

YOLOv5를 이용한 안전모 검출 Detection Of Safety Helmets Using YOLOv5

송 석 진, 꺾 내 정*

배재대학교, 배재대학교*

Song seokjin, Kwak neajoung*

Beajea Univ., Beajea Univ.*

요약

본 논문은 근로자가 산업현장에서 개인보호장비 착용 여부를 자동으로 감지하는 방법을 제안한다. YOLOv5는 영상 속 작업자 보호장비 중 안전모 착용을 실시간으로 감지한다. 실험은 15887개의 학습용 데이터와 4641개의 검증용 데이터를 훈련에 사용한다. 실험 결과 Epochs91을 기준으로 mAP 96.2% 성능이 45FPS 속도로 나타났다.

I. 서론

우리나라의 산업안전 의무를 준수하지 않아 사고와 부상이 많다. 특히 건설업은 2020년 6월 현재 전체 산업재해 분포의 54%를 차지하고 있다. 건설현장과 중장비가 자주 사용되는 산업현장은 근로자들이 각종 위험에 노출돼 안전이 항상 중요한 곳이다. 개인보호장비는 추락사고나 낙하물에 의한 부상 위험 등 사람을 보호하는 데 큰 역할을 한다. 특히 안전모는 사람의 신체에서 중요한 부분인 머리를 위험으로부터 보호해준다. 따라서 본 논문은 객체 감지 딥 러닝 알고리즘인 YOLOv5[1]를 사용하여 안전 헬멧에 대한 자동 감지 알고리즘을 제안한다. 딥 러닝 알고리즘을 활용해 작업자의 보호장비를 실시간으로 인식함으로써 관리자가 모니터링에 필요한 시간과 인력을 줄일 수 있다.

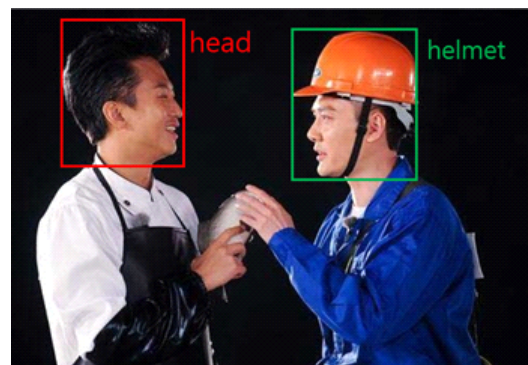
II. YOLOv5

객체 탐지 모델 중 YOLOv5 모델을 사용하여 안전모를 감지하여 착용 여부를 검사한다. YOLO는 이미지를 동일한 크기의 그리드로 나누고 위치와 클래스를 한 번에 결정하는데, 이는 Two-Stage 검출기에 비해 속도가 빠른 것이 특징이다[2]. Two-Stage 기법 보다는 정확도가 낮다는 단점에도 불구하고 YOLOv5는 빠르고 현저하게 향상된 정확도를 보여준다[1]. YOLOv5는 모델의 레이어(model depth multiple) 및 레이어 당 채널 수(layer channel multiple)에 따라 S, M, L, X의 네 가지 모델이 있다. 모델의 크기가 커질수록 파라미터는 깊어지고 mAP는 증가하지만 속도는 감소한다. 본 논문에서는 산업 현장에서 실시간 처리가 가능해야 한다는 점을 고려하여 YOLOv5 모델 4개 중 가장 빠른 S 모델을 사용하여 안전모를 감지한다.

III. 실험 환경

1. 데이터 셋

본 논문은 Kaggle의 "YOLO helmet/head"[3] 데이터 셋을 사용했다. 데이터 셋은 15887개의 훈련용 데이터, 4641개의 검증용 데이터 및 2261개의 테스트용 데이터로 구성되어 있다. Class는 Head와 Helmet이라는 라벨이 붙어 있으며, Ground Truth는 8만 4천개의 Head와 4만 1천개의 Helmet을 갖는다. 본 논문에서는 안전모가 없는 모든 머리를 Head라 표기하고, 안전모를 착용한 근로자는 Helmet이라 표기한다. Ground Truth 영역은 안전모만 있을 때에 객체 검출을 피하기 위해 모자를 포함한 머리 부분을 경계 상자로 설정했다.



▶▶ 그림 1. Ground Truth 예시

2. 개발환경

개발환경은 Ubuntu 18.04.5 LTS의 Jetson AGX Xavier로 수행되었다.

표 1. Jetson AGX Xavier 사양

OS	Ubuntu 18.04.5 LTS
GPU	512-Core Volta GPU with Tensor Cores
CPU	8-Core ARM v8.2 64-Bit CPU, 8 MB L2 + 4 MB L3
RAM	32 GB 256-Bit LPDDR4x 137 GB/s
CUDA	10.2
cuDNN	8.0

IV. 실험 및 결과

실험은 2261개의 테스트 세트로 이루어졌고, 객체 감지 알고리즘은 빠른 YOLOv5s를 사용했다. 입력이미지의 사이즈는 640x640이고 배치 사이즈를 다르게 하여 차이를 비교하고 Learning rate를 다르게 설정하여 수행되었다. 표 2는 배치 크기에 따른 실험 결과를 보여주고, 표 3은 Learning rate(LR)에 따른 차이를 보여준다.

표 2. Batch-size에 따른 차이

Model	Batch-size	Epochs	mAP	FPS	Runtime
M1	16	50	0.961	45	9.0h
M2	64	50	0.961	45	8.2h

표 3. Learning rate에 따른 차이

Model	LR	Batch-size	Epochs	mAP	FPS
M3	0.01	64	100	0.961	45
M4	0.001	64	600	0.965	45

모델 M1과 M2는 배치 사이즈를 다르게 하여 실험한 결과이며 mAP와 FPS(Frame Per Second)에는 영향이 없었고, 배치 사이즈가 클 때 모델의 훈련시간이 줄어드는 것을 확인할 수 있다. 모델 M3과 M4는 학습률에 변화를 주었을 때, FPS에는 변화가 없었지만 학습률이 낮은 M4 모델은 모델 수렴에 대한 에폭(epoch)이 더 오래걸렸고 M3 모델보다 좋은 mAP를 보였다.



▶▶ 그림 2. 안전모 검출 결과

V. 결론

본 논문에서는 안전모 착용 모델을 YOLOv5s 모델로

구현하여 배치사이즈와 학습률에 변화를 주어 비교해보았다. 그 결과 배치 크기가 클수록 학습 시간이 단축된다는 것을 알 수 있었다. 따라서 GPU가 모델 학습을 허용하는 한 큰 배치 크기를 사용하는 것이 좋다. 또 학습률이 낮을 때 mAP가 높아지는 결과를 얻었다. 향후에는 속도와 성능을 절충한 결과를 위해 처리속도와 입력 이미지의 크기에 의한 연구를 진행할 것이다.

■ 참고 문헌 ■

- [1] YOLOv5 Github. <https://github.com/ultralytics/yolov5>
- [2] Redmon, J. Divvala, S. Girshick, R., & Farhadi, A., "You only look once: Unified, read-time object detection," IEEE CVPR Las Vegas, NV, pp. 779-788, 2016
- [3] Kaggle YOLO safety helmet/head dataset. <https://kaggle.com/vodan37/yolo-helmethead>
- [4] Sangyoon, Park, Sanghyun, Yoon, Joon, Heo., "Image-Based Automatic Detection of Construction Safety helmets Using R-FCN and Transfer Learning," KOREAN SOCIETY OF CIVIL ENGINEERS, Vol. 39, No. 3, pp. 399-407. 2019