

컴퓨터교육학회 논문지 2026년 제29권 제3호
https://doi.org/10.32431/kace.2026.29.3.004



멀티모달 기반 시각장애인 일상지원 앱 개발 연구*

A Study on the Development of a Multimodal Daily Life Support App for People with Visual Impairment

전준혁[†] · 이영주[‡] · 김규리[†] · 최준희[†] · 유수진^{††}

Junhyuck Jeon[†] · Youngjoo Lee[‡] · Kyuri Kim[†] · Joonhee Choi[†] · Sujin Yoo^{††}

요약

본 연구는 시각장애인의 자립성과 안전한 일상생활의 지원을 위해 멀티모달 기반 보조 어플리케이션 “소리눈”을 개발하였다. “소리눈”은 객체 인식, 경로 안내, 문서 인식 및 음성 안내, 보호자와의 위치 공유 및 긴급 연락 기능을 통합적으로 제공한다. 이를 통해 시각장애인의 이동 편의성, 정보 접근성, 위기 상황 대응력을 향상시키고자 하였다. 실시간 객체 인식과 TTS, GPS 등의 기술을 연계하여 다양한 상황에 대응 가능한 환경을 구현하였다. 다만 낮은 조도에서의 인식 성능 저하, 개인정보 보호 이슈, 실제 사용자 대상 테스트 부족 등의 한계도 존재한다. 향후 사용자 피드백 기반의 개선과 공공서비스 연계를 통해 실용성과 확장성을 높일 수 있을 것으로 기대된다.

주제어 시각장애인 보조, 멀티모달 어플리케이션, 음성 안내, 객체 인식, 경로 안내

ABSTRACT

This study developed a multimodal assistive application, “SoriNoon”, to support the independence and safe daily living of individuals with visual impairments. “SoriNoon” integrates various functions, including object detection, route guidance, document recognition with voice output, real-time location sharing with caregivers, and emergency contact features. The system aims to improve the mobility, information accessibility, and emergency responsiveness of visually impaired users. By combining real-time object detection, text-to-speech (TTS), and GPS technologies, the application provides a responsive environment suitable for diverse real-world situations. However, there are limitations such as decreased recognition accuracy under low-light conditions, concerns over privacy in public spaces, and a lack of sufficient testing with actual users. Future improvements based on user feedback and integration with public services are expected to enhance the practicality and scalability of the system.

Keywords Assistive Technology for the Visually Impaired, Multimodal Application, Voice Guidance, Object Detection, Route Navigation

[†]정회원 한성대학교 컴퓨터공학부 재학
^{††}정회원 한성대학교 컴퓨터공학부 조교수
 (교신저자)
논문투고 2025년 07월 25일
심사완료 2025년 11월 17일
게재확정 2025년 11월 19일
발행일자 2026년 03월 19일

* 본 연구는 2025년 한성대학교 학술연구비 지원과제임

1. 서론

인간의 오감 중 시각은 가장 핵심적인 감각기관이다. 감각수용체 중 70% 이상이 눈에 존재하며 뇌에서 전달받은 정보 중 90% 이상이 시각을 통해 전달되기 때문이다[1].

2024년 장애인 등록 현황[2]에 따르면 한국의 등록 시각장애인은 약 25만 명이다. 해당 집단은 정보 접근 및 물리적 이동에서 구조적 제약을 경험하여 사회적 고립감과 상실감을 느낄 수 있다[3]. 또한 등록 시각장애인의 95.4%가 돌이 후에 시각손상을 겪었으며, 92.4%는 후천적 장애발생에 의한 시각장애로 보고되어[4], 초고령화가 진행 중인 한국 사회에서 시각장애인을 위한 사회적·제도적 지원 필요성이 가중되고 있음을 시사한다.

시각장애인은 시각 정보를 활용할 수 없기 때문에, 주로 청각과 촉각을 통해 외부 환경을 인지하며 일상생활을 영위한다[5, 6]. 하지만 이러한 대체 감각만으로는 다양한 상황에서 충분한 정보를 얻기 어려운 경우가 많다. 특히 보행 중 주변 장애물이나 돌발적인 위험 요소를 식별하는 데 큰 어려움을 겪는다[7, 8]. 뿐만 아니라 안내 표지, 상품 정보, 공공 문서와 같은 시각적 정보에 대한 접근성 역시 제한적[9-11]이며, 이는 시각장애인의 자립적인 생활과 사회 참여를 제약하는 주요 원인이 된다.

이러한 문제를 해결하기 위해 다양한 보조기술이 개발되어 왔으며, 최근에는 인공지능 기술과 모바일 기기의 발전으로 정보 접근 방식에 혁신적인 변화를 일으켰다[12]. GPS를 이용한 실시간 위치 기반 경로 안내 기술, 광학 문자 인식(Optical Character Recognition, OCR) 기술, TTS(Text-to-Speech)를 통한 음성 정보 제공, 자연어 처리 기술과 지식 검색, 그리고 YOLO(You Only Look Once)와 같은 딥러닝 기반 실시간 객체 인식 기술은 시각장애인의 주변 환경 인식 능력을 보완하고, 보다 안전하고 효율적인 이동을 가능하게 한다.

이러한 기술들은 단독으로도 유용하지만, 상호 연동되어 통합적으로 활용될 때 더욱 높은 효과를 발휘할 수 있다. 예를 들어, 객체 인식을 통해 탐지된 위험 요소를 음성으로 안내하거나, 사용자의 위치 정보와 연계해 경로상의 장애물을 사전에 경고하는 등의 융합 서비스는 실제 보행 환경에서 시각장애인의 대응력을 높이는 데 기여할 수 있다.

본 논문에서는 멀티모달[13]을 통해 이러한 기술들의 연계와 응용 효과에 대해 논의하며, 시각장애인을 위한 보조 시스템 개발의 방향성과 활용 방안을 모색하고자 한다.

2. 관련 연구

시각장애인은 신체의 감각 중 청각과 촉각을 사용하여 정보를 인지한다. 보편적으로 촉각은 점자와 같은 정적인 정보를 이해하는데 활용되며, 청각은 정보를 보완하거나, 동적인 정보를 전달하는 임무를 수행한다. 이러한 특성에 기

반하여 다양한 보조 기술들이 개발 및 상용화되고 있다.

2.1 멀티 모달(Multimodal)

멀티 모달은 하나의 매체 안에 텍스트, 이미지, 오디오 등 여러 모드를 적용하는 것을 의미한다[14, 15]. Parcalabescu 외(2021)는 기존의 멀티모달리티 개념이 머신러닝 실무에 비해 협소하다고 지적하며, 단순히 입력 데이터 유형의 차이가 아니라 특정 머신러닝 task에서 필요한 정보의 최소 단위(atomic units of information)에 두고, 다양한 모달리티 종류가 아니라 작업상 꼭 필요한 정보 단위 간의 차이가 존재할 때 해당 작업을 '멀티모달'로 정의해야 한다고 주장한다[13]. Liang 외(2024)는 모달리티(modality)를 정보가 인코딩돼 표현·전달되는 방식으로 정의[16]하며, 언어, 시각, 청각, 촉각 등 이질성(heterogeneity), 의미적·정보적 연결과 부분적 중첩이 존재하는 연결(connections), 모달리티들이 결합될 때 개별 모달리티만론 얻기 어려운 추가 정보를 도출할 수 있는 상호작용(interactions)을 핵심 원칙으로 제시한다. 즉, 단순한 데이터 종류의 결합이 아니라, 각기 다른 정보 구조가 결합해 시너지와 새로운 의사결정 논리를 만들어내는 시스템이다.

본 연구에서는 스마트폰을 활용할 수 있으며 일정 수준의 색상 구분이 가능한 시각장애인 및 저시력자를 주요 대상 사용자로 설정하였다. 이들의 보행 시 전송되는 실시간 영상, 제품 정보 확인을 위한 촬영 이미지, 사용자의 음성 명령 및 텍스트 데이터를 수용하여, 이를 기반으로 시각장애인의 일상생활을 지원하는 어플리케이션을 구현하였다.

2.2 YOLO(You Only Look Once) 기반 안내 기술

YOLO 이전의 객체 탐지 모델로는 R-CNN[17], Fast R-CNN[18], Faster R-CNN[19]과 같은 모델들이 존재했다. 이 모델들은 객체를 탐지하기 위해 내부적으로 여러 단계를 거쳐야 했다. R-CNN의 경우 먼저 영역을 제안(Region Proposal)한 후, CNN으로 영역을 분류하며, Fast R-CNN은 Region Proposal Network(RPN)를 도입하여 속도가 개선되었으나 여전히 여러 단계를 거쳐야만 했다.

그러나 YOLO는 이미지를 그리드로 나누어 한 번의 신경망 패스로 객체 존재 여부, 위치, 클래스를 동시에 예측한다. 이에 속도가 이전 모델에 비해 훨씬 개선되어 실시간으로 객체를 탐지하는 것에 적합하다[20]. 그로 인해 단 한 번(You Only Look Once) 보고 객체를 탐지할 수 있다는 의미로 YOLO로 명명되었다. 이러한 특성 덕분에 YOLO는 카메라 영상을 통해 사람, 자동차, 장애물 등의 위험 요소를 빠르고 정확하게 탐지 및 분석할 수 있으며, 단순한 길 안내 시스템에서 더 발전된 상황 인지 기반 안내 시스템 구현이 가능하다.

기존 연구는 주로 비장애인을 대상으로 한 보행 경고 시스템이나 자동 운전 분야에 집중되었으며, 시각장애인을 위

한 음성 기반 위험 경고에서 활용한 연구는 드물다.

본 연구에서는 YOLO v11을 사용하여 시각장애인 및 저시력자의 보행 중에 빠른 속도로 다가오는 사람이나 물체를 우선적으로 인식하고 미리 알림을 주어 시각장애인의 안전을 확보할 수 있도록 구현하였다.

2.3 AI 기반 시각장애인 지원 연구

2.3.1 정보 접근성 향상 연구

본 섹션에서는 시각장애인이 일상에서 필요한 정보를 보다 효율적, 직관적으로 접근할 수 있도록 수행된 연구들을 다룬다.

권경수 외(2024)는 ‘설리번 플러스’라는 앱을 통해 시각장애인의 시각보조 기능을 구현하였다[12]. 해당 앱은 문자 인식, 장면 인식, 화폐 인식, 얼굴 인식 기능을 제공하며, 인식 결과는 사용자에게 음성으로 전달된다. 클라우드 AI, 온디바이스 AI[21, 22]를 모두 사용함으로써 사용자에게 보다 빠른 피드백을 제공한다. 하지만 “소리눈”에서와 같이 장애인의 이동성을 향상시키는 길찾기와 같은 기능은 별도로 제공하지 않는다.

이소라 외(2020)는 시각장애인 중 전맹이 아닌 저시력인을 대상으로 이미지 분석, 텍스트 분석 기능을 제공하는 음성지원 앱을 개발하였다[23]. 저시력인의 경우 약간의 도움으로도 주변의 정보 획득이 가능하다는 특징이 있다. 이미지 분석은 이미지 캡셔닝(Image Captioning) 기술을 통해 어떤 이미지인지 자연어로 설명하는 방식으로 수행된다. 텍스트 분석은 OCR(Optical Character Recognition) 기능으로 문자열 추출 후 사용자에게 안내하는 TTS(Text-to-Speech)와 사용자의 음성을 변환하여 요청을 수행하는 STT(Speech-to-Text)를 모두 사용한다. 사용자의 음성에 대한 결과를 반환하기 위해, 웹 데이터를 크롤링하거나 실시간 감성분석을 이용한다. 앱은 음성으로 모든 기능 구동이 가능하며, 결과 역시 음성으로 제공한다. 하지만 이미지 캡셔닝 시 사용하는 API가 한국어 기반의 결과를 제공하지 않으므로, 네이버 파파고 번역 결과를 한 번 더 거침으로써 다소 어색한 문장이 나올 수 있다는 점, 검색 대상이 위키피디아로 다소 정확도가 떨어질 수 있다는 한계가 있다.

강술지 외(2020)는 식, 의약품 어플리케이션을 제안하였다[24]. 시각장애인의 경우 약품 정보를 잘못 인식해서 의약품 오남용 경험이 있고, 식품의 유통기한을 틀리게 인지하고 있어 질병에 노출되는 경우가 많은 것으로 나타났다. 이에 강술지는 기존 김하영 외(2019)의 연구[25]를 개선하여 음성검색, 이미지 인식 기능을 통해 의약품, 식품의 상세 정보를 제공하도록 어플리케이션을 구현하였다. 그러나 데이터 수집을 위해 수행된 인터뷰 인구가 총 2명으로 전체 시각장애인 집단의 다양한 요구와 경험을 정확히 대표하지 못할 수 있으며, 프로토타입 테스트가 충분하게 이루어지지 않았다는 한계점이 있다.

임문철 외(2024)는 YOLO v5와 라즈베리파이 기반의 상

품 식별 시스템을 제안하였다[26]. 해당 시스템은 시각장애인의 상품 정보 접근성과 편의성 향상을 목적으로 한다. 구현을 위해 라즈베리파이 4, Pi 카메라, 배터리 모듈, 버튼으로 구성되어 있으며 라즈베리파이가 YOLO 기반 객체 인식과 이미지 처리를 수행하고 결과에 대한 음성을 출력한다. 사용자가 카메라 앞에 인식을 원하는 상품을 두고 촬영 버튼을 누르면, 촬영된 이미지가 라즈베리파이에 전송되고 YOLO 기반 알고리즘을 통해 분석된다. 인식된 상품 이름은 음성으로 출력된다. 하지만 각 상품당 500장의 이미지만으로 10가지 상품에 대해 학습이 진행되었다. 그로 인해 모델의 객체 인식의 정확도가 다소 떨어지며, 객체 인식의 다양성이 부족하다는 한계점이 있다.

2.3.2 IoT 및 센서 기반 보행 보조

본 섹션에서는 주변 환경에 대한 인지 능력을 보완하면서 시각장애인의 이동성을 향상시키는 연구들을 다룬다.

정예진 외(2020)는 초음파, 조도, 진동 센서, 그리고 비콘을 통한 스마트 지팡이를 구현하였다[27]. 각각의 센서는 사용자가 장애물 위치를 더욱 정확하게 알 수 있도록 알림을 전달하며 조도센서를 통해 주변에 다른 보행자가 인식되면 시각장애인의 존재를 알린다. 비콘은 신호등과 음향 신호기 간의 거리를 측정하여 1m 이내가 되면 휴대전화, 스마트 지팡이 모두 진동이 울리도록 하였고, 블루투스로 스마트폰과 지팡이의 연결을 구현하였다. 그러나 진동이나 음향을 통해 알림을 제공하므로, 시각장애인 사용자가 구체적으로 어떤 장애물이 있는지 인지하기는 어렵다는 한계가 있다. 또한 빠른 속도로 이동하는 사람이나 사물에 대한 부분은 해당 연구에서는 고려하지 않았다.

박세영 외(2020)는 카메라 센서와 IoT 초음파 센서를 부착한 보행 보조기를 통해 사용자 주변의 장애물이 있는지 감지하고 물체의 위치에 따라 경고음을 다르게 제공함으로써 사용자가 근접 정도를 구별할 수 있도록 하였다[28]. 현재 위치 오차를 줄이기 위해 스마트 폰의 GPS 뿐 아니라 초음파 센서와 카메라 센서를 함께 이용한다. 해당 장비에는 인텔의 OpenVINO를 사용하였고, 보행자 경로 탐색을 위해 Tmap API를 활용하였다. 이를 바탕으로 음성 인식, TTS, 큰 아이콘을 사용하여 길안내 앱을 구현하였다. 길안내 앱은 시각장애인 사용자가 위급상황이 되었을 때 보호자에게 영상을 발송한다. 하지만 보행 중 마주칠 수 있는 객체들 중 ‘person’, ‘car’, ‘bicycle’, ‘dog’, ‘motorbike’, ‘bus’ 등 22개 객체만 인식토록 학습시켰으며, 해당 장비가 시각장애인 사용자의 스마트폰과 별도로 몸에 부착하여 휴대해야 한다는 한계점이 있다. 또한 물체 인식 속도의 향상이 필요하다.

정인철 외(2017)는 시각장애인을 대상으로 수행한 설문조사 결과로 파악된 생활 속 어려움들을 해결할 수 있도록 IoT 인지정보 보조 시스템을 설계하였다[29]. NFC가 부착된 사물의 정보를 스마트폰이나 스마트워치를 통해 인지하고, 그 정보를 셋톱박스를 통해 음성으로 출력해준다. 또한

사용자가 새로운 사물 정보의 추가가 필요할 경우 웹 기반 관리 서버에 접속하여 입력할 수 있다. 하지만 의약품, 판매 물품을 비롯하여 모든 사물에 NFC를 부착하기 어려우며, 새로운 사물 정보를 일일이 관리서버에 추가하는 것이 현실적으로 쉽지 않다는 한계점이 있다.

2.3.3 환경 인식 및 길 안내

정용섭(2024)은 시각장애인 안내견을 대신할 수 있는 로봇 기반 길 안내 시스템 “가이드 독”을 제안한다[3]. YOLO v8과 CenterPoint[30]를 사용하여 동적 객체를 검출, 추적하고 계단이나 연석과 같은 요소들을 극복할 수 있는 사족보행 로봇을 채택하였다. 이외에도 경로 설정 시 자동문, 신호등, 에스컬레이터, 지하철 입구, 엘리베이터 문과 버튼, 지하철 개찰구, 스크린도어의 설정이 가능하다. 사용자의 음성 명령을 인식하고 처리할 수 있으며, VQA(Visual Question Answering), VLM(Visual Language Model)을 활용하여 간결하고 정확한 상황 설명을 지원한다. 출발지에서 목적지까지 전역 경로를 계획하고, 해당 경로를 바탕으로 실시간 환경 변화에 대응이 가능하도록 지역 경로 계획을 세운다. 안내견 한 마리의 훈련에 소요되는 시간, 낮은 성공률, 실제 안내견의 관리 및 의사소통 문제 등을 극복하고 보다 효율적인 보조 수단을 제시하였다. 해당 연구를 위하여 사용한 레인보우 로보틱스 사의 RBQ-3[31]는 약 2천만원 이상의 금액으로 시각장애인 안내견의 총 훈련비용인 2억원에 비해 저렴한 편이지만, 여전히 높은 비용으로 시각장애인 사용자들에게 접근성이 떨어진다는 한계가 있다.

시각장애인들은 유도블록과 점자블록을 통해 곧은 길인

지, 엘리베이터나 횡단보도 앞인지 식별할 수 있다. 오세랑과 배영철(2021)은 YOLO v3 기반으로 유도 블록을 인식하는 알고리즘을 구현하였다[32]. 이전의 유도블록 인식 시스템과 달리 오세랑의 연구에서는 스마트폰 환경에서 사용이 가능하도록 유도 블록 인식 알고리즘을 제시하였고, 인식 속도가 다른 기존의 방법보다 개선되었다. 하지만 해당 연구에서는 “소리눈”과 달리, 유도블록과 점자블록처럼 보행로에 고정된 위치에 있는 블록들을 대상으로 연구를 진행하였으므로 보행 중 돌발상황에 대해서는 고려하지 않았다. 뿐만 아니라 이미지 인식과 같은 이동중이 아닐 때 실생활에서 필요한 별도의 기능에 대해서는 구현하지 않았다.

2.4 연구의 차별점

본 연구는 시각장애인 및 저시력자를 위한 일상생활 지원 어플리케이션으로 YOLO 기반 객체 인식 기술과 TTS 기술을 통합한 음성인식 길 안내, 의약품이나 식품 성분을 포함하여 OCR을 통해 인식된 글자를 읽어주고 안내하는 시스템을 구현한다.

기존 안내 시스템이 단순히 목적지 방향에만 초점을 맞추었다면 본 시스템은 위험 요소를 우선 탐지하여 TTS로 경고 메시지를 출력하는 구조로 설계되었다. 예를 들어, 길 안내 중 위험 요소를 감지했다면 위험 요소를 우선적으로 사용자에게 음성으로 전달함으로써, 보다 안전한 보행이 가능하도록 지원한다. 위급 상황 발생시 사용자가 버튼 또는 음성 명령을 통해 보호자에게 실시간으로 알림 전송도 가능하다. 단순한 길안내를 넘어선 위험 대응 및 보호자 연동형 시스템이라는 점에서도 기존 연구와 차별화된다.

Table 1. Key Differences Between Prior Research and This Study

Feature	Kwon et al. (2024)	Lee et al. (2020)	Kang et al. (2020)	Lim et al. (2024)	Jeong et al. (2020)	Park et al. (2020)	Jeong et al. (2017)	Jeong (2024)	Oh&Bae (2021)	This research
TTS(Text-To-Speech)	O	O	O	O	O	O	O	O	X	O
OCR/Text Recognition	O	O	X	X	X	X	X	X	X	O
Image Captioning / Scene Recognition	O	O	O	O	X	X	X	△	X	O
Currency / Face Recognition	O	X	X	X	X	X	X	X	X	X
Voice Command & Control	X	O	O	X	X	O	X	O	X	O
Web Data / Real-time Analysis	X	O	X	X	X	X	X	O	X	O
Navigation	X	X	X	X	△	O	X	O	△	O
Hazard Detection	X	X	X	X	△	△	X	O	X	△
Caregiver / Family Notification	X	X	X	X	X	O	X	X	X	O
IoT/Sensor Integration	X	X	X	X	O	O	O	X	X	X
Smartphone-Only Operation	O	O	O	X	△	△	△	X	O	O
Medicine / Food Information	X	X	O	X	X	X	X	X	X	△

뿐만 아니라 OCR을 통하여 의약품이나 식품의 성분 및 유통기한 안내를 통해, 건강한 의약품 사용 및 식품 섭취를 도움으로써 시각장애인이 독립적으로 일상생활을 영위할 수 있도록 돕는다는 점에서 차별성이 있다. 이를 위해 별도의 장비나 센서가 필요하지 않고, 기존에 사용하던 스마트폰을 그대로 이용할 수 있다는 점 또한 선행 연구와는 다른 점이다. 식의약품 이외에 OCR을 통해 인식되는 모든 형태의 글자에 대해 원문 그대로 음성 안내를 제공하거나 사용자가 요청하면 요약해서 안내를 제공할 수도 있다는 차이점이 있다. Table 1은 2장에서 다룬 선행연구들과 본 연구를 비교하여 정리한 것으로, 시각장애인 대상 서비스에서 주로 등장하는 기능들과 연구별 차이점을 보여준다.

3. 시스템 구조

본 연구에서 제안하는 어플리케이션 “소리눈 (SoriNoon)”은 시각장애인 및 저시력자(이하 ‘사용자’)의 자립적인 생활을 돕기 위한 모바일 기반 보조 시스템이다. 사용자는 휴대폰 카메라와 음성 명령을 통해 다양한 정보를 입력할 수 있으며, 시스템은 실시간 객체 탐지, 경로 안내, 문서 인식 기능을 통해 시각적 정보를 음성으로 전달한다. 이러한 기능은 모두 Android 및 iOS를 지원하는 크로스플랫폼 환경(Flutter)[33] 위에서 동작하며, 다양한 AI 모듈과 서버 기술과 연동되어 구성된다.

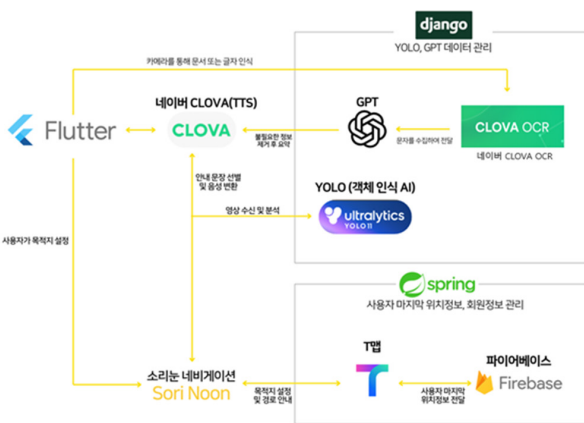


Figure 1. “SoriNoon” System Architecture

“소리눈”은 모바일 프론트엔드(Flutter), AI 처리 서버(Django)[34], 사용자 및 보호자 관리 서버(Spring)[35] 등 다양하게 구성된 클라이언트-서버 기반 아키텍처를 따른다. Fig. 1은 주요 기능의 흐름을 통합적으로 보여준다.

구현에 사용한 도구들은 Table 2와 같다. 함께 연동한 외부 서비스로는 네이버 CLOVA TTS[36], OCR[37]과 Tmap API[38], OpenAI의 GPT 3.5-turbo[39], Firebase[40]이 있다.

Table 2. Technologies to develop “SoriNoon”

Technology	Version
Django	5.1.3
YOLO	v11
GPT	3.5-turbo
Flutter	3.29.3
Dart	3.7.2
Spring Boot	3.4.3
Java	21

Fig 1은 “소리눈”의 전체 구조와 흐름에 대해 보여준다. 먼저 로그인 후 사용자가 길찾기를 수행하고자 하면 초기 화면의 가장 하단의 ‘안내 모드’를 터치하거나 마이크를 눌러 음성 명령을 내린 후 목적지를 입력하게 된다. Flutter는 설정한 목적지를 “소리눈” 네비게이션으로 보내게 된다. 실제로 길을 안내하게 될 때는 Tmap을 통해 경로 안내를 수행하게 된다. 이 과정에서 시각장애인 사용자의 위치 정보가 주기적으로 Firebase에 전송이 되어, 보호자가 원할 때 시각장애인의 가장 최근 위치를 확인할 수 있다.

보행 도중 갑자기 빠른 속도로 어떤 물체나 사람이 다가오게 될 경우, 경고 메시지를 음성으로 출력하게 되는데 이를 위해 YOLO v11이 사용되었다. 음성 출력은 네이버 CLOVA의 TTS를 통해 이루어지게 된다.

“소리눈”의 기능은 길찾기 이외에도 인식모드가 있다. “소리눈” 로그인 후 화면 중간의 ‘인식 모드’를 터치하거나 마이크를 통해 음성 명령을 내리면, 사용자의 휴대폰 카메라가 활성화된다. 사진 촬영 후 Flutter는 해당 사진을 네이버 CLOVA의 OCR로 전송하게 된다. 이를 통해 이미지 내의 글자가 추출되고, 해당 결과를 GPT에 전송하면 사용자가 이해하기 쉬운 형태로 설명이 생성된다. 생성된 설명은 네이버 CLOVA의 TTS를 통해 음성으로 변환되어 “소리눈”에 전송된다. “소리눈”은 해당 음성을 사용자의 휴대폰을 통해 출력하게 된다.

4. 시스템 기능

“소리눈” 어플리케이션은 시각장애인과 보호자를 위한 모바일 보조 기술로, 사용자는 카카오톡 연동 로그인 방식을 통해 간편하게 가입할 수 있다. 이는 아이디·비밀번호 기반의 본인 인증 절차를 생략하고 접근성을 높이기 위한 방안이다. 로그인 후에는 시각장애인 사용자와 보호자에 따라 서로 다른 UI가 제공된다. Fig. 2는 시각장애인 사용자가 로그인을 하면 나타나는 첫 화면이다.

시각장애인 모드에서는 인식 모드, 설정, 안내 모드 버튼이 상·중·하단에 위치하고 있으며, 좌측 하단에는 음성 인식 버튼이 배치되어 있어 주요 기능을 손쉽게 음성으로 실행할 수 있다. Fig. 3은 시각장애인 사용자가 설정 버튼을 눌렀을 때 나타나는 화면이다.



Figure 2. Visually Impaired User Home Screen



Figure 3. Visually Impaired User Setting Screen

4.1 음성 기반 경로 탐색 및 네비게이션 안내

시각장애인 사용자가 출발지와 도착지를 직접 입력하거나 음성 명령을 통해 설정하면, 시스템은 Tmap API의 도로 경로 탐색 기능을 활용하여 최적 경로를 계산하고, 실시간으로 음성 안내를 제공한다. 음성 기반 안내 모드의 경우, 안정적인 카메라 위치를 보장하기 위해 사용자가 자신의 스마트폰을 목걸이에 걸고 후면카메라를 사용자가 바라보는 방향에 둔 채로 사용하는 것을 가정하였다. 경로 안내는 Kotlin[41] 기반 로직을 통해 처리되며, 사용자 위치 기준의 출발지·도착지 마커, 선형 경로, 소요 시간 등의 정보가 XML(eXtensible Markup Language) 응답으로 전달된다. Fig. 4는 안내 모드로 들어가서 음성으로 출발지를 입력하는 화면이다.

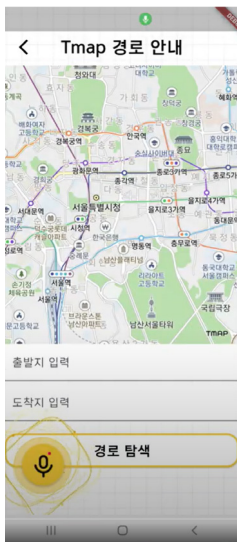


Figure 4. Navigation Screen

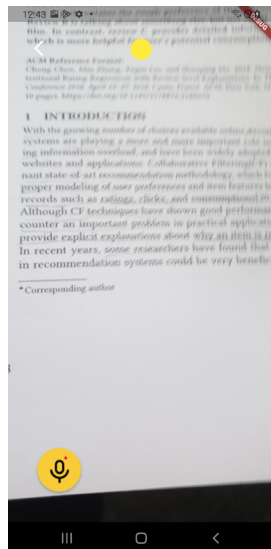


Figure 5. Recognition Mode (OCR)

시스템은 음성 인터페이스에 적합한 안내 문장을 추출하여, 예컨대 “오른쪽으로 이동하세요”, “목적지에 도착했습니다” 등의 텍스트를 TTS로 전달함으로써 시각 정보에 의존하지 않고도 이동 경로를 따라갈 수 있도록 돕는다. 또한, 장시간 사용을 고려해 저전력 모드를 함께 구현하였다.

4.2 객체 탐지 및 음성 경고 기능

사용자의 음성 명령으로 전면 카메라가 활성화되면, 실시간으로 객체 탐지 기능이 수행된다. 본 기능은 YOLO 기반의 딥러닝 알고리즘을 활용하여 사람, 자전거, 자동차, 전봇대, 소화전 등 주변의 위험 가능성이 있는 객체를 탐지하고, Tensor 기반 추론을 통해 가장 가까운 객체를 선별한다.

객체가 사용자의 3m 이내에 위치할 경우, Clova TTS 엔진을 통해 “사람이 앞에 있습니다”와 같은 경고 메시지를 음성으로 안내한다. 동일 객체에 대해 반복 경고를 방지하기 위해 시간 기반(5초) 중복 필터링도 적용되어 있다. 객체 정보는 WebSocket을 통해 Flutter 프론트엔드와 실시간 통신되며 응답성 또한 유지된다.

4.3 문서 인식 및 요약

시각장애인 사용자가 “전송” 또는 “요약” 음성 명령을 내리면, 앱은 현재 화면을 캡처하여 Django 백엔드의 OCR 요약 API로 이미지를 전송한다. 서버에서는 Clova OCR을 통해 텍스트를 추출하고, OpenAI GPT API를 통해 요약 또는 문장 정리를 수행한 뒤 결과를 앱으로 반환한다. Fig. 5는 인식 모드에 진입했을 때 나타나는 화면이다.



Figure 6. Guardian Mode Home Screen



Figure 7. Guardian QR Code Screen

사용자는 Flutter TTS를 통해 음성으로 해당 정보를 안내받는다. 이 기능은 상품 성분표, 공공기관 안내문 등 시각 정보 접근이 어려운 문서 인식에 활용될 수 있으며, 음성 중

심의 인터페이스를 통해 자유로운 정보 획득이 가능하다

4.4 보호자 위치 공유 기능

시각장애인 사용자는 어플리케이션 내에서 보호자 등록을 수행할 수 있으며, 등록된 보호자는 보호자 모드를 통해 시각장애인의 위치 정보를 확인할 수 있다. Fig. 6은 보호자로 로그인했을 때 보게 되는 홈 화면이다. 보호자는 QR 코드를 통해 인증받아야 하며, 사용자가 앱을 종료하기 전까지의 마지막 GPS 위치가 Firebase에 저장된다. Fig. 7은 QR 코드로 보호자 등록할 때 보게되는 화면이다.

이를 통해 보호자는 대상자의 최근 위치와 이동 경로를 확인할 수 있으며, 가족뿐 아니라 장애인 활동 지원사도 보호자로 설정할 수 있다. 보호자는 복수의 시각장애인을, 시각장애인은 복수의 보호자 설정이 가능하다. Fig. 8은 보호자가 복수의 시각장애인을 확인할 수 있는 화면이며, Fig. 9는 보호자가 시각장애인의 현재 위치 확인 시 나타나는 예시 화면이다.

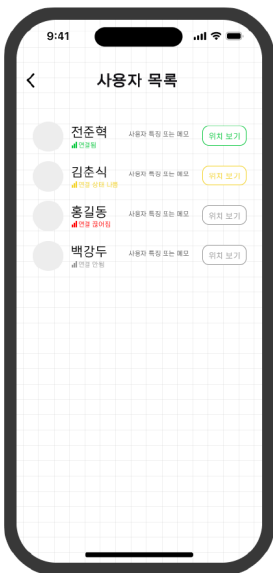


Figure 8. Protected Users List



Figure 9. Current Location of Protected User

5. 실험 및 결과

본 연구에서는 제안한 어플리케이션인 “소리눈(SoriNoon)”의 사용성을 평가하기 위하여 사용자 테스트를 계획하였다. 초기에는 실제 시각장애인을 대상으로 테스트를 수행하고자 하였으나, 연구진이 소속된 대학에는 시각장애인 재학생이 없었고, 시각장애인 사회복지관을 비롯한 관련 기관에 협조를 요청하였음에도 참여자를 확보하기 어려웠다. 이에 본 연구에서는 대안적으로 시각장애인 진료 경험이 풍부한 안과 전문의 1인(이하 전문가 참여자)을 대상으로 시연 및 사용성 테스트를 실시한 후 심층 인터뷰를 진행

하였다. 연구진 4인이 참여하여 “소리눈” 앱의 기획 배경을 설명하고, 튜토리얼 소개, 시연 및 테스트, 인터뷰를 순차적으로 진행하였으며, 전체 과정은 약 70분이 소요되었다.

인터뷰에 사용한 질문 목록은 다음과 같다.

Table 3. Interview Question List

#	Category	Question
1	Impressions and User Fit	What were your first impressions of the “SoriNoon” app after trying it today?
2		Do you think the app is suitable as a daily life assistance tool for low-vision or visually impaired users?
3		Which aspects of the app seem most helpful or beneficial for your patients?
4		Conversely, which parts of the app might feel challenging or difficult for users with low vision?
5	Visual Impairment Characteristics and UI/UX	Do you think the app’s color contrast (e.g., yellow logo and buttons) is sufficiently visible and easy to distinguish for low-vision users?
6		Is the font size adjustment feature adequate for typical low-vision patients, or would additional options be needed?
7		Do you believe the haptic feedback (vibration when touching buttons) would meaningfully help low-vision users?
8		Are the voice prompts and commands intuitive and easy to understand? How could they be improved?
9		Is the login and setup process simple enough for older or low-vision users?
10	Recognition Mode (OCR)	How useful would the text recognition and voice reading feature be for patients (e.g., reading medication labels or food packaging)?
11		Between the “read exactly as shown” and the “summarize” options, which one do you think patients would find more valuable?
12	Navigation Mode (T map Integration)	Do you think the voice-based input for starting point and destination is practical for low-vision users?
13		What additional safety features would you recommend for navigation? (e.g., crosswalk alerts, obstacle detection)
14		What do you think would be the biggest barrier for users when using navigation mode? (e.g., noisy environments, sudden pedestrians or obstacles)
15	Guardian Mode	Is the ability for caregivers to track the user’s real-time location useful for patients and their families?
16		What additional information or features could make the caregiver mode more helpful? (e.g., travel history, battery status)
17		Is the QR-based registration process intuitive enough for older caregivers?
18	Clinical Application and Recommendations	Would introducing this app during low-vision rehabilitation or patient education sessions be beneficial?
19		From a physician’s perspective, what should be considered before recommending this app to patients?
20		Do you see potential for collaboration with medical or rehabilitation centers? (e.g., low-vision specialists, orientation & mobility trainers)

#	Category	Question
21	Clinical Application and Recommendations	Are there any medical accuracy or safety concerns that should be addressed before wider clinical use?
22		What additional features could better reflect the diverse needs of low-vision patients? (e.g., mode by vision level, enhanced contrast options)
23	Future Improvements and Research Directions	In your opinion, what should be the top priority for improvement to make this app more practical for low-vision users?
24		Considering patients' varying digital literacy levels, what kind of support (e.g., tutorials, voice guidance style) would help them learn to use this app?
25		How do you evaluate the app's potential to enhance patients' quality of life (QoL) in the long term?
26		As a medical professional, do you have any additional advice or suggestions for improving the app?

전문가 참여자는 인터뷰 과정에서 다음과 같은 점을 “소리눈”의 긍정적인 기능으로 평가하였다.

“글자가 작은 경우에도 잘 읽어줘서, 일상생활은 할 수 있지만 시력이 많이 나빠서 돋보기를 들고 글자를 읽는 사람들에게 이 기능이 무척 좋을 것 같습니다.”

“요약 기능과 원문 그대로 읽어주는 기능은 사용자들이 필요에 따라 적절하게 선택하면 되니 좋을 것 같습니다. 병원에 내원하는 환자들 중 시각장애인의 경우, ‘동의서’를 읽어줘야 하는 경우가 있는데 이 어플을 사용하면 사용자가 스스로 카메라로 동의서를 인식시키면 해당 내용을 전부 읽어주니 편리할 것 같습니다.”

“길 안내 모드에서 ‘출발지’와 ‘도착지’를 각각 음성으로 입력하게 하는 방식은 실용적인 것 같습니다.”

“안내모드를 이용하면서 이동하던 도중, 사용자가 위험할 때 보호자에게 곧바로 전화를 걸 수 있는 SOS 모드가 있는 것은 좋은 것 같습니다.”

“QR 코드를 이용해서 보호자를 추가할 수 있도록 되어있는 점은 사용하기 쉽게 잘 만들어진 것 같아요.”

“일상 생활이 가능한 경증 시각장애인(4~6급)의 경우, 작은 글자를 읽기 힘들어한다면 인식 모드를 사용하는 것이 도움이 될 것 같습니다. 일상생활이 어려운 중증 시각장애인(1~3급)의 경우, 길 안내 모드도 도움이 될 수 있을 것 같습니다. 또한 장애인이 아니라고 하더라도, 시력이 많이 안 좋은 분들에게 인식 모드는 정말 도움이 될 것 같아요.”

이어 전문가 참여자는 다음과 같은 점을 아쉬운 점, 개선이 필요한 점으로 평가하였다.

“마이크 버튼이 너무 작습니다. 그리고 노란색이 이것보다는 조금 더 시각장애인 사용자의 눈에 들어오는 색으로 바뀌면 좋을 것 같습니다. 마이크 이외에 앱에서 (시각장애인) 사용자 모드에서 나오는 버튼들은 기술적인 이슈가 없다면 크게 할 수 있는 만큼 최대한 크게 하면 좋을 것 같습니다.”

“노란색이 다른 색에 비해 시각장애인이 잘 인식할 수 있는 색인 것은 맞지만, 지금 ‘소리눈’ 앱의 배경색인 하얀색과 같이 있을 때는 한 눈에 들어오는 색은 아닌 것 같아요. 흰색

배경을 유지하고 싶으면 더 눈에 들어오는 색깔로 버튼 색을 변경하고, 기존의 버튼 색인 노란색을 유지하고 싶다면 배경색은 흰색보다는 검정색이 나올 것 같습니다.”

“처음 실행하고 나서 인식 모드, 안내 모드 둘 다 글자 혹은 물체를 인식하는데 시간이 조금 오래 걸리는 것 같습니다.”

“길 안내 모드나 인식 모드의 경우 일반적으로 ‘밝은’ 상태에서 할 가능성이 높아 보이는데, 밝은 화면에서 노란색 마이크 버튼은 눈에 띄지 않습니다. 밝은 화면에서 눈에 띄는 색으로 바꾸는게 좋겠습니다.”

“화면의 글자 크기를 키우는 기능은 그냥 직관적으로 엄지와 검지 손가락 두 개로 확대하면 안되는지 궁금합니다.”

“길 안내에서 횡단보도 안내를 추가해주면 좋을 것 같습니다. 길을 가다가 갑자기 오토바이가 나오는 상황에, 어디쯤 가고 있는지 안내가 없는 상태이면 사용자가 불안할 것 같습니다.”

“측면에서 오는 차량을 인식할 수 있으면 좋겠습니다. 카메라가 측면을 인식할 수 있도록 광각모드로 안내하면 좋을 듯 합니다.”

이외에 제안 사항은 아래와 같았다.

“시각, 청각 장애를 가진 경우 그렇지 않은 사람들에 비해 치매에 빨리 걸린다는 연구 결과가 있습니다. 그래서 이런 어플이 있다고 하더라도 비장애인에겐 너무 쉬운 기능들이 그들에게는 어려울 수 있습니다. 키오스크 사용법을 교육하듯이 ‘소리눈’ 어플을 어떻게 사용할 수 있는지 교육을 해주면 더 잘 사용할 수 있을 것 같습니다.”

“보호자 모드에서 단순히 위치 추적만 된다면 기존에 있는 위치 추적 어플리케이션이 더 기능이 좋을 수도 있습니다. 보호자 입장에서 이 어플을 쓰는게 더 좋을 이유가 무엇이 있을지 고민해보면 좋을 것 같습니다.”

“시각장애인 안내견이 정말 똑똑하지만 모든 시각장애인들이 안내견과 함께 다니는 것은 아니기에, 앞서 언급했던 한계점들이 잘 개선되면 안내견을 대체할 수 있을 것 같습니다. 그리고 안내견은 글자를 읽어주진 못하죠. 그러면 안내견 훈련에 사용되는 비용이 절감될 수 있을 것이라고 기대되네요. 또 스마트폰만 있으면 사용 가능하니 접근성이 더 좋겠네요.”

6. 논의 및 결론

본 연구는 시각장애인의 자율성과 안전한 일상생활을 지원하기 위한 보조 어플리케이션 “소리눈”을 개발하고, 그 주요 기능과 기대 효과를 살펴보았다. “소리눈”은 시각장애인의 이동, 정보 접근, 보호자와의 소통을 하나의 플랫폼에서 지원함으로써 일상생활의 편의성과 안전성을 높이고자 설계되었다. 이를 위해 기존 시각장애인 지원 기술에서 지적된 여러 한계점을 개선하였다. 먼저, 기존 시각장애인 보조

기기가 스마트 지팡이, 라즈베리파이, 외부 센서 등 별도의 하드웨어에 의존했던 문제를 해결하고 스마트폰만으로도 사용할 수 있는 보조 시스템을 구현하였다. 또한 단순히 목적지까지의 경로만 안내하던 기존 길 안내 시스템과 달리 전방 물체를 탐지하고 위험 요소를 우선적으로 음성으로 경고함으로써 보행 안전성을 강화하였다. 정보 접근성 측면에서도 약품이나 식품과 같은 일상 물품의 텍스트를 OCR로 인식하여 음성으로 안내하고, 사용자의 필요에 따라 전체 읽기와 요약 안내를 선택할 수 있도록 하여 편의성을 높였다. 아울러 보호자와의 실시간 위치 공유 및 긴급 호출 기능을 제공함으로써 시각장애인의 안전 확보와 보호자와의 소통을 동시에 지원하였다.

그러나 본 연구에는 여전히 몇 가지 한계가 존재한다. 객체 인식의 범위가 제한적이어서 사람, 자동차, 자전거, 오토바이 등 일부 대상만 학습되어 있으며, 측면에서 접근하는 장애물을 감지하지 못하거나 동일 객체에 대해 반복적으로 경고하는 문제가 남아 있다. 또한 낮은 조도 환경이나 저해상도 카메라 사용 시 객체 인식 및 OCR 정확도가 저하될 수 있으며, 공공장소에서의 카메라 촬영은 주변인의 개인정보를 침해할 가능성이 있다. 더불어 실제 시각장애인을 대상으로 한 대규모 현장 테스트가 이루어지지 못해 사용자 경험 기반의 개선이 충분히 이루어지지 않은 점도 한계로 남는다.

본 연구는 시각장애인의 실질적 일상 문제를 통합적으로 해결하고자 한 시도라는 점에서 의미가 크다. 본 시스템은 물리적 이동을 지원하는 길 안내와 전방 물체 탐지, 정보 접근성을 높이는 문서 인식 및 음성 안내, 그리고 보호자와의 실시간 소통을 하나의 어플리케이션으로 통합하여 기존의 단일 기능 중심 보조 기술을 넘어섰다. 특히 스마트폰만으로 구현이 가능하여 별도의 고가 장비나 안내견 훈련과 같은 사회적·경제적 비용을 크게 절감할 수 있으며, 안내견이 제공하지 못하는 문서 인식 및 음성 안내 기능까지 제공함으로써 시각장애인의 정보 접근성을 실질적으로 확장할 수 있다. 또한 위험 요소를 우선 감지해 음성으로 경고하고, 사용자가 필요할 때 보호자에게 긴급 전화를 즉시 연결할 수 있도록 지원함으로써 자율성과 안전한 보행을 동시에 확보할 수 있다는 점에서도 가치가 크다. 이러한 특징은 향후 본 시스템이 안내견을 부분적으로 대체하거나 보완하는 실질적 대안이 될 수 있음을 시사하며, 사회적 비용 절감과 시각장애인의 독립적 생활을 동시에 촉진할 잠재력을 보여준다.

특히 본 연구는 영상 기반 객체 탐지(YOLO), 광학 문자 인식(OCR), 음성 인식(STT), 음성 안내(TTS), 모바일 내비게이션(T map API) 등 서로 다른 기술을 멀티모달 방식으로 유기적으로 연계하여 하나의 사용자 경험으로 통합했다는 점에서 의미가 크다. 기존 연구가 단일 기능을 개별적으로 제공하거나 특정 장비에 의존한 반면, 본 시스템은 다양한 입력(영상, 이미지, 음성, 텍스트)을 실시간으로 처리하고 상호 보완적으로 결합하여 보다 풍부하고 직관적인 지원을 가능하게 했다. 이는 멀티모달 통합이 시각장애인 보조

기술의 사용성, 접근성, 응용 가능성을 동시에 확장할 수 있음을 보여준다.

본 연구를 통해 구현된 “소리는” 어플리케이션은 시각장애인의 이동과 정보 접근성을 지원하는 통합적 보조 도구로서 가능성을 확인하였으나, 전문가 피드백을 바탕으로 여러 개선 과제가 도출되었다. 첫째, 사용자 인터페이스(UI)의 가시성과 직관성을 높이기 위한 디자인 개선이 필요하다. 현재 마이크 버튼이 작고 밝은 화면에서 노란색이 충분히 눈에 띄지 않는다는 의견이 제기되었으며, 버튼 크기를 가능한 한 크게 조정하고 배경색과 버튼 색의 대비를 재설계하여 저시력 사용자의 인지성을 향상시킬 필요가 있다. 둘째, 인식 모드와 안내 모드 모두 초기 인식 속도가 다소 느리다는 지적이 있어, 객체 탐지와 OCR 처리 속도를 향상시키기 위한 모델 최적화와 경량화가 요구된다. 셋째, 글자 크기 확대를 위한 제스처 기반 조작(예: 두 손가락 확대)을 지원하는 등 사용자 경험을 보다 직관적으로 개선할 필요가 있다. 넷째, 길 안내 기능 측면에서 횡단보도 안내, 측면에서 접근하는 차량이나 오토바이 인식, 광각 카메라 모드를 활용한 시야 확장 등 안전성 강화를 위한 보행 지원 기능의 고도화가 필요하다. 다섯째, 시각장애인 및 저시력자의 디지털 리터러시 수준을 고려한 체계적 사용 교육이 요구된다. 일반 사용자에게는 단순해 보이는 기능이라도 시각장애인에겐 학습이 필요할 수 있으므로, 키오스크 교육과 유사한 튜토리얼 및 교육 콘텐츠를 제공함으로써 앱의 활용성을 높일 수 있을 것이다. 여섯째, 보호자 모드의 기능적 차별화를 강화할 필요가 있다. 단순 위치 추적만 제공할 경우 기존 위치 추적 앱 대비 경쟁력이 낮을 수 있으므로, 긴급 호출 외에도 사용자 상태 모니터링, 경로 이탈 알림, 배터리 잔량 안내 등 보호자에게 실질적으로 도움이 되는 기능을 추가하는 방향이 제안된다.

이와 같은 개선을 통해 “소리는”은 보다 안전성, 접근성, 실시간성을 강화하고, 시각장애인과 보호자 모두에게 기존 솔루션보다 차별화된 가치를 제공할 수 있을 것이다. 특히 사용자 중심의 반복적 현장 검증을 통해 UI/UX와 기능을 지속적으로 고도화한다면, 본 시스템은 안내견의 역할을 보완하거나 일부 대체할 수 있는 실질적 보조 도구로 발전할 가능성이 있으며, 사회적 비용 절감과 시각장애인의 독립적 생활 지원에 기여할 수 있을 것이다.

참고문헌

- [1] Jeong, C., Kim, M., Yun, S., Moon, K. & Shin, H. (2024). Human Visual Ability Enhancement Technology Trends and Development Prospects. *ETRI Electronics and Telecommunications Trends*, 39(4), 63-72. <https://doi.org/10.22648/ETRI.2024.J.390407>
- [2] Doo, Y. (2025, April 17). Statistics on registered persons with disabilities, 2024. Ministry of Health and Welfare. https://www.mohw.go.kr/board.es?mid=a10503000000&bid=0027&list_no=1485363&act=view
- [3] Jeong, Y. (2024). AI-Based Multimodal Sensor Fusion Technology for Guiding Robots for the Visually Impaired. *Journal of the Institute of Electronics and Information Engineers*, 51(8), 23-29.
- [4] Lee, J. (2018, September 5). 2017 survey on the living conditions of persons with disabilities. Ministry of Health and Welfare. https://www.mohw.go.kr/board.es?mid=a10411010100&bid=0019&tag=&act=view&list_no=345972
- [5] Kim, M., Jeong, C., Moon, K., Park, Y., Bach, J. & Lee, J. (2020). Analysis of Preference Information on Visual-to-Audio Sensory Substitution for the Visually Impaired. *Journal of Information Technology and Applied Engineering*, 10(1), 51-61. <https://doi.org/10.22733/JITAE.2020.10.01.007>
- [6] Kim, K., Shin, G. & Cho, J. (2010). Visual Information to Tactile Transmission Device Development for Visually Impaired People. *Annual Conference of KIPS*, 17(1), 53-55. <https://doi.org/10.3745/PKIPS.y2010m04a.53>
- [7] Khan, W., Hussain, A., Khan, B., & Crockett, K. (2023). Outdoor mobility aid for people with visual impairment: Obstacle detection and responsive framework for the scene perception during the outdoor mobility of people with visual impairment. *Expert Systems with Applications*, 228, 120464. <https://doi.org/10.1016/j.eswa.2023.120464>
- [8] nkaranarayanan, R., Manjushree, R., Harshini, P., & Jeeshitha, V. (2024). Obstacle detection for visually impaired. *International Research Journal on Advanced Engineering and Management*, 2(4), 1200-1203. <https://doi.org/10.47392/IRJAEM.2024.0160>
- [9] Marston, R., & Gollidge, G. (1998). Improving transit access for the blind and vision impaired. *Intellimotion*, 7(2), 4-5 & 11.
- [10] Arditi, A. (2017). Rethinking ADA signage standards for low-vision accessibility. *Journal of vision*, 17(5), Article 8. <https://doi.org/10.1167/17.5.8>
- [11] Bambridge-Sutton, A. (2024, January 17). The consequences of poor communication in packaging for the blind and partially sighted. *Food Navigator*. <https://www.foodnavigator.com/Article/2024/01/17/Packaging-for-blind-and-partially-sighted-fails-to-impress/>
- [12] Kwon, K., Lim, Y., Park, C. & Kim, J. (2024). A Case Study on Information Accessibility of the AI-Based Visual Assistance App 'Sullivan Plus' for the Visually Impaired. *Information and Communications Magazine*, 41(11), 68-75.
- [13] Parcalabescu, L., Trost, N., & Frank, A. (2021). What is multimodality?. *arXiv preprint arXiv:2103.06304*.
- [14] University of Illinois Springfield. (n.d.). What is multimodal? *Learning Hub*. <https://www.uis.edu/learning-hub/writing-resources/handouts/learning-hub/what-is-multimodal>
- [15] Kress, G., & Van Leeuwen, T. (2020). *Reading images: The grammar of visual design*. Routledge.
- [16] Liang, P., Zadeh, A., & Morency, L. (2024). Foundations & trends in multimodal machine learning: Principles, challenges, and open questions. *ACM Computing Surveys*, 56(10), 1-42. <https://doi.org/10.1145/3656580>
- [17] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587). <https://doi.org/10.1109/CVPR.2014.81>
- [18] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448). <https://doi.org/10.1109/ICCV.2015.169>
- [19] Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6), 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [20] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788). <https://doi.org/10.1109/CVPR.2016.91>
- [21] Deng, S., Zhao, H., Fang, W., Yin, J., Dustdar, S., & Zomaya, A. (2020). Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence. *IEEE Internet of Things Journal*, 7(8), 7457-7469. <https://doi.org/10.1109/JIOT.2020.2984887>
- [22] Sipola, T., Alatalo, J., Kokkonen, T., & Rantonen, M. (2022). Artificial intelligence in the IoT era: A review of edge AI hardware and software. In *Proceedings of the 31st Conference of Open Innovations Association (FRUCT)* (pp. 320-331). IEEE. <https://doi.org/10.23919/FRUCT54823.2022.9770931>
- [23] Lee, S., Shin, H., An, S. & Yu, K. (2020). Development of Voice Assistant Apps for Low Vision People Based on Artificial Intelligence. *Journal of Digital Contents Society*, 21(2), 277-284. <http://dx.doi.org/10.9728/dcs.2020.21.2.277>
- [24] Kang, S., Park, S., Lee, J., Lee, Y. & Lee, W. (2020). A Proposal of Food and Medicine Application for the Visually Impaired. In *Proceedings of HICK 2020*, 1162-1165.
- [25] Kim, H., Kim, J., Song, G., Kim, E. & Lee, W. (2019). Study of Voice-Based Crowdsourcing Platform for the Enhancement of Self-Support for the Visually Impaired. In *Proceedings of HCKI 2019*, 187-192.
- [26] Lim, M., Back, S., Lee, J., Kwak, T., Kim, B. & Kim, C. (2024). YOLO-based learning object recognition system for the visually impaired. In *Proceedings of the 2024 Fall Conference of the Korean Institute of Communications*

and Information Sciences, 1290-1291.

- [27] Jeong, Y., Park, N. & Kim, D. (2020). Implementation of Smart Cane using Beacon Communication for Visually Impaired People. *Journal of Digital Contents Society*, 21(3), 453-561. <http://dx.doi.org/10.9728/dcs.2020.21.3.453>
- [28] Park, S., Park, J., Ryu, H., Lee, H. & Lee, K. (2020). Situation Awareness Auxiliary IoT Device and Pedestrian Navigation Service Application for Visually Impaired Persons. *Journal of Digital Contents Society*, 21(2), 269-275. <http://dx.doi.org/10.9728/dcs.2020.21.2.269>
- [29] Jeong, I., Ahn, E., Joo, H., Seo, Y., Jung, J. & Cho, J. (2017). Design of information support system with IoT for low vision people and blind. In *Proceedings of the 2017 Summer Conference of the Korean Institute of Communications and Information Sciences*, 1369-1370.
- [30] Yin, T., Zhou, X., & Krahenbühl, P. (2021, June). Center-Based 3D Object Detection and Tracking. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp.11784-11793). <https://doi.org/10.1109/CVPR46437.2021.01161>
- [31] RBQ-3. (n.d.). Rainbow Robotics. https://www.rainbow-robotics.com/en_rbq
- [32] Oh, S. & Bae, Y. (2021). Braille Block Recognition Algorithm for the Visually Impaired Based on YOLO V3. *Journal of Korean Institute of Intelligent Systems*, 31(1), 60-67. <http://dx.doi.org/10.5391/JKIIIS.2021.31.1.060>
- [33] Flutter. (2025, April 15). (Version 3.29.3). [Computer software]. <https://flutter.dev/>
- [34] Django. (2024, November 5). (Version 5.1.3). [Computer software]. <https://www.djangoproject.com/>
- [35] Spring Boot. (2025, February 20). (Version 3.4.3). [Computer software]. <https://spring.io/projects/spring-boot>
- [36] Naver Cloud Platform. (n.d.). CLOVA Voice [Text-to-speech service]. Retrieved from <https://www.ncloud.com/product/aiService/clovaVoice>
- [37] Naver Cloud Platform. (n.d.). CLOVA OCR [Optical character recognition service]. Retrieved from <https://www.ncloud.com/product/aiService/ocr>
- [38] TMAP Mobility Co. Ltd. (n.d.). TMAP API [Location-based services API platform]. Retrieved from <https://tmapapi.tmapmobility.com/>
- [39] OpenAI. (2025, January 25). GPT 3.5-turbo (Version gpt-3.5-turbo-0125) [Computer software]. <https://platform.openai.com/docs/models/gpt-3.5-turbo>
- [40] Firebase. (n.d.). Firebase [Platform and SDKs documentation]. Retrieved from <https://firebase.google.com/>
- [41] Kotlin. (n.d.). Kotlin. <https://kotlinlang.org/>



전준혁

- 2020년~현재 한성대학교 컴퓨터공학부 디지털콘텐츠 가상현실트랙 재학
- 2020년~현재 한성대학교 컴퓨터공학부 재학

✚ 관심분야 : User Experience & interaction Design, Mobile Application Development, Game Development

✉ onlymes1932@gmail.com



이영주

- 2019년~현재 한성대학교 컴퓨터공학부 웹공학트랙 재학
- 2019년~현재 한성대학교 컴퓨터공학부 재학

✚ 관심분야 : Mobile Application Development, Web Service Engineering, Cloud-Based Backend Technologies

✉ zeroweek4324@gmail.com



김규리

- 2022년~현재 한성대학교 컴퓨터공학부 웹공학트랙
- 2022년~현재 한성대학교 컴퓨터공학부 재학

✚ 관심분야 : User Interface / User Experience (UI/UX) & Interaction Design, User-Friendly Experience, Component-Based Design

✉ kuyrik1@naver.com



최준희

- 2022년~현재 한성대학교 컴퓨터공학부 빅데이터 트랙
- 2022년~현재 한성대학교 컴퓨터공학부 재학

✚ 관심분야 : NLP, ML/DL, Recommender Systems

✉ 2271314@hansung.ac.kr



유수진

- 2010년 고려대학교 컴퓨터교육과(이학사)
- 2012년 고려대학교 일반대학원 컴퓨터교육학과 (이학석사)
- 2020년 고려대학교 일반대학원 컴퓨터학과 (공학 박사)
- 2015년 2월~2017년 5월 롯데카드 IT기획팀, IT비즈니스
- 2017년 10월~2018년 12월 Stipop 폴스택 개발자, CTO
- 2021년 3월~2021년 12월 한양대학교 SW융합원 SW 교육전담교수
- 2022년 2월~2022년 12월 성균관대학교 소프트웨어 융합대학 초빙교수
- 2023년 1월~2024년 8월 고려대학교 데이터과학원 연구교수
- 2024년 9월~현재 한성대학교 컴퓨터공학부 웹공학 트랙 조교수

✚ 관심분야 : 지식그래프, 링크드데이터, 컴퓨터교육, 인공지능 윤리, 디지털격차

✉ sujin.yoo@hansung.kr