

YOLO를 이용한 공사 현장 작업자의 보호장구 착용여부 및 위험지역 진입여부 검출

김현용¹, 박태형²

¹충북대학교 산업인공지능연구센터, ²충북대학교 지능로봇공학과

Detection of Wearing a Protective Gear and Entering Hazardous Areas of Worker in Construction Site using YOLO

Hyun-Yong Kim¹ and Tae-Hyoung Park²

¹Industrial AI Research Center, Chungbuk National University

²Department of Intelligent Robot Engineering, Chungbuk National University

Abstract - 공사장에서 허술한 안전관리로 인해 안전모나 마스크 등의 안전장비를 착용하지 않아 발생하는 사망사고가 지속적으로 발생하고 있다. 본 연구에서는 YOLOv5를 이용하여 공사장 CCTV 영상을 분석하여 작업자의 안전모와 마스크 착용을 실시간 검사할 수 있는 모델을 제시한다. 기존 연구가 안전모와 마스크의 객체검출에 초점을 둔 반면, 본 연구에서는 안전모와 마스크의 검출뿐만 아니라 그것들이 각각 머리와 얼굴에 올바르게 착용되었는지까지 판단하였다. 또한 작업자가 위험지역에 진입했는지 여부도 실시간으로 탐지하여 안전사고를 방지하도록 하였다. 이를 위해 안전모와 마스크는 Kaggle Datasets을 이용하고, 인체(사람, 머리, 얼굴) 데이터는 Open Image Dataset을 활용하였다. 안전모 검출 위치와 마스크 검출 위치 및 착용여부 판정을 위한 인체 검출 위치를 종합한 안전모 및 마스크 착용여부 검사 정확도는 95%, 속도는 40fps로 실시간 검사가 가능한 것으로 나타났다.

1. 서 론

2020년 산업재해 사고사망자 중 절반 이상이 건설업 사고사망자일 정도로 건설 현장은 고위험 산업으로 작업자의 안전관리가 중요하다. 그러나 건설 현장의 70% 정도가 안전모 미착용 등 개인 보호구도 제대로 갖추지 않는 등 안전관리가 부실한 실정이다. 따라서 공사 현장의 작업자 안전관리를 강화하기 위해서는 기존의 단일한 수동방식을 대체할 효율적인 시스템의 도입이 절실하다.

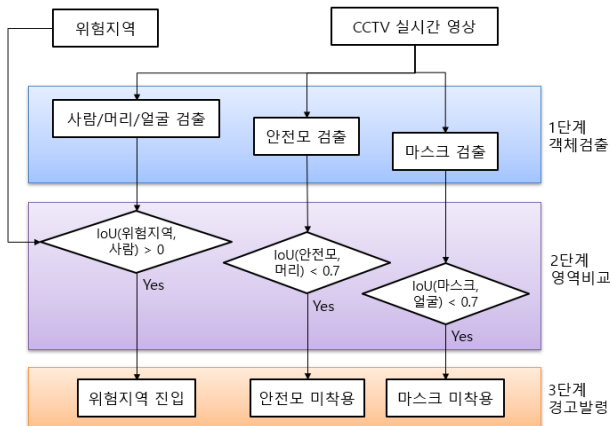
떨어짐, 부딪힘, 물체에 맞음, 낙상 등의 경우 머리의 상해는 매우 심각하고 종종 치명적이기 때문에 작업장 내에서는 반드시 안전모를 착용해야 한다. 마스크는 공사 현장의 분진으로부터 작업자를 보호할 뿐만 아니라, 최근 COVID-19 팬데믹에 따른 확산을 방지하기 위해 의무적으로 착용해야 한다. 또한 공사 현장에서는 낙하물 발생 위험구간이나 건설장비 작업중 충돌 발생 위험구간 등에 작업자가 출입하지 않도록 감시할 필요가 있다.

본 연구에서는 작업자의 허술한 안전관리를 개선하기 위해 CCTV 영상을 실시간으로 분석하여 공사 현장 작업자의 안전에 가장 중요한 보호장구인 안전모와 마스크가 올바르게 착용되었는지 검출하고, 작업자가 지정한 위험구역에 출입하지 않도록 실시간으로 감시하는 모델을 개발하고자 한다.

2. 본 론

2.1 시스템 구성

CCTV에서 전송된 실시간 스트리밍 속에서 프레임을 추출하여 객체 검출을 수행하는 시스템 구성을 그림 1에 나타내었다. 먼저 1단계에서는 객체를 검출하고, 2단계에서 검출된 객체의 좌표를 이용해 IoU(Intersection over Union)를 계산하여 임계값에 따라 3단계 경고를 발령하게 된다. 1단계에서 3번의 객체검출을 수행한 것은 기존의 공개된 데이터셋을 활용하기 위함이다.



〈그림 1〉 시스템 구성도

그림 1에서 IoU(사람, 위험지역)은 검출된 사람 객체의 Bounding box와 위험지역에 대한 IoU값을 말하며, IoU는 그림 2와 같이 정의된다.

$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}}$$

〈그림 2〉 IoU의 정의

2.2 객체 검출 알고리즘

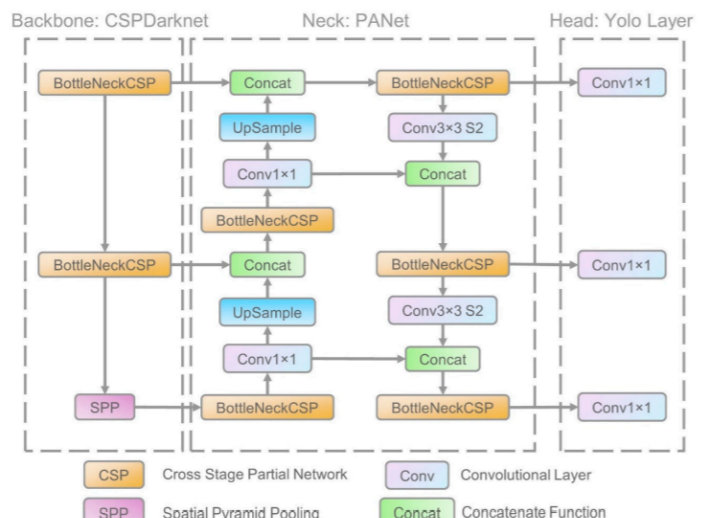
딥러닝 기술의 비약적인 발전으로 인해 사회 각 분야에서 영상처리를 활용하는 사례가 늘어나고 있다. 특히 저렴한 CCTV의 보급과 고속의 네트워크 인프라 구축으로 인해 지능형 CCTV를 통해 자동차 번호판 인식, 화재 감지, 침입탐지 등에 객체검출(Object detection) 기술이 활발히 연구되고 있다.

객체검출에는 R-CNN, Fast R-CNN, Faster R-CNN 계열처럼 영역 제안(Region proposals) 단계와 이렇게 도출된 ROI(Region of Interest)에 대한 CNN 분류 단계로 구성된 Two-stage 기법과 YOLO 계열처럼 One-stage 기법이 있다. 정확도 면에서는 R-CNN 계열이 우수하지만 YOLO 계열은 탁월한 처리속도의 강점이 있다. YOLOv2는 K-Means Clustering을 통한 Anchor box 선정 기법을 도입하였고, YOLOv3에서는 FPN(Feature Pyramid Network)를 적용하여 작은 객체의 검출 정확도를 향상시키는 등 최근의 YOLOv5[1]까지 지속적인 성능향상으로 속도와 정확도 측면에서 많은 개선이 있었다.

CCTV를 통한 객체검출은 실시간 처리를 통해 영상을 분석하기 때문에 정확도 못지않게 신속한 처리가 필요하다. 따라서 본 연구에서는 속도와 정확도 면에서 우수한 YOLOv5를 이용하였다.

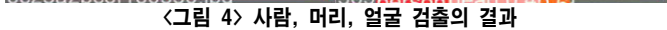
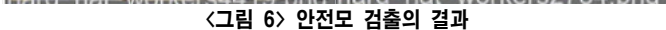
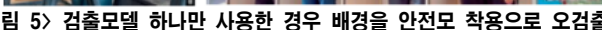
2.2.1 YOLOv5 네트워크 아키텍처

YOLOv5의 네트워크 아키텍처는 그림 3과 같이 이미지의 특징을 추출하는 Backbone과 추출된 특징을 혼합 및 결합시키는 Neck, 그리고 Neck의 특징들로부터 클래스와 Bounding box 좌표를 예측하는 Head로 구성되어 있다. Neck은 CSP(Cross Stage Partial Network)를 사용하여 그래디언트 연산속도를 향상시켰고, Neck에서는 PANet(Path Aggregation Network)을 통해 먼저 추출되는 low-level 특징이 Head까지 잘 전달될 수 있도록 하여 보다 섬세한 검출이 가능하게 하였다.



〈그림 3〉 YOLOv5 네트워크 아키텍처[2]

작은 데이터로 성능을 높이기 위해 COCO 데이터셋으로 사전 훈련된 (pre-trained) YOLOv5의 'large' 모델로 전이학습을 수행하였다. 100 epoch의 학습시간으로 mAP@0.5=0.84의 정확도를 보였다. YOLOv5 모델의 경우 32fps의 추론속도를 보였다. CCTV용 범용 컴퓨터를 이용한 추론시간을 예측하기 위해 Intel I7, RAM 24GB, GPU는 NVIDIA GTX-1650Ti 4GB 컴퓨터를 사용하였다. 추론 결과의 예를 그림 4에 나타내었다.

[illegible]

1. $\frac{1}{2} = \frac{1}{2} \hat{=} \frac{1}{2} \hat{=} \frac{1}{2}$



이 책 기 여 기 이 타 기 이 -

그러나 작업자 1처럼 건설장비에 탑승하여 지면과 접촉하지 않는 경우에는 여전히 오탐지하는 문제가 발생하므로 이에 대한 추가적인 연구가 필요하다.

[illegible][illegible]

1000

.....