

# 딥러닝을 이용한 마스크 착용자의 안전모 착용 감지 연구

이재성, 구윤경, 김선형, 최유라, 권영우

경북대학교 컴퓨터학부

lee01042000@gmail.com, hellokyk20@gmail.com, kimsh9510@gmail.com, yura0812@knu.ac.kr, ywkwon@knu.ac.kr

## A Study on Detecting a Helmet Wearing a Mask via Deep Learning

Jaeseong Lee, Yoonkyung Gu, Seonhyeong Kim, Yoola Choi, Young-Woo Kwon

School of Computer Science and Engineering

Kyungpook National Univ.

### 요약

개인형 이동 장치의 보급으로 편리한 이동성이 제공되고 있지만 안전모 미착용으로 인한 안전사고가 빈번히 발생하고 있다. 이에 본 논문에서는 딥러닝을 사용하여 안전모 착용을 감지할 수 있는 모델을 제안한다. 정확성과 추론 시간을 고려해서 MobileNetV2-SSD를 활용한 안전모 객체 탐지 모델을 개발하고 공개된 안전모 착용 데이터를 활용하여 여러 객체 탐지 모델에서 훈련을 진행하였다. 그리고 COVID-19의 영향으로 마스크 착용이 의무화됨에 따라 마스크 착용시 안전모 착용을 감지하기 위하여 마스크를 착용한 데이터셋을 추가하여 학습을 하였으며 모바일 환경에서 82.4%의 성능을 보였다.

### I. 서론

개인형 이동 장치(이하 PM)가 도심 곳곳에 등장하면서 공유형 전동 킥보드 시장 규모가 커짐에 따라 관련 교통사고도 해마다 증가하는 추세이다. 이로 인해 대구시의회에서는 '대구광역시 개인형 이동장치 이용안전 증진 조례'를 제정했다[1]. 하지만 조례가 제정됨에도 불구하고 보호장구 착용이 제대로 이루어지지 않고 있다.

본 논문에서는 조례 중 주요 내용인 '안전사고 예방을 위해 안전모 등 보호 장구 착용[1]'을 확인하기 위한 안전모 착용 탐지 모델을 소개한다. 공유형 전동 킥보드 사용자는 스마트폰 앱을 실행하고 스마트폰에 내장된 카메라를 이용해 안전모 착용 여부를 확인받는다. 안전모를 착용하면 전동 킥보드를 이용할 수 있지만 착용하지 않은 경우는 킥보드를 이용할 수 없게 된다. 안전모 착용 감지를 위하여 딥러닝 기반의 이미지 분류 방식을 사용하고 모바일 환경에서도 동작할 수 있는 경량화 모델을 이용하므로 충분한 FPS를 보장할 수 있고 적은 용량으로도 모델을 구동할 수 있다.

공개된 데이터로 모델을 학습시킨 결과 COVID-19(코로나바이러스감염증-19)로 인해 마스크 착용이 의무화되면서 마스크 착용으로 인해 안전모 착용 감지 성능이 하락했다. 따라서 본 논문에서는 기존의 안전모 착용 이미지에 마스크를 합성한 데이터셋과 1인칭 시점(POV)의 마스크 및 안전모 착용 이미지 데이터셋을 수집해서 학습시켜 모델의 성능을 약 82.4%로 향상시켰다.

### II. 관련 연구

다음과 같은 영상 처리 기반 안전모 탐지에 관한 연구들이 진행되었다. 공사장, 작업장에서의 근로자의 안전사고를 예방하기 위한 안전 보호 장비 착용 여부를 판단하는 연구[2]가 있다. 약 3,000장의 다양한 건설 현장에서 공사장 안전모를 착용한 데이터셋을 구성해 사용했다. 본 논문의 안전모 착용 감지 대상이 공유형 전동 킥보드 사용자임을 고려하면 해당 데이터셋을 그대로 사용하는 것은 한계가 있다. 따라서 본 논문에서는 자전

거용 안전모를 착용한 이미지 데이터셋을 활용한다.

영상 인식 알고리즘을 이용한 안전모 탐지에 관한 연구[3]가 있다. 해당 연구에서는 객체 탐지 알고리즘으로 YOLO 시리즈 알고리즘(YOLOv2, YOLOv3 등)을 제시했다. 이 모델들은 스마트폰의 하드웨어 사양으로는 구동하기 적절치 않다. 따라서 본 논문에서는 모바일 환경에서 구동하기 적합한 여러 모델로 실험을 진행하기로 한다.

### III. 본론

#### 3.1 데이터셋

본 논문에서는 우선 Kaggle에 공개되어있는 Bikes Helmet Dataset을 사용하였다. 해당 데이터셋에는 764장의 이미지에 With Helmet, Without Helmet 두 가지 클래스에 해당하는 정보를 포함하고 있다. 그림 1은 마스크 이미지를 추가로 합성한 데이터이다.



그림 1. 안전모 착용 원본과 마스크 합성 이미지

앞서 언급한 데이터셋의 경우 마스크를 착용한 경우가 거의 없어 모델 실행 시 편향이 발생할 것으로 예상되어 현재 코로나 상황에서 실용성 저하가 우려되었다. 따라서 마스크 이미지들을 각 데이터셋 이미지에 합성하여 데이터를 생성하였다. 형태와 색상이 다

양한 마스크 이미지를 사용하여 마스크로 인한 추가적인 편향이 발생하는 것을 방지하였다. 또한, 안전모 감지 모델 사용 시 전면 카메라를 사용해 촬영할 경우가 많을 것으로 예상되어 1인칭 시점(POV) 이미지를 추가로 확보하였다. 그림2는 1인칭 시점에 마스크를 합성하고 추가로 안전모까지 합성한 데이터이다.



그림 2. 마스크와 안전모를 합성한 1인칭 이미지

합성한 안전모는 여러 키보드, 자전거용 안전모 이미지를 사용하였으며 형태가 유사한 경우는 대조적인 색상으로 변경시킨 이미지를 사용하여 추가적인 편향이 발생하는 것을 방지하였다. 또한 각 이미지에는 이전 데이터셋과 동일하게 안전모 착용 여부 정보를 추가하였다.

### 3.2 실험 모델

본 실험에서는 적합한 모델을 찾기 위해 3가지의 객체 탐지 모델을 사용하여 실험을 진행하였다. MobileNetV2[4]에 기반한 Single Shot Detector(SSD)[5] 모델의 경우 컴퓨팅 자원이 한정적인 모바일 기기에서 효과적으로 작동하는 장점이 있어 모바일 환경에서 실용적인 모델을 구현하고자 하는 본 논문의 목표와 일치하기에 먼저 선택하였다. 하지만 타 모델보다 정확도가 떨어진다는 단점이 있고 이를 비교하기 위해서 RFCN-ResNet 101[6] 모델과 R-CNN[7]모델을 사용하였다. 각 모델은 Tensorflow Framework를 사용하였으며 배치 크기 등의 하이퍼파라미터는 동일한 환경에서 훈련을 진행하였다. 기본 Epoch은 200으로 설정하고 Early Stopping에서 Patient를 20으로 설정하여 각 훈련 시에 가장 좋은 모델을 선택하고자 하였다.

## IV. 모델 훈련 결과

본 논문에서는 안전모 착용 여부를 판단하기 위해 정확도와 정밀도를 기준으로 평가하였다. 또한, 각 기준에서 전체 이미지 대비 목표 객체의 범위를 기준으로 small, medium 그리고 large로 나누어 각각 평가를 진행하였다. 그리고 모바일 환경에서 각 이미지의 평균 추론 시간을 측정하였다. 정확도와 정밀도는 With Helmet 클래스에 대한 객체 탐지 모델의 추론결과가 50%이상인 경우를 기준으로 측정하였다. 추론 시간의 경우 Tensorflow로 훈련된 모델을 Tensorflow Lite로 변환한 후 Galaxy S10 5G 모델을 이용하여 측정하였다. 실험 결과로 RFCN-ResNet101 모델이 가장 높은 결과를 도출했으며 Faster R-CNN InceptionV2 와 MobileNetV2-SSD가 그 뒤를 이었다. MobileNetV2-SSD모델이 타 모델보다 비교적 정확성은 낮지만, 추론 시간이 상당히 짧아 모바일 환경에서 사용하는데 용이할 것으로 보인다. 표1은 여러 모델의 테스트 결과이다.

표 1. 각 모델에 대한 성능 측정 결과

Model	Average Precision			Average Recall			Inference Time
	small	medium	Large	small	medium	Large	
MobileNetV2-SSD	0.021	0.483	0.764	0.08	0.514	0.824	82ms
RFCN-ResNet101	0.081	0.523	0.874	0.182	0.624	0.927	435ms
FR-CNN InceptionV2	0.072	0.511	0.839	0.13	0.597	0.896	320ms

## V. 결론

본 논문에서는 안전모를 반드시 착용하고 주행해야 하는 키포드 관련 법률로 인한 안전모 착용을 감지해야 할 요구를 충족하기 위하여 딥러닝을 이용한 모바일용 객체 탐지 모델을 제작하였다. 기존 안전모 데이터를 활용하고 코로나 시기를 반영하여 마스크를 합성하여 데이터 생성을 진행하였다. 또한 여러 객체 탐지 모델을 사용하여 정확성과 추론 시간을 비교 및 평가하였다. MobileNetV2-SSD 모델의 경우 타 모델보다 정확성은 조금 낮지만 추론 시간이 매우 짧아 최종적으로 이를 사용하게 되었다. 향후 안전모 착용 판단 모델을 공유형 전동 키포드와 연동한 서비스를 개발할 예정이다.

## ACKNOWLEDGMENT

본 논문은 교육부 및 한국연구재단의 BK21 플러스 사업(경북대학교 컴퓨터학부 Smart Life 실현을 위한 SW 인력양성사업단, 21A20131600005)의 지원으로 수행된 연구임.

## 참 고 문 헌

- [1] 대구광역시 개인형 이동장치 이용안전 증진 조례 제 5517호.
- [2] Jixiu, W., Nian, C., Wenjie, C., Huiheng, W., and Guotian, W., "Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset," *Automation in Construction*, Vol. 106, No. 102894, 2019.
- [3] 노천명, 김기관, 이수봉, 강동훈, 이재철, "영상 인식 알고리즘을 이용한 안전 보호구(안전모) 탐지에 관한 연구," *한국CDE학회 논문집*, Vol. 25, No. 4, pp. 350-357, 2020.
- [4] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L. C., "Mobilenetv2: Inverted residuals and linear bottlenecks," *IEEE conference on computer vision and pattern recognition*, pp. 4510-4520, 2018.
- [5] Liu, W. et al. "Ssd: Single shot multibox detector," *European conference on computer vision*, pp. 21-37, 2016.
- [6] Huang, J. et al. "Speed/accuracy trade-offs for modern convolutional object detectors," *IEEE conference on computer vision and pattern recognition*, pp. 7310-7311, 2017.
- [7] Ren, S., He, K., Girshick, R., and Sun, J., "Faster r-cnn: Towards real-time object detection with region proposal networks," *arXiv preprint arXiv*, 2015.