BiFPN을 이용한 LiDAR 2D 객체 탐지에 관한 연구

박재영, 한동석*
경북대학교 전자전기공학부
endimion123@knu.ac.kr, *dshan@knu.ac.kr

A Study on LiDAR 2D Object Detection Using BiFPN

Jae Young Park, Dong Seog Han*
Kyungpook National Univ.

요 약

본 논문은 YOLOv8을 사용해 2D LiDAR 파노라마 이미지에서 효율적인 객체 탐지를 위해 BiFPN(Bidirectional Feature Pyramid Network) 사용을 제안한다. 기존 PANet(Path Aggregation Network) 기반 neck의 한계를 극복하고자, 정보 손실을 줄이는 스킵 연결(skip-connection)과 중요도 기반 특징 통합 방식을 적용하였다. BiFPN의 반복 적용 실험(2~4회)을 통해 2회 반복이 가장 높은 정확도를 보였으며, 그 이상에서는 오히려 성능 저하가 나타났다. 본 연구를 통해 실시간 처리 성능을 유지하면서도 소형 객체에 대한 탐지 정확도를 향상시킬 수 있는 경량 탐지 구조를 제시함으로써, 향후 자율주행 시스템에 안정적이고 실용적인 LiDAR 기반 인식 모듈로 활용될 수 있을 것으로 기대된다.

Ⅰ. 서 론

최근 자율주행 기술의 상용화가 가속화됨에 따라, 다양한 센서를 활용한 실시간 객체 탐지 기술에 대한 중요도가 높아지고 있다. 실시간 객체 탐지 기술은 주변 환경의 급격한 변화에 빠르게 대응하고, 주행 경로 설정이나 충돌 회피와 같은 의사결정을 즉시 수행해야 하는 시스템에서 필수적이다.

YOLO(You Only Look Once)는 객체 탐지를 위한 대표적인 단일 단계 탐지 모델로, 높은 추론 속도와 경량성을 바탕으로 실시간 응용 분야에서 널리 활용되고 있다.[1] 특히 YOLOv8은 이전 버전들에 비해 backbone 및 neck 구조가 개선되었으며, 다양한 하드웨어 환경에서도 최적의 속도와 정확도를 달성할 수 있도록 설계되었다.[2] 이러한 특징은 실시간성이 중요한 LiDAR 기반 시스템과의 결합에 적합하다. 따라서 본 연구에서는 YOLOv8을 기반으로 구조 개선 실험을 수행하고, 이를 통해 2D 객체 탐지 성능을 향상시키고자 한다.

본 연구에서는 YOLOv8에 기반한 LiDAR 파노라마 이미지에서의 객체 탐지의 정밀도를 향상시키기 위해 BiFPN(Bidirectional Feature Pyramid Network)과 가중합 기반 연산(weighted addition)의 결합 구조를 제안한다.[3] 이는 기존 YOLOv8 구조의 단점을 보완하고 LiDAR 이미지의 특성을 효과적으로 반영하기 위한 목적을 가진다.

Ⅱ. 본론

2.1 데이터 구성

연구에서 사용된 데이터는 실제 도로 환경에 탑재된 차량의 루프 상단에 장착된 LiDAR 센서로부터 획득된 근적외선 파노라마 이미지이다. 클래스는 실제 주행 상황에서 볼 수 있는 객체들로 차, 트럭, 버스, 자전거, 오토바이 그리고 사람이다. 이미지의 크기는 가로 1024픽셀, 세로 128픽셀으로 8:1의 종횡비를 가진다. 객체 탐지 시 목표로 설정한 최소 크기는 차량의 경우 600픽셀, 사람의 경우 400픽셀이다.

2.2 BiFPN 구조

연구에 다양한 해상도의 특징 맵을 효과적으로 결합하여 탐지 정확도를 향상하기 위해, BiFPN에 가중합 기반 연산을 통합하여 YOLOv8의 Neck 구조를 개선하였다. 기존 YOLOv8 Neck은 PANet (Path Aggregation Network)을 기반으로 하나, 상·하향 양방향 경로를 제공함에도 실제로는 한 차례의 상향식 진행 이후 하향식 진행만 수행되며 각 노드가 정적 원소별 합으로 특징을 합성한다. 이러한 방식은 모든 입력 특징을 동일한 중요도로 취급하여 상대적으로 중요하지 않은 정보까지 함께 반영되고, 이는 불필요한 정보의 강조 또는 중요한 정보의 희석으로 이어져 탐지 정확도를 제한한다. 이 문제를 해결하기 위해 BiFPN을 활용한 neck 구조를 설계하였다. BiFPN은 스킵 연결(skip-connection) 구조를 사용해 동일 해상도 간의 정보가 전달되어 정보 표현력을 높인다. 여기에 가중합 기반 연산을 도입하여 각 해상도 간의 기여도를 동적으로 조절함으로써, 특정 해상도에 민감한 객체 탐지 성능을 보완하였다. 이 연산은 연산량을 적게 유지하면서도 정보 보존력을 향상시키는 장점을 가지며,

그림 1. LiDAR 근적외선 파노라마 이미지

특히 수직 해상도가 낮은 파노라마 이미지에서 작은 객체의 탐지 성능 향상에 도움을 준다.

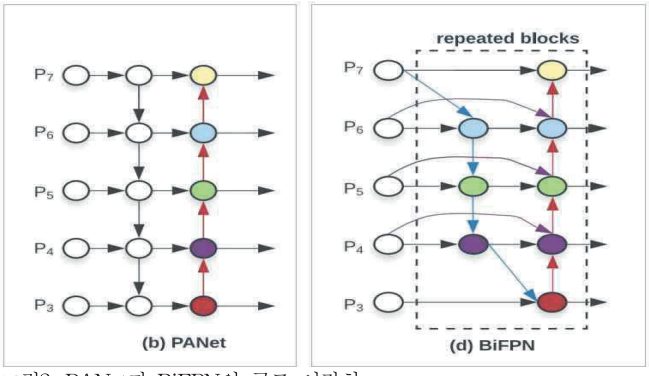

그림2. PANet과 BiFPN의 구조 시각화

2.3 연구 실험

실험은 윈도우 환경에서 Python 3.10, Conda 11.8을 기반으로 진행되었으며, 하드웨어는 NVIDIA RTX 4080 Super GPU를 사용한다. 학습에서 모델은 YOLOv8중 x모델을 사용한다. 학습 설정으로 에포크 100, 배치 크기 16, 그리고 학습률을 0.01로 지정한다. 이미지의 증강에는 좌우반전과 대비조절을 사용한다. 이미지 입력 크기는 YOLOv8에서 원래 정사각형을 사용하지만, 모델에서 제공하는 하이퍼 파라미터인 rect설정을 이용해 입력 사이즈를 직사각형인 1024×128으로 고정한다. 비교 실험은 BiFPN 구조가 탐지 정확도에 미치는 영향을 분석하기 위해, BiFPN 모듈의 반복 횟수를 달리한 실험을 수행한다. YOLOv8x 기본 모델과, 본 연구에서 제안한 BiFPN 레이어를 각각 2회, 3회, 4회로 쌓은 모델을 구성하여 비교한다. 성능의 평가는 mAP50, mAP50-95, Precision(P), Recall(R)를 기준으로 한다. 실험 결과는 다음과 같다.

표1. BiFPN 레이어에 따른 성능 비교

Model	mAP50	mAP50-95	P	R
v8x	0.78	0.46	0.74	**0.78**
v8x+(BiFPN layer 2)	**0.79**	**0.50**	**0.76**	0.74
v8x+(BiFPN layer 3)	0.74	0.44	0.75	0.72
v8x+(BiFPN layer 4)	0.72	0.45	0.70	0.68

BiFPN 레이어를 2번 적용한 모델이 mAP50, mAP50-95, Precision에서 가장 높은 점수를 기록한다. 이것은 해당 모델의 정확도가 실험에서 사용한 모델 중 가장 높다는 것을 의미한다. 기존의 YOLOv8x 모델에서는 Recall 값이 가장 높아 물체를 가장 잘 검출하지만, 정확도는 BiFPN을 2번 사용한 모델에 비해 떨어진다. BiFPN을 3번, 4번 사용한 모델은 모든 지표가 앞의 모델들보다 떨어진다. 이것은 BiFPN에서 레이어를 추가 했을 때 표현력은 늘어나지만, 과적합 또는 스케일의 불균형이 커질 수 있음을 시사한다. 또한 모델의 실제 추론시간은 기존의 모델은 평균 13ms, BiFPN 레이어를 1번 적용할 때마다 약 2ms씩 증가한다. LiDAR에서 생성되는 이미지가 10fps이기 때문에, 실험에 사용한 모든 모델은 실시간 객체 탐지에 충분히 쓰일 수 있다. 따라서, BiFPN 레이어를 2번으로 제한했을 때 기존의 YOLOv8x 대비 정확도를 개선하였으며, 추가로 레이어를 늘리는 것은 오히려 성능 저하로 이어질 수 있다. 따라서 본 연구의 최종 모델은 YOLOv8x + BiFPN 레이어를 2번 적용한 모델이 가장 합리적이라는 결론에 도달한다.

Ⅲ. 결론

본 연구는 YOLOv8 기반의 LiDAR 2D 객체 탐지에서 BiFPN과 가중합 기반 연산을 결합한 구조를 제안하고 실험적으로 검증하였다. BiFPN을 적절한 깊이로 활용하면 다중 해상도 특징 맵 간의 효과적인 정보 융합이 가능하여 소형 객체 탐지 및 탐지 정확도를 실질적으로 개선할 수 있음을 확인한다. 향후에는 더욱 다양한 도로 환경의 데이터를 활용하여 연구의 일반화 가능성을 검증하고, 추가적인 모델 최적화 및 경량화를 통해 좀 더 범용적이고 효율적인 실시간 객체 탐지 시스템을 개발할 계획이다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부와 한국산업기술진흥원의 지역혁신클러스터 육성 (R&D, P0025274 사업의 지원을 받아 수행된 연구결과임

참 고 문 헌

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, real-time object detection,” in Proc. IEEE CVPR, pp. 779-788, 2016.

[2] Ultralytics, “YOLOv8 Documentation,”

https://docs.ultralytics.com, 2023.

[3] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition pp. 10781-10790