JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 27, No. 8, pp. 2263-2269, Aug. 2026

Check for updates

전동 퀵보드 객체 탐지 성능 향상을 위한 상황 기반 데이터 증강 기법

황 현 성[1] · 홍 기 훈[2*]
[1]서원대학교 AI소프트웨어학과 학사과정
[2]서원대학교 AI소프트웨어학과 교수

A Context-Based Data Augmentation Method for Improving Electric Scooter Object Detection Performance

Hyeon-Seong Hwang[1] · Ki-Hun Hong[2*]

[1]Undergraduate, Department of AI Software, Seowon University, Cheongju 28674, Korea
[2]Professor, Department of AI Software, Seowon University, Cheongju 28674, Korea

[요 약]

본 연구는 도심 보도에 방치 및 불법 주정차된 공유 전동 퀵보드를 안정적으로 탐지하기 위해 YOLOv8m 기반 객체 검출 모델을 구축하고 방치된 도로 상황에 필요한 데이터 증강 기법을 적용한 연구이다. 전동 퀵보드 데이터셋 원본과 기본 증강 (conventional augmentation) 그리고 쓰러짐/기울어짐 (fallen/tilted), 강한 그림자 (shadow), 부분 가림 (occlusion)을 모사하는 맞춤형 증강 모듈을 설계 및 적용한 상황 기반 증강 (proposed) 방법을 적용하여 3종 데이터셋을 구성하고 학습하였다. 동일한 학습 설정에서 제안하는 방법은 Precision 0.9295, Recall 0.8829, mAP@50-95 0.6442로 가장 우수한 성능을 달성하였다. 결과적으로 도심 환경의 다양한 비정상적 상황을 데이터 수준에서 반영하는 전략이 미탐 감소와 일반화 성능 향상에 효과적임을 보여주고 있다.

[Abstract]

This study develops a YOLOv8m-based object detection system to reliably detect shared electric scooters abandoned or illegally parked on urban sidewalks. To address the complexities of real-world road environments, a context-aware augmentation strategy was designed and implemented. Three distinct datasets were constructed and trained: the original electric scooter dataset, a conventionally augmented dataset, and a proposed dataset incorporating customized augmentation modules that simulate specific scenarios such as fallen/tilted orientations, strong shadows, and partial occlusions. Under identical training configurations, the proposed method achieved superior performance with a Precision of 0.9295, a Recall of 0.8829, and an mAP@50-95 of 0.6442. These results demonstrate that a strategy reflecting various abnormal urban conditions at the data level is highly effective in reducing missed detections and enhancing the generalization performance of the model.

색인어 : 데이터 증강, 상황, 이미지 합성, 객체 탐지, 전동 퀵보드
Keyword : Data Augmentation, Context, Image Synthesis, Object Detection, Electric Scooter

http://dx.doi.org/10.9728/dcs.2026.27.8.2263

Received 12 May 2026; **Revised** 29 June 2026
Accepted 21 July 2026

***Corresponding Author; Ki-Hun Hong**
Tel: +82-43-299-8777
E-mail: khong@seowon.ac.kr

Ⅰ. 서 론

최근 공유 모빌리티 확산으로 전동 퀵보드 이용이 일상화되었으나 보도, 횡단보도 인근 방치와 무질서한 주정차는 보행권 침해 및 안전사고 위험을 키우는 사회적 문제로 부각되고 있다. 기존 관리 방식은 위성 항법 장치(GPS; global positioning system)와 앱 기반 방안에 크게 의존하는데, 실제 현장에서는 위치 오차, 건물 반사, 지형 영향 등으로 정밀한 상태 판단이 어렵고, 결국 인력 단속과 수거에 의존하게 되어 비용과 대응 속도 측면에서 한계가 발생한다. 영상 기반 탐지는 기울어진 퀵보드와 보도 점유 등 물리적 상태를 직접 확인할 수 있다는 점에서 유효하지만, 도심 영상은 시간대별 명암 변화, 그림자, 부분 가림, 다양한 촬영 각도, 촬영 거리 등 복합적인 변수가 존재해 모델이 쉽게 혼동한다. 따라서 실사용을 위해서는 일반적인 학습 데이터만으로는 부족하며, 실패가 잦은 상황을 데이터 차원에서 충분히 재현해 학습시키는 접근이 필요하다.

본 연구는 탐지할 객체가 처한 쓰러짐, 그림자, 가림과 같은 특정 상황 기반의 증강 기법이 기존의 일반적인 증강 기법보다 객체 탐지에 효과적임을 보이고자 한다.

Ⅱ. 관련 연구

논문[1]에서는 도심 환경의 취약 도로 이용자인 전동 퀵보드 탑승자의 탐지 성능을 개선하기 위한 연구를 수행하였다. 해당 연구는 특히 부분 가림(partial occlusion) 상황에서 탐지 정확도가 급격히 저하되는 문제에 집중하여, 가림 정도를 0%에서 99%까지 정량화한 테스트 벤치마크를 구축하고 가림 인지형(occlusion-aware) 탐지 파이프라인을 제안하였다. 방법론적으로는 2단계 구조를 채택하여, CenterNet을 통해 사람 후보군을 먼저 검출한 후 바운딩 박스(bounding box)의 종횡비에 따라 ROI(region of interest)를 기하학적으로 확장하여 전동 퀵보드 본체 정보를 포함하도록 하였다. 이후 ResNet101 분류기를 통해 탑승 여부를 최종 판별함으로써, 가림이 심한 구간에서도 기존 모델 대비 약 15.93%의 정확도 향상을 달성하였다.

전동 퀵보드와 보행자를 검출하여 법규 위반을 단속하고자 하는 시도는 최근 딥러닝 기반 객체 검출 기술을 통해 활발히 이루어지고 있다. 논문[2]에서는 YOLOv5s(small)모델을 기반으로 탑승 여부를 판별하는 RIDE 모델과 법규 위반을 판별하는 VIOLATION 모델의 2단계 파이프라인을 제안하였다. 해당 연구는 라즈베리 파이와 같은 엣지 디바이스 환경에서의 실시간 성능을 확보하기 위해 경량화 모델인 YOLOv5s를 채택하였으며, RIDE 모델을 통해 검출된 객체 영역(ROI)을 크롭(crop)하여 후속 모델의 입력값으로 사용하는 방식을 통해 배경 잡음을 효과적으로 제거하였다. 실험 결과, RIDE 모델은 mAP50 기준 99.475%라는 높은 정확도를 기록하여 실제 도로 환경에서의 실용성을 입증하였다.

기존 방법들은 전동 퀵보드의 탑승자에 초점을 두고 있고 이동하는 탑승자를 잘 구분 및 탐지하여 각 논문의 목적을 달성하고자 하였다. 그러나 본 연구의 목적은 주로 방치되고 쓰레기 봉투나 자전거, 기타 적재물 등으로 가려진 상황에서 전동 퀵보드를 잘 탐지하려는 목적이다. 또한 분리된 개별 모델 혹은 모델의 계층적 구조를 통해 개선하기 보다는 전동 퀵보드의 상황에 맞는 데이터 증강 기법을 적용하여 이를 해결하고자 하였다. 이러한 방법은 타겟 객체가 바뀌고 그 객체가 처한 상황이 바뀌어도 객체 데이터셋과 상황 이미지 샘플을 변경하면 바로 개선된 학습이 가능하다는 점에서 기술적 차별성을 가진다[3]-[6].

Ⅲ. 제안하는 데이터 증강 방법

본 연구는 도심 환경에서 방치, 불법 주정차된 공유 전동 퀵보드를 안정적으로 탐지하기 위한 객체 탐지 시스템을 설계하고, 모델 구조 변경보다 데이터셋 구성과 상황 기반 증강 전략이 탐지 성능에 미치는 영향을 정량적으로 분석하였다. 전체 처리 절차는 도심 전동 퀵보드 이미지를 수집하고, 전처리 및 전통적인 기본 증강과 상황 기반 증강을 수행하여 각각 학습 데이터셋을 구축한다. 각 증강 방법별 데이터셋들을 YOLOv8m(medium) 기반으로 학습, 추론, 평가, 시각화 단계로 구성하였다. 이러한 데이터 증강 기법을 통해 탐지 성능을 향상시키는 것을 목표로 하였다[7].

3-1 전동 퀵보드 이미지 데이터 수집

데이터 수집은 전동 퀵보드가 많이 배치되는 도심 보도, 건물 출입구 주변, 자전거 도로, 보도와 차도의 경계부 등을 중심으로 선별하였고, 직사광선이 강한 정오, 건물 그림자가 길게 드리우는 오후, 흐린 날 등 다양한 조명 조건을 포함하도록 계획하였다. 데이터셋은 1,231개 이미지 샘플로 Roboflow 데이터셋 1,031개 샘플[8]과 한국의 전동 퀵보드 이미지를 반영하기 위해 웹에서 크롤링한 샘플 200개를 사용하였다. 모델 평가의 신뢰성을 위해 데이터는 학습, 검증, 테스트로 7:2:1 분할했고, 동일 위치에서 촬영된 이미지가 서로 다른 데이터셋에 동시에 포함되지 않도록 하였다.

3-2 전처리

전처리 모듈은 수집된 이미지를 YOLOv8m 모델의 입력 형식에 최적화하면서도 원본 데이터의 기하학적 특성을 보존하기 위한 변환을 수행하였다. 먼저 모든 이미지는 긴 변을 기준으로 리사이즈한 뒤 빈 공간에 패딩을 추가하는 Letterbox 방식을 사용하여 640×640 해상도로 통일하였다.

이는 YOLO 계열 모델의 입력 특성을 반영함과 동시에, 이미지 왜곡을 최소화하여 객체의 종횡비를 유지하기 위함이다. 이 과정에서 바운딩 박스 좌표 역시 동일 비율로 스케일링하였으며, 패딩 영역에 따른 오프셋을 적용하여 좌표계를 재산출하였다.

3-3 기본 증강 방법(Conventional Augmentation Method)

기본 증강 단계에서는 좌우 반전(horizontal flip), 밝기·대비 조절(brightness/contrast), 약한 가우시안 블러 등의 연산을 제한적으로 적용하였다. 밝기·대비 조절은 ±20% 범위 내에서만 변화시키고, 과도한 노이즈 추가나 색조 변경은 지양하였다. 이는 뒤에서 설명할 상황 기반 증강이 훨씬 강한 변형을 수행하기 때문에, 기본 증강 단계에서 지나치게 큰 변형을 가하면 전체 데이터 분포가 인위적으로 왜곡될 위험이 있기 때문이다. 또한 YOLOv8m이 기본적으로 제공하는 Mosaic, MixUp 등 고수준 증강 옵션에 대해서는 실험적으로 영향을 확인한 뒤, 본 연구에서는 Mosaic의 빈도와 종료 시점만 조정하고 나머지는 비활성화하였다. Mosaic은 네 장의 이미지를 하나로 합성하여 다양한 배치와 스케일 변화를 제공하지만, 전동 퀵보드가 작은 객체로 등장하는 경우 지나치게 작아져 학습에 불리하게 작용할 수 있다. 이에 따라 학습 초기 50% 구간까지만 Mosaic을 적용하고, 이후 에포크(epoch)에서는 Mosaic을 해제하여 안정적인 수렴을 유도하였다. MixUp과 CopyPaste는 미사용하였고 HSV(hue, saturation, value) augmentation은 hsv_h 0.015, hsv_s 0.7, hsv_v 0.4를 사용하였다.

3-4 상황 기반 증강 방법(Proposed Augmentation Method)

상황 기반 증강 모듈은 본 연구의 핵심 구성 요소로 Fallen/Tilted, Shadow, Occlusion 세 가지 하위 모듈로 구성된다. 이 모듈의 목적은 탐지할 객체의 실제적 상황을 반영한 데이터 증강 기법과 오탐 혹은 미탐의 원인인 가리는 객체를 데이터에 반영하여 이를 학습함으로써 강건하게 만드는 것이다. 실제 도심 환경에서 전동 퀵보드가 처하는 다양한 상황을 데이터 수준에서 적극 반영함으로써, 모델이 “쓰러진 상태, 강한 그림자, 부분 가림”과 같은 객체 탐지가 어려운 케이스에서도 안정적으로 탐지하도록 만들었다.

1) Fallen/Tilted 모듈

Fallen/Tilted 모듈은 원본 이미지와 그에 대응하는 바운딩 박스를 동일한 회전 변환 행렬에 의해 변환한다. 회전 각도는 −60°에서 +60° 사이에서 균일하게 샘플링하였으며, 회전 후 이미지 경계 밖으로 나가는 영역은 잘라낸 뒤 빈 영역은 검은색으로 채운다. 이때 바운딩 박스 좌표는 네 꼭짓점 좌표를 모두 회전시킨 뒤, 다시 축 정렬된 최소 외접 사각형을 계산하여 새로운 박스로 사용하였다. 결과적으로 기존에는 세워진 상태로만 존재하던 퀵보드가, 다양한 쓰러짐·기울기 자세로 학습 데이터에 포함되도록 하였다.

2) Shadow 모듈

Shadow 증강 모듈은 전동 퀵보드 주변에 불규칙한 형태의 그림자를 합성하여 가변적인 조명 조건을 시뮬레이션한다. 먼저, 객체 바운딩 박스를 기준으로 해당 영역 내에 무작위 다각형 마스크를 생성한다. 생성된 마스크는 0.4에서 0.8 사이의 알파(alpha) 값을 가지며, 원본 이미지와의 선형 보간(linear interpolation)을 통해 자연스럽게 합성된다. 이때 광원의 방향성(예: 좌우 또는 상하 방향)을 재현하기 위해 다각형의 기울기와 좌표를 정밀하게 제어하였다. 이러한 기법은 강한 직사광선에 의한 명암 대비나 건물 그림자에 투영된 퀵보드의 외형 변화를 모사함으로써 모델의 탐지 성능에 긍정적 영향을 줄 것으로 보인다.

3) Occlusion 모듈

Occlusion 모듈은 사진 촬영과 웹을 통해 수집한 occluder 이미지(쓰레기봉투, 가로등 기둥, 도로 표지판, 화분 등)를 전동 퀵보드 바운딩 박스 영역에 합성하는 기능을 수행한다. 합성 절차는 다음과 같다. 첫 번째, occluder 이미지의 투명 배경을 유지한 채, 전동 퀵보드 박스의 20~60% 정도를 가리도록 크기를 조정한다. 두 번째, occluder를 배치할 위치를 박스 내부에서 랜덤하게 선택하되, 완전히 퀵보드를 가리거나 전혀 가리지 않는 경우는 제외한다. 마지막으로 알파 채널을 이용해 자연스럽게 합성하고, 필요 시 밝기와 색상을 약간 조정하여 배경과 이질감이 줄어들도록 한다. Occlusion 모듈을 통해 생성된 이미지는 실제 도심에서 관찰되는 “쓰레기봉투 뒤에 반쯤 숨은 퀵보드, 나무 기둥에 가려진 퀵보드”와 유사한 형태를 갖는다.

3-5 객체 탐지 모델 학습 및 시스템 구현

그림 1은 시스템의 전반적인 데이터 처리와 학습 흐름을 보여주고 있다. 학습 및 탐지 모델은 Ultralytics 공식 구현인 YOLOv8m을 사용하였는데 속도와 정확성 부분에서 균형을 이루는 모델로 연구용으로 많이 활용되는 모델이다. 세 실험 모두 입력 이미지 크기는 640×640 픽셀, 학습 횟수는 120 에포크, 배치 사이즈는 8로 통일하였다. 초기 학습률 0.01에서 코사인 스케줄러(cosine scheduler)를 통해 점진적으로 감소시켰으며, SGD(stochastic gradient descent) 옵티마이저(momentum 0.937, weight decay 0.0005)를 사용하여 모델의 안정적인 수렴을 도모하였다.

내장 증강 하이퍼 파라미터도 비교의 일부로 간주해 동일하게 유지했고, 회전, 이동, 스케일, HSV 변환 등을 설정하

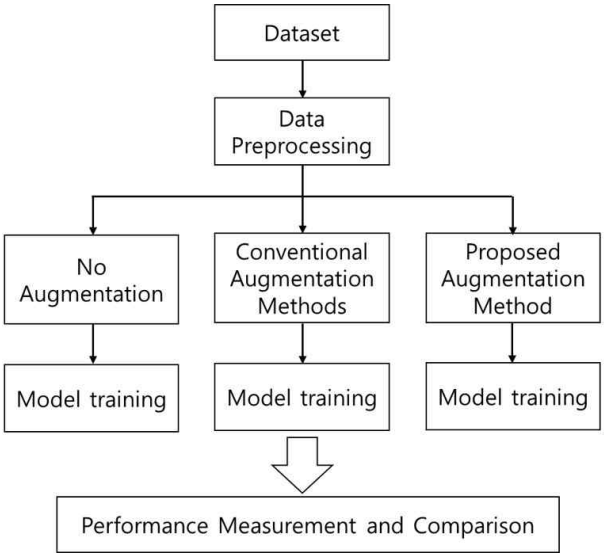

그림 1. 시스템의 데이터 처리 흐름
Fig. 1. Data processing flow of the system

되, MixUp과 CopyPaste는 비현실적 겹침을 줄이고 제안 증강의 효과를 명확히 보기 위해 적극적으로 사용하지 않았다. Mosaic은 데이터 다양성을 제공하지만 작은 객체가 더 작아지는 문제가 있어 학습 초기 일부 구간에서만 제한적으로 사용하고 후반에는 비활성화하는 방식으로 정밀도를 보완했다. 평가는 공통 테스트셋에서 Precision, Recall, mAP@50-95를 측정하고, 동일 테스트 이미지를 입력했을 때 세 모델의 박스 출력 결과를 한 화면에서 비교할 수 있는 웹 기반 UI를 구현하였다.

Ⅳ. 실험 및 성능 평가

4-1 실험 환경 및 사용 데이터셋

실험 환경(컴퓨터 사양)은 표 1과 같다. 실험에 사용된 데이터셋은 Roboflow[8] 1,031개와 웹 크롤링을 통해 수집한 200개를 합친 총 1,231개로 구성하였다. 이 중 862개를 학습 데이터셋으로 분리하였으며, 이를 기반으로 증강 기법을 적용하였다.

상황 기반 증강은 각 모듈의 적용 여부와 빈도를 확률적으로 제어하였다. 특정 이미지에 'Fallen/Tilted'와 'Occlusion'을 동시에 적용할 확률을 0.3, 'Shadow'를 단독 적용할 확률을 0.2로 설정하는 등, 다양한 상황 조합이 전체 데이터셋에 균형 있게 반영되도록 하였다.

데이터셋은 크게 세 가지로 구축하였다. 첫째, 증강을 하지 않은 862장의 원본 데이터셋이 있다. 둘째, 기본 증강 기법을 적용하여 2,287장의 데이터셋을 구축하였으며, 여기에는 원본 862장이 포함된다. 셋째, 상황 기반 증강 기법을 적용하여

표 1. 실험용 컴퓨터 사양
Table 1. Spec. sheet of simulation computer

CPU	Intel i7-12700 2.1GHz
RAM	32 GB
GPU	NVIDIA GeForce RTX 5060
OS	Ubuntu 22.04.5 LTS

4,300장의 데이터셋을 구축하였으며, 이는 기본 증강 데이터셋(2,287장)을 포함하도록 하였다.

실험의 공정성을 확보하기 위해 모든 학습 과정에서 1 에포크당 샘플 수를 862개로 고정하였다. 증강 데이터셋을 활용한 학습 시에는 해당 데이터 풀(Pool)에서 매 에포크마다 862개를 무작위로 추출(random sampling)하여 모델을 학습시켰다. 성능 테스트 시에는 증강 전에 미리 7:2:1(학습:검증:테스트)로 분할해 놓은 테스트셋을 사용하여 증강 데이터가 포함되지 않는다.

4-2 객체 탐지 성능 지표

각 증강 기법별 성능을 비교하기 위해 주로 사용된 성능 지표는 Recall, Precision, mAP@50-95 등이며 이에 대한 정의 및 설명은 다음과 같다.

수식 (1)은 정밀도(precision)를 의미하며 객체 탐지에서 모델이 확신이 있는 경우 검출 결과를 내놓는다. (TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative)

$$\mathrm{Pr}ecision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

수식 (2)는 재현율 (Recall)을 의미하며 객체 탐지에서 실제 물체를 놓치지 않고 최대한 많이 탐지하는지를 나타낸다.

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{3}$$

또한 객체 탐지의 성능 측정에서 많이 활용되는 것은 IoU (intersection over union)로 수식 (3)과 같이 객체 탐지 모델이 그린 박스와 실제 정답 박스가 겹치는 비율을 표현한다. 일반적으로 IoU가 0.5 즉, 50% 이상 겹치면 맞힌 것(TP)으로 인정한다. 객체 탐지에서 AP (average precision, 평균 정밀도)는 모델이 특정 클래스(예: 전동 퀵보드)를 얼마나 잘 찾아내는지 나타내는 단일 요약 지표이며 Precision (정밀도)과 Recall(재현율)은 서로 상충 관계에 있기 때문에, 이 둘을 종합하여 모델의 성능을 하나의 숫자로 표현하기 위해 AP를 사용한다. AP는 PR 곡선(precision-recall curve)을

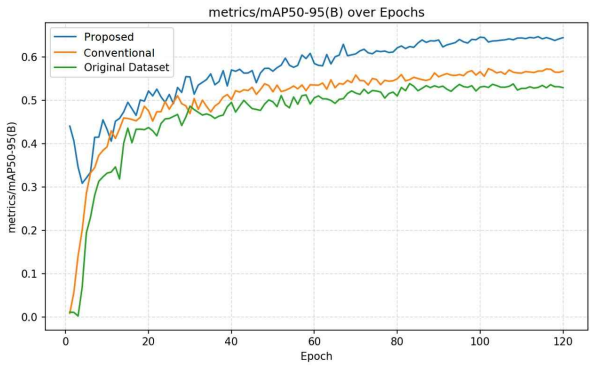

그림 2. 데이터 증강 방법별 mAP@50-95 그래프
Fig. 2. Comparison of mAP@50-95 across different data augmentation methods

그리고 PR 곡선 아래의 면적(AUC: area under the curve)을 의미하며 면적이 1에 가까우면 성능이 좋은 모델이다.

$$mAP@.50:.95 = \frac{1}{10}\sum_{i=0}^{9} AP_{IoU=0.50+0.05\times i} \quad (4)$$

mAP(mean average precision)는 여러 클래스의 객체를 탐지할 때, AP의 모든 클래스에 대해 평균을 내어 계산한 값이다. mAP@50은 IoU 임계값이 50%임을 나타내며 박스가 겹치는 구간이 50% 이상이면 정답으로 인정한다는 의미이다. 수식 (4)는 mAP@50-95를 나타내며 임계값 0.5부터 0.95까지 0.05 단위로 증가한 임계값에 대한 AP값의 평균을 의미한다. 본 연구에서는 단일 객체(전동 퀵보드)를 대상으로 하므로 AP와 mAP는 같으나 성능의 정밀한 분석을 위해 COCO(common objects in context) 데이터셋 가이드라인을 따른 mAP@50과 mAP@50-95를 주요 평가지표로 사용하였다. 이는 다양한 IoU 임계값에 대해 강건한 성능을 보일 수 있는지를 평가하는 지표로 사용된다.

4-3 제안 방법의 성능

그림 1과 같이 3종의 데이터 증강 기법으로 만들어진 학습 데이터셋을 기반으로 반복 학습하면서 모델 성능이 120 에포크 정도에서 수렴했다. 실험 결과인 그림 2와 표 2를 보면, 데이터 증강 기법을 적용하지 않은 방법은 Precision 0.9246, Recall 0.6683, mAP@50-95 0.5292로 정밀도는 높지만 미탐이 많은 보수적 탐지 양상을 보였고, 기본 증강 기법(conventional)은 Recall 0.7270, mAP@50-95 0.5670으로 개선되었으나 Precision이 0.8563으로 낮아져 오탐 증가 가능성이 관찰되었다. 반면 제안하는 방법은 Precision 0.9295, Recall 0.8829, mAP@50-95 0.6442로 가장 우수했으며, 특히, Recall의 유의미한 향상은 방치, 불법 주정차 및 쓰레기 등에 의해 부분적으로 가려진 전동 퀵보드에 대한 탐지 누락을 줄이고, 실제 환경에서의 객체 탐지 능력을 향상시켰음을 의미한다. 표 2에서 Recall 부분만 비교해 보면 고도화된 증강 기법이 적용될수록 목표 객체를 놓치지 않고 잘 탐지하는 경향을 볼 수 있다.

표 2. 데이터 증강 기법별 성능 비교
Table 2. Performance result by data augmentation technique

Aug. Methods	Precision	Recall	mAP@50-95
No Aug.	0.9246	0.6683	0.5292
Conventional	0.8563	0.7270	0.5670
Proposed	0.9295	0.8829	0.6442

그림 3은 증강 기법별 탐지 결과를 보여 주는 샘플 그림으로 맨 우측의 그림에서 검은 봉투와 쓰레기들 그리고 벽면의 창살이 있는 상황에서 제안하는 방법이 가장 우수한 탐지 성능을 보여 주고 있다. 그림 4는 학습 횟수에 따른 각 증강 기법별 Precision, Recall, mAP50, mAP50-95 성능의 변화 과정을 보여주고 있다. 본 제안 기법을 적용한 결과를 보면 mAP@50뿐만 아니라 높은 정밀도를 요구하는 mAP@50-95 지표에서도 유의미한 성능 향상을 확인하였다.

종합하면, 전통적인 기본 증강만으로는 쓰러짐, 강한 그림자, 부분 가림 같은 도심 특수 상황을 충분히 재현하기 어렵기 때문에 목표 상황을 직접 모사하는 상황 기반 증강 기법을 데이터셋에 적용할 때 미탐이 유의미하게 감소하며 전반적인 성능이 향상되었다. 이는 전동 퀵보드에 국한되지 않고 도심 환경에서 정지, 방치될 수 있는 다른 객체 탐지에도 확장 가능한 데이터 증강 기법이라고 할 수 있을 것이다.

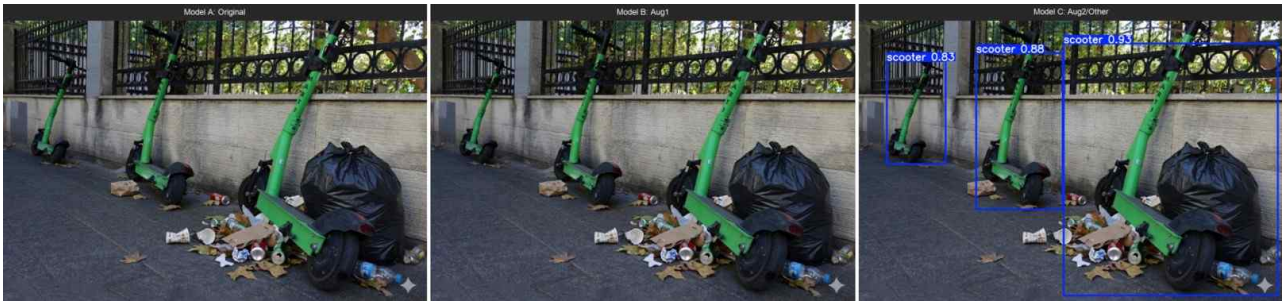

그림 3. 데이터 증강 기법별 탐지 결과 샘플 (Model A: No Aug., Model B: Conventional, Model C: Proposed)
Fig. 3. Sample detection results by data augmentation technique (Model A: No Aug., Model B: Conventional, Model C: Proposed)

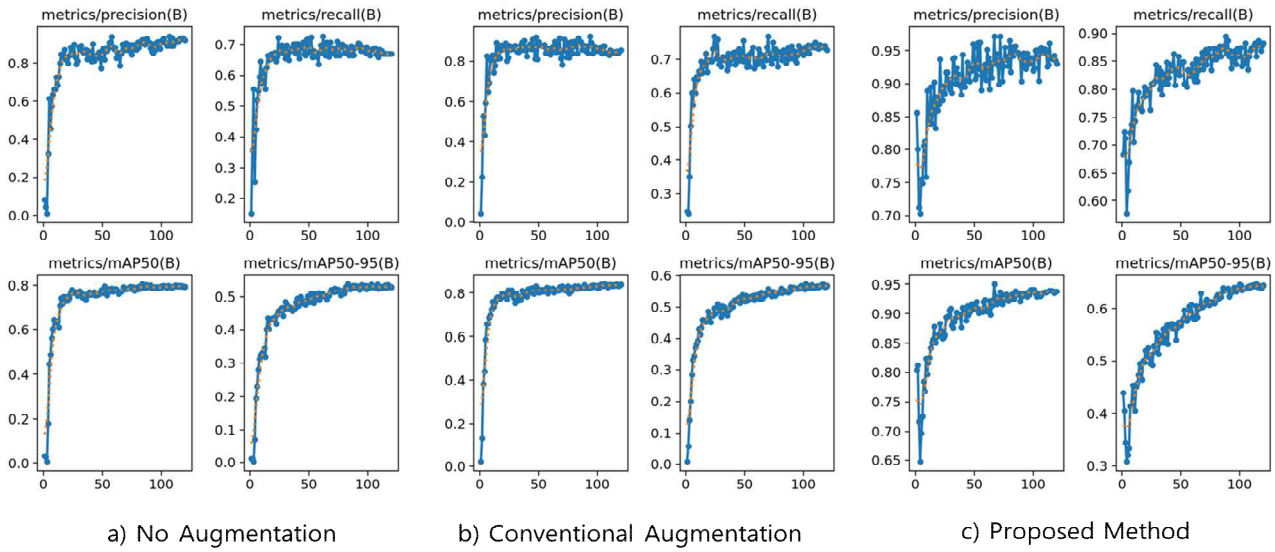

그림 4. 데이터 증강 방법별 Precision, Recall, mAP50, mAP50-95 성능 비교
Fig. 4. Performance comparison of Precision, Recall, mAP50, and mAP50-95 across different data augmentation methods

V. 결 론

본 연구는 도심 환경에서 방치, 불법 주정차된 공유 전동 퀵보드를 안정적으로 탐지하기 위해 객체가 위치한 상황을 기반으로 증강 기법을 제안하여 성능이 개선됨을 실험적으로 확인하였다. 원본 데이터만으로 학습한 경우, 정밀도는 높았지만 쓰러짐, 그림자, 부분 가림과 같은 실제 현장 조건에서 미탐이 증가해 재현율과 mAP@50-95가 제한되는 문제가 나타났다. 기본 증강은 일정 수준의 성능 향상에 기여했으나 정밀도 하락이 동반될 수 있었고, 현실 세계에서 자주 발생하는 복잡하고 탐지가 어려운 상황을 충분히 반영하지 못하였다. 반면 본 연구에서 제안한 실제 상황에서 많이 발생하는 Fallen, Tilted, Shadow, Occlusion 기반 증강 기법을 적용한 데이터셋은 재현율과 mAP@50-95를 유의미하게 향상시키면서도 정밀도를 안정적으로 유지해, 실서비스 관점에서 중요한 미탐 감소 효과를 보여 주었다. 이는 도심 특수 상황을 목표로 설계한 데이터 중심 접근이 탐지 모델의 일반화 성능을 강화하고, 제한된 원본 데이터에서도 활용 가능성을 높인다는 것을 의미한다. 또한 세 모델의 결과를 동일 화면에서 이미지로 비교할 수 있는 시각화 구현 환경을 구성함으로써 성능 차이를 정성적으로 검증하고, 현장 적용 시 고려해야 할 실패 사례를 구체적으로 확인하여 선별할 수 있다.

향후 연구에서는 더 다양한 도시, 기상, 야간 환경으로 데이터 범위를 확장하고, 전동 퀵보드의 상태 분류(정상, 쓰러짐, 보도 점유 등)와 같은 후처리 로직을 결합해 단속 기준에 가까운 판단까지 자동화할 필요가 있다. 더 나아가 영상 스트림 기반 실시간 처리, 경량 모델 적용, 오탐 억제를 위한 추가 학습 전략을 통해 실제 운영 시스템으로의 확장 가능성을 높일 수 있을 것이다. 종합하면 본 연구는 공유 전동 퀵보드 관리 문제를 영상 기반으로 해결하기 위한 실증적 근거를 제시했으며, 데이터 설계와 증강 전략이 도심 객체 탐지 성능을 좌우하는 핵심 요인임을 보여 준다.

감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원(IITP-2026-RS-2020-II201602, 100%)을 받아 수행된 연구임.

참고문헌

[1] S. Gilroy, D. Mullins, E. Jones, A. Parsi, and M. Glavin, “E-Scooter Rider Detection and Classification in Dense Urban Environments,” *Results in Engineering*, Vol. 16, 100677, 2022. https://doi.org/10.1016/j.rineng.2022.100677

[2] S.-H. Lee, S.-H. Oh, and J.-G. Kim, “YOLOv5-Based Electric Scooter Crackdown Platform,” *Applied Sciences*, Vol. 15, No. 6, 3112, 2025. https://doi.org/10.3390/app15063112

[3] E. Syahrudin, E. Utami, and A. D. Hartanto, “Augmentation for Accuracy Improvement of YOLOv8 in Blind Navigation System,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, Vol. 8, No. 4, pp. 579-588, 2024. https://doi.org/10.29207/resti.v8i4.5931

[4] H. S. Park, J. A. Ha, S. P. Shin, S. G. Park, and K. M. Hong, Development of Technology to Detect Illegal Parking of

Shared PM (Personal Mobility) Based on Video AI Technology, Korea Institute of Civil Engineering and Building Technology (KICT), Goyang, KICT 2023-107, 2023.

[5] Y.-H. Woo and M.-Y. Yoo, “Performance Analysis of Temporary Structure Recognition Using YOLOv8 and YOLOv11 Segmentation Models,” *Journal of the Architectural Institute of Korea*, Vol. 41, No. 10, pp. 339-348, 2025. https://doi.org/10.5659/JAIK.2025.41.10.339

[6] J.-S. Jang, D. M. Chu, S. A. Lee, H. G. Noh, and K. S. Chung, “A Study on Data Preprocessing and Augmentation for Robust Object Detection in Traffic CCTV Under Environmental Variations,” in *Proceedings of the 2025 Korea Society of Computer and Information Summer Conference*, Vol. 33, No. 2, pp. 949-952, 2025.

[7] H. S. Hwang and K. H. Hong, “A Study on Improving Object Recognition Accuracy Using Contextual Data Augmentation Techniques,” in *Proceedings of the 2026 Korean Institutes of Communications and Information Sciences Winter Conference*, Yongpyong, Vol. 89, pp. 1346-1347, 2026.

[8] Roboflow. Scooter Computer Vision Model Dataset [Internet]. Available: https://universe.roboflow.com/computer-vision-dibh7/scooter-31wb6, 2025.

황현성(Hyeon-Seong Hwang)

2026년 : 서원대학교
AI소프트웨어학과

※관심분야 : 객체탐지(Object Detection), 영상 처리(Image processing)

홍기훈(Ki-Hun Hong)

2000년 : 숭실대학교
전자전기정보통신공학부
졸업

2002년 : 숭실대학교 정보통신공학과
(공학석사)

2006년 : 숭실대학교 정보통신공학과
(공학박사-사이버 보안)

2006년～2007년: 미국 UC Davis, Post-Doc.

2008년～2022년: ㈜시큐아이 정보보안연구소 수석연구원

2022년～2025년: 숭실대학교 전자정보공학부 IT융합전공
산학협력중점교수

2025년～현 재: 서원대학교 AI소프트웨어학과 교수

※관심분야 : AI 보안, AI 응용, 적대적 공격 방어, 사이버 보안 등