

## 2-input 딥러닝 기반 Multi-task 전동 킥보드 안전 주행 시스템

염다민\*, 윤희준\*, 이소연\*, 이가현†

이화여자대학교 전자전기공학부

e-mail : *damin321@ewhain.net*, *hjyoon9808@gmail.com*, *whatevudo@ewhain.net*,  
*kh.lee@ewha.ac.kr*†

2-input deep learning based multi-tasking  
safe driving system for e-scooter

Da-Min Yeom\*, Hee-Jun Yoon\*, So-Yeon Lee\*, Ka-Hyun Lee†  
The Department of Electronic and Electrical Engineering,  
Ewha Womans University

### I. 서론

#### Abstract

The electric scooter-sharing service has experienced rapid growth with the limelight of personal mobility devices. However, there is a growing concern in society about increasing cases of e-scooter accidents. To address this problem, we propose a deep learning-based electric scooter safe driving system that considers both user images and road images. In our framework, we use CNN-based models to detect whether the user is using a helmet or keeping eyes forward using a user image. At the same time, we check whether the driver is riding on the road and then estimate the distance between the user and the closest pedestrian within a road image. Finally, we determine the final velocity of the e-scooter using the result value of each model. Through a fast and accurate deep learning model, we improve efficiency for real-time inference and achieve high confidence in the system by considering multiple dangers simultaneously.

2018년 최초 도입된 전동킥보드 공유서비스는 개인형 이동장치의 각광과 함께 빠르게 성장하였다. 2021년 한국 소비재원이 발표한 <전동킥보드 공유서비스 안전실태조사>에 의하면 국내 전동킥보드 이용자는 20년 기준 3배 이상 급증하였으나, 교통사고로 인한 사망·부상자 또한 전년 대비 2배 이상 증가하였다.[1] 이러한 안전사고 예방을 위해 도로교통법 등 관련 규정이 존재하지만, 2021년 교통문화지수 실태조사를 통해 이용자 대다수가 관련 법규를 알고 있지만 준수하지 않음을 확인할 수 있다.[2]

기존 선행 연구는 헬멧 착용 여부, 혹은 전방주시여부 확인과 같이 단일 태스크에 집중되어 있었다.[3][4] 또한, 모델 학습 시 사용하는 이미지 데이터셋이 전동킥보드를 탑승하는 사용자들에 대한 거리나 촬영 각도가 일정하지 않아 실제 거리에 설치되는 카메라의 위치에 따라 모델의 추론 결과가 크게 영향을 받게 되므로 실생활에 도입하기 어렵다는 한계점이 존재한다. [5]

\* 공동 1저자

† 교신 저자

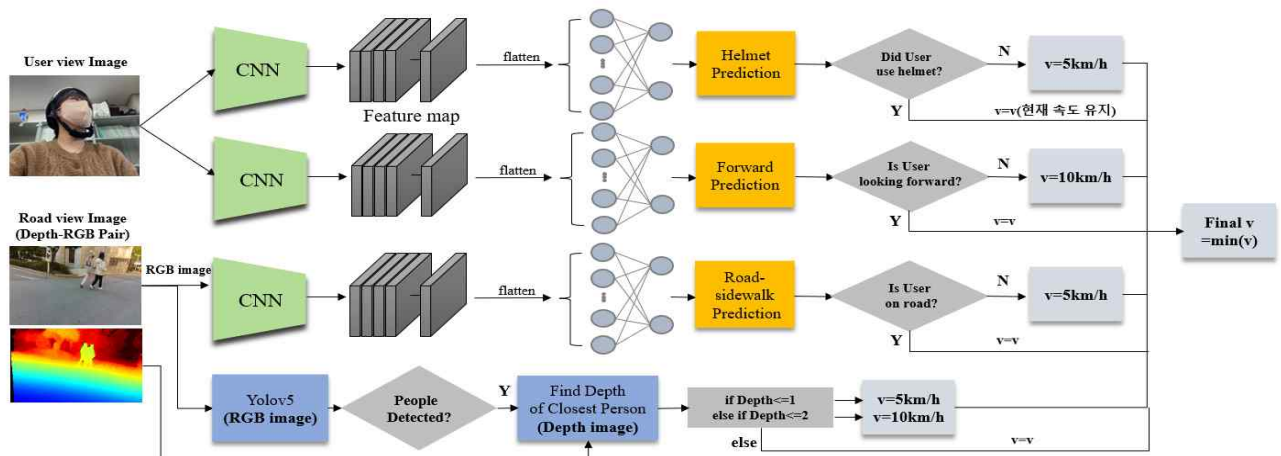


그림 1. 시스템 전체 흐름도

하지만, 안전사고 발생 양상이 헬멧 미착용, 전방 주시 미비, 보도 주행으로 인한 충돌 등 다양한 요소들로 이루어져 있는 만큼 다양한 요소를 고려하여 시스템을 개발할 필요가 있다. 이러한 점에 착안하여 본 논문에서는 전동킥보드 이용자 측의 이미지와 도로 측 이미지를 모두 활용한 종합적 안전 시스템을 제안하고자 한다.

## II. 제안 시스템

본 논문에서 제안하는 시스템은 그림 1 과 같이 전동킥보드 주행 중 발생할 수 있는 여러 상황을 고려하고자 킥보드 사용자 측과 도로 측 총 2 가지 입력을 모두 활용하여 최종 결과를 도출한다.

사용자 측 카메라의 RGB 영상으로 헬멧 착용여부 및 전방 주시여부를 확인하는 이진분류(binary classification)가 진행된다. 이와 동시에 도로 측 카메라의 RGB 이미지를 받아 보도 주행여부 확인을 위한 이진분류 및 보행자를 감지하는 물체 인식이 함께 이루어진다. 만일 보행자가 감지되었다면, 감지된 객체의 좌표를 바탕으로 depth 값을 활용해 거리를 측정하여 보행자와의 충돌 위험을 감지한다.

예측한 결과에 따라 불법으로 지정된 헬멧 미착용 혹은 보도 주행 상태라면 5km/h, 전방 미주시의 경우 10km/h 로 킥보드 주행 가능 속도를 제한한다. 또한, 보행자와 킥보드 사용자 사이의 거리가 1m 이하라면 5km/h, 2m 이하는 10km/h 로 속도를 제한하였다.

최종적으로 4 가지 기능에 대한 결과 중 최솟값으로 킥보드를 감속할 수 있게 한다. 현행법상 공유킥보드의 최고속도는 25km/h 이나 21 년 12 월 제도 개선안에 20km/h 미만으로 낮추는 내용이 언급되어 본 프로젝트에서는 킥보드의 최고 속도를 20km/h 로 지정하였다.

## III. 데이터셋 구축 및 학습 모델



그림 1. 사용자 측 학습 데이터셋 획득 장소

그림 2. 사용자 측 학습 데이터셋 예

### 3.1 데이터셋 구축

사용자 측 학습 데이터셋의 경우, 실제 전동킥보드 주행 환경을 반영한 데이터셋을 얻기 위해 그림 2와 같이 인도 및 골목길이 많고 유동 인구가 많은 서초, 교대, 강남역에서 촬영을 진행하였다. 총 2명의 인물이 그림 3 과 같이 헬멧 착용, 전방 주시 두 가지 라벨에 대해 총 4가지 조합으로 동영상으로 촬영한 후 6fps로 이미지를 추출하여 약 15000장의 학습 데이터를 수집하였다. 또한 모델이 외양의 특정 부분만 학습하는 것을 방지하고자 촬영 중간 안경을 쓰거나 머리를 묶는 등의 변화를 주었다. 테스트 데이터셋은 다양한 환경과 인물에 대한 모델 성능 확인을 위해 학습 데이터 수집 환경과 다른 실내외 공간에서 촬영하고, 학습 데이터셋에 포함되지 않은 인물의 사진을 포함해 총 100장가량 수집하였다.

도로 측 학습데이터셋은 AI 허브에서 제공된 인도 보행 영상 공공 데이터셋[6]과 42dot 에서 제공하는 SDLane 데이터셋[7]을 이용해 직접 보도와 차도 여부를 라벨링하여 약 2000 장의 데이터셋을 사용하였다. 테스트 데이터셋의 경우, 다양한 환경에서 인도와 차도를 촬영하여 약 100 장가량 수집하였다.

추가적으로, 학습 시 모델의 과적합을 막고 일반화 성능을 향상시키기 위해 사용자 측, 도로 측 이미지의 학습 데이터셋에 대해 무작위로 수평 대칭 이미지를 생성시키는 Random Horizontal Flip 데이터 증강 기법을 사용하였다.

### 3.2 학습 모델

모델 선정을 위해 CNN 기반 모델 중 ResNet18, ResNet34, ResNet50, GoogLeNet, MobileNetV2 총 다섯가지 기본 모델을 이용하여 실험을 진행하였다. 해당 모델을 모두 ImageNet 데이터셋에 대해 사전 학습된 파라미터를 불러와 전이 학습을 진행했다.

사용자 측의 경우 헬멧 착용, 전방 주시 총 두가지 클래스에 대한 정답 값을 예측해야 하기에 각 경우를 따로 판단하는 이진분류(2-class), 전체 네 가지 경우를 한꺼번에 판단하는 다중분류(4-class), 백본(CNN&pooling)은 공통으로 사용하여 동일한 feature 를 공유하되 서로 다른 라벨을 판단하는 multi-feature classification[8] 방식 총 세가지로 실험을 진행하였다.

## IV. 실험 결과 및 시스템 구현

### 4.1 사용자 측 학습 모델 테스트 결과

	2 Class		4 Class		Multi-feature	
	Acc(%)	Param	Acc(%)	Param	Acc(%)	Param
ResNet18	<b>97.5</b>	11,177,538	89.6	11,178,564	65.9	11,178,564
ResNet34	<b>97.1</b>	21,285,698	88.3	21,286,724	72.5	21,286,724
ResNet50	<b>95.3</b>	23,512,130	90.6	23,516,228	78.3	23,516,228
GoogLeNet	<b>87.9</b>	5,601,954	86.0	5,604,004	69.0	5,604,004
MobileNetv2	<b>90.0</b>	<b>2,226,434</b>	81.4	2,228,996	71.5	2,228,996

표 1. 사용자 측 학습 모델 테스트 결과 정확도 및 파라미터 수

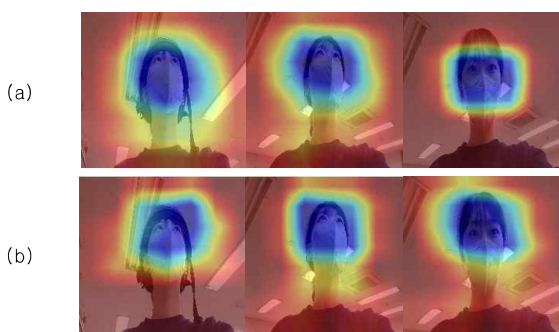


그림 3. 2-class classification 방식으로 학습된 MobileNetV2의 테스트 이미지의 Grad-CAM. 순서대로 (a) head pose classification (b) helmet classification 모델에 대한 결과

헬멧 착용 및 전방 주시 분류에 대한 최적 모델을 찾기 위해 표 1 과 같이 총 5 가지 기본 모델과 3 가지 학습 모델 기법(총 15 종류)을 활용하여 학습 및 테스트를 진행하였다. 추가로 테스트 데이터셋에 대한 정확도 뿐만 아니라 실시간성을 확인하기 위해 모델의 총 파라미터의 수를 함께 비교하였다. 이때, 최종

정확도는 모든 라벨에 대한 정확도의 평균을 이용하여 산출한 결과이다.

multi-feature 를 이용한 방식이 가장 정확도가 낮았으며 2-class 를 이용한 방식이 전반적으로 가장 높은 정확도를 보였다. 2-class 방식 중에서도 ResNet18 모델을 이용할 때 가장 성능이 좋았으나 최종적으로 실시간 처리를 위한 연산량 및 하드웨어 스펙에 따른 모델의 유동성을 고려해 보았을 때, MobileNetV2 를 사용한 2-class 방식이 가장 적합하다고 판단되어 해당 모델을 최종 실험에 사용하였다. 또한, 그림 4 와 같이 해당 모델의 Grad-CAM 결과를 확인해 보았을 때 헬멧 착용여부 모델에선 헬멧을, 전방주시여부 확인 모델에선 얼굴 전체를 판별 근거 영역으로 이용한 것을 통해 전이 학습이 잘 진행되었음을 알 수 있다.

### 4.2 도로 측 학습 모델 테스트 결과

	sidewalk-road	
	Acc(%)	Param #
ResNet18	92.7	11,177,538
ResNet34	92.7	21,285,698
ResNet50	96.4	23,512,130
GoogLeNet	96.4	5,601,954
MobileNetv2	<b>98.8</b>	<b>2,226,434</b>

표 2. 도로 측 학습 모델 테스트 결과 정확도 및 파라미터 수

최적의 인도·차도 분류 모델을 찾기 위해 총 5 가지의 기본 모델을 사용하여 학습을 진행하였다. 표 2. 의 결과를 바탕으로 파라미터 수 대비 성능이 가장 좋은 MobileNetV2 모델을 선정하여 본 논문의 실험에 사용하였다.

물체 인식 모델의 경우 기존 COCO128 데이터셋으로 사전 학습된 YOLOV5 모델을 사용하여 빠르고 정확한 객체 인식을 진행하였다.



그림 4. 하드웨어 및 LCD 표시 예

### 4.3 하드웨어

하드웨어는 그림 5 와 같이 Nvidia Jetson Nano 에 사전 학습된 모델을 임베딩하고, intel Realsense 카메라 및 CSI 카메라 모듈을 연결하여 실 상황에서 테스트가 가능하도록 구현하였다. 도로 측의 경우

Realsense 카메라를 사용하여 RGB 와 Depth 값을 받아왔고 사용자 측의 경우 CSI 카메라 모듈을 통해 RGB 인풋을 입력 받았다. 실제 추론(Inference) 과정에서는 지정된 샘플링 비율로 두 인풋을 동시에 받아와 젯슨 나노 하드웨어에서 인도·차도 구분, 헬멧 착용 및 전방주시 여부, 사람 인식 및 거리 측정 기능을 한 번에 수행하여 최종 위험도를 판단하였다. 감속이 필요한 상황이라 판단되면 킥보드의 속도를 낮추는 것이 해당 프로젝트의 최종 목표이지만 본 논문에서는 SW 에 더 집중해 I2C LCD 모듈을 활용하여 각 위험 상황에 경고 메시지를 띄우는 것으로 대체하였다.

## V. 결론 및 향후 계획

본 논문에서는 공유킥보드 안전 사고 예방을 위해 도로 측, 사용자 측 이미지를 모두 이용한 사고 예방 시스템을 제안한다. 사용자 측 이미지를 통해 헬멧 착용 여부, 전방 주시 여부를 확인하고 도로 측 이미지를 이용하여 보도 주행여부와 전방 충돌 위험성이 높은 보행자 여부를 동시에 확인한다. 빠르고 정확한 딥러닝 모델을 통해 추론 시 효율성을 높였고, 여러 상황을 동시다발적으로 고려함으로써 효과적으로 관련 안전 사고를 예방할 것으로 기대한다. 이러한 시스템은 공유 전동킥보드 뿐만 아니라 자전거, 오토바이 등 모든 퍼스널 모빌리티로 적용 가능해 전반적 안전사고 방지 시스템에 이용될 수 있으며, 공유 전동킥보드 시장 활성화 및 이용자 증가에도 큰 기여를 할 것으로 기대한다. 향후에는 교차로에서 사고 발생이 잦다는 점을 고려하여 CNN 기반 모델을 이용해 교차로가 포함된 도로 측 이미지를 학습시킨 후, 주행 중 교차로가 인식될 시 사전에 사용자에게 경고하는 시스템을 만들고 이에 대한 정확도를 향상시키는 연구를 수행하고자 한다.

## 참고문헌

- [1] 김현중, "전동킥보드 공유서비스 안전실태조사", 한국소비자원 안전감시국. 2021년 10월 11일 수정, 2022년 10월 20일 접속, <https://www.kca.go.kr/smartconsumer/sub.do?menukey=7301&mode=view&no=1003197854&page=5>.
- [2] 한국교통안전공단 교통조사평가처, "PM관련 법규 인지도 및 주행실태 조사 결과", 한국교통안전공단. 2021년 12월 2일 수정, 2022년 10월 19일 접속.

- [https://www.kotsa.or.kr/portal/bbs/report\\_view.do?menuCode=05010200&bbsCode=report&bbsSeqn=17862](https://www.kotsa.or.kr/portal/bbs/report_view.do?menuCode=05010200&bbsCode=report&bbsSeqn=17862)
- [3] 강성인, 장길수, 이수석 외 3명. 전동 킥보드 안전 보조 장치 및 그 제어 방법. 특허 출원번호 1020190023573, 출원일 2019년 2월 28일, 등록일 2021년 1월 6일.
- [4] Zhou, Fangbo, Huailin Zhao, and Zhen Nie. "Safety helmet detection based on YOLOv5." 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA). IEEE, 2021.
- [5] Gilroy, Shane, et al. "E-Scooter Rider Detection and Classification in Dense Urban Environments." arXiv preprint arXiv:2205.10184 (2022).
- [6] AI 허브 인도보행 영상, "인도 데이터셋", [https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=real\\_m&dataSetSn=189](https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=real_m&dataSetSn=189), (2022-09-21)
- [7] Jin, Dongkwon, et al. "Eigenlanes: Data-driven lane descriptors for structurally diverse lanes." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.
- [8] Lee, Michelle A., et al. "Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks." 2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019.