

시각장애인을 위한 XR 글래스 기반 보행 장애물 감지 및 실시간 원격 도우미 통합 보행 지원 시스템

고민준*, 조진수**

XR Glasses-based Integrated Walking Assistance System with Obstacle Detection and Real-Time Remote Support for the Visually Impaired

Minjun Ko*, Jinsoo Cho**

이 논문은 정부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2025-2396318312982076870001).

요 약

본 연구는 시각장애인의 보행 안정성과 이동 자율성 향상을 위한 XR 글래스 기반 통합 보행 지원 시스템을 제안하였다. 제안 시스템은 XR 글래스를 통해 사용자 시점 영상을 획득하고, 스마트폰 기반 온디바이스 AI에서 객체 탐지 및 위험도 분석을 수행한 뒤 TTS 기반 음성 안내를 제공한다. 또한 WebRTC 기반 실시간 통신을 통해 원격 도우미와 영상·음성 연결을 지원하여 예외 상황에 대응할 수 있게 한다. 객체 탐지 모델은 경량 모델 비교를 통해 YOLOv8-Nano를 선정하였으며, 모바일 환경에서 실시간 처리 성능을 확보하였다. 실험 결과, 전체 처리 시간은 약 0.16초, 객체 탐지 정확도는 95.4%, 통신 응답 시간은 0.04초로 나타났다. 사용성 평가에서는 보행 안정성과 원격 도우미 서비스에서 높은 만족도를 보였으며, 평균 4점 이상의 사용성을 보였다.

Abstract

This study proposes an XR glasses-based integrated walking assistance system for visually impaired individuals. The system captures first-person video and performs object detection and risk analysis using an on-device AI module, providing voice guidance via TTS. A WebRTC-based framework enables real-time communication with a remote helper for handling unexpected situations. The system integrates perception and remote assistance to support safe walking in complex environments. YOLOv8-Nano was selected as the optimal lightweight model for real-time performance in mobile environments. Experimental results show a processing time of 0.16 seconds, 95.4% detection accuracy, and 0.04-second communication latency. User evaluation indicates high satisfaction in walking safety and remote assistance, with overall usability above 4.0.

Keywords

visually impaired, XR glasses, on-device AI, object detection, YOLOv8, remote assistance

* 가천대학교 컴퓨터공학과 학사과정
- ORCID: <https://orcid.org/0009-0005-0408-5368>
** 가천대학교 컴퓨터공학과 교수(교신저자)
- ORCID: <https://orcid.org/0000-0002-8119-8068>

• Received: Mar. 31, 2026, Revised: Apr. 23, 2026, Accepted: Apr. 26, 2026
• Corresponding Author: Jinsoo Cho
Dept. of Computer Engineering, Gachon University, Korea
Tel.: +82-31-750-5820, Email: jscho@gachon.ac.kr

I. 서 론

시각장애는 전 세계적으로 상당한 규모의 인구에게 영향을 미치고 있다. GBD(Global Burden of Disease) 연구에 따르면 2020년 기준 약 4,330만 명이 전맹이며 약 2억 9,500만 명이 중등도에서 중증의 시각장애가 있는 것으로 보고되었다[1]. 이러한 시각장애 부담은 인구 고령화와 만성질환 증가 등으로 인해 앞으로 더욱 증가할 것으로 전망된다.

보행은 공간 인지, 장애물 탐지, 위험 회피, 방향 결정을 포함하는 복합적 활동이다. 그러나 시각 정보가 제한되면 이러한 과정 수행이 어려워지며, 시각장애인은 일반 인구보다 교통사고 위험이 약 46% 높고 보행 중 충돌이나 낙상 사고도 빈번한 것으로 보고되었다[2]. 이에 따라 다양한 보행 방식과 보조 수단이 활용된다.

최근에는 보행 지원의 필요성에 따라 센서 기반 장애물 탐지, 딥러닝 기반 객체 인식, 촉각 및 음성 피드백을 활용한 다양한 보행 보조 기술이 제안되고 있다[3]-[5]. 또한 원격 도우미와 연결하여 시각 정보를 제공하는 연구도 진행되고 있다[6]. 그러나 이러한 기술들은 각각 특정 기능에 초점을 맞추고 있어 보행 중에 발생할 수 있는 다양한 상황에 종합적으로 대응할 수 있는 통합형 보행 지원 시스템은 아직 부족한 실정이다.

이러한 한계를 보완하기 위해 본 연구에서는 XR 글래스를 활용한 보행 장애물 감지, 음성 피드백, WebRTC 기반[7] 실시간 원격 도우미 서비스를 통합한 시각장애인 보행 지원 시스템을 제안한다.

본 논문의 핵심 기여는 다음과 같다.

- 객체 탐지, 위험도 분석, 음성 피드백으로 구성된 자동 인식 파이프라인과 WebRTC 기반 원격 도우미 서비스를 통합 운영하는 시스템 아키텍처를 설계한다.
- 경량 객체 탐지 모델 비교를 통해 YOLOv8-Nano[8]를 선정하고, 온디바이스 AI 추론을 모바일 환경에서 안정적으로 구현하여 실시간 처리 성능을 확보한다.
- 시각장애인을 대상으로 한 2단계 시범서비스를 통해 제안 시스템의 실시간 처리 성능과 사용성을 정량적으로 검증한다.

본 논문의 구성은 다음과 같다. 2장에서는 시각장애인의 보행 방식을 살펴보고, 보행 보조 기술과 XR 글래스 기술 동향을 분석한다. 3장에서는 제안한 시스템의 전체 아키텍처를 설명하고, XR 글래스 기반 보행 장애물 인식 모듈의 설계와 실시간 원격 도우미 서비스의 구조 및 동작 방식을 기술한다. 4장에서는 시스템 성능 평가와 사용성 평가를 통해 제안 시스템의 유효성을 검증하고 결과를 분석한다. 마지막으로 5장에서는 연구 내용을 종합하여 결론을 도출하고 향후 연구 방향을 제시한다.

II. 관련 연구

2.1 시각장애인 보행 방식

시각장애인의 보행 방법은 다양하며, 전문적인 보행 훈련을 통해 안전하고 효율적인 이동이 가능하다. 시각장애인의 보행 방식은 그림 1과 같이 안내보행, 흰지팡이 보행, 안내견 보행, 전자기기 기반 보행으로 구분된다[9].



그림 1. 시각장애인의 보행 방식

Fig. 1. Mobility methods for visually impaired individuals

그림 1(a)에서 안내보행은 보호자나 주변인의 도움을 받는 방식으로 안정성이 높지만, 타인 의존성이 높아 독립성 확보에 한계가 있다. 또한 인건비

등 운영 비용이 높아 지속적인 지원에 제약이 존재한다. 그림 1(b)에서 흰지팡이 보행은 시각장애인이 가장 널리 사용하는 독립 보행 방식으로, 일정 기간의 보행 훈련을 통해 환경 탐지와 방향 감각을 향상시킬 수 있다[10]. 그러나 탐지 범위가 주로 지면 중심으로 제한되며, 머리 높이의 장애물이나 일정 거리 이상의 객체를 인지하기 어렵다는 한계가 있다[11]. 그림 1(c)에서 안내견 보행은 훈련된 안내견이 시각장애인의 이동을 보조하는 방식으로 높은 이동 안정성과 심리적 안정감을 제공한다. 하지만 장기간의 훈련과 높은 양성 비용으로 인해 공급 가능한 안내견 수가 제한적이라는 현실적 제약이 따른다[12]. 그림 1(d)에서 전자기기 기반 보행은 다양한 센서를 활용하여 장애물을 탐지하고 정보를 사용자에게 전달하는 방식으로, 지팡이, 안경, 목걸이 등 다양한 웨어러블 형태의 보행 보조장치로 개발되고 있다.

2.2 시각장애인 지원 보조 기술 동향

센서 및 인공지능 기술의 발전에 따라 전자기기 기반 보행 보조 기술이 빠르게 확장되고 있으며, 기존 연구는 객체 인식, 감각 피드백, 원격 영상 지원 등 기능 단위로 진행되어 왔다.

기존 연구를 살펴보면, 음성 및 촉각 피드백을 결합하여 보행 중 장애물 회피 성능을 향상시키고자 한 연구가 있으며, 이는 멀티모달 피드백을 통해 사용자 반응 속도를 개선하는 데 기여하였다[3][4]. 그러나 이러한 연구는 가상 환경 기반 실험이 중심이거나, 장애물의 존재 여부 전달에 초점을 맞추어 실제 환경에서의 종합적인 상황 인지까지는 확장되지 못한 한계가 있다.

또한 초음파, ToF(Time-of-Flight) 센서 등을 활용한 촉각 중심의 웨어러블 장치 연구에서는 장애물까지의 거리 정보를 진동으로 전달하는 방식이 제안되었으나[5], 객체의 종류나 맥락 정보는 제공되지 않아 정보 해석의 직관성이 제한될 수 있다.

이러한 인식 기반 시스템의 한계를 보완하기 위한 접근으로, 원격 지원 기반 연구 역시 제안되고 있다. 증강현실 또는 영상 스트리밍을 활용하여 시각장애인이 원격 도우미와 연결되고, 실시간 음성 안내를 받을 수 있는 서비스 구조가 제안되었다[6].

그러나 이러한 접근은 자동화된 환경 인식 기능과의 통합이 이루어지지 않았으며, 보행 중 지속적인 지원 체계로 활용하기에는 한계가 존재한다. 이와 같은 한계를 고려할 때, 자동 환경 인식과 원격 도우미 지원을 통합한 보행 지원 시스템 연구는 여전히 제한적이며, 본 연구는 이러한 통합적 접근에서 차별성을 가진다.

2.3 XR 글래스 기술 및 응용

XR은 가상현실, 증강현실, 혼합현실을 포괄하는 몰입형 기술로, 현실 환경과 가상 정보를 결합하여 확장된 상호작용 환경을 제공한다[13]. 이러한 XR 기술은 디스플레이 장치, 공간 인식 센서, 컴퓨터 비전, 음성 인터페이스 등 다양한 요소의 융합을 통해 발전해왔으며, 산업, 교육, 의료 등 여러 분야에서 활용된다.

특히 XR 글래스는 사용자 시점의 영상을 실시간으로 획득할 수 있는 안경형 웨어러블 장치로, 컴퓨터 비전 및 온디바이스 AI 기술과 결합하여 객체 인식과 환경 분석을 실시간으로 수행할 수 있어 보행 환경 인식 분야에서의 활용 가능성이 높다.

보행 보조기술에서는 환경 정보의 획득, 분석, 그리고 피드백 전달 과정이 연속적으로 이루어지기 때문에 지연을 최소화하고 환경 인식 결과를 사용자에게 신속하게 전달하는 것이 중요하다. 이러한 측면에서 XR 글래스는 사용자의 시점 영상을 직접 획득하고 온디바이스 AI와 결합하여 환경 정보를 빠르게 분석할 수 있어 시각장애인을 위한 보행 보조기술에 적용될 수 있는 잠재력을 가진다.

III. XR 글래스 기반 보행 장애물 감지 및 실시간 원격 도우미 통합 보행 지원 시스템

3.1 전체 시스템 구조 및 구성

제안 시스템은 보행 환경 인식과 원격 지원 기능을 통합하여 시각장애인의 보행 안전성과 이동 자율성을 지원한다. 사용자 주변 환경을 실시간으로 인식하면서 필요 시 원격 도우미의 지원을 받을 수 있도록 구성하였다. 전체 구조는 그림 2와 같다.

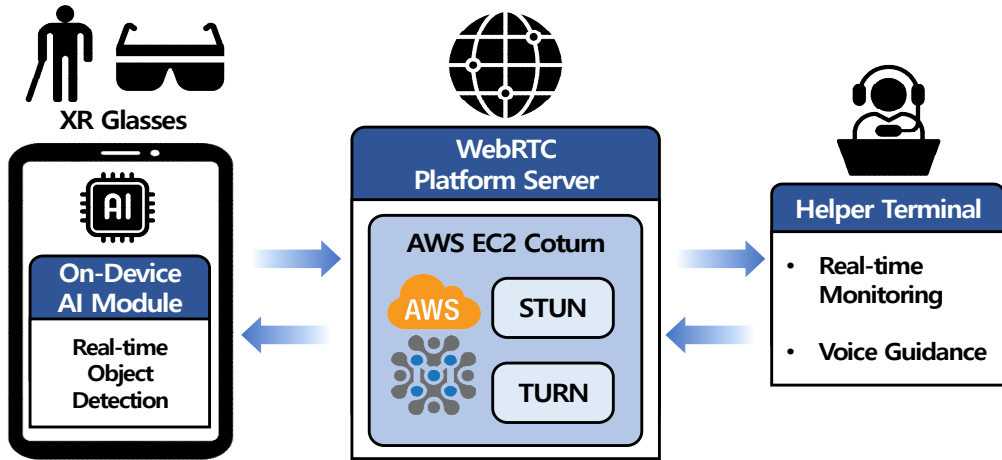


그림 2. XR 글래스 기반 통합 보행 지원 시스템 구조

Fig. 2. Architecture of the XR Glasses-based integrated walking assistance system

그림 2에 나타난 바와 같이, 제안 시스템은 XR 글래스, 스마트폰(온디바이스 AI 모듈), WebRTC 플랫폼 서버, 도우미 단말의 네 가지 구성 요소로 이루어진다. XR 글래스로 획득한 사용자 시점 영상은 스마트폰의 온디바이스 AI 모듈에서 처리되어 음성 안내로 사용자에게 제공되며, 동일 영상은 WebRTC를 통해 도우미 단말로 전달되어 필요 시 원격 지원으로 연결된다.

3.2 보행 장애물 감지 모듈

보행 장애물 감지 모듈은 XR 글래스를 통해 획득한 사용자의 시점 영상을 기반으로 보행 환경의 위험 요소를 탐지하고, 탐지된 결과를 사용자에게 전달하는 역할을 한다. 본 모듈은 객체 탐지, 위험도 분석, 음성 피드백 생성의 세 단계로 구성되며, 이를 통해 시각장애인이 보행 중 주변 환경의 위험 요소를 실시간으로 인지할 수 있도록 설계하였다.

모델은 YOLOv8-Nano를 사용하였으며, 보행 중 충돌 위험이 있는 고정 및 이동 장애물을 탐지 대상으로 선정하였다. 고정 장애물은 블라드, 전봇대 등 8개 클래스, 이동 장애물은 자동차, 보행자 등 7개 클래스로 구성된다. 또한 각 클래스별로 5만 장 이상의 데이터를 라벨링하여 데이터셋을 구축하고, 이를 기반으로 모델을 학습하였다.

그림 3에 나타난 바와 같이, 인공지능 모델은 모바일 환경에서 동작하는 구조로 이루어진다. 모델은 PyTorch로 학습한 후 양자화 및 ONNX 변환을 거쳐 스마트폰에서 실행된다. XR 글래스가 수집한 영상은 Unity 엔진으로 전달되어 Android 환경에서 처리 가능한 형태로 전처리된 뒤 모델에 입력되며, 모델은 신뢰도(Confidence) 0.5 이상의 후보 객체에 대해 NMS(Non-Maximum Suppression) IoU 임계값 0.45를 적용하여 중복 박스를 제거하고, 클래스·바운딩 박스·신뢰도 정보를 출력한다. 출력 결과는 Unity에서 데이터 가공 및 사용자 인터페이스 처리를 거쳐 그림 4와 같이 XR 글래스 화면에 실시간으로 시각화된다.

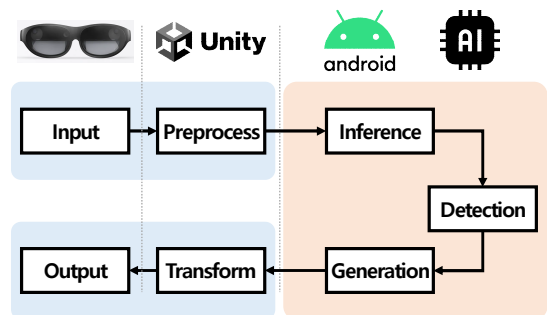


그림 3. 인공지능 모바일 탑재 및 연동 구조

Fig. 3. Mobile deployment architecture of the AI model

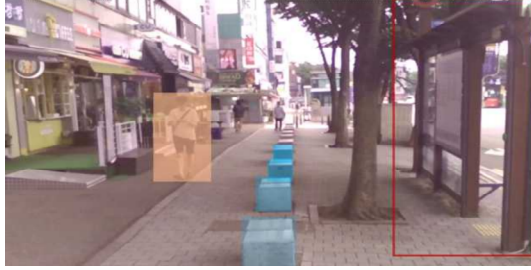


그림 4. 객체 인식 결과 시각화 예시
Fig. 4. Visualization of object detection results

탐지된 객체는 단순 표시를 넘어 보행 안전을 고려한 위험도 분석에 활용한다. 장애물의 종류, 위치, 크기, 거리 등을 반영하여 위험도를 산출하며, 여러 객체가 동시에 탐지될 경우 과도한 증가를 방지하기 위해 RSS(Root Sum Square)[14] 기반 결합 방식을 적용하였다. 이를 통해 다수의 객체가 존재하는 환경에서도 위험도를 안정적으로 계산할 수 있으며, 종합 위험도는 다음 식 (1)과 같이 정의된다.

$$R_{total} = \sqrt{\sum_{i=1}^n (W_{type}^{(i)} \times W_{pos}^{(i)} \times W_{prox}^{(i)})^2} \quad (1)$$

여기서 n 은 탐지된 장애물의 개수를 의미하며, W_{type} 은 장애물 유형에 따른 가중치, W_{pos} 는 사용자의 보행 방향 대비 객체의 위치에 따른 가중치, W_{prox} 는 객체의 크기와 사용자와의 거리를 반영한 근접도 가중치를 의미한다. 각 가중치는 각 요소가 보행 안전에 미치는 상대적 영향을 고려하여 설정

하였다. 이러한 요소를 종합하여 보행 환경의 위험도를 산출하며, 계산된 위험도는 안전, 주의, 위험, 심각한 네 단계로 분류된다. 위험도가 일정 임계값 이상일 경우 해당 객체 정보를 음성 피드백 생성 과정으로 전달한다.

음성 피드백은 객체 탐지 결과를 기반으로 생성된 정보를 사전에 정의된 [위치]-[객체]-[위험도] 구조의 안내 템플릿에 적용한 후, TTS(Text-to-Speech) 기술을 통해 음성으로 변환하여 제공한다. 예를 들어 “전방에 차량이 감지되었습니다. 위험 단계입니다.”와 같은 안내를 생성한다. 이를 XR 글래스의 스피커를 통해 사용자에게 전달하며, 사용자는 보행 중 주변 환경의 위험 요소를 청각적으로 인지할 수 있다.

3.3 실시간 원격 도우미 서비스

실시간 원격 도우미 서비스는 자동 장애물 인식만으로 대응하기 어려운 상황에서 원격 도우미와 연결하여 사용자 시점 영상을 공유하고 음성 안내를 받을 수 있도록 지원한다. 사용자 단말과 도우미 단말은 WebRTC 기반 P2P 통신 구조로 연결하며, 초기 연결 협상을 위해 그림 5와 같이 Firebase Realtime Database를 시그널링 서버로 활용하여 SDP(Session Description Protocol)와 ICE Candidate 정보를 교환한다.

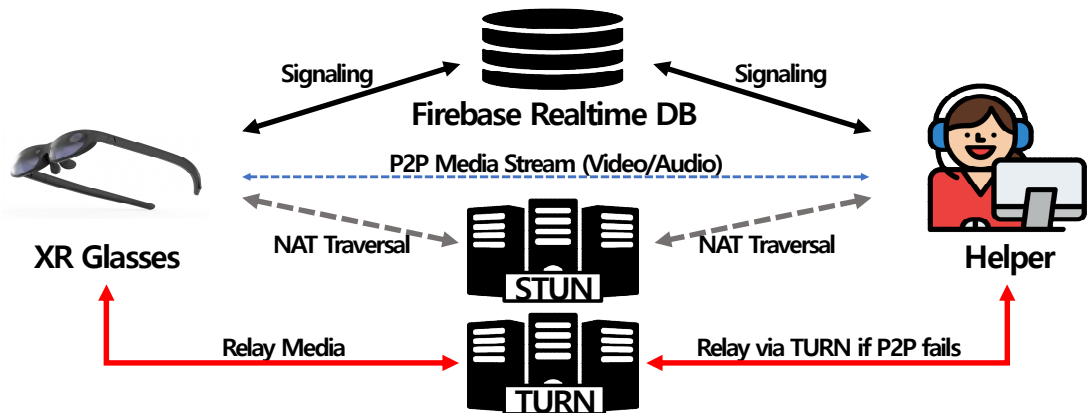


그림 5. WebRTC 기반 원격 도우미 서비스 통신 구조
Fig. 5. WebRTC-based communication architecture for the remote helper service

WebRTC 연결은 P2P 직접 연결을 우선 시도하며, 대칭 NAT(Symmetric NAT) 또는 방화벽 환경에서는 STUN/TURN 중계로 자동 전환한다. 이를 통해 이동 환경에서도 안정된 통신 연결을 유지한다.

XR 글래스로 촬영된 사용자 시점 영상은 전용 미들웨어인 NRSDK의 RGB 카메라 인터페이스로 이미지 센서에 직접 접근하여 처리 지연을 최소화한다. 이후 NRKernel의 RGB 텍스처 변환과 Android MediaCodec[15] 기반 하드웨어 가속 인코딩을 거쳐 압축한다. 압축된 영상을 WebRTC를 통해 도우미 단말로 전송하며, 기기 성능과 네트워크 상태에 따라 해상도를 동적으로 조절하여 시스템 부하를 감소시킨다.

사용자는 보행 중 추가 지원이 필요한 경우 스마트폰 애플리케이션을 통해 원격 지원을 요청할 수 있다. 요청 시 WebRTC 기반 통신으로 도우미 단말과 연결하고, 실시간 영상·음성 안내를 통해 예외 상황에 대응할 수 있다. 도우미 단말에서는 그림 6과 같이 사용자 시점 영상을 실시간으로 확인하며, 이를 기반으로 보행 환경을 파악한다.

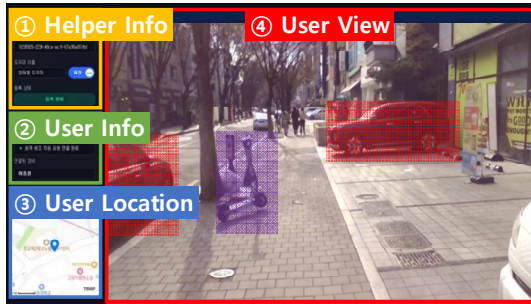


그림 6. 도우미 단말 실시간 영상 확인 화면
Fig. 6. Real-time video view of the helper terminal

이러한 구조를 통해 복잡한 교차로, 공사 구간 등 자동 인식 시스템이 대응하기 어려운 상황에서도 사용자와 도우미가 양방향으로 상호작용할 수 있다. 사용자는 도우미의 안내를 즉시 인지하고 반응할 수 있어, 단독 보행 시의 예외 상황 대응력이 향상된다.

IV. 시스템 성능 평가 및 사용성 평가

본 연구의 시스템 성능 평가에 사용된 실험 환경 및 주요 장비 사양을 표 1에 정리하였다.

표 1. 실험 환경 및 시스템 사양

Table 1. Experimental setup and system specifications

Category	Item	Specification
PC	CPU	Intel Core i7-9700K
	RAM	128GB
	GPU	NVIDIA RTX 2080 Super
	OS	Ubuntu 18.04.5 LTS
	Pytorch	1.10.2 + cu102
XR Glasses (XREAL Light)	Display resolution	1920×1080 pixels
	Field fo View	52°
	Refresh rate	120 Hz
Smartphone (Samsung Galaxy S23)	Processor	Qualcomm Snapdragon 8 Gen 2
	Memory	8 GB RAM
	OS	Android 13
Server	Platform	AWS EC2 Coturn

4.1 객체 탐지 모델 성능 평가

모바일 환경에서 실시간 객체 탐지를 위해 정확도, 처리 속도, 자원 사용량을 고려한 모델 선택이 필요하다. 이에 경량 객체 탐지 모델을 대상으로 성능 평가를 수행하고 최적 모델을 선정하였다. 성능 평가를 위해 YOLO, MobileNet, EfficientDet, PP-YOLO 계열 및 TinyML 모델을 비교 대상으로 선정하였다. 각 모델은 COCO 데이터셋[16]에서 약 35,000장의 학습 데이터와 5,000장의 검증 데이터를 샘플링하여 학습 및 평가를 수행하였으며, mAP, FPS, 모델 크기, RAM 사용량을 기준으로 성능을 측정하였다.

최종 모델 선정을 위해 각 성능 지표를 정규화한 후 가중치를 부여하는 평가 방식을 적용하였으며, 식 (2)와 같이 최종 점수를 산출하였다.

$$S = 0.4F + 0.35M + 0.15R + 0.1C \quad (2)$$

여기서 F 는 FPS, M 은 mAP, R 은 RAM 효율, C 는 모델 크기 효율을 의미한다. 가중치는 보행 지원 도메인의 우선순위에 따라 설정하였다. 안전한 안내를 위해 약 30 FPS 이상의 처리 속도가 요구되므로 응답성에 직결되는 FPS에 0.4를 부여하였고, mAP는 후단의 위험도 분석·안내 템플릿이 일부 오차를 보정할 수 있는 점을 고려하여 0.35로 설정하였다. RAM 효율과 모델 크기 효율은 최근 스마트폰 사양 향상으로 제약이 완화된 점을 반영하여 각각

0.15, 0.1의 보조 지표로 두어, 안전·실시간성과 자원 효율을 3:1 비중으로 평가하였다. 가중치 기반 평가를 통해 산출된 결과는 표 2에 나타난다.

표 2에서 확인할 수 있듯이, YOLOv8n 모델은 약 45FPS의 처리 속도와 73.5%의 mAP를 기록하여 모바일 환경에서 실시간 처리에 적합한 성능을 나타냈다. 반면, PP-YOLOv2 모델과 EfficientDet 계열 모델은 상대적으로 높은 정확도를 보였으나 처리 속도는 낮은 경향을 보였다. 또한 TinyML 모델은 약 50FPS로 가장 높은 속도와 200MB의 RAM 사용량, 4.0MB의 모델 크기를 기록하여 자원 효율성 측면에서는 우수한 성능을 보였으나, 정확도가 낮아 전체적인 성능 균형 측면에서는 한계를 보였다.

모델별 성능 측정 결과를 종합적으로 비교한 결과, YOLOv8n 모델이 89.2점으로 가장 높은 점수를 기록하였다. 이는 전체 모델 평균 점수(약 82.2점) 대비 약 8.5% 높은 수준으로, 정확도, 속도, 자원 효율성 측면에서 균형 잡힌 성능을 보여 모바일 환경에서의 객체 탐지에 적합한 모델로 판단된다.

표 2. 모델별 성능 측정 결과

Table 2. Model performance evaluation results

Model	mAP (%)	FPS	Model size (MB)	RAM usage (MB)	Final score
YOLOv8n	73.5	45	6.2	350	89.2
YOLOv8s	78.2	38	11.2	450	87.0
MobileNetV2	70.5	40	12.0	400	83.5
MobileNetV3 Large	72.1	35	14.3	420	81.6
MnasNet	70.5	38	13.0	420	81.7
EfficientDet D0	77.0	30	15.5	500	79.8
EfficientDet D1	79.0	25	20.1	600	76.1
PP-YOLO Tiny	74.2	40	10.0	350	85.4
PP-YOLOv2	80.0	32	18.0	550	79.5
TinyML	65.0	50	4.0	200	78.6

4.2 시스템 성능 평가

보행 보조 시스템에서 신속하고 안정적인 정보 전달은 사용자 안전과 직결되는 요소이다. 따라서 보행 보조 시스템은 사용자에게 위험 정보를 인지

및 대응 가능한 시간 범위 내 제공할 수 있는 실시간성을 확보해야 한다.

제안 시스템의 성능을 분석하기 위해 처리 시간과 정보 추출 정확도, 데이터 전송 정확도 및 시스템 신뢰성을 평가하였다. 보행 장애물 감지 과정은 객체 탐지, 위험도 분석, 안내 문장 생성, TTS 변환의 4단계로 구성되며, 각 단계의 처리 시간을 독립적으로 측정하였다.

평가는 XR 글래스로 촬영된 이미지 1,000장을 기준으로 수행하였으며, 각 이미지에 대해 전체 처리 과정을 1회씩 수행하여 성능을 측정하였다. 각 처리 단계의 평균 소요 시간은 그림 7과 같다.

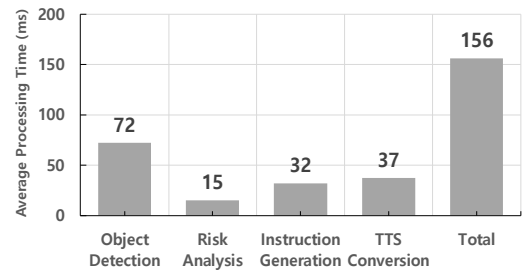


그림 7. 보행 장애물 감지 단계별 평균 처리 시간

Fig. 7. Average processing time per stage for obstacle detection

객체 탐지 단계는 평균 72 ms, 위험도 분석 단계는 15 ms, 안내 문장 생성 단계는 32 ms, TTS 변환 단계는 37 ms의 처리 시간이 소요되는 것으로 나타났다. 이를 종합한 전체 시스템 처리 시간은 약 0.16 s로 확인되었다. 또한, 본 연구에서 구축한 보행 장애물 데이터셋을 기반으로 학습한 모델의 보행 장애물 감지 정확도는 약 95.4%로 나타났다. 이는 제안한 모듈이 실시간 처리 성능과 인식 정확도를 동시에 확보하였음을 보여주며, 실제 보행 환경에서의 적용 가능성을 시사한다.

비대면 도우미 서비스 플랫폼의 응답 성능 분석 결과는 표 3과 같다. 동일한 사양의 XR 글래스와 스마트폰 5대를 이용하여 다중 클라이언트 환경을 구성하고 성능 평가를 진행하였다. 각 단말에서는 실제 서비스 환경을 모사하기 위해 영상 및 음성 데이터를 전송함과 동시에 일정 주기로 다수의 요청을 반복적으로 발생시켜 다중 사용자 환경에서의 서버 부하 상황을 재현하였다. 또한 네트워크 기반

성능 지표는 측정 환경에 따라 변동이 발생할 수 있으므로, 동일 조건에서 수집된 결과를 바탕으로 평균값을 산출하여 성능을 평가하였다.

표 3. 모델별 성능 측정 결과

Table 3. Model performance evaluation results

Metric	Value	Description
Average response time	0.04s	End-to-end latency between request and response
Throughput	3,125 req/s	Number of processed requests per second
Data transmission accuracy	99%	Ratio of correctly received data
System reliability	99%	Ratio of successfully processed requests

평균 응답 속도는 요청 발생 시점부터 서버 응답 수신 시점까지의 지연 시간을 기준으로 산출하였으며, 약 0.04초 수준으로 나타났다. 이는 실시간 콘텐츠 제공 서비스에서 요구되는 응답 지연 기준을 충족하는 수준으로, 사용자와 원격 도우미 간 즉각적인 정보 전달이 가능함을 의미한다. 동시 접속 처리량은 단위 시간당 처리된 요청 수를 기준으로 계산하였으며, 평균 3,125건/s로 분석되었다. 이를 통해 다수의 사용자가 동시에 접속하는 환경에서도 안정적인 서비스 제공이 가능함을 확인할 수 있다. 또한 데이터 전송 정확도와 시스템 신뢰성은 각각 약 99% 수준으로 나타나, 다양한 환경에서도 안정적인 동작 성능을 확보한 것으로 판단된다.

4.3 사용성 평가

제안 시스템의 활용 가능성과 사용자 수용성을 검증하기 위해 시범서비스 기반 사용성 평가를 수행하였다. 평가 대상은 각 차수별로 10명의 시각장애인으로 구성하였으며, 1차와 2차 시범서비스로 구분하여 단계적으로 실험을 진행하였다.

1차 시범서비스는 XR 글래스 기반 정보 제공 기능에 대한 사용자 이해도 검증에 중점을 두었으며, 보행 보조 및 안내 기능 체험을 통해 정보 인지 여부와 사용 편의성을 평가하였다. 2차 시범서비스는 실제 보행 환경에서 수행하여 시스템의 실효성과

안정성을 검증하였으며, 장애물 인지와 원격 도우미 서비스 기능을 중심으로 사용자 경험을 평가하였다.

그림 8과 9는 시범서비스 수행 장면과 주요 설문 결과를 나타낸다.



그림 8. 시범서비스 수행 장면

Fig. 8. Field test scene of pilot service

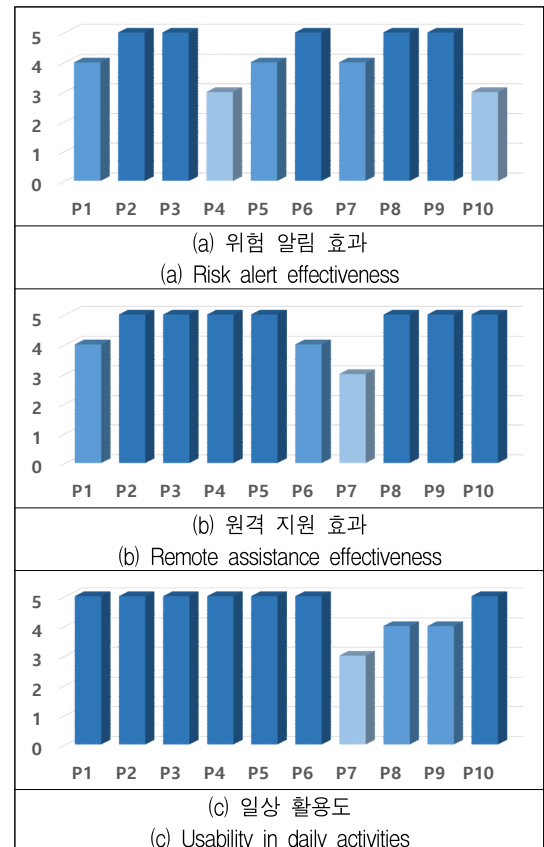


그림 9. 주요 사용성 평가 결과

Fig. 9. Key Usability Evaluation Results

설문 평가는 5점 척도 기반으로 수행하였으며, 정보 인지 지원, 보행 안전, 원격 도우미 서비스, 하드웨어 착용성 등의 항목으로 구성하였다.

평가 결과, 보행 중 위험 상황 인지를 지원하는 기능은 평균 약 4.3점으로 나타나 전반적으로 긍정적인 평가를 받았으며, 위험 상황을 사전에 인지하는 데 효과적인 것으로 나타났다. 또한 원격 도우미 서비스는 평균 약 4.6점으로 높은 만족도를 보였으며, 예외 상황 대응과 심리적 안정감 제공 측면에서 유의미한 효과를 나타냈다. 특히 음성 기반 안내 방식은 직관적인 정보 전달 수단으로 작용하여 높은 수용성을 보였다. 2차 시범서비스에서는 음성 품질(약 4.2점)과 음성 지연(약 4.0점)이 전반적으로 양호한 수준으로 평가되었다.

반면, XR 클래스의 하드웨어 요소에 대한 만족도는 상대적으로 낮게 나타났다. 기기 무게는 평균 약 3.1점, 착용감은 약 2.6점으로 평가되어 장시간 사용 시 피로감과 착용 부담이 존재하는 것으로 분석되었으며, 이는 기능적 완성도와 별개로 하드웨어 설계 개선의 필요성을 시사한다.

V. 결론 및 향후 과제

본 연구에서는 시각장애인의 보행 안정성과 이동 자율성 향상을 위해 XR 클래스 기반 보행 장애물 감지 및 실시간 원격 도우미 서비스 통합 보행 지원 시스템을 제안하였다. 제안 시스템은 사용자 시점 영상 기반 객체 탐지 및 위험도 분석, 음성 안내, WebRTC 기반 원격 도우미 서비스를 통합한 구조로 구성된다.

시스템 성능 평가 결과, 전체 처리 시간은 약 0.16초, 객체 탐지 정확도는 약 95.4%로 나타났으며, 통신 성능 또한 약 0.04초의 낮은 지연과 높은 신뢰성을 확보하여 실시간 상호작용이 가능한 수준을 확인하였다.

사용성 평가에서는 보행 안전 지원, 정보 인지 보조, 원격 도우미 서비스 전반에서 평균 4점 이상의 만족도를 나타내어 실제 보행 환경에서 활용 가능한 수준의 사용성을 확인하였다. 특히 장애물 감지 기능과 원격 도우미 서비스는 사용자 이동 안전

성 향상에 기여하는 핵심 요소로 나타났다.

다만 XR 클래스의 착용성 및 일부 네트워크 환경에서의 지연과 같은 한계를 확인하였다. 이에 따라 향후 연구에서는 하드웨어 경량화, 온디바이스 처리 성능 향상, 사용자 인터페이스 개선을 중심으로 시스템 고도화가 필요하다. 또한 다양한 보행 환경을 반영한 데이터셋 확장과 함께, 평가 대상자 규모 확대 및 다양한 사용자군을 대상으로 한 장기 실사용 기반 평가를 통해 시스템의 실효성을 정량적으로 검증할 필요가 있다. 아울러 기존 보행 보조 시스템과의 비교 실험 및 위험도 분석 모듈 적용 전후 비교를 통해 제안 시스템의 우수성을 보다 객관적으로 입증할 필요가 있다.

더 나아가 본 시스템은 보행 경로 안내 기능을 포함한 내비게이션 시스템으로 확장될 수 있으며, 다양한 장애 유형을 고려한 통합 보행 지원 서비스로 발전할 수 있을 것으로 기대된다.

References

- [1] GBD 2019 Blindness and Vision Impairment Collaborators, "Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study", *Lancet Global Health*, Vol. 9, No. 2, pp. e130-e143, Dec. 2020. [https://doi.org/10.1016/S2214-109X\(20\)30425-3](https://doi.org/10.1016/S2214-109X(20)30425-3).
- [2] P. Piyasena, V. O. Olvera-Herrera, V. F. Chan, M. Clarke, D. M. Wright, and G. MacKenzie, "Vision impairment and traffic safety outcomes in low-income and middle-income countries: a systematic review and meta-analysis", *Lancet Global Health*, Vol. 9, No. 10, pp. e1411-e1422, Oct. 2021. [https://doi.org/10.1016/S2214-109X\(21\)00303-X](https://doi.org/10.1016/S2214-109X(21)00303-X).
- [3] F. S. Ricci, L. Liguori, E. Palermo, J.-R. Rizzo, and M. Porfiri, "Navigation training for persons with visual disability through multisensory assistive technology: Mixed methods experimental study", *JMIR Rehabilitation and Assistive Technologies*,

- Vol. 11, pp. e55776, Nov. 2024. <https://doi.org/10.2196/55776>.
- [4] G. Ren, Z. Huang, W. Lin, T. Huang, G. Wang, and J. H. Lee, "Enhancing street-crossing safety for visually impaired pedestrians with haptic and visual feedback", *Applied Sciences*, Vol. 15, No. 7, Art No. 3942, Apr. 2025. <https://doi.org/10.3390/app15073942>.
- [5] S.-H. Liu and G.-H. Feng, "Wireless embedded wearable haptic assistance navigation device based on networked technology for the visually impaired", *Computers and Electrical Engineering*, Vol. 126, Art No. 110540, Aug. 2025. <https://doi.org/10.1016/j.compeleceng.2025.110540>.
- [6] R. Salameh, J. Lanir, and S. F. A. Szpiro, "Remote assistance with augmented reality: A flexible accessibility solution for low vision museum visitors", *International Journal of Human-Computer Interaction*, Vol. 42, pp. 2988-3011, Jul. 2025. <https://doi.org/10.1080/10447318.2025.2531268>.
- [7] WebRTC (Web Real-Time Communication), <https://webrtc.org/>. [accessed: Mar. 25, 2026]
- [8] Ultralytics, "YOLOv8," Ultralytics YOLO Docs, <https://docs.ultralytics.com/ko/models/yolov8>. [accessed: May 20, 2026] - yolov8
- [9] Korea Blind Union (Orientation and Mobility), http://www.hsb.or.kr/client/visually/visually3_3.asp. [accessed: Mar. 10, 2026]
- [10] Elsevier ScienceDirect, "White cane - an overview", <https://www.sciencedirect.com/topics/medicine-and-dentistry/white-cane>. [accessed: Mar. 10, 2026]
- [11] K. Shinohara and J. O. Wobbrock, "In the shadow of misperception: Assistive technology use and social interactions", *International Journal of Design*, Vol. 5, No. 3, pp. 49-64, 2011. <http://www.ijdesign.org/index.php/IJDesign/article/view/1209/559>.
- [12] M. Mims, "Guide dogs are expensive and scarce?could robots do their job?", *Communications of the ACM*: <https://cacm.acm.org/news/guide-dogs-are-expensive-and-scarce-could-robots-do-their-job/>. [accessed: Mar. 12, 2026]
- [13] S. Li, Y. Zhang, H. Wang, and X. Liu, "Lightweight vision-based obstacle detection for visually impaired people using deep learning", *Applied Sciences*, Vol. 15, No. 6, Art No. 3282, Mar. 2025. <https://doi.org/10.3390/app15063282>.
- [14] F. Schenkelberg, "Root Sum Squared Tolerance Analysis Method", *Accendo Reliability*, <https://accendoreliability.com/root-sum-squared-tolerance-analysis-method/>. [accessed: Mar. 25, 2026]
- [15] Android Developers, "MediaCodec", <https://developer.android.com/reference/android/media/MediaCodec>. [accessed: Mar. 25, 2026]
- [16] COCO - Common Objects in Context, <https://cocodataset.org/>. [accessed: Mar. 25, 2026]

저자소개

고민준 (Minjun Ko)



2021년 3월 ~ 현재 : 가천대학교
컴퓨터공학과 학사과정
관심분야 : 영상처리, 보조공학,
인공지능, 빅데이터

조진수 (Jinsoo Cho)



1998년 2월 : 콜롬비아대학교
전기공학과(공학석사)
2003년 2월 : 조지아공과대학교
전기공학과(공학박사)
2006년 4월 ~ 현재 : 가천대학교
컴퓨터공학과 교수
관심분야 : 보조공학, 스마트

헬스케어, IoT, 인공지능, 영상처리