

# 시각장애인의 라이프 사이클을 지원하는 인공지능 웨어러블 플랫폼

<sup>1,2</sup>박시웅, <sup>3</sup>김정은, <sup>4</sup>강현서, <sup>5\*</sup>박형준

## Artificial intelligence wearable platform that supports the life cycle of the visually impaired

<sup>1,2</sup>Siwoong Park, <sup>3</sup>Jeung Eun Kim, <sup>4</sup>Hyun Seo Kang, <sup>5\*</sup>Hyoung Jun Park

### 요약

본 논문에서는 시각장애인의 라이프 사이클을 사전에 학습하여 시각장애인의 자립생활을 돕는 적정기술로 음성인식 기반 스마트 웨어러블 디바이스, 스마트 기기 및 웹 AI서버를 포함하는 음성, 사물 및 문자 인식 플랫폼을 제안하였다. 시각장애인을 위한 웨어러블 기기는 착용편의성과 사물인식기능 효율을 높이기 위해 리버스 넥밴드 구조로 설계하여 제작하였으며, 웨어러블 기기에 부착된 고감도 소형 마이크와 스피커는 웨어러블 기기와 연동된 스마트기기의 앱으로 구성된 음성인식 인터페이스 기능을 지원하도록 구성하였다. 음성, 사물 및 광학문자 인식 서비스는 웹 AI 서버에서 오픈소스 및 구글 API를 활용하였고, 서비스 플랫폼의 음성, 사물 및 광학문자 인식 정밀도는 실험을 통하여 평균 90%이상 달성하였음을 확인하였다.

### Abstract

*In this paper, a voice, object, and optical character recognition platform including voice recognition-based smart wearable devices, smart devices, and web AI servers was proposed as an appropriate technology to help the visually impaired to live independently by learning the life cycle of the visually impaired in advance. The wearable device for the visually impaired was designed and manufactured with a reverse neckband structure to increase the convenience of wearing and the efficiency of object recognition. And the high-sensitivity small microphone and speaker attached to the wearable device was configured to support the voice recognition interface function consisting of the app of the smart device linked to the wearable device. From experimental results, the voice, object, and optical character recognition service used open source and Google APIs in the web AI server, and it was confirmed that the accuracy of voice, object and optical character recognition of the service platform achieved an average of 90% or more.*

**Keywords:** *Appropriate Technology, The Visually Impaired, Artificial Intelligence, Infrared Image, Wearable Device*

<sup>1</sup> 한국전자통신연구원 호남권연구센터 연구원(swp@etri.re.kr)

<sup>2</sup> 광주과학기술원 전기전자컴퓨터공학부 박사과정(swp@gm.gist.ac.kr)

<sup>3</sup> 한국전자통신연구원 호남권연구센터 선임연구원(j.kim@etri.re.kr)

<sup>4</sup> 한국전자통신연구원 호남권연구센터 책임연구원/실장(hskang87@etri.re.kr)

<sup>5</sup> 교신저자 한국전자통신연구원 호남권연구센터 책임연구원(spacegon@etri.re.kr)

## I. 서론

최근 초고령화 사회 진입으로 후천적 안질환으로 인해 시각장애인 인구가 급속하게 증가하고 있다. 세계보건기구 WHO(World Health Organization)와 국제실명예방기구 IAPB(International Agency for the Prevention of Blindness)에 따르면 전 세계적으로 시각 장애를 갖고 있는 것으로 추정되는 인구는 최소 22억명에 달하는 것으로 보고되고 있다[1]. 스마트폰, 스마트 패드 등 디지털 정보기기의 급속한 확산으로 저시력자가 2050년까지 약 2배 이상 증가할 것으로 전망되는 연구결과도 발표되었다[2]. 보건복지부에 따르면 2018년 기준 국내 약 252만명의 장애인 중 약 9.8%에 해당하는 24.6만명이 시각장애인이며, 이 중 장애 정도에 따라 경증 장애인은 19.9만명, 중증 장애는 4.8만명으로 통계 결과가 발표되었다[3].

시각장애인은 신체적인 제약으로 인해 주변 상황 및 사물에 대한 인지력이 저하됨에 따라, 일상 생활을 돕는 보조도구를 사용한다. 대다수의 시각장애인들은 점자 보도블록과 지팡이를 이용하여 장애물 유무, 도면 상태 등의 정보를 얻거나 안내견의 안내를 받아 야외 보행을 해왔다. 하지만, 특정 장소에서의 안내견 출입 제한, 점자 보도블록이 없는 보행 도로, 지팡이에 의해 하방으로만 제한된 좁은 탐색 영역, 지팡이로 인한 양손이 자유롭지 못한 점, 야간보행의 어려움 등의 문제로 인해 독립보행과 자립생활에 불편함을 겪어왔다. 이러한 문제점을 인식한 많은 연구자들이 첨단 센서 및 인지기술을 이용하여 시각장애인의 기반으로 일상 생활을 돕기 위한 보조공학기기들을 개발해왔다[4-7]. 모자, 안경, 벨트와 같이 착용하는 형태와 손전등, 지팡이와 같이 들고 다니는 형태, 스마트폰 기반 어플리케이션 보조 서비스 등 다양한 보조공학기기들이 출시되었으나, 시각장애인을 대상으로 사용편의성과 이동보조의 불편함을 이유로 소수의 보조공학기기들만 사용되어 왔다. 게다가, 딥러닝 기반 인공지능 개발에 힘입어 보조공학기기도 인공지능을 접목하여 카메라를 통한 글자인식, 사물인식, 사람인식 등의 기능이 내장된 보조공학기기들이 개발되었으나[8-14], 시각장애인의 라이프 사이클에 대한 맞춤형 사전학습기반의 적정기술이 적용된 보조공학기기의 부재에 따라, 시각장애인의 눈높이에 맞는 편의성을 제공하고, 개인의 일상 삶을 돕는 보조공학기기의 필요성이 요구되고 있다.

본 논문에서는 시각장애인 맞춤형 서비스를 제공하기 앞서, 음성인식 기반의 사물인식, 광학문자 인식 기능을 지원하는 웨어러블 기기, 스마트기기 및 웹 인공지능 서버를 포함하는 시각장애인용 인공지능 플랫폼을 제안한다. 제작한 웨어러블 기기는 착용편의성과 사물인식기능 효율을 높이기 위해 리버스 텍밴드 구조로 설계하여 제작하였으며, 웨어러블 기기에 부착된 고감도 소형 마이크와 스피커는 웨어러블 기기와 연동된 스마트기기의 앱으로 구성된 음성인식 인터페이스 기능을 지원하도록 구성하였다. 본 논문에서는 제안한 인공지능 플랫폼의 음성, 사물 및 광학문자 인식 정밀도는 분석 툴을 이용하여 실험 데이터 인식결과로 보고한다.

## II. 음성인식 인터페이스 기반의 시각장애인용 인공지능 플랫폼

### 2.1 시각장애인 대상 인공지능 서비스 플랫폼 구성

제안한 서비스 플랫폼은 인공지능 서버, 스마트 디바이스 및 웨어러블 기기로 그림 1과 같이 구성하였다. 시각장애인용 웨어러블 기기에서 취득한 음성, 사물 및 문자 데이터는 블루투스 및 Wi-Fi 모듈을 이용하여 스마트폰이나 스마트 패드와 같은 스마트기기로 전송하고, 스마트 기기는 서비스 시나리오에 따라 취득한 데이터를 인공지능 서버로 전송한다. 인공지능 서버는 스마트단말의 데이터를 수집하여 저장하는 파트, 시나리오에 따른 서비스 처리부, 스마트 단말과 통신을 주고받는 인터페이스 부 및 스마트 단말에서 받은 음성, 사물 및 광학문자를 판별할 수 있는 딥러닝 기반의 강화학습 모델이 포함된 인공지능 엔진부로 구성된다. 서버로 전송한 데이터는 인공지능 엔진을 적용하여 음성, 사물 및 광학문자별로 인식결과를 스마트 기기로 전송하고, 스마트 기기는 웨어러블 기기에 음성으로 인식결과값을 변환하여 사용자에게 음성으로 결과를 전달하는 구조로 시스템을 구성하였다.

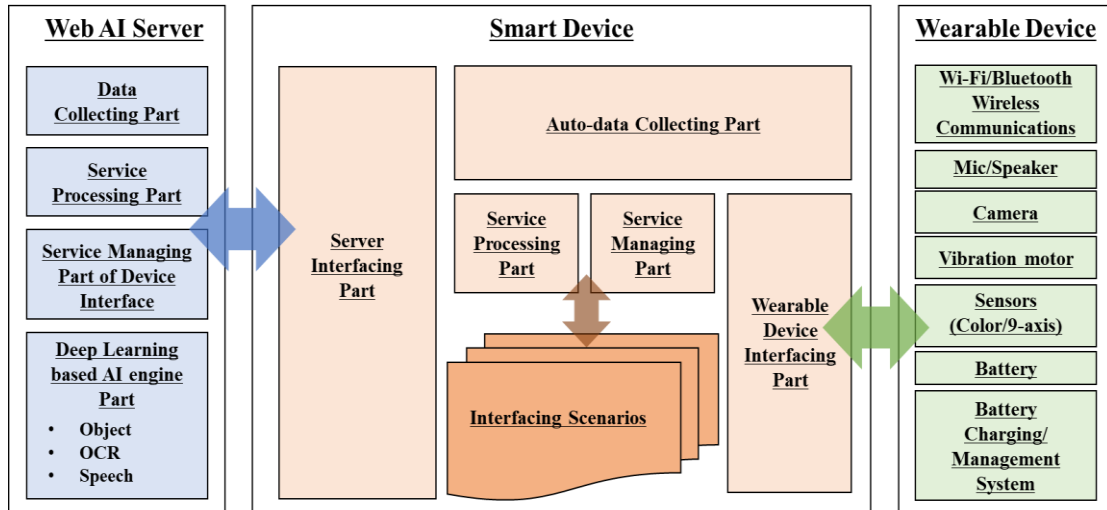


Figure 1. Configuration of proposed system

2.2 시각장애인 적용 웨어러블 디바이스 설계 및 제작

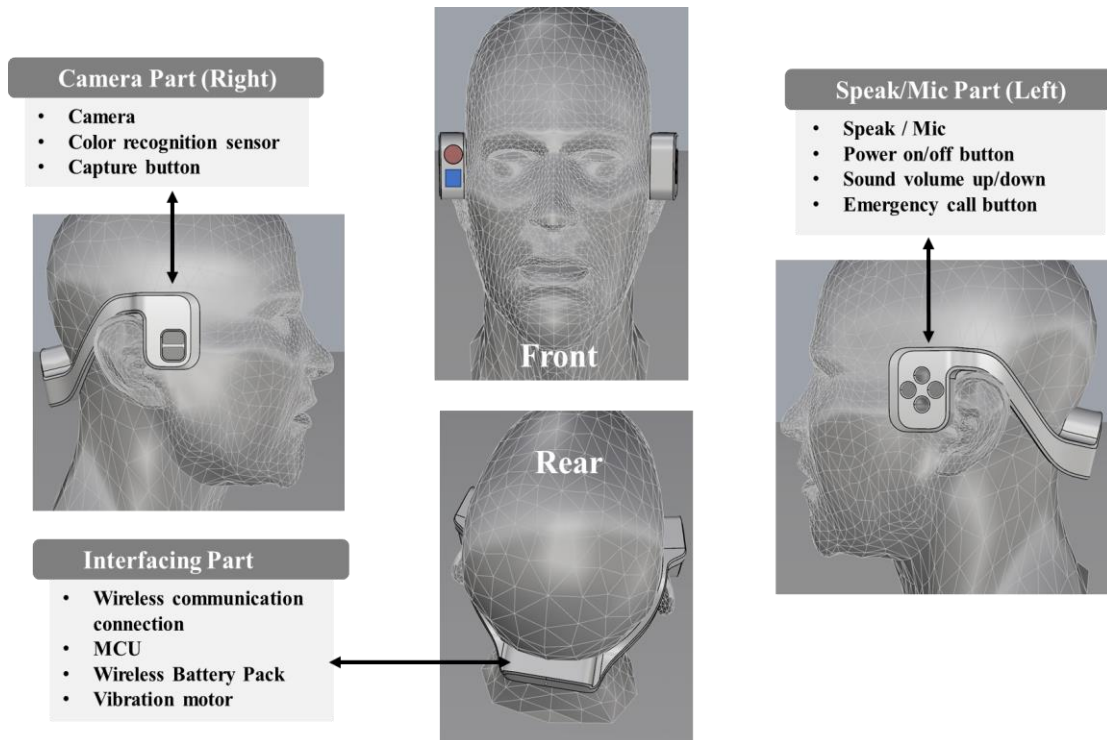


Figure 2. Design for a wearable device

시각장애인 적용 웨어러블 디바이스는 그림 2와 같이 카메라 부, 음성인식 부 및 인터페이스 부로 구성하고 형상을 설계하였다. 카메라 부에는 사물인식 및 광학문자 인식을 위한 소형 카메라 모듈과 시각 장애인들의 양말, 신발 및 옷과 같은 사물의 단순 색상을 구별 가능하도록 색상 구별 센서를 실장하고, 카메라 부 측면에는 각각의 모듈을 동작을 지원하는 물리버튼이 제공된다. 스피커 및 마이크 부에는 음성인식을 위한 고감도 소형 마이크와 인공지능 판단 결과값을 음성으로 지원하는 소형 스피커를 실장하고, 측면에는 기기 전원 버튼, 음성 볼륨버튼

및 응급상황을 전파할 수 있는 버튼이 제공된다. 뒷면에 위치한 스마트 단말과 연동하는 인터페이싱 부는 무선으로 스마트 단말과 연결할 수 있는 Wi-Fi 및 BT 5.0 모듈, 진동으로 착용자에게 정보를 전달할 수 있는 진동모터, 웨어러블 기기의 수집된 데이터를 처리할 있는 마이크로 컨트롤러 및 전원을 공급하는 무선충전 배터리 팩으로 구성된다. 30 cm 이상의 사물에 포커싱을 맞추기 위하여 렌즈 설계와 더불어 웨어러블 기기의 카메라 모듈을 디자인 형상에 따라 아래 그림 3 과 같이 제작하였고, 제작한 모듈의 성능 규격은 표 1 과 같다.

Table 1. Specifications of camera module for a wearable device

Item	Unit	Description
Dimensions (W x H x T)	mm	38x16x23
Resolution	-	1920x1080
Image sensor size	Inch	1/2.7
Image Transfer Rate	FPS	30@FHD
AEC/AGC/White Balance	-	Auto
Sensitivity	V/lux-sec	3.3@550 nm
Mini illumination	lux	0.05

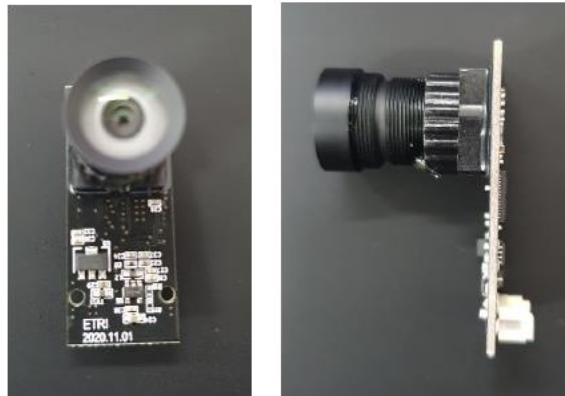


Figure 3. Lensed camera module for a wearable device

그림 4 는 시각장애인을 위한 웨어러블 장치의 시제품을 착용한 그림이다. 리버스 넥밴드 구조로 제작하여 기기로부터 하중을 최소화하였으며, 시각장애인의 정면 방향의 사물 및 문자 인식이 가능하도록 카메라를 위치하였다.



Figure 4. A proto-type for a wearable device

### III. 실험 및 고찰

본 논문에서 음성인식 서비스는 머신러닝 활용이 가능한 Google STT(Speech-To-Text)를 활용하였고, 문자인식은 Google Vision API 를 사용하여 테스트 추출하였고, 사물인식은 Darknet 오픈소스의 하나인 YOLO(You Only Look Once) v3 를 이용하여 웹 AI 서버에서 시험을 수행하였다. 웹기반 AI 서버는 CPU 3.6GHz, 8 core, RAM 16GB, GPU GTX-1080Ti 환경으로 구축하였고, 그림 5 와 같이 취득한 음성데이터를 2468 개 데이터 샘플로 나누고 비정상 음성데이터와 정상 음성데이터를 구별하기 위한 임계값을 2.11 로 설정하였다. 음성인식 수행을 위해 그림 Batch size 를 32 로, Epoch 를 10 회 적용하여 학습한 결과, 음성인식 시험 데이터 셋에 대해 Recall=0.893, Precision=0.948 을 결과값을 얻었다.

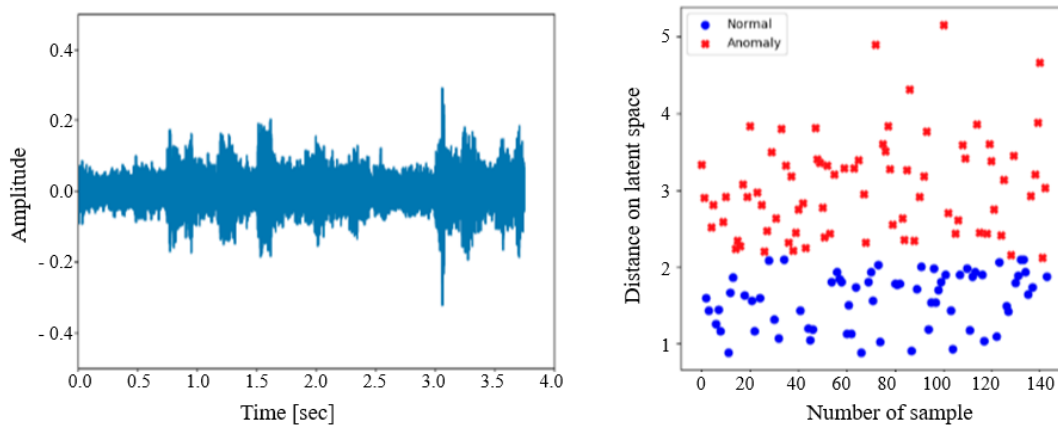


Figure 5. Anomaly test of voice recognition

문자인식 시험은 시각장애인의 일상 생활에 필요한 라벨이미지를 웨어러블 기기의 카메라를 이용하여 얻었고, 스마트기기를 통해 이미지를 웹 AI 서버로 전송한 후, Google Vision API 를 활용하여 문자를 그림 6 과 같이 추출하였다.



Figure 6. Optical character detection and recognition in capture images of a wearable device

또한, 사물인식을 위해 Microsoft COCO (Common objects in context) 2014 의 훈련 데이터셋으로 학습된 가중치(weight)를 기반으로 사물영상을 사용하여 전이학습(transfer learning)을 진행하였다. 40 개의 class 에 Batch size 를 16 하여 273 회 Epoch 에 대해 학습하였고, 시각장애이용 보조기기로 촬영한 영상의 객체검출을 테스트한 결과, mAP=0.561, Recall=0.834, F1-score=0.895, Precision=0.966 으로 나타내었다. 그림 7 은 학습한 인공지능 플랫폼을 이용하여 시각장애이용 웨어러블 카메라로 캡처한 이미지로부터 모니터, 키보드 등은 모두 100 %, 컵과 핸드폰은 각각 82 %, 94 %로의 사물인식 정확도를 보여준다.

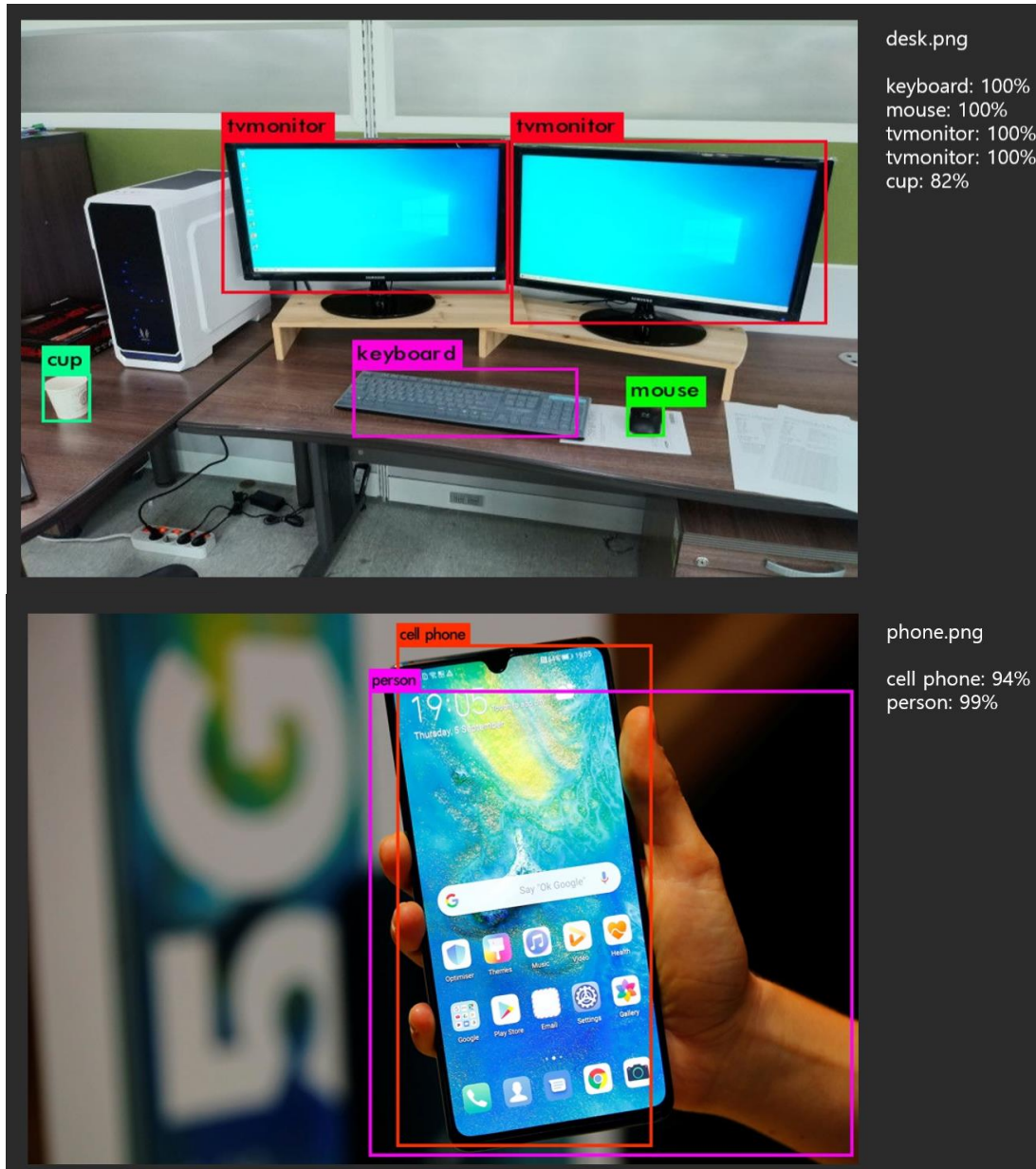


Figure 7. Object detection and recognition in capture images of a wearable device

#### IV. 결론

본 논문에서는 시각장애인의 자립을 돕는 적정기술로 활용하기 위해 웨어러블 기기, 스마트기기 및 인공지능 서버를 포함하는 시각장애인 맞춤형 인공지능 플랫폼을 제안하였다. 시각장애인을 위한 웨어러블 기기는 착용편의성과 사물인식기능 효율을 높이기 위해 리버스 벡밴드 구조로 설계하여 제작하였으며, 웨어러블 기기에 부착된 고감도 소형 마이크와 스피커는 웨어러블 기기와 연동된 스마트기기의 앱으로 구성된 음성인식 인터페이스 기능을 지원하도록 구성하였다. 원활한 서비스 지원을 위한 음성, 사물 및 광학문자 인식 인공지능 서비스 플랫폼을 구성하였고, 구성한 서비스 플랫폼의 음성, 사물 및 광학문자 인식 정밀도는 인공지능 서버 분석 프로그램을 이용하여 평균 90%이상 달성하였음을 확인하였다. 향후, 추가

연구개발을 통해, 개발한 플랫폼을 기반으로 음성명령어를 이용하여 사물 및 광학문자 인식 서비스 이외에도 긴급알림 서비스, 낙상알림 서비스, 일정관리 서비스 및 보행자 안내서비스등 시각장애인의 삶을 지원하는 적정기술 서비스를 개발할 예정이다.

## V. 감사의 글

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.1711117076, 시각장애인의 독립보행과 자립생활을 위한 음성인식 인터페이스 기반의 웨어러블 디바이스 개발)

## VI. 참고문헌

- [1] World Health Organization (WHO), "World report on vision", Aug. 2019. Available: <https://www.who.int/publications/i/item/world-report-on-vision>
- [2] Bourne, R., Flaxman, S. R., Braithwaite, T., Cicinelli, M. V., Das, A., Jonas, J. B., Keeffe, J., Kempen, J. H., Leasher, J., Limburg, H., Naidoo, K., Pesudovs, K., Resnikoff, S., Silvester, A., Stevens, G. A., Tahhan, N., Wong, T. Y., Taylor, H. R., Vision Loss Expert Group, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," *The Lancet Global Health*, Vol. 5, No. 9, pp. e888-e897, Sep. 2017.
- [3] The Ministry of health and welfare, "2020 Disabled people's life in statistics," Jul. 2020. Available: [http://www.mohw.go.kr/react/al/sal0301vw.jsp?PAR\\_MENU\\_ID=04&MENU\\_ID=0403&page=1&CONT\\_SEQ=356153](http://www.mohw.go.kr/react/al/sal0301vw.jsp?PAR_MENU_ID=04&MENU_ID=0403&page=1&CONT_SEQ=356153)
- [4] M. M. Islam, M. Sheikh Sadi, K. Z. Zamli and M. M. Ahmed, "Developing walking assistants for visually impaired people: A review," *IEEE Sensors Journal*, Vol. 19, No. 8, pp. 2814-2828, Apr. 2019.
- [5] K. Yang, K. Wang, L. M. Bergasa, E. Romera, W. Hu, D. Sun, J. Sun, R. Cheng, T. Chen, E. López, "Unifying terrain awareness for the visually impaired through real-time semantic segmentation," *Sensors*, Vol. 18, No. 5, pp. 1506, May 2018.
- [6] Sh. Al-khalifa, M. Al-Razgan, "Ebsar: Indoor guidance for the visually impaired," *Computer and Electrical Engineering*, Vol. 54, pp.26-39, Aug. 2016.
- [7] A. Bhowmick, S. M. Hazarika, "An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends," *Journal on Multimodal User Interfaces*, Vol. 11, pp. 149-172, Jan. 2017.
- [8] S. -Y. Bea, "Research trends on related to artificial intelligence for the visually impaired: focused on domestic and foreign research in 1993-2020," *Journal of the Korea Contents Association*, Vol. 20, No. 10, pp. 688-701, Oct. 2020.
- [9] S. Park, T. Jeon, S. Kim, S. Lee, J. Kim, "Deep learning based symbol recognition for the visually impaired," *Jorunal of Korea Institute of Information, Electronics, and Communication Technology*, Vol. 9, No. 3, Jun. 2016.
- [10] J. Oh, C. Bong, J. Kim, "Design of immersive walking interaction using deep learning for virtual reality experience environment of visually impaired people," *Journal of Korea Computer Graphics Society*, Vol. 25, No. 3, pp. 11-20, Jul. 2019.
- [11] H. W. Seol, C. Poleak, J. W. Kwon, "Deep learning based character-oriented image captioning method for visually impaired," *Journal of Rehabilitation Welfare Engineering & Assistive Technology*, Vol. 13, No. 2, pp. 143-149, May 2019.
- [12] T. Jeon, S. Lee, "Deep learning based sign detection and recognition for the blind," *Journal of the Institute of Electronics and Information Engineers*, Vol. 54, No. 2, Feb. 2017.
- [13] S.-H. Lee, M.-S. Kang, "Implementation of object detection and voice guidance system for the visually handicapped using object recognition technology," *Journal of the Institute of Electronics and Information Engineers*, Vol. 55, No. 11, pp. 65-71, Nov. 2018.



- [14] C.-J. Kim, M.-S. Park, M.-H. Kim, "A development architecture research of intelligent object recognition and voice service for the visually impaired," Journal of Knowledge Information Technology and Systems, Vol. 13, No. 4, pp. 441-450, Aug. 2018.

## 저자 소개



박시웅(*Siwoong Park*)

2016년 8월 광주과학기술원 전기전자컴퓨터공학부 석사  
2020년 3월 ~ 현재 광주과학기술원 전기전자컴퓨터공학부 박사과정  
2016년 10월 ~ 현재 한국전자통신연구원 호남권연구센터 연구원

관심분야: IT 융합, 인공지능, 광융합 센서



김정은(*Jeong Eun Kim*)

2004년 2월 한국과학기술원 대학원 물리학과 석사  
2011년 5월 베를린공과대학교 대학원 물리학과 박사  
2012년 9월 ~ 현재 한국전자통신연구원 호남권연구센터 선임연구원

관심분야: IT 융합, 인공지능, 광융합 센서



강현서(*Hyun Seo Kang*)

1996년 2월 성균관대학교 대학원 전자공학과 석사  
2000년 2월 성균관대학교 대학원 전자공학과 박사  
2001년 9월 ~ 현재 한국전자통신연구원 호남권연구센터 책임연구원/실장

관심분야: IT 융합, 인공지능, 광융합 센서



박형준 (*Hyoung Jun Park*)

2004년 2월 전북대학교 대학원 제어계측공학과 석사  
2009년 2월 전북대학교 대학원 전자공학과 박사  
2010년 7월 ~ 현재 한국전자통신연구원 호남권연구센터 책임연구원

관심분야 : IT 융합, 인공지능, 광융합 센서