

딥러닝 모델에 따른 버스 주변 교통환경 인식 성능 비교 및 분석

강신재, 은승우, 이재우, 한동석*

경북대학교

kangsj129@knu.ac.kr, sweunwave@knu.ac.kr, jeu@auto-it.co.kr *dshan@knu.ac.kr

Comparative Analysis of Traffic Environment Recognition Around Buses Based on Deep Learning Models

Kang Sin Jae, Eun Seung Woo, Lee Jae Woo, Han Dong Seog*

School of Electronic and Electrical Engineering, Kyungpook National Univ. Automotive Information Technology Inc.

요 약

도심에서는 다양한 유형의 차량과 교통 흐름을 조절하는 신호 및 제어 시스템이 복잡한 환경을 형성한다. 차량과 보행자, 비차량 간의 상호작용이 빈번하게 발생하며, 특히 버스는 보행자, 자전거, 킥보드 등과 자주 상호작용한다. 이로 인해 버스 주변에서 움직이는 객체를 탐지하는 것은 사고 예방과 효율적인 교통 운영에 필수적이다. 딥러닝 기반의 최신 SOTA(state of the art) 모델을 도입하면 버스 주위 환경에서 객체를 효과적으로 탐지할 수 있으며, 이는 승하차하는 승객들의 안전 보호에 기여할 수 있다. 본 논문에서는 다양한 최신 SOTA 모델을 적용하여 성능을 비교하고 분석한다. 또한, 이러한 분석 결과를 바탕으로 계산 비용과 정확도 측면에서 최적의 성능을 가진 모델을 개발하여 실차량에 적용할 수 있다. 이를 통해 더 쾌적한 도로 환경과 승객 안전성을 강화하는 데 기여하기를 기대한다.

I. 서 론

도심 내 교통 흐름은 차량과 보행자, 그리고 비차량 간의 빈번한 상호작용으로 인해 복잡한 환경을 형성한다. 특히 버스는 보행자, 자전거, 킥보드 등과 자주 상호작용하며, 이들 객체를 탐지하는 것은 사고 예방과 효율적인 교통 운영에 필수적이다. 본 연구는 딥러닝 기반의 최신 기술을 적용하여 버스 주변 객체를 탐지함으로써 교통 안전성을 향상시키는 것을 목표로 한다.

딥러닝 기반의 최신 SOTA(state of the art) 모델의 적용은 여러 가지 이점을 제공한다. 첫째, 도로 위에서 차량과 비차량 간의 상호작용으로 인한 사고 위험을 감소시킨다. 둘째, 교통 흐름을 개선하여 대중교통의 시간 효율성을 높인다. 셋째, 딥러닝을 통해 주변 환경을 탐지함으로써 안정적인 대중교통 시스템을 구축하고, 자율주행 버스 개발에도 활용할 수 있다. 마지막으로, 버스에 장착된 카메라와 딥러닝 기술을 통해 얻은 객체 탐지 데이터를 바탕으로 도시의 교통 흐름을 분석하고, 이를 통해 교통 정책 및 인프라 계획 수립에도 유용하게 활용할 수 있다.

II. 본론

가. 딥러닝 기반 모델 학습 환경

본 연구에서는 오토아이티(주)에서 제공한 비디오 형식의 데이터셋을 학습에 활용하였다. 이 데이터셋은 보행자, 자전거, 킥보드 세 가지 객체를 확장한 낮 시간에 다양한 각도에서 촬영한 영상으로 구성되어 있다. 저자는 비디오 데이터를 개별 이미지로 변환한 후, 보행자, 자전거, 킥보드의 세 클래스로 라벨링을 진행하였다. 데이터는 전체 82의 비율로 훈련 데이터(1,561장)와 검증 데이터(356장)로 분할하였으며, 데이터 전처리에는 기본적인 이동 및 플립 기술만을 사용하였다.

모델 훈련에는 NVIDIA RTX-3070 GPU를 사용하여 단일 GPU 환경에서 학습을 진행하였다. 모든 모델의 훈련은 200 에포크로 설정되었으며,

CUDA 메모리에 맞추어 결과물의 해상도는 640, 배치 사이즈는 4로 설정하였다. YOLOv5 이후 도입된 너비/깊이 배수는 모두 1로 고정하였고, 초기 가중치는 난수 생성 방식을 통해 설정하였다. 또한 COCO 데이터셋에서 사전 훈련된 가중치를 활용하여 학습하였다.

나. 딥러닝 기반의 SOTA 모델

본 연구에서 훈련된 모델은 1단계 탐지기(1-stage-detector)인 YOLO 시리즈로, 2단계 탐지기(2-stage-detector) 모델인 R-CNN 계열 모델 [1, 2, 3, 4]들에 비해 빠른 속도를 자랑하며, 실시간 탐지에 유리하다. 최근 주목받고 있는 SOTA 모델인 YOLOv8, YOLOv9 [5], GELAN[5], YOLOv10 [4] 시리즈는 뛰어난 성능뿐만 아니라 백본-넥-헤드 구조의 유연함이 특징인 확장성이 높은 모델이다.

YOLOv8은 YOLOv5와 유사하지만, 아키텍처를 구성하는 모듈에서 차이가 있다. YOLOv8은 연산량이 적으면서도 높은 성능을 유지한다. 또한, YOLOv8의 헤드는 디커플드 헤드(decoupled head) 구조로, 회귀(regression) 부분과 분류(classification) 부분이 분리되어 있어, 단일 헤드 구조에 비해 속도가 빠르고 성능도 우수하다.

YOLOv9은 PGI를 사용하여 정보 병목 현상을 완화하고, 더 나은 그라디언트 정보를 통해 모델을 업데이트할 수 있다. 또한, 가역 함수를 사용한 보조 네트워크가 도입되어 더 나은 일반화 성능을 보이며 수렴 속도도 향상된다. 추가적으로, GELAN은 YOLOv9에서 보조 네트워크 제거한 버전으로, 커스텀 데이터셋에서 YOLOv9보다 우수한 성능을 발휘한다.

가장 최근에 주목받고 있는 YOLOv10 모델은 NMS(non-maximum suppression) 후처리 과정을 개선하였으며, 효율성과 정확성을 모두 고려하여 재설계되었다. 이 모델은 공개 데이터셋에서 가장 높은 성능을 기록하였다.

다. 딥러닝 모델들의 훈련 비교 및 분석

표 1~4는 오토아이티㈜에서 제공한 데이터셋을 사용하여 딥러닝 모델을 학습시킨 결과를 클래스별로 정밀도(precision), 재현율(recall), 평균-평균 정확도(mAP)로 나누어 나타낸 것이다. 각 클래스별 모델의 성능을 보다 명확하게 비교할 수 있도록 구성되었다.

표 1. YOLOv8 학습 결과

Model	Class	Precisi on	Recall	mAP50	mAP50-95
YOLOv8	All	0.961	0.98	0.974	0.74
	Bicycle	0.98	0.988	0.986	0.73
	Pedestrian	0.977	0.959	0.979	0.753
	Kickboard	0.927	0.992	0.956	0.736

표 2. YOLOv9 학습 결과

Model	Class	Precisi on	Recall	mAP50	mAP50-95
YOLOv9	All	0.971	0.977	0.978	0.754
	Bicycle	0.988	0.988	0.988	0.753
	Pedestrian	0.984	0.96	0.984	0.772
	Kickboard	0.94	0.984	0.961	0.738

표 3. GELAN 학습 결과

Model	Class	Precisi on	Recall	mAP50	mAP50-95
GELAN	All	0.974	0.979	0.983	0.758
	Bicycle	0.979	0.975	0.987	0.747
	Pedestrian	0.976	0.978	0.99	0.773
	Kickboard	0.966	0.984	0.972	0.753

표 4. YOLOv10 학습 결과

Model	Class	Precisi on	Recall	mAP50	mAP50-95
YOLOv10	All	0.961	0.98	0.974	0.74
	Bicycle	0.98	0.988	0.986	0.73
	Pedestrian	0.977	0.959	0.979	0.753
	Kickboard	0.927	0.992	0.956	0.736

이들을 분석한 결과, 정밀도 측면에서는 YOLOv9와 GELAN 모델이 우수한 성능을 보였으며, 재현율 측면에서는 YOLOv8이 뛰어난 성능을 나타냈다. 평균-평균 정확도(mAP)에서는 GELAN이 가장 우수했다. 재현율이 높다는 것은 모델이 가능한 많은 긍정 사례를 포착할 수 있음을 의미한다. 그러나 이는 부정율 긍정으로 오인할 가능성이 높아질 수 있어, 정밀도와 상충 관계가 존재한다. YOLO 시리즈는 이러한 정밀도와 재현율의 균형을 맞추며 성능을 발전시켜왔다. 따라서, YOLOv8이 재현율에서 강점을 보이지만, 정밀도에서는 YOLOv9와 GELAN에 뒤처지는 것을 관찰할 수 있다. 최신 SOTA 모델인 YOLOv10은 본 연구에서 활용된 데이터셋에 대해서 다른 모델들에 비해 다소 아쉬운 성능을 보여주었다. 일반적으로 공개 데이터셋에서 YOLOv10이 성능과 연산량에서 좋은 모습을 보이거나 커스텀 데이터셋에서 그렇지 못했는데, 저자는 그 이유를 YOLOv10에 있는 어텐션 모듈에 있을 것이라 추측한다. 어텐션 모듈은 트랜스포머 모델의 주요 메커니즘으로 객체 탐지에 있어 뛰어난 성능을 보이지만 파라미터 수가 많아 상대적으로 적은 데이터셋에서는 과적합되어 성능을 감소시킬 수 있다. 그러나 기본적으로 템스-와이즈 컨볼루션을

활용하기 때문에 다른 모델들 보다 파라미터 수, 연산량 부분에서 좋은 모습을 보여주었다.

이와 같이 각 모델은 서로 다른 강점을 가지고 있어, 특정 사용 사례에 따라 적합한 모델을 선택할 필요가 있다.

III. 결론

본 논문에서는 버스 주변 상황을 인지하고 대처하여 승객의 안전성과 교통 흐름의 효율성을 높이기 위해 딥러닝 기반의 최신 SOTA 모델들을 적용하여 비교 및 분석하였다. 오토아이티㈜에서 제공받은 커스텀 데이터셋으로 훈련된 모델들은 전반적으로 좋은 성능을 보였으나, 정밀도에서는 GELAN 모델이, 재현율에서는 YOLOv8이, 파라미터 수와 연산량 측면에서는 YOLOv10이 우수한 결과를 나타냈다.

본 연구는 향후 높은 정확도와 재현율을 가진 성능 좋은 모델을 개발하기 위해 진행된 실험으로, 다른 객체 탐지 모델을 추가로 학습하여 비교한 다음 최종 목적에 부합하는 향상 가능성이 있는 모델을 선택하여 개선된 모델을 개발할 계획이다.

모델의 성능을 더욱 향상시켜 본 연구의 개선된 모델이 교통 흐름의 효율성, 안전성, 그리고 자율주행 버스 개발에 기여할 수 있을 것으로 기대한다.

ACKNOWLEDGEMENT

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술평화원의 지원(P0024162, 2023년 지역혁신클러스터육성)과 정부(과학기술정보통신부)의 재원으로 과학기술사업화진흥원의 지원을 받아 수행된 연구임(‘학연협력플랫폼구축 시범사업’ RS-2023-00304695)

참 고 문 헌

[1] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).

[2] Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

[3] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.

[4] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-time end-to-end object detection. arXiv preprint arXiv:2405.14458.

[5] Wang, C. Y., Yeh, I. H., & Liao, H. Y. M. (2024). YOLOv9: Learning what you want to learn using programmable gradient information. arXiv preprint arXiv:2402.13616.