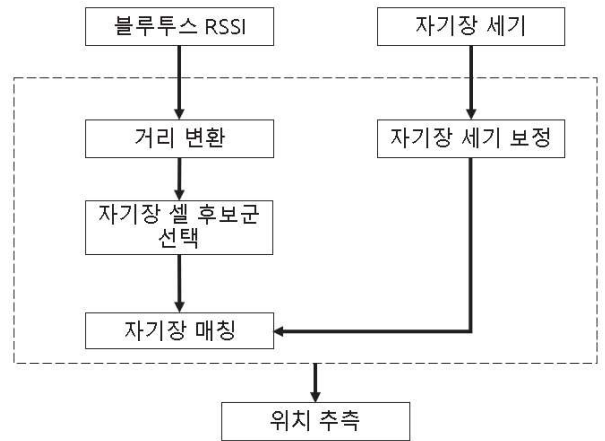


그림 3. 제안 알고리즘 개요

### 4. 실험

실험은 총 4층 높이 건물의 지하 1층에서 진행하였다. 실험 공간은 가로 8m, 세로 11m이고 1m 단위로 총 88개의 셀로 분할하였다. 각 셀에서 같은 방향을 향하게 한 뒤 스마트폰을 지면과 수평이 되게 파지하고 자기장 세기를 수집하였다. 실내 공간에는 총 3개의 비콘을 부착하고 각 셀에서 0.5초 동안 각 비콘의 세기와 자기장 세기를 수집하였다. 측위 단계에서는 자기장 맵을 생성할 당시와 같은 위치, 방향, 파지 자세로 0.5초 동안 비콘 신호와 자기장 세기를 수집하였다. 총 88개의 셀 모두에서 비콘 신호와 자기장 세기를 수집하여 측위 결과를 냈고, 실제 위치와 유클리드 거리를 이용해 오차 거리를 계산했다. 자기장 맵 작성에는 n981n 기기를 사용했고, 측위에는 g930s 2개, x600s, n981n을 사용했다. 그림 3은 3m, 5m, 7m 이내에 비콘이 있을 때의 오차를 기기 별로 상자 그림으로 나타낸 것이고, 표 1은 표로 나타낸 것이다. 4개의 기기 모두에서 오차 거리가 가장 짧을 때는 0m였지만 비콘이 3m 이내에 있음에도 10m 이상의 오차가 발생하는 경우도 있었다. 이러한

오차를 발생시키는 원인으로서는 다수 개의 비콘 중에서 가장 강한 신호를 보냈다고 인식한 비콘이 가장 가까이 있지 않은 경우가 있었다.

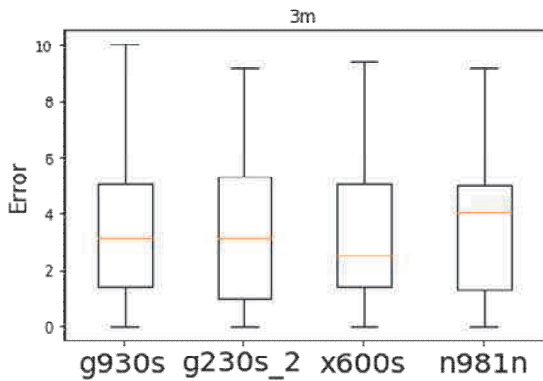


그림 2 (b). 보정 후 셀 별 자기장 세기

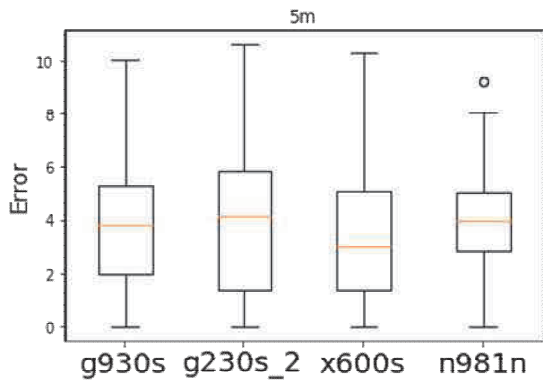


그림 3 (b). 5m 거리의 상자 그림

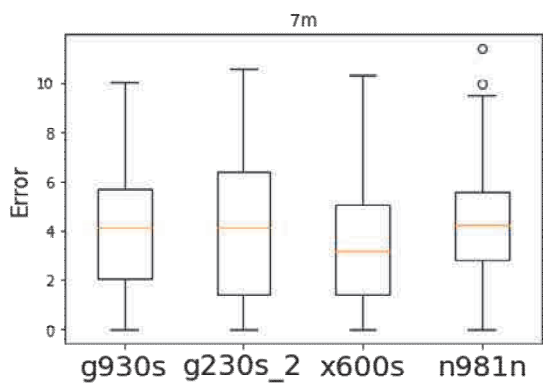


그림 3 (c). 7m 거리의 상자 그림

표 1 (a). 3m 거리에서 최대, 최소, 평균 오차

2. 기기	3. 최대	4. 최소	5. 5% 6. 절사 평균	7. 10% 8. 절사 평균
9. n981n	10.9.21m	11.0m	12.3.4m	13.3.34m
14.g930s	15.10.04m	16.0m	17.3.52m	18.3.47m
19.g930s_2	20.9.2m	21.0m	22.3.53m	23.3.45m
24.x600s	25.9.43m	26.0m	27.3.12m	28.2.97m

표 1 (b). 5m 거리에서 최대, 최소, 평균 오차

29.기기	30.최대	31.최소	32.5% 33.절사 평균	34.10% 35.절사 평균
36.n981n	37.7.28m	38.0m	39.3.84m	40.3.79m
41.g930s	42.10.04m	43.0m	44.3.75m	45.3.7m
46.g930s_2	47.10.63m	48.0m	49.3.89m	50.3.83m
51.x600s	52.10.29m	53.0m	54.3.38m	55.3.31m

S

표 1(c). 7m 거리에서 최대, 최소, 평균 오차

56.기기	57.최대	58.최소	59.5% 60.절사 평균	61.10% 62.절사 평균
63.n981n	64.11.4m	65.0m	66.4.31m	67.4.26m
68.g930s	69.10.04m	70.0m	71.4.01m	72.3.96m
73.g930s_2	74.10.63m	75.0m	76.4.22m	77.4.18m
78.x600s	79.10.29m	80.0m	81.3.57m	82.3.49m

## 5. 결론

본 논문에서는 스마트폰의 블루투스 센서와 자기장 센서를 혼합한 실내측위 방법을 제안하였다. 블루투스 RSSI를 이용한 방법은 다수개의 비콘을 사용하므로 일부 기기의 오동작이나 오차 발생 시 측위 결과에 오차를 발생시킬 수 있고, 자기장을 사용하는 방법은 다수 개의 셀에서 유사한 값을 가지게 되므로 측위 단계에서 자기장 세기 참조 영역을 최소화할 필요가 있다. 제안하는 방법은 하나의 비콘만 사용하므로 다수 개의 비콘을 사용하는 것에 비해 기기 의존도를 낮출 수 있으나 현재 위치로부터 가장 가까이 있는 비콘을 식별하고 비콘으로부터의 거리를 정확하게 계산되어야 측위 정확도를 높일 수 있을 것으로 예상된다.

## 참고 문헌

- [1] Jianyong, Z., Haiyong, L., Zili, C., & Zhaohui, L. “RSSI based Bluetooth low energy indoor positioning,” in International Conference on Indoor Positioning and Indoor Navigation, p. 526-533, 2014.
- [2] 명현, 이승목, 오택준, 김형진, 정종대, “영상 및 지

자계를 이용한 실내 측위 기술 동향,” in Information and Communications Magazine, vol.8, p.24-31, 2015.

[3] Bramhe, M. V., Gan, J., Ghodpage, N., Nawale, A., & Bahe, G. “Indoor Positioning System using Magnetic Positioning and BLE beacons,” Int. Res. J. Eng. Technol, 4(3), 1031-1036, 2017.

# 프로그래밍 교육 피드백을 위한 누적 학습 데이터 기반

## 학습자 수준 예측

오승민<sup>o</sup> 김덕엽 이우진

경북대학교 IT대학 컴퓨터학부

tmdals179@naver.com, ejrduq77@naver.com, woojin@knu.ac.kr

## Learner Level Prediction based on Cumulative Learning Data for Programming Education Feedback

Seung Min Oh<sup>o</sup> Deok Yeop Kim Woo Jin Lee

School of Computer Science and Engineering, Kyungpook National University

### 요 약

프로그래밍 교육에서 학습자가 이론적 지식을 습득하는 것을 넘어 적용하고 문제를 해결하기 위해서는 일반적인 이론 교육보다 더 많은 학습 과정이 필요하다. 따라서 목표로 하는 학습 과정을 달성할 수 있도록 효과적인 학습을 하려면 학습자별 수준에 맞는 피드백이 필요하다. 하지만 강사가 학습자의 수준을 정확하게 파악하기는 어렵다. 학습자가 학습 과정 중 어느 수준에 있는지 확인할 수 있는 정보나 근거가 부족하기 때문이다. 본 논문에서는 프로그래밍 강사에게 학습자 수준 판단에 도움이 될 수 있는 정보들을 제공한다. 이를 위해 학습자들의 수준이 반영되는 누적 학습 데이터를 활용해 학습자들을 그룹화한다. 그리고 그룹별 특징을 분석해 그룹들을 수준별로 정렬한다. 그룹화 데이터는 학습자들의 그룹 예측 모델의 학습에 사용되며 모델의 실습별, 그룹별 예측 성능을 통해 활용 가능성을 확인한다. 강사는 예측한 학습자의 그룹과 그룹 특징 정보를 통해 학습 진행 중에 학습자의 수준을 판단할 수 있다. 또 판단한 학습자의 수준을 바탕으로 학습자에게 학습 과정에 맞는 적절한 피드백을 제공해 효과적인 학습을 할 수 있다.

### 1. 서 론

프로그래밍 교육에서 학습자가 이론적 지식을 습득하는 것을 넘어 적용하고 문제를 해결하기 위해서는 일반적인 이론 교육보다 더 많은 학습 과정이 필요하다. 그래서 프로그래밍 교육은 이론적인 이해 부분과 적용 및 활용을 위한 실습 부분으로 나뉜다. 따라서 각 부분에서 목표로 하는 학습 과정을 달성하려면 학습자별 수준에 맞는 피드백을 통한 효과적인 학습이 이루어져야 한다[1]. 하지만 강사가 학습자의 수준을 정확하게 파악하기는 어렵다. 학습자가 학습 과정 중 어느 수준에 있는지 확인할 수 있는 정보나 근거가 부족하기 때문이다. 또 학습자의 이론 수준과 실습 수준이 일치한다는 보장이 없다[2-3]. 시험 성적을 활용할 수 있지만 단기 정보라서 학습자 수준이 제대로 반영되지 않고, 시험 후에만 사용할 수 있다는 한계가 있다.

본 논문에서는 프로그래밍 강사가 효과적인 피드백을 할 수 있도록 학습자 수준 판단에 도움이 될 수 있는 정보들을 제공

하고자 한다. 이를 위해 먼저 프로그래밍 강의에서 수집한 기존 학습자들의 학습 활동 데이터 중에서 학습자 수준이 반영되는 핵심 누적 학습 데이터들을 선정한다. 그리고 그룹별 특징이 두드러지는 최적 그룹 수를 탐색하고, 최적 그룹 수로 그룹화 한 그룹들의 특징을 분석해 수준별로 정렬한다. 마지막으로 그룹 예측 모델을 학습시켜 학습자가 속하는 그룹을 실습별, 그룹별로 예측한다. 강사는 학습자가 속하는 그룹과 그룹 특징 정보를 통해 강의 시기와 상관없이 학습자의 수준을 판단할 수 있다. 이렇게 판단한 수준을 바탕으로 학습자에게 학습 과정에 맞는 적절한 피드백이 제공되어 효과적인 학습이 이루어질 수 있다.

### 2. 관련 연구

#### 2.1 프로그래밍 교육에서 피드백 그룹에 따른 학습 효과

프로그래밍 교육에서 학습자에게 제공되는 피드백 차이가 학습 효과에 어떤 영향을 주는지 확인해본 연구에서는 피드백 그룹을 크게 세 그룹으로 나눈다[4]. 세 그룹은 학습자들에게 TA의 피드백이 제공되는 TA 그룹, 학습자끼리 피드백하는 Non-TA 그룹, 그리고 피드백이 없는 No 그룹으로 구성된다. 그리고 35회의 프로그래밍 교육 진행 후 세 그룹에 대한 학습 효과를 소요 시간, 제출 수, 완료한 실습 수 등의 데이터로 비

\* 이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. NRF-2018R1A6A1A03025109).



교한다. 비교 결과에서 TA 그룹의 학습자들이 다른 그룹보다 훨씬 더 많은 코드를 제출하며 평균적으로 테스트 및 디버그의 반복 횟수가 많게 나타난다. 그다음으로 No-TA 그룹 학습자들이 분포하며 No 그룹이 가장 저조한 것으로 나타난다. 즉 피드백은 학습자가 효과적인 학습을 통해 최대한의 학습 과정을 달성하는 데 도움이 된다는 것을 의미한다. 또 TA 그룹이 No-TA 그룹보다 학습 효과가 훨씬 좋게 나타났기 때문에 학습자가 스스로 해결하도록 하는 피드백보다 강사가 수준별 피드백을 제공할 수 있도록 하는 것이 학습 효과 향상에 더 효과적인 것으로 판단된다.

## 2.2 프로그래밍 학습자 수준 판단을 위한 기존 연구

학습자 수준을 Recognize, Design&Model, Implement, Test&Debug, Refactor, Present의 6개 학습 단계로 나눈 연구에서는 학습자를 특정 학습 단계에 매핑하기 위해 다양한 데이터들을 활용한다[5]. 데이터에는 강의자료 읽음 여부, 컴파일 과정에서 입력 데이터, 코드 및 출력 등이 있다. 그러나 연구에서는 각 학습 단계와 매핑되는 데이터들이 왜 그 학습 단계에 적합한지에 대한 설명이 없다. 또 프로그래밍 활동에 대한 데이터들을 중점으로 수집해 이론적인 이해 수준은 제대로 반영되지 않을 수 있다. 따라서 이론적인 이해 수준을 반영할 수 있는 추가 데이터를 활용하면 더 정확한 학습자 수준 판단에 도움이 될 수 있다.

## 3. 누적 학습 데이터 기반 그룹화와 그룹별 특징 분석

학습자의 수준을 예측하기 전에 먼저 학습자들을 수준이 반영되는 여러 그룹으로 나눈다. 이를 위해 활용하는 데이터의 강의와 학습자에 대한 정보는 표 1과 같다.

표 1. 대상 강의 및 학습자 관련 정보

구분	내용
강의	프로그래밍 기초
언어	C 언어
강의 기간	한 학기, 18회
강의 구성	퀴즈, 이론 강의, 실습 (총 4시간)
학습자	컴퓨터 전공, 대학 1학년 (20살)
학습자 수	110명 (3개 강의)

대상 강의는 복습을 위한 퀴즈와 강사의 이론 강의, 실습으로 구성되며 실습은 온라인 채점 시스템을 통해 온라인에서 자동으로 채점된다. 따라서 온라인 채점 시스템에는 학습자들의 학습 활동에 대한 다양한 데이터들이 수집된다. 이러한 학습 활동 데이터 중 학습자의 수준 반영 지표로 퀴즈 점수, 실습 완료 후 남은 시간, 완료한 실습 수를 선택했다. 그리고 학습자의 실습 경향을 판단하기 위해 제출 수 및 4개 오류에 대한 각 발생 수까지 총 8개의 변수를 선정했으며 표 2와 같다. 데이터들은 모두 강의를 진행됨에 따라 학습자별 특징이 명확하게 구분될 수 있도록 누적 데이터 형태로 사용한다.

데이터는 110명 학습자, 18개 실습에 대한 것이기 때문에 총 1,980개이며 이것을 모두 학습자 그룹화에 사용하지는 않는다. 그룹화는 110명 학습자를 대상으로 하는 것이기 때문에 기준이 되는 실습 데이터가 필요하다. 따라서 학습자별 특징이 누적되어 가장 잘 구분될 수 있는 마지막 실습을 기준으로 그룹화한다.

표 2. 학습자 수준을 반영하는 학습 활동 데이터

구분	설명	형태
퀴즈 점수	복습 목적의 퀴즈 점수	누적
남은 시간	실습 완료 후 남은 시간	
완료 실습 수	완료한 실습 수	
제출 수	제출 코드의 수	
컴파일 오류	컴파일 오류 수	
시간초과 오류	시간초과 오류 수	
런타임 오류	런타임 오류 수	
실행 오류	실행 오류 수	

학습자들의 그룹화에는 K-means 클러스터링을 활용하며 먼저 가장 잘 구분되는 최적의 그룹 수를 파악하기 위해 Elbow 기법을 적용한다. 군집의 응집도를 나타내는 Inertia 값의 변동 폭이 많이 감소하는 시점인 Elbow를 이상적인 클러스터 개수로 판단한다. 실습 18의 110명 데이터를 대상으로 한 적용 결과는 그림 1에 나타난다.

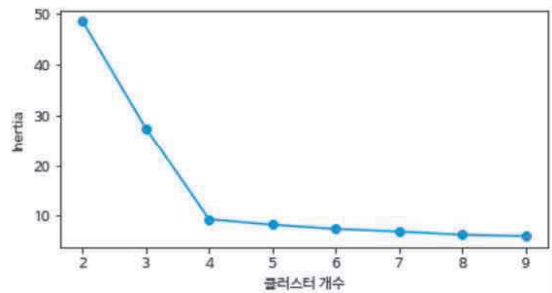


그림 1. 최적 그룹 수 탐색 결과

클러스터 개수를 최대 9까지 설정해서 탐색해 본 결과 학습자 그룹화를 위한 최적 그룹 수는 4로 확인된다. 이제 그룹별 특징과 수준을 확인하기 위해 그룹화 결과를 산점도를 활용해 분석한다. 각 학습 활동 변수에 대한 관계를 2차원 좌표상에 시각화하여 분석하고 그룹별 분석 결과를 통해 그룹별 특징과 수준에 따라 순서대로 정렬한다. 그림 2는 산점도를 활용한 분석 예시를 나타내며 표 3은 분석 결과에 따른 그룹별 특징과 수준별 정렬 결과를 나타낸 것이다.

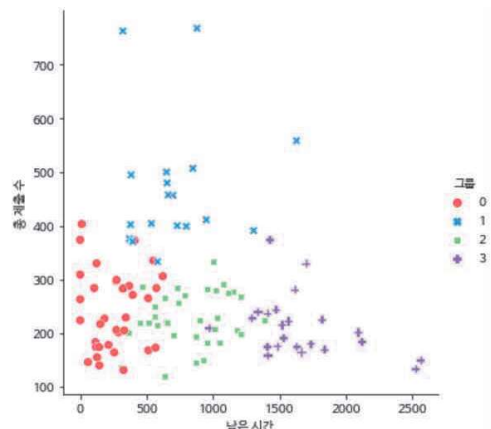


그림 2. 그룹별 특징 분석

표 3. 그룹별 특징과 수준별 정렬 결과

특징	그룹 0	그룹 1	그룹 2	그룹 3
학습자 수	32	17	35	26
수준	하	중(하)	중(상)	상
이론적 이해	하	중	중	상
솔루션 생성	하	중	중	상
코드 구현	하	하	중	상
제출 수	중	상	중	하
오류 발생	중	상	중	하

전체적인 수준에 따라 그룹들을 오름차순으로 정렬했기 때문에 그룹 0이 하위 그룹, 그룹 3이 상위 그룹이다. 상위 그룹은 점수가 높고 많은 실습을 완료했지만, 코드 제출 및 발생하는 오류는 적게 나타난다. 하위 그룹은 상위 그룹과 완전히 반대로 나타나며 그룹 1과 그룹 2는 비슷한 중위권에 분포한다. 그러나 그룹 1의 경우 다른 그룹들에 비해 훨씬 많은 코드 제출이 이루어지며, 그에 따라 많은 오류도 발생하는 특징을 보인다. 즉 그룹 2와 같은 중위권이지만 상대적으로 프로그래밍 능력이 떨어진다고 볼 수 있다. 따라서 그룹들의 순서는 그룹 1을 중하위권, 그룹 2를 중상위권으로 판단한 결과이다.

#### 4. 학습자 그룹 예측

학습자 그룹화 결과를 활용해 학습자 그룹을 예측해보기 위해 먼저 그룹화 기준이 된 실습18의 데이터를 제외한 나머지 17개 누적 실습 데이터에 학습자별 그룹 정보를 포함한다. 그리고 학습 데이터와 테스트 데이터를 9:1 비율로 랜덤 샘플링한다. 그리고 샘플링 시드를 변경해가며 반복적으로 테스트한다. 학습자 그룹 예측에 활용한 모델은 K-NN(K-Nearest Neighbor)이며 적용 결과 이웃 수에 따른 모델의 평균 정확도는 그림 3과 같다.

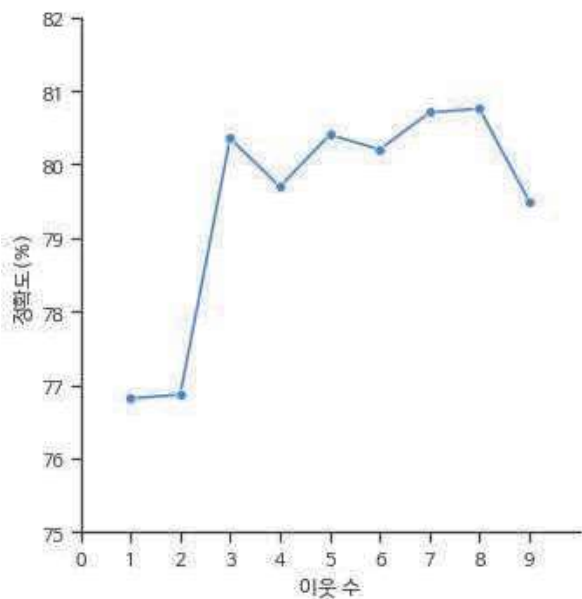


그림 3. 이웃 수에 따른 그룹 예측 평균 정확도

테스트 결과 모델의 그룹 예측 평균 정확도는 이웃 수가 8일 때 약 81%로 가장 높게 나타난다. 따라서 이웃 수를 8로 고정하고 시드를 1에서 10까지 변경해가며 실습별 정확도를 확인해 본 결과는 그림 4와 같다.

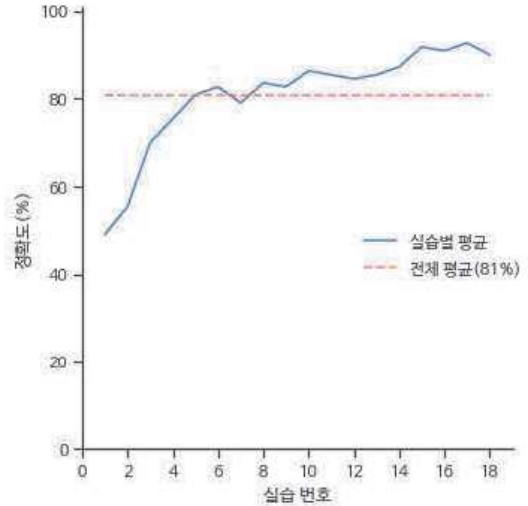


그림 4. 실습별 그룹 예측 정확도

실습별 학습자 그룹 예측 테스트 결과 전체 실습 평균은 약 81%의 정확도를 보인다. 그리고 초기 구간 실습들에서는 약 50%~70%의 낮은 예측 성능을 보이지만 실습이 진행됨에 따라 정확도가 높아져 최종적으로 약 85% 정도의 예측 성능을 달성한다. 초기 구간의 낮은 예측 성능은 데이터가 충분히 누적되지 않았기 때문에 발생할 수 있다. 따라서 강의 초기에는 학습자 그룹 예측 모델의 활용성이 떨어지지만, 대부분의 실습에서는 81%의 예측 성능으로 활용 가능성이 충분하다.

다음은 실습 진행에 따른 그룹별 예측 성능을 확인한다. 학습자 수가 110명으로 적기 때문에 랜덤 샘플링을 적용할 경우 중복 가능성이 커 모든 학습자의 개별 예측 결과에 대한 정확도 평균을 확인한다. 그룹별 예측 결과는 그림 5에 나타난다.

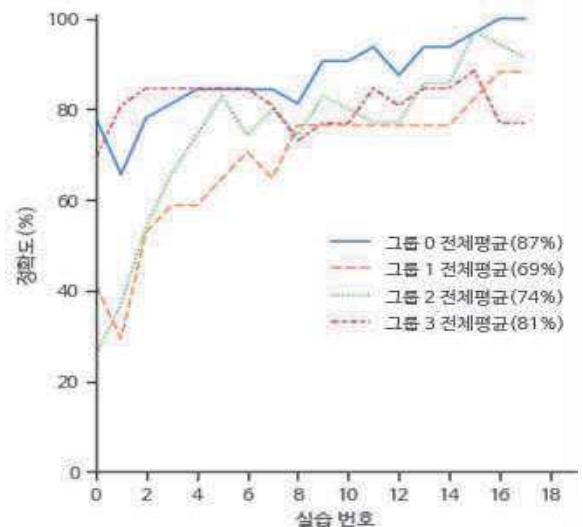


그림 5. 그룹별 그룹 예측 정확도

그룹별 예측 결과 그룹 3을 제외한 나머지 그룹들은 실습이 진행됨에 따라 예측 성능이 좋아진다. 그리고 상위권 및 하위권에 속하는 그룹 3과 그룹 0의 예측 정확도 평균은 약 84%로 높게 나타나지만, 중위권에 속하는 그룹 1과 그룹 2의 평균 정확도는 약 72%로 낮게 나타난다. 중위권 그룹들은 실습 중간 구간까지 70%보다 낮은 예측 정확도를 보인다. 즉 상대적으로 중위권 학습자들의 정확한 수준 파악이 어렵다는 것을 확인할 수 있다. 하지만 실습 초기 구간을 제외하면 80%에 가까운 예측 정확도를 보여주기 때문에 수준 판단에 활용할 수 있다.

## 5. 결론

프로그래밍 교육의 학습 과정은 이론 이해 부분과 실습 부분으로 나뉘기 때문에 학습 목표 달성을 위해 더 효과적인 학습이 필요하다. 그러나 강사가 학습자의 수준 판단을 정확하게 하지 못해 효과적인 학습을 위한 수준별 피드백이 이루어지기 어렵다. 따라서 강사가 학습자의 수준을 정확하게 판단할 수 있도록 누적 학습 데이터로 학습자들을 그룹화한 결과 학습자들이 특징 있고 수준이 반영되는 4개의 그룹으로 구분된다. 또 강의 초기 구간의 그룹 예측 정확도는 떨어지지만, 평균 81%의 정확도로 그룹을 예측할 수 있다. 그룹별로도 상위권 및 하위권 그룹의 경우 약 84% 정확도, 중위권 그룹은 약 72% 정확도로 중위권 학습자의 예측 정확도는 떨어지지만, 강의 초기 일부 구간을 제외할 경우 학습자의 수준 판단에 충분히 활용 가능한 성능이라고 판단된다. 따라서 강사는 예측한 학습자 그룹 정보와 그룹별 특징 정보를 바탕으로 학습자 수준을 판단하고 수준에 맞는 피드백을 제공할 수 있다. 수준별 피드백은 학습자의 학습 효과를 향상해 학습 목표를 달성할 수 있다.

## 참 고 문 헌

- [1] Min-young Kam, "Relationship between Feedback Types and Task persistence: Moderating Effects of Feedback Acceptance and Academic Self-Efficacy", Educational Psychology and Evaluation of Kyungpook National University Graduate School of Education, 2015.
- [2] Lister, R. and Leaney, J., "Introductory programming, criterion-referencing and Bloom", Proceedings of the 34th SIGCSE technical symposium on Computer science education, Reno, Nevada, USA, ACM Press, 2003.
- [3] Winslow, L.E. "Programming Pedagogy - a Psychological Overview", SIGCSE Bull, No. 28, pp. 17~22, 1996.
- [4] Suin Kim, Jungkook Park, Jae Won Kim and Alice Oh, "Elice: An online CS Education Platform to Understand How Students Learn Programming", Learning at Scale, pp. 225~228, 2016.
- [5] Suin Kim, Jungkook Park, Jae Won Kim and Alice Oh, Elivate: A Real-Time Assistant for Students and Lecturers as Part of an Online CS Education Platform, 2016.

# 한글 손글씨 문장 인식에 대한

## OCR 오픈 소스 성능 비교 분석

신찬규<sup>○\*</sup>, 김세연<sup>\*</sup>, Barathi Subramanian, 김정홍  
경북대학교 컴퓨터학부

scv1702@knu.ac.kr, kimseyun314@naver.com, achu\_samriti@yahoo.com, jhk@knu.ac.kr

## Performance Analysis of OCR Open Source

### for Korean Handwritten Sentence

Shin Changyu<sup>○\*</sup>, Kim Seyeon<sup>\*</sup>, Barathi Subramanian, Kim Jeonghong  
Kyungpook National University Department of Computer Science

#### 요 약

본 연구는 대표적인 오픈 소스 기반의 OCR 모델인 PP-OCRv3와 EasyOCR를 한글 손글씨 문장 인식에 적용해보고, 1-NED 평가 방법을 통해 각 모델의 성능을 평가했다. AI Hub의 한국어 글자체 이미지 데이터를 적용한 결과, PP-OCRv3는 1-NED 평가 방법에서 0.721점, EasyOCR은 0.671점을 달성했다. 본 연구는 기존의 글자 단위의 한글 손글씨 인식에서 문장 단위로 확장했다는 점에서 의의가 있다.

#### 1. 서 론

오늘날 인간이 생성하는 데이터 양은 매우 방대하기 때문에, 데이터를 효율적으로 이해하고 분석하기 위해선 컴퓨팅 기술이 필수 불가결하다. 컴퓨터 비전을 활용한 OCR(Optical Character Recognition) 기술이 등장함에 따라, 인간의 수작업 없이 자동으로 문서, 영수증, 명함 등을 인식해 이를 전자적 형태로 변환하는 것이 가능해졌다. 이를 통해 문서 및 도서 자료 번역, 기록물 데이터베이스화, 전문 의료 정보 해석 및 공공 서비스 지원 등 다양한 분야에서 OCR 기술이 활용되고 있다.[1]



그림 1. OCR의 동작 과정

OCR의 동작 과정(그림 1)은 이미지에서 문자가 속한 영역을 검출하는 문자 검출(Text Detection) 과정과 해당 영역의 문자가 실제로 어떤 문자인지 인식하는 문자 인식(Text Recognition) 과정으로 나눌 수 있다. 또한 OCR의 성능을 높이기 위해 전처리 및 후처리 단계를 추가하기도 한다.

본 연구는 전자 문서 애플리케이션을 개발하기 위한 목적으로, 대표적인 OCR 오픈 소스인 EasyOCR과 PP-OCRv3[2]를 활용해 한글 손글씨 문장에서 단어를 검출 및 인식하여 이를 전자적 형태로 변환하고 두 오픈 소스 간 성능에 대한 데이터를 제시한다.

기존의 한글 손글씨 인식에 대한 OCR 오픈 소스에는

IBM에서 개발한 tensorflow-hangul-recognition[3]이 있는데, 이는 단일 CNN(Convolutional Neural Network)을 사용하기 때문에 글자 단위의 인식만 가능하다. 이는 대부분 문장으로 이루어진

전자 문서 애플리케이션과 적합하지 않기 때문에, 문자 검출과 문자 인식이 모두 가능한 EasyOCR과 PP-OCRv3를 활용했다. 또한 최신 OCR 오픈소스에 대해 한글 손글씨 문장 인식을 접목한 연구가 전무하기 때문에, 향후 유사한 애플리케이션 개발 시 EasyOCR과 PP-OCRv3를 적용하고자 할 때 본 연구의 실험 결과를 참고할 수 있을 것이다.

#### 2. 관련 연구

##### 2.1 EasyOCR

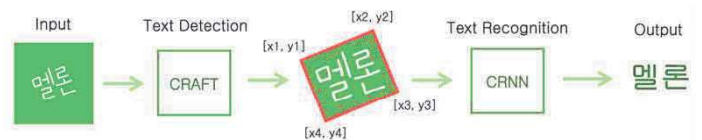


그림 2. EasyOCR 구조도

EasyOCR(그림 2)의 문자 검출 모델인 CRAFT[4]는 Batch Normalization을 사용한 VGG16 기반의 FCN(Fully Convolutional Network) 모델이며, Decoding 부분에 Skip Connection을 이용하여 낮은 레벨의 특성들을 통합한다. CRAFT는 어떤 픽셀이 글자의 중심일 확률을 의미하는 Region Score와 어떤 픽셀이 인접한 두 글자의 중심일 확률을 의미하는 Affinity Score를 이용해 인접한 개별

글자들을 상향식 방식으로 연결하여 단어 단위로 검출하기 때문에, 곡선이나 기형적 형태의 문자들 또한 잘 검출한다. 이로 인해 CRAFT는 다양한 형태의 문자들을 인식해야 하는 Scene Text Detection 문제에서 주목받고 있다.

최근 문자 인식 모델들은 개별 글자가 아닌 주로 단어 단위로 인식하는데, CNN을 이용해 문자 이미지의 특성을 추출하거나 LSTM(Long Short-Term Memory) 또는 GRU(Gated Recurrent Unit)가 있는 Recurrent 모델을 사용한다.[1] EasyOCR의 문자 인식 모델인 CRNN[5]은 CNN과 RNN(Recurrent Neural Network)을 이용해 단어 단위 인식을 수행한다. CRNN은 입력 이미지로부터 Feature Sequence를 추출하는 Feature Sequence Extraction 단계와 추출한 Sequence의 각 Frame에 대한 Label Distribution을 예측하는 Sequence Labeling 단계, 이들을 최종 문자로 변환하는 Transcription 단계로 이루어진다.

## 2.2 PP-OCRv3

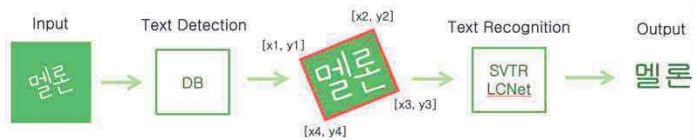


그림 3. PP-OCRv3 구조도

PP-OCRv3(그림 3)는 DB(Differentiable Binarization)[6] 기반의 문자 검출 모델을 사용한다. PP-OCRv3는 모바일 환경에서 동작하는 것을 가정하기 때문에, 모델의 경량화를 목표로 하며 이를 위해 다양한 방법을 사용했다. 딥러닝 모델은 Backbone의 크기가 대부분을 차지하는데, PP-OCRv3는 모바일 기기를 대상으로 하는 MobileNetV3[7]을 Backbone으로 사용했으며 MobileNetV3에서 많은 수를 차지하는 SE(Squeeze and Excitation) 모듈을 제거해 모델을 경량화했다. 또한 객체 탐지에 많이 사용되는 FPN(Feature Pyramid Network)와 유사한 Head를 사용했으며, 1x1 Convolution 필터를 사용해 Inner Channel을 줄임으로써 모델을 경량화했다.

PP-OCRv3는 CRNN과 같이 Feature Extraction을 위한 Visual 모델과 Text Transcription을 위한 Sequence 모델로 이루어진 많은 Scene Text Recognition 모델과는 다르게, 하나의 Visual 모델만을 사용한 SVTR[8]와 경량 CNN 기반의 PP-LCNet[9]을 결합한 SVTR-LCNet을 문자 인식 모델로 사용한다. PP-OCRv3는 SVTR 중 작은 크기의 SVTR-Tiny를 선택했는데, SVTR-Tiny는 구조적 한계로 인해 Intel의 수학 커널 라이브러리인 MKLDNN을 사용하는 CPU에서 PP-OCRv2의 문자 인식 모델인 CRNN보다 10배나 느리다는 문제점이 있다. 이를 해결하기 위해 SVTR-Tiny 속도의 대부분을 차지하는 Mix 모듈을 제거하고 PP-LCNet의 일부분으로 교체했다. 또한 Global Mix 모듈의 속도가 입력 특성의 크기와 관련 있다는 점을 발견해, 이를 Pooling Layer 다음으로 옮겨 예측 속도를 향상시켰다.

## 3. 한글 손글씨 문장에 대한 성능 비교 분석

### 3.1 데이터셋 및 실험 환경

본 실험에서는 한글 손글씨 문장 인식을 위해 한국지능정보사회진흥원 AI Hub의 “한국어 글자체 이미지” 데이터셋을 각 모델에 학습시켰다. 해당 데이터셋에는 현대 한글 11,172자를 사용한 폰트 50종의 글자체와 성별, 연령층 별로 직접 제작한 이미지 파일 구축 및 간판, 상표, 교통 표지판 등의 한글이 포함된 이미지 10만 장을 구축한 이미지들이 제공된다. 본 실험에서는 손글씨 단어 데이터 일부를 학습 및 검증 데이터로, 손글씨 문장 데이터 일부를 시험 데이터로 사용했다. 실험은 Google Colab Pro+에서 진행했으며, 운영 체제는 Ubuntu 18.04.6 LTS, GPU는 NVIDIA A100-SXM4-40GB이다.

### 3.2 성능 평가 기준

자주 사용되는 문자 인식 성능 평가 기준에는 WEM(World based Exactly Matching) 평가 방법과 1-NED(1-Normal Edit Distance) 평가 방법이 있다. WEM 평가 방법은 정답 단어와 예측 단어가 정확히 일치하는지 평가하는 방법인데, 이는 문장 단위의 평가 시 정확히 예측한 단어 수를 반영하지 못하기 때문에 본 실험에 적합하지 않다. 이로 인해, 본 실험에서는 1-NED 평가 방법을 사용했다. 1-NED 평가 방법은 두 단어 간 편집 거리를 긴 단어의 길이로 정규화한 값을 평가하는 방법이다. 1-NED는 정답과 예측 단어 간 편집 거리가 작을수록 1에 가깝다. 1-NED는 전체 단어에서 글자 단위의 점수를 평가하는 방법이기 때문에, WEM보다 본 실험에 적합하다.

### 3.3 성능 평가

Model	Model Size(M)	1-NED	Speed(ms) (A100 GPU)
EasyOCR (Pretrained Only)	24.8	0.481	0.158
PP-OCRv3 (Pretrained Only)	<b>15.6</b>	0.402	0.097
EasyOCR (Fine-tuned Recognition)	24.8	0.671	0.139
PP-OCRv3 (Fine-tuned Recognition)	<b>15.6</b>	<b>0.721</b>	<b>0.096</b>

표 1. EasyOCR과 PP-OCRv3의 성능 비교표

표 1은 EasyOCR과 PP-OCRv3의 문자 인식 모델을 학습시키기 전(Pretrained Only)과 후(Fine-tuned Recognition)에 대한 한글 손글씨 문장 인식 성능을 나타낸다. 학습 후 EasyOCR은 동작 속도가 0.139 ms로 0.158 ms에서 1.13배 상승했고, PP-OCRv3는 0.097 ms에서 0.096 ms로 1.01배 상승했다. 또한 1-NED 평가 방법에 대해 EasyOCR이 0.481에서 0.671로 1.40배 상승했고, PP-OCRv3이 0.402에서 0.721로 1.80배 상승했다. 여기서 주목할 점은 PP-OCRv3가 EasyOCR에 비해 모델 크기가 1.59배 더 작음에도 불구하고 1.07배 더 좋은 성능을 보여주었으며 속도 또한 1.45배 더 빨랐다.

## 4. 논의

문자 인식의 성능 평가 기준은 단어뿐만 아니라 띄어쓰기까지 평가 범위에 포함되기 때문에, 띄어쓰기를 검출하지 못하면 낮은 점수를 받게 된다. 문장은 띄어쓰기가 적지 않은 비율을 차지하며, 손글씨의 경우 사람에 따라 띄어쓰기 간격이 매우 좁거나 일정하지 않기 때문에 인식이 잘되지 않는 경우도 발생했다. 이러한 경우 모든 문자를 잘 인식했음에도 불구하고 충분한 점수를 받지 못하기 때문에, 이에 대한 성능 평가 기준의 개선이 필요하다고 판단된다.

## 5. 결론

본 연구는 최근 OCR 분야에서 많이 사용되는 EasyOCR과 PP-OCRv3에 한글 손글씨 문장 인식을 접목시켜 보고, 이에 대한 성능 비교 분석을 진행했다. PP-OCRv3는 모델 구조 상 모바일 환경에 적합할 뿐만 아니라, EasyOCR에 비해 더 좋은 성능을 보여주기 때문에 앞으로 많은 한글 손글씨 문장 인식 OCR 분야에서 적극적으로 활용될 것으로 예상된다.

## 사 사 문 구

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2021-0-01082)

## 참 고 문 헌

- [1] 민기현 외, “딥러닝 기반 광학 문자 인식 기술 동향”, 전자통신동향분석, 37권, 5호, pp. 25-30, 2022.
- [2] Chenxia Li et al., “PP-OCRv3: More Attempts for the Improvement of Ultra Lightweight OCR Sytesm”, arXiv preprint, CoRR, arXiv:2206.03001v2, 2022.
- [3] <https://github.com/IBM/tensorflow-hangul-recognition>
- [4] Youngmin Baek et al., “Character Region Awareness for Text Detection”, IEEE Conf. Comput. Vis. Pattern Recognit., pp. 9365-9374, 2019.
- [5] Baoguang Shi, Xiang Bai, Cong Yao, “An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition”, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 39, No. 11, pp. 2298-2304, 2016.
- [6] Minghui Liao et al., “Real-Time Scene Text Detection with Differentiable Binarization”, AAAI Conf. Artif. Intell., Vol. 34, No. 7, pp. 11474-11481, 2020.
- [7] Andrew Howard et al., “Searching for MobileNetV3”, IEEE Int. Conf. Comput. Vis., pp. 1314-1324, 2019.
- [8] Yongkun Du et al., “SVTR: Scene Text Recognition with a Single Visual Model”, Int. Joint Conf. Artif. Intell., pp. 884-890, 2022.
- [9] Cheng Cui et al., “PP-LCNet: A Lightweight CPU Convolutional Neural Network”, arXiv preprint, CoRR, arXiv:2109.15099, 2021.

# 가상 환경을 통한 커스텀 학습 데이터 생성 및 기계 학습의 유효성 검증

배진호<sup>o</sup>, 구은성, 김근찬, 김영호, 최희정, 이효경(교신저자)

경북대학교 컴퓨터학부

{jin\_void1, kooes903, mouse4918, kkyh12180, chj00, hokyoungee}@knu.ac.kr

## Verification of creating custom training dataset using virtual environment and machine learning's validation

Jinho Bae<sup>o</sup>, Eunsung Koo, Geunchan Kim, Younghyo Kim, Heejung Choi, Hokyoung  
Lee(Corresponding author)

School of Computer Science and Engineering, Kyungpook National University

### 요약

가상 환경을 통해 평면적 데이터를 다양하게 수정하여 생성한 데이터 셋이 실제 기계 학습에서 유효한지를 탐구하였다. 연구 결과, 이미지를 배치하는 과정에서 무작위성만 보장된다면 실제 기계 학습에서 유의미한 결과가 나타나는 것을 확인했다. 본 연구 결과는 추후 평면적 데이터를 사용하는 기계 학습에서 가상 환경을 사용하는 것이 더 효율적일 수 있음을 나타내는 데 의의를 가진다.

### 1. 서론

실시간 영상 객체 탐지 모델을 학습시키기 위해서는 인식하고자 하는 객체에 대한 학습 데이터가 필요하다. 학습 데이터는 대량으로 준비되어야 하며, 여러 변수가 반영되어야 한다. 학습 데이터의 양과 다양성이 객체 탐지의 정확성과 연결되기 때문이다. 그러나 대량의 학습 데이터를 수집하기 위해서는 인력, 시간 등 여러 비용이 든다는 문제점이 존재한다.

본 연구에서는 효율적인 학습 데이터 수집을 위해 가상 환경을 이용한 학습 데이터 가공을 수행하고자 한다. 적은 양의 학습 데이터로 여러 변수를 반영해 수많은 학습 데이터를 생성하여 효율성을 높이는 것을 목표로 한다. 또한 생성한 학습 데이터를 이용해 모델을 학습시킨 결과를 가시화하고자 한다.

### 2. 데이터 연구

#### 2.1 기존 학습 데이터 수집 및 연구 방법

0	0.5	0.5	0.7	0.45
type	CenterX	CenterY	Width	Height

그림 1. 이미지의 학습 레이블 형태

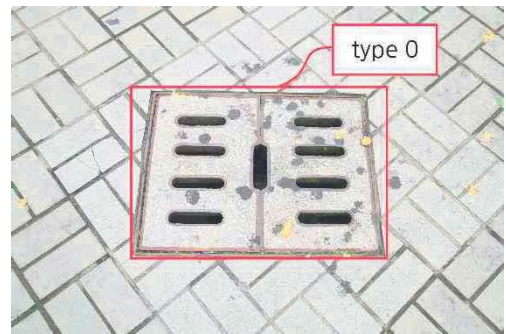


그림 2. 학습 이미지 정면에 레이블 처리된 모습

그림 1은 딥러닝 라이브러리 yolov5에서 한 개의 이미지에 대응되는 객체들의 좌표 정보가 담긴 레이블 파일 한 개의 포맷을 나타낸 것이다. 각 행은 하나의 객체 정보를 의미하고 열은 각각 객체 구분 번호, 중심 좌표 X, 중심 좌표 Y, 너비, 그리고 높이를 의미한다. 객체를 제대로 탐지하기 위해서는 각 이미지에 대응되는 레이블이 대량으로 필요하고, 객체가 존재하는 좌표 정보를 수작업으로 기입하여야 하기에 학습 이미지 및 데이터의 구성에 시간적, 인력적, 비용적인 부담이 발생한다. 예를 들어, 그림 2와 같이 배수구 모양을 학습하고자 하는 경우, 이미지에 대한 하수구 레이블을 필요로 하며, 촬영 각도, 크기, 촬영 거리, 태양광 세기, 날씨 등과 같은 변수에 따른 각각의 이미지 학습 자료를 준비해야 한다는 문제점이 있다. 따라서 본 장에서는 학습하고자 하는 객체의 이미지 및 레이블을 가상 3D 공간에서 계산 및 대량 자동 생성하고 이를 학습하는 방법을 제시한다.

## 2.2 가상 3D 기반 인공 학습 데이터 제작 및 개선

본 연구에서 감지하고자 하는 학습 객체의 경우, 평면상에 존재하는 객체(예 배수구 위치)를 감지하는 것이 주 목표이므로 각종 환경변수(빛, 반사색상, 각도, 거리 등)와 함께 가상 3D 이미지로 자동 구성하여 학습 자료로 활용할 수 있다.

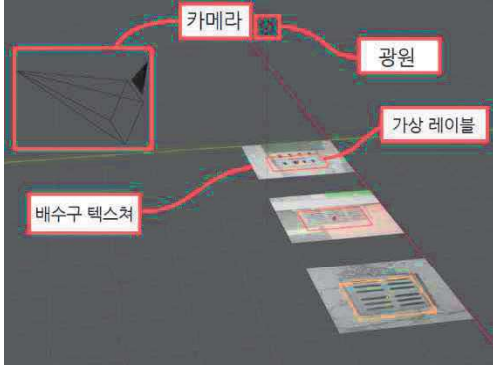


그림 3. 가상 공간에 배치된 학습 이미지 및 레이블

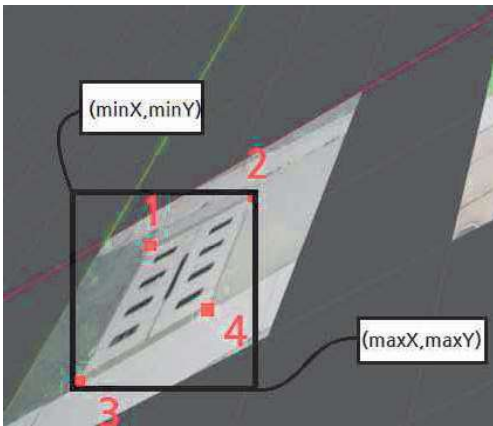


그림 4. 3D 레이블을 2D 좌표계로 투영시킨 모습

그림 3은 공간에 학습하고자 하는 이미지 자료의 배치를 나타낸 것이다. Blender 3D 모델링 소프트웨어를 사용해 가상의 환경에서 탐지할 객체인 배수구를 나타내기 위해, 배수구의 이미지를 평면 상에 텍스처로 적용하고 그 밑에 배수구 위치만 나타내고 있는 가상의 레이블 객체를 구성하였다. 각 3D 좌표상에 존재하는 객체의 각 정점들은 월드 좌표(World coordination)기준 3차원(X,Y,Z) 좌표를 가지고 있기에, 이를 실제 렌더링 되는 화면상의 2D 좌표에 투영(Projection)시키는 방식으로 2D 렌더링 된 정점 좌표를 계산할 수 있다. 그림 4과 같이 가상 공간에서의 3D 좌표를 2D 좌표계로 변환 된 위치 좌표는 각 (X,Y) 정점 4개를 가지는데, yolov5 라이브러리에서 요구하는 정규화 된 직사각형 중심 좌표 및 너비, 높이로 변환 하는 과정을 거쳐야 한다. 직사각형의 좌표를 구하기 위하여 정점들의 (x최솟값, y최솟값)와 (x최댓값, y최댓값)을 구하고, 중심 좌표를 계산한 뒤, 이를 이미지의 크기에 맞게 정규화 하여 2D 이미지 상의 바운딩 박스를 추출하여 학습 레이블을 자동 생성할 수 있다.



그림 5. 애니메이션 기능을 활용하여 자동 추출한 상황 별 3D 배수구 이미지 및 레이블링 형태

배수구를 바라보는 각도, 크기, 방향을 난수로 지정하여 자동화된 이미지 및 객체 레이블 학습 자료를 가상 배수구 장면을 상황에 맞춰 대량 추출할 수 있다. 그림 5는 실제 가상 공간에 생성된 배수구 텍스처를 카메라 각도, 좌표를 변경해 가며 추출한 레이블 파일 및 이미지 파일 집합이다. 하지만 상기 방법으로 렌더링 된 이미지들은 고정적으로 생성된 텍스처 위치에 바라보는 카메라 벡터만 바꾼 것으로, 미리 생성된 배수구 텍스처의 고정된 순서에 의존하여 학습될 수 있다는 문제점이 있다. 따라서, 배수구에 대한 이미지를 정면으로 대량 촬영하고 그에 대한 배수구 레이블을 구성한 뒤, 렌더링 장면을 생성할 때 미리 만들어진 정면 이미지와 객체 레이블을 랜덤으로 3D 좌표에 배치함으로써, 각종 변수와 함께 여러 각도에서 바라볼 때의 학습 이미지를 생성할 수 있다.

## 2.3 학습 장면 및 환경 변수 개선 기법



그림 6. 일반 학습 이미지 사진과 Normal Map 데이터



그림 7. Normal Map와 함께 렌더링 된 학습 이미지

특히, 이미지 학습에서 용이하도록 사실적이고 환경 동적인 이미지를 다양하게 추출하기 위해 Normal Map 기법을 사용할 수 있다. 그림 6의 Normal Map 데이터는 배수구 사진에 대한 법선 반사 정보를 담은 이미지 자료이다. 각 색상 R,G,B 데이터에 빛이 반사해야 할 법선 벡터 정보가 담겨 있으며, 해당 맵을 사용하여 텍스처를 렌더링하면 가상 3D 환경에서 빛의 위치에 따라 질감이 사실적이고 동적인 이미지를 생성할 수 있다. 이렇게 Normal Map이 적용되면 그림 7처럼 사실적인 배



수구 묘사가 가능해지며, 인공지능 이미지 학습에 최대한 현실과 동일한 질감 조건을 줄 수 있다.

해당 연구에서는 아침, 점심, 저녁 등 외부 빛에 대한 변수를 생성하고, 하수구에 해당 태양광(SunLight) 및 백열등(PointLight) 광원을 적용함으로써 그림 7과 같이 동적이고 사실적인 배수구 화면을 학습에 유리하도록 추출하였다. 결과적으로 광원의 위치, 밝기, 반사 광과 같이 다양한 환경 변수를 매번 다르게 지정함으로써 이미지 학습에 유리하도록 대량 생성하였다.

### 3. 시뮬레이션 및 평가

#### 3.1 시뮬레이션 결과

##### (1) 이미지를 배치한 뒤 다양한 시간대의 필터를 적용한 경우



그림 8. 학습에 사용된 이미지 예시

실제 이미지를 가상 3D환경에 배치한 뒤, 가상 환경의 이점을 활용하고, 다양한 환경에서의 정확도를 올리기 위해 다양한 필터를 사용하여 데이터를 처리하였다.

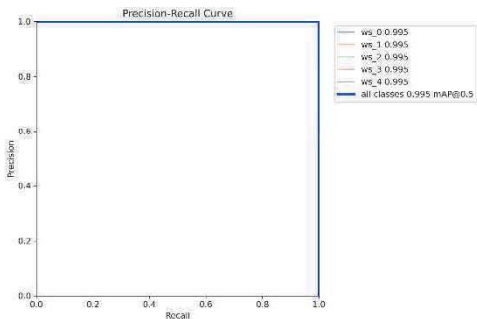


그림 9. yolov5m.pt로 60회 학습 결과

yolov5m.pt로 60회 학습한 결과, 그림5와 같이 PR 곡선이 직각의 형태로 나타났다. 실제로 위의 모델을 사용하여 다른 이미지를 예측했을 때 이를 인식하지 못하는 모습을 보였다.

이러한 결과에 대해 두가지 원인이 예측된다. 먼저 모델의 복잡성으로 인해 이러한 결과가 나타났다고 볼 수 있다. 데이터의 개수에 비해 복잡성이 높은 모델을 사용하여 과적합(overfitting)이 발생한 것으로 예상된다. 더해서 규칙성이 있는 데이터를 사용하다 보니, 학습 과정에서 해당 규칙성을 학습한 것으로 예측된다.

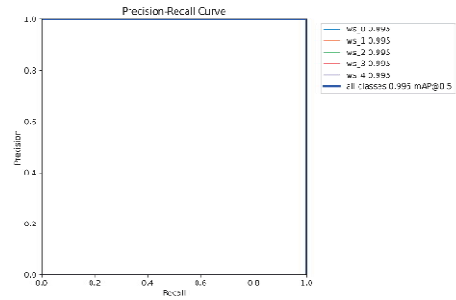


그림 10. Yolov5s.pt로 60회 학습 결과

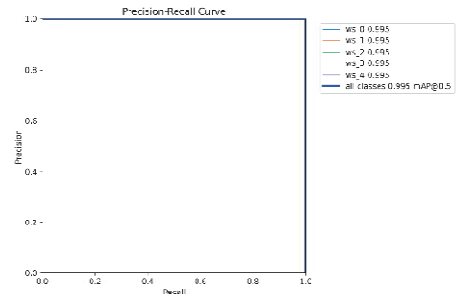


그림 11. yolov5n.pt로 60회 학습 결과

두 결과 전부 60회 학습에 대한 결과이며, 속도의 차이만 존재할 뿐 앞서 보였던 직각의 형태를 보이고 있다. 이는 현재 이루어지는 학습이 데이터 내의 패턴을 학습 중이라는 것을 보여주며, 이미지 내 무작위성이 추가되어야 함을 나타내고 있다.

##### (2) 단일 레이블링에 대해 이미지 배치에 무작위성을 추가한 경우

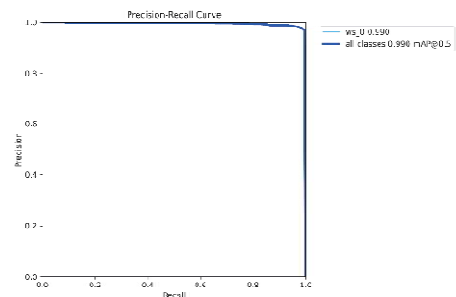


그림 12. yolov5s.pt로 25회 학습 결과(PR Curve)

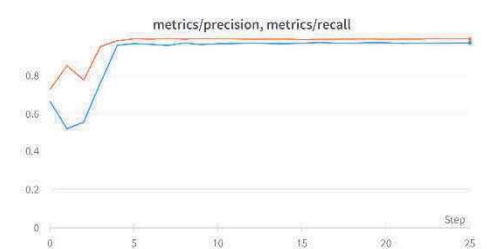


그림 13. yolov5s.pt로 25회 학습 결과 (Epoch에 따른 Precision과 Recall)

PR 곡선의 형태가 앞의 학습 결과와는 다르게 극단적인 직각의 형태가 아닌 것을 볼 수 있다. 실제로 학습 과정에서 정

밀도(precision)와 재현율(recall) 수치가 지속적으로 진동하는 것을 확인할 수 있는데, 이는 이전 데이터로 학습했을 때와는 다른 양상을 보인다.

그럼에도 불구하고 과적합(overfitting)과 유사한 높은 정확도를 보이는 이유는 데이터셋의 크기가 작기 때문에 발생한 것으로 보인다.

### (3) 이미지 배치의 무작위성은 유지하고, 복수의 레이블링을 사용한 경우

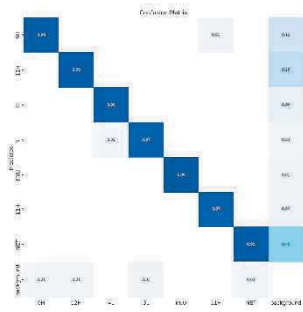


그림 14. yolov5s.pt로 230회 학습 결과 (Confusion matrix)

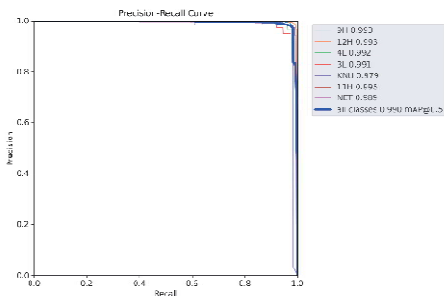


그림 15. yolov5s.pt로 230회 학습 결과 (PR Curve)



그림 16. 위의 모델을 통해 예측한 이미지

PR 곡선의 형태가 (2)의 결과와 유사한 모습을 볼 수 있다. 또한 위의 결과로 나온 모델을 통해 실제 이미지를 테스트한 결과도 정확도가 높은 것을 볼 수 있다. 이는 복수 레이블링에 대해서도 이미지의 무작위성만 보장된다면 가상 3D 환경을 사용해도 무방하다고 볼 수 있다.

## 4. 결론

### 4.1 연구 결과

본 연구는 학습 데이터 확보를 위한 시간적, 인력적, 그리고 비용적 부담 문제 해결을 위해 가상 환경을 통한 기계 학습의

유효성 검증을 진행하였다. Blender 소프트웨어를 통해 가상 환경에서 여러 변수 조절을 통해 데이터 셋을 생성했고, 이를 yolov5 오픈소스를 통해 학습을 진행했다.

총 4번의 실험 중 앞의 2번의 실험을 통해 무작위성이 없는 데이터에 대한 학습을 통해 이는 실제 환경에 적용되지 않는다는 결론을 얻었으며, 이후 나머지 실험을 통해 무작위성이 있는 데이터에 대한 학습은 실제 환경에서 사용할 수 있다는 결과를 도출하였다.

그러나 이러한 결과에도 여전히 과적합(overfitting)에 대한 검증이 필요하다. 세번째 실험의 경우 약 1600장의 데이터를 사용하여 학습을 진행하였는데, 적은 시행 횟수에도 금세 재현율(precision) 수치가 증가하고, 학습이 진행됨에 따라 mAP(mean Average Precision) 수치만 개선됨을 확인할 수 있었다. 다만, 네번째 실험의 경우에는 세번째 실험과 레이블링의 개수에만 차이를 뒀을뿐, 진행 자체는 유사했다. 네번째 실험에서는 과적합으로 보이는 징후가 크게 없었기에, 데이터의 다양성 차이에 의한 결과로 결론지었다.

### 4.2 연구의 의미

본 연구의 의의는 평면적인 물체를 학습시킬 때, 적은 데이터로 최대한의 효율을 증가시킬 수 있다는 것이다.

사물이나 대상을 인식시키기 위해서는 아주 많은 양의 학습 데이터가 요구된다. 실제로 yolov5 공식 문서에 따르면, 최적의 결과를 위해서는 각 클래스 별로 1,500장 이상의 이미지를 권장하고 있다. 그러나 네번째 실험을 진행할 때 실질적으로 사용한 단일 이미지는 약 679장에 불과하며, 가상 환경을 통해 679장의 이미지들을 다양한 각도, 거리, 환경에서 무작위로 배치하여 재촬영한 것을 모아 데이터셋으로 활용했다. 결과적으로 총 2,000여장의 렌더 이미지를 사용하였으며, 2,000장의 이미지 내에 25장의 세부 이미지가 들어갔다는 점을 고려하면 수치적으로 약 75배의 효율성을 보인다.

이는 단순히 이미지의 개수 뿐만 아니라, 레이블링을 위한 시간적인 면에서의 효율성도 의미한다. 실험을 진행하는 과정에서 레이블링도 스크립트를 통해 자동으로 진행했으며, GTX 460 기준 장당 2초의 시간이 걸렸다. 총 50,000개의 레이블링에 대해 4,000초의 시간 정도가 소요되었고, 이는 사용 GPU에 따라 더 절약될 수 있는 부분이다.

즉, 연구를 통해 데이터가 평면적인 경우 가상 환경을 사용하면 적은 양의 데이터로 시간적, 인력적, 비용적인 부분에서 큰 효율적 이득을 볼 수 있음을 알 수 있었다.

## 사 사 문 구

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음”(2021-0-01082)

## 참 고 문 헌

1. Tips for Best Training Results. (n.d.). Retrieved November 19, 2022, from YOLOv5 Documentation: <https://docs.ultralytics.com/tutorials/training-tips-best-results/>

# 입원 환자 도움 프로그램\*

김재우<sup>01</sup> 전찬용<sup>1</sup> 안동혁<sup>1</sup> 최도진<sup>1</sup>

<sup>1</sup>창원대학교 컴퓨터공학과

lgapt5@naver.com, w7e43958@naver.com, donghyeokan@changwon.ac.kr, dojinchoi@changwon.ac.kr

## An Inpatient Assistance Program

Jaewoo Kim<sup>01</sup> Chanyong Jeon<sup>1</sup> Donghyeok Ahn<sup>1</sup> Dojin Choi<sup>1</sup>

<sup>1</sup>Department of Computer Engineering, Changwon National University

### 요약

대형 병원은 간호사 한 명당 담당하는 환자의 수가 많아, 입원 환자가 제대로 된 간호를 받기 어려운 것이 현실이다. 2021년 서울특별시 공공 보건 의료 재단의 기사에 따르면, 대학 병원의 경우 간호사 1명당 환자 12 ~ 20명, 요양 병원의 경우에는 환자 40명에 이른다. 본 논문에서는 입원 환자를 위한 프로그램을 제안한다. 제안하는 입원 환자 도움 프로그램은 5개의 다양한 기능 제공을 통해 입원 환자들의 간호의 질을 높이고자 한다.

### 1. 서론

대형 병원은 간호사 한 명당 담당하는 환자의 수가 많아, 입원 환자가 제대로 된 간호를 받기 어려운 것이 현실이다. 2021년 서울특별시 공공 보건 의료 재단의 기사에 따르면, 대학 병원의 경우 간호사 1명당 환자 12 ~ 20명, 요양 병원의 경우에는 환자 40명에 이른다[1]. 따라서 기존에 이를 효율적으로 지원하기 위한 다양한 간호용 애플리케이션들이 제안되었다[2-4]. 기존에 선행되었던 연구는 ‘거리 감지 기술 IoT를 이용한 낙상 예방기기’ 등 프로그램이 단일 기능을 수행하는 경우가 많았다[2]. 본 논문에서는 입원 환자를 위한 간호의 질을 높이기 위하여 해당 프로그램을 기획하게 되었다.

본 논문에서 새롭게 다양한 기능을 동시에 할 수 있도록 제작하여 구매 비용 대비 효율을 증가시키고자 하였다. 해당 프로그램은 총 5가지의 기능을 포함하고 있다. 기능을 구현하는데 라즈베리파이4와 덤러닝, OpenCV 구동을 위해 랩탑을 사용하였다.

### 2. 제안하는 입원 환자 도움 프로그램

#### 2.1 전체 프로그램 기능도

제안하는 입원 환자 도움 프로그램의 전체 기능도는 다음 그림 1과 같다. 재활 환자를 위한 이동 거리 측정, 주사 위치 조

절 알림, 도지플로우 조절기, 심부전 환자 사망률 예측기, 혈액 역류 탐지기이다. 라즈베리파이를 통해 3개의 기능을 구현 및 검증하였으며, 2가지 기능은 랩탑을 활용하여 구현하였다.

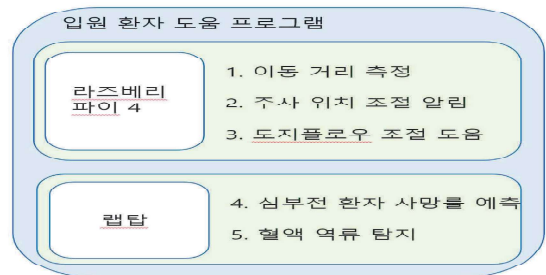


그림 1 전체 프로그램 기능도

#### 2.2 이동 거리 측정

첫 번째 기능은 재활 환자를 위한 이동 거리 측정 기능이다. 해당 기능은 라즈베리파이의 가속도 센서를 이용하였다. 병원 측에서는 만보기를 따로 구입하지 않아도 되고, 만보기에 측정되지 않는 경우도 방지할 수 있다. 가속도 센서의 x축 가속도 값과 y축 가속도 값을 0.1초 마다 받아 엑셀에 저장한다. 측정된 가속도 값을 한 번 적분하여 속도를 계산한다. 그 속도를 다시 한 번 더 적분을 하면 거리가 나오게 되어 환자가 얼마나 이동을 하였는지 알 수 있다.

그림 2는 첫 번째 기능인 재활 환자를 위한 이동 거리 측정 기능 결과이다. 다음과 같이 0.1초마다 가속도 값이 x축과 y축으로 구분 되어 저장 된다. 제안하는 프로그램은 같은 시간에 들어온 가속도 값을 이용하여 속도로 변환한다. 최종적으로 같은 시간의 거리 값을 이용하여 입원 환자의 누적 움직임 거

\* 이 논문은 2022학년도 자기주도 연구역량 지원사업 연구비에 의하여 연구되었음

리를 측정할 수 있다. 이를 재활 프로그램에 활용하여 환자의 재활 상태 및 운동량을 측정할 수 있다.

가속도		속도		거리	
x	y	x	y	x	y
1.676239	1.657082	0.083812	0.082854	0.004191	0.004143
1.005744	0.996165	0.217911	0.215516	0.019277	0.019061
-0.08621	1.331413	0.263888	0.331895	0.043367	0.046432
-0.3544	0.325669	0.241857	0.414749	0.068654	0.083764
-0.01916	-0.65134	0.223179	0.398466	0.091906	0.124425
1.101529	-1.52298	0.277298	0.28975	0.11693	0.158836
2.710718	-1.0249	0.46791	0.162356	0.15419	0.181441
2.672404	-0.17241	0.737066	0.10249	0.214439	0.194683
1.973173	1.015322	0.969345	0.144636	0.299759	0.20704
1.781603	0.795016	1.157084	0.235152	0.406081	0.226029
0.852487	1.034479	1.288789	0.326627	0.528375	0.254118
-0.23946	1.034479	1.31944	0.430075	0.658786	0.291953
-0.34483	0.201149	1.290225	0.491857	0.789269	0.33805
0.028736	-0.2682	1.274421	0.488504	0.917502	0.387068
1.321834	-0.86207	1.341949	0.431991	1.04832	0.433092

그림 2 재활 환자를 위한 이동 거리 측정 결과

### 2.3 주사 위치 조절 알림

두 번째 기능은 주사 위치 조절 알림 기능이다. 수액은 수액 비닐과 환자와 연결된 주사 바늘간의 높이 차이를 이용하여 수액이 몸 속으로 이동하는 원리다[5]. 오랫동안 풀대 상단을 잡고 있는 경우, 침상 위에 장시간 올라가 있는 경우, 보호자가 어린 아이를 들어서 안고 있는 경우 등 입원 환자가 장시간 수액을 맞으면 손의 위치는 가변적일 수 있다. 만약 수액 비닐과 주사 간의 높이 차가 일정하게 유지되지 않는다면 수압에 문제가 생기고, 이로 인하여 수액의 정상적인 투여가 이루어지지 않을 수 있다. 더 나아가 극단적인 경우에는 환자의 혈액이 호스를 따라 역류할 수도 있고, 이를 사이펀 현상[6]이라고 부른다. 본 기능에서는 사이펀 현상을 방지하고, 일정량의 수액이 환자에게 지속적으로 투여될 수 있도록 하고자 하였다. 그림 3은 주사 위치 조절 알림 기능 모형을 나타낸다. 점적통은 높이 10cm로 기존의 점적통보다 큰 것으로 마련하였으며, 점적통의 천장에서 바닥을 수직으로 바라보도록 하여 점적통 속 수위를 인식할 수 있게 초음파 센서를 부착하였다.

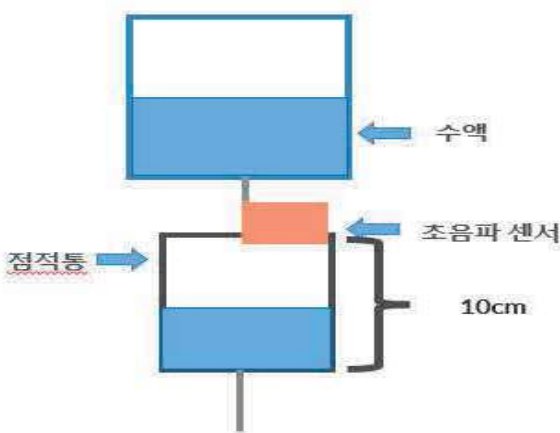


그림 3 주사 위치 조절 알림 기능 모형도

그림 4는 주사 위치 조절 알림 기능의 흐름도를 나타낸다. 구현은 라즈베리파이4를 기반으로 하였다. 주사 위치 조절 알림 기능은 총 3개의 하위 기능을 포함하고 있다. 첫 번째 하위 기능은 모든 환자에게 적용될 수 있는 기본 기능이다(2-1). 두 번째 하위 기능은 수액이 일정하게 투여되는 것이 정말 중요한 당뇨병, 고혈압, 심장병이 있는 주의 환자들을 위한 기능이다(2-2). 세 번째 하위 기능은 주의 환자가 아닌 일반 환자에 해당 될 경우에 작동하는 기능이다(2-3).

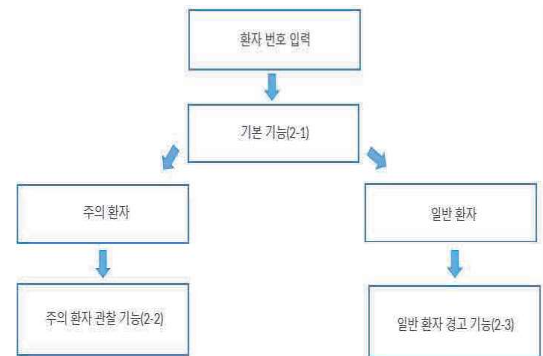


그림 4 주사 위치 조절 알림 기능 흐름도

그림 5는 기본 기능 결과이다. 기본 기능의 기준을 충족하면 알림이 작동된다. 기본 기능은 초음파 센서와 수면과의 거리가 2.5cm 이하가 되면 환자에게 높이 조절이 필요하다는 것을 알리기 위하여 빨간 led가 켜지고 부저가 울린다. 그리고 초음파 센서와 부표와의 거리가 7.6cm 초과, 즉 수면이 경통의 바닥에 아주 가까워지면 수액이 거의 다 소진되었다고 판단하여 초록 led가 켜지고 부저가 울림으로써 수액 교체의 필요성을 알린다. 수액이 거의 다 소진된 채로 장시간 방치되는 경우도 수압이 일정하게 유지되지 않아 혈액의 역류가 일어날 수 있어 해당 기능을 넣었다.

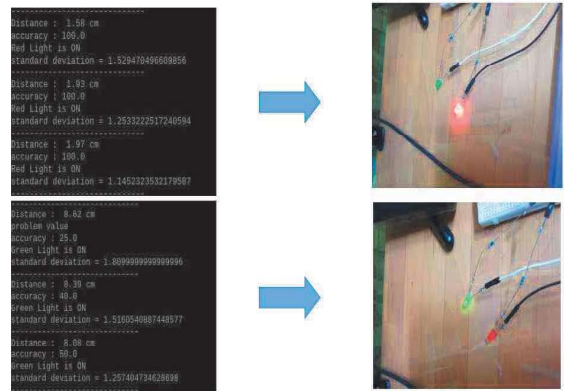


그림 5 기본 기능

그림 6은 주의 환자 관찰 기능 결과를 나타낸다. 환자가 주의 환자에 해당되면 그래프를 그리고, 기준 구간을 벗어나면

알림 및 문자가 자동으로 발송되는 것을 볼 수 있다. 주의 환자 관찰 기능은 입력 받은 환자 번호가 주의 환자에 해당 되면 그래프가 그려지면서 해당 기능이 작동하게 된다. 점적통의 중간 범위인 3.5cm 이상 ~ 6.5cm 이하 구간을 벗어나게 되면 파란 led가 켜지고 부저가 울리며, 담당 간호사의 휴대폰에 자동으로 문자가 발송된다. 그래프는 시각적으로 어떻게 진행되고 있는지 보여줄 뿐만 아니라, 간호사가 문자를 받고 왔을 때 점적통 속 수위의 변화가 어떻게 진행되었는지 확인할 수 있다.

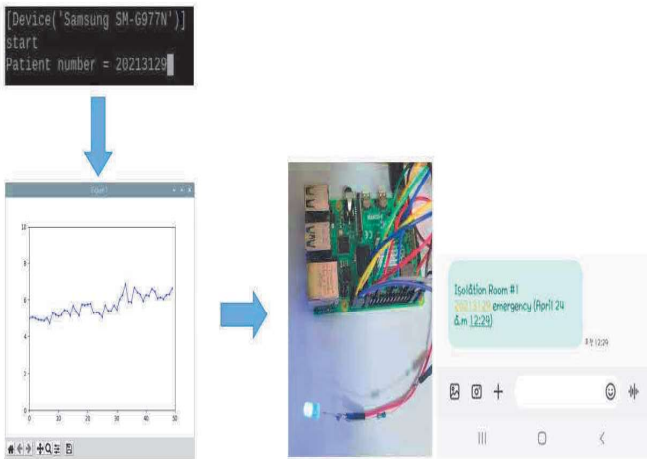


그림 6 주의 환자 관찰 기능

그림 7은 일반 환자 경고 기능 결과이다. 분포 기준과 표준 편차 기준 중 하나라도 만족하면 파란 led가 켜지고 부저가 울린다. 일반 환자 경고 기능은 기본 기능에서 빨간 led와 초록 led가 켜지는 구간을 제외한 구간을 동일하게 세 구간으로 나누었다. 첫 번째 하위 기준은 세 구간 중 가장 상단에 위치한 구간의 분포가 전체 세 구간의 절반 이상이면 경고한다. 그리고 측정된 값이 30개 이상일 때부터 작동한다. 두 번째 하위 기준은 세 개의 구간 중 정중앙에 해당하는 5cm로부터 측정값들의 표준 편차가 얼마나 되는지를 계산한다. 표준 편차가 1이상이면 환자에게 경고한다.

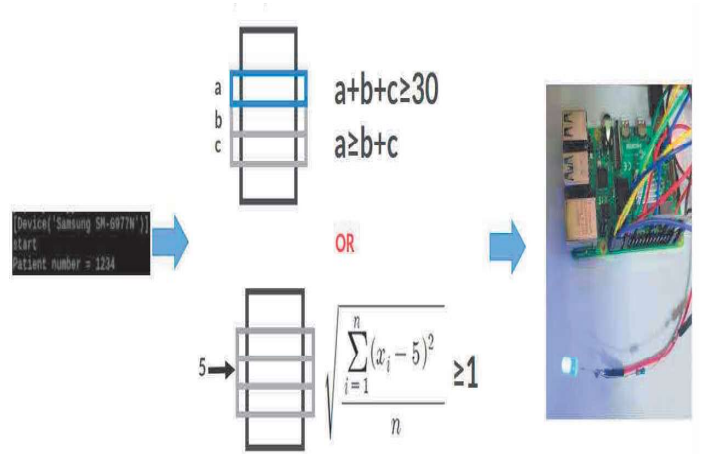


그림 7 일반 환자 경고 기능

## 2.4 도지플로우 조절 도움

세 번째 기능은 도지플로우 조절 도움 기능이다. 일반적으로 간호사의 경험에 의하여 도지플로우를 조절하게 되는데, 점적통에 한 방울이 맺히고 떨어지는데 몇 초가 걸려야 하는지 알려주고 해당 시간에 맞게 소리가 나와 보다 정확하게 조절할 수 있게끔 도움을 준다. 수액 한 방울은 약 0.05ml이다. 라즈베리파이에 부착되어 있는 스위치를 누르면 창이 뜨고, 수액을 맞는 시간과 용량을 입력할 수 있다. 다음 그림 8은 세 번째 기능인 도지플로우 조절 도움 기능을 나타낸다. 10시간 동안 3L를 맞으려면 한 방울이 맺히고 떨어지는데 0.6초가 걸려야 한다는 것을 보여주고, 해당 시간에 맞는 소리가 나온다.



그림 8 도지플로우 조절 도움 기능

## 2.5 심부전 환자 사망률 예측

네 번째 기능은 심부전 환자 사망률 예측 기능이다. 심부전 환자의 검사 결과를 통하여 해당 환자의 사망률이 얼마나 되는지 계산하여 예측해 볼 수 있다. 실제 사망한 심부전 환자와 사망하지 않았던 심부전 환자들의 데이터를 학습시킨 후 해당 환자의 사망 확률을 예상한다. 확률을 통하여 진료 및 간호의 강도를 결정하는데 도움을 줄 수 있다. 혈청 나트륨, 혈청 크레아틴, 혈소판, 혈압 등 총 10가지의 데이터를 학습하였다. 기본 DNN(Deep Neural Network) 모델을 사용하였으며, 활성화 함수로 ReLU(Rectified Linear Unit)와 sigmoid를 이용

하였다. 학습시킨 후 테스트 파일을 이용하여 검증 한 결과, 최대 95.67% 사망률 예측을 나타내었다.

## 2.6 혈액 역류 탐지

다섯 번째 기능은 혈액 역류 탐지 기능이다. 환자가 침상에 누워 있는 경우에도 혈액이 역류할 수 있는데, 이를 탐지하기 위한 기능이다. 해당 기능은 카메라를 이용한다. 구현에는 OpenCV를 이용하였으며, 혈액과 유사한 빨간색이 검출되었을 시에 ‘warning’ 이라는 경고음이 나온다. 다른 빨간 물건은 카메라 범위 내에 나오지 않는 것을 전제로 한다. 그림 9는 다섯 번째 기능인 혈액 역류 탐지 기능 결과이다. 왼쪽 사진과 같이 팔만 보였을 때는 화면에 아무런 반응이 없고 경고음이 나오지 않는다. 하지만 오른쪽 사진과 같이 혈액과 유사한 색깔이 나타나면 화면에 초록색 선이 그려지며 인식하고, 경고음이 나온다.



그림 9 혈액 역류 탐지 기능 결과

## 3. 결 론

본 논문에서는 입원 환자 도움 프로그램을 제안하였다. 제안한 프로그램을 통하여 입원 환자에게 더 나은 간호 환경을 제공할 수 있을 것이라고 기대한다. 프로그램 기능 중 주사 위치 조절 알림으로 해당 프로그램을 구현하였다. 자동적으로 모터를 이용하여 수액 비닐의 위치를 조절하는 방식이면 환자가 주사 위치 조절에 신경 쓰지 않아도 수압이 유지될 수 있다. 향후에는 모형의 소형화와 입원 환자를 위한 간호의 질을 높이기 위하여 추후에 더 유용한 어플리케이션을 기능을 추가할 예정이다.

## 참 고 문 헌

- [1] <http://www.seoulhealth.kr/board/notice/read?menuId=17&searchBbsCd=5&searchSeq=3800>
- [2] 김동근, 강라경, 김예령, 신동희, “거리 감지 기술 및 IoT를 이용한 낙상 예방 기기,” ATRAC 3D, 2021.
- [3] 모문희, “항암화학요법을 받는 폐암 환자의 피로 자가 간호 스마트폰 애플리케이션 개발 및 적용 효과,” Asian Oncology Nursing, Vol. 21, No.4, pp. 183-193, 2021.
- [4] 박가은, 이해정, 강아름, “제2형 당뇨병 환자를 위한 자동 맞춤형 셀프케어 프로그램 개발,” 대한간호과학회지, Vol. 52, No.5, pp. 535-549, 2022.
- [5] <https://stdict.korean.go.kr/search/searchView.do?pageSize=10&searchKeyword=%EB%A7%81%EA%B1%B0>
- [6] <https://stdict.korean.go.kr/search/searchView.do?pageSize=10&searchKeyword=%EC%82%AC%EC%9D%B4%ED%8E%80%EC%9E%91%EC%9A%A9>

# 쿠버네티스를 활용한 MLOps 플랫폼 ‘MeLonKube’ 개발

박보경, 박종민, 신시온, 이에림, 이동욱, 고석주

경북대학교

kshmith3@gmail.com, qkrqhr@naver.com, pjmin0101@gmail.com, sion@knu.ac.kr, yaeeee1004@knu.ac.kr,

dwlee@datastreams.co.kr, sjkoh@knu.ac.kr

## Development of MLOps platform ‘MeLonKube’ using Kubernetes

SeokHee Kim, Bokyung Park, JongMin Park, Sion Shin, YeRim Lee,

83.DongWook Lee, SeokJoo Koh

School of Computer Science and Engineering, Kyungpook National University

### 요약

실제 서비스 단계에서 머신러닝과 딥러닝의 코드가 차지하는 부분은 크지 않다. 따라서 모델의 성능만이 아닌 개발과 운영을 유기적으로 통합하기 위해 MLOps가 등장하였다. MeLonKube는 하이퍼파라미터 튜닝을 수행하는 MLOps 플랫폼이다. 본 논문에서는 MNIST 분류모델의 학습결과를 제시하여, 쿠버네티스를 사용하는 이유를 보인다. 쿠버네티스는 리소스 관리와 컨테이너 오케스트레이션 등의 장점으로 시간과 비용을 절감할 수 있다. MeLonKube는 이러한 쿠버네티스 기반의 플랫폼이지만 사용자는 쿠버네티스에 대한 전문적인 지식 없이 사용할 수 있다. 이를 통해 효율적인 MLOps의 기능을 수행하는 MeLonKube 플랫폼의 중요성을 보인다.

## 1. 서론

최근 인공지능을 활용한 데이터 분석에 대한 수요가 높아지고 있다. 기업과 연구에서는 대용량 컴퓨팅 능력을 탑재한 서버를 사용해 대용량의 데이터 분석을 수행한다. 대용량의 데이터에 대한 모델의 하이퍼파라미터를 수동으로 조절하는 하이퍼파라미터 튜닝 과정은 시간과 비용이 많이 든다. 이런 하이퍼 파라미터 튜닝 작업을 자동화한다면, 최적의 하이퍼파라미터 값을 찾기 위해 사람이 수많은 모델들을 직접 학습시켜야 하는 수고를 덜 수 있다. 이처럼 쿠버네티스를 사용하여 ML을 수행하면 비용을 절감할 수 있고, 클러스터를 편리하게 관리하는 이점을 얻을 수 있다.

MeLonKube는 효율적인 ML 개발을 위한 하이퍼파라미터 튜닝 서비스를 제공해 줌으로써 머신러닝 개발자들의 금전적, 시간적 비용을 절감해 주고 개발을 용이하게 해준다. 본 논문에서는 컨테이너와 컨테이너 오케스트레이션 기술을 활용한 하이퍼파라미터 튜닝 플랫폼의 개발과정과 MLOps의 필요성 및 본 플랫폼의 중요성을 보인다.

## 2. Kubernetes

컨테이너는 어떤 환경에서든지 실행될 수 있는 소프트웨어 실행 유닛으로써 애플리케이션 코드가 해당

라이브러리 및 종속 항목과 함께 담겨 있는 하나의 독립된 프로세스 공간이다. [1] 컨테이너 기술의 도입 이후, 많은 기업들이 서비스의 구성 요소로 컨테이너를 사용하면서 컨테이너의 수가 급속도로 많아졌다. 수많은 컨테이너의 관리 및 유지하는 관리자에게 상당한 부담이 된다. 컨테이너 관리자가 개별 컨테이너를 조작하는 등의 작업을 줄이고 관리성과 편의를 제공하기 위하여 컨테이너 오케스트레이션 기술이 등장했다.

컨테이너 오케스트레이션은 컨테이너의 배포, 관리, 확장, 네트워킹을 자동화하는 것을 말한다. 컨테이너 오케스트레이션은 대표적으로 프로비저닝 및 배포, 리소스 할당, 컨테이너 상태 모니터링, 리소스 할당, 로드 밸런싱 등의 기능을 제공한다. [2]

그중, 쿠버네티스는 구글에서 만든 오픈소스 플랫폼으로 점유율이 높아 사용 가능한 라이브러리 또는 레퍼런스가 많다. CNCF 애플리케이션 대부분이 쿠버네티스와 높은 효율을 보이며, 주요 클라우드 벤더와 On-premise 솔루션 업체와 같은 업체에서 쿠버네티스를 지원하고 있다. 따라서, 쿠버네티스만 설치되어 있으면 서비스를 쉽게 올릴 수 있다. [3]

## 3. MeLonKube

### 3.1 쿠버네티스를 활용한 하이퍼파라미터 튜닝 플랫폼 MeLonKube

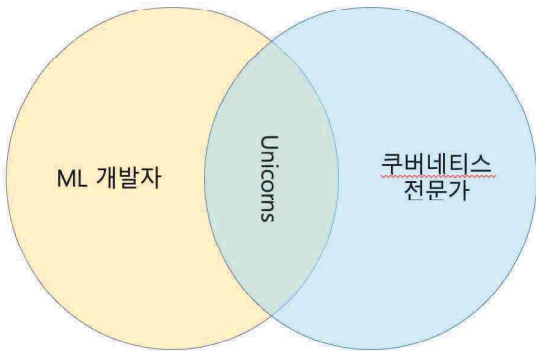


그림 1. Unicorns

ML 개발자는 쿠버네티스를 사용하여 개발을 하는 과정을 거쳐야 한다. 하지만, ML개발자는 쿠버네티스에 대한 모든 지식을 모두 이해하여 사용하는 것은 매우 어렵고, 두 분야를 모두 잘 아는 사람은 unicorn이라 불릴 만큼 매우 흔치 않은 경우이다. ML 개발자가 ML개발에만 집중할 수 있게 해주는 툴의 필요성이 대두된다.

본 논문에서는 ML 개발자가 쿠버네티스에 대한 이해 없이도 개발 과정을 쉽게 해주는 하이퍼파라미터 튜닝 플랫폼인 MeLonKube를 제안한다. MeLonKube는 추상화를 통해 ML 코드와 하이퍼파라미터만 가지고 ML 개발을 할 수 있게 해주는 서비스이다.[4]

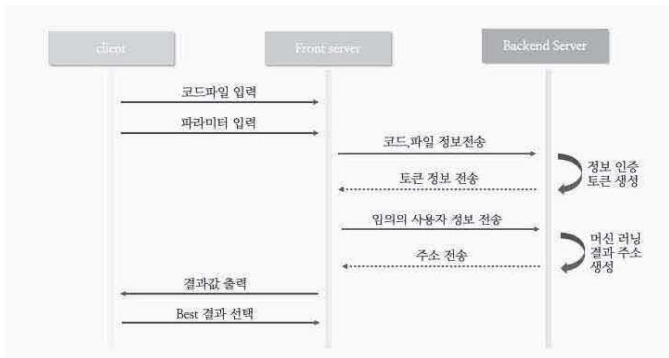


그림 2. MeLonKube의 Sequence Diagram

그림 2는 MeLonKube의 Sequence Diagram을 보여준다. 사용자가 ML 코드, 데이터, 하이퍼파라미터를 Frontend-server에 주면, 그 정보를 Backend-Server에 전송한다.

Backend-Server는 정보 인증 토큰 정보를 생성해 Frontend-server에 전송한다. 실행하기 전 사용자가 입력한 데이터가 올바른지 검사하는 과정을 거치고 코드를 실행하면 빠르게 사용자와 소통할 수 있다.

토큰 정보를 받은 Frontend-server는 사용자 정보를 다시 Backend-Server로 전송한다.

위의 과정에서 받은 정보들을 바탕으로 MeLonKube의 서버에서 쿠버네티스 기술을 활용해 여러 컨테이너를 생성하여 병렬적으로 ML을 수행한다.

수행한 ML의 결과 주소를 생성해 Frontend-server에 전송하고, Frontend는 plot형태로 결과값을 출력한다.

사용자는 그 값 중 최선의 결과를 선택할 수 있다.

### 3.2 MeLonKube의 차별점

하이퍼파라미터 튜닝을 지원하는 서비스로써 MeLonKube가 가지고 있는 장점은 첫 번째로, 사용자에게 쿠버네티스 관련 지식을 요구하지 않는다는 점이다. 쿠버네티스로 하이퍼파라미터 튜닝을 돕는 다른 서비스로는 Kubeflow가 있다. Kubeflow는 쿠버네티스를 통하여 클러스터를 관리하며 배포하지만, 사용자가 직접 쿠버네티스를 조작해야 한다는 단점이 존재한다. Kubeflow에서 제공하는 Katib, PyTorchJob 등의 컴포넌트는 모두 사용자가 쿠버네티스의 기능을 정확하게 파악하고 이해하고 있다는 가정하에 존재한다. [4] 본 서비스는 ML 개발자들의 그런 고충을 인지하여 쿠버네티스에 대한 지식 없이 원하는 하이퍼파라미터 튜닝을 수행할 수 있도록 개발되었다.

두 번째 장점은 모든 것이 웹에서 서비스된다는 점이다. 구름IDE처럼 웹상에 존재하는 하나의 사이트이기에, 별도의 프로그램 설치나 환경설정 없이 언제든지 하이퍼파라미터 튜닝을 할 수 있다. 장비나 개발환경이 충분하지 않은 1인 개발자와 소규모 스타트업 개발자들은 쿠버네티스를 비롯한 튜닝 환경이 요구하는 CPU, GPU 사양을 충족하기 어렵다. 본 서비스는 사용자 대신 연산을 수행하기 때문에 장비 사양이 충분하지 못한 사용자들에게 적절한 서비스를 제공한다.

## 4. MeLonKube를 활용한 MNIST 분류모델 학습 결과

본 논문에서는 플랫폼의 예시를 보여주기 위해 간단한 분류 모델을 만들어 보여주려 하였다. 누구나 사용할 수 있는 keras에서 제공하는 MNIST데이터셋을 사용하고, layer는 Multi-Layer Perceptron을 사용하였다.

다양한 하이퍼파라미터를 조합하여 볼 수 있지만, 본 예시에서는 num\_layer, learning rate와 optimizer를 튜닝하는 예시를 보였다. batch\_size는 64로 진행하였으며, optimizer는 sgd, adam, ftrl을 사용하여 진행하였다.

MNIST 분류모델을 학습하기 위해 AWS ec2를 이용해 서버를 생성하였으며, 마스터 노드와, 워커 노드의 사양은 다음과 같다.

	vCPU	메모리	디스크
마스터 노드	4	16GiB	100GiB
워커 노드	2	4GiB	10GiB

표 1. 쿠버네티스 클러스터 구성

쿠버네티스를 이용해 12개의 조합을 3개씩 병렬적으로 수행을 하였으며, 그렇지 않은 경우는 12개의 조합을 1개씩 수행하였다.

전자의 경우, 총 3번의 측정 결과 평균이 406.6초가 나왔으며, 후자의 경우는 평균 1247.6초로 사용자가 직접 하이퍼파라미터 튜닝을 하는 경우에 약 3배의



시간이 더 걸리는 것을 확인하였다. 결과는 아래 표와 같다.

	HPT by Kubernetes			HPT		
	Start	End	Total	Start	End	Total
1 차 측정	22s	7m 13s	391s	22s	20m 21s	1199s
2 차 측정	21s	7m 53s	431s	22s	22m 36s	1334s
3 차 측정	22s	7m 20s	398s	22s	20m 32s	1210s

표 2. 하이퍼파라미터 튜닝 시간측정 결과

본 논문에서는 12개의 조합으로 실험을 진행하였지만 실제 ML 모델의 수많은 하이퍼파라미터 조합으로 하이퍼파라미터 튜닝을 진행한다면 쿠버네티스를 통해 수행한 HPT와 비교했을 때 더 많은 비용이 발생할 것임을 예상할 수 있다.

## 5. 결 론

본 논문에서는 쿠버네티스를 활용한 하이퍼파라미터 튜닝 자동화 플랫폼인 MeLonKube를 소개하였다. 하이퍼파라미터 튜닝 과정에서 클라우드를 사용하여 리소스를 효율적으로 관리하였고, 쿠버네티스를 활용하여 시간과 비용을 절감한 것을 보였다.

MeLonKube는 사용자가 최적의 하이퍼파라미터 조합을 선택하는 것뿐만 아니라 배포까지 할 수 있는 플랫폼을 제공함을 목표로 한다.

### 사 사 문 구

“이 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 통해 수행되었음” (2021-0-01082)

### 참 고 문 헌

[1] IBM, “컨테이너란?“, 2021년 6월 23일 수정, 2021년 6월 23일, <https://www.ibm.com/kr-ko/cloud/learn/containers>

[2] Red Hat, “컨테이너 오케스트레이션이란?“, 2019년 12월 2일, <https://www.redhat.com/ko/topics/containers/what-is-container-orchestration>

[3] 달사쿠, “쿠버네티스(Kubernetes)한번에 정리하기“, 2020년 8월 3일 수정, <https://dalsacoo-log.tistory.com/entry/쿠버네티스-Kubernetes#recentEntries>

[4] Neil Conway, David Hershey, “Data Scientists Don’t Care About Kubernetes”, Determined AI, 2020년 11월 30일 <https://www.determined.ai/blog/data-scientists-dont-care-about-Kubernetes>

# 효과적인 기상 정보 안내를 위한

## 딥러닝 기반 수어 교육 시스템

손혜정<sup>o</sup> 윤선아 임가은 조건형 손윤식

동국대학교 컴퓨터공학전공

thswjd0609@gmail.com gw07167@naver.com lipaomp@naver.com sims7141@dgu.ac.kr

sonbug@dongguk.edu

### For effective weather information guidance Sign language education system based on deep learning

Hyejeong Son<sup>o</sup> Suna Yoon Gaeun Im Gunhyung Cho Yunsik Son

Dept, of Computer Science and Engineering, Dongguk University

#### 요 약

본 연구는 농아인들이 생활속에서 겪는 의사 소통의 문제를 해결하기 위한 연구를 기상정보 도메인을 중심으로 소개한다. 수어의 효과적인 인식을 위해 OpenPose와 BLSTM을 활용하였으며, 데이터 셋은 AI-HUB에서 제공하는 한국수어 영상 데이터의 정면, 위, 아래, 왼쪽, 오른쪽 총 5각도 촬영 영상을 사용하였다. 주요 기상 상황과 이에 적합한 서술 단어를 포함하여 20개의 class를 결정하였다. 각 영상의 효과적인 분석을 위해 대상 영상을 20 프레임으로 요약하고, 각 프레임 별로 OpenPose를 이용하여 영상내의 손가락 위치에 대해 21개의 키포인트를 생성한다. 생성된 키포인트는 BLSTM의 입력으로 사용되며, 이를 이용하여 수어 영상 - 한국어 관계를 학습한다. 제안하는 연구를 통해 청각 장애인 의사소통에 있어서 편리함을 제공하고 일상에서 자주 쓰이는 기상관련 수화에 대한 효과적인 교육이 가능할 것으로 판단된다.

#### 1. 서 론

장애인들의 생활 개선 문제는 최근 사회적 이슈로 크게 드러나고 있다. 최근 장애인을 고용할 시 세금 혜택을 주는 등 장애인 고용 문화를 정착시키려는 정책은 그 중요성이 점점 대두되고 있음을 반증한다. 청각장애인과 타인 간의 의사소통을 위해서는 전문 수어 통역사가 필요하지만, 전문 수어 통역사의 수가 부족하기 때문에 청각 장애인과 비장애인 상호 간의 의사소통이 어려워 많은 불편함이 발생한다. 또한 수어를 필요에 의해 사용해야 하는 청각 장애인 혹은 수어를 통해 청각 장애인에게 정보를 전달할 필요가 있는 비장애인은 정확한 수화 동작을 학습할 수 있어야 한다.

수어와 같이 순차적으로 진행되는 데이터를 학습할 수 있는 다양한 딥러닝 모델이 존재하는데, 그 중 가장 기반이 되는 모델이 바로 RNN이다. 이번 연구는 OpenPose를 이용한 손의 키포인트 좌표를 기반으로 동영상의 프레임을 양방향으로 학습한 BLSTM을 통한 분류 방법을 통해 수화를 자연어로 번역하려 한다.

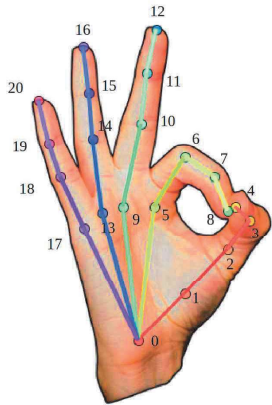
#### 2. 이론적 배경

##### 2.1 OpenPose

OpenPose는 Carnegie Mellon University에서 OpenCV와 Caffe를 사용해서 C++ 기반으로 개발된 오픈소스 라이브러리이다[1]. 단일 이미지에서 실시간으로 25개의 몸체 키포인트, 왼손과 오른손 각각 21개에 해당하는 키포인트와 70개의 얼굴 키포인트, 총 137개의 키포인트들을 인식할 수 있는 모델로 키포인트들을 검출하여 이를 입력된 이미지 또는 비디오 위에 시각화 하여 생성한다. 사람의 동작을 인식해 장비를 제어할 뿐만 아니라, 의사소통에도 도움을 줄 수 있는 기술로 기존의 데이터로 학습된 모델을 함께 제공한다. 또한, 2차원으로 실시간 영상 데이터에서 여러 사람의 키포인트 추출이 가능하다.

입력으로는 이미지, 비디오, 웹캠, IP 카메라와 c++ 형태의 사용자 입력 값을 지원하고 출력으로는 배경 이미지와 키포인트가 표시된 이미지(PNG, JPG, AVI 등)와 키포인트를 JSON, XML, YML 형태로 저장할 수 있다. OpenPose는 손의 위치를 추출하기 위해 [그림 1]과 같이 한 손에는 손목(1) + 손가락(5) \* (손가락 관절(3) + 손가락 끝(1))으로 키포인트 총 21개

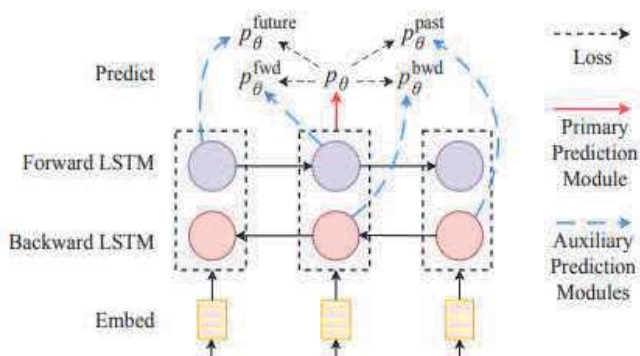
의 2차원 좌표를 따로 받아낼 수 있다[2].



[그림 1] Hand Output Format

## 2.2 BLSTM

BLSTM은 step도 반영한 모델인 양방향 LSTM으로 두 개의 LSTM으로 구성된 시퀀스 처리 모델이다. 한 LSTM은 입력을 순방향으로, 다른 LSTM은 역방향으로 사용한다. step마다 두 모델에서 나온 2개의 hidden 벡터는 학습된 가중치를 통해 하나의 hidden 벡터로 만들어 학습한다[3].



[그림 2] Bidirectional LSTM 구조

## 3. 실험과정

### 3.1 데이터 셋

본 논문에서는 AI-Hub의 한국수어 영상 데이터 셋에서 선정한 기상 관련정보 도메인 데이터로 ‘춥다’, ‘덥다’, ‘먹구름’, ‘바람’, ‘비’, ‘온도’, ‘장마’, ‘햇빛’, ‘싫다’, ‘좋다’, ‘양산’, ‘체온’, ‘하늘’, ‘강풍’, ‘따뜻하다’, ‘홍수’, ‘예정’, ‘사계절’, ‘눈’, ‘피약별’ 총 20개의 수화 단어로 한정한다[4]. 한 class당 정면, 위, 아래, 양옆 총 5각도에서 촬영한 80개의 영상을 확보

하고 총 1600개의 영상을 학습시키는데 사용한다. 영상은 3~4초가량의 길이와 1920x1080의 해상도를 갖는다. 또한 30fps로 촬영하여 영상 길이에 비례하므로 총 프레임은 90~120 프레임이다.

[표 1] 데이터 라벨

Class Num	Class Name	Class Num	Class Name	Class Num	Class Name	Class Num	Class Name
0	춥다	5	온도	10	양산	15	홍수
1	덥다	6	장마	11	체온	16	예정
2	먹구름	7	햇빛	12	하늘	17	사계절
3	바람	8	싫다	13	강풍	18	눈
4	비	9	좋다	14	따뜻하다	19	피약별

### 3.2 데이터 전처리

먼저 OpenPose 라이브러리를 통해 실시간으로 손가락의 관절을 검출하는 과정이 필요하다. Colab Pro 환경에서 수화 영상들을 Google Drive에 저장하고 이와 연동하여 진행했다. 본 연구는 pytorch-openpose의 demo.py를 통해 영상의 키포인트를 추출하였다[5]. 한 영상마다 20 프레임으로 규격을 맞추고 수화 동작에 대한 손의 키포인트 추출을 위해 손을 감지한다.

한 손에 21개의 키포인트에 대한 x, y좌표를 csv파일에 저장한다. csv파일에는 왼손 x좌표 21개, 왼손 y좌표 21개, 오른손 x좌표 21개, 오른손 y좌표 21개가 0부터 83번째 열까지 순서대로 출력되고 84열에는 프레임 번호, 85열에는 class번호가 출력된다.

영상 하나의 정보는 왼손과 오른손의 키포인트 84개와 프레임 20개의 값으로 저장한다. 각 수어 영상에 대한 순차적인 키포인트 정보를 입력으로 넣기 위해 20 프레임에 해당하는 키포인트를 한 행으로 재배열한다. 재배열한 영상 정보 파일을 모델 입력으로 사용한다.



[그림 3] OpenPose를 이용한 수어 영상의 분석 예시

### 3.3 데이터 증강

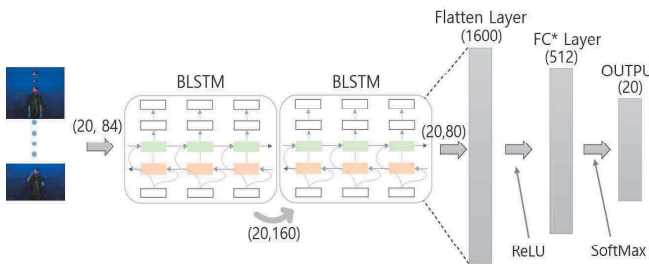
두 가지 문제로 인하여 데이터 셋에 데이터 증강을 적용하였다. 첫 번째는 데이터 셋에 각 단어에 대한 데이터가 충분하지 않다는 것이고, 두 번째는 비슷한 동작에 혼동이 발생하는 것이다.

모든 class의 각 80개의 영상에 대하여 16\*9의 비율인 기존 비디오의 비율을 4\*3으로 변경하여 추가함으로써 데이터를 증강시켰다. 또한, 비슷한 동작인 '1-덥다', '3-바람'과 '8-싫다', '9-좋다' 4개의 class에 대해 20 픽셀만큼 왼쪽으로 shifting시켜 각 데이터 셋을 증가시키는 작업을 하였다.

따라서 각 class 당 1600개의 영상에 비율을 조정한 영상 1600개, 4개의 class에 대하여 shifting한 영상 320개 총 3520개의 영상 데이터를 모델을 훈련시키는데 사용한다.

### 3.4 Model

전처리과정에서 얻어진 1680차원의 feature를 (84, 20)차원으로 변환하여 설계한 BLSTM 모델의 입력으로 사용한다. BLSTM layer를 두 번 반복하고 그 결과를 fully-connect layer를 통해 dataset 클래스의 개수인 20개로 줄여 출력한다. activation function은 relu와 softmax 함수를, loss function은 Categorical Cross Entropy를, optimizer는 adam을 사용하였다.

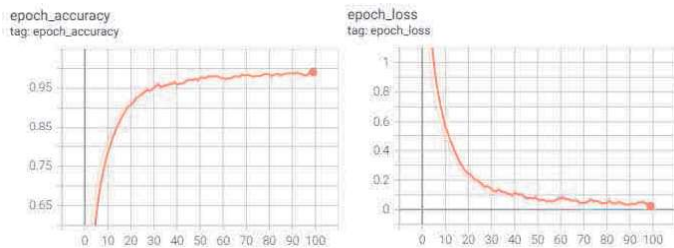


[그림 4] 수어 영상 정보를 입력으로 받아 클래스를 예측하는 Model의 구조

## 4. 실험 결과 및 분석

실험을 위해 OpenPose 라이브러리는 Colab Pro를 통하여 구현했다. 데이터 셋 영상들을 4:1의 비율로 train과 test로 나누고 정규화 시켜 실험을 진행했다. 모델 배치 사이즈는 16으로 설정하였고 100 epoch으로 학습시켰다.

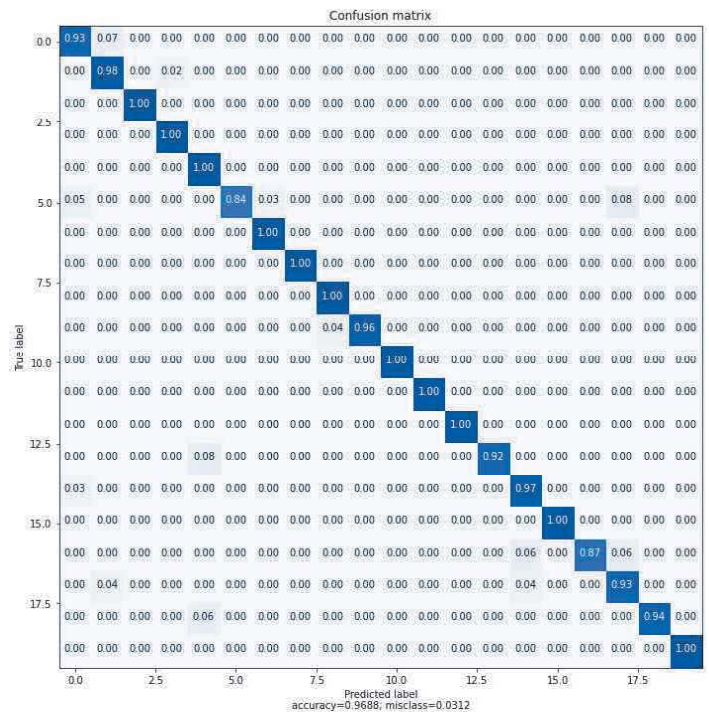
[그림 5]에서 epoch가 지날수록 Accuracy가 지속적으로 증가하는 모습을 볼 수 있으며, 마지막 step에서의 train Accuracy는 99.92%이다. Loss Function인 Categorical Cross Entropy를 사용한 결과 [그림 5]에서 epoch가 지날수록 손실이 감소했음을 알 수 있다. 마지막 step에 도달했을 때의 손실 값이 0.018%로 매우 낮게 측정되었다.



[그림 5] Epoch Accuracy, Loss

[그림 6]은 각 class별로 Test Accuracy를 나타낸 Confusion Matrix이다. Confusion Matrix에 표기된 숫자는 '0-춤', '1-덥다', '2-떡구름', '3-바람', '4-비', '5-온도', '6-장마', '7-햇빛', '8-싫다', '9-좋다', '10-양산', '11-체온', '12-하늘', '13-강풍', '14-따뜻하다', '15-홍수', '16-예정', '17-사계절', '18-눈', '19-피약별'의 특정 기상 도메인 수화 단어이다.

모든 class의 정확도가 매우 높은 것을 확인할 수 있다. 총 평균 Test Accuracy는 96.88%로 비슷한 동작의 프레임이 포함되어 있음에도 불구하고 데이터 증강을 통해 높은 확률로 올바르게 예측하여 분류하였다.



[그림 6] Test Accuracy를 나타낸 Confusion Matrix

## 5. 결론

청각장애인의 의사소통에 있어서 편리함을 제공하기 위해 일상에서 자주 쓰이는 기상관련 수화 번역 시스템을 제작했다. 수화를 인식하는 것은 농아인들의 삶의 질 향상에 중요하다. 비장애인이 본 수화 번역 시스템을 이용하여 기상 관련 분야의 수화를 학습하는 교육 시스템으로도 사용할 수 있다. 이를 통해 일상생활에서 기상을 알고자 하는 장애인들에게 정보를 전달할 수 있으며, 청각장애인의 수화도 번역함으로써 비장애인과의 원활한 소통에 기여하고 대화의 벽을 허물어 줄 것으로 기대된다.

향후 본 연구를 통하여 OpenPose를 활용하여 비수지 요소를 포함한 수화 인식을 할 수 있음을 알고, 더 나아가 기상 도메인을 늘려 기상 전문 분야에서 활용할 수 있을 것으로 기대된다.

## 사 사 문 구

이 논문은 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2022-2020-0-01789)

## 참 고 문 헌

[1] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. IEEE transactions on pattern analysis and machine intelligence, 2019.

[2][https://cmu-perceptual-computing-lab.github.io/openpose/web/html/.github/media/keypoints\\_hand.png](https://cmu-perceptual-computing-lab.github.io/openpose/web/html/.github/media/keypoints_hand.png)

[3] Kevin Clark, Minh-Thang Luong, Christopher D. Manning, and Quoc Le. Semi-Supervised Sequence Modeling with Cross-View Training. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pages 1914–1925, 2018.

[4] AI DATA 사업본부, 2020. 수어 영상, 한국지능정보사회  
진 흥 원 ,  
<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=103>

[5] “pytorch-openpose”, Hzzone github, 2021년 6월 21일  
수정, 2022년 09월 07일 접속,  
<https://github.com/Hzzone/pytorch-openpose>

# IoT Hub 보안성 향상을 위한 취약점 분석 프로세스에 관한 연구

장지원<sup>o</sup> 윤원희 김성은 김지연 안영준 강준혁 강인욱 김홍진

한국정보기술연구원 차세대 보안 리더 양성 프로그램 (Best of the Best)

[nebulaboratory@naver.com](mailto:nebulaboratory@naver.com), [creampenne@gmail.com](mailto:creampenne@gmail.com), [jinna0319@gmail.com](mailto:jinna0319@gmail.com),  
[jyeoun5713@gmail.com](mailto:jyeoun5713@gmail.com), [zerogeas1234@gmail.com](mailto:zerogeas1234@gmail.com), [kangsecu@gmail.com](mailto:kangsecu@gmail.com),  
[hkkiw0823@gmail.com](mailto:hkkiw0823@gmail.com), [hjkim0892@gmail.com](mailto:hjkim0892@gmail.com)

## A Study on the Vulnerability Analysis Process for Improving IoT Hub Security

Jiwon-Jang<sup>o</sup> Wonhee-Yoon Sungeun-Kim Jiyeoun-Kim Yeongjun-An  
Junhyuk-Kang Inwook-Kang Hongjin-Kim

Best of the Best, Korea Information Technology Research Institute

### 요 약

IoT(Internet of Things) 기술은 우리 삶 깊숙하게 자리 잡고 있다. 편리한 주거 환경을 만들어 주는 스마트 홈부터 자동화된 공정을 관리를 위한 스마트 팩토리, 나아가 이용자와 관리자의 편리성을 추구할 수 있는 무인 매장까지 구성할 수 있게 해준다. 각각의 IoT 기기들은 중심이 되는 IoT Hub에 연결되어 통신을 주고받는다.

IoT 환경을 안전하게 이용하기 위해서는 IoT Hub의 보안성이 보장되어야 한다. 이를 위해서는 제품 출고 전 취약점 분석을 수행하여 보안 위협 발생 가능성을 낮추는 것이 중요하다. 그러나 임베디드 시스템으로 설계된 IoT Hub는 기기의 특수성으로 인해 취약점을 식별하고 분석하는데 어려움이 많다. 이에 따라 본 논문에서는 IoT Hub의 보안성을 향상시키기 위해 특수한 환경을 고려한 취약점 분석 프로세스를 연구하여 제안하고, 보안성을 높일 수 있는 방안을 제시한다.

### 1. 서 론

스마트 홈, 팩토리, 상점 등 점차 스마트 환경이 조성되어 감에 따라 센싱 및 통신 기능을 내장한 IoT 기기들이 만연해지고 있다. 이러한 흐름은 시간과 공간에 제약 없이 정보통신 서비스를 제공할 수 있음을 의미하는 '유비쿼터스'를 가능하게 한다. 이때 IoT 기기들을 관리·관제하기 위해 IoT Hub(Gateway)를 중심에 두어 사용한다.

IoT Hub는 다양한 IoT 기기들을 네트워크로 연결하는데, 이때 다양한 데이터의 통로 역할을 한다[1]. 이러한 IoT 기기들은 소형화, 경량화, 저전력 및 저비용 등의 특징으로 강력한 암호화를 수행하지 못하거나 주요키를 하드코딩 하는 등의 방식을 채택하여 보안성을 보장할 수 없는 경우가 많다.

그러나 많은 기기 간의 상호 연결 관계가 깊어진 유비쿼터스 사회에서 IoT Hub의 보안 위협은 단일의 침해 사고로 끝나지 않는다. 가령, 과거 CVE-2020-6007 취약점에서는 Philips Hue IoT Hub에서 발생한 취약점을 시작으로 홈 네트워크 전체를 장악할 수 있었다[2].

따라서 안전한 스마트 환경을 위한 IoT Hub를 향한 보안성 고찰은 필요하며, 일반에 공개되기 전 철저한 취약점 분석을 수행하여 사고 발생 가능성을 낮추어야 한다. 하지만 IoT Hub

의 취약점 분석 수행은 일반적인 웹/앱 애플리케이션 분석에 비교하여 상대적으로 난도가 높으며, 비용과 시간이 더 많이 들게 된다[3]. 따라서 본 논문에서는 최신 IoT Hub에 적용할 수 있는 일련의 취약점 분석 프로세스에 관해 연구하여 제안하고, 연구 결과에 따른 보안성 강화 방법을 도출하여 제시한다.

### 2. 배경 지식

#### 2.1 임베디드 시스템

임베디드 시스템이란 특수한 목적으로 제어를 필요로 하는 시스템을 제어하기 위해 특정 기능을 수행하도록 설계된 시스템이다. 임베디드 시스템은 기본적으로 냉장고, 에어컨, 리모컨, 카메라 등과 같이 단일 목적을 가지도록 시스템을 설계하기 때문에 목적에 따라 임베디드 시스템의 요구사항이 달라진다. 임베디드 시스템은 저전력, 저사양으로 설계되는 경우가 많아 많은 기능을 필요로 하지 않는 기기에 사용하기에 적합하다. 이러한 임베디드 시스템에 네트워크가 연결되어 그 역할이 확장된 것이 IoT 이다. 그러나 현재는 IoT 기기들이 네트워크 환경뿐만 아니라 센싱 기능과 클라우드 컴퓨팅, 빅데이터와 같이 다양한 기능을 요구하고 있어 고사양의 처리 능력

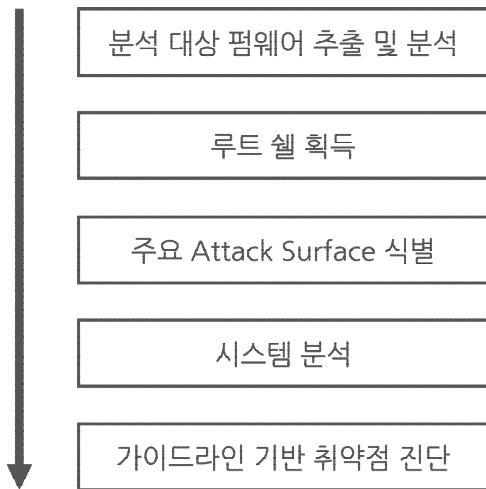
을 보유한 임베디드 시스템을 도입하여 이러한 IoT 기기들을 총괄하는 IoT Hub를 중심으로 더 다양한 방식으로 IoT 기기들을 활용할 수 있게 되었다.

## 2.2 펌웨어

펌웨어란 특정 하드웨어를 제어하는데 용이하도록 만든 소프트웨어이다. 펌웨어는 특정 하드웨어를 위해 제작되어 해당 하드웨어를 통해서만 정상 동작하는 경우가 많아 분석 시 내부 로직을 파악하여 하드웨어 종속성을 식별하는 작업이 필요하다[4]. 펌웨어는 부트로더, 커널, 파일 시스템으로 구성되어 있다. 부트로더는 주요 하드웨어 자원을 초기화하고 시스템 시작을 위한 작업을 수행한 뒤 커널을 트리거 하는 역할을 한다. 커널은 하드웨어와 응용 프로그램 사이의 인터페이스를 제공하는 역할을 하며 가장 핵심이 되는 요소이다. 또한 서비스와 응용 프로그램이 동작할 때 필요한 시스템 자원을 할당하고 관리한다. 파일 시스템은 펌웨어가 사용하는 대부분의 정보를 포함하고 있어 이를 분석하면 펌웨어 시스템 내부 구조 파악에 있어 유용하다.

## 3. 취약점 분석 프로세스

IoT Hub의 특징에 기반하여 취약점 분석 프로세스를 도출하였다. 분석 과정은 다음 [그림 1]과 같다.



[그림 1] 제안하는 취약점 분석 프로세스 순서도

### 3.1 분석 대상 펌웨어 추출 및 분석

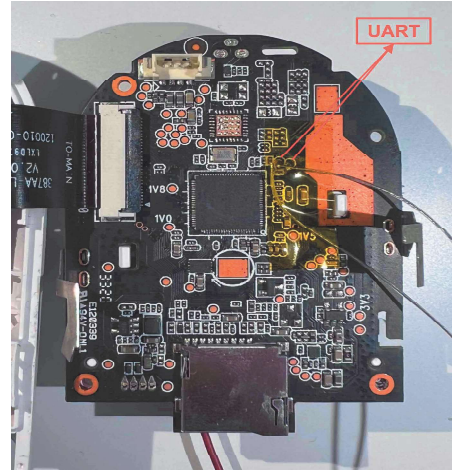
IoT Hub에 관한 취약점 분석을 수행하기 위해서는 내장된 시스템의 펌웨어를 추출할 수 있어야 한다. 펌웨어를 확보할 수 있는 가장 간단한 방법은 제조사에서 실제 바이너리를 공개한 경우이다. 그러나 바이너리를 제공하지 않는 경우에는 다양한 방식으로 추출할 수 있다. 대표적으로 디버깅 포트를 통한 시리얼 통신과 네트워크 패킷을 스니핑하는 방법이 있다.

시리얼 통신은 시스템의 디버깅을 위해 탑재해 둔 디버깅 포트를 식별하여 접근하는 방법이다. 대표적인 예시로 UART

(Universal Asynchronous Receiver Transmitter), JTAG(Joint Test Action Group) 포트가 있다. 이러한 시리얼 통신을 통해 펌웨어 정보를 파악 및 추출하거나, 내부 시스템에 접근하여 셸 획득, 나아가 동적 분석 등이 가능하다[5].

이러한 디버깅 포트가 물리적인 접근이 차단되어 있을 경우 플래시 메모리 디스더링 작업을 통해 펌웨어를 추출하는 방법도 있다.

유무선 통신으로 시스템 업데이트를 하는 경우, 해당 과정에서 MITM(Man-in-the-Middle) 공격 등으로 내부 패킷을 스니핑하여 펌웨어를 획득할 수 있다.



[그림 2] 내부 시스템 접근을 위해 UART 디버깅 포트를 찾아 연결한 모습

### 3.2 루트 셸 획득

셸(Shell)은 펌웨어의 핵심인 커널과 사용자를 연결하는 명령행 인터페이스로 사용자의 명령을 해석하고 프로그래밍 기능과 사용자 환경 설정 기능을 제공하는 역할을 한다. 따라서 시스템 분석에서의 셸 획득은 시스템과 직접적인 상호작용이 가능하여 다양한 것을 가능하게 한다. 임베디드 리눅스 시스템에서 셸은 크게 2가지가 존재하며, 각 부트 셸(Boot Shell), 루트 셸(Root Shell)이다. 부트 셸은 부트 로더가 부팅 작업을 수행할 때 인터럽트를 발생시켜 중간에 부팅 시퀀스를 멈추게 하고, 그 권한을 취득하는 것을 말한다. 이를 통해 하드웨어와 관련된 정보를 검증할 수 있다. 루트 셸은 bash, zsh 등과 같이 리눅스 환경에서 사용하는 셸이 Root 권한을 획득한 상태를 말하며, 해당 시스템 전반을 장악하고 내부 구조를 파악할 수 있으며, 추출할 수 있게 한다.

### 3.3 주요 Attack Surface 식별

효과적인 취약점 탐색을 위해서는 전체 바이너리를 분석하기보다는 취약점 내재 가능성이 높은 부분을 식별하여 우선 순위화하는 과정이 필요하다. 주요 분석 Surface로는 일반 사용자의 입력값을 받아 내부 연산이 수행되는 부분을 선정할 수 있다. 일반 유저와 상호작용하는 인터페이스 부분은 예외가 발생할 가능성이 상대적으로 높기 때문이다.

두 번째로 근거리 저전력 통신 프로토콜 혹은 Wi-Fi 등의 통신을 위한 API 호출 부나 외부 스트림 값을 받아와 파싱하는 부분이 될 수 있다.

세 번째로 부팅 과정 중 초기화 프로그램 및 스크립트를 확인할 수 있다. IoT 기기에 전원 공급이 이루어질 때 펌웨어 시스템 중 부트로더가 동작하게 되고, 이 부트로더는 하드웨어를 초기화하고 플래시 메모리에 저장된 커널과 디스크 이미지를 로딩한다. 이때의 과정을 분석하게 되면 초기화 스크립트 조작과 같이 IoT 기기의 공격 벡터를 식별할 수 있게 된다.

마지막으로 실제 IoT 기기 내부에서 실행되는 프로세스들을 확인하면서 공격 벡터가 될 수 있는 불필요한 포트와 데몬들을 식별할 수 있다.

### 3.4 시스템 분석

획득한 펌웨어 시스템을 분석할 수 있는 방법으로는 크게 정적 분석과 동적 분석이 있다. 정적 분석(static analysis)이란 펌웨어 바이너리 파일을 실행하지 않고 분석하는 것을 말한다. 많은 경우 역어셈블러, 디컴파일러 등을 통하여 소스코드 단에서 분석하며, 취약한 함수 사용 여부나 logic 오류를 찾는 것을 목적으로 한다[6]. 하지만 정적 분석은 동적 요소를 고려할 수 없어 한정된 분석 범위를 가진다는 단점이 있다.

반면 동적 분석(dynamic analysis)이란 실제 프로그램을 실행시킨 후, 그 반응을 보며 분석하는 것을 말한다. 펌웨어를 구동시킬 때는 하드웨어 장치에 영향을 받는 경우가 많으며, 이를 해소하기 위해 에뮬레이션 기술을 접목한다. 동적 분석에는 퍼징(Fuzzing) 기술이나 오염 분석(Taint Analysis), 기호 실행(Symbolic Execution) 등이 있다. 다양한 테스트 케이스 등이 사용되며, 정적 분석을 통해 파악한 정보를 바탕으로 동적 분석하면 더 효과적으로 오류를 발견할 수 있다.

### 3.5 가이드라인 기반 취약점 진단

다음 [표 1]에 표기한 가이드라인은 IoT를 대상으로 보안성 강화 방법을 제시하고 취약점 점검을 권고하는 내용을 포함한 문서이다. 이를 활용하면 취약점을 진단하는데 더욱 효율적이고 체계적으로 진행할 수 있다. 이러한 문서를 활용하여 대상 기기에 맞춘 점검 항목을 도출하여 취약점을 진단 및 점검한다.

[표 1] 취약점 분석 시 활용 가능한 가이드라인 목록

No.	발행 연도	주관	가이드라인 명
1	2016	KISA	사물인터넷 소형 스마트 홈/가전 보안 가이드
2	2017	KISA	홈·가전 IoT 보안 가이드라인
3	2021	OWASP	OWASP Top 10 IoT

데이터의 흐름을 도출하여 DFD(Data Flow Diagram)를 제작하고, 이를 토대로 위협 모델링을 수행하거나 어택 시나리오를 구상해 위협을 식별한다면 효과적으로 취약점을 분석할 수 있다.

## 4. 논의 및 보안성 강화 방법 제안

IoT 기기를 출고할 때 내부적으로 본 논문에서 제안하는 바와 같이 취약점 검토를 더욱 세밀하게 진행해야 한다. 취약점이 내재된 기기가 상용화된다면 기기와 적합하게 설계된 시스템에 보안 패치를 적용하기는 어렵기 때문이다. 그러나 실제 취약점 검토가 이루어지더라도 패치를 적용하기에 소형화, 저전력, 원가 절감 등의 이유로 trade-off 하는 경우가 많다.

출고 시 장비에 내부적으로 접근하기 어렵게 장비 자체에 실리콘, 글루건, 아크릴 등으로 보호하는 작업이 필요하고, UART와 JTAG와 같은 디버깅 포트의 흔적을 지우는 작업을 수행해야 한다.

더불어 사용자는 IoT Hub를 IoT 환경에 구성할 때 초기 비밀번호를 반드시 변경하여 사용하고, 일정 강도 이상의 패스워드를 설정하려는 노력이 필요하다.

## 5. 결론

본 논문에서는 IoT Hub의 보안성 강화를 위하여 사전에 취약점을 식별할 수 있는 과정을 연구하여 소개하였다. IoT Hub는 임베디드 장비이며, 일반적이지 않은 리눅스 펌웨어 시스템을 사용한다는 점이 분석 시 고려해야 할 특징이다. 이에 따라 일반 웹/앱을 분석할 때와는 다른 접근 방식을 취해야 한다.

IoT Hub는 향후 발전을 거듭함에 따라 다양한 종류의 IoT 기기와 연결될 것이고, 그 기능은 더욱 다양해질 것으로 전망된다. 따라서 IoT Hub의 내부 구조를 파악하고 그 보안성을 강화하려는 노력은 지속적으로 필요하다. 향후 본 연구의 분석 프로세스를 실제 기기에서 실현하여 다양한 취약점을 찾을 수 있도록 연구할 계획이다.

## 참고 문헌

- [1] J. Park, S. Kang, and S. Kim, "Study of Security Requirement of Smart Home Hub through Threat Modeling Analysis and Common Criteria," Journal of the Korea Institute of Information Security & Cryptology, Vol. 28, No. 2, p. 513-528, Apr. 2018. (in Korean)
- [2] Common Vulnerabilities and Exposures, "CVE-2020-6007" [Internet], <https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2020-6007>
- [3] Y. Cho, "A Study on System Development to Improve IoT Firmware Vulnerability Analysis Efficiency," Korean Journal of Industrial Security, Vol. 12, No. 1, pp. 31-50, Feb. 2022. (in Korean)
- [4] E. Im, "Firmware Analysis of Embedded Devices," Communications of the Korean Institute of Information Scientists and Engineers, Vol. 34, No. 1, pp. 16-21, Jan. 2016. (in Korean)
- [5] Y. Kim, J. Gim, and S. Kim, "A Study on Systematic Firmware Security Analysis Method for IoT Devices," Journal of the Korea Institute of Information Security & Cryptology, Vol. 31, No. 1, pp. 31-49, Feb. 2021. (in Korean)
- [6] S. Lee, W. Oh, S. Park, E. Cho, and I. Baek, "Binary



Vulnerability Analysis Framework Combining Static and Dynamic Analysis," Journal of KIISE, Vol. 45, No. 12, pp. 1217-1226, Dec. 2018. (in Korean)

# 유니티를 이용한 학습용 콘텐츠 기반 게임 설계\*

강지윤<sup>1</sup> 안진수 이수민 김재우 최도진

창원대학교 컴퓨터공학과

dkswlstn6321@naver.com, looloo404@hanmail.net, 20193081@gs.cwnu.ac.kr

lgap5@naver.com, dojinchoi@changwon.ac.kr,

## A Design of Learning-Contents Based Game Using Unity2D

Jeeyoon Kang<sup>1</sup> Jinsoo Ahn Sumin Lee Jaewoo Kim Dojin Choi

<sup>1</sup>Department of Computer Engineering, Changwon National University

### 요약

턴제 게임은 실시간 게임과는 달리 플레이어들이 차례대로 돌아가면서 명령을 수행하는 전략 게임의 한 장르로써 한 플레이어의 차례가 끝나면 다음 플레이어의 차례가 돌아오는 식으로 플레이하게 된다. 본 논문에서는 학습용 턱제 게임을 제안한다. 각종 게임 제작을 위한 편리한 기능을 제공하는 유니티 엔진을 활용하였다.

### 1. 서론

턴제 전략 게임은 플레이어들이 차례대로 돌아가면서 명령을 수행하는 전략 게임의 한 장르이며 이에 해당하는 대표적인 예시로는 체스, 장기와 틱택토 같은 보드 게임, 고스톱, 포커 같은 카드 게임에 주로 쓰이는 방식이다. 현재 PC, 모바일 게임의 예시로는 다키스트 던전, 프린세스 커넥트, 하스스톤, 새도우 버스 등의 게임이 있다. 턱제 전략 게임은 플레이어의 반응 속도나 현란한 기술에 의존하지 않는 특징을 지니고 있다. 현재 스마트 폰을 이용한 모바일 게임 시장이 성장하는 한편, 현란한 컨트롤이 제한되는 모바일 환경에서 AOS,RPG 장르와 비교하여 가볍게 즐길 수 있는 턱제 전략 게임이 모바일 게임 시장을 위주로 빠르게 성장하고 있다. 글로벌 게임 산업 동향 분석 매체 ‘WePC’는 모바일 게임 시장 매출이 부상할 것으로 예상하였으며 2016년 기준으로 67억 달러였던 모바일 전략 게임 시장 규모가 2023년까 133억 달러로 약 2배 성장할 것으로 내다봤다[1]. 유니티는 3D 및 2D 비디오 게임의 개발 환경을 제공하는 게임엔진이며 3D 애니메이션, 건축 시각화, 가상현실 등 반응형 콘텐츠 제작을 위한 도구이며 현재 게임 개발자들이 가장 많이 사용하는 엔진이기도 하다[2-3]. 본 논문에서는 유니티2D의 이미지 기능인 Sprite, 물리 기능인 Collider, Physics2D, 애니메이션 기능을 위한 Animator, Animation 등을 이용하여 싱글 플레이 타일 기반의 턱제 전략 게임을 설계하고자 한다. 제안하는 게임은 학습용 콘텐츠가 추가된 형태로 구성되었다. 다양한 퀴즈 활동을 통해 게임 진행에 이점을 부여하고 게임 사용자에게 학습 콘텐츠를 제공한다.

\* 이 논문은 2022학년도 자기주도 연구역량 지원사업 연구비에 의하여 연구되었음

### 2. 제안하는 학습용 콘텐츠 기반 게임

#### 2.1 시나리오

그림 1은 본 논문의 콘텐츠의 시나리오를 나타내고 있다. Loadscene과 Loadscene\_additive의 차이점은 다른 scene으로 이동 시 Loadscene을 사용할 경우 기존의 scene의 객체가 모두 사라지지만 Loadscene\_additive는 기존 scene의 객체를 보존한 상태로 다른 scene으로 이동할 수 있다[4]. 그렇기 때문에 다시 한 번 진행될 필요가 없는 scene의 경우, Loadscene을 통해 기존의 객체를 삭제하도록 제작하였다. Title 화면에서 start 버튼을 클릭하여 턱제 게임 scene으로 이동하며 이 턱제 게임 scene이 본 논문의 콘텐츠의 main scene이다. 턱제 게임 scene에서 카드를 클릭함으로써 퀴즈 게임 scene으로 이동할 수 있으며 퀴즈 게임 scene에서 특정 점수를 넘었을 경우 이로운 효과를, 그렇지 않을 경우, 다음턴에 다시 도전할 수 있다. 주 게임의 흐름은 턱제 게임 scene과 퀴즈 게임 scene의 이동으로 이루어지며 아군이나 적의 킹 유닛이 사망 시 Win/Lose UI가 출력된다. Win/Lose UI는 따로 scene으로 만들지 않고 유니티에서 제공하는 canvas UI로 제작하였으며 retry 버튼을 클릭하면 LoadScene 함수를 통해 턱제 게임 scene을 재호출하며 게임이 처음부터 다시 실행된다.

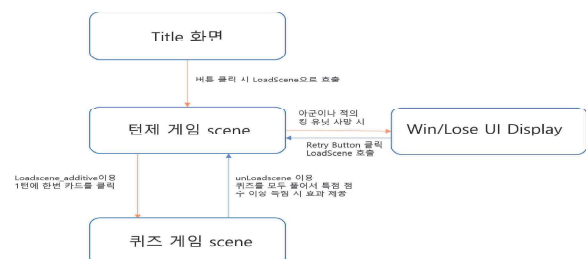


그림 1 시나리오

## 2.2 전체 프로그램 기능도

그림 2와 3은 제안하는 콘텐츠의 주요 클래스들을 나타낸다. 턴제 게임 주요 클래스로는 GameManager, Units, Tile이 있으며 각 클래스는 유니티의 클래스인 MonoBehaviour를 상속받고 있으며 이를 통해 유니티의 start, update 등의 주요 함수들을 사용할 수 있다. GameManager는 게임의 전체적인 UI 초기화 및 파라미터의 초기화, 적 유닛의 행동을 담당하며 인스턴스는 한 개만 생성된다. Tile은 유닛이 움직일 수 있는 지면 부분이며 highlight 기능을 통해 색감을 바꿈으로써 시각적 효과를 나타내며 장애물 존재 여부를 통해 이동가능한 타일인지 식별한다. Units는 Units 인스턴스의 공격, 피격, 사망, 이동등의 역할을 담당하며 Units 인스턴스의 체력 공격력 같은 스테이터스를 조정할 수 있다.



그림 2 턴제 게임 주요 클래스 기능도

퀴즈 게임 Scene은 턴제 게임에서 Card를 클릭하면 이동이 가능하다. QuestionSO, Quiz, QuizGameManager 등의 주요 클래스가 있으며 QuestionSO는 유니티에서 제공하는 컨테이너로 성능적 이점과 유니티에서의 인터페이스적 이점 때문에 사용하였으며 컨테이너는 퀴즈 게임을 위한 문제와 답안으로 구성되어 있다. Quiz 클래스는 Quiz 씬에서의 플레이어가 문제를 선택할 때마다 UI 초기화 및 문제, 답안, 결과 화면의 Display를 총괄한다. QuizGameManager는 턴제 게임 Scene과 퀴즈 게임 Scene을 연결해주는 매개체 역할을 하며 점수에 따라 유닛의 각종 이로운 효과나 턴제 게임에서의 기믹 해제를 효과를 제공한 뒤 턴제 게임으로의 Scene 이동을 하는 클래스이다.

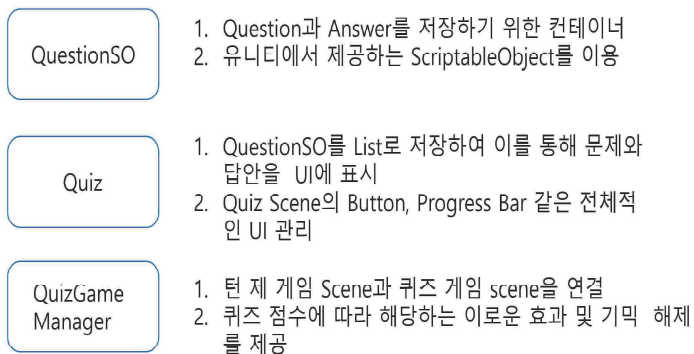


그림 3 퀴즈 게임 주요 클래스 기능도

## 2.3 GameManager

GameManager 클래스는 턴 종료, 턴제 게임에서의 각종 UI 초기화, 싱글 게임에서의 적 움직임 구현 등의 기능을 가지며 1개의 인스턴스만 생성된다. 이 논문의 게임에서의 승리 조건

은 적의 킹 유닛을 쓰러뜨리는 것이며 이를 위해 GameManager는 킹 유닛의 움직임을 추적하기 위한 인수도 내포하고 있다. GameManager의 핵심 함수로는 EndTurn함수 및 AIAttack 함수가 있으며 플레이어가 턴을 마치기 위한 EndTurn버튼을 누를 시 EndButton 함수가 동작하며, 각종 UI의 초기화와 누구의 턴이냐에 따라 AIAttack의 함수가 실행되며 적의 공격 및 이동이 실행된다. AIAttack 함수는 유닛들이 차례로 이동 및 공격 할때의 시간 간격을 위해 코루틴 함수로 구현하며 트리를 이용하여 최대 몇 턴까지의 움직임을 볼 것인지를 결정할 수 있다. 본 논문에서는 유닛들의 최대 이동 거리를 고려하여 트리의 레벨을 3으로 설정하였으며 본 논문의 게임에서는 사용하지 않았으나 TileBased Turn based Game의 특성의 특성을 고려하였을 때 체스 등에서 쓰이는 minimax 알고리즘을 변형해서 사용할 수도 있을 것이다.

## 2.4 Tile

Tile은 유닛들이 움직이는 지면이며 collider2D를 이용하여 Obstacle의 유무를 판단한 뒤 이동 가능한가를 판단할 수 있으며 highlight 기능과 hovering Effect를 통한 시각적 효과 및 유닛들의 플래그 여부에 따라 타일을 마우스 오른쪽을 클릭 시 Units 인스턴스가 내포하고 있는 Move함수 호출 또는 유닛을 구매 시 boolean 변수를 이용하여 Units 인스턴스를 생성시킨다.

## 2.5 Units

Units 인스턴스는 보드 게임에서의 장기 말들에 해당하며 게임 캐릭터의 체력 공격력, 이동 범위 같은 스테이터스 인수들을 포함하고 있으며 유닛의 공격, 피격 사망 및 각 유닛들의 애니메이션 효과와 시각적 효과들을 위한 함수 및 파라미터를 가지고 있다. 시각적 효과를 위한 함수로는 GetWalkableTiles가 있으며 캐릭터 클릭 시 Units인스턴스의 이동거리를 기준으로 거리를 계산하여 이동할 수 있는 타일들의 highlight 함수를 실행시켜 플레이어가 Units인스턴스의 이동할 수 있는 타일들을 알 수 있게 한다. Units 클래스의 이동 거리는 캐릭터와 타일들간의 유클리드 거리 기준으로 산출했으며 각각의 타일들간의 중심 간격은 1로 설정했다. 이 중심으로 유닛들이 이동하며 좌우 이동은 1, 대각선 이동은 2의 비용이 든다. 이를 위한 Move 함수에서는 이동할 때의 시간 간격을 위해 코루틴 함수를 사용하였다. 또한 그 외의 공격할 수 있는 유닛들을 저장하는 GetEnemies 함수, 공격과 피격, 사망을 처리하는 Attack 함수 등이 있다. 각 유닛들은 한 턴에 한 번씩 이동 및 공격을 할 수 있으며 이를 위한 플래그는 hasMoved와 hasAttacked이며 health, attackDamage, defenseDamage, armor 등은 유닛의 공격 피격등의 데미지 산출을 위한 파라미터이다. 또한 이는 유니티의 인스펙터 기능을 통해 밸런스 조절을 위한 변경이 가능하다. 그 외에 particle Effect, camera shaking, 음향 효과를 위한 인수와 적과 아군을 구별하는 playerNumber 등의 인수들을 볼 수 있다.

## 2.6 Prefab

유니티에서는 prefab으로 씬 안에 있는 게임 오브젝트를 얼마든지 만들 수 있으며 앞에서 서술한 클래스들을 삽입하여 제작하였다. 이러한 prefab을 수정할 경우 수정한 내용이 그 prefab으로부터 생성된 인스턴스에도 반영이 되며 개별 인스

턴스 또한 수정이 가능하다[5].

prefeb을 제작하는데 필요한 예셋은 유니티의 예셋스토어와 Black Thorn prod라 불리는 개발자의 교육용 예셋 및 유니티의 교육 사이트인 GameDev의 무료 배포용 예셋을 사용하였으며 이에 더해 필요에 따라 유니티에서 제공하는 collider, AudioSource, Animator 등의 컴포넌트를 추가하여 제작하였다. 단순히 캐릭터 뿐만 아니라 아이콘, particle Effect 같은 시각적 효과들, Text Mesh Pro 같은 유니티 UI 또한 prefeb으로 제작이 가능하다. 또한 컴포넌트 추가의 경우 Add Componets 버튼을 클릭 함으로써 간편하게 추가할 수 있다.

### 2.7 개발 결과

그림 4.5 는 제작한 게임의 완성된 형태이며 그림 4의 우측 화면에 보이는 카드들을 클릭함으로 그림 5의 퀴즈 scene을 Load할 수 있으며, 이를 위해 유니티의 additive\_scene 함수를 이용하였다. 퀴즈 scene에서는 비트들의 And와 Or 연산,정보 처리 기사의 문제와 컴퓨터 공학적 상식문제를 기반으로 문제를 만들었으며, 오른쪽상단의 원형의 객체는 타이머이며 SCORE 옆의 막대는 퀴즈의 progress Bar이다. 본 퀴즈 scene을 통해 플레이어들의 컴퓨터 공학적 지식에 대한 흥미를 고취 시키는 것을 목표로 하고 있다.



그림 4 완성된 형태

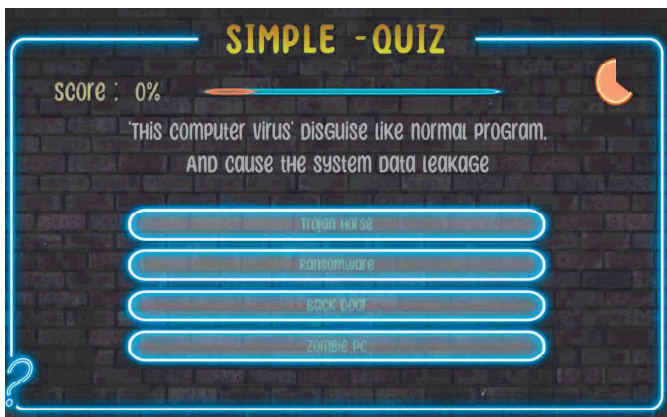


그림 5 퀴즈 게임

### 3. 결 론

본 논문에서 제작한 게임에서는 유니티의 기초적인 개념만을 주로 사용하였으며 애니메이션 particle Effect 및 퀴즈 scene에서 Unity의 UI 등에 대해서 숙련도의 문제가 있었으나 향후 더욱 많은 기능을 담고 있는 게임을 개발함으로써 해당 게임에 대한 다양한 방향을 제시하고자 하며 이를 통해 더욱 높은 수준의 게임 개발 방향을 연구할 계획이다.

### 참 고 문 헌

- [1]<https://gametoc.hankyung.com/news/articleView.html?idxno=69031>
- [2] 이한성, 류승택, 서상현. “유니티와 언리얼 엔진 4 의 구조와 구현 방식 비교 연구,” *Journal of the Korea Computer Graphics Society*, Vol. 25, No. 4, pp. 17-24, 2019.
- [3] Capundan, Paolo Angelo P., et al. ““TAKIPSILIM” : A Fantasy-Themed 2D Turn-Based RPG Developed Using Unity Engine.” *International Journal of Computing Sciences Research* 6 ,2022.
- [4] Unity Documentation version(2020.3)
- [5] 오노 고지,웁진이 김범준 『유니티 2D 게임 공작소』 길벗 2015 Prefeb Chapter

# 생산라인의 이상유무와 예지보전 판단을 위한 인공지능

## 기반 모바일 애플리케이션

김명균, 전왕수, 이상용

경남대학교 컴퓨터공학부

kmg5524@naver.com, jws2218@naver.com, syrhee@kyungnam.ac.kr

## A Mobile Application based on Artificial Intelligence to Determine Whether the Production Chain is Anomaly and Predictive Maintenance

Myoung-Gyun Kim, Wang-Su Jeon, Sang-Yong Rhee

Dept. Computer Engineering, University of Kyungnam

### 요약

본 논문에서는 현장 근무자들이 사용할 수 있도록 인공지능을 통해 구현된 예지보전을 모바일 앱을 개발하였다. 이상유무를 판단하기 위해 DNN, XGBoost, Random forest 모델을 사용하였고, 각 모델의 정확도는 99.71%, 99.99%와 99.99%의 정확도를 보였다. 그리고 예지보전을 예측하기 위해 seq2seq 모델을 사용하여 예측한 결과 MSE에서  $5.76e-05$ , MAE에서 0.005,  $r^2$ score에서 0.85의 정확도를 보였다. 우리가 개발한 앱을 현장 작업자들이 활용하도록 실시간 데이터 시각화, 공장라인의 상태, 설비현황과 상세화면이 나오도록 구현했다. 이 앱을 통해 실시간으로 작업자에게 알람을 제공함으로써 생산라인의 이상유무와 예지보전이 용이할 것으로 판단되고 차후 예지보전기술의 현장 적합성이 증가할 것으로 기대한다.

### 1. 서론

세계 스마트 팩토리 시장은 2019년부터 2024년까지 연평균 성장률은 9.8%이며, 국내 시장 또한 11.4% 성장할 전망이다. 국내에선 과거 스마트 팩토리를 2018년도부터 2022년까지 3만 개를 보급했으나, 최근엔 수요산업 특성에 맞춘 질적 고도화 전략으로 전환되었다[1]. 질적 고도화 전략에선 공장의 설비효율을 나타내는 지표, 설비종합효율(Overall Equipment Effectiveness)이 중요하다.

Deloitte insights에 따르면 예지보전기술이 설비종합효율을 90% 상승 시킬 수 있다고 한다[2]. 해외에는(히타치제철소, 아사히 카세이 엔지니어링, JFE스틸 등) 이미 예지보전을 적극 활용하고 있다[3]. 또한 국내 공장들은 비스포크 상품의 생산 및 공급을 지향하고 있으며, 이러한 비스포크 상품생산은 기존의 예방보전 기술보다 예지보전 기술이 더욱 적합하다[3]. 현재 예지보전 기술은 비현장직 전문가들이 데이터를 관리하지만, 현장직 근무자의 작업환경 및 경험을 고려한 데이터를 관리한다면, 예지보전의 데이터 적합성을 높일 수 있다[4-5].

에어릭스 사의 U-SCADA시스템은 모니터링 대시보드 화면에 설비이상(전류이상, 모터이상, 레일이상) 정보를 확인하는 것이 가능하지만, 화면에 설비 이상 정보를 띄우기만 해서 사고에 대한 즉각적 대처와 작업환경이 고려된 데이터를 파악하는 것이 어렵다.

본 논문에서는 이상유무 판단과 예지보전을 위한 애플리케이션을 개발하였다. 이 앱은 DNN(Deep Neural Network),

XGBoost, Random forest 모델을 사용하여 공장 설비들의 상태를 실시간 확인하고, seq2seq모델을 사용하여 설비들을 예지보전 하며, 대시보드를 모바일 앱으로 제공하면, 예지보전의 현장 적합성과 현장작업의 안정성을 높일 수 있다.

### 2. 스마트 팩토리 예지보전 애플리케이션

#### 2.1 실시간 데이터 시각화

실시간 데이터 모니터링 화면은 서버로부터 데이터가 수신되면 그 즉시 그래프에 데이터 값을 시각화 할 수 있다.

#### 2.2 가동률 표기

시간가동률은 공장에서 설비가 실제 작동하는 시간을 나타낸다. 이를 앱 화면에 표기하여, 공장의 다운타임(Down Time)에 대해 효과적으로 대응할 수 있다. 식 1의 시간가동률은 설비의 정지를 제외한 실제가동시간(부하시간-정지시간)에서 부하시간의 비율을 산출한 것이다.

$$\text{시간가동률(\%)} = \frac{(\text{부하시간} - \text{정지시간}) * 100}{\text{부하시간}} \quad (1)$$

## 2.3 예지보전 시각화

### 2.3.1. 전동기 이상 유무 예측 모델

전동기의 전류 및 진동 데이터 파일 233만개를 사용하여 전동기의 이상유무를 예측할 수 있는 모델로 DNN, XGboost, Random forest를 사용했다.

### 2.3.2 설비 예지보전 모델

Seq2seq(sequence-to-sequence)는 RNN을 조합한 구조로, 언어 번역에 자주 쓰이는 모델이다. 구조는 encoder와 decoder를 이어놓은 구조로, encoder에서 시계열데이터를 압축한 context vector를 생성하고, 생성된 context vector를 decoder에서 결과를 출력한다.

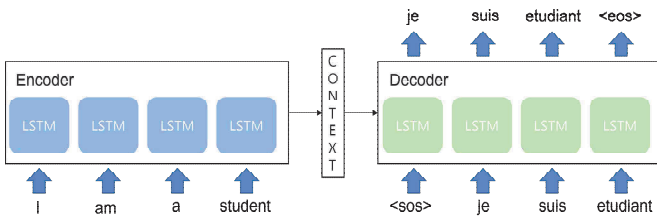


그림 1. seq2seq모델

## 3. 실험방법 및 결과

본 논문의 실험환경은 다음과 같다. 운영체제는 Linux 18.04, CPU는 Xeon Silver 4214, 메모리는 256GB, GPU는 RTX3090 1개를 사용하여 실험을 진행했다. 전동기의 이상유무와 예지보전 시점을 예측하기 위한 데이터는 다음과 같다. 이상유무를 예측하기 위해 사용된 데이터는 표 1과 같다. 총 233만개의 데이터에서 정상데이터는 약 109만개, 베어링불량은 36만개, 회전체불평형은 약 22만개, 축정렬 불량은 약 33만개, 벨트 느슨함은 약 31만개로 각각 구성되어 있다. 예지보전을 위해 사용된 데이터는 주기가 12,000인 데이터 1,000개를 사용한다.

이상유무 예측모델은 DNN, XGBoost와 Random forest를 사용하고, DNN은 64개의 은닉층으로 구성된 하나의 레이어를 사용하고 XGBoost와 Random Forest는 기본 파라미터로 학습했다. 표 2는 모델별로 성능 평가한 결과를 나타낸 것으로 DNN모델보다 XGBoost, Random forest가 99.99%로 0.2% 정도의 성능차이가 있었다. 그러나 XGboost와 Random forest는 정확도가 같으므로 혼돈행렬(Confusion Matrix)를 이용하여 결과를 나타낸다.

표 3~4는 XGBoost와 Random forest의 혼돈행렬 결과를 나타낸다. 세 모델 전부 정상 검출이 가장 낮은 성능을 띄는데, 이는 정상/불량을 검출하는 것보다, 정상인데 불량으로 검출하는 빈도가 다른 상태(베어링 불량, 회전체 불평형, 축 정렬 불량, 벨트 느슨함) 검출 때 보다 약 3~4배 높은 것을 알 수 있다. 즉, 정상으로 예측했지만, 불량 상태를 일정부분 포함된 것으로 해석된다.

예지보전 모델 학습을 위해 주기가 12,000인 데이터에서 윈도우사이즈를 100으로 설정하여 사전데이터와 사후 데이터로

각각 나누어 학습하고 예측하였으며, 성능평가는 표 6과 같다. 그림 2는 테스트 데이터를 사용하여 예측한 결과를 나타내고, 파랑색은 테스트 데이터, 주황색은 예측결과이다.

그림 3은 예지보전을 위한 모바일 앱 화면으로 실시간 확인, 공장라인, 설비현황, 설비현황 상세 순으로 총 4개의 화면으로 구성된다. 실시간 확인 화면에서는 서버로부터 받은 데이터를 그래프를 통해 실시간 확인이 가능하다. 화면의 아래는 이슈 관리자 창이 있으며, 설비데이터에 이상이 생기면 이슈 관리자 창에 이슈가 갱신되고 사용자는 이슈 관리자 창에서 상세정보를 확인할 수 있다. 공장라인 화면에서는 라인별 가동률을 확인할 수 있으며, 라인마다 안전도를 계산하여 라인구역에 색으로 표시를 하여 라인별 작업 환경 위험을 인지할 수 있다. 설비현황 화면에서는 라인내부에 구축되어있는 설비들의 상세정보를 확인할 수 있다.

표 1. 전동기 데이터 구성

	데이터 파일 개수
정상데이터	1,098,327
베어링불량	367,929
회전체불평형	223,502
축정렬 불량	330,651
벨트느슨함	312,763
합계	2,333,172

표 2. 모델별 이상유무 정확도 성능평가 결과(%)

	DNN	XGBoost	Random Forest
Accuracy	99.71	99.99	99.99
Precision	99.71	99.99	99.99
Recall	99.71	99.99	99.99
F1-Score	99.71	99.99	99.99

표 3. XGBoost 혼돈행렬

	정상	베어링 불량	회전체 불평형	축정렬 불량	벨트 느슨함
정상	92523	0	1	0	2
베어링 불량	1	27002	0	0	0
회전체 불평형	2	0	21733	0	0
축정렬 불량	1	0	0	28260	1
벨트 느슨함	0	1	0	0	33059

표 5. Random forest 혼돈행렬

	정상	베어링 불량	회전체 불량형	축정렬 불량	벨트 느슨함
정상	92520	0	1	2	2
베어링 불량	0	27002	0	0	0
회전체 불량형	0	0	21734	0	0
축정렬 불량	1	0	0	28259	0
벨트 느슨함	5	1	0	1	33058

표 6. Seq2seq 모델을 이용한 예지보전 성능평가

	R <sup>2</sup> score	MSE	MAE
seq2seq	0.85	0.005	5.76e-05

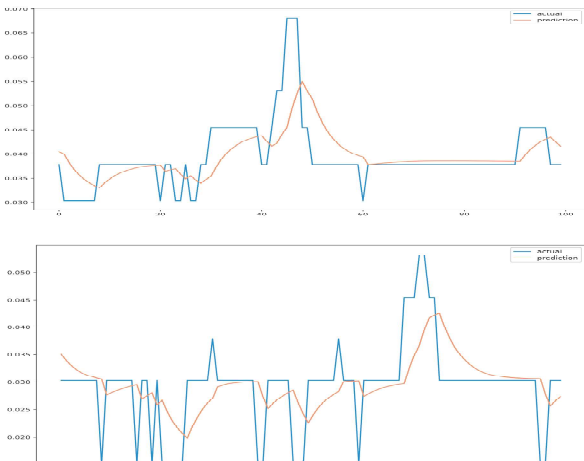


그림 2. Seq2seq모델 예측 그래프

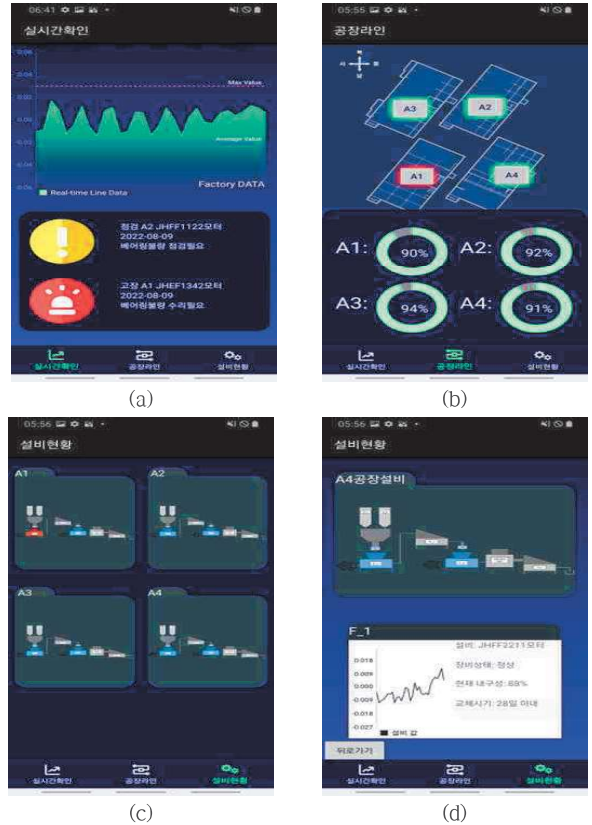


그림 3. 모바일 앱 화면

(a) 실시간 확인 화면, (b) 공장라인 화면, (c) 설비현황 화면 (d) 설비현황 상세 화면

#### 4. 결론

본 논문에서는 전동기 이상 유무를 예측하기 위해 AI-hub에서 구한 데이터를 사용하여 DNN, XGBoost와 Random forest 모델을 사용하였다. 각 모델을 이용하여 성능평가를 수행한 결과 DNN은 99.71%의 정확도를 XGBoost와 Random forest는 99.99%의 정확도를 보였다.

예지보전을 위해 사용된 데이터는 1000개를 임의로 뽑아 seq2seq모델의 학습에 사용하였다. 예지보전 모델의 성능은 MSE는 0.005, MAE는 5.76e-05, R<sup>2</sup>score는 0.85의 정확도를 보였다.

예지보전의 데이터 적합성을 높이기 위해 현장 작업자들이 활용할 수 있는 스마트폰 앱을 프로토타입 형태로 개발하였다. 제안하는 모바일 앱에서는 공장 모니터링이 가능하며, 이를 위해 실시간 확인, 공장라인, 설비현황, 설비현황 상세 화면들을 구현하여 현장 작업자들이 쉽게 활용할 수 있을 것으로 예상된다.

향후 연구에서는 실제 공장데이터를 활용하여 실시간으로 데이터수집과 시각화, 분석이 가능하고, 현장적합성이 뛰어난 예지보전 기술을 개발하고자 한다.

## 사 사 문 구

본 논문은 과학기술정보통신부 지원을 받아 경남 SW융합 클러스터2.0 특화산업 강화사업 과제로 수행한 연구 결과입니다. 본 과제(결과물)는 2021년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역 혁신 사업의 결과입니다(2021RIS-003).

## 참 고 문 헌

- [ 1 ] Chong HR, Bae KH, Lee MK, Kwon HM, & Hong SH. 2020. Quality strategy for building a smart factory in the fourth industrial revolution. J Korean Soc Qual Manag. 48(1):87-105.
- [2] Arturo Garza-Reyes, Jose, Steve Eldridge, Kevin D. Barber, and Horacio Soriano-Meier. "Overall equipment effectiveness (OEE) and process capability (PC) measures: a relationship analysis." International Journal of Quality & Reliability Management 27, no. 1 (2010): 48-62.
- [3] 유기성, 문영민. (2020). 스마트공장 설비진단 및 예지보전 기술의 발전 동향. 한국통신학회지(정보와 통신), 37(7), 36-42.
- [4] 이상원. (2022). 스마트 팩토리 품질요인이 스마트 팩토리 고도화 의도에 미치는 영향: 지각된 가치 및 스마트 팩토리 성과의 매개효과. 한국콘텐츠학회논문지, 22(5), 367-382.
- [5] 권대훈, 오창현. (2021). 스마트 팩토리를 위한 예지보전 기술. 한국정보통신학회 종합학술대회 논문집, 25(1), 172-174.



# 저장 매체에 따른 카프카 처리량 분석<sup>1)</sup>

이동영<sup>o</sup> 강동현

국립창원대학교 컴퓨터공학과

20193076@gs.cwnu.ac.kr, donghyun@changwon.ac.kr

## Analysis of Kafka Throughput by Storage Media

Dongyeong Lee<sup>o</sup> Donghyun Kang

Changwon National University, Dept. of Computer Engineering

### 요약

본 논문에서는 저장 매체가 아파치 카프카에 미치는 영향을 분석하기 위해 HDD와 SSD 환경에서 카프카의 성능을 테스트한다. 실험 환경은 카프카를 설치할 서버와 메시지를 송수신할 로컬 컴퓨터를 준비했다. 메시지의 생산자인 프로듀서와 소비자인 컨슈머를 1개에서 64개까지 늘려서 테스트를 진행했고 브로커의 개수 또한 1개에서 2개까지 늘려서 진행했다. HDD, SSD 환경 모두 프로듀서, 컨슈머의 수가 늘어날수록 처리량이 감소하는 경향은 동일했다. 하지만, HDD 환경의 경우 브로커의 수가 적을수록 프로듀서의 처리량에 크게 영향을 미쳤고 SSD 환경에서도 정도의 차이는 존재했지만 비슷한 경향을 보인다. 하지만, 컨슈머의 경우에는 HDD, SSD 모두 브로커의 수에는 처리량에 큰 영향을 미치지 않았다. 이로써 저장 매체는 카프카의 성능에 큰 영향을 미치지 않는다는 것을 알게되었다. 이러한 결과는 카프카 서버 구축에 도움이 될 것이라 생각한다.

## 1. 서론

최근 IoT 기술, SNS 등의 증가로 실시간 데이터의 생성 양이 증가하며 이를 저장 및 처리하기 위한 기술들이 주목받고 있다. 분산 저장을 위한 하둡[1], Ceph[2], 데이터 스트림을 처리하기 위한 아파치 스파크(Apache Spark)[3] 등이 이러한 기술들이다. 또한, 데이터를 안정적으로 전송 및 처리하기 위한 기술로 메시지 큐 시스템이 존재한다. 그러나 RabbitMQ[4]와 같은 기존의 메시지 큐 시스템은 단일 노드 기반으로 구성되어 현재 늘어나는 데이터의 양을 감당하기엔 한계가 존재한다. 아파치 카프카(apache kafka)[3]는 기존 메시지 큐 시스템과 달리 여러 개의 노드를 사용하는 분산 메시지 큐 시스템이다. 아파치 카프카는 메시지 수신자와 송신자 사이의 메시지를 로그의 형태로 분산해서 처리하여 기존의 시스템에 비해 처리량이 높다. 또한, 네트워크 통신 실패로 인한 데이터의 손실을 예방할 수 있는 장점이 있다. 본 논문에서는 카프카의 로그를 저장하는 저장 매체의 차이가 카프카의 처리량에 미치는 영향을 분석하기 위하여, HDD와 SSD 환경에서 프로듀서와 컨슈머의 수에 따라 변화하는 처리량을 분석하였다. 실험 결과, 카프카의 성능이 저장 매체에 영향을 받지 않는다는 것을 확인하였다.

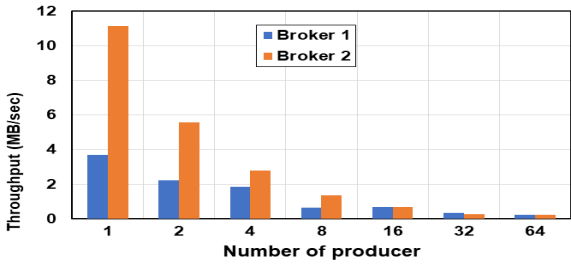
1) 이 논문은 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. NRF-2021R111A3047006).

본 논문의 구성은 다음과 같다. 2장에서 카프카에 대해 살펴보고, 3장에서는 저장 매체에 따른 카프카의 처리량을 관측하고 분석한다. 4장에서는 본 연구의 결론을 정리한다.

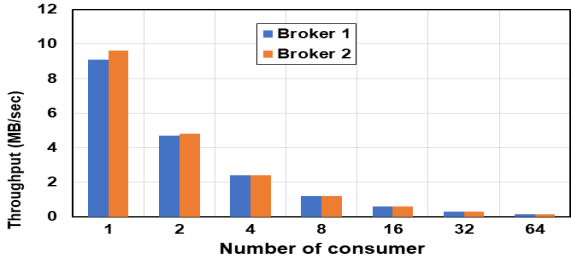
## 2. 배경지식

아파치 카프카는 실시간으로 데이터 스트림을 게시, 구독, 저장 및 처리할 수 있는 분산 데이터 스트리밍 플랫폼이다[3]. 카프카의 구조는 데이터를 만들어 내는 프로듀서(producer), 데이터를 소비하는 컨슈머(consumer), 클라이언트(client) 사이에서 데이터를 분산 저장하는 브로커(broker)로 구성된다[3]. 프로듀서가 브로커를 통해 메시지를 게시하게 되면 컨슈머에서 브로커를 구독하여 메시지를 가져가는 형태로 진행된다. 브로커는 전달받은 메시지를 파일 시스템에 로그 형태로 메시지를 저장하게 되고 토픽(topic)과 파티션(partition)을 통해 메시지를 관리한다. 토픽은 카프카에서 데이터를 관리할 때 기준이 되는 개념으로 데이터를 관리하는 그룹이며 1개 이상의 파티션으로 구성된다. 파티션은 실제로 데이터가 저장되는 공간이며 각 토픽 당 데이터를 분산 처리하는 단위이다.

## 3. 실험 분석



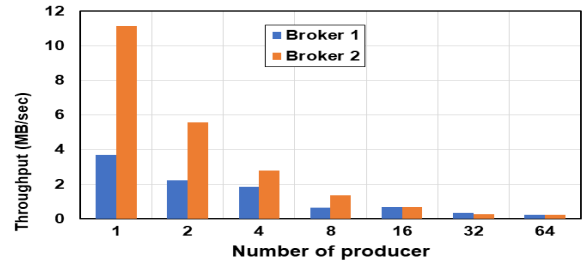
(a) 프로듀서의 처리량



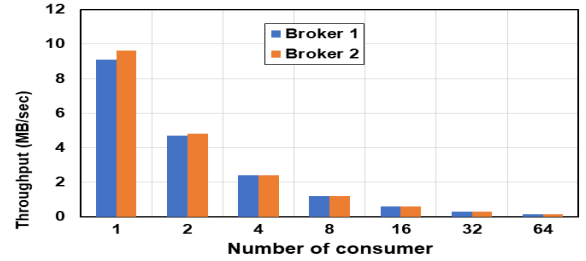
(b) 컨슈머의 처리량

그림 1. HDD 환경에서의 처리량

본 논문에서는 카프카의 로그가 저장되는 저장 매체의 종류에 따른 카프카의 메시지 처리량 차이를 분석하고자 하였다. 실험에는 카프카에서 기본으로 제공하는 실험용 셸 스크립트 파일인 kafka-producer-perf-test.sh와 kafka-consumer-perf-test.sh를 사용하였다[5]. 실험에 사용된 HDD는 500GB HDD인 WDC WD5000AAKX-2이며, SSD는 1TB SSD인 Samsung SSD 980 PRO 1TB이다. 토픽의 파티션의 개수는 브로커의 개수와 동일하게 설정하였다. 총 10만 개의 10kb 메시지를 토픽에 게시하여 실험하였으며, 메시지는 균등하게 분배되었다. 프로듀서와 컨슈머의 수는 1부터 64까지 2의 배수로 증가시키며 진행하였으며 브로커의 개수도 2까지 늘려서 실험을 진행하였다. 그림 1은 HDD를 저장소로 설정했을 때 프로듀서와 컨슈머의 실험 결과이다. 그림 1(a)에서 브로커의 수에 상관없이 프로듀서의 수가 늘어날수록 처리량이 감소하는 경향을 보인다. 또한 프로듀서의 개수가 8개 이전에서 브로커의 수에 따라 처리량의 차이가 크게 나타난다. 그림 1(b)에서도 브로커의 수에 상관없이 컨슈머의 수에 따라 처리량이 감소하는 경향을 보이나, 프로듀서와 달리 브로커의 수에 따른 처리량의 차이는 크게 드러나지 않았다. 그림 2은 SSD를 저장소로 설정했을 때의 프로듀서와 컨슈머의 실험 결과이다. 그림 2(a)는 그림 1(a)와 비슷한 경향을 보이나, 브로커의 수가 1개일 때, 그림 2(a)의 처리량이 그림 1(a)의 처리량 대비 2.13배가량 향상되었다. 그러나, 그 외의 경우에는 그림 1(a)와 그림 2(a)의 처리량이 유사한 것을 확인할 수 있다. 그림 2(b) 또한 HDD와 유사한 결과를 보인다. 이러한 실험 결과를 통해서



(a) 프로듀서의 처리량



(b) 컨슈머의 처리량

그림 2. SSD 환경에서의 처리량

카프카의 로그 저장에 사용된 저장 매체가 카프카의 처리량에 끼치는 영향은 크지 않다고 생각할 수 있다.

## 4 결론

본 논문에서는 저장 매체에 따른 카프카의 데이터 처리량의 차이를 살펴보았다. HDD 상황에서 카프카의 처리량은 프로듀서의 수가 적을수록 브로커의 수에 영향을 많이 받았지만, SSD는 그런 경향이 적게 나타났다. 하지만, 브로커의 수를 늘렸을 때의 처리량에는 차이가 거의 나타나지 않으므로 저장 매체가 아닌 브로커의 설정이 처리량에 더 큰 영향을 미치는 것으로 판단된다. 보통 브로커의 수를 단일로 설정하지는 않음으로 저장 매체는 카프카의 성능에 큰 영향을 미치지 않는다는 것을 알게 되었다. 이러한 결과는 카프카 서버를 구축할 때 저장 매체에 대한 효율성을 높일 수 있다고 생각한다.

## 참고문헌

- [1] Shvachko, K., Kuang, H., Radia, S. and Chansler, R., "The hadoop distributed file system," 2010 IEEE 26th symposium on mass storage systems and technologies, pp.1-10, 2010.
- [2] Weil, S.A., Brandt, S.A., Miller, E.L., Long, D.D. and Maltzahn, C., "Ceph: A scalable, high-performance distributed file system," Proceedings of the 7th symposium on Operating systems design and implementation, pp.307-320, 2006.
- [3] APACHE KAFKA, <https://kafka.apache.org/>
- [4] RabbitMQ, <https://www.rabbitmq.com/>

[5] kafka-\*-perf-test,  
<https://docs.cloudera.com/runtime/7.2.10/kafka-managing/topics/kafka-manage-cli-perf-test.html>

# 바코드를 이용한 사물인식 시스템 설계

권순성<sup>0</sup>, 이 셋, 강호윤, 김우현, 정중수  
안동대학교 정보통신공학과  
whrekrdl@naver.com, jschung@andong.ac.kr

## Design of Object Recognition System Using Barcode

Sun-Seong Gwon<sup>0</sup>, Set Lee, Ho-Yun Kang, Woo-Hyun Kim, Jung-Su Chung  
Dept. of Information-Communication Engineering, An-dong National University

### 요약

본 연구에서는 시각장애인들의 편의를 위해 바코드를 이용한 사물 인식 시스템을 설계하였다. 안드로이드 스튜디오 기반 자바 언어를 활용하여 서버와의 통신을 구현하였고, 서버를 구축하여 톰캣을 구동한 PC에 Spring Framework 기반의 서버-API(Application Programming Interface) 간 통신을 구현하였다. 이 기술을 이용하여 실시간으로 바코드값을 읽고, 그 바코드에 대한 상품 정보를 알 수 있게 구현 및 검증하였다

## 1 서론

시각장애인들은 비장애인과 같이 자유롭게 물건을 사며 경제활동을 할 권리가 있다. 하지만 2022년 현재까지도 국내에서는 그러한 시각장애인들의 권리가 제대로 보호되고 있지 않은 실정이다. 시각장애인은 물건을 선택하고 구매하는 상황에 있어 한계점이 존재하는 반면, 구매하는 과정에 가장 큰 도움이 되는 점자가 표기된 물품은 극히 일부에 불과하다[1]. 식품 점자 표시가 법적 의무 사항이 아니기 때문이다. 최근 조사 결과 시각장애인들의 식품 점자 표시에 대한 불만족도는 71.9%에 달했다[2]. 이러한 상품 구매 과정에서 시각장애인들의 불편함을 해소하고, 독립성을 증대시키고자 하는 것이 본 연구의 목표이다.

바코드를 이용한 사물인식 방식은 YOLO나 Tensorflow 등을 통한 이미지를 학습시켜 구현하는 방법과 다르게 새로운 대상이 추가될 시에도 별도의 이미지 라벨링과 학습 과정, 애플리케이션의 업데이트 등이 필요하지 않다는 장점을 지닌다[3]. 하지만 바코드가 없는 면의 사물 인식에 있어 인식이 불가능하다는 한계점을 가진다.

본 연구에서는 Google의 ML KIT, Jsoup 등의 오픈소스 라이브러리, 안드로이드 스튜디오, 톰캣 등을 활용하여 사물인식 시스템을 설계하였다. 효율적인 연구 진행을 위해 바코드 인식의 범위는 국내 EAN-13 형식의 상품 바코드 데이터를 제공하는 코리안넷 유통물류진흥원과 소비자 24에 등록된 상품들로 한정하였다[4][5].

## 2 사물인식 시스템 구조

본 논문에서 제안된 사물인식 시스템의 구조는 그림 1과 같다. 구현한 애플리케이션 시스템 구조는 총 6가지 과정으로 구성된다. 그림은 좌측부터 각각 상품, 스마트폰, 서버, API로 구성되어 있다. 사용 방법 및 처리 과정은 다음 순서와 같다.

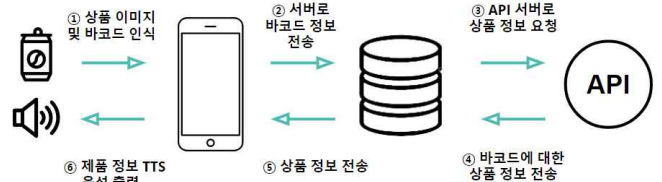


그림 1. 시스템 구조

- ① 사용자가 애플리케이션을 켜고 상품을 적정거리(10~20 cm)에 위치시킨다. 카메라를 통해 들어온 정보를 바탕으로 바코드의 타입을 식별하고 타입에 맞는 데이터 처리를 진행한다.
- ② 바코드가 식별되면 프레임별로 리스트에 바코드값을 할당한다. 할당된 리스트 3개의 값이 모두 일치하면 정상적으로 잘 읽힌 데이터라고 판단, 서버로 바코드 값을 전송한다. 만약 3개 중 하나라도 바코드 정보가 다를 시, 제대로 읽히지 않은 정보라고 판단, 사용자에게 TTS(Text to speech, 텍스트 음성 변환)로 상황을 알린다.
- ③ 서버는 애플리케이션으로부터 전달받은 바코드 정보를 바코드 정보 검색 사이트에 질의한다.
- ④ 서버는 결과로 나온 값을 크롤링(웹 페이지에서 추출, 파싱하여 필요한 데이터를 가져오는 행위)한다. 크롤링된 정보에는 제품의 회사명과 제품명이 포함된다.
- ⑤ 서버는 크롤링한 정보를 HTML의 body 항목에 담아 클라이언트에게 전송한다.
- ⑥ 애플리케이션의 백그라운드에서 서버 정보를 크롤링(정보 수집) 및 파싱(분석 및 변환)함으로써 상품정보를 사용자에게 TTS(Text to Speech)로 전달하게 된다.

### 3. 상품인식 시스템 설계

#### 3.1 설계 과정

애플리케이션이 켜졌을 때, 설정 창으로부터 돌아올 때, 다이얼로그 창에서 확인을 눌렀을 때 실시간 영상 캡처 및 분석을 진행한다. QR Code, EAN-13, ISBN 등 바코드의 타입에 따라 case 구문을 통해 처리 과정을 달리하였다. 본 연구에서 가장 중점을 둔 것은 서론에서 언급한 목표인 시각장애인들의 상품 선택에 있어서의 불편함 해소 및 편의성과 독립성 증대이므로, 상품 바코드 데이터의 표준인 EAN-13를 중점으로 진행하였다. 만약 캡처 및 분석 과정에서 상품정보, 그 중에서도 값이 값이 13인 EAN-13이 들어올 경우 그 값을 리스트에 추가한다. 리스트의 사이즈가 3이 될 경우, 카메라를 통한 바코드 캡처 및 분석 활동을 일시 중지시킨다. 만약 그렇게 연속적으로 찍힌 리스트의 값이 모두 일치할 경우에는 정상적으로 찍힌 바코드 값이라고 판단, 서버로 바코드의 정보를 보낸다[6]. 만약 일치하지 않는 경우에는 초점이 안 맞거나 비스듬한 경우이므로 사용자에게 재촬영 안내 TTS를 통해 상황을 전달한다. 일치 여부를 검사한 이후에는 다음 캡처를 위해 바코드 값들이 저장되어 있던 리스트를 초기화한다.

그림2는 서버가 애플리케이션으로부터 바코드 정보를 받았을 때의 서버-API 동작 흐름도이다. 애플리케이션이 HTTP의 GET 방식으로 바코드 정보를 서버에 전달하면 서버는 그림 2의 과정을 거쳐 전달받은 바코드 정보에 대해 바코드 정보 API 서버에 질의를 요청하게 된다.

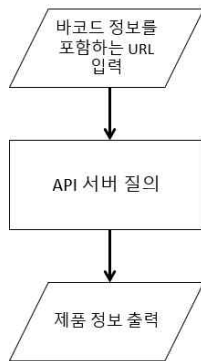


그림 2. 서버와 API 동작 흐름

바코드 정보 API 서버로부터 전달받은 결과에는 제품에 대한 회사명과 제품명이 담겨있다. 서버는 HTML 본문에 회사명과 제품명 정보만을 담아 요청했던 애플리케이션에 전달한다.

그림 3은 애플리케이션이 서버로부터 제품 정보를 받아오기 위해 Jsoup 라이브러리를 활용한 HTTP 크롤링 & 파싱을 활용하여 정상적으로 인식이 이뤄진 결과이다. 서버가 API에 요청한 바코드 정보가 API가 갖고 있지 않는 정보일 때에는 서버가 not found를 표시한다. 애플리케이션은 서버로부터 받은 값이 not found일 때, 사용자에게 API 에 없는 상품임을 TTS를 통해 알린다. 이후 다시 바코드 캡처 및 분석을 시작한다. API 서버가 바코드 정보에 대한 상품 데이터를 갖고 있는 경우에는 정상적

으로 상품명을 받아오게 된다. 이 경우 TTS로 사용자에게 상품명을 알려주고, 다이얼로그를 띄워 다시 인식을 시작할 것인지, 설정 창으로 이동할 것인지를 선택하게끔 한다.



그림 3. 바코드 인식 전후

설정 창에서는 TTS 속도 조절, 바코드 정보를 받아오는 서버 주소의 변경, 애플리케이션의 정보 노출 등의 부가적 기능들을 구현하였다.

#### 3.2 검증

기존에 활용 했었던 Zxing 오픈소스 라이브러리[7]는 회전 된 바코드의 인식이 잘 되지 않고, 굴곡이 있을 시 정확한 인식이 어려우며, 인식 속도도 느리다는 문제점이 존재했다. 이러한 문제점을 해결하기 위해 회전 시에도 문제 없이 사용 가능한 Camera X Core 라이브러리를 사용하는 ML Kit 바코드 스캔 방식으로 바꾸어 구현하였다.

그림 4는 물건에 회전 각도를 주어 바코드 인식을 시도한 경우이다. 이와 같은 상황에도 정확하게 값을 읽어올 수 있었다.

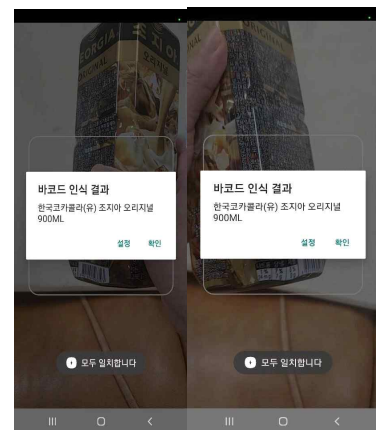


그림 4. 기울어진 바코드

그림 5(좌 - Zxing, 우 - ML kit)는 바코드를 회전시켜 바코드 인식을 시도한 경우이다. 기존에 오픈소스 바코드 인식 라이브러린 Zxing 라이브러리를 이용하여 바코드 인식 및 캡처를 구현한 경우에는 정상적으로 인식

되지 않았으나, ML Kit 라이브러리로 인식 방법을 바꿈으로써 이와 같은 회전된 상황에서도 정확하고 빠르게 인식이 가능하게 되었다. 그림 6은 뒤집어서 바코드 인식을 시도한 경우이다. 이 경우에도 정상적으로 값을 읽어 올 수 있다.

시각장애인의 경우 바코드의 위치를 파악하고, 스마트폰 카메라에 평행하게 바코드를 일치 시키는데에는 불가능에 가깝다. 하지만 평행하게 일치시키지 않고도 이와 같이 정상적으로 바코드를 인식할 수 있으므로 시각장애인에게 더 높은 활용성과 편의성을 제공한다.

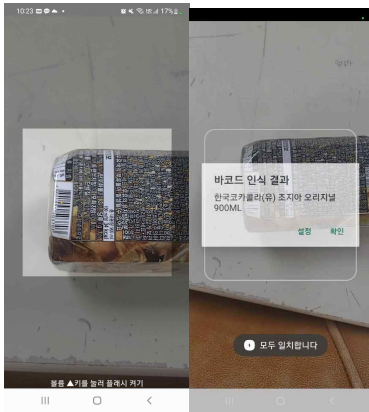


그림 5. 회전된 바코드



그림 6. 뒤집어진 바코드

실제 애플리케이션 사용에 있어서 상품 정보를 불러오기까지 걸리는 시간은 상당히 중요하다. 카메라를 통한 첫 번째 리스트 캡처부터 서버로 상품 정보를 요청하여 수신하기까지의 속도를 Java에서 지원하는 시스템 함수인 System.currentTimeMillis()를 이용해 측정하였다.

해상도를 1280\*720으로 설정하고 리스트의 크기는 3으로 설정하여 바코드를 측정한 결과는 그림 7과 8과 같다. 측정한 결과값은 요청한 서버에 따라 그 처리 속도가 크게 나뉘었다.

그림 7은 코리아넷 대한상공회의소 유통물류진흥원 GTIN에 상품정보를 요청하는 형태로, 총 300회의 테스트를 진행하였다. 차트의 가로 축이 횟수, 세로 축이 소요되는 시간(ms)을 의미한다. 최소 지연 시 253ms, 최대 지연 시 520ms, 평균 301.51ms로 측정되었다.

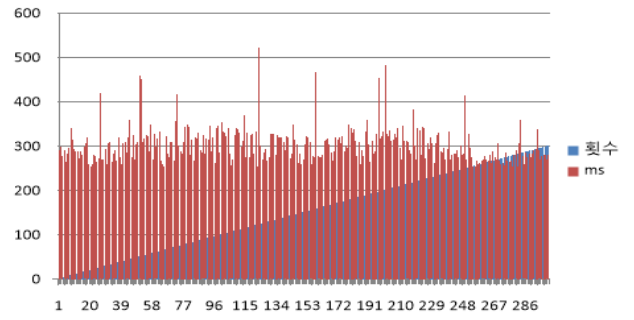


그림 7. 코리아넷 정보 요청 지연속도 분석

그림 8은 소비자 24 물품정보검색에 상품정보를 요청하는 형태로, 총 80회의 테스트를 진행하였다. 최소 지연 시 4,890ms, 최대 지연 시 6,313ms, 평균 5657.75ms로 측정되었다.

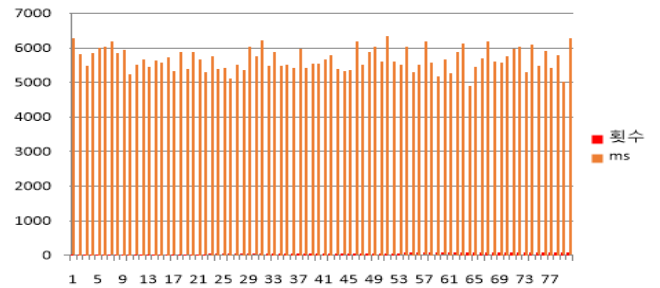


그림 8. 소비자 24 정보 요청 지연속도 분석

그림 7과 8은 각각 Jsoup 라이브러리와 Selenium 라이브러리는 다른 크롤링 방식을 사용한다. 이에 따라 지연속도 결과값의 큰 차이가 나타났다.

소요 시간의 차이로 실제 사용에 있어서는 코리아넷에 정보 요청을 하는 방법이 더 적합한 것으로 판단된다. 그림 8의 경우, 서버에 데이터베이스를 구축하여 데이터 크롤링이 아닌 데이터베이스에서 불러오는 방식으로 시간 지연 문제를 개선할 수 있다.

#### 4 결론 및 향후 연구 방향

본 연구에서는 안드로이드 기반 애플리케이션을 이용해 스마트폰의 카메라를 통해 바코드 인식 및 서버와의 통신으로 상품에 대한 정보를 실시간으로 얻을 수 있는 시스템을 구현하였다.

이 바코드 사물인식 시스템을 통해 시각장애인들이 주변의 도움 없이 독립적으로 물품 구매가 가능하게 됨으로써 사회적으로 긍정적인 영향을 끼칠 것이라고 기대한다.

본 논문에서는 구현하지 않았으나, EAN-13과 QR Code, ISBN 형식의 바코드뿐만 아니라 더 많은 형식을 지원하고, 바코드 이미지 커스텀 데이터 세트를 학습시켜 바코드의 거리 정보, 회전 정도, 반사 정도 등을 파악하여 사용자에게 상황에 적합한 조절 방법을 알려줌으로써 사용성을 증가시킬 수 있을 것으로 보인다.

#### 참고 문헌

- [1] 이난초, “음료 34개 중 ‘점자’ 제품명 있는 건 단 하나뿐”, 국민일보, <http://news.kmib.co.kr/article/view.asp?arcid=0015398836&code=61171811>, 2021.1.
- [2] 한국소비자원, “시각장애인 식품 점자표기\_실태조사\_보고자료”, 2022. 9.
- [3] 한상혁, 박다수, 임채민, 정지운, “시각장애인을 위한 편의점 제품 인식 애플리케이션”, 한국정보처리학회 추계학술발표대회(28권 2호) 2021.02.
- [4] 대한상공회의소 유통물류진흥원 바코드 검색 <http://www.koreannet.or.kr/home/hpisSrchGtin.gs1?>
- [5] 소비자 24에 등록되어 있는 바코드 수는 물품 및 가공식품이 3,273,140건, 도서는 1,764,047 건이다. 2022.11.14
- [6] [http://서버IP:8080/barcode/숫자13자리\(EAN-13\)](http://서버IP:8080/barcode/숫자13자리(EAN-13)), [http://서버IP:8080/barcode2/숫자13자리\(EAN-13\)또는 ISBN](http://서버IP:8080/barcode2/숫자13자리(EAN-13)또는 ISBN)
- [7] Kulyukin, Vladimir, and Aliasgar Kutiyawala. “From ShopTalk to ShopMobile: vision-based barcode scanning with mobile phones for independent blind grocery shopping.” Proceedings of the 2010 Rehabilitation Engineering and Assistive Technology Society of North America Conference (RESNA 2010), Las Vegas, NV. Vol. 703. 2010.

# 스마트밴드를 이용한 위험 심박수 감지 모바일앱

강동현, 정근모, 송하주

부경대학교 컴퓨터공학부

k.d.hyeon97@gmail.com, rmsah0612@gmail.com, hajoosong@pknu.ac.kr

## A Mobile App for Detecting Dangerous Heart Rate using A Smart Band

Donghyeon Kang Keunmo Jeong Hajoo Song

Division of Computer Engineering

Pukyong National University

### 요약

일상 생활에서 자신의 심박수를 잘 알고 있는 것은 질병 관리에 도움이 되며, 심박수에 이상이 발생할 경우 미리 병원을 방문하여 심혈관계 질환에 적절한 대처를 할 수 있다. 특히, 급성 심혈관계 질환의 경우 증상 발생 후 수 분 이내에 사망에 이를 수 있는 위험한 질환이므로 심박수가 정상적인 수치를 넘어 갔을 때 빠른 시간 내에 보호자나 주치의, 혹은 119 안전신고센터에 연락해야 한다. 웨어러블 기기의 발전으로 심박수를 자동으로 감지하여 위험한 수치에 도달했을 경우 즉시 보호자 또는 119에 신고하는 기능을 탑재하고 있다. 하지만 기기의 가격이 높아 보급에는 한계가 있다. 본 연구에서는 상대적으로 저렴한 스마트밴드를 사용하여 사용자 심박수를 측정하고 위험 심박의 발생 여부를 관찰하는 모바일앱을 개발하였다. 제안하는 모바일앱은 구글 Fitness API를 사용하여 사용자의 심박 정보를 주기적으로 모니터링하고 이상 현상이 예상되면 사용자에게 경고한다. 또한 급격한 상황이 발생한 것으로 추정되면 지정된 연락처로 위험 신호를 전달하여 도움을 받을 수 있도록 한다.

### 1. 서론

심장질환은 사망률이 높은 질환이다. 통계청 보도자료 [1]에 따르면, 2021년 주요 만성질환 중(암 제외) 사망률이 가장 높은 것은 심장질환으로 인구 10만 명당 61.5명으로 나타났다. 심장질환은 안정 시 심박수가 비슷한 연령대의 평균 심박수와 비교했을 때 높거나 낮을 때 발생할 확률이 증가하므로 평소 자신의 심박수를 숙지하고 있을 필요가 있다. 평소 자신의 심박수를 알고 있는 것이 중요하며, 일상생활에서 심박수를 측정하는 것은 질병 관리에 도움이 될 수 있다고 알려져 있다[2]. 따라서 평소 심박수를 주기적으로 확인하는 것은 매우 중요하며, 심박수에 이상이 발생하였을 경우 미리 병원을 방문하여 심혈관계 질환을 조기에 예방할 수 있다.

### 2. 관련 연구

웨어러블 기기의 발전으로 착용하는 것만으로도 사용자의 심박수를 측정할 수 있다. 대표적인 기기로는 삼성전자의 갤럭시워치 시리즈와 애플의 애플워치 시리즈가 있다. 해당 제품들은 심박수 측정, 심전도 측정, 낙상감지기능 등을 갖추어 사용자가 위험한 순간에 이를 감지하여 빠른 시간내에 119 안전신고센터 등으로 신고할 수 있는 기능을 갖추었다. 하지만, 해당 제품들은 판매가격이 높게 책정되어 있어 구매하기 부담스러울 수 있다. 대체품으로 샤오미의 미밴드 시리즈와 같은 저렴한 스마트밴드가 있으나, 높은 심박수에 대한 경고만 할 뿐, 보호자나 119 안전신고센터 등으로 도움을 요청하는 기능은 갖추고 있지 않다.

저렴한 스마트밴드의 경우 대부분 Google Fitness와 연

동되어있다. Google Fitness는 연동된 웨어러블 기기가 측정된 데이터를 가공할 수 있도록 Google Fit API를 제공한다. 이 API를 활용하여 사용자의 심박수를 손쉽게 가져올 수 있다. 본 연구에서 사용한 “미밴드 6”는 저렴하면서도 안정적으로 심박 정보를 제공하는 기기이다.

이에 본 논문에서는 샤오미의 미밴드 시리즈와 같은 저렴한 스마트밴드를 활용하여 사용자의 심박수를 주기적으로 측정된 후 위험한 심박수를 판단하여 보호자 및 119에 도움을 요청하는 기능을 제공하는 모바일앱을 제안한다.

### 2. 스마트 밴드를 사용한 위험 심박 감지 모바일 앱의 설계와 구현

#### 2.1 위험 심박의 감지

애플리케이션을 제작하기에 앞서 위험한 심박수를 정의할 필요가 있다. 119급대원 현장응급조치 표준지침 [3]에 따르면 맥박이 분당 100회 이상 뛰는 것을 빈맥, 맥박이 분당 60회 미만인 경우 서맥으로 정의하며, 분당 심박수가 150회 이상 또는 50회 미만이 되면 이상 증상이 발생한다고 알려져 있다. 지침상의 내용을 검증하기 위해 실제로 스마트밴드를 착용하고 일주일간 심박수를 측정하였고, 그 결과를 반영하여 그림 1과 같이 작동 방식을 구성하였다. 심박수는 개인별 차이가 있으므로 위험한 심박수 범위는 자신에게 맞는 심박수 범위로 설정할 수 있도록 한다.

#### 2.2 애플리케이션 제작

모바일 앱의 실행과 함께 주의사항을 안내하고 초기화





기능을 사용할 수 있도록 하였다. 해당 애플리케이션은 Google Fitness와 연동하여 다양한 웨어러블 기기에서 해당 기능을 사용할 수 있다. 이를 통해 사용자는 기기 구매 시 비용을 절감할 수 있으며, 다양한 기기의 선택을 할 수 있다. 추가적인 효과로는 심장질환을 조기에 발견하고 예방할 수 있다는 점이다. 심장질환은 평소 안정 시 심박수가 높거나 낮을 때 발생할 확률이 증가하기 때문에 평소 자신의 심박수 변화를 주기적으로 확인해주는 것이 좋다. 사용자는 해당 애플리케이션이 보낸 경고 메시지를 통해 본인의 현재 상태를 인지할 수 있다. 본인의 상태를 인지한 사용자는 병원에 방문하여 진단을 받을 수 있고 이를 통해 심장질환을 조기에 예방할 수 있다.

### 참 고 문 헌

- [1] 통계청, 2021년 사망원인통계 결과, 국가통계포털, 2022년 9월 22일 수정, 2022년 11월 12일 접속, [https://www.kostat.go.kr/portal/korea/kor\\_nw/1/6/2/index.board?bmode=read&bSeq=&aSeq=420715&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=](https://www.kostat.go.kr/portal/korea/kor_nw/1/6/2/index.board?bmode=read&bSeq=&aSeq=420715&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=)
- [2] 박동혁 and 전용관. “건강예측 요인으로서의 안정시 심박수의 가치와 의미“ 운동과학 29, no.1 (2020) : 24-33.
- [3] 소방청, 119구급대원 현장응급처치 표준지침, 2021년 9월 16일 수정, 2022년 11월 12일 접속, [https://www.nfa.go.kr/nfa/publicrelations/legalinformation/archives/?boardId=bbs\\_000000000000018&mode=view&cntId=26&category=&pageIdx=](https://www.nfa.go.kr/nfa/publicrelations/legalinformation/archives/?boardId=bbs_000000000000018&mode=view&cntId=26&category=&pageIdx=)

# 양상블 기법을 이용한 PCB 불량 검출 방법에 대한 연구

최성혁<sup>1</sup>, 배완기<sup>1</sup>, 유병주<sup>1</sup>, 김동현<sup>1</sup>, 정설영<sup>1</sup>, 이문용<sup>2</sup>

<sup>1</sup>경북대학교 컴퓨터학부, <sup>2</sup>JS SYSTEM Co., Ltd

che980529@naver.com<sup>1</sup>, shiningcityuponahill7408@gmail.com<sup>1</sup>, qudwn8712@gmail.com<sup>1</sup>,

water\_beetle@naver.com<sup>1</sup>, snowflower@knu.ac.kr<sup>1</sup>, my@js-system.co.kr<sup>2</sup>

## PCB Defect Detection based on Ensemble Method

Choi, Seong Hyeok<sup>1</sup>, Bae, Wan Gi<sup>1</sup>, Yu, Byeong Ju<sup>1</sup>, Kim, Dong Hyun<sup>1</sup>, Jung, Seol Yeoung<sup>1</sup>, Lee, Moon Yong<sup>2</sup>

<sup>1</sup>College of Computer Science and Engineering, KyungPook National University,

<sup>2</sup>JS SYSTEM Co., Ltd

### 요 약

본 논문에서는 AFVI 장비에서 검출된 PCB 불량 의심 이미지들을 CNN 기반 알고리즘에 학습시킨 후 양상블 기법 중 소프트 보팅 방법을 이용해 불량을 검출하고 불량 종류를 분류하는 기법을 제안한다. 모델은 ResNet101, InceptionV3, VGG16 모델을 사용하였다. 본 연구 결과를 활용하면, 현재 PCB 제조 공정 과정 중 불량 검출 효율성을 높이는 데 도움을 줄 수 있을 것이다.

## 1 서론

한국 전자 통신 연구소에 따르면 현재 일부 대기업을 제외하고는 독일, 일본 등의 선진국에 비해 한국의 제조업 혁신도는 월등히 낮은 편이다. 글로벌 대기업을 경우, 제조업에 자동화 시스템, 생산관리 시스템 등을 확보해 나가고 있지만, 중소기업의 경우에는 여전히 제조업 혁신에 많은 어려움을 갖고 있다.[1]

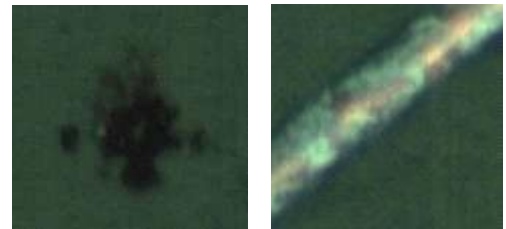
스마트 팩토리 사업을 통해 AI 기반 알고리즘을 적용하여 새로운 기술을 확보해야 할 필요성이 대두되었기에, 현재 PCB 제조 공정 검사 중 AI 기반 머신 비전을 통해 불량 검출 효율성을 높이기 위한 방법을 제시하였다.

본 논문에서는 AFVI 장비로부터 얻은 PCB 불량 의심 이미지들을 불량 요인에 따라 효과적으로 분류하기 위해, 합성곱 신경망(Convolutional Neural Network, CNN) 기반 모델들을 활용한 양상블 기법(Ensemble Method, EM)의 PCB 불량 분류 방법을 제안한다.

## 2 본론

### 2-1 데이터 전처리

본 연구에 사용된 이미지 데이터는 AFVI로부터 얻어진 불량 의심 이미지들을 10개의 불량 클래스로 임의로 분류하고 라벨링하였다. 총 2127장의 512 x 512 해상도 이미지들을 사용하였다. ImageNet의 224 x 224 해상도의 이미지들로 사전 학습된 CNN 모델들을 사용하기 때문에, 2127장의 이미지들을 224 x 224의 해상도로 변환하였다.



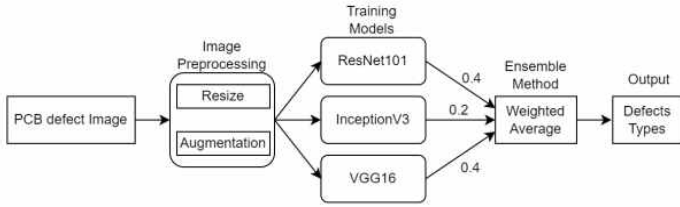
[그림 1] - 불량 PCB 이미지 예시

그림 1은 데이터 해상도 변환 과정을 마친 후의 PCB 불량 이미지이다.

모델의 과적합을 방지하고 일반화 성능을 극대화하기 위해 데이터 증강 기법을 적용하였다. 적용된 기법들은 확대, 좌우 전환, 상하 전환이고, 검증에 사용된 이미지의 수는 전체 수집 데이터 수의 20%이다.

### 2-2 시스템 구성

그림 2는 전체 시스템 구조이다. 그림 2와 같이 AFVI로부터 얻은 2127장의 이미지들을 2-1에서 언급한 Resize와 Augmentation 방법을 적용한 후, CNN 기반 모델들인 ResNet101[2], InceptionV3[3], 그리고 VGG16[4]에 입력으로 넣어준다. 각각의 모델에서 얻은 클래스 별 확률에 가중치를 곱한 후 양상블 기법 중 하나인 소프트 보팅 기법[5]을 사용해 최종 불량 클래스를 판정한다.



[그림 2] - 시스템 모델

### 2-3. 모델 학습 환경

Augmentation과 Resize가 적용된 2127장의 이미지들을 ImageNet에서 사전 훈련된 ResNet101, InceptionV3, VGG16 모델에 입력으로 넣은 후, 과대 적합을 방지하기 위해 Patience가 10로 설정된 Early Stopping을 적용하여 100 epoch만큼 훈련을 진행하였다. 3개 모델과 앙상블 기법 모두 RTX 3080 GPU로 구성된 시스템에서 학습, 검증, 그리고 테스트를 진행하였다.

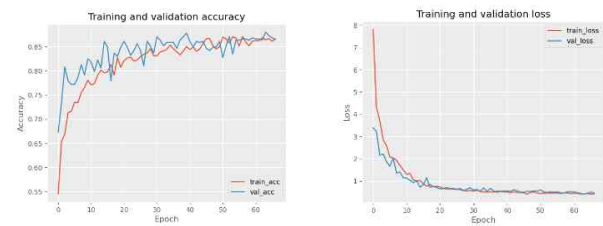
학습에 사용된 이미지는 1706장, 검증에 사용된 이미지는 421장이다.

[표 1] - 학습 파라미터 설정 값

Optimization Function	Learning Rate	Batch Size	Loss	Activation
Adam	1e-4	32	Cross entropy	ReLU

각각의 모델의 학습 파라미터들을 위 표 1의 내용과 같이 값을 설정하였다.

### 3. 학습 결과

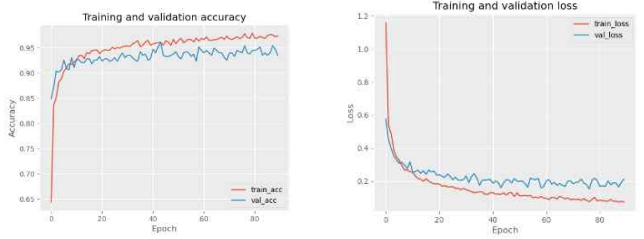


[그림 3 - a, b] InceptionV3의 accuracy와 loss

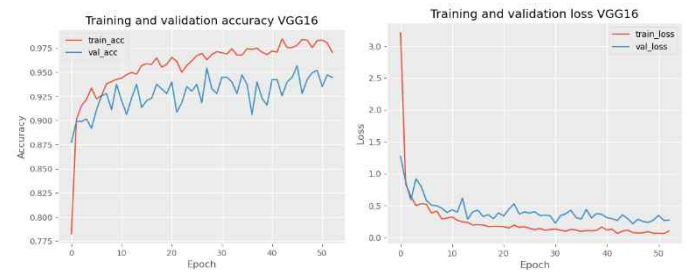
InceptionV3 모델은 67 epoch에 학습이 종료되었다.

그림 3-a는 InceptionV3의 accuracy를 나타낸 그래프로 86.5%의 validation accuracy를 보였다.

그림 3-b는 InceptionV3의 loss를 나타낸 그래프로 training loss는 0.4770, validation loss는 0.5256이 나왔다. 세 모델 중 validation accuracy가 가장 낮게 나왔다.



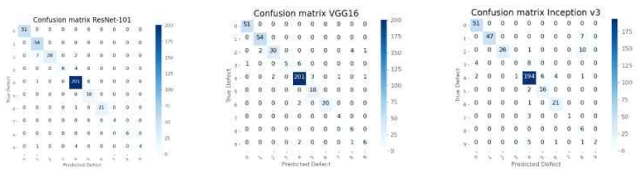
[그림 4 - a, b] ResNet101의 accuracy와 loss  
ResNet101 모델은 90 epoch에 학습이 종료되었고, 그림 4-a는 ResNet101의 accuracy를 나타낸 그래프로 93.8%의 validation accuracy를 보였다. 그림 4-b는 ResNet101의 training loss는 0.0739, validation loss는 0.2110이 나왔다.



[그림 5 - a, b] VGG16의 accuracy와 loss

VGG16 모델은 53 epoch에 학습이 종료되었고, 그림 5 - a는 VGG16의 accuracy를 나타낸 그래프로 93.8%의 validation accuracy를 보였다.

그림 5 - b는 VGG16의 training loss는 0.0993, validation loss는 0.2692이 나왔다.

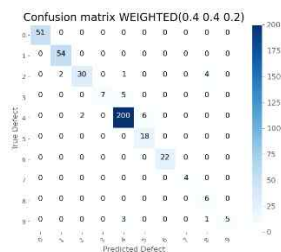


[그림 6 - a] [그림 6 - b] [그림 6 - c]

그림 6 - a, b, c는 각각 ResNet101, VGG16, Inception v3의 Confusion Matrix이다.

모든 모델이 클래스 3과 9에 대해서 분류 성능이 떨어진다.

InceptionV3는 클래스 7에 대해서도 분류 성능이 떨어져 가장 낮은 분류 성능을 보였다.



[그림 7] - Ensemble Method의 Confusion Matrix

Ensemble Method는 기존의 각 모델과 동일하게 클래스 3, 9의 분류 성능이 떨어지지만 기존의 개별 모델에 비해 향상된 정확도를 보였다.

[표 2] - Ensemble Method의 분류성능 평가지표

	Precision	Recall	F1-Score	Support
<b>0</b>	1.000	1.000	1.000	51
<b>1</b>	0.964	1.000	0.982	54
<b>2</b>	0.938	0.811	0.870	37
<b>3</b>	1.000	0.583	0.737	12
<b>4</b>	0.957	0.962	0.959	208
<b>5</b>	0.750	1.000	0.857	18
<b>6</b>	1.000	1.000	1.000	22
<b>7</b>	1.000	1.000	1.000	4
<b>8</b>	0.545	1.000	0.706	6
<b>9</b>	1.000	0.556	0.714	9
<b>Accuracy</b>	0.943			421
<b>Macro Avg</b>	0.915	0.891	0.882	421
<b>Weighted Avg</b>	0.951	0.943	0.942	421

표 2는 가중치를 사용한 소프트 보팅을 적용한 앙상블 메소드의 분류 성능 평가지표이다.

세 개의 모델을 앙상블한 최종 결과가 94.3%로 기존의 각 모델 중 가장 높은 정확도인 93.8%보다 더 높은 분류 정확도를 보여준다.

## 4 결론

본 논문에서는 PCB 제조 공정 중 불량 검출의 효율성을 높이기 위해, 3개의 CNN 기반 모델인 ResNet101, InceptionV3, VGG16을 가중치를 사용한 소프트 보팅에 접목시켜 AFVI로부터 얻은 PCB 불량 의심 이미지를 불량 원인에 따라 분류해보았다. 본 연구에서는 충분한 양의 이미지셋이 아닌 2127장의 작은 이미지셋과 PCB 불량 의심 이미지 라벨링 과정에서 전문인력의 도움을 받지 못해 시행착오를 겪었다. 그럼에도 준수한 분류 성능을 보였고, 추후 불량 클래스 라벨링의 개선, 더 큰 이미지셋 확보, 그리고 파라미터 튜닝을 통해 실제 공정에서도 적용될 수 있을 것으로 기대된다.

### 사사문구

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음”

### 참고 문헌

- [1] 최선미. “중소 제조업, 지능화로 혁신의 스피드를 높여라.” 기술정책연구본부. (2019): 1-8
- [2] He, Kaiming, et al. “Deep residual learning for image recognition.” Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE. pp.770-778. 2016.
- [3] Xia, Xiaoling, Cui Xu, and Bing Nan. “Inception-v3 for flower classification.” 2017 2nd international conference on image, vision and computing (ICIVC). IEEE. pp.783-787. 2017.
- [4] Simonyan, Karen, and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition.” The 3rd International Conference on Learning Representation (ICLR2015). arXiv preprint arXiv:1409.1556.
- [5] Wang, Haishen, et al. “Soft-voting clustering ensemble.” International Workshop on Multiple Classifier Systems. MCS. Lecture Notes in Computer Science, vol 7872. pp.307-318. 2013.

# IoT 기반 화재 정보 감지 시스템의 설계

이셋<sup>0</sup>, 정중수, 유은상\*\*

\*안동대학교 정보통신공학과, \*\*(주) 라성에너지

e-mail : dltp99@gmail.com, [jschung@andong.ac.kr](mailto:jschung@andong.ac.kr)

## System Design for Fire Sensing Information based on IoT System

Set Lee<sup>0</sup>, Joong-Soo Chung, Eun-Sang Yu\*\*

\*Dept of Information-communications Engineering, An-Dong National University.

### 요 약

본 논문에서는 IoT 기반의 화재 정보 감지 시스템의 중요 요소가 되는 화재 발생 위치 및 시간과 더불어 경도, 위도 등 좌표 처리기능과 그에 관계되는 데이터 처리 구조를 설계 및 구현하였다. 하드웨어는 드론에 라즈베리파이를 접속하였다. 경도, 위도 등 좌표 처리를 위한 GPS 모듈, 구축될 서버와 인터넷 통신을 위한 LTE 라우터를 라즈베리파이와 접속하였다. 라즈베리파이 측 클라이언트 프로그램은 파이썬 언어, 서버는 Flask 프레임워크로 설계되었다. 라즈베리파이에서 서버로 전달되는 데이터 구조 처리는 HTTP POST 방식으로 제시되었다.

키워드 : 서버, 화재 인식, 센서, IoT, 텐서플로우

Server, Fire Detection, Sensor, IoT, Tensorflow

해 기술하였다.

### 1. 서론

대한민국에서 산불은 많이 발생하는 재난 중 하나로, 국토의 70% 이상이 산지를 이루고 있어 산불의 빈도가 높다. 그리고 자주 발생하는 산불로 인해 받는 피해는 적지 않으니 대한민국에서 산불에 관한 관심은 높다.

본 논문에서는 화재 감지 시스템과 IoT 시스템의 융합을 기반으로 만든 드론을 이용한 산불 화재 감지 시스템의 구조를 제시한다. 또한 시스템과 서버간 데이터 송수신 처리 기법으로 HTTP의 POST 방식을 제시하였다. 구성된 시스템은 드론에 부착된 GPS 모듈과 라즈베리파이 간 ttyUSB 통신을 통해 좌표 값을 수신한다. 그리고 카메라 모듈을 이용해 촬영한 이미지를 프로그램에서 압축하게 된다. 라즈베리파이의 마지막 과정으로 좌표 값과 이미지를 웹 형식으로 구축된 시스템 서버에 전달한다. 서버는 전달받은 이미지에 대해 텐서플로우로 학습된 모델을 이용해 화재 여부를 판단하고 결과를 드론에 전달하고 좌표 값과 이미지, 화재 여부 등을 데이터베이스에 저장한다.

본 논문의 구성으로 2장에서는 화재 감지 IoT의 시스템 구조, 3장에서는 화재 감지 시스템 설계를 서술하였다. 4장에서는 본 연구의 결론 및 향후 연구 계획에 관

### 2. 화재 감지 IoT의 시스템 구조

화재 감지 IoT의 시스템 구조는 그림 1과 같다. 드론에 라즈베리파이와 LTE 모듈, GPS 모듈, 카메라 모듈과 함께 부착되어 있고, 라즈베리파이의 프로그램에서 카메라 모듈을 이용해 사진을 촬영하고, GPS 모듈을 이용해 위도, 경도, 고도의 값을 수신한다. 촬영된 사진과 좌표 값, 라즈베리파이 ID, 현재 시각을 서버로 전송하는데 이때 HTTP 프로토콜의 POST 방식을 사용한다. 서버는 수신받은 라즈베리파이 ID를 이용해 어떤 드론에서 보낸 사진인지 구분하고, 수신받은 이미지를 텐서플로우 화재 인식 모델을 활용하여 화재 여부를 판별하고 이에 대한 응답을 HTML의 body 항목에 담아 라즈베리파이에게 응답한다. 그리고 서버의 데이터베이스 내에 수신받은 드론 ID, 좌표 값, 이미지가 저장된다. 라즈베리파이는 서버로부터 전달받은 HTML 내 body 항목을 통해 화재 여부를 식별하고 만약 화재 상황이라면 장착된 스피커를 이용해 사이렌 소리를 출력하여 주변에 화재 상황을 알리게 된다.

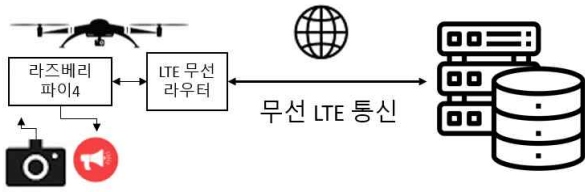


그림 30. 화재 감지 시스템의 구조

### 3. 화재 감지 시스템 설계

IoT 시스템과 융합된 화재 감지 시스템은 라즈베리파이의 프로그램이 서버로 이미지를 전송하고, 서버는 화재 여부 데이터를 응답하는 것이 주된 목적이다. 따라서 라즈베리파이와 서버 사이의 원활한 통신 및 정보제공이 필수이다. 이 장에서는 IoT 기반 데이터 전송 설계, 서버의 이미지를 이용한 화재 식별 설계로 분류하여 서술하였다.

#### 3.1 IoT 기반 데이터 전송 설계

드넓은 산의 환경과 드론에 부착되어 계속해서 위치가 변경되는 라즈베리파이와 Wi-Fi는 통신에 활용하기 적절하지 않아 LTE 모듈을 장착해 LTE 환경에서 통신을 수행한다. 화재 감지 시스템 설계 시 드론의 관리를 위해 드론마다 고유의 ID를 사용한다.

라즈베리파이의 프로그램은 그림 2와 같은 구조로 동작하는데, 먼저 카메라 모듈을 이용해 카메라 촬영 명령을 내린다. 촬영되어 이미지 형태로 저장된 파일을 Imagemagick 프로그램을 이용하여 파일의 크기를 줄인다. 다만 과도하게 줄이게 되면 이미지의 화질이 낮아져 화재 식별의 오답률이 높아질 가능성이 있어 본 연구에서는 40%의 크기로 줄이게 설계하였다. 라즈베리파이에 사용되는 LTE 모듈은 LTE Cat.M1으로, 저전력 광대역 전송 기술을 활용하기 때문에 월별 사용 한도가 정해져 있고, 전송 속도 또한 300Kbps이므로 서버로 전송하는 이미지의 용량은 적을수록 유리하다. 그리고 GPS 모듈을 이용해 위도, 경도, 고도 값을 수신한다. 서버로 HTTP의 POST 방식으로 드론 번호(라즈베리파이 ID), 크기 조정된 이미지, GPS 값, 촬영 시간을 전송한다. 이후 서버는 HTML body 항목에 화재 여부를 담아 응답한다. 이를 통해 라즈베리파이는 화재의 유무를 식별하게 된다. 만약 응답이 “Fire”라면 부착된 스피커 모듈을 이용해 사이렌, 대피 안내 방송을 출력하도록 구성하였다.

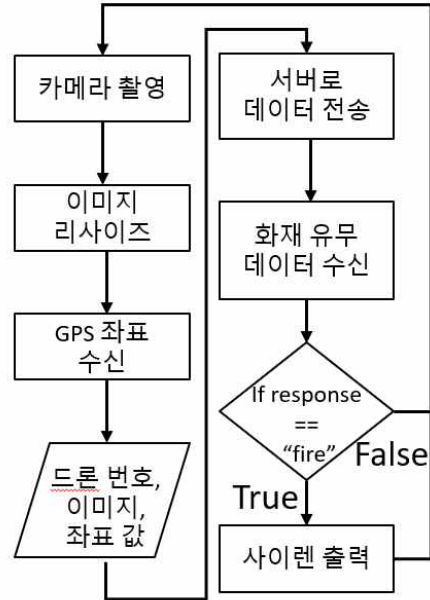


그림 31. 라즈베리파이 프로그램 구조

#### 3.2 서버의 화재 식별 설계

서버는 파이썬의 웹서버 프레임워크 Flask를 이용해 구축하였으며, 라즈베리파이로부터 데이터 수신을 위한 HTTP의 POST 항목과 데이터베이스를 조회 및 관리하기 위한 관계 페이지가 존재한다. 서버는 라즈베리파이가 전송한 드론 번호, 이미지, 경도, 위도, 고도 값의 데이터 처리를 수행한다. 드론 번호를 이용해 어떤 드론이 전송한 데이터인지 구분하고, 이 드론 번호는 추후 이미지 파일을 구분하여 저장하는 데 사용된다. 서버는 이미지를 통해 화재 유무를 식별할 수 있도록 구성된 텐서플로우 Keras 모델을 이용해 수신받은 이미지에 대해 화재 유무를 추정하고 결과를 라즈베리파이에 HTML의 body 항목으로 전송한다. 이후 드론 번호, 이미지, 경도, 위도, 고도 값, 촬영 시간, 화재 여부를 데이터베이스에 저장한다. 그리고 관계 화면을 웹상에 구현하여 어디서든 접속하여 관제할 수 있도록 하였다.

### 4. 검증 결과

화재 검증을 위한 테스트 이미지로, 무료 이미지 공유 사이트인 pixabay에서 검색한 그림 3의 산불 이미지를 이용하여 화재 상황을 촬영한 상황을 가정하고 검증을 진행한다.



그림 32. 테스트 산불 이미지[8]

라즈베리파이에 해당 이미지를 내려받고, 실험을 진행하였다. 라즈베리파이에서 작성한 프로그램을 실행하였고, 결과는 fire로 출력되었다. 즉, 서버가 전송받은 사진을 화재 상황으로 인식하고 라즈베리파이에 전송한 것이다.

```
pi@raspberrypi:~/20181093 $ python post_test.py
fire
pi@raspberrypi:~/20181093 $
```

그림 33. 라즈베리파이 프로그램 실행 및 결과 화면

image 파일이 저장된 경로, 위도, 경도, 고도, 촬영시간이 저장된 것을 확인할 수 있다. 실제 이미지는 서버 측에 저장되는데, 이 저장된 경로가 데이터베이스에 저장되는 것이다. 이미지를 DB에 저장하는 것은 base64 와 같은 인코딩을 이용한 직렬화 과정을 거치면서 파일의 크기가 더 커지게 된다. 반면 이미지의 경로를 저장하는 방식은 원하는 이미지를 서버 측에서 바로 열어볼 수 있다는 장점 때문에 이미지는 서버의 하드디스크에 저장하고 이미지의 경로를 데이터베이스에 저장하는 방식을 제시한다.

id	image	latitude	longitude	altitude	time	fire
1	./static/images/1/2022-11-16 20-40-21.jpg	36.5435000000000000	128.7959000000000000	170.1000000000000000	2022-11-16 20:40:21	fire

그림 34. 데이터베이스 조회 화면

그림 6은 웹상에서 구현한 관제 화면이다. 드론 ID, 이미지, 촬영 시간, 좌표, 화재 여부를 확인할 수 있다.

Fire Detection System

Drone ID	Image	Time	Location	Fire
1		2022-11-16 20:40:21	36.5435000000000000 128.7959000000000000 170.1000000000000000	fire

그림 35. 관제 화면

## 5. 결론 및 향후 연구

본 논문에서는 화재 감지 기반 IoT 시스템의 동작과 데이터 구조 및 서버와의 인터페이스 설계를 제시하였다. 또한 화재 정보가 포함될 수 있는 이미지 데이터와 위도, 경도, 고도 등 좌표 값을 수신하여, 피해 상황을 신속히 파악하고 피해를 줄이는 데 집중하였다. 향후 본 논문에서 제시된 설계를 바탕으로 관제 화면의 시각화 개선, 사용자 요구에 의한 데이터베이스구축과 웹서버 설계가 요구된다. 이러한 환경에 적합한 화재 탐지 IoT 시스템의 데이터셋을 이용해 요구 환경에 최적화된 Keras 훈련 모델을 제작하여 응용 범위와 정확도를 높여야 할 것이다.

## 참고문헌

- [1] 정중수 외, “농업용 IoT 기반의 관수 시스템 설계”, 한국 정보과학, UCWIT 2015, 2015.
- [2] 정중수 외, “IoT 기반의 교육 시스템 설계 및 구현”, UCWIT 2016, 2016.
- [3] 김규성 외, “교육용 IoT 기반 시스템의 서버 설계”, UCWIT 2016, 2016.
- [4] 김규성 외, “교육용 IoT 시스템의 앱 관리 설계”, UCWIT 2016, 2016.
- [5] 이현동 외, “균형 기반 의료 IoT 시스템 설계 및 구현”, UCWIT 2017, 2017.
- [6] 이현동 외, “균형 기반 의료 IoT 시스템 설계 및 구현”, UCWIT 2018, 2018.
- [7] 이현동 외, “균형 기반 의료 IoT 시스템의 웹 응용 시스템 설계”, UCWIT 2018, 2018.
- [8] pixabay, 산불 이미지 <https://pixabay.com/images/id-5458901/>



# Streamlit 프레임워크를 이용한 TTS 및 SVS 기반 웹 서비스 개발

김예원 이충현 김성재 김인중  
한동대학교 전산전자공학부

21800132@handong.ac.kr , chlee@handong.ac.kr, sjkim@handong.ac.kr, ijkim@handong.ac.kr

## Developing Web Service based on TTS and SVS Using Streamlit Framework

Yewon Kim, Choonghyeon Lee, Sungjae Kim, and Injung Kim  
School of Computer Science and Electrical Engineering, Handong Global University

### 요약

TTS(Text-to-Speech)와 SVS(Singing-Voice-Synthesis)는 음성처리 분야의 핵심 연구 주제 중 하나이다. TTS/SVS 기술을 실용적 서비스에 응용하기 위해서는 GPU 서버에서 동작하는 TTS와 SVS 엔진을 웹과 연동해야 한다. 본 연구에서는 파이썬 기반 오픈소스 웹 프레임워크 Streamlit을 이용하여 웹기반 인터페이스를 통해 텍스트, 화자, 및 감정을 입력받은 후 GPU 서버의 TTS/SVS 엔진을 원격으로 호출해 음성을 합성하고 웹 인터페이스를 통해 출력하는 웹 서비스를 개발하였다. 개발 과정에서 Streamlit이 제공하는 디자인에는 색상, 글꼴의 종류 및 크기 변경이 제한적으로만 허용된다는 제약이 있었으나, HTML과 CSS를 이용해 이러한 제약을 극복하고 시각적으로 보기 좋은 인터랙티브 웹 인터페이스를 개발하였다.

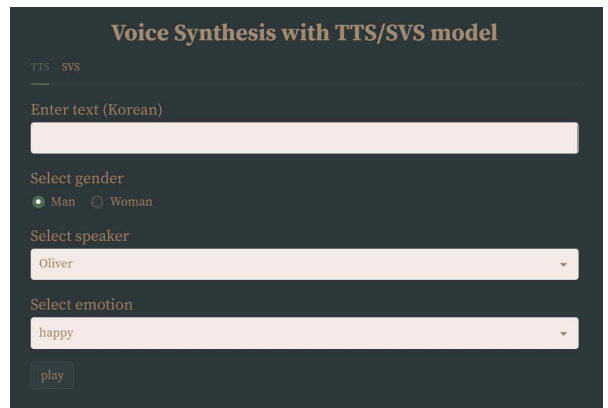
## 1. 서론

TTS(Text-to-Speech)는 텍스트를 입력받아 음성을 합성하는 기술이며 SVS(Singing Voice Synthesis)는 텍스트 외에 추가적으로 각 음표의 음정과 길이가 포함된 악보 정보를 입력받아 노래음성을 합성하는 기술이다. 최근 수 년 동안 딥러닝 기반 TTS/SVS 기술은 크게 발전하여 성별, 화자, 감정 등 다양한 스타일 속성을 컨트롤 하면서도 고품질의 음성을 합성하는 TTS/SVS 모델들을 구현할 수 있게 되었다[1,2,3,4]. 이러한 TTS/SVS 기술을 실용적인 서비스에 응용하기 위해서는 파이토치나 텐서플로우 등 딥러닝 프레임워크 상에서 개발된 TTS/SVS 모델을 웹 기반 인터페이스에 연동해야 한다.

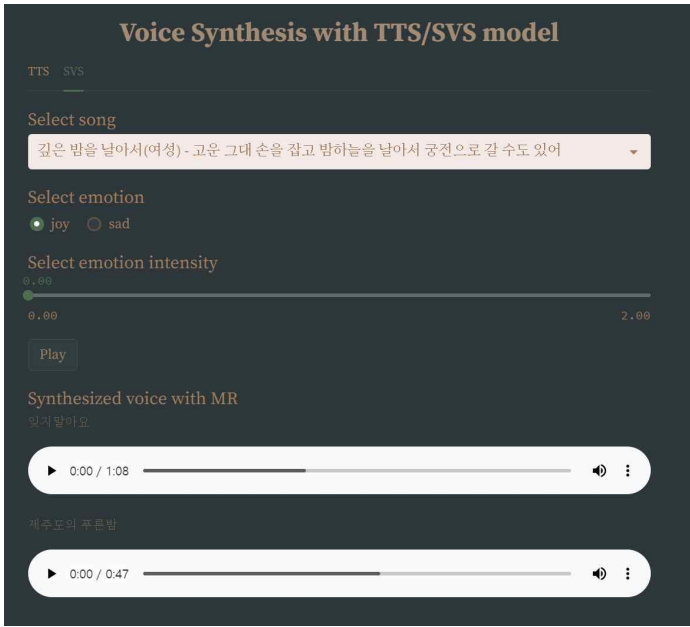
본 연구에서는 웹 인터페이스를 통해 사용자로부터 텍스트, 또는 노래(곡)과 복수의 스타일 속성을 입력받아 GPU 서버에서 동작하는 TTS/SVS 엔진을 연동해 일반음성, 또는 노래음성을 들려주는 웹 서비스를 개발하였다. 일반음성 합성을 위해서는 그림 1과 같이 음성으로 변환할 텍스트와 함께 성별, 화자ID, 감정ID 등 세 가지 스타일 속성을 지정할 수 있는 인터페이스를 제공한다. 노래음성 합성을 위해서는 그림 2와 같이 노래와 감정의 종류(기쁨 또는 슬픔)와 감정의 강도 등 두 가지 스타일 속성을 지정할 수 있는 인터페이스를 제공한다. 노래음성 합성을 위해서는 가사, 음정, 음길이가 포함된 악보정보가 요구된다. 그러나, 웹 화면에서 악보정보를 입력하는 것은 불편하기 때문에 사전 등록된 곡 중 하나를 선택하면 미리 입력된 악보 정보를 이용해 노래음성을 합성하도록 구현하였다.

## 2. 음성 합성을 위한 백엔드 및 프론트엔드 개발 2.1 시스템 구성

전체 시스템은 그림 1과 같이 웹 서버에서 동작하는 프론트엔드 인터페이스와 GPU 서버에서 동작하는 백엔드 TTS/SVS 엔진으로 구성된다. 프론트엔드에서는 웹 인터페이스를 통해 사용자로부터 음성으로 합성할 텍스트, 또는 악보정보를 입력받는다. 추가적으로, 성별, 화자(또는 가수), 감정, 감정 강도 등 복수의 스타일 속성을 함께 입력받아 GPU 서버에서 동작하는 백엔드 TTS, 또는 SVS 엔진으로 전달한다. TTS, 또는 SVS 엔진은 프론트엔드에서 전달받은 텍스트(또는 악보정보) 및 스타일 속성들을 전달받아 이를 반영한 일반음성, 또는 노래음성을 합성하여 프론트엔드로 전달하며, 프론트엔드는 이를 전달받아 사용자에게 들려준다.



<그림1 일반음성합성(TTS) 인터페이스 화면>



<그림2 노래음성합성(SVS) 인터페이스 화면>

## 2.2. 백엔드: TTS/SVS 엔진

본 서비스의 개발을 위해서는 다양한 스타일 속성을 반영해 음성을 합성할 수 있는 TTS 및 SVS 엔진이 요구된다. TTS 엔진으로는 화자와 감정을 지정해 일반음성을 합성할 수 있는 UniTTS[1]를 사용하였으며, SVS 엔진으로는 화자, 감정의 종류 및 강도를 지정해 노래음성을 합성할 수 있는 U-Singer[2]를 사용하였다. 두 모델은 모든 음성 스타일 요소들을 통합된 임베딩 공간에서 표현한다. 스타일 속성 간 조건부 확률을 학습한 후 복수의 스타일 속성을 공동확률에 의해 표현하기 때문에 서로 중첩되어 분리하기 어려운 스타일 속성 간 간섭을 최소화 하면서 고품질의 음성을 합성한다.

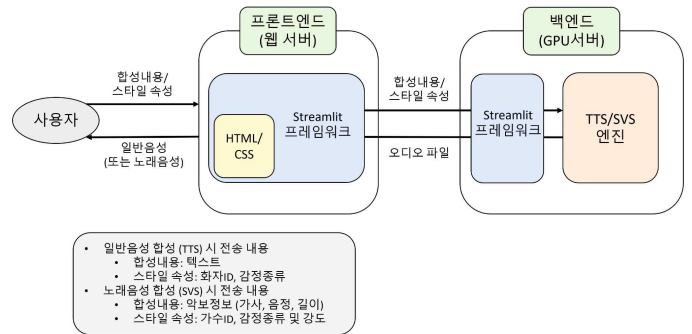
## 2.3. 백엔드와 프론트엔드 간 연동

본 연구에서 백엔드와 프론트엔드를 연결하기 위해 세 가지 프레임워크를 검토하였다. 첫째는 파이썬 기반 웹 프레임워크 Flask이다[5]. Flask는 복수의 웹페이지를 제작하는데 유용하지만, HTML 및 Flask 자체에 대한 많은 지식을 요구한다. 둘째는 HTML과 파이썬 간 연동을 위한 프레임워크 PyScript이다[6]. HTML에서 구현하기 어려운 기능들을 파이썬 코드를 이용해 쉽게 구현할 수 있다는 장점이 있으나, 사용 가능한 라이브러리의 수가 제한될 뿐 아니라 TTS/SVS 모델의 파라미터를 로드하기 위한 기능이 제공되지 않는다. 마지막으로 세 번째 검토한 파이썬 기반 웹 프레임워크는 Streamlit이다[7]. 파일 접근이 가능해 TTS/SVS 모델의 파라미터를 로드할 수 있으며 다른 파이썬 프로그램과 연동도 가능하다. 특히 사용자로부터 TTS 및 SVS의 입력 정보를 받거나 웹 페이지에서 오디오를 재생하기 위한 API가 제공되기 때문에 본 연구에서 웹과 TTS/SVS엔진을 연동하기 위한 프레임워크로 Streamlit을 선택하였다.

사용자가 웹 인터페이스를 통해 TTS 또는 SVS를 선택하면 기능에 따라 합성내용(텍스트 또는 곡)과 스타일 속

성들을 입력할 수 있는 화면으로 전환된다. 사용자가 합성내용과 스타일 속성들을 입력하고 실행버튼을 누르면 입력 정보를 백엔드로 전송한다. 'st.text\_input', 'st.radio', 'st.selectbox' 등 Streamlit이 제공하는 위젯들을 이용해 인터페이스를 구현하면 Streamlit 프레임워크가 입력정보를 자동으로 백엔드로 전달한다.

백엔드에서는 프론트엔드에서 전달받은 정보를 TTS/SVS 엔진에 입력하기 위한 포맷으로 변환한다. 예를 들어, 곡명, 화자/가수의 ID, 감정의 종류 및 강도는 각각 고유 번호로 변환된다. 그 후 TTS, 또는 SVS 엔진을 호출해 음성을 합성해 파일로 저장한다. 이와 같이 백엔드 서버에 저장된 음성 파일은 Streamlit의 'st.audio' 위젯을 통해 프론트엔드로 연결되어 플레이된다. 음성이 플레이된 이후에는 서버의 음성 파일을 삭제한다. 백엔드에서 TTS/SVS 엔진이 음성을 합성하는 동안 프론트엔드에서는 로딩 표시를 통해 음성합성이 진행 중임을 사용자에게 알린다.



<그림3. 시스템 구조>

## 2.4. 프론트엔드 구현

### 2.4.1. TTS 및 SVS의 웹기반 사용자 인터페이스 개발

Streamlit, HTML 그리고 CSS를 이용해 사용자가 원하는 내용과 스타일을 지정할 수 있는 웹 기반 인터페이스를 제작하였다. 일반음성을 합성할 때는 그림 1과 같이 사용자로부터 텍스트, 성별, 화자 및 감정 종류를 입력받는다. 각각의 정보를 입력 받기 위해 Streamlit이 제공하는 'st.text\_input', 'st.radio', 'st.selectbox' 위젯을 사용하였다. 노래음성을 합성할 때는 그림 2와 같이 곡, 감정의 종류 및 강도를 입력받는다. 일반적인 사용자가 웹 환경에서 악보 입력하는 것은 불편하기 때문에 미리 제작된 악보 중에서 선택하도록 하였으며, 감정 종류와 강도를 지정하면 SVS 엔진이 이를 반영해 노래음성을 합성할 수 있게 하였다. 악보, 감정 종류를 입력 받을 때는 'st.selectbox', 'st.radio' 위젯을 사용하였다. 특히, 감정 강도 같이 연속적으로 증가 또는 감소할 수 있는 정보를 입력 받기 위해 'st.slider' 위젯을 사용하였다.

### 2.4.2. Streamlit 디자인적 제약 및 극복방안

Streamlit은 프론트엔드와 백엔드 간 연동에 용이하지만 색상과 폰트 설정은 제한적으로만 허용된다. 설정 파

일 ‘config.toml’ 을 통해 웹 화면의 색상과 폰트를 지정할 수 있다. 그러나, 이렇게 지정한 색상 및 폰트는 모든 컨트롤에 대해 일괄적으로 적용되며 각 컨트롤 별 색상과 폰트를 지정하는 것은 허용되지 않는다. 이러한 제약은 시각적으로 보기 좋은 웹 화면을 구현하는데 장애가 된다.

본 연구에서는 ‘st.markdown’ 함수를 이용해 HTML을 직접 이용해 웹 화면을 구성함으로써 이러한 제약을 극복하였다. ‘st.markdown’ 함수는 마크다운[8] 형식의 문서를 스트링 형식의 ‘body’ 파라미터로 입력받아 화면에 표시한다. 이때 ‘unsafe\_allow\_HTML’ 파라미터를 true로 지정하면 HTML로 작성된 문서도 화면에 표시할 수 있다. 웹 화면의 색상 및 폰트는 ‘body’ 파라미터에 CSS를 포함시킴으로써 설정할 수 있다. st.markdown에 작성된 CSS는 해당 st.markdown에 개별적으로 적용되는 것이 아니라, 다른 st.markdown과 유기적으로 동작한다. 또한 CSS를 기존의 HTML문서에서처럼 식별자에 따라 개별적으로 적용할 수 있다. CSS의 모든 디자인 요소가 python 파일 내부에서 적용하는 것이 가능해지기 때문에, Streamlit의 디자인적 제약을 극복할 수 있다. 따라서 본 논문에서는 파이썬 파일에 HTML 코드를 삽입하여 st.markdown함수를 통해 연동함으로써 다양한 색상과 폰트를 활용한 웹 화면을 구현하였다.

[8] Markdown 홈페이지 <https://www.markdownguide.org>

### 3. 결론

본 논문에서는 Streamlit 프레임워크를 이용하여 TTS 및 SVS를 이용한 웹 서비스를 개발하였다. TTS 엔진으로는 UniTTS, SVS 엔진으로는 U-Singer를 사용하였다. TTS 및 SVS모델과 연동하기 위해 Streamlit 프레임워크를 사용하였으며, HTML과 CSS를 사용해 Streamlit의 디자인적 제약을 극복하였다.

### 사사의 글

본 연구는 과학기술정보통신부의 소프트웨어중심대학 지원사업 (2017-0-00130)의 지원을 받아 수행하였음.

### 참고 문헌

- [1] Minsu Kang, Sungjae Kim and Injung Kim, UniTTS: Residual Learning of Unified Embedding Space for Speech Style Control, arxiv, 2021
- [2] Sungjae Kim, Kihyun Na, Choonghyeon Lee, Jehyeon An and Injung Kim, U-Singer: Multi-Singer Singing Voice Synthesizer that Controls Emotional Intensity, arxiv, 2022
- [3] Lee, Juheon, Hyeong-Seok Choi, and Kyogu Lee. Expressive Singing Synthesis Using Local Style Token and Dual-path Pitch Encoder. arXiv preprint arXiv:2204.03249 (2022).
- [4] Kim, Jaehyeon, Jungil Kong, and Juhee Son. Conditional variational autoencoder with adversarial learning for end-to-end text-to-speech. International Conference on Machine Learning. PMLR, 2021.
- [5] Flask 홈페이지 <https://flask.palletsprojects.com>
- [6] PyScript 홈페이지 <https://pyscript.net>
- [7] Streamlit 홈페이지 <https://streamlit.io>

# 반려동물 돌봄을 위한 모바일 앱 설계 및 구현<sup>1)</sup>

김도선<sup>01</sup>, 정민욱<sup>1</sup>, 박희원<sup>1</sup>, 정원준<sup>2</sup>, 김동균<sup>1\*</sup>, 임호만<sup>3</sup>

<sup>1</sup>경북대학교 컴퓨터학부

<sup>2</sup>경북대학교 건설방재공학부

<sup>3</sup>(주)에스엠에스

ehtjsv2@knu.ac.kr, alsdnrdl01@knu.ac.kr, phwon7@knu.ac.kr, museoul@knu.ac.kr,

dongkyun@knu.ac.kr, mjhm@smsinfo.co.kr

## Design and Implementation of Mobile App for Pet Care

Doseon Kim<sup>01</sup>, Minuk Jung<sup>1</sup>, Heewon Park<sup>1</sup>, Wonjun Jung<sup>2</sup>, Dongkyun Kim<sup>1</sup>, Homan Lim<sup>3</sup>

<sup>1</sup>Department of Computer Science, Kyungpook National University

<sup>2</sup>School of Disaster Prevention and Environmental Engineering

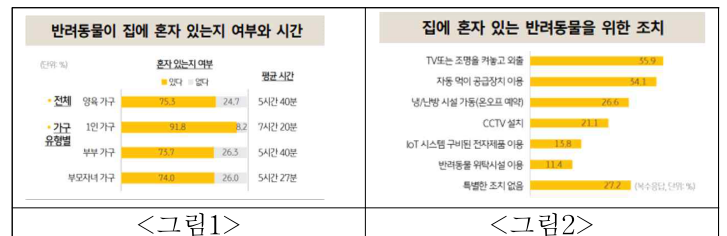
<sup>3</sup>Security Management Solution

### 요 약

반려동물 키우는 인구가 늘어남에 따라 각자의 반려동물 특성에 맞추어 돌봐주는 펫시터라는 직업이 새로 생겨났다. 이러한 수요의 증가에 따라 펫시터와 고객을 연결해주는 어플리케이션이 많아졌다. 그러나 기존의 앱의 경우, 제공하는 서비스의 가격이 낮지 않아 서비스를 이용하는 고객들에게 금전적인 부담이 있다는 문제점이 존재한다. 이러한 문제점을 해결하기 위하여 본 앱은 비슷한 상황에 있는 반려인들이 모여 보다 저렴하고 편리하게 본인의 반려동물을 돌볼 수 있게 하기 위해 기존 돌봄 서비스에 추가로 ‘함께 돌봄’ 서비스를 제공하여 고객의 금전적인 부담을 덜 수 있게 하였다. 또한 실시간 스트리밍 및 업로드 영상 확인, 실시간 위치 공유 기능으로 고객이 펫시터의 서비스 진행 상황을 바로 파악할 수 있으며, 필터링 기능을 통해 고객이 원하는 조건에 맞는 펫시터를 보다 손쉽게 검색해서 사용할 수 있다.

## 1 서 론

자신의 곁을 지켜주는 동물을 이제는 ‘애완동물’ 이 아닌 ‘반려동물’ 이라고 부르는 시대가 다가왔다. 2021년, 한국 반려동물 보고서(KB 연구보고서)에 따르면, 최근 반려동물을 키우는 가구가 매우 많아져 전체 가구의 29.7%인 604만 가구에서 반려동물을 키우고 있다고 한다. 그러나 “반려동물이 혼자 집에서 보내는 시간이 있는가?” 라는 질문에 약 78.7%가 그렇다고 응답했다고 한다. 아래는 반려동물이 혼자서 보내는 평균 시간을 나타낸 그래프이다. 혼자 시간을 보내는 반려동물을 위해 어떤 조치를 해주고 있는가? 라는 질문에는 아래와 같은 결과가 나왔다.



이러한 결과를 통해 반려동물을 키우는 가구 수는 증가했으나, 외벌이 가구가 줄고 맞벌이 가구가 증가하는 추세로 인해 집에 혼자 남겨져 있는 반려동물이 상당히 많음을 알 수 있다. 이러한 문제를 해결하기 위해 주인이 안심할 수 있게 반려동물 특성에 맞추어 돌봐주는 펫시터(petsitter) 라는 직업이 생겨났다. 다양한 사회, 문화적 변화가 TV나 인터넷 등의 매체들을 통해 발생하고, 반려동물의 교육과 돌봄 방식이 중요시되는 시대를 맞이한 것이다. 이에 따라 반려동물의 케어에 어려움을 겪고 있는 사람들은 신뢰할 수 있는 서비스를 제공하는 시스템이 필요하게 되었다.

따라서 본 연구에서 제안하는 앱은 비슷한 상황에 있는 반려인들이 모여 보다 저렴하고 편리하게 본인의 반려동물을 돌볼 수 있게 하기 위해 기존 돌봄 서비스에 추가로 ‘함께 돌봄’ 서비스를 제공하여 고객의 금전적인 부담을 덜 수 있다<sup>[1]</sup>. 602 실시간 스트리밍 및 업로드 영상 확인, 실시간 위치 공유 기능으로 고객이 펫시터의

1) “본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음”(2021-0-01082)

서비스 진행 상황을 바로 파악할 수 있으며, 필터링 기능을 통해 고객이 원하는 조건에 맞는 펫시터를 보다 손쉽게 검색해서 사용할 수 있도록 하였다.

## 2 App 구조 및 사용된 기술

### 2.1 App 구조

#### 2.1.1 화면흐름도

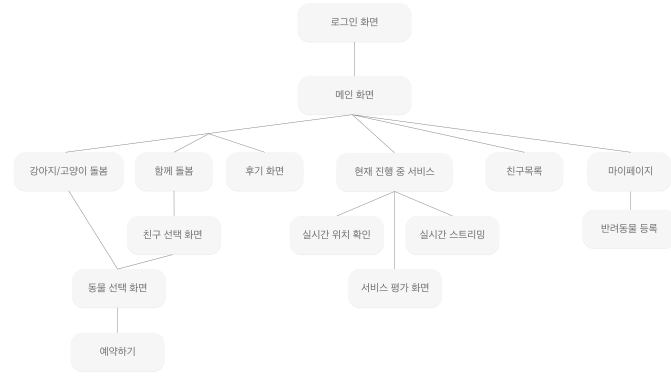
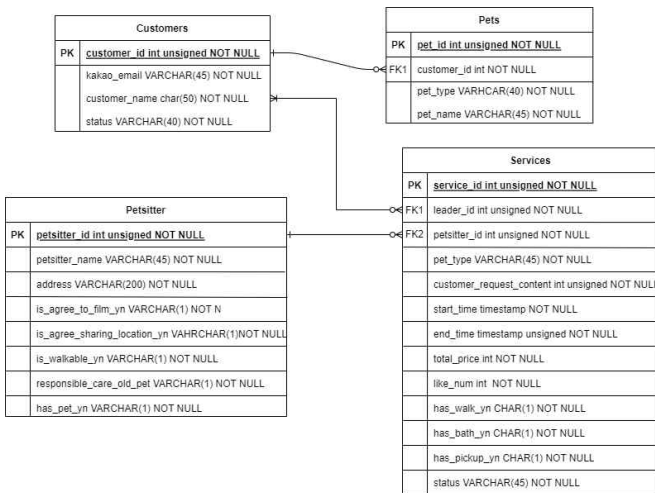


그림 38

앱을 처음 로딩했을 때 카카오 로그인을 통해서 메인화면으로 들어간다. 메인화면에는 총 4가지의 ImageButton이 있다. 각각의 돌봄서비스 버튼을 누르면 옵션 선택화면을 거쳐서 예약하기 화면으로 넘어간다. 진행 중 서비스에는 실시간 위치 확인 화면과 실시간 스트리밍 화면으로 넘어갈 수 있는 버튼이 존재하며, 서비스 이용 종료 버튼을 누를 때 서비스 평가 화면으로 이동한다. 마이페이지 화면에서는 반려동물 등록하기 화면으로 넘어갈 수 있다.

#### 2.1.2 ERD 설계



<그림 39 >

위 그림은 본 연구에서 설계한 데이터베이스의 RED 문서 중 핵심적인 부분만 간추린 그림이다. 크게 위에 4가지로 분류할 수 있다.

Customers 테이블은 고객의 정보를 저장하는 테이블이며 카카오톡을 정보를 통해 처음 로그인한 유저는 자동으로 Customers 테이블에 정보가 저장된다. status는 고객의 회원가입 진행상태를 나타내며 서비스를 받기 위한

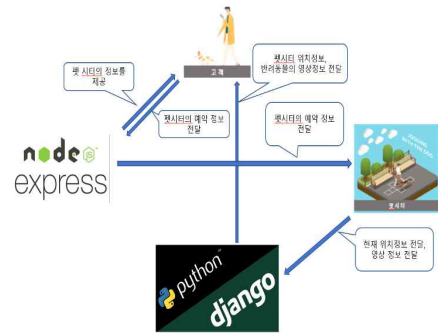
필수 정보를 추가로 입력한 유저만이 회원가입을 최종 완료한 것으로 간주한다. 회원가입이 완료되지 않은 고객은 status 칼럼을 통해 구분하며 이러한 회원들은 서비스를 이용하는데 일부 제한이 있다.

Pets는 유저가 등록한 본인의 반려동물들을 저장하는 테이블이다. 유저와 펫은 1:N 관계에 있어 테이블을 분리하였다.

PetSitters는 펫시터 이용자의 정보를 저장하는 테이블로 펫시터의 정보와 서비스를 제공하는 데 있어 필요한 정보들을 저장한다.

Services 테이블은 실제 서비스 정보를 저장하는 테이블이다. status 필드는 이 서비스가 예약상태인지 완료 상태인지 등을 저장하여 구분한다.

### 2.2 서버와의 통신



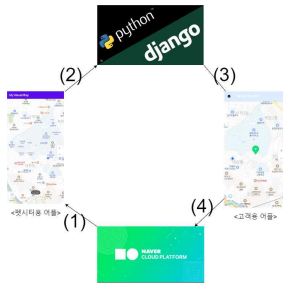
<그림 40 >

Django 서버는 펫시터가 갤러리에서 선택한 영상파일을 form-data로 전송받는다. 장고 서버는 전송받은 영상을 저장 후 저장된 경로를 SQLite 데이터베이스에 저장한다. 고객이 영상데이터를 요청하면 SQLite에서 경로를 추출해서 다운로드 링크를 전송한다. 고객은 이 링크를 통해서 영상파일을 받을 수 있다.

Node 서버는 API 서버로써 역할을 한다. 고객, 펫, 펫시터, 서비스에 대한 정보를 저장하며 각 화면에 필요한 데이터들을 가공해서 전달하는 역할을 한다. 이외에도 회원가입과 새로 생산된 정보를 DB에 저장하는 등 파일과 위치정보를 제외한 데이터를 주고받는 기능을 모두 담당한다.

Node와 Django는 서비스에서 필요한 역할을 위와 같이 분담해서 기능하고 있다. 이를 통해 하나의 서버에서 가졌어야 할 부담을 나누어 가짐으로써 서버의 부담을 덜어줄 것이라 기대한다. 빈번히 이뤄지는 호출은 대부분 Node에서 처리되며 파일을 주고받는 것과 같이 큰 용량의 데이터를 주고받는 행위는 Django에서 담당하여 처리된다.

### 2.3 위치 기반 서비스(네이버 지도 API사용)



(1) 펫시터는 Mobile Dynamic Map API를 이용하여 Naver Cloud Platform으로 본인의 위치 정보를 전달하고, Naver Cloud Platform은 이를 통해 사용자의 지도정보를 다시 펫시터에게 보낸다.

(2) 펫시터는 본인의 위치 정보를 장고 서버로 전송하고 장고는 이 위치 정보를 저장한다.

(3) 고객용어플은 장고로부터 펫시터의 위치 정보를 전달받는다.

(4) 고객용 어플은 Naver Cloud Platform으로부터 지도 정보를 전달받고 장고에게 전달받은 위치 정보를 활용해 지도에 표시한다.

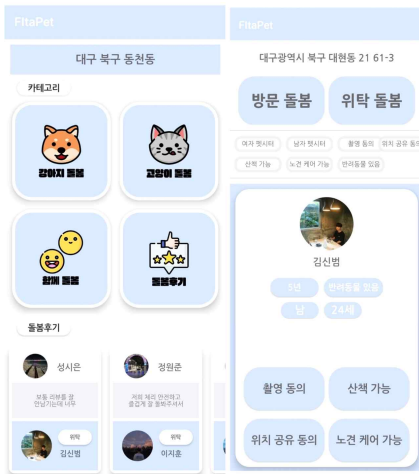


<그림 43>

기본적으로 반려동물을 등록을 한 마리라도 하지 않으면 돌봄서비스에 접근할 수 없다. 반려동물 등록에서는 기본적으로 종과 품종, 성별, 중성화 여부, 이름, 생일, 무게 정보, 반려동물의 사진을 받게 되어있고 상반되는 정보를 담은 Button은 selected 값을 true, false를 조정하며 한 버튼만 활성화 되게 구현했다. 강아지는 동물등록을 해야 하므로 강아지 선택 시에만 외장 칩, 내장 칩 버튼의 visibility를 'gone' 에서 'visible' 로 전환하여 보이도록 구현하였다. 그리고 펫시터가 참고할 수 있는 정보를 받게 하기 위해 반려동물에 대한 정보를 설문조사식으로 받는다.

### 3. App 구현

#### 3.1 펫시터 리스트 화면



<그림 42>

메인화면에서 각각의 이미지 버튼을 클릭하면 버튼에 맞게 제공하는 돌봄서비스를 이용할 수 있는 화면으로 넘어간다. 서비스 이용의 첫 단계는 나에게 서비스를 제공해 줄 수 있는 펫시터 선택이다. 이때, 내가 받고 싶은 서비스의 종류를 선택, 펫시터에게 요구하는 조건을 선택하여 필터링할 수 있다. 필터링은 버튼의 색, 혹은 버튼 stroke의 색을 click 속성에 따라 바꾸어 나타내었다. 펫시터 리스트의 전체 화면은 카드 형태와 버튼을 적절히 사용해 고객이 펫시터의 조건들을 한눈에 알아보기 쉽게 구현하였다.

#### 3.2 반려동물 등록 화면

#### 3.3 실시간 위치 확인 화면

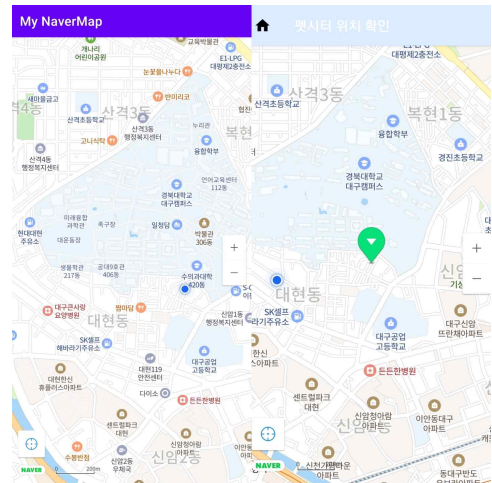


그림 44 (좌-펫시터화면, 우-고객화면)

고객이 돌봄 서비스를 신청할 경우 고객의 신뢰감 확보를 위해 기본적으로 고객이 펫시터의 위치를 실시간으로 확인할 수 있도록 하는 서비스를 제공한다. 펫시터의 위치를 확인하고 표기하기 위해 네이버 지도 API를 사용하여 화면에 지도를 띄웠다. 그리고 펫시터의 현재 위치를 기준으로 지도의 기본 화면이 설정되고, 펫시터의 현재 위치에 마커가 표시된다.

#### 3.4 실시간 영상 스트리밍 및 업로드 영상 목록 화면

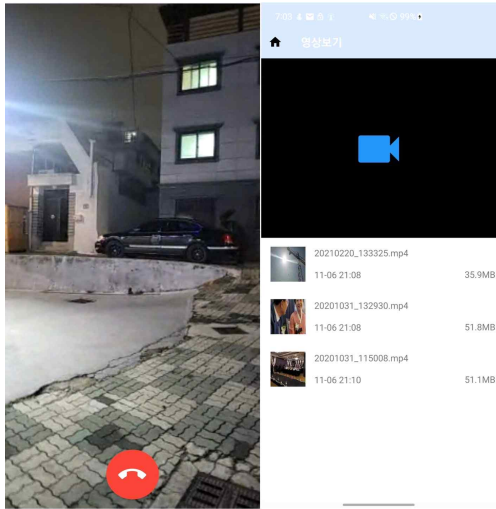
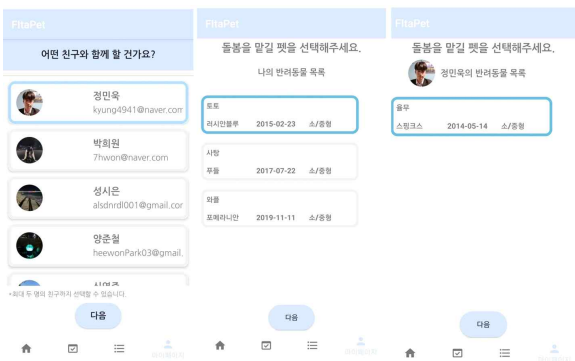


그림 45 좌-스트리밍, 우-업로드 영상

고객이 돌봄 서비스받는 동안에 현재 펫시터의 돌봄 상황을 실시간으로 확인할 수 있게 하도록 실시간 영상 스트리밍 기능을 구현하였다. 이 기능을 구현하기 위해서 Kakao i connect Live에서 제공하는 SDK를 사용하였다. 이를 통하여 영상 확인 화면에서 <그림9> 우측 화면의 원형 버튼을 클릭하면 실시간 영상 스트리밍이 시작하게 되고 밑의 종료 버튼을 클릭 시 양쪽에서 스트리밍이 동시에 종료되도록 구현하였다. 돌봄 서비스 동안 펫시터는 자신의 돌봄 상황을 영상으로 찍어 업로드 할 수 있다. 고객은 영상확인 화면에서 수시로 펫시터가 업로드 한 영상의 목록들을 썸네일과 함께 확인할 수 있고, 그 영상을 클릭하면 고객의 핸드폰에 영상이 다운로드되게 된다.

### 3.5 함께 돌봄 서비스 예약



<그림 46>

함께 돌봄은 최대 2명의 친구까지 선택할 수 있게 되어 있다. 신청하는 사람의 집 기준으로 서비스를 제공한다. 고양이는 영역 동물로, 자기 영역에서 벗어나게 되면 불안해하고 다시 돌아가기 위해 애쓰는 동물이다. 따라서 함께 돌봄은 서비스 제공 시 방문 돌봄만 제공하며, 고양이와 강아지를 함께 돌봄 신청하기 위해서는 고양이의 반려인이 신청해야만 서비스를 이용할 수 있게 제한했다. 최대 2명의 친구를 고를 수 있고 반려동물이 받는 스트레스를 고려해 본인 반려동물을 포함해 최대 3마리까지 신청할 수 있도록 했다. 함께 돌봄서비스를 이용하

면 고객은 통상적인 가격의 20% 할인가로 이용이 가능하고 펫시터는 짧은 시간에 비교적 높은 수당을 받을 수 있다는 장점이 있다. 친구목록과 반려동물 목록은 서버에 데이터 요청을 하여 데이터를 받아와 RecyclerView를 활용해 구현하였다.

## 4. 결론 및 향후 과제

결론에 앞서 언급한 설문조사를 다시 인용하자면 현재 반려동물을 키우고 있는 사람들은 1448만 명 정도이고, 그중에서 78.7%가 반려동물이 집에서 혼자 보내고 있다고 응답했다. 본 앱은 많은 사람이 현재 공통으로 겪고 있는 문제인 ‘집에 홀로 남겨진 반려동물’을 해결할 수 있게 펫시터의 서비스를 실시간 스트리밍 및 영상으로 업로드, 실시간 위치 공유 기능으로 고객이 서비스 진행 상황을 바로 파악할 수 있게 하였으며, 필터링 기능을 통해 고객이 원하는 조건에 맞는 펫시터를 보다 손쉽게 검색해서 사용할 수 있도록 하였다. 본 앱만의 차별화된 기능인 ‘함께 돌봄’이 가격 부문의 이점과 편리함을 이유로 같은 지역 내 반려동물을 키우는 직장인들에게 유용하게 사용될 수 있다고 여겨진다. 또한, 향후 펫시터용 예약 스케줄링 기능을 추가로 구현함으로써 펫시터 UX 개선 및 유저들의 편리함을 높일 수 있게 개발할 예정이다.

## 5. 참고문헌

- [1] [김경자 교수의 탐스러운 소비]펫시터 ‘1인 가구’까지 돌보네, 더스쿠프  
<https://www.thescoop.co.kr/news/articleView.html?idxno=21602>
- [2] 반복되는 국립공원 고양이 갈등...먹이 주지 말랬더니 민원 빗발, 연합뉴스  
<https://www.yna.co.kr/view/AKR20221111115600530?input=1195m>

# 머신러닝을 활용한 인지기능 상태 예측

구동준<sup>○</sup> 이아라 원경재 김일곤<sup>2)</sup>  
경북대학교

dongjunkoo02@gmail.com, eara0367@gmail.com, yui06031@gmail.com, ikkim@knu.ac.kr

## Prediction of Cognitive Status using Machine Learning

Dongjun Koo<sup>○</sup> Ah Ra Lee Kyung Jae Won Il Kon Kim  
Kyungpook National University

### 요약

치매는 노인 집단이 가장 두려워하는 질병으로 치료제가 없어 현재로서는 예방이 최선이다. 치매 진단에는 간이정신상태검사(Mini-Mental State Examination)와 같은 다양한 인지기능 검사가 사용되고 있다. 노인의 경우 건강 상태가 빠르게 변하기 때문에 지속적인 인지기능 상태 파악이 필요하지만, 간이정신상태검사는 질문 특성상 전문가의 도움이 필요하므로 자주 검사를 받기에는 어려움이 있다. 본 연구에서는 머신러닝을 활용하여 인지기능 상태를 예측하고자 하며, 검사방법은 개인이 스스로 진행할 수 있는 설문지 형태로 개발하여 언제 어디서든 자신의 인지기능을 측정할 수 있도록 하고 치매 예방을 돕고자 한다.

### 1 서론

전 세계적인 고령화로 기대수명이 증가하면서 노인 의료비 부담 부담 또한 심화되고 있다. 특히 치매는 국내 60세 이상 노인 집단이 가장 두려워하는 질병 1위로 언급되기도 했는데 치매에 걸리면 기억력, 인지기능, 행동 등에 문제가 생겨 일상생활이 힘들어질 뿐 아니라 주변인에게도 병간호로 인한 시간적, 경제적 부담을 초래하기 때문이다. 세계보건기구(World Health Organization, WHO)에 의하면 2019년 기준 전 세계 치매 환자는 약 5,500만 명으로 2030년에는 7,800만 명, 2050년에는 1억 3,900만 명까지 늘어날 것으로 전망되고 있다[1]. 하지만 현재 치매는 완벽하게 치료할 방법이 없어 조기에 이상 징후를 발견하고 예방하는 것이 최선으로 여겨지고 있다. WHO에서는 인지기능 및 치매 위험 감소를 위한 권고사항을 발표하기도 했는데, 특히 인지기능 감소 및 장애를 높이는 인자들을 일상생활에서 조기에 관리함으로써 치매 발병 위험을 효과적으로 줄일 수 있다고 언급한 바 있다[2].

치매 진단과 관련해 다양한 인지기능 검사가 사용되고 있는데 간이정신상태검사(Mini-Mental State Examination, MMSE)가 그 대표적인 예이다[3]. MMSE는 다른 인지기능 검사법에 비해 비교적 짧은 시간 내에 인지 장애 진행

정도를 파악하기 위해 임상 및 연구 환경에서 널리 사용되는 30점 만점의 설문 검사법으로 전체 점수가 24점 이하일 경우 치매 혹은 경도인지장애를 의심할 수 있다. 하지만 MMSE의 경우 임상 전문가의 해석이 필수적이기 때문에 검사를 시행할 수 있는 특정 기관에 방문해야만 한다. 최근에는 65세 이상뿐 아니라 비교적 젊은 40-50대에서도 조발성 치매가 많이 발견되고 있는데 우리나라의 경우 만 60세 이상에만 무료로 MMSE 검사를 지원하고 있어 치매 위험을 조기에 진단하거나 예방하는 데 어려움이 있다.

MMSE 검사의 한계점을 극복하기 위해 전문가 도움 없이 개인이 인지기능 저하 및 치매 위험을 판별할 수 있도록 Information Questionnaire on Cognitive Decline in the Elderly(IQCODE)나 Seoul Information Report Questionnaire for Dementia(SIRQD)와 같은 치매 선별용 정보 제공자 보고형 설문 방식의 평가 도구가 개발되어왔다[4,5]. 이러한 설문 방식의 평가 도구는 기관에 방문하지 않고 가정에서 사용할 수 있어 시행이 간편할 뿐 아니라 노인이 직접 응답하지 않더라도 일상생활을 함께하는 가족 혹은 노인의 건강 상태를 잘 알고 있는 보호자가 대신 설문에 응답하여 치매 위험을 판별할 수 있다는 장점이 있다. 또한, MMSE와 비교했을 때도 조기에 치매 징후를 발견하는 능력이 우수한 것으로 보고되어 치매가 발생하기 전 미리 예방 및 관리하는데 효과적인 지표로 사용될 수 있다[6].

하지만 MMSE는 현재 시점의 인지기능을 검사하는 방식인 반면 IQCODE와 SIRQD는 10년에 걸친 인지기능의 변화를 평가하는 방식으로 답변하기 어려운 항목이 존재한다. 과거와 현재 상태를 비교하는 유형의 설문 문항은 인지기능이 저하되지 않은 사람도 10년 전 건강 상태를 정확하게 기억하는 것이 쉽지 않을 뿐 아니라 10년 동안 노인의 건강 상태를 계속해서 지켜본 가족 등 대신 답변해줄 수 있는 보호자가 존재하지 않을 수도 있어 정확한 응답을 수집하기 어렵다는 한계점이 있다.

2) This work was supported by the Technology Innovation Program(20011061, Personalized digital health framework for infectious diseases and frailty prevention) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea). This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2018R1A6A1A03025109). This research was supported by a grant from the Korea Health Technology R&D Project through the Korean Health Industry Development Institute, which is funded by the Ministry of Health & Welfare, Republic of Korea(grant number: HI15C3153).



따라서 본 연구에서는 의료기관에 방문하지 않고도 개인 혹은 보호자가 직접 가정에서 현재 건강 상태와 관련해 간단하게 응답할 수 있는 몇 가지 설문 문항을 통해 인지기능 상태를 평가함으로써 조기에 치매 징후를 예측할 수 있는 머신러닝 모델을 개발하고자 한다.

## 2 연구 방법

머신러닝 모델 학습에는 한국노인노쇠코호트 사업단에서 수집한 한국인 노쇠 코호트 데이터를 사용한다[7]. 해당 데이터 셋에는 한국 70~84세 노인 집단을 대상으로 인구 사회학적 정보, 노쇠와 관련된 항목에 대한 검사 결과 및 설문 응답 등이 수집되어 있다. 본 연구에서는 인지기능 상태를 평가하기 위한 MMSE 검사 결과와 개인 스스로 응답이 가능한 116개의 설문 문항을 사용한다.

2016년 및 2017년 조사에 응답한 참가자 총 3,016명을 대상으로 MMSE 검사 결과에 따라 24점 미만으로 치매 혹은 경도인지장애를 의심할 수 있는 사람과 확정적 정상 판정을 받은 사람으로 분류한다. 잠재적 요인 변수로 사용하는 설문 항목에 대한 결측값 및 모름으로 응답한 경우를 제외한 후 치매 혹은 경도인지장애를 의심할 수 있는 사람 461명, 확정적 정상 1,858명으로 불균형한 분포를 이루고 있음을 확인할 수 있다. 적은 수의 클래스를 기준으로 언더샘플링을 수행하여 실험에는 총 922개의 데이터(확정적 정상: 461명, 치매 혹은 경도 인지장애 의심: 461명)를 사용한다. 학습에 사용할 116개의 자가 진단 설문 항목에 대해서는 변수 선택 기법을 적용한다. MMSE 검사 점수에 대한 상관계수를 계산하여 높은 상관관계를 보이는 50개를 선정한 후 그림 1과 같이 샐플리 값(Shapely value)을 기반으로 가장 영향력이 높은 변수 8개를 선정한다 (표 1). 선정한 변수에 대해 최소-최대 정규화를 적용하고 전체 데이터는 학습 데이터와 테스트 데이터(9:1)로 분리한다. 학습 데이터는 Stratified K-fold 방식을 사용해 학습 및 검증 데이터로 분리하여 교차검증을 수행한다.

모델 학습 및 평가에는 다양한 분류 문제에서 고르게 좋은 성능을 보이는 네 가지 모델인 Support Vector Machine(SVM), Random Forest(RF), Extreme Gradient Boosting(XGboost)을 사용하여 Precision, Recall 및 F1-score를 사용해 모델별 최종 성능을 평가한다[8].

표 1. 인지기능 저하 예측에 사용된 변수

요소	내용
age	나이
height	키
weight	몸무게
k_abc5	물건에 닿기 위해 발뒤꿈치를 든다
k_abc15	에스컬레이터를 난간을 잡지 않고 탄다
sarcf4	보조기구나 타인의 도움이 없이 10개의 계단을 쉬지 않고 오르는 것이 얼마나 어려운가요?
sf12_2_1	다소 힘든 활동
sn5	자주 보고 이야기를 나누는 친구와 얼마나 자주 보거나 이야기를 나누십니까?

## 3 연구 결과

머신러닝 모델별 성능 측정 결과, 표 2와 같이 다양한 모델에서 전반적으로 고르게 좋은 성능을 보였으며 그중에서도 Random Forest 모델이 평균 F1-score 기준 0.82로 가장 좋은 성능을 보였다. 확정적 치매(Class 0)에 대해서는 정밀도(precision)와 재현율(recall) 모두 0.71 이상의 성능을 보였으나 치매 혹은 경도인지장애를 의심할 수 있는 클래스(Class 1)에 대한 재현율은 0.66~0.77 사이로 비교적 낮은 결과를 보였다.

표 2. 머신러닝 모델 별 성능 평가 결과

모델		Precision	Recall	F1-score
SVM	Class 0	0.71	0.85	0.77
	Class 1	0.81	0.66	0.73
	Average	0.76	0.75	0.75
DT	Class 0	0.69	0.74	0.72
	Class 1	0.73	0.68	0.70
	Average	0.71	0.71	0.71
RF	Class 0	0.78	0.87	0.82
	Class 1	0.86	0.77	0.81
	Average	0.82	0.82	0.82
XGB	Class 0	0.71	0.87	0.78
	Class 1	0.84	0.66	0.74
	Average	0.78	0.76	0.76

## 4 결론 및 향후 연구

본 연구에서는 자가 진단 설문을 바탕으로 개인이 가정에서 인지기능 저하를 예측할 수 있는 머신러닝 모델을 제안했다. 병원이나 보건소 등 특정 기관에 방문하지 않고도 개인의 건강 상태와 관련된 설문 응답을 통해 인지기능 이상 징후를 발견함으로써 치매가 발생하기 전 의료기관에 방문하여 정밀 검사를 받을 수 있다. 또한, 치매는 인지기능 외에도 일상생활 기능 저하, 기저 질환 등 다양한 요인에 의해 진단되는 질환이기 때문에 향후 보다 복합적인 측면에서의 영향 요인을 활용해 인공지능 모델을 개발한다면 치매의 조기 진단 및 진행 속도를 늦추기 위한 예방적 건강 관리에 기여할 수 있을 것으로 기대한다.

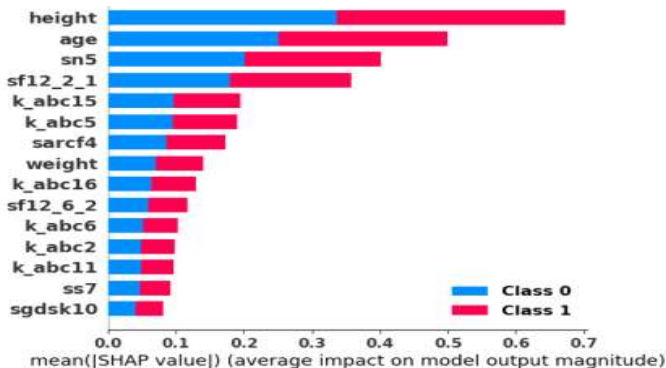


그림 47. 잠재적 요인 변수에 대한 Shapley value 측정 결과

\* Class 0: 확정적 정상 (MMSE 점수 24점 이상)

\* Class 1: 치매 혹은 경도 인지장애 의심 (MMSE 점수 24점 미만)

## 참고문헌

1. World Health Organization, “Dementia” [Online] Available: <https://www.who.int/news-room/fact-sheets/detail/dementia> (Accessed: May. 6, 2022)
2. World Health Organization, “Global action plan on the public health response to dementia 2017 - 2025.” [Online] Available <https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025> (Accessed: Nov, 2022)
3. Marshal F. Folstein, Lee N. Robins, John E. Helzer, “The Mini-Mental State Examination,” in *Arch Gen Psychiatry*, 40 (7):812. doi:10.1001/archpsyc.1983.01790060110016, 1983
4. Lee, Dong Woo, et al. “Validity of the Korean version of Informant Questionnaire on the Cognitive Decline in the Elderly (IQCODE).” in *Journal of the Korean Geriatrics Society* 9.3: 196-202, 2005
5. 이동영, 김기웅, 윤종철, 주진형, 이정희, 우종인, “치매 선별용 정보 제공자 보고형 설문지,” *대한신경정신의학회*, vol. 43, no.2, pp. 209-218, 2004
6. Ritchie K, Fuhrer R. “A comparative study of the performance of screening tests for senile dementia using Receiver Operating Characteristics analysis.” *J Clin Epidemiol*, 45:627-637, 1992
7. Won CW., et al. “Korean frailty and aging cohort study (K FACS): cohort profile.” in *BMJ Open*, 10(4), 2020
- 8, Olson RS, Cava WL, Mustahsan Z, Varik A, Moore JH. “Data-driven Advice for Applying Machine Learning to Bioinformatics Problems.” in *arXiv:1708.05070 [q-bio.QM]*, 2018

# 아두이노를 이용한

## 무호흡 감지 및 증상 완화 모듈 설계 및 구현

김대영, 서민석, 이현로<sup>o</sup>  
경북대학교 컴퓨터학부

E-mail: eodud5730@naver.com, tjaltstjr111@naver.com, youro0007@naver.com

### Design and Implementation of Apnea Detection and Symptom Relief Module, using Arduino

Daeyoung Kim, Minseok Seo, Hyeonro Lee<sup>o</sup>  
Dept. of Computer Science Engineering, Kyungpook National University

#### 요 약

잠을 자는 도중 일시적으로 숨을 멈추는 현상인, 수면 무호흡증 환자가 최근 급증하고 있다. 무호흡 상태가 지속될 경우 산소 부족으로 인한 뇌 손상으로 치매, 뇌졸중 등의 합병증을 초래할 수 있다. 본 논문에서는 아두이노 생체신호 감지 모듈을 사용하여 수면 중 무호흡 증상을 감지하고, 진동이나 작은 소리를 통해 무호흡 증상을 멈추고자 한다. 나아가 진동이나 작은 소리로 막을 수 없는 경우, 스마트폰을 통해 보호자나 병원에 알려 즉각적인 조치를 받도록 돕는다.

## 1. 서 론

잠을 자는 도중 일시적으로 숨을 멈추는 현상인, 수면 무호흡증 환자가 최근 급증하고 있다. 국민건강보험공단으로부터 제출받은 자료에 따르면 수면 무호흡증 진료환자는 2015년 2만 9255명에서 2019년 8만 6006명으로 5만 6751명(194%)이 급증했다고 한다.[1]

10초 이상 호흡을 멈추는 증상을 무호흡이라고 하는데, 이 증상이 지속될 경우 산소 부족으로 인한 뇌 손상으로 치매, 뇌졸중 등의 합병증을 초래할 수 있다. 이외에도 심정지에 의한 돌연사도 발생할 수 있는데, 수면 중 무호흡이 염증을 발생시켜 고혈압, 관상동맥 질환, 심부전 등을 증가시키고 이로 인해서 사망까지 발생할 수 있다는 것이다.

이러한 수면무호흡의 발생원인에는 크게 두 가지가 있다.

먼저, 수면 중에는 근육들이 이완되어 늘어지기 때문에 부분적으로 공기 통로가 좁아지는데, 기도 부위의 과도한 근육 이완이나 기타 원인으로 인해 공기 통로가 일시적으로 완전히 막혀, 수면무호흡이 발생할 수 있다.[2]

두 번째 경우 뇌간이라는 뇌의 일부에서 호흡 조절에 문제가 있을 때 발생한다. 즉, 호흡에 관여하는 신경이 일시적으로 기능을 하지 않기 때문에 숨 쉬려는 노력을 하지 않게 되는 것을 말하는 것이다.[3] 따라서, 자극을

통해 뇌를 활성화하여 호흡을 유도하는 과정이 필요하다.

그래서, 본 연구에서는 아두이노 생체신호 감지 모듈을 사용하여 수면 중 무호흡 증상을 감지, 진동 및 소리를 통해 뇌 신경을 자극하여 자발적인 호흡을 유도하고자 한다. 나아가 여러 자극에도 여전히 호흡 반응이 감지되지 않을 시, 스마트폰을 통해 보호자나 병원에 알려 즉각적인 조치를 받을 수 있게 한다.

## 2. 시스템 구성

### 2.1 아두이노 구성

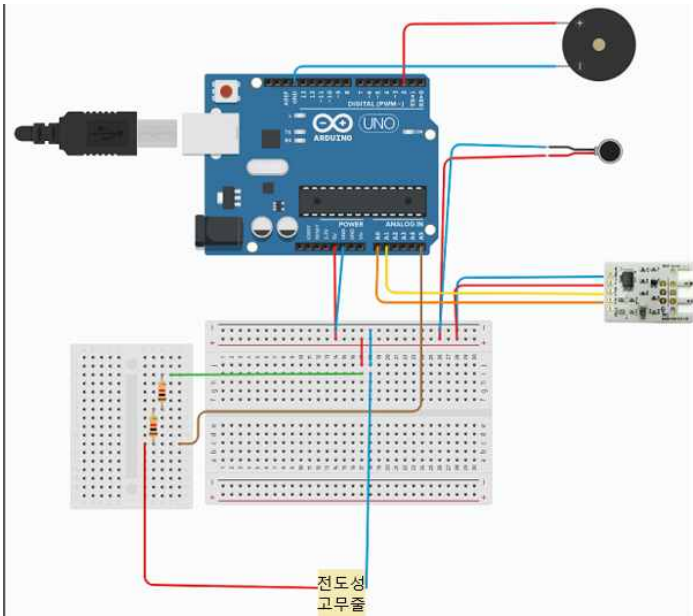
전체적인 구성은 아두이노 Uno R3 보드에 호흡 감지 센서, 호흡 유도 장치, HTTP 통신을 위한 WIFI 모듈을 연결하였다.

호흡 감지 센서에는 바람 센서(ZAS-MD-0550), 장력 센서(전도성 고무줄)가 있다. 바람 센서를 통해 코나 입을 통한 호흡 유무를 판단한다. 장력 센서를 이용하여 들숨, 날숨에 의해 배가 부푸는 정도, 즉 장력 센서 값의 변화량을 추적하고, 이를 통해 호흡 유무를 판단한다.

호흡 유도 장치로는 진동 모터(DM159), 피에조 부저(DM657)를 사용하였다.

WIFI 통신 모듈은 esp8266 esp-12e uart를 사용하였다. 이를 통해 사용자 정보를 받고, API호출 및 아두이노

측정값을 서버로 전송한다.



[그림1] 아두이노 회로도

## 2.3 안드로이드 애플리케이션 구성

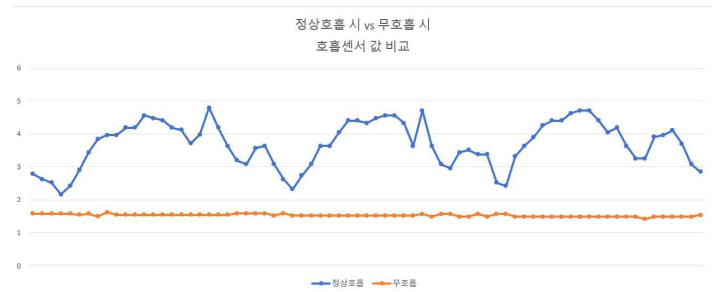
안드로이드는 웹 뷰를 통해 웹페이지를 띄워준다. 반응형 웹을 사용하여 사용자의 기기에 맞는 웹페이지를 구성하여 준다. 사용자는 로그인 및 회원가입, 마이페이지를 통한 측정 데이터 및 그래프 보기, 아두이노 측정을 시작하기 위한 측정 시작 및 종료 버튼, SMS API를 위한 설정 페이지를 구현하였다.

## 3. 구현 및 실험 결과

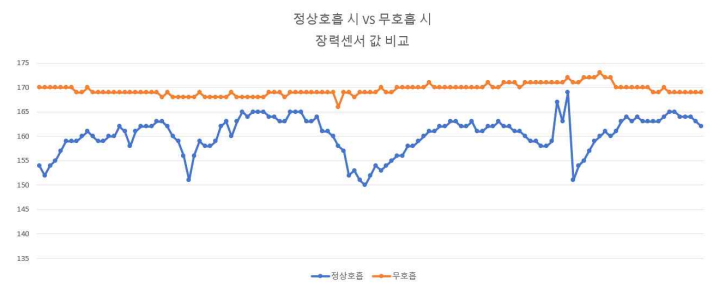
먼저 사용자의 호흡 패턴을 파악하기 위해 초기 측정을 진행한다. 초기 측정을 진행하는 이유는 호흡량(바람센서값)과 그에 따른 복부의 팽창도(장력센서값)는 사람마다 제각각 다르기 때문이다.

측정은 수시로 변화하는 호흡을 보다 정확하게 측정하기 위해 0.1초마다 진행한다.

아래 그림2와 그림3은 센서를 착용하고 측정한 호흡센서 및 장력센서 변화 그래프이다.



[그림3] 호흡 센서값 비교(정상 호흡 vs 무호흡)

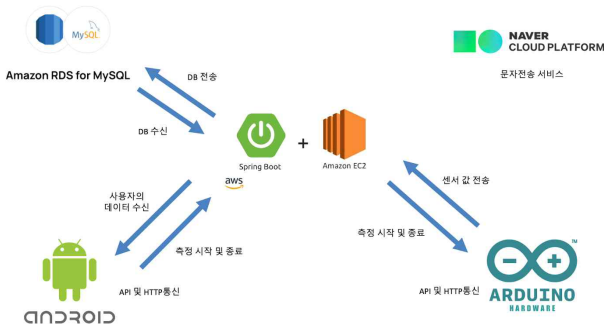


[그림4] 장력 센서값 비교(정상 호흡 vs 무호흡)

두 그래프의 값을 보면 정상 호흡 시 들숨과 날숨에 따라 값의 변화가 규칙적인 패턴으로 나타남을 볼 수 있다. 이를 기반으로 무호흡을 판단하는 기준값을 정하였다.

호흡 센서의 경우, 정상적인 호흡 시에는 2~5mph, 무호흡 시에는 0~2mph 사이의 값이 나타난다. 호흡에 따

## 2.2 서버 및 웹 페이지 구성



[그림2] 아두이노 및 서버 구성도

웹서버로 자바 기반 프레임워크인 Spring Boot를 사용하며, 아마존 웹 서비스 중 가상 컴퓨터(EC2) 인스턴스를 통해 서버 배포를 진행한다. 아두이노와 HTTP 통신을 진행하고, JPA를 사용하여 MySQL과 연동을 진행한다.

아두이노로부터 센서값을 HTTP 통신 및 API 통신을 통해 받고, 이를 DB에 저장하고자, MySQL을 사용한다.

웹 페이지는 자바 서버 템플릿 엔진인 mustache를 통해 구현한다.

웹 페이지를 통해 사용자는 아두이노의 측정 시작과 종료를 진행할 수 있고, 네이버 클라우드 플랫폼 SMS API를 통해 응급상황 시 자동 문자메시지 전송이 진행된다.

또한, 사용자 개인만의 수집된 데이터로 수면무호흡 측정 현황을 보여주는 개인 페이지를 구성하고, 로그인 및 세션을 관리하고자 Spring Security를 사용한다.

른 바람의 세기를 측정하는 것이기 때문에 단순히 최소 호흡 기준값만 넘으면 호흡이 이루어졌다고 판단한다. 그 기준값은 정상호흡 시의 호흡량을 기반으로 정해지며, 다음과 같다.

$$P = \sum_{i=1}^n B/n \times 0.8$$

[그림4] 호흡 센서 기준값 수식

P는 호흡판단 기준값이고, B는 호흡 센서값, n은 초기 측정 표본 수를 의미한다. 초기 호흡량 측정 결과의 평균에 0.8을 곱한 값을 기준값으로 지정했다. 보통 호흡량의 40% 정도라고 볼 수 있다.

장력 센서의 경우, 0에서 1023 사이의 아날로그 값이 센서값으로 출력되는데, 정상적인 호흡 시에는 150~170의 값이 나타난다. 호흡을 판단하는 데 있어서 복부 팽창도는 단순히 그 값의 크기가 아닌 변화량이 중요하다. 그렇기 때문에 호흡 센서의 기준값과 달리 값의 고저, 즉 변화량에 주목하였고 그 기준은 다음과 같다.

$$P = (Max + Min) \times 0.4$$

[그림5] 장력 센서 기준값 수식

P는 호흡판단 기준값이고 Max는 측정구간의 최댓값, Min은 측정구간의 최솟값을 의미한다. 기준값은 보통 호흡 변화량의 40%로 지정했다. 매 측정 시 버퍼(저장공간)에 센서값을 저장해두고, 그 변화량이 초기 측정 시 설정된 기준값, 즉 평균 호흡량의 40%를 넘지 못하면 무호흡이라고 판단한다.

이러한 기준을 바탕으로 무호흡 지속시간을 측정하고, 무호흡이 얼마나 오래 지속되는지에 따라 위험도를 3단계로 나누었다. 이때, 호흡 센서와 장력 센서 두 센서가 동시에 측정을 진행하고, 둘 중 하나라도 위험을 감지하면 다음과 같은 조치를 취한다.

무호흡 지속시간이 10초 이상일 경우 위험도 1에 해당하며, 진동 자극으로 사용자의 호흡을 유도한다. 지속시간이 15초 이상인 경우 위험도 2 상황이며, 소리 자극을 통해 호흡을 유도한다. 지속시간이 30초를 넘어가면 위험도 3단계가 되고, HTTP 통신을 통한 서버로의 API 호출을 통해 보호자에게 긴급 문자 알림이 가도록 구현했다.

측정 종료 시, 수면 중 감지된 총 무호흡 횟수를 사용자 데이터베이스에 저장하여 추후 앱을 통해 확인할 수 있다.

## 4. 결 론

아두이노를 이용하여 무호흡 증상을 감지하고 증상 완화를 돕는 IoT 생체신호 감지 모듈을 설계 및 구현해 보았다. 본 시스템은 병원이 아닌 집에서, 단순히 무호흡 감지뿐만 아니라 즉각적인 조치가 이루어진다는 점에서 의미가 있다.

하지만 아두이노 센서가 인체 친화적이지 않아 착용에 다소 불편함이 있고, 실제로 본 시스템이 수면 무호흡증 환자들의 증상 완화에 유의미한 영향을 주는지에 대한 결과분석이 제대로 이루어지지 않았다는 한계가 있다.

따라서 착용하기 편한 센서 개발 및 개선과 함께, 수면 무호흡 환자들을 대상으로 한 결과분석 실험이 향후 연구의 주제가 될 것이다.

## 사 사 문 구

**"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음" (2021-0-01082)**

# WebRTC 기반 실시간 영상공유 및 모니터링 시스템의 설계 및 구현

이원근<sup>o</sup>, 이성원, 천용태, 최명빈  
경북대학교 IT 대학 컴퓨터학부

wongun0409@gmail.com, poi933@naver.com, ide2143@naver.com, bbin852@knu.ac.kr

## Design and Implementation of a WebRTC-Based Real-Time Image Sharing and Monitoring System

Wongun Lee<sup>o</sup>, Seongwon Lee, Myoungbin Choi, Yongtae Cheon  
Kyungpook National Univ. Department of Computer Science

### 요약

RTC(Real-time Communication)란 웹 브라우저를 확장한 새로운 표준으로, 인터넷을 통한 소셜 미디어, 화상회의와 같은 정보에 접근할 수 있도록 하고, 통합된 소통을 가능하게 한다. WebRTC는 RTC를 Web 기반으로 처리하는 기술의 세계 표준으로, 기본적으로 일대일(P2P, peer to peer) 통신을 지원하지만, 미디어 서버를 구축한다면 1인 라이브 스트리밍(1:N)이나 다자간 통신(N:N) 또한 구현이 가능하다. 본 연구는 WebRTC와 미디어 서버를 이용한 다자간 화상 통신에 대해 소개하고, 이를 응용한 웹서비스를 설계 및 구현해 사용자의 접근성과 사용성을 높였다.

## 1 서론

현장에서의 안전에 대한 관제, 모니터링은 과거부터 중요시 되어왔다. 이러한 사고에 의한 인명 피해를 줄이기 위해 빠른 대처 및 실시간 모니터링의 중요성이 대두되고 있다. 또한 최근 코로나19로 인해 언택트 시기를 겪으면서 화상 회의 등비대면 커뮤니티 활동이 활성화되었다.<sup>3)</sup> 웹 기반 화상회의에 많이 사용되는 표준은 WebRTC(Web Real-time Communication)인데, 이는 높은 실시간성과, 웹을 통한 음성 및 화상 통화가 특징이므로 원격 CCTV 모니터링과 동시에 화상회의가 가능하며 별도의 플러그인이나 추가 설치가 필요하지 않아 사용자 편의성이 높다. 이에 본 논문은 CCTV와 화상회의 플랫폼을 결합하여 실시간 영상공유 및 모니터링 시스템을 설계 및 구현하였다.

## 2 관련

WebRTC를 이용해 다대다 통신을 구현하기 위해서 사용하는 서버는 다음과 같다. 먼저 Node.js를 사용한 시그널링 서버에서 각 클라이언트를 WebSocket에 연결해 서로 신호를 보낼 수 있게 한다. 이때 NAT(Network Address Translation) 환경을 사용하여 Private IP에서 Public IP로 변환해야 하는데, COTURN을 활용한 STUN(Session Traversal Utilities for NAT) 서버와 TURN 서버(Traversal Utilities for NAT)를 활용해 NAT 환경에서도 P2P로 데이터 교환이 가능하도록 한다. 이후 Kurento 미디어 서버를 이용해 WebRTC의 P2P 통신을 다대다 연결로 확장하고, CCTV와도 연결해 시그널링 서버와 CCTV의 데이터를 관제 및 통합한다.

### 2.1. WebRTC

WebRTC는 위에서 언급한 Zoom이나 Webex같은 서비스와는 다르게 별도의 소프트웨어나 플러그인 등이 없어도 실시간 통신이 가능하게 해주는 기술이다. 하지만 기본적으로 WebRTC는 P2P 형식으로 데이터를 주고받는 개념이기 때문에 요구되는 서비스를 제공하는 것이 어렵다. 이러한 P2P 형식으로 되어있는 WebRTC로 N:N의 통신을 구현하기 위한 방법이 3가지 있는데 Mesh, MCU(Multipoint Conferencing Unit), SFU(Selective Forwarding Unit)이다.

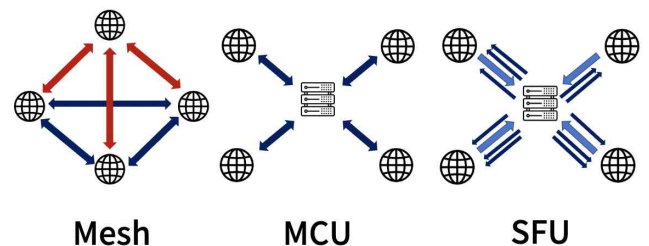


그림 1. 미디어 서버의 설계 방법

#### A. Mesh

Mesh 방식은 모든 클라이언트가 다른 클라이언트와 연결되는 것으로, 따로 서버를 두지 않기 때문에 구현이 가장 간단하다. 하지만 클라이언트의 수가 n개로 늘어날 경우 한 클라이언트당 n-1개, 총 n(n-1)/2개의 연결이 필요하므로 네트워크 스트림의 중복과 더불어 각 클라이언트 브라우저의 메모리나 CPU 사용량도 커져 확장성이

떨어진다. 이러한 단점 때문에 대규모 미디어 서버에 크게 사용되지는 않는다.4)

### B. MCU

MCU(Multipoint Control Unit) 방식은 서버에서 모든 클라이언트의 데이터를 모은 후 한 번에 처리하는 것을 말한다. 클라이언트는 서버로의 연결만 하면 되기 때문에 각 클라이언트 브라우저의 메모리나 CPU 사용량도 적고, 클라이언트가 많을 때도 모든 클라이언트를 연결할 필요가 없기 때문에 Mesh보다 확장성이 좋다. 그러나 중앙 서버에서 모든 클라이언트의 데이터를 처리해야 하기 때문에 서버의 큰 대역폭과 뛰어난 성능의 CPU와 메모리를 요구한다는 단점이 있다.

### C. SFU

SFU(Smart Forwarding Unit)은 클라이언트들이 하나의 병합된 스트림이 아닌 각각 2개(오디오, 비디오)씩의 트랙을 가진 단일 스트림을 수신한다. 중간 서버는 따로 인코딩이나 디코딩 없이 클라이언트에 전달하는 기능만 수행한다. 그러므로 MCU보다 요구되는 서버의 성능이 높지 않아도 되는 것이 큰 장점이다.

본 연구에서는 MCU 방식이 클라이언트의 확장성이 뛰어나고 각 클라이언트 브라우저에 부담이 없다는 점을 고려해 MCU로 미디어 서버를 구현하였다.

## 2.2. Kurento

### A. Kurento-Media-Server

N:N 통신을 구현하기 위하여 MCU 방식을 사용하려면 중앙 미디어 서버가 필요해진다. Kurento(Kurento Media Server, KMS)란 WebRTC 플랫폼을 개발할 때 사용되는 멀티미디어 서버 패키지로, 오픈소스 플랫폼이다5).

본 연구에서는 Kurento-Media-Server라는 오픈 소스 플랫폼을 활용하여 이를 구현할 수 있었다.

### B. Kurento의 모듈

Kurento는 Media Elements와 Media Pipeline이라는 두 가지 개념에 기반을 두고 있으며 미디어를 모듈식으로 처리한다. Elements를 Pipeline에 연결하여 데이터를 주고받는다는 개념을 통하여 아주 직관적인 접근이 가능했다. Media Elements 개념 안에 Endpoint가 존재하며 우리가 필요로 한 Endpoint는 WebRtcEndPoint와 PlayerEndpoint이다. WebRtcEndPoint가 바로 WebRTC 기술을 구현하여 통신하는 Endpoint이며 구현에서 각 Client들에게 할당되어 미디어 스트리밍을 지원하여 줄 것이다. PlayerEndpoint는 CCTV(IP 카메라)의 RTSP(Real-Time Streaming Protocol)을 Media Pipeline에 삽입하는 Endpoint이며 구현에서 RTSP URL을 할당받아 다른 모든 WebRtcEndpoint(유저)들에게 CCTV의 영상을 스트리밍 해줄 것이다.

### 2.3. TURN(Traversal Using Relays around NAT) 서버

클라이언트들과 미디어 서버가 통신할 때, 서로 같은 로컬 네트워크에 있지 않다면 방화벽이나 NAT 환경에 의하여 서로의 통신이 불가능할 수 있다. 직접적인 시그널링을 위하여 TURN 서버를 구축하여 경유할 필요가 있었으며, 오픈 소스인 COTURN 프로젝트를 활용하여 TURN 서버를 구축할 수 있었다.

### 2.4. 구현 환경

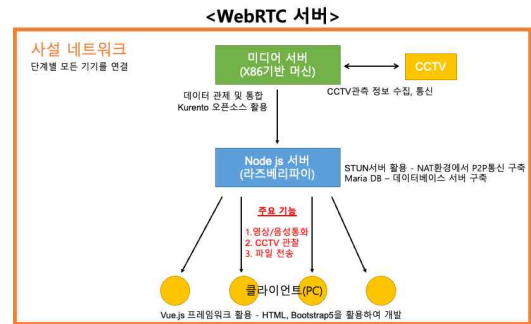


그림 2. 다대다 통신에 사용된 WebRTC 서버 개요도

개발은 일반 가정의 KT 인터넷을 연결한 IPTime 공유기를 통하여 각 서버에 네트워크 환경을 구축해주었다. X86 머신은 맥북에 Ubuntu 18.04 버전을 탑재하여 KMS를 구축하였고 TURN 서버를 활용하여 외부에서 접근할 수 있도록 설정해주었다. CCTV(IP 카메라)는 VLC라는 미디어 플레이어를 통하여 RTSP로 실시간 스트리밍을 하였고 RTSP URL을 통하여 미디어 서버의 Pipeline에 삽입될 것이다. 웹서버와 시그널링 서버는 라즈베리 파이를 사용하였다. 구현은 백엔드는 Node.js로, 프론트엔드는 HTML, Bootstrap 프레임워크를 활용하여 개발하였다. 각 서버를 이렇게 Window에서의 WSL이나 가상 머신을 통하지 않고 네이티브 리눅스를 활용함으로써 Window에서의 복잡한 네트워크 환경 세팅 과정을 생략할 수 있었다.

### 2.5. 구현 방법

시그널링 기능은 WebRTC에서 자체적으로 지원하는 기능이 아니기 때문에 WebSocket API를 사용하여 시그널링을 구현하였다. 클라이언트가 웹에 접속한 뒤에 통신을 시작하는 순간, 개별 클라이언트는 webRtcEndPoint, KurentoClient, MediaPipeline을 생성해 연결한다. 미디어 서버, 그리고 미디어 서버와 통신하기 위한 pipeline은 모든 클라이언트가 동시에 사용하는 공용 객체이기 때문에, 생성되어 있지 않은 경우에만 새로 생성하고, 아닐 경우에는 이미 생성된 pipeline과 kurentoClient를 가져와 본인의 webRtcEndPoint를 생성해 연결한다.

이 과정을 통해서 오직 1대1 혹은 1대다 통신만이 가능하다. 이 프로젝트에서 다수 이용자의 데이터를 mixing

하여, 하나의 link를 통해 통신하는 MCU(multi-point control unit) 방식을 채택하였다. 즉, 개별 클라이언트는 송수신을 위한 총 두 개의 link만이 필요하다. 위의 과정은 클라이언트가 본인의 영상을 미디어 서버에 송출하는 하나의 uplink를 생성한 과정이며, 이는 다대다 통신을 위해서는 downlink 작업이 추가적으로 필요함을 암시한다. 이를 위한 객체가 composite과 hubPort이다.

최초 클라이언트가 통신을 시작하였을 때 위의 과정을 수행한 후 추가로 composite과 hubPort를 생성한다. Composite은 각각 클라이언트의 데이터를 mixing 하여, 한 번에 보내기 위한 공용 객체로 이 역시, 최초 접속한 클라이언트가 대표로 생성하고, 나머지 클라이언트는 생성된 composite를 불러와 이용한다. 이후에 각 클라이언트는 본인의 개별 hubPort를 생성해 composite에 연결한 후, 본인의 webRTCEndpoint와 HubPort를 연결한다. 이로써 WebRTCEndpoint와 pipeline 사이의 uplink 그리고, HubPort와 WebRTCEndpoint간의 downlink, 총 2개의 링크만으로 다대다 통신을 할 수 있는 시그널링 서버를 구축하였다.

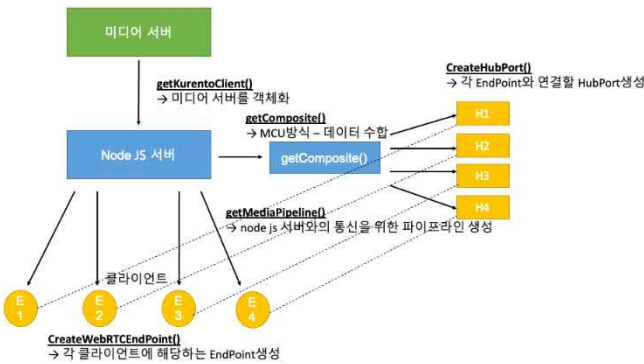


그림 3. 시그널링 서버 구조 개요도

### 3. 결론 및 향후 연구



위 사진은 상기 명시한 대로 구현한 결과 화면으로, 실시간 CCTV 화면 모니터링과 화상통화 기능을 제공하며, pc와 모바일 환경에서 크롬 브라우저를 기준으로 구현하

였다. Start/Stop으로 클라이언트 본인의 스트림을 On/Off 할 수 있고 IpCam Start/IpCam Stop으로 CCTV 영상을 On/Off 할 수 있다.

이로써 WebRTC 기반 실시간 영상 공유 및 모니터링 시스템의 프로토타입 구현에 성공하였으며 WebRTC의 장점인 높은 실시간성과 별도의 플러그인을 필요치 않는 것은 큰 강점이라고 할 수 있다.

### 참고문헌

- [ 1 ] 배영임, 신혜리 “코로나19, 언택트 사회를 가속화하다”, 『이슈&진단』, 416호, p.10
- [ 2 ] “Welcome to Kurento”  
<https://dockurento.readthedocs.io>
- [ 3 ] Thomas Sandholm, “SnoW: Serverless n-Party callsover WebRTC”, June 28, 2022

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음”(2021-0-01082)

“This research was supported by the Korean MSIT (Ministry of Science and ICT), under the National Program for Excellence in SW(2021-0-01082) supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)“(2021-0-01082)



# 아두이노 압력센서 데이터 기반 딥러닝 자세 예측 의자

김주형<sup>o</sup>, 심규혁, 우용태  
창원대학교 컴퓨터공학과

wngud0711@naver.com, shimguh@naver.com, ytwoo@sarim.changwon.ac.kr

## Posture Prediction Chair using Deep Learning based on Arduino Pressure Sensor Data

Ju-Hyeong Kim<sup>o</sup>, Gyu-Hyeok Shim, Yong-Tae Woo  
Department of Computer Engineering, Changwon National University

### 요약

국내에서 자세 교정 및 학습 시스템은 인식이나 사용 방법이 매우 미미한 수준이다. 기존 접촉식이 아닌 비접촉식으로 자세 분석 기술은 정확도 면에서 아직 많은 기술 발전이 요구된다. 사람이 앉은 자세는 수학적으로 고차원이며, 동일한 동작을 하더라도 착석자의 몸무게, 키 등 스펙을 고려해서 자세 인식과 관련된 탐색 및 최적화 과정을 수행하는데 기술적으로 부족하다. 이에 본 논문에서는 착석자의 앉은 자세 데이터를 저장하고 인공지능 딥러닝을 통해 최적의 올바른 앉은 자세의 모델을 학습할 수 있는 서비스를 개발하였다.

## 1 서론

작업 및 학업의 환경이 변화하면서 서있거나 움직이는 활동보다 앉아서 생활하는 시간이 증가하게 되었다. 오랜 시간 컴퓨터를 사용하는 사무직 종사자나 컴퓨터 게임을 즐기는 사람들, 현대인들에게 척추측만증이나 허리 디스크 등의 문제가 증가하고 있다. 이는 생활 습관과 나쁜 자세 등의 환경적인 요인이 원인으로 파악되며 특히, 성장기 청소년들에게서 척추측만증 비율이 빠르게 증가하고 있다.

척추측만증의 원인은 대부분 의자에서 생활하는 인원들이 잘못된 자세로 의자에서 생활하여 발병하게 되는 것이 원인이다. 척추측만증의 환자는 대부분 책상과 의자에서 활동하는 시간이 많은 10대 청소년으로 구성되어 있고, 또한 오랜 시간 앉아서 근무하는 버스나 택시기사 등의 직업군에서도 많이 발병하고 있다.

의자에서 활동하는 시간이 길어지면서 의자에 잘못된 자세가 아닌 바른 자세로 앉는 것의 중요성이 점차 커지고 있다[1]. 하지만 대부분의 사람들은 의자에 앉을 때 잘못된 자세로 앉아있는 것을 인지하지 못한 상태로 생활하게 된다.

본 논문에서는 압력센서를 사용하여 현대인의 자세 불안정의 심각성을 인지시켜 주기 위해 6가지 자세로 실험을 진행하였다. 자세를 인식하고 여러가지 자세를 알려주며, 딥러닝으로 압력센서를 기반으로 한 자세를 규정하여 바른 자세와 잘못된 자세를 판단한다. 또한, 받은 데이터의 통계 값을 시각화를 통해 전달하고 잘못된 자세의 비율이 높을 경우 위험성을 인지시키고 사용자의 평소 습관을 확인 및 개선이 가능하도록 한다.

## 2 배경지식

### 2.1 아두이노

아두이노[2]는 오픈 소스를 기반으로 한 단일 보드 마이크로컨트롤러로 완성된 보드(상품)와 관련 개발 도구 및 환경을 말한다. 아두이노는 다수의 스위치나 센서로부터 값을 받아들여, LED나 모터와 같은 외부 전자 장치들을 통제함으로써 환경과 상호작용이 가능한 물건을 만들어 낼 수 있다. 임베디드 시스템 중의 하나로 쉽게 개발할 수 있는 환경을 이용하여, 장치를 제어할 수 있다. 아두이노 통합 개발 환경(IDE)을 제공하며, 소프트웨어 개발과 실행코드 업로드도 제공한다. 또한 어도비 플래시, 프로세싱, Max/MSP와 같은 소프트웨어와 연동할 수 있다. 오픈소스이기 때문에 아두이노를 기반으로 여러 가지 프로젝트를 수행할 수 있다.

### 2.2 딥러닝

심층학습 또는 딥러닝(deep structured learning, deep learning, hierarchical learning)[3]은 여러 비선형 변환기법의 조합을 통해 높은 수준의 추상화(abstractions, 다량의 데이터나 복잡한 자료들 속에서 핵심적인 내용 또는 기능을 요약하는 작업)를 시도하는 기계 학습 알고리즘의 집합으로 정의되며, 큰 틀에서 사람의 사고방식을 컴퓨터에게 가르치는 기계학습의 한 분야라고 이야기할 수 있다.

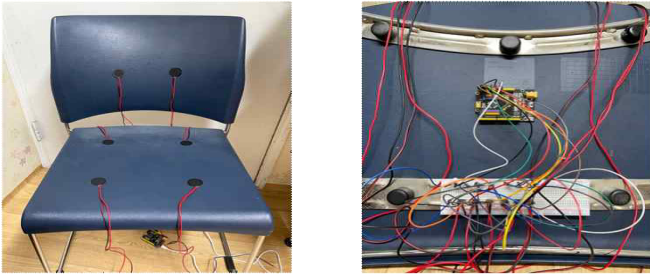
심층 신경망(Deep Neural Network, DNN)은 입력층(input layer)과 출력층(output layer) 사이에 여러 개의 은닉층(hidden layer)들로 이뤄진 인공신경망(Artificial Neural Network, ANN)이다. 심층 신경망은 일반적인 인공신경망과 마찬가지로 복잡한 비선형관계(non-linear relationship)들을 모델링할 수 있다. 예를 들어, 물체 식

별모델을 위한 심층 신경망 구조에서는 각 물체가 영상의 기본적 요소들의 계층적 구성으로 표현될 수 있다. 이때, 추가 계층들은 점진적으로 모여진 하위 계층들의 특징들을 규합시킬 수 있다. 심층 신경망의 이러한 특징은 비슷하게 수행된 인공지능망에 비해 더 적은 수의 유닛(unit, node)들만으로도 복잡한 데이터를 모델링할 수 있게 해준다.

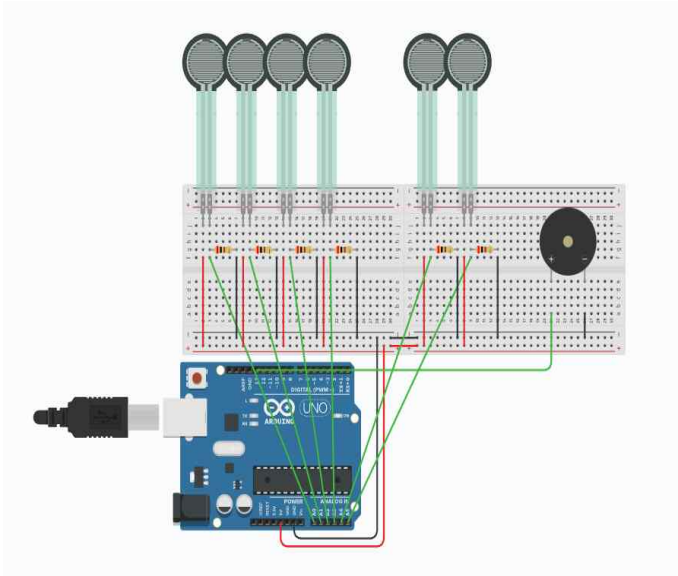
### 3. 구현 및 결과

#### 3.1 제작 의자

의자에 압력센서, 부저, 아두이노를 부착하였다.



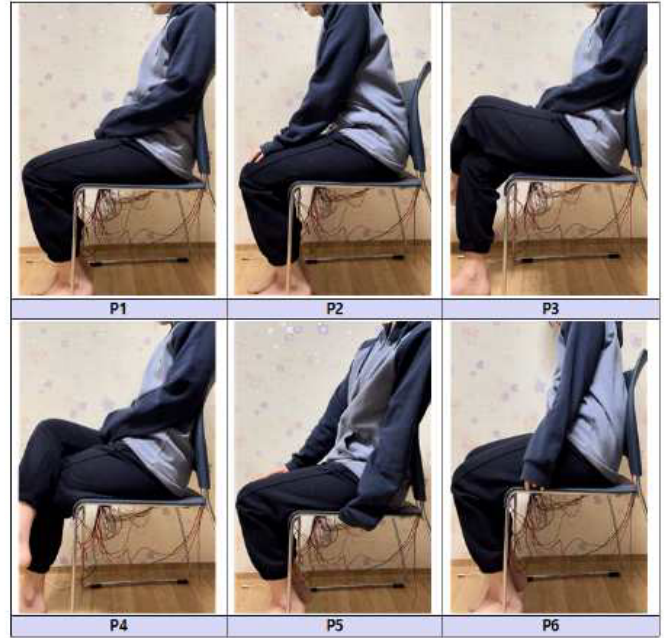
의자의 등 부분과 받침 부분에 각각 동등한 비율로 등분하여 등 부분에 2개, 받침 부분에 4개의 압력센서를 부착하여 아두이노(Arduino)로 압력값을 얻는다. 압력센서는 RA30P 압력센서, 부저는 수동형 피에조 부저, 아두이노(Arduino)는 우노(Uno) 보드[4]를 사용하였다.



#### 3.2 지정 자세

실험을 위해 P1~P6의 자세를 임의로 정의하여 차례로 자세를 바꿔가며 압력을 측정한다. P1은 정자세이고, P2는 상반신을 앞으로 기울인 자세, P3는 왼쪽 다리를 꼰 자세, P4는 오른쪽 다리를 꼰 자세, P5는 몸을 왼쪽으로 기울인 자세, P6은 몸을 오른쪽으로 기울인 자세로 정의하였다. 단, P0은 아무도 앉지 않았을 때를 말한다.

1.



2.

#### 3.3 데이터 수집

데이터 수집 방법은 P1~P6의 자세를 차례로 바꿔가며 측정하게 되고 자세를 바꾸는 동안은 의자에서 떨어지지 않고 자세를 모두 취하고 나면 일어난 후 다시 앉아서 자세를 취한다. 이를 10회 반복한다. 이렇게 얻은 압력값을 모두 사용하지 않고 안정적인 부분의 압력값만 사용한다.

3. 좌 앞	4. 우 앞	5. 좌 뒤	6. 우 뒤	7. 등 좌	8. 등 우	9. 자 세
10.95 6	11.96 2	12.89 3	13.93 6	14.88 5	15.71 9	P1
16.96 9	17.95 4	18.96 0	19.98 9	20.0 0	21.0 6	P2
22.0 0	23.98 0	24.10 23	25.96 8	26.95 0	27.71 6	P3
28.95 0	29.0 0	30.85 0	31.10 19	32.90 3	33.80 1	P4
34.88 5	35.90 6	36.87 1	37.96 0	38.82 2	39.53 9	P5
40.91 6	41.91 2	42.90 7	43.91 6	44.68 4	45.75 1	P6

자세별 압력 데이터 예시를 보면 몸을 기울이거나 다리를 꼰 자세 등 특정 압력센서에서는 값이 들어오지 않는 것을 볼 수 있다.

총 30명을 대상으로 실험하여 전체 5,598개의 데이터를 얻을 수 있었다.

46.

#### 3.4 실험 결과

47. 총 30명을 대상으로 실험하여 얻은 전체 5,598개의 데이터 중 75%인 약 4,200개로 학습하고 25%인 약 1,400개로 테스트하였다.



51. 테스트 결과 98%의 정확도를 보였고 P1-P6의 자세 중 하나의 자세로 예측할 수 있었다.

52. 1399/1399 [=====] - 12s 8ms/step - loss: Accuracy = 0.98

#### 4 결론

53. 본 논문에서는 앉은 자세의 압력값을 측정하기 위한 센서 시스템을 개발하였다. 이에 의해 착석자의 앉은 자세가 올바른 자세인지 아닌지 판별하고 올바르지 않을 경우 부저가 울리도록 하였다. 본 논문에서는 6개의 압력 센서를 사용하여 앉은 자세의 압력값을 측정하고 이를 통해 올바른 자세인지 올바르지 않은 자세인지 판별할 수 있도록 하였다.

54. 실험 결과, 임의로 지정한 6가지의 자세는 예측할 수 있었고, 올바른 자세(P1)에서는 부저가 울리지 않았고, 올바르지 않은 자세(P2~P6)에서는 부저가 울렸다. 또한 임의로 지정한 자세 외의 다른 자세들도 6가지 자세 중 하나로 인식되어 개선이 필요함을 느꼈다.

사람마다 신체적 특성이나 의자에 앉는 자세가 다르기 때문에 많은 사람들이 의자를 이용하여 자신에게 맞는 올바른 자세를 찾기를 기대하고 압력센서의 성능과 개수가 많아지면 보다 정밀한 압력값으로 실험하여 더 정확한 결과가 나올 수 있을 것이라고 예상하고 자세에 따른 부저 알람의 소리를 변경하여 착석자가 자신이 앉은 자세의 종류까지 알게 하여 자세교정에 더 많은 도움이 될 수 있다.

#### 참고문헌

[1] 박진석, “의자에 부는 4차산업혁명... IoT 의자”, <https://www.e-patentnews.com/4744>  
 [2] 아두이노, <https://ko.wikipedia.org/wiki/아두이노>  
 [3] 딥러닝, [https://ko.wikipedia.org/wiki/딥\\_러닝](https://ko.wikipedia.org/wiki/딥_러닝)  
 [4] Uno 보드, <https://m.blog.naver.com/eduardo/220863694128>

# 게임에서 NFT의 활용

김정훈<sup>o</sup>, 이수현

창원대학교 컴퓨터공학과

ggl1874@naver.com, suhyun@sarim.changwon.ac.kr

## Using NFT in a Game

Jeong-Hoon Kim<sup>o</sup>, Su-Hyun Lee

Department of Computer Engineering, Changwon National University

### 요약

NFT와 게임을 접목시킴으로써 특정 게임 내에서만 존재하던 아이템이나 재화를 NFT라는 방식을 통해 다른 게임이나 다른 공간에서도 그 아이템이나 재화를 소유하고 가질 수 있기에 NFT와 게임을 결합시킨 서비스를 제공하는 게임이 많아질수록 한 게임에서 얻은 아이템을 다른 게임에서 사용할 수 있고 거래를 통해 기존에 서비스하던 게임이 서비스를 종료하더라도 그 NFT만의 가치와 기능을 가지고 여러 생태계를 만들 수 있다. 이는 게임을 이용하는 게이머에게도 기존의 게임에서 즐기는 오락의 즐거움에 더해 자신만의 고유한 가치를 지닌 NFT를 직접 게임을 통해 만들고 소유하며 거래하는 즐거움을 더해 더욱 게임에 흥미와 관심을 가지게 한다.

## 1 서론

현재 전 세계적으로 NFT의 인기가 증가하고 있으며 NFT관련 시장이 매우 커지고 있는 상태이다. 급격한 성장을 이루고 있는 NFT 시장에서 ‘NFT가 익숙하지 않은 사람들이 NFT라는 디지털 자산에 자연스럽게 관심을 가지고 참여하며 익숙해질 수 있을까?’ 라는 질문에 대한 하나의 결론으로 사람들이 익숙하지 않은 NFT를 사람들이 익숙한 개념안에 접목시켜서 자연스럽게 알 수 있도록 결합시켜보는 방법을 택하게 되었다. 사람들이 익숙한 개념과 NFT가 연결되기에 가장 적합한 항목을 찾아보던 와중 P2E(Play to Earn)이라는 개념을 알게 되었다. 이는 ‘게임을 하면서 돈을 번다’ 라는 뜻이다. 이는 게임 내 재화, 게임 내 아이템을 현실의 가치와 교환하여 게임 내에서만 가치가 존재하던 것들을 현실의 재화와 교환함으로써 가치의 증폭을 일으키고 게임을 단순히 즐기는 오락의 용도에서 벗어나 현실 재화를 벌기위한 수단으로 사용할 수 있도록 만든 개념이다.

현재 게임 내 재화와 아이템을 일명 현금화를 시키기 위해서는 개인간의 거래를 통하여 교환을 하여야하고 사기를 당할 위험이 도사리고 있다. 여기에서 아이디어를 얻어 게임 내 재화나 아이템을 NFT화 시키면 블록체인 상에서 안전하게 아이템을 거래할 수 있고 이를 바로 현금화할 수 있기에 NFT라는 개념이 익숙하지 않은 사람들도 익숙한 게임을 하면서 그 게임 내 재화를 현실의 재화로 안전하게 변환시키고 거래를 통해 그 가치를 인정받는 NFT의 개념에 더 쉽게 다가갈 수 있을 것이라 생각하였다.

## 2 배경지식

### 2.1 NFT

NFT란 Non-Fungible Token의 약자이며 대체불가능한 토큰이라는 의미이다. NFT는 블록체인 기술 상에 디지털 정보를 입력하여 보관하는 것으로 각각의 NFT들이 블록체인으로 서로 연결되어 현재로선 해킹이 불가능하고 내용 역시 변경이 불가능하며 소유자 이전 시 그 이력이 기록되기에 디지털 자산으로써는 가장 안전한 형태의 자산이라고 할 수 있다. 각각의 NFT 안에 저장된 기록물의 내용이 모두 같더라도 각각의 NFT는 각각 고유한 가치를 지니며 대체가 불가능하기에 같은 내용의 NFT라도 서로의 가치가 다르고 블록체인에 기록된 순서가 다르기에 따로 취급하며 각각이 서로 다른 가치와 자산으로 운용되어진다.

### 2.2 P2E

P2E란 Play to Earn의 약자로서 게임을 구분하는 여러 가지의 수많은 형태 중에 하나이다. 어떤 게임의 특성이 MMORPG이던 FPS이던 RTS등 어느 것이라고 해도 그 게임의 서비스 형태에 P2E 게임이라고 붙일 수 있다. 비슷한 용어로 PVP, PVE 등이 있다. 이들 역시 게임의 특성을 구별해주는 용어이나 사용법이 다르다. 예를 들어 인기 게임인 롤(LOL)의 경우 MOBA 게임으로 TeamPlay, RTS, RPG를 합한 개념이며 전투방식은 PVP이다. 이런 식으로 어떤 P2E 게임이 있다면 그 게임의 예시는 MMORPG 형태의 게임이며 PVE 방식을 채택하고 있으며 이 게임은 P2E게임이다.

### 2.3 핑크랜드

핑크랜드(전. 네코랜드)란 핑크랜드 회사 자체의 소개를 인용하자면 ‘게임 개발자와 NFT 제작자가 풍부한 메타버스 경험을 만들 수 있도록 지원하는 Web3 게임 플랫폼입니다. 또한 플레이어는 NFT 자산을 사용하여 연속성을 가지고 플레이하는 새로운 방법을 제공하는 2D MMO 세계에 액세스 할 수 있습니다.’ 라고 설명하고 있다. 핑크랜드는 게임 개발자에게 게임 개발에 필요한 Studio를 자체적으로 제공해주고 있다. Studio 안에는 자체적인 DB와 Server를 메인 핑크랜드 DB와 Server에서 트래픽을 공유하여 제공해주고 있다. 게임 개발 툴 역시 Interface가 직관적이고 이미지화가 되어있어서 초보 개발자들도 손쉽게 접근하여 게임 개발이 용이하다는 장점을 가지고 있다. 또한 제작한 게임들을 핑크랜드 자체적으로 서비스해주는 게임 유통망도 가지고 있는 회사이다. 즉, 초보 게임 개발자들이 손쉽게 게임을 개발하고 게임을 배포하기에 굉장히 접근성과 편리성을 갖춘 회사인 것이다. 하지만 회사 자체의 인지도가 떨어지고 커뮤니티 구성이 활발하지 않은 상태이므로 게임 개발과 리소스 공유에 있어서 접근성이나 정보의 다양성이 부족하다는 점이 존재한다.

### 2.4 메타마스크와 오픈씨

NFT 역시 디지털 종류의 한 자산이다. 자산은 보관을 해야하며 시장경제에 따라 형성된 가치에 맞춰 수요와 공급이 진행되어야 한다. 그러기 위해서는 보관을 위한 지갑이 있어야하며 거래를 위한 거래소가 있어야 한다. 현재 전 세계적으로 NFT의 보관과 거래를 위한 지갑 및 거래소 사이트가 많이 있지만 본 논문에서는 ‘메타마스크(MetaMask)’ 라는 NFT 지갑과 ‘오픈씨(OpenSea)’ 라는 NFT 전시 및 거래 사이트를 사용하였다. 메타마스크는 이해하기 쉽게 표현하자면 디지털 지갑이다. 자신의 디지털 자산을 인터넷 상에 안전하게 보관하도록 서비스하는 지갑으로 자신의 지갑 주소를 통해 NFT 물품들의 거래가 이루어지고 보관이 가능하다. 또한 가상화폐인 일명 코인들을 자신의 지갑으로 보내 보관하거나 NFT 거래에 쓰는 등 가상화폐를 인터넷 상에서 실질적인 화폐로 사용할 수 있다.

오픈씨는 자신이 소유한 NFT를 전시하거나 거래를 하기 위한 거래소의 서비스를 제공하는 사이트로 국내에서 가장 유명한 NFT 거래소 사이트이다. 오픈씨 사이트와 자신의 지갑과 연동을 시키면 자신이 보유한 NFT 항목들을 볼 수 있으며 그 항목들을 OpenSea 사이트에서 다른 이용자가 관람할 수 있도록 전시할 수 있다. 또한 ‘Sell’ 항목을 이용하여 본인이 원하는 코인을 선택한 후 가격을 책정하여 판매를 할 수도 있다. 하지만 모든 코인을 거래에 사용할 수 있는 것은 아니고 제공되는 코인들로만 거래가 가능하다. 가장 많이 쓰이는 코인은 ‘이더리움(Ethereum)’ 이며 점점 서비스 코인이 늘어가는 추세이다.

## 3. 개발 내용

주요 구현 기능으로는 게임 내 유저와의 대화와 공지 및 알림을 표시하기 위한 ‘대화 Text창’ 과 채팅 및 이모티콘을 입력하기 위한 ‘대화 시작 창’, 그리고 캐릭터의 상태를 나타내는 ‘HP, MP Bar’, ‘Level창’, ‘경험치 Bar’ 가 있으며 그 외적인 요소들을 파악하기 위한 ‘주요 옵션’ 창이 구현되어 있다.



다음 그림에서 우측에 있는 아이템 창이 게임 내에서 쓸 수 있는 아이템에 대한 전체적인 인벤토리이고 왼쪽에 있는 ‘핑크랜드 NFT 아이템’ 창이 NFT 아이템을 NFT 창고에 저장하여 NFT화 및 거래에 사용할 수 있는 인벤토리 창이다.



다음 그림에서 왼쪽 NFT 창고에 NFT 아이템이 들어가 있는 것을 확인할 수 있다. 해당 창고에는 모든 아이템이 들어가는 것이 아니라 오직 NFT화가 가능하도록 DB가 제작된 아이템만이 들어갈 수 있다. 해당 창고에 들어간 아이템은 핑크랜드에서 제공하는 ‘핑크랜드 Wallet’ 에서 NFT화 및 아이템화를 시킬 수 있으며 OpenSea 사이트를 통해 거래 및 전시가 가능하다.



핑크랜드에서 제공하는 온라인 지갑에 본인의 메타마

스크 지갑을 연동하면 인게임에서 넣은 NFT 창고 아이템을 확인할 수 있으며 아이템의 NFT화 및 재아이템화를 진행할 수 있다. 펑크랜드 지갑에서 아이템을 NFT화시키면 본인의 메타마스크 지갑과 연동된 오픈씨 사이트에서 아이템이 확인 가능하다.

게임 내 아이템이 NFT화 되어서 거래가 전시 및 거래가 가능하고 NFT화 된 아이템은 삭제 및 게임 내 유저간 이동을 막기위해 인게임에서 사용이 불가하게 된다. 하지만 다시 NFT 아이템을 재아이템화 한다면 정상적으로 인게임 내에서 기존의 아이템으로써의 기능을 동작하게 된다.

## 4 결론

본 논문에서는 NFT와 게임의 결합을 이용한 P2E 게임을 개발하였다. 게임과 NFT를 엮는데 성공하였고 인게임의 아이템을 NFT화 시켜 저장하는 데에도 성공하였다. 게임 개발과 NFT의 연동에 있어 P2E 게임 개발에 펑크랜드 엔진을 사용하였고 제작된 게임 역시 '2022.10월' 기준 펑크랜드 자체 게임 유통망에 'OpenBeta' 로 등록되어 누구나 이용이 가능한 상태이다.

현재 게임의 서비스 상태가 2022년 10월 기준 베타서비스 상태이므로 정식 오픈이 아니기에 만들어진 NFT의 가치를 파악하기에는 무리가 있다. NFT의 가치는 순전히 게임의 퀄리티와 이용자의 수에 달려있기에 지속적인 게임자체의 재미와 다양한 콘텐츠 개발로의 발전이 필요하고 이용자 확보에도 노력을 기울여야한다.

NFT와 게임을 접목시킴으로써 특정 게임 내에서만 존재하던 아이템이나 재화를 NFT라는 방식을 통해 다른 게임이나 다른 공간에서도 그 아이템이나 재화를 소유하고 가질 수 있기에 NFT와 게임을 결합시킨 서비스를 제공하는 게임이 많아질수록 한 게임에서 얻은 아이템을 다른 게임에서 사용할 수 있고 거래를 통해 기존에 서비스하던 게임이 서비스를 종료하더라도 그 NFT만의 가치와 기능을 가지고 여러 생태계를 만들 수 있을 것으로 기대한다. 이는 게임을 이용하는 게이머에게도 기존의 게임에서 즐기는 오락의 즐거움에 더해 자신만의 고유한 가치를 지닌 NFT를 직접 게임을 통해 만들고 소유하며 거래하는 즐거움을 더해 더욱 게임에 흥미와 관심을 가지며 애정을 느낄 수 있게 될 것이라 생각한다.

## 참고문헌

- [1] 펑크랜드, <https://docs.punkland.io/>
- [2] 메타마스크, <https://metamask.io/>
- [3] OpenSea, <https://opensea.io/>
- [4] 장문경, "P2E(Play-to-Earn) 게임 지속이용의도에 대한 연구." 지식경영연구 23.3, 2022

# Social IoT traffic prediction

Dong Hu

The Department of Data Convergence Computing  
Kyungpook National University  
Daegu 41566, Republic of Korea  
Email: hudong@knu.ac.kr

Limei Peng\*

School of Computer Science and Engineering  
Kyungpook National University  
Daegu 41566, Republic of Korea  
\* : corresponding author  
Email: aurorapl@knu.ac.kr

**Abstract**—Prediction of Social Internet of Things (SIoT) data traffic helps understand daily user behavior and helps administrators manage devices. One of the most popular traffic prediction methods in noisy environments are Kalman filtering (KF), which is very simple and versatile. However, KF requires a dynamic traffic and measurement model, which introduces additional overhead and is often difficult to obtain. It is difficult to obtain. In contrast, deep learning models in contrast, deep learning models with recurrent neural network (RNN) structure have been widely used for dynamic flow modeling. They have been widely used in modeling dynamic models that evolve over time. On the other hand, the long short-term memory network LSTM (long short-term memory) is a variant of RNN which avoids the gradient disappearance of RNN, which overcomes the effect of short term memory. Therefore, it is more promising in predicting SIoT data traffic. In this paper, we propose an LSTM-EKF, which is a new deep learning unit that uses the LSTM to predict the system parameters of the extended Kalman filter (EKF). Note that the EKF is suitable for predicting SIoT traffic in a nonlinear, noisy environment. The proposed LSTM-EKF can improve the performance of time series data prediction. Prediction of nonlinear SIoT data traffic. Numerical experiments were conducted to evaluate the SIoT traffic prediction performance of the proposed LSTM-EKF method. The performance of the proposed LSTM-EKF method is evaluated on a real dataset. Wikipedia web access traffic data. The proposed method shows better performance than existing prediction methods in terms of metrics of root mean square error (RMSE) and determination coefficient ( $R^2$ ).

**Index Terms**—Extended Kalman filter (EKF); normalized mean square error (NMSE); time-series prediction; LSTM;

## I. INTRODUCTION

The Internet of things (IoT) is an interconnected system of computing devices, machinery, and digital machines that digitize the real world. The IoT has already affected people's lives, including transportation, housing, food, clothing, health, and remote monitoring. Many home appliances can be controlled through mobile phones and voice. Many applications allow users to improve their quality of life and even enable the elderly and the disabled to live more conveniently. MGI's report shows that starting from 2025, the Internet of Things will create an output value of 3.9 trillion to 11.1 trillion US dollars in nine environments, including factories, retail, and cities, and the number of IoT devices is expected to grow to 754 100 million, which is equivalent to adding 127 IoT devices every second in the world starting in 2020 [1]. In the next generation of IoT, the objects are integrated with our social dimension, making them smart and social objects.

Social Internet of Things (SIoT) is a new concept of combining such social network relations with IoT [2]. Under such development, we have found a trend of social relationships explosion which includes device-device relationships, user-user relationships, and user-device relationships. To realize the social interactions, objects should start establishing social relationships based on the object profile, shared hobbies, and social activities. As shown by Fig. 1, these social interactions can be classified into the following types: a 'co-location' relationship, i.e., objects are close in geographical location; a 'co-work' relationship, i.e., objects work together for some work; a 'parental' relationship, i.e., objects have common device attributes, like devices model; a 'social' relationship, i.e., objects are associated by social interactions; and a 'co-ownership' relationship, i.e., objects are owned by the same users.

However, the increasing number of these relationships and their heterogeneous social attributes have introduced various computing and communication challenges that prevent the IoT network from taking advantage of these relationships to improve the offered services and personalize the delivered content. It is helpful to forecast the SIoT data traffic to predict the above relationship and social behavior. For example, a network traffic-aware mobile application recommendation system based on social network interaction traffic is proposed [3]. We also can understand further about users' social behaviors by predicting web traffic generated by a user [4]. However, there are very few works dedicated in this direction.

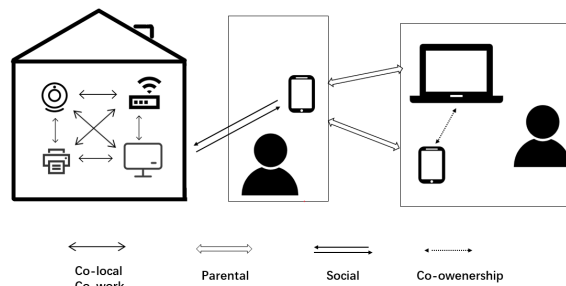


Fig. 1. Types of social relationships among objects.

Autoregressive Integrated Moving Average (ARIMA) is a common machine learning approach [5] used for traffic prediction. In [6], the authors try to combine ARIMA prediction scheme with Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) and show that it provides better prediction accuracy than ARIMA method. In order to make prediction over long-range, ARIMA is also combined with the Markov Modulated Poisson process (MMPP) [7]. All these conventional models require a large extent on the stationary data used for prediction and the prediction time range. Therefore, these models are only able to be trained by a particular time series, and they are not easy to be generalized to predict a set of time-series with a single model training. Deep learning based models are proper to solve this problem since deep learning methods are proper to deal with nonlinear characteristics of traffic data. Recurrent Neural Network (RNN) is first being applied to network traffic prediction by work [8]. RNNs are particularly proper for tasks that involve time series. In practice, Long short-term memory (LSTM) model is often used, as opposed to a vanilla (or standard) RNN, because it is more computationally effective. In fact, the LSTM was introduced to solve a problem that standard RNNs suffer from [9].

Another problem facing the SIoT traffic prediction problem is how to make predictions in noisy environments. Because the above mentioned machine learning or deep learning based methods must assume that all training data are clean enough to extract traffic information, however, this assumption is not practical in real SIoT environments. Extended Kalman filtering (EKF) is an important method for learning the true state based on noisy observations, and its application in traffic prediction can be found in [10]. Based on these observations, this paper proposes the use of the extended Kalman filter (LSTM-EKF), a new deep learning unit suitable for predicting nonlinear SIoT traffic in noisy environments. To verify the effectiveness of the proposed LSTM-EKF, numerical experiments are conducted to evaluate the SIoT traffic prediction performance of the proposed new LSTM-EKF method on real datasets. Wikipedia web access traffic. The proposed method has better performance than the existing prediction methods by using the following performance metrics: Root Mean Square Error (RMSE), and coefficient of determination ( $R^2$ )

## II. SYSTEM MODEL

In this section, we will review construction components required to build the proposed LSTM-EKF model. The Kalman filter (KF) and Extended Kalman filter (EKF) will be introduced in Section II-A. The neural unit LSTM will be reviewed in Section II-B.

### A. LSTM

LSTM is a special type of recurrent neural network (RNN) designed to solve the gradient disappearance and gradient explosion problems during long sequence training. LSTM has a recursive structure similar to RNN, but unlike RNN, LSTM uses four processing modules that interact in a special way to

achieve long term dependency as shown in Fig. 2. Specifically, each neuron in LSTM is a memory cell containing an input gate, a forget gate, and an output gate. These gates control the save, read and write operations of the memory cell. LSTM neural networks are widely used in time series predictions such as predictions for electricity consumption, traffic flow, air quality, etc.

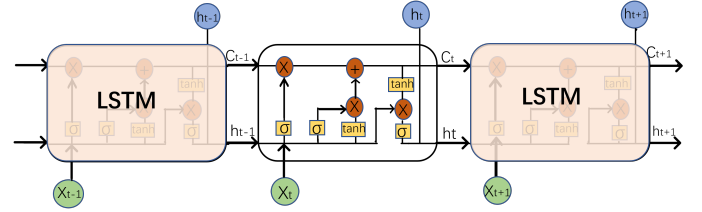


Fig. 2. LSTM network model.  $X_i$ : input time series at moment  $i$ ;  $h_i$ : the hidden state of memory cells at moment  $i$ ;  $C_i$ : the cell state at moment  $i$ ;  $\sigma$  &  $\tanh$ : neural network layers for learning;  $\times$  &  $+$ : point-by-point multiplication & summation.

In LSTM, the forget gate layer using the sigmoid function determines the information to discard based on the node states. Specifically, it outputs a fraction value in  $[0,1]$  based on  $h_{t-1}$  and  $x_t$  for each cell state  $c_{t-1}$ , as shown in Fig. 2. The extreme values of one and zero indicate maintaining the complete information and discarding the complete information, respectively. The forget gate output is denoted by

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

where  $\sigma(\cdot)$  is the sigmoid function,  $x_t$  denotes the input at time  $t$ ,  $h_{t-1}$  is the output of the last neuron structure, and  $[h_{t-1}, x_t]$  denotes the operation of joining two vectors into a longer one;  $W_f$  and  $b_f$  denote the weight and bias associated with the forget gate, respectively.

The next step is to determine what new information to store in the unit node state. Regarding this, the input gate layer applying the sigmoid function first determines the values to be updated. Then, the  $\tanh$  layer creates a vector of new candidate values, i.e.,  $\tilde{c}_t$ , to be added to the state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

where  $i_t$  is the portion of new information retained at moment  $t$ ,  $W_i$  and  $W_c$  are the weight matrices, and  $b_i$  and  $b_c$  are the bias vectors for  $i_t$  and  $\tilde{c}_t$ , respectively. Then we combine the results obtained from (2) (3) above to update the old cell state  $c_{t-1}$  to a new one  $c_t$  by Eq. (4):

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (4)$$

Finally, the output  $h_t$  is obtained based on the cell state  $c_t$ . Specifically, we run a sigmoid layer to determine the parts of the cell state to output and then place the  $c_t$  in  $\tanh$  to restrict



the value in a range of  $[-1, 1]$ , as shown in Eqs. (5) and (6):

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (6)$$

Note that LSTM uses a back propagation algorithm during training. When we get a predicted value, the loss value between this predicted value and the true value is passed to the neural network by back propagation, and then the gradient of the loss value is used to update the parameters of the neural network for weight update.

### B. Kalman Filter and Extended Kalman Filter

For dynamic and control theory, Kalman filtering, is an algorithm that uses a series of measurements observed over time, including background noise and other inaccuracies, and produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone, by estimating a joint probability distribution over the variables for each timeframe. The filter is named after Rudolf E. Kálmán, who was one of the primary developers of its theory.

Mathematically, if we represent our state, i.e., the real-time state data at time  $t$ , as vector  $x_t$ , and our measurement, i.e., the error data obtained through observation at time  $t$ , as  $z_t$ , the state dynamics of KF can be formulated as

$$x_{t+1} = \Phi x_t + w_t, \quad (7)$$

where  $w_t$  is the associated white noise process with known covariance  $Q$ . The observation dynamics of KF can be formulated as

$$z_t = Hx_t + v_t. \quad (8)$$

The covariances of the two noise models are assumed stationary over time and are given by

$$Q = E[w_t w_t^T], \quad (9)$$

and

$$R = E[v_t v_t^T]. \quad (10)$$

The error covariance matrix at time  $t$  can be defined as

$$P_t = E[(x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})^T], \quad (11)$$

where  $\hat{x}_{t|t}$  is the estimated state at time  $t$ . We assume that the prior estimate of  $\hat{x}_{t|t}$  is denoted as  $\hat{x}_{t|t-1}$ , the relationship between  $\hat{x}_{t|t}$  and  $\hat{x}_{t|t-1}$  can be expressed as

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (z_t - H\hat{x}_{t|t-1}), \quad (12)$$

where  $K_t$  is the Kalman gain.

The expression  $z_t - H\hat{x}_{t|t-1}$  in Eq. (12) is known as the *innovation residual*, which can be expressed as

$$i_t = z_t - H\hat{x}_{t|t-1}. \quad (13)$$

By applying Eq. (13) into Eq. (12), we will get

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (Hx_t + v_t - H\hat{x}_{t|t-1}). \quad (14)$$

Similarly, by applying Eq. (14) into Eq. (11), we obtain

$$P_t = E \left[ \begin{aligned} & [(I - K_t H) (x_t - \hat{x}_{t|t-1}) - K_t v_t] \\ & [(I - K_t H) (x_t - \hat{x}_{t|t-1}) - K_t v_t]^T \end{aligned} \right] \quad (15)$$

Since  $x_t - \hat{x}_{t|t-1}$  is the error of the prior estimate, we can express Eq. (15) as

$$\begin{aligned} P_t &= (I - K_t H) E \left[ (x_t - \hat{x}_{t|t-1}) (x_t - \hat{x}_{t|t-1})^T \right] \\ &\cdot (I - K_t H) + K_t E[v_t v_t^T] K_t^T \end{aligned} \quad (16)$$

By applying Eqs. (10) and (11) into Eq. (15), we obtain the following:

$$P_t = (I - K_t H) P_t' (I - K_t H)^T + K_t R K_t^T, \quad (17)$$

where  $P_t'$  is the prior estimate of  $P_t$ .

In summary, we have the following update about Kalman filter

$$\begin{aligned} K_t &= P_t' H^T (H P_t' H^T + R)^{-1} \quad (\text{Kalman Gain}) \\ \hat{x}_{t|t} &= \hat{x}_{t|t-1} + K_t (z_t - H\hat{x}_{t|t-1}) \quad (\text{Post. State Update}) \\ P_t &= (I - K_t H) P_t' \quad (\text{Cov. Update}) \\ \hat{x}_{t+1|t} &= \Phi \hat{x}_{t|t} \quad (\text{Pri. State Update}) \\ P_{t+1} &= \Phi P_t \Phi^T + Q \quad (\text{Cov. Update}) \end{aligned} \quad (18)$$

If the state transit relations  $\Phi$  and  $H$  in Eqs. (7) and (8) do not follow linear mapping relations, we have to consider extended KF. For nonlinear mapping relations, EKF can be formulated as

$$x_{t+1} = f(x_t, u_t) + w_t, \quad (19)$$

where  $f(\cdot)$  is a nonlinear function for the previous state and control variable  $u(t)$ . The observation dynamics of KF can be formulated as

$$z_t = h(x_t) + v_t, \quad (20)$$

where  $h(\cdot)$  is a nonlinear function for the observation. To deal with EKF, we can perform the following linearization to get parameters  $\Phi$  and  $H$  in Eqs. (7) and (8) as:

$$\Phi_t \approx \left. \frac{\partial f}{\partial x} \right|_{x_t}, \quad (21)$$

and

$$H_t \approx \left. \frac{\partial h}{\partial x} \right|_{x_t}. \quad (22)$$

### C. The proposed model and the flow chart of the method

We will use LSTM to model  $f$ , by considering input  $z_{t-1}$ , which is the network status vector obtained by variational autoencoder from the network status. After LSTM trains required parameters for EKF, it will be used to generate the

output  $\hat{x}_{t|t}$ . At each time step  $t$ , the LSTM takes the previous prediction  $\hat{x}_{t-1|t-1}$  as input and generates the intermediate state  $f(\hat{x}_{t-1|t-1})$ . Note that the term is independent of the current measurement. The estimate of the process covariance,  $Q_t$  is also updated based on the intermediate state update as the output. Similarly, based on the observations from the network  $z_{t-1}$  an estimate of the measurement covariance  $R_t$  is generated as an output. Finally, the  $f(\hat{x}_{t-1|t-1})$  and  $z_{t-1}$ , along with our covariance estimates  $Q_t$  and  $R_t$  are fed to the extended Kalman filter for updating  $t$  is discussed in Section II-B. The LSTM-EKF expands time, which can be trained end-to-end by backpropagation through time as shown in Fig. 3.

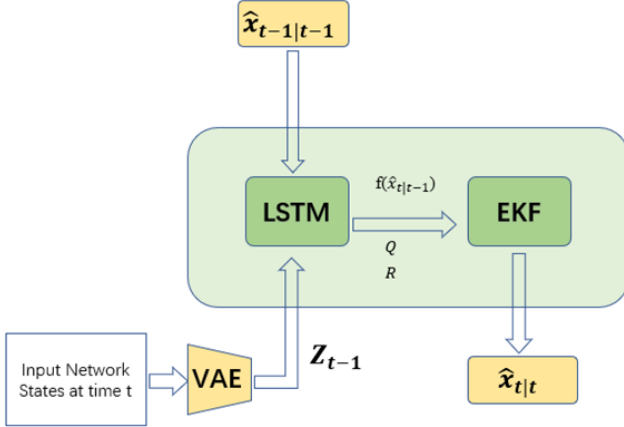


Fig. 3. LSTM-EKF MODLE

### III. NUMERICAL RESULTS

This section will apply the proposed LSTM-EKF model to predict SIoT related datasets. we will apply LSTM-EKF to predict Wikipedia webpages traffic based on the web traffic time series dataset provided by [11]. Following metrics are adopted in our prediction performance evaluation: the root means square error (RMSE) and the coefficient of determination ( $R^2$ ). These metrics can be defined as

$$\begin{aligned} \text{RMSE} &= \sqrt{E((Y - \hat{Y})^2)}; \\ R^2 &= 1 - \frac{\sum_{i=0}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=0}^N (Y_i - E(Y))^2}, \end{aligned} \quad (23)$$

where  $\hat{Y}, \hat{Y}_i$  are predicted values and  $Y, Y_i$  are true (actual) values.

#### A. Dataset

The real dataset under investigation in this work is the Wikipedia traffic [11]. The training dataset consists of approximately 145k time series. Each of these time series represents a number of daily views of a different Wikipedia article, starting from July, 1st, 2015 up until December 31st, 2016. For each time series, you are provided the name of the article as well as the type of traffic that this time series represents (all, mobile, desktop, spider). We compare the proposed LSTM-EKF with

other three prediction methods: EKF and LSTM used in Section Numerical Results by metrics RMSE and  $R^2$  defined by Eq. (23). In table I, we show the prediction performance based on the traffic visiting the AND wikipedia webpage [12]. These results show that the proposed novel LSTM-EKF method has better performance than the other two existing deep-learning prediction approaches in terms of RMSE and  $R^2$ .

For the given dataset, 70% of the data is randomly sampled as sequences for training and the other 30% is sampled for testing. The training dataset is then divided into  $X_{\text{train}}$  and  $Y_{\text{train}}$ , where  $X_{\text{train}}$  contains all the feature values except the target ones to be predicted and  $Y_{\text{train}}$  contains the features that must be predicted. For the proposed LSTM-EKF model using either univariate or multivariate dataset, we use the mean square error (MSE) loss function and an efficient Adam version of stochastic gradient descent. The model will fit 30 training epochs with a batch size of 32. Once the model is fitted, we can predict for the entire test dataset.

#### B. Numerical Results

We have used EKF, LSTM, LSTM-EKF models to predict the Wikipedia data respectively. The prediction results are shown in Fig. 4 and Table. Specifically Figure 4 shows the comparison of our prediction results with the true values. Both prediction plots show that the predicted results of the proposed LSTM-EKF model are more stable than those of the LSTM and EKF models, indicating the superiority of the proposed LSTM-EKF model. The table compares the prediction accuracy of the three models in terms of MSE and  $R^2$ . It can be seen that the LSTM-EKF is more accurate

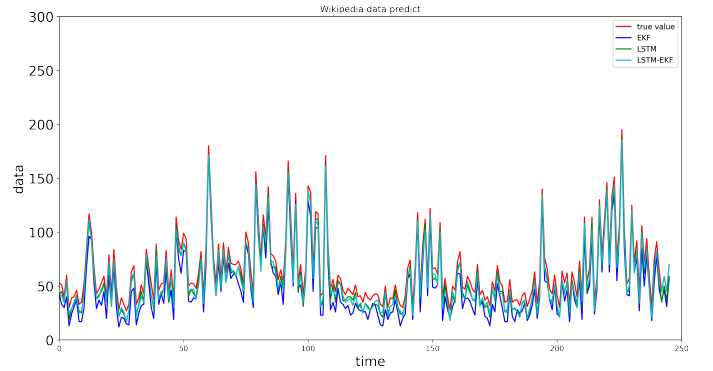


Fig. 4. Prediction comparison of EKF, LSTM and LSTM-EKF schemes on Wikipedia data

TABLE I  
THE ALL EXPERIMENTAL RESULTS OF  $R^2$  AND RMSE INDICATORS

Model	Performance	$R^2$	RMSE
EKF		0.86	4.3
LSTM		0.87	3.5
LSTM-EKF		<b>0.88</b>	<b>1.38</b>

As can be seen in Table 1, the Kalman filter can help improve the prediction accuracy when using the same data set type. For example, the value of  $R^2$  increases from 0.88 to 0.89, and the RMSE decreases when only the LSTM is used. From the observed results, the combination of LSTM and EKF can improve the prediction accuracy for the dataset we used.

Obviously, the proposed LSTM-EKF model shows a significant improvement in prediction accuracy compared to the LSTM, EKF model. We can conclude that Training the required parameters for EKF coupled with the noise reduction of the data by the EKF, greatly improves the prediction effect.our proposed new LSTM-EKF method obviously outperforms the two existing deep-learning prediction methods

#### IV. CONCLUSION

To predict the traffic of SIoT, we propose a new prediction method called LSTM-EKF. The proposed CARU-EKF can help to determine the EKF parameters to perform traffic prediction in noisy environments. We validate the LSTM-EKF by applying We validate the proposed prediction method by applying the LSTM-EKF to a real dataset on SIoT. We validate the proposed prediction method by applying the CARU-EKF to two real datasets on SIoT: The dataset is the dataset of visiting traffic. is the dataset of traffic visiting Wikipedia pages. We use four performance metrics: mean absolute error (MAE), mean absolute error (MAE)

#### REFERENCES

- [1] P. Gokhale, O. Bhat, and S. Bhat, "Introduction to IOT," *International Advanced Research Journal in Science, Engineering and Technology*, vol. 5, no. 1, pp. 41–44, 2018.
- [2] F. Al-Turjman, "5G-enabled devices and smart-spaces in social-IoT: an overview," *Future Generation Computer Systems*, vol. 92, pp. 732–744, 2019.
- [3] X. Su, Y. Zheng, J. Lin, and X. Liu, "A network traffic-aware mobile application recommendation system based on network traffic cost consideration," *International Journal of Computational Science and Engineering*, vol. 19, no. 2, pp. 259–273, 2019.
- [4] K. Xu, F. Wang, and L. Gu, "Behavior analysis of internet traffic via bipartite graphs and one-mode projections," *IEEE/ACM Transactions on Networking*, vol. 22, no. 3, pp. 931–942, 2013.
- [5] H. Z. Moayedi and M. Masnadi-Shirazi, "ARIMA model for network traffic prediction and anomaly detection," in *2008 international symposium on information technology*, vol. 4. IEEE, 2008, pp. 1–6.
- [6] B. Zhou, D. He, Z. Sun, and W. H. Ng, "Network traffic modeling and prediction with ARIMA/GARCH," in *Proc. of HET-NETs Conference*, 2005, pp. 1–10.
- [7] A. Sang and S.-q. Li, "A predictability analysis of network traffic," *Computer networks*, vol. 39, no. 4, pp. 329–345, 2002.
- [8] R. Vinayakumar, K. Soman, and P. Poornachandran, "Applying deep learning approaches for network traffic prediction," in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2017, pp. 2353–2358.
- [9] H. D. Trinh, L. Giupponi, and P. Dini, "Mobile traffic prediction from raw data using LSTM networks," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2018, pp. 1827–1832.
- [10] C. P. Van Hinsbergen, T. Schreiter, F. S. Zuurbier, J. Van Lint, and H. J. Van Zuylen, "Localized extended Kalman filter for scalable real-time traffic state estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 385–394, 2011.
- [11] "Web traffic time series forecast," <https://www.kaggle.com/c/web-traffic-time-series-forecasting>, accessed: 2018-09-30.
- [12] "The and wikipedia webpage," [https://zh.wikipedia.org/wiki/A\\_N\\_D](https://zh.wikipedia.org/wiki/A_N_D).

언제나 한결같은 모습으로  
연구자를 위한 오아시스가 되겠습니다.



**hbn** 고급두뇌를 위한 네트워크  
하이브레인넷



교수/연구원  
채용정보제공



대학원생  
모집정보제공



학술연구정보  
제공



해외연수/  
유학정보제공



연구자  
커뮤니티운영



# (주)큐텍코리아

- 국회 산업통상자원중소벤처기업위원장 표창(2020년)
- 위치정보사업허가기업(제102호), 기술평가우수기업

## O2O Platform Service 전문기업

- PDS (Proxy Driving Service)
- TMS (Truck Management System)

## 생활밀착형 위치기반서비스

- 배달대행, 대리운전, 퀵서비스, 화물배차관제

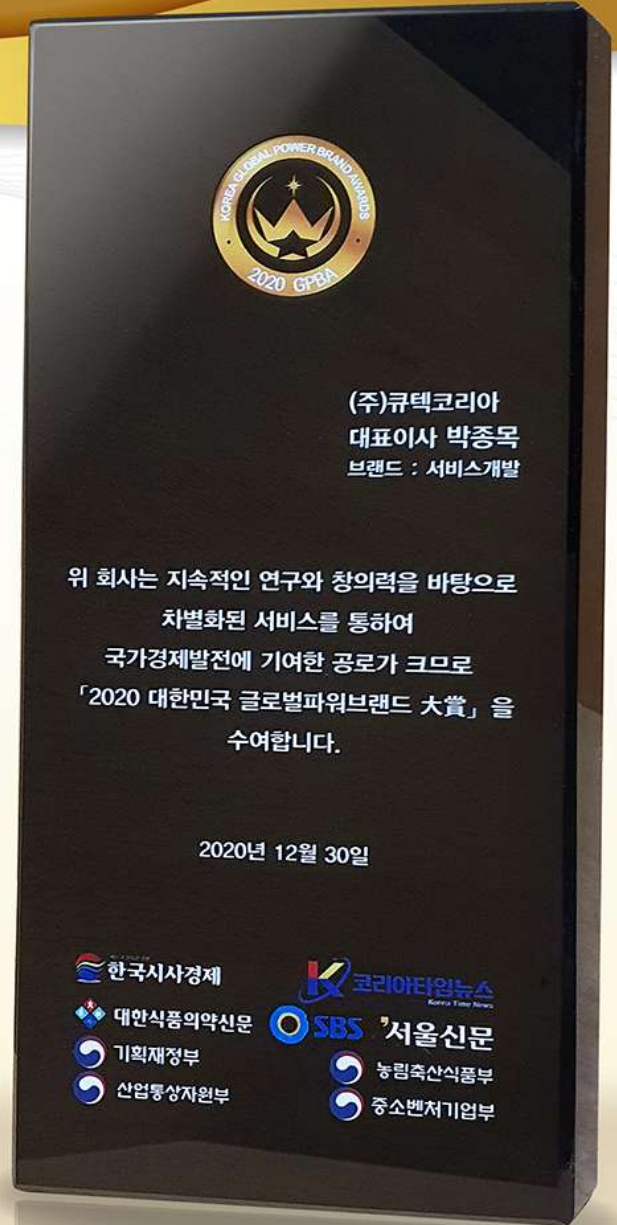
## 기업의 꿈을 이루는 기술

큐텍코리아는 혁신적인 솔루션 기술을 바탕으로 기업의 성공을 이끌어드립니다.

Location-Based Service, AI-Big Data • PC,POS,Mobile integrated development solutions • SMS solution



① <http://www.qtec.kr>    ② [qtec@qtec.kr](mailto:qtec@qtec.kr)  
 ③ 대구광역시 북구 경대로 17길 47 IT융합산업빌딩 401호  
 ④ Tel. 053-622-8572    ⑤ Fax.053-623-8572



(주)큐텍코리아  
대표이사 박종묵  
브랜드 : 서비스개발

위 회사는 지속적인 연구와 창의력을 바탕으로 차별화된 서비스를 통하여 국가경제발전에 기여한 공로가 크므로 「2020 대한민국 글로벌파워브랜드 大賞」을 수여합니다.

2020년 12월 30일



# UCWIT2022



06704 서울특별시 서초구 방배로 76, 401호 (방배동, 머리재빌딩)  
전화 1588-2728 팩스 02-521-1352  
<http://www.kiise.or.kr> [kiise@kiise.or.kr](mailto:kiise@kiise.or.kr)