

시각장애인 안전을 위한 영상 기반 저비용 보행 공간 인지 알고리즘[☆]

Vision-based Low-cost Walking Spatial Recognition Algorithm for the Safety of Blind People

강 성 현¹ 이 세 훈¹ 안 준 호^{1*}
Sunghyun Kang Sehun Lee Junho Ahn

요 약

현대사회에서 시각장애인들은 도보, 승강기, 횡단보도 등 일반적인 환경에서 보행을 하는데 어려움이 있다. 시각장애인의 불편 해소를 위한 연구로 영상이나 음성을 이용한 연구가 있으며, 이런 연구는 고비용의 웨어러블 장치, 고성능 CCTV, 음성 센서 등을 사용하여 실생활에 적용하는 데는 한계가 있다. 본 논문에서 시각장애인이 보행 중에 안전한 이동을 위해서 스마트폰에 포함된 저비용의 영상 센서를 활용하여 주변 도보 공간을 인지하는 인공지능 융합 알고리즘을 제안한다. 제안된 알고리즘은 이동 중인 사람 탐지를 위해서 모션 캡처 알고리즘과 장애물 탐지를 위한 객체 탐지 알고리즘을 융합하여 개발하였다. 모션 캡처 알고리즘으로 mediapipe를 사용하여 이동 중에 있는 주변 보행자들을 모델링 및 탐지하였다. 객체 탐지 알고리즘을 사용했으며 도보 중에 발생하는 다양한 장애물을 모델링 하였다. 실험을 통하여 인공지능 융합 알고리즘을 검증했으며, 정확도 0.92, 정밀도 0.91, 재현율 0.99, F1 score 0.95로 결과를 얻어서 알고리즘의 성능을 확인하였다. 본 연구로 보행 중에 발생하는 블라드, 공유 킥보드, 자동차 등의 주변 장애물 및 이동 중인 보행자 회피하여 시각장애인들의 통행에 도움을 줄 수 있다.

☞ 주제어 : 모션 캡처, 객체 탐지, 보행, 시각장애인, 실시간 영상

ABSTRACT

In modern society, blind people face difficulties in navigating common environments such as sidewalks, elevators, and crosswalks. Research has been conducted to alleviate these inconveniences for the visually impaired through the use of visual and audio aids. However, such research often encounters limitations when it comes to practical implementation due to the high cost of wearable devices, high-performance CCTV systems, and voice sensors. In this paper, we propose an artificial intelligence fusion algorithm that utilizes low-cost video sensors integrated into smartphones to help blind people safely navigate their surroundings during walking. The proposed algorithm combines motion capture and object detection algorithms to detect moving people and various obstacles encountered during walking. We employed the MediaPipe library for motion capture to model and detect surrounding pedestrians during motion. Additionally, we used object detection algorithms to model and detect various obstacles that can occur during walking on sidewalks. Through experimentation, we validated the performance of the artificial intelligence fusion algorithm, achieving accuracy of 0.92, precision of 0.91, recall of 0.99, and an F1 score of 0.95. This research can assist blind people in navigating through obstacles such as bollards, shared scooters, and vehicles encountered during walking, thereby enhancing their mobility and safety.

☞ keyword : Motion Capture, Object Detection, Walking, Blind People, Real-time Video

1. 서 론

한국장애인고용공단 고용개발원 조사통계부에 따르면 2022년 시각장애인 인구는 25.1만 명으로 등록장애인 수의 약 10%를 차지하고 있다. 연령대별 비율은 15-29세는 2.8%, 30-54세는 22%, 55세 이상은 74.6%이다[1]. 국토부의 조사에 따르면 교통약자의 보행환경 만족도는 시내버스, 시외버스, 철도 등의 교통수단 만족도보다 낮게 조사되었다. 특히 시각장애인의 교통수단 사용 비율은 다른 유형의 교통약자에 비해 버스, 기차 등 대중교통 사용이

¹ AI Computer Science and Engineering, Kyonggi University, Suwon, Gyeonggi-do, 16227, South Korea.

* Corresponding author (Junho Ahn, jha@kyonggi.ac.kr)

[Received 7 September 2023, Reviewed 16 September 2023 (R2 17 October 2023), Accepted 24 October 2023]

☆ This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)" (2021-0-01393)

낮고 자가용, 장애인 택시 사용이 높게 조사되었다[2]. 시각장애인은 대중교통 사용에 불편함을 느끼고 국토부 조사 결과 보행환경 만족도가 가장 낮음에 따라 보행에도 많은 불편함이 있다. 실제 시각장애인 11명을 대상으로 연구해 본 결과 다양한 장애물이 보행에 방해되었으며 장애물로는 불법 주정차 차량, 블라드, 일반 보행자, 자전거 등 다양했다. 실제 연구 도중 블라드에 부딪히고 짐차 블록 위 불법 주정차 차량에 보행이 막히고 속도가 빠른 자전거의 위협과 스마트폰을 보면서 걷는 일반 보행자와 부딪히는 결과가 있었다[3]. 시각장애인들의 보행 공간에서 각종 장애물과 일반 보행자를 인지하기 위해 스마트폰을 활용하려 한다. 한국지능정보사회진흥원의 통계에 따르면 2022년 시각장애인의 스마트폰 사용률은 85.5%로 전년 대비 11.7% 높아졌다[4-5]. 시각장애인의 스마트폰 사용률이 높아질 것으로 전망되는 이유로 시각장애인을 위한 스마트폰 개발과 보이스오버[6]와 같은 기술이 발전하면서 스마트폰 사용의 어려움이 점차 해소되고 있다. 시각장애인들의 스마트폰 사용률이 증가하여 소유한 스마트폰의 영상 센서로 주변 도로 공간을 인지할 시 장애물과 일반 보행자를 탐지해 보행에 도움을 줄 수 있다.

시각장애인 보행에 방해되는 장애물을 인식하는 다양한 보행 연구가 있다[7-9]. 장애물을 인식하기 위해 객체 탐지를 이용하여 장애물 모델을 만들어 인식하고, 장애물 구별도 가능하며 이를 음성 알림으로 알려준다. 스마트폰 영상 센서를 활용해 장애물을 인식하는 연구[10]와 스마트폰의 프로세서의 한계 때문에 YOLO 최적화에 있는 어려움을 줄이는 방법에 관한 연구[11]도 있다. 하지만 이 연구들 모두 단순 객체 인식만 가능해 움직임이 있는 일반 보행자는 인식에 한계가 있다. Lidar 센서를 이용해 시각장애인의 보행을 도와주는 연구[9]나 웨어러블 장치를 사용해 보행을 도와주는 연구[12]도 있다. 이 연구들은 추가 고비용 장치가 필요해 상용화 및 사용성에 어려움이 있다.

본 논문에서는 시각장애인의 안전한 보행을 위해서 일반적으로 소지한 스마트폰을 활용해 추가적인 비용 없이 스마트폰에 내장된 카메라 센서를 활용하여 주변 보행 공간 인식 융합 알고리즘을 제안한다. 블라드, 트래픽콘, 벤치, 공유 자전거, 공유 킥보드, 불법 주정차 차량과 같은 시각장애인의 보행에 방해되는 움직임이 없는 장애물과 움직임이 있는 일반 보행자를 모델로 만들었다. 시각장애인 보행에 있어 충돌 위험이 높은 것은 움직임이 있는 장애물이다. 일반 보행자를 대표적인 움직임 있는 장애물로 정해 모델을 만들었다. 일반 보행자를 인식하기

위해 mediapipe[13]를 사용해 사람의 몸에 33개의 landmark를 지정해 사람을 인식되게 만들었다. mediapipe 모델을 검증해 본 결과 보행자의 precision 값이 0.99가 나왔으며, recall 값도 0.95가 나왔다. 움직임이 있는 보행자뿐만 아니라 움직임이 없는 장애물도 같이 인식하여 시각장애인의 보행을 효과적으로 보조할 수 있게 해준다. 움직임이 없는 장애물은 YOLO[14]를 사용해 약 4000개의 이미지 데이터로 만들어진 모델을 통해 움직임이 없는 장애물을 인식, 구별 되게 만들었다. 만들어진 객체 탐지 모델을 검증해 본 결과 정한 장애물의 평균 precision 값이 0.95가 나왔으며, 평균 recall 값은 0.99가 나왔다. mediapipe만 가지고 모션 캡처 시 오답이 발생하는 경우가 존재하기에 mediapipe와 YOLO를 융합한 알고리즘 사용을 제안한다. 융합 알고리즘을 사용해 본 결과 단일 알고리즘 때보다 오답이 줄어드는 것을 확인할 수 있었다.

본 논문 2장에서는 mediapipe와 객체 탐지 기술을 이용해 시각장애인의 보행을 보조하는 연구 사례에 대해 서술하였다. 3장에서는 모션 캡처 알고리즘, 객체 인식 알고리즘, 융합 알고리즘을 비교 및 설명하였다. 4장에서는 실험 및 실험 결과에 대해 서술하였다. 마지막 5장에서는 본 연구에서 나아가 향후 연구 방향과 결론에 대해 서술하였다.

2. 관련 연구

인공지능을 이용하여 시각장애인의 보행 보조 연구가 진행 중이다[3, 7-10, 12, 23]. 미국 스탠퍼드대학교 연구팀과 미국국립과학재단이 협업하여 인공지능을 이용한 지팡이를 개발했다. 이는 자율 주행 자동차의 장애물 인식 방식을 활용하여 시각장애인 보행 시 장애물을 인식해서 회피할 수 있게 해준다[15]. 국내 시각장애인의 보행 보조 기술들의 사례로는 머신러닝을 이용하여 버스의 번호판을 인식해서 특정 버스의 도착을 알려주는 앱 ‘버스스로’[16]가 있다. 단순 장애물 인식뿐만 아니라 장애물 구별도 동시에 가능하면 보행이 안전해진다. COCO 데이터 세트와 AI-hub 데이터 세트로 Mask-RCNN 모델 학습을 진행하여 보행 중의 장애물들을 인식하는 것뿐만 아니라 장애물을 구별할 수 있는 연구가 있다[7]. 다른 객체 탐지 기술인 SSD 모델을 활용한 연구도 있다. 차량의 인도 진입을 막기 위한 장애물 블라드는 시각장애인의 보행에 있어서 어려움을 주는 장애물이다. 이를 효과적으로 인식하기 위해 SSD mobilenet v1을 이용해 장애물 인식과 구별을 동시에 하는 연구가 있다[8]. 시각장애인들의 실

내 보행을 돕는 연구도 있다. Lidar와 비주얼 오도메트리를 통해 사용자의 위치 정보를 파악한다. 파악한 위치 정보를 기반으로 사용자 전방의 실내 장애물을 딥러닝 기반 R-CNN을 통해 인식하고 이를 음성 인식으로 사용자에게 알려주는 연구가 있다[9].

mediapipe는 구글이 개발한 AI 프레임워크로 사람 자세, 손 등 다양한 비전 AI 기능이 있다. mediapipe를 통해 걷는 사람의 자세를 인식하고 이를 토대로 정형외과적 질환을 유추하는 연구가 있다. LSTM-Autoencoder는 딥러닝 기반 알고리즘으로 정상인의 보행 데이터를 학습시켜 근감소증과 같은 정형외과적 질환 때문에 이상 보행하는 사람을 구분 지을 수 있게 했다. 이 과정에서 mediapipe를 이용해 정상 보행자의 데이터와 이상 보행자의 데이터를 추출해서 병원 진료 없이 정형외과적 질환을 알아내는 연구가 있다[17]. 사고의 위험이 높은 독거 노인들의 실내 사고를 대처하기 위한 연구도 있다. mediapipe의 pose AI 프레임워크를 사용해 사람의 모습을 33개의 landmark 좌표로 전처리한다. 독거노인의 낙상사고 발생 시 크게 변동하는 몸의 좌표들을 감지한다. 정확도를 높이기 위해 좌표의 변화를 감지하고 1분 이상 독거노인들의 움직임이 없으면 이를 낙상사고로 판단하고 애플리케이션을 통해 보호자에게 알림을 주는 연구가 있다[18]. 독거노인의 낙상사고뿐만 아니라 이상행동도 감지하는 연구가 있다. mediapipe를 통해 데이터 전처리 후 GRU 학습 데이터로 사용하기 위한 추가 전처리 과정으로 슬라이딩 윈도우 방식으로 행렬로 표현한다. 그 후 LSTM보다 연산속도가 빠른 GRU로 데이터를 학습시켰다. 추가로 라즈베리파이 기반으로 음성 인식도 추가하여 독거노인의 낙상사고, 이상행동 탐지 정확도를 높이는 연구가 있다[19]. 수어의 한글 지문자를 효과적으로 인식할 수 있게 새로운 방식을 제안하여 청각장애인의 학습과 번역을 도와주는 연구도 있다. mediapipe hand API를 사용해 손의 좌표점을 추출하고 scikit-learn을 통해 한글 지문자를 학습시켜 모델을 만들었다. 손목과 손 가운데 지점과 중지와 손바닥 가운데 지점 두 지점의 landmark 추출해 두 지점 간의 각도를 통해 손의 무게중심의 이동을 계산할 수 있다. 손의 무게중심 이동을 인식하는 방식을 토대로 한글 지문자를 인식 실험을 진행했고 인식을 100%를 얻었다[20]. 수어를 배우려는 수어 학습자를 돕기 위한 학습 보조 시스템을 개발하는 연구가 있다. 수어의 지숫자 9개와 지문자 24개를 Labeling 과정으로 사전학습을 한다. 수어 학습자의 수어 동작을 mediapipe hand를 통해 손의 Landmark를 1초당 10 프레임으로 추출한다. 사전학습 시킨 수어 동작과 일치하

면 “O”, 일치하지 않으면 “X”를 보여주어 수어 학습자의 학습을 보조하는 연구가 있다[21]. mediapipe를 통해 시각장애인을 위한 보행자 인식 연구에 대해서는 아직 연구하는 바가 미흡해 이 부분을 보충하려고 한다.

YOLOv3-tiny-PRN 모델을 이용해 시각장애인의 손을 인식하여 물건을 잡은 상태와 안 잡은 상태를 구별하고, AI-hub 상품 데이터 세트로 잡은 물건을 구분해서 시각장애인이 잡은 물건의 정보를 알려주는 연구가 있다[22]. YOLO와 웨어러블 기기와 접목하여 활용성을 높인 연구도 있다. 소형 카메라가 부착된 안경을 통해 보행 중의 횡단 보도를 tiny-YOLOv2 알고리즘을 통해 인식하게 했다. 전방에 횡단보도가 나타나면 이를 인식해서 골전도 이어폰으로 시각 장애인에게 정보 전달하고 보행자 신호등의 초록 불과 빨간 불을 인식하고 구별하여 횡단보도를 보행할 때의 어려움을 줄이는 연구가 있다[12]. YOLOv5를 이용해 시각장애인의 보행에 도움을 주는 점자블록을 인식하는 연구도 있다. 점자블록을 인식하는 과정에서 점자블록과 일반 도보 블록과의 색상 차이를 이용해 점자블록을 효과적으로 인식하기 위해 이진화 처리 과정을 거침으로써 점자블록의 방향 정보를 정확히 전달할 수 있게 하는 연구가 있다[23]. 대부분의 사람들이 가지고 있는 스마트폰 카메라를 이용한 연구도 있다. 앱을 통해 실시간 보행 영상을 촬영하고 이를 서버에 전송하면 서버에서 OpenCV를 통해 전달받은 영상 데이터를 변환하고 객체 탐지 알고리즘을 사용해서 장애물을 인식하고 구별한다. 장애물을 인식, 구별하고 장애물과의 거리를 거리 당 알림 방식을 설정해 이를 서버에서 다시 앱으로 전달해 앱 사용자에게 장애물 간의 거리도 알려주는 연구가 있다[10]. 이처럼 스마트폰을 활용하면 사용자의 접근성을 높일 수 있어 시각장애인이 로봇, 웨어러블 기기와 같은 장치를 구매하지 않고도 보행 보조 서비스를 이용할 수 있다. 하지만 스마트폰의 프로세서의 성능상의 한계가 있으며 객체 탐지 알고리즘을 스마트폰과 접목해서 사용할 때의 프로세서의 한계를 극복하고자 최적화 연구하고 있다. 매개변수 저장을 위한 비트를 낮춤으로써 YOLO 모델 경량화를 하는 양자화 방식과 depth_multiple과 width_multiple 변수를 스마트폰 환경에 적합하게 설정하는 방식이 있다. android 기반인 스마트폰에서 알고리즘 구동을 도와주는 api ‘NNAPI’를 이용하는 방법도 있다. 위 3가지 방식을 소개하고 테스트함으로써 스마트폰 프로세서의 성능으로도 객체 탐지 알고리즘을 효과적으로 구동하는 방법을 알려주는 연구가 있다[11]. 객체 탐지 기술을 융합하여 사용한 연구가 있으며 YOLOv5 모델 중에서 성능

비교를 통해 YOLOv5-small 모델을 이용했다. 객체 인식에서는 Attention-OCR 모델을 이용해 Attention-OCR 모델의 인코더 과정에서 Sliding CNN을 활용해서 feature vector를 추출하는 융합 시스템을 개발해 신호등을 인식한 연구가 있다[24].

3. 제안 알고리즘

3.1 모션 캡처 알고리즘



(그림 1) mediapipe 활용한 모션 인식
(Figure 1) Motion Recognitions Using the Mediapipe

일반적으로 보행 시 많이 발견되는 것은 길거리에 있는 일반보행자이다. 그렇기에 보행 중 가장 충돌이 많으며 그만큼 시각장애인에게 위험을 줄 수 있는 빈도가 가장 높다. 시각장애인에게 가까이 접근하는 사람의 동작을 탐지하기 위해 mediapipe를 활용한다. mediapipe는 google에서 제공하는 AI 프레임워크로서, 영상 데이터에서 모델에 해당되는 객체를 비전 AI 기능을 파이프라인 형태로 인식한다. 본 논문에서는 여러 가지 mediapipe 모델 중 pose를 통해 사람의 모션 캡처를 할 것이다. mediapipe는 안드로이드, 웹, 파이썬에서 사용할 수 있어 접근성이 좋기 때문에, 본 논문에서 평소 보행 중에 저비용 영상 센서와 스마트폰을 활용 시 좋은 시너지가 기대된다. mediapipe pose 모델은 사람의 신체의 33개의 지점을 탐지한다. 실제 사람 영상으로 실험 시 모델에서 지점을 탐지하고 이를 파이프라인 형식으로 출력한다.

3.2 객체 탐지 알고리즘

장애물 객체로는 블라드, 트래픽콘, 벤치, 자전거, 키포드, 자동차를 선정했다. 블라드, 트래픽콘, 벤치의 경우는 길거리에서 쉽게 찾아볼 수 있는 고정 장애물이기에 보

행 시 방해가 된다. 자전거, 키포드는 최근 공유 자전거, 키포드 서비스가 발전하여 대한민국 대부분의 도심에서 길거리에서 방치되어 있거나 사용하는 자전거와 키포드를 손쉽게 찾아볼 수 있고 서로 충돌 시 사고가 일어날 수 있어 위험성이 높다. 자동차는 인도가 없는 골목길에서는 사각지대에서 나오는 자동차나 불법 주 정차된 자동차는 시각장애인에게 큰 위협으로 다가온다.

그림 2처럼 원하는 객체를 바운딩 박스 형태로 탐지하는 알고리즘인 YOLO를 활용한다. 객체 탐지 알고리즘을 이용해 바운딩 박스 형태 객체로 인식하기 위해서 탐지할 객체의 사진을 이용해 사각형 형태로 라벨링하여 객체 탐지 알고리즘을 통해 학습시켜 모델을 만든다.



(그림 2) 장애물 객체 인식
(Figure 2) Obstacle-Related Object Recognitions

3.3 융합 알고리즘

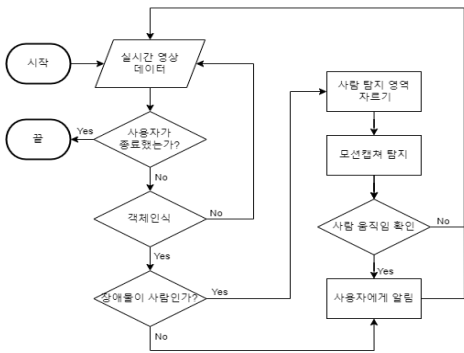
각 알고리즘의 실험 결과값은 정밀도, 민감도, 정확도 등이 높은 점수로 측정돼 신뢰도가 높지만, 실험데이터는 방해가 없는 최적 상황의 결과이기에 실사용 시에는 오차가 발생할 수 있다. 이런 오차를 줄이기 위해 두 알고리즘 융합을 제안한다.

사람 이외의 객체가 다수 포함된 영상을 mediapipe pose 단독 실행 시 오탐률이 급격하게 상승하여 그림 3 왼쪽과 같이 사람 이외의 객체에 파이프라인이 출력되고, 객체 탐지 알고리즘은 사람 탐지 시, 사람의 동작을 인식하지 못해 동적인 탐지하기 어려움이 있기에 움직임이 있는 사람에 대한 신뢰성이 떨어진다. 떨어진 신뢰성을 올리기 위해 객체 탐지 알고리즘과 mediapipe를 같이 사용하여 각 알고리즘의 단점을 보완하여 그림 3 오른쪽과 같은 신뢰성 높은 결과값 출력이 가능했다.



(그림 3) mediapipe 오탐 예시(왼쪽), 융합 알고리즘 기반 모션 인식(오른쪽)

(Figure 3) Mediapipe Misdetection Example(Left), Motion Recognition Using Fusion Algorithm Recognition(Right)



(그림 4) 융합 알고리즘 Flowchart
(Figure 4) Fusion Algorithm Flowchart

mediapipe의 오탐률을 줄이기 위해 먼저 오탐률이 낮은 객체 탐지 알고리즘을 통해 탐지 후 YOLO의 바운딩 박스를 이용하여 사람이 찍힌 부분만 mediapipe를 실행시켜 동적인 사람인가를 판단하여 오탐률을 낮춘다. 그림 4로 실행한 결과 약 3배의 오탐이 측정된 데이터가 오탐이 없어진 것이 실험을 통해 확인할 수 있었다. 결과 데이터를 통해 mediapipe와 YOLO를 융합하여 단독 사용보다 더 신뢰성 있는 탐지 알고리즘을 구현했다.

4. 실험 결과 및 분석

4.1 실험데이터 수집

실험을 진행하기 위해 대략 4,500장 사진과 150개의

영상을 직접 촬영하고 수집하여 진행했다. 실사용과 비슷한 환경 구축을 위하여 실사용 시 빠른 객체 탐색과 전력 효율을 위해 720p로 영상 촬영했다. 720p 화질에서 장애물과 사람을 탐지할 수 있는지 확인하기 위해 여러 상황의 영상을 mediapipe와 객체 탐지 알고리즘으로 실행시켜 장애물 객체와 사람의 탐지율을 실험한다.

모션 캡처 모델 실험을 위해 사람이 걸을 때를 다양한 각도에서 촬영하여 실험데이터를 수집했고, 객체 탐지 실험을 위해 본 논문에서 정한 다양한 장애물들의 객체를 다양한 각도에서 촬영하여 실험데이터로 수집했다. 통합 실험에 사용될 데이터는 보행자가 길거리에서 볼 수 있는 횡단보도, 공원 등과 위험에 노출되기 쉬운 공사장, 사각지대 등의 영상을 포함해 총 18개의 실내외 상황을 촬영해 수집했다.

실험데이터의 성능을 나타내기 위한 지표로 머신러닝 평가지표인 F1 스코어를 활용할 예정이다. F1스코어는 모델의 출력된 데이터를 활용해 True-Positives, True-Negative, False-Positives, False-Negative를 분류하여 정밀도와 민감도를 계산한다. 정밀도는 모델이 인식한 객체가 실제 객체와 일치하는지를 나타내는 지표이며, 민감도는 정밀도와 반대로 실제 객체 중 모델이 인식한 객체의 비중을 나타내는 비중이다. F1 스코어는 민감도와 정밀도의 조화 평균값을 의미한다. 정확도는 전체 데이터 중 모델이 인식한 객체의 비율을 의미한다.

4.2 모션 캡처 실험 결과

모션 캡처 탐지율 실험을 위해 사람의 걷는 동작을 다양한 각도에서 5초의 짧은 시간으로 나누어 영상을 촬영해 실험데이터를 수집했다.

(표 1) 모션 캡처 실험 결과

(Table 1) The Experimental Result for the Motion Capture Algorithm

정밀도	민감도	정확도	F1스코어
0.81	0.89	0.83	0.85

수집한 영상으로 mediapipe pose를 실행한 결과, 사람이 단독으로 촬영되고, 다른 객체가 적은 영상의 결과는 결괏값이 높게 측정되었다. 하지만 다양한 객체가 다수로 출현하는 영상의 경우에는 오탐률이 높아 표 1과 같이 정밀도, 민감도, 정확도, F1 스코어가 0.9 이하로 측정되는 아쉬운 결과물을 볼 수 있었다.

4.3 객체 탐지 실험 결과

객체 탐지 실험을 위해 YOLO 모델에 학습시켰던 다양한 객체들의 사진을 촬영과 수집하여 실험을 진행했다. 장애물 분류에 대한 실험과 장애물 인식에 대한 실험을 동시에 시행했다.

(표 2) 객체 탐지 실험 결과
(Table 2) The Experimental Result for the Obstacle Algorithm

	정밀도	민감도	정확도	F1 스코어
블라드	0.93	0.94	0.93	0.94
꼬깔	0.92	0.95	0.92	0.93
벤치	0.97	0.84	0.98	0.90
자전거	0.97	0.80	0.98	0.88
키보드	0.92	0.80	0.93	0.86
자동차	0.98	0.95	0.98	0.96
평균	0.95	0.89	0.96	0.92
장애물	0.99	0.95	0.99	0.97

표 2의 결과를 보면 정밀도, 민감도, 정확도, F1 스코어가 대부분 0.9 이상 값으로 측정되어 신뢰도 높은 모델을 보여줄 수 있었다. 객체 형태가 다양하고 시야 위치에 따라 생김새가 달라지는 객체인 벤치, 자전거, 키보드는 블라드, 트래픽콘에 비하면, 민감도가 0.8로 비교적 적은 수치로 측정되었다. 특히, 자전거와 키보드는 F1 스코어가 0.9 이하인 각 0.88, 0.86으로 아쉬운 결과를 보여줬다. 하지만, 객체를 분류하지 않고 장애물로서 탐지한 데이터는 모든 데이터가 0.95 이상의 값을 측정되어 보행 공간 인지에 신뢰도 높은 출력값을 보여줄 수 있다.

4.4 융합 알고리즘 실험 결과

움직이는 장애물과 고정된 장애물을 탐지하기 위한 알고리즘을 개발하기 위해서 융합 알고리즘을 개발했다. 융합 알고리즘의 실험 결과는 개별 알고리즘보다 더 높은 실험 결과를 얻을 수 있었다. 모션 캡처에서 발생하는 오차를 줄이고 정밀도, 민감도, 정확도, F1 스코어를 더 높일 수 있었다.

(표 3) 융합 알고리즘 실험 결과
(Table 3) The Experimental Result for the Fused Recognition Algorithm

정밀도	민감도	정확도	F1스코어
0.91	0.99	0.92	0.95

모션 캡처와 객체 탐지에서 수집된 데이터를 이용해 융합 알고리즘으로 실험을 진행한 결과로는 표 3의 값을 측정할 수 있었다. 모든 값이 0.9 이상이 나오면서 모션 캡처의 아쉬운 결함값이 융합 알고리즘을 통해 개선되었다고 판단할 수 있었다.

4.5 상황별 실험 결과

수집한 다양한 상황 영상으로 본 논문에서 학습시킨 모델로 실행시켰다.

(표 4) 통합 실험 탐지 결과
(Table 4) Integrated Experiment Detection Results

	객체 탐지	mediapipe	융합
차도인접거리	o	x	o
골목길 사거리	x	o	o
일반 횡단보도	o	o	o
x자 횡단보도	o	o	o
공원	x	o	o
실내 사각지대	o	x	o
사각지대 계단	o	o	o
직선 계단	o	o	o
실내 복도	o	x	o
일반 골목길	o	o	o
오르막길	o	o	o
내리막길	x	o	o
승강기	o	o	o
육교	o	o	o
자리 찾을때	x	o	o
주차장	o	x	o
공사장	o	o	o
지하철 출입구	o	o	o

실험을 진행한 결과 단독 알고리즘으로 실행에는 모든 객체가 탐지가 안 되는 경우가 발생했지만, 이를 보완하기 위한 융합 알고리즘 모델로 실행 시 18개의 상황에서 모든 객체를 탐지하여 높은 신뢰성을 표 4에서 확인할 수 있었다.

5. 결론 및 향후 연구과제

시각장애인들의 이동권을 보장하기 위해 모션 캡처와 객체 탐지 알고리즘을 활용한 보행 공간인지 융합 알고리즘을 개발했다. 보행 중 일반인에게는 대부분 상황은 문제가 없는 상황이지만, 시야 확보가 어려운 시각장애인은 발에 걸리는 장애물에 넘어지거나, 경사가 큰 내리막 길에서 균형을 맞추지 못해 미끄러지는 등의 위험한 보행 상황이 일어날 수 있다. 시야 확보가 어려운 경우 눈을 대신해 카메라를 통해 촬영된 실시간 영상 데이터를 통하여 장애물 객체를 탐지하여 시야 확보가 어려운 사용자에게 청각적인 신호로 피드백하여 보행에 방해되는 장애물 객체 탐지 시 선제적으로 회피할 수 있게 보장할 것을 기대한다.

본 논문은 저비용 영상 센서를 이용하여 보행 중 장애물 객체 선제 탐지를 연구하고 실험했다. 실험 결과는 대부분의 상황 속에서 탐지 성공했다. 하지만 길거리에는 학습시키지 못한 장애물 객체들도 존재한다. 이러한 아쉬운 점을 개선하기 위해 객체별 학습뿐만 아니라 물리적인 장애물을 탐지하는 알고리즘을 추가 연구할 예정이며, 학습시킨 이외의 물리적인 객체는 피드백을 통해 추가 학습시켜 더 안전한 모델을 제작 진행할 것이다.

참고문헌(Reference)

- [1] Korea Employment Agency for Persons with Disabilities(KEAD), 2023, <https://www.kead.or.kr/>
- [2] minisrty of land, infrastructure and transport, 2022, <http://www.molit.go.kr/portal.do>
- [3] JH Kim, CW Oh, YH Nam, "A Qualitative Study of the Independent Walking Experience of the People with Visual Impairments," The Korean Journal of Visual Impairment, Vol. 37, No. 1, pp. 63-81, 2021. <http://dx.doi.org/10.35154/kjvi.2021.37.1.63>
- [4] National Information Society Agency (NIA), 2021, https://www.nia.or.kr/site/nia_kor/main.do
- [5] National Information Society Agency (NIA), 2022, https://www.nia.or.kr/site/nia_kor/main.do
- [6] Apple VoiceOver, 2023, <https://www.apple.com/kr/accessibility/vision/>
- [7] JH Kim, SJ Kim, DY Lee, MG Jeon, JH Kim, SB Jeon, DM Seo, "Artificial Intelligence-based Forward of Object Detection for Visually Impaired People," Korea Computer Congress 2021 (KCC2021), Vol. 2021, No. 6, pp. 1978-1980, June 2021. <https://www.dbpia.co.kr/Journal/articleDetail?nodeId=N ODE10583501>
- [8] SB Ou, JW Lee, "Implementation of a Bollard Recognition System for Safe Walking of the Visually Impaired," Korea Computer Congress 2019 (KCC2019), Vol. 2019, No. 6, pp. 901-903, June 2019. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=N ODE08763367>
- [9] SH Jin, KD Kim, MS Kim, JY Shin, SH Chae, "Lidar and VO Based Wireless Walk Guide System for Blind People," ICT Convergence Korea 2020, Vol. 2020, No. 11, pp. 417-418, November 2020. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=N ODE10501269>
- [10] SH Cho, HJ Kim, SS Park, YJ Choi, SW Lee, "A Study on Mobile-based Obstacle Detection for Blinds," The Kips Spring Conference 2021, Vol. 28, No. 1, pp. 433-436, May 2021. <https://kiss.kstudy.com/Detail/Ar?key=3893687>
- [11] BH Ahn, JU Kim, "YOLOv5-based Optimization Study Considering Speed and Accuracy in Mobile Environment," Korea Software Congress 2022, Vol. 2022, No. 12, pp. 1604-1606, December 2022. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=N ODE11224537>
- [12] ST Oh, KD Jeong, HM Kim, YK Kim, Young-Keun Kim, "Development of Street Crossing Assistive Embedded System for the Visually-Impaired Using Machine Learning Algorithm," Journal of the HCI Society of Korea, Vol. 14, No. 2, pp. 41-47, May 2019. <https://doi.org/10.17210/jhsk.2019.05.14.2.41>
- [13] Mediapipe, August 18th 2023, <https://developers.google.com/mediapipe>
- [14] YOLO, August 18th 2023,

- <https://pjreddie.com/darknet/yolo/>
- [15] HY Kim, SY Lee, DY Kim, SJ Jo, "An Analysis on O&m Ability Status and Needs of People with Visual Impairments for the Development of AI Guide Robot," *The Korean Journal of Visual Impairment*, Vol. 39, No. 2, pp. 99-121, 2023.
<http://dx.doi.org/10.35154/kjvi.2023.39.2.99>
- [16] US Ji, KS Kim, CJ Ahn, "Mobility Service Improvement for the Visually Impaired," *Gyeonggi Research Institute*, Vol. 2022, No. 17, pp. 1-89, July 2022.
<https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE11155084>
- [17] HR Yoon, SJ Ryu, EA Jo, JI Yoo, JH Kim, "LSTM-Autoencoder-Based Anomaly Detection for Orthopedic Diseases Using MediaPipe Gait Models," *Korea Software Congress 2022*, Vol. 2022, No. 12, pp. 613-615, December 2022.
<https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE11224187>
- [18] ES Ju, HG Im, SM Lee, SI Park, CH Jeon, YS Jung, "Implementation of Fall Accident Detection System," *Proceedings of KSCI Conference 2022*, Vol. 30, No. 2, pp. 461-462, July 2022.
<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11140451>
- [19] CM Kwak, YS Kim, JB Lee, JM Jung, CJ Yoo, "Deep Learning-Based Abnormal Behavior Detection and Monitoring Service for the Senior Living Alone," *Proceedings of KIIT Conference 2022*, Vol. 2022, No. 6, pp. 461-465, June 2022.
<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11082577>
- [20] JY Kim, ES Kang, "Korean Finger Spelling Recognition Using Hand Landmarks," *The Journal of KINGComputing*, Vol. 18, No. 1, pp. 81-91, 2022.
<http://doi.org/10.23019/kingpc.18.1.202202.008>
- [21] JY Kim, H Sim, "Development of a Sign Language Learning Assistance System using Mediapipe for Sign Language Education of Deaf-Mutlity," *The Journal of The Korea Institute of Electronic Communication Sciences*, Vol. 16, No. 6, pp. 1355-1362, December 2021.
<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11317341>
- [22] DY Park, SB Lim, "Object Detection Algorithm for Explaining Produce to the Visually Impaired," *The Journal of the Korea Contents Association*, Vol. 22, No. 10, pp. 1-10, October 2022.
<https://doi.org/10.5392/JKCA.2022.22.10.001>
- [23] JK Kang, V Bajenezza, SM Ahn, MW Sung, YS Lee, "A Method to Enhance the Accuracy of Braille Block Recognition for Walking Assistance of the Visually Impaired: Use of YOLOv5 and Analysis of Vertex Coordinates," *Journal of KIISE (JOK)*, Vol. 49, No. 4, pp. 291-297, April 2022.
<https://doi.org/10.5626/JOK.2022.49.4.291>
- [24] HS Cho, MJ Lee, YJ Han, "Kubernets-based Framework for Improving Traffic Light Recognition Performance: Convergence Vision AI System based on YOLOv5 and C-RNN with Visual Attention," *Annual Conference of KIPS 2022*, Vol. 29, No. 2, pp. 851-853, November 2022.
<https://kiss.kstudy.com/Detail/Ar?key=3988580>

● 저 자 소 개 ●



강 성 현(Sunghyun Kang)

2021년~현재 경기대학 컴퓨터공학부 학부생
관심분야 : 아키텍처, 인공지능, 물리엔진.
E-mail : tsugu00@kyonggi.ac.kr



이 세 훈(Sehun Lee)

2019년~현재 경기대학 컴퓨터공학부 인공지능학과 학부생
관심분야 : 데이터베이스, 인공지능, 컴퓨터비전.
E-mail : redkye2@kyonggi.ac.kr



안 준 호(Junho Ahn)

2009년~2013년 University of Colorado, Boulder, Computer science, Ph.D.
2013년~2017년 ETRI 국가보안기술연구소
2017년~2023년 한국교통대학교 컴퓨터학부 교수
2023년~현재 경기대학교 AI컴퓨터공학부 인공지능전공 교수
관심분야 : 인공지능, 딥러닝, 비전.
E-mail : jha@kyonggi.ac.kr