

객체 탐지 모델 구현을 통한 MediaPipe와 YOLO를 이용한 작업자 쓰러짐 감지 성능 비교

손지훈, 신인용, 박진영, 권순량

동명대학교 전자공학과

son5063278@naver.com, song230248@gmail.com, jyp10311@naver.com, srkwon@tu.ac.kr

Performance Comparison of Worker Fall Detection Using MediaPipe and YOLO Through Object Detection Model Implementation

Son Ji Hun, Park Jin Yeong, Shin In Yong, Kwon Soon Ryang

Department of Electronic Engineering, Tongmyong Univ.

요약

산업 및 건설 현장에서 다양한 요인으로 인해 작업자가 쓰러질 때 능동 대처하게 되면 인명 손상 및 손실이 발생하고 기업의 생산성 저하로도 이어질 수 있어 이러한 경우에 신속히 대처할 수 있는 수단이 마련되어야 한다.

본 논문에서는 그 해결 방안으로 감시카메라와 컴퓨터 비전 기술을 활용한 작업자 쓰러짐 감지 모델을 구현한다. 사람의 골격 구조를 추정하여 쓰러짐을 감지하는 MediaPipe를 이용한 포즈 추정 기법과 이미지에서 사람의 자세와 위치를 파악하여 쓰러짐을 감지하는 YOLO를 이용한 객체 검출 기법을 구현 모델에 적용하고 두 방법에 대해 비교 분석한다.

I. 서론

산업 및 건설 현장에서 작업자의 쓰러짐 사고는 예상치 못하게 발생할 수 있다^[1]. 이는 치명적인 인명 피해로 이어질 수 있고 기업의 생산성 저하를 초래할 수 있어 능동적으로 대처하기 위한 방법이 요구된다.

본 논문에서는 쓰러짐 감지 시스템의 실용적 구현에 필요한 기술을 채택하기 위해 인체의 주요 관절 위치를 추정하여 자세 변화를 분석하는 MediaPipe 기반 포즈 추정 기법과 이미지에서 사람의 전체적인 자세와 위치를 실시간으로 검출하는 YOLO 기반 객체 검출 기법을 비교 분석하고, 효과적인 시스템 구현을 위한 최적의 접근 방식을 제시하고자 한다.

연구 방법으로는 각 기술을 이용한 쓰러짐 감지 알고리즘을 구현하고, 정확성, FPS(Frames Per Second), 효율성 등의 성능 비교를 진행한다.

본 논문에서는 두 기법의 이론적 배경, 구현 방법, 실험 및 비교 분석 결과를 간략히 제시하고, 쓰러짐 감지 시스템의 실용적 구현을 위한 제언을 도출한다.

II. 본론

1. 객체 탐지 기법

1.1 MediaPipe

그림 1은 MediaPipe 포즈 추정의 골격좌표를 나타낸 것이다.

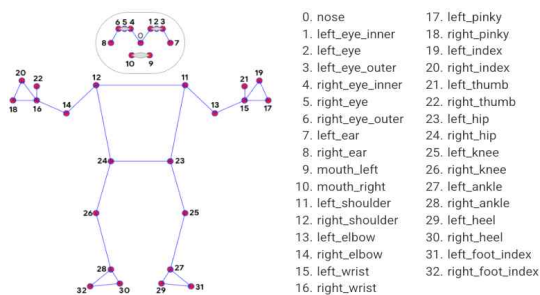


그림 1. MediaPipe의 골격좌표

MediaPipe는 구글에서 개발한 오픈소스 프레임워크로, 인체의 주요 관절 위치를 실시간으로 추정하는 포즈 추정 기술을 제공한다^[2]. MediaPipe 기반 포즈 추정 프로세스는 그림 2와 같은 단계로 이루어진다.

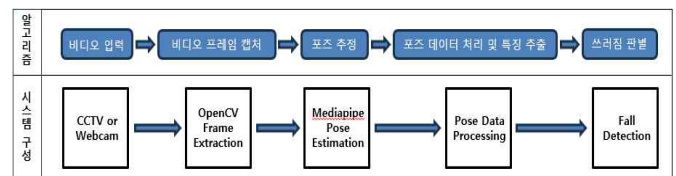


그림 2. MediaPipe 기반 포즈 추정 프로세스

MediaPipe 기반 포즈 추정 프로세스는 그림 2에서 보는 바와 같이 촬영된 실시간 영상을 대상으로 전처리 과정을 진행하여 사람의 영역을 추출하고, 변환된 이미지 데이터를 사용하여 사람의 포즈 좌표(어깨, 팔, 무릎 등)의 위치를 예측한다. 이 좌표를 기반으로 인체 구조를 형성한 후 추정된 관절 정보를 기반으로 자세를 분석하여 특정 동작(쓰러짐)을 판별한다.

1.2. YOLO

YOLO(You Only Look Once)는 실시간 객체 검출을 위한 딥러닝 알고리즘으로, 이미지나 비디오 프레임에서 다양한 객체를 빠르고 정확하게 탐지할 수 있는 특징을 가지고 있다^[3].



a) Labeling b) Bounding Box

그림 3. YOLO 기반 객체 탐지 과정

그림 3은 YOLO 기반 객체 탐지 과정을 나타낸 것으로 YOLO가 다양한 이미지에서 여러 객체를 탐지하고, 각 객체에 대해 바운딩 박스를 설정한 후 이를 라벨링하여 해당 클래스('사람', '쓰러짐')를 예측하는 과정을 거친다. 이 과정은 그림 4와 같은 단계로 이루어진다.

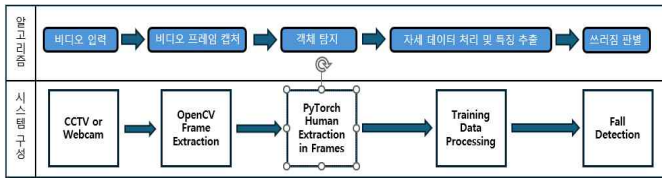


그림 4. YOLO 기반 객체 검출 프로세스

YOLO 기반 객체 검출 프로세스는 그림 4에서 보는 바와 같이 실시간 영상을 입력받아 전처리 과정을 진행한다. 전처리된 영상 데이터를 YOLO 모델에 입력하여 미리 학습된 가중치를 사용하여 영상 내의 객체를 감지하고 분류하며 이 과정에서 영상내 사람의 위치, 경계 상자, 사람의 자세 등을 분류한다. 이후 YOLO가 감지한 객체들 중 '넘어짐' 클래스로 분류된 객체를 식별한다.

2. 객체 탐지 모델 구현

2.1. MediaPipe 기반 포즈 추정

본 논문에서 MediaPipe를 이용한 포즈 추정 기법을 적용한 객체 탐지 모델을 구현하였다.

하드웨어 구성으로 프로세서는 라즈베리파이4를 사용하였고, 카메라 모듈로는 Pi 카메라(V2)를 사용하였다. 소프트웨어 환경은 Raspberry Pi OS에서 Python 3.8과 MediaPipe라이브러리 0.8.9.1 버전을 사용하였다.

구현 방법으로는 Python 스크립트에 MediaPipe 라이브러리를 import하여 포즈 추정 파이브라인을 구성하였다. 실시간 영상 처리를 위해 OpenCV 라이브러리를 함께 사용하였으며 MediaPipe의 Pose 솔루션을 활용하여 33개의 신체 랜드마크를 추출하였다.

그리고 넘어짐 감지를 위해 추출된 랜드마크를 이용하여 상체와 하체의 중심값을 구해 발 사이의 거리를 비교하여 특정 임계값과 비교하여 넘어짐을 분석하는 알고리즘을 구현하였다.

2.2. YOLO 기반 객체 검출

본 논문에서 YOLO를 이용한 객체 검출 기법을 적용한 객체 탐지 모델을 구현하였다.

하드웨어 구성으로 Mini PC(N100), PC 카메라, 그리고 WiFi Dongle을 사용하였다. 소프트웨어 환경은 Windows 11 운영체제에, Python 3.8과 YOLOv5, PyTorch2.8.1 버전을 사용하였다.

구현 방법으로는 미리 학습된 YOLO 모델을 로드하여 실시간 영상에서 사람 객체를 검출하도록 설정하였다. 각 프레임에서 감지된 사람 객체의 경계 상자 정보를 이용하여 넘어짐 상태를 판단하는 알고리즘을 구현하였다.

3. 실험 및 결과 비교 분석

3.1. MediaPipe와 YOLO의 정확도 성능 비교

[표 - 1]은 MediaPipe와 YOLO를 실행하여 정확도를 분석한 결과를 나타낸 것이다.

탐지할 객체의 상태가 정상일 때에는 두 기법 모두 높은 검출 성능을 보여주었다. 그러나 탐지할 객체가 불안정할 상태일 때에는 정확도가

MediaPipe는 40%, YOLO는 63%가 나와 정확도 측면에서는 YOLO가 우수한 성능을 보여주었다.

표-1. 정확도 비교 테이블

N = 50 (EA)	Mediapipe (EA)		Yolov5 (EA)	
Detection	Fall	Success	Fall	Success
Standing (10EA)	1	9	0	10
Falling (40EA)	24	16	15	25

3.2. MediaPipe와 YOLO의 FPS 비교

MediaPipe와 YOLO의 FPS 성능에서는 MediaPipe의 경우 단일 객체에 대해서만 자세 검출이 가능하여 실시간 영상에서 30~120 FPS로 높은 성능을 보여주었고, YOLO의 경우는 여러 개의 객체를 동시 추적하면서 실시간 영상에서 5~60 FPS로 비교적 낮게 나오는 모습을 보여주었다. 결과적으로 단일 객체에 대한 추적에서는 MediaPipe가 더욱 높은 성능을 보여주었다.

3.3. MediaPipe와 YOLO의 효율성

MediaPipe는 포즈 추정 기능을 제공하여 작업자의 자세와 각 관절의 위치를 실시간으로 추적한다. 이를 통해 자세의 변화를 감지하고, 쓰러짐과 같은 급격한 자세 변화를 정밀하게 인식할 수 있다. 고도로 최적화된 파이프라인을 통해 높은 FPS를 유지하므로, 복잡한 환경에서 작업자의 자세를 지속적으로 추적하는 데 적합하다.

YOLO는 빠른 객체 검출을 통해 작업자의 쓰러짐을 신속하게 인식할 수 있다. 사전에 학습된 모델을 사용하여 시스템을 구축할 수 있으며, 특정 영역 내 여러 객체를 동시에 추적할 수 있는 이점이 있다. 이러한 특성으로 인해 YOLO는 다수의 작업자나 객체를 동시에 모니터링해야 하는 환경에 적합하다.

III. 결론

본 논문에서는 MediaPipe와 YOLO를 이용한 쓰러짐 감지 기법을 적용한 객체 탐지 모델 구현을 통해 두 기법의 성능을 비교 분석하였다.

연구 결과, 각 기법은 서로 다른 환경에서 강점을 보이는 것으로 나타났다. MediaPipe는 단일 작업자의 세밀한 자세 분석이 필요한 환경, 복잡한 동작이나 자세 변화를 지속적으로 추적해야 하는 상황에서 효과적이다. 예를 들어, 개인 운동 모니터링이나 재활 치료 현장에서 유용하게 활용될 수 있다. 반면 YOLO는 여러 작업자나 객체를 동시에 모니터링해야 하는 환경, 넓은 공간에서 빠른 객체 검출이 필요한 상황에 적합하다. 공장 작업장이나 공공장소 안전 모니터링과 같은 환경에서 YOLO의 장점이 두드러진다.

향후 연구방향으로 두 기법의 장점을 결합한 하이브리드 시스템 개발을 제안한다. MediaPipe의 정확한 자세 추정 능력과 YOLO의 빠른 객체 검출 능력을 통합함으로써, 다양한 환경에서 더욱 효과적인 쓰러짐 감지가 가능할 것으로 예상된다.

참 고 문 헌

- [1] <https://www.youtube.com/watch?v=s5hOa69TuUg&t=1s> : 40도 넘는 곳에서 작업하다...외주 노동자 쓰러져 숨져 / SBS뉴스, 2020.6.11.
- [2] <https://ai.google.dev/edge/mediapipe/solutions/guide?hl=ko>
- [3] <https://pjreddie.com/darknet/yolo/>