

# YOLO 기반 이륜차와 헬멧 객체 검출 방법에 관한 연구

박세연, 장혜령  
동국대학교

seyeon0627@gmail.com, hyeryung.jang@dgu.ac.kr

## A Study on the YOLO-based Two-wheeled Vehicle and Helmet Detection

Seyeon Park, Hyeryung Jang  
Dongguk Univ.

### 요약

전동킥보드, 자전거, 오토바이를 포함한 이륜차의 이용률 증가에 따라 관련 사고 또한 증가하는 추세이며, 사고 위험을 줄이기 위해 헬멧 착용이 필수화 되었으나 정작 이를 지키는 운전자는 적다. 헬멧 미착용 주행자 단속은 주로 경찰이 도로에 나가 직접 판별하는 수동 검출을 사용하고 있으나 이는 정확도와 인력 및 시간비용 측면에서 효율적이지 못하다. 본 논문에서는 딥러닝을 이용해 주행 관련 이미지로부터 이륜차의 종류 및 헬멧 착용 여부를 판별하는 과정을 자동화하여 앞서 언급한 문제점을 해결하는 YOLO 기반 이륜차와 헬멧 객체 검출 방법을 제안하고자 한다. 이를 위해 다양한 이륜차와 주행자의 이미지 데이터를 수집하고, 실시간 객체 검출에서 주로 사용되는 YOLO 기반 모델을 적용하여 학습하였으며, 데이터 증강을 적용하여 학습 성능이 향상됨을 확인하였다.

### 1. 서론

최근 이륜차의 이용률이 증가하면서 매년 2 만 건 이상의 관련 사고가 발생하고 있다[1]. 이륜차는 운전자에 대한 보호장치가 없어 사고 시 부상의 위험이 다른 이동수단보다 크기 때문에 헬멧을 필수로 착용해 위험 요소를 보호할 필요가 있다. 도로교통법상 자전거 및 개인형 이동장치를 운전할 때는 헬멧을 의무 착용해야 하지만, 실상 이를 지키는 운전자는 절반 정도에 그친다[2].

기존의 경우 경찰이 거리에 나가 헬멧 미착용 운전자를 단속하는 수동 검출 방법을 사용했다. 그러나 이 방법은 투자한 인력과 시간 대비 효과가 미미하고 골목 등은 단속이 어려운 것과 같이 범위가 제한적이기 때문에 한계점을 가진다. 관련 선행 연구[3]에서는 YOLOv5 를 이용해 오토바이와 헬멧 착용 여부를 검출하는 방법을 제안하였으나 오토바이에만 국한된 데이터로 검증하였으며 특히 헬멧 미착용 클래스에 대하여는 성능이 저조하였다.

본 논문에서는 이륜차 종류 및 헬멧 검출을 자동화해 인력과 시간을 절약하고 더 넓은 범위의 검출이 가능한 딥러닝 기반의 검출 방법을 제안하고자 한다. 이를 위해 이륜차의 대표적인 종류인 오토바이, 자전거, 전동킥보드의 다양한 이미지 데이터에 대해 빠르고 정확한 학습과 예측을 할 수 있는 YOLO 알고리즘을 이용해 학습하고 mAP 성능지표를 이용해 사용 가능성을 검증하였으며 더 나아가 데이터 증강을 활용하여 성능 향상을 확인하였다.

### 2. 본론

#### 2.1 YOLO

YOLO 는 객체의 종류와 위치를 한 번에 파악하는 1-stage 알고리즘을 사용하기 때문에 R-CNN 등 2-stage 기반의 알고리즘보다 검출 속도가 빠르다[4]. 본 논문에서는 영상 속 헬멧과 이동 수단을 빨리 검출해야 하므로 실시간 검출에 특화된 YOLO 를 주 학습

알고리즘으로 채택하였다. YOLO 알고리즘은 다음과 같다. 먼저 이미지가 입력되면  $S \times S$  개의 그리드 셀로 이미지를 나누어준 후 2 가지를 동시에 진행한다. 첫째로 바운딩박스와 confidence 를 계산한다. 바운딩박스의 정보로 박스의 중심  $x, y$  좌표와 박스의 가로, 세로 길이인  $w, h$  를 구한다. confidence 는 박스에 객체가 존재하는지에 대한 확률값인  $\Pr(\text{Object})$ 와 박스 검출의 정확도를 나타내는 지표인  $\text{IoU}(\text{pred}, \text{true})$ 값의 곱으로 계산한다. 둘째로 각 그리드 셀에 대한 분류를 동시에 진행한다. 각 셀에서  $C$  개의 클래스에 대해  $\Pr(\text{class} | \text{object})$  조건부확률을 예측한다. 그중 가장 높은 확률값을 가진 클래스를 해당 셀의 클래스로 선택한다. 마지막으로 여러 바운딩 박스 중 예측력이 좋은 하나를 남기는 NMS 기법을 사용해 최종 결과를 나타낸다.

#### 2.2 학습 데이터 및 데이터 증강

이륜차와 헬멧 객체 검출을 위한 학습을 위해서는 다양한 이륜차와 주행자의 이미지 데이터가 필요하다. 이륜차로는 도로에서 흔치 볼 수 있는 오토바이, 자전거, 전동킥보드를 선택했으며, 헬멧으로는 이륜차 주행 시 주로 사용하는 풀 페이스 헬멧, 하프 페이스 헬멧 등의 다양한 색깔과 모양의 헬멧 이미지를 선택하였다. 우선 Kaggle 사이트를 통해 헬멧, 자전거, 오토바이 클래스가 라벨링된 이미지 데이터 609 장을 수집했다[5]. 전동킥보드의 경우 구글 이미지 크롤링 및 유튜브 영상 캡처를 통해 헬멧과 함께 나온 이미지 171 장을 수집한 후, `labelImg` 도구를 이용해 직접 객체에 대한 라벨을 생성하였다. 총 880 장에 대한 학습 결과 헬멧 착용과 미착용 클래스의 데이터 개수 불균형으로 해당 클래스에 대한 구별 정확도가 낮다는 문제점을 파악했다. 이를 해결하기 위해 Github 의 헬멧 미착용 이미지 약 50 장을 추가 수집하고[6], 구글 이미지 크롤링으로 헬멧 이미지를 수집하고 라벨링 과정을 거쳐 총 1,016 장의 이미지를 획득하였다.

본 연구에서는 모델의 성능을 향상시키기 위해 데이터 증강 기법을 사용하였다. 이미지의 각도와 위치를 고려해 전체 좌우 반전, 바운딩 박스 좌우 반전, 회전(-30° to +30°) 기법을 적용하였다. 나아가 날씨와 시간대에 따른 영상 및 이미지 밝기를 고려해 -20° to +20°만큼 밝기 변화를 주었다. 마지막으로 모자이크 기법을 사용해 4 개의 다른 이미지를 하나로 합치는 효과를 랜덤하게 적용해 총 2,642 개의 이미지 데이터셋을 생성하였다. 라벨 클래스는 {With Helmet, Without Helmet, Motorcycle, Bicycle, Scooter} 총 5 개로, 전체 라벨의 개수는 29,394 개가 생성되었으며, 이 데이터셋에 대해 학습과 검증 데이터셋을 9:1 비율로 나누어 학습하였다.

### 2.3 학습 결과

다양한 크기의 여러 객체를 검출하기에 성능과 속도가 우수하고, 목적에 맞게 환경 설정이 가능한 Pytorch 기반의 YOLOv5 를 사용해 헬멧과 이륜차의 객체 검출 모델을 학습하였다. YOLO 네트워크는 크게 백본(backbone), 넥(neck), 헤드(head)로 나뉜다. 백본은 합성곱신경망으로 이미지의 특징을 추출하며, YOLOv5 에서는 CSPNet 을 사용한다. 넥에서는 백본에서 추출한 특징을 결합하고, 마지막 헤드 부분에서는 넥에서의 특징을 이용해 객체에 대한 박스를 생성하고 클래스를 예측한다. 메인 모델로 사용할 YOLOv5m 에서는 depth\_multiple 은 0.67 로, width\_multiple 은 0.75 로 설정하고 3 개의 아웃풋 레이어를 생성해 총 369 개의 레이어와 약 2 천만 개의 파라미터를 사용하였다. 데이터셋에 맞춰 image size 640×640, batch size 16, epochs 300 으로 파라미터를 설정하고, 클래스 개수를 5 개로 바꾸어 학습을 수행하였다.

Model	Precision	Recall	mAP
YOLOv5m	0.94	0.95	0.98

표 1. YOLOv5m 학습 결과

성능 지표로는 precision, recall, mAP 를 사용한다. Precision 이란 정밀도로, 모델이 positive 로 예측한 것 중 실제 positive 인 비율로 계산한다. Recall 은 재현율로 실제 positive 값 중 positive 로 예측한 비율을 말한다. 이 둘을 종합해 평가한 것이 평균 예측 정밀도, 즉 mAP 이다. AP 란 precision 과 recall 을 각각 x, y 축으로 그린 그래프의 안쪽 면적을 의미하며 이의 평균이 mAP 이다. 본 논문에서는 IoU 값이 0.5 일 때의 mAP(mAP@0.5, 이하 mAP)를 성능 지표로 사용하였다. 실험 결과, 표 1 과 같이 클래스 전체에 대한 precision 은 94%, recall 은 95%, mAP 98%로 우수한 성능을 확인하였다. 또한 검증데이터 이미지에 대한 객체 검출 결과 예시는 그림 1 과 같다. 다양한 크기와 클래스에 해당하는 객체가 높은 confidence 로 검출됨을 알 수 있다.



그림 1. 객체 검출 결과 예시

표 2 는 객체 검출에 대표적으로 사용되는 타 알고리즘과의 mAP 비교를 나타내며, 데이터 증강으로 인한 성능 향상을 보여준다. YOLOv5 에서는 모델의 크기에 따라 5s, 5m, 5l 의 모델을 사용하였고, YOLOv7 도 적용하여 파라미터 수가 많아짐에 따라 검출 성능이 향상됨을 확인하였다. YOLOv5m 의 FPS 또한 74 로 속도와 성능 측면에서 타 모델에 비해 우수한 결과를 보여준다. 또한, 데이터 증강 전 후에 따른 학습 성능을 비교함으로써, 모든 모델에 대해 데이터 증강이 효과적임을 확인하였다.

Model	파라미터 수	mAP	
		증강 전	증강 후
YOLOv5s	7.0M	0.89	0.96
YOLOv5m	20.8M	0.89	0.98
YOLOv5l	46.1M	0.90	0.98
YOLOv7	37.2M	0.87	0.99
MobileNetSSD	6.8M	0.60	0.89
Faster-RCNN	13.3M	0.49	0.62

표 2. 타 모델과의 성능 비교 및 데이터 증강 전후 비교

### 3. 결론

본 논문에서는 이륜차 주행 중 헬멧의 착용 여부 탐지 시스템을 구축하기 위해 딥러닝을 이용한 효율적인 객체 검출 방법을 제시하였다. 실시간 영상 처리에 적합한 YOLOv5 의 파라미터를 수정해 이륜차와 헬멧을 검출하는 모델을 만들어 실험하였고, 플립, 모자이크 등 다양한 이미지 증강 기법을 적용해 검출 정확도를 높이는 데 성공하였다. 그 결과 YOLOv5m 를 이용한 모델의 경우 mAP 98%로 큰 폭 향상된 결과를 얻을 수 있었으며 실제 시스템에 활용하기에 적합할 것으로 보인다. 향후 이를 CCTV 영상에 검출 기법을 실제 적용함으로써 이륜차와 헬멧 수의 차이를 계산하여 이륜차 주행 안전 시스템을 구축하고자 한다.

### ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터육성지원사업 (IITP-2022-2020-0-01789) 및 한국연구재단의 지원을 받아 수행되었음 (No.2021R1F1A1063288).

### 참고 문헌

- [1] TAAS 교통사고분석시스템, 교통사고 추세 (<http://taas.koroad.or.kr/>)
- [2] 한국소비자원, “전동킥보드 공유서비스 안전실태조사,” 2021, pp. 19-20.
- [3] 김현호, 박태형, “YOLO 를 활용한 오토바이 교통법규 위반 감지 시스템 개발,” 대한전기학회 학술대회 논문집, 2021, pp. 348-349.
- [4] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.
- [5] Kaggle, Helmet Detection (<https://www.kaggle.com/datasets/andrewmvd/helmet-detection>)
- [6] Github, Safety-Helmet-Wearing-Dataset (<https://github.com/njvisionpower/Safety-Helmet-Wearing-Dataset>)