

단일 깊이 센서를 이용하는 자율주행 시나리오 기반의 깊이 데이터 객체 감지

*김명균, 정진우, 김성제

*한국전자기술연구원

*{foulpg99, jw.jeong, sungjei.kim}@keti.re.kr

Depth data object detection based on autonomous driving scenario using a single depth sensor

*Myeong-kyun Kim, Jinwoo Jeong, Sungjei Kim

*Korea Electronics Technology Institute (KETI)

요 약

본 논문에서는 단일 깊이 센서를 사용하는 자율주행 시나리오에서 거리 계산에 주로 이용되는 깊이 데이터만 이용하는 객체 감지 기술을 제안한다. 우선, 해당 시나리오에서 객체 감지 학습 데이터는 깊이 데이터가 사용되지만 상대적으로 객체 감지 성능을 비교하기 위해 동일한 시간의 색상, 깊이 데이터를 함께 획득하여 학습에 이용한다. 학습 모델은 객체 감지 분야에서 최근 주목 받고 있는 YOLOv5로 선정하여 색상, 깊이 데이터의 객체 감지 학습의 결과를 각각 확인하였다. 결과적으로 색상과 깊이 데이터 사이에서 객체 감지 학습 결과의 차이를 확인하며 본 논문에서 제안하는 자율주행 시나리오에 깊이 영상만 이용하는 객체 감지 기술의 문제점과 향후 자율주행 기술 발전에 기여 가능성을 확인할 수 있다.

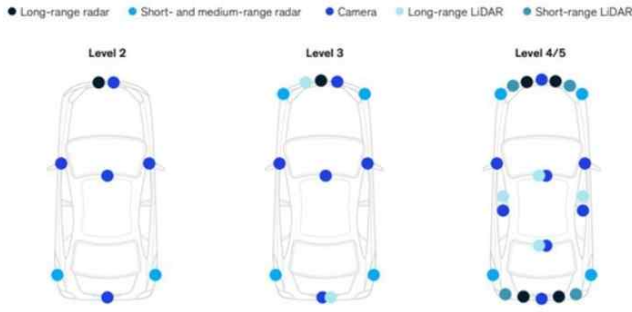
1. 서론

최근 딥러닝은 인공지능망을 토대로 GPU 등 하드웨어 기술의 발전과 인공지능 기술을 통해 기존 신경망 학습 과정에서 발생하는 여러 문제를 해결하기 위해 연구되고 있다. 기존 연구들의 성과로 다양한 딥러닝 모델이 구현되고, 스마트 시티, 자율주행 자동차, CCTV 모니터링 등 다양한 분야에서 놀라운 성능을 보이는데 특히 컴퓨터 비전 분야에서 좋은 성과를 보여주며 기존의 방식들을 대체할 수 있는 새로운 대안으로 여겨지고 있다.

컴퓨터 비전 분야에서 객체 감지(Object detection) 기술은 이미지 또는 비디오 상의 객체를 식별하는 기술인데, 딥러닝 네트워크를 기반으로 객체 인식 및 검출을 목적으로 가진 뛰어난 성능의 다양한 알고리즘들이 발표되고 있다. 객체 감지를 목적으로 하는 대표적인 딥러닝 알고리즘으로는 YOLO(You Only Look Once)[1]가 있다. YOLO는 단일 단계 방식의 객체 탐지 알고리즘

으로 이미지 내에 존재하는 객체와 해당 객체의 위치를 한번의 입력 이미지로 예측할 수 있지만 YOLO 알고리즘으로 검출한 객체의 위치는 정확히 말하면 입력 이미지 내의 위치를 나타내기 때문에 실제 영상이 촬영되는 카메라의 위치와 객체 사이의 실제 거리는 차이가 있다.

객체 감지와 실제 거리 계산 정보는 자율주행 시나리오에서 차선, 차량, 보행자 인식 등 안전과 관련된 중요한 정보다. 현재 자율주행 시나리오는 차량에 장착된 각종 센서를 통해 외부 정보를 감지하여 차량 스스로 주행제어를 수행하는 첨단 운전자 지원 시스템인 ADAS(Advanced Driver Assistance System)의 형태로 구현되는데, ADAS에 사용되는 기술 중 많은 응용 기술들은 자율주행 실현을 위해 실제 거리 측정 정보를 중요하게 이용한다. 또한 자율주행 시나리오의 레벨(Level)이 발전되면서 거리 측정 정보가 중요해짐에 따라 거리 측정 가능한 센서의 사용 비중이 그림 1과 같이 높아지고 있다.



<그림 1 - 자율 주행 레벨에 따른 센서 수>

Image count	총 18729 개 (train-16856, validation-1873)
Resolution	1920 x 1080
Data format	PNG (Color: RGB 24bit, Depth: GrayScale16bit)
Class	총 8 개 (person, bicycle, car, motorcycle, bus, truck, motorcycle_person, bicycle_person)

<표 1 - 데이터 세트 구성>

본 논문에서는 단일 깊이(Depth) 센서를 사용하는 자율주행 시나리오를 목표로 정확한 거리 측정에 용이한 깊이 데이터만 사용하여 객체 감지할 수 있는 기술을 제안한다. 깊이 데이터만 이용하는 객체 감지 기술의 딥러닝 알고리즘은 YOLOv5를 사용하고, 학습 데이터는 색상 데이터와 깊이 데이터의 객체 감지 성능을 비교하기 위해 ZED 카메라[2]를 차량에 부착하여 직접 촬영한 데이터 세트(Custom Dataset)로 구성한다. 직접 촬영한 데이터 세트는 색상, 깊이 모두 동일한 시간 정보를 쌍으로 포함하고, 이 데이터 세트를 이용한 객체 감지 학습 결과를 비교하여 깊이 데이터만 이용한 객체 감지 기술 성능을 확인할 수 있다.

본 논문의 구성은 다음과 같다. 2절에서 본 논문에서 사용한 데이터 세트 구성을 설명하면서 YOLOv5를 객체 감지 모델로 사용한 이유와 학습 방법에 대해 알려준다. 3절은 깊이 데이터만 이용하는 객체 감지 성능을 색상 데이터의 객체 감지 성능과 비교하여 정량적 결과로 보여주고, 4절에서 본 논문에 대한 결론을 맺는다.

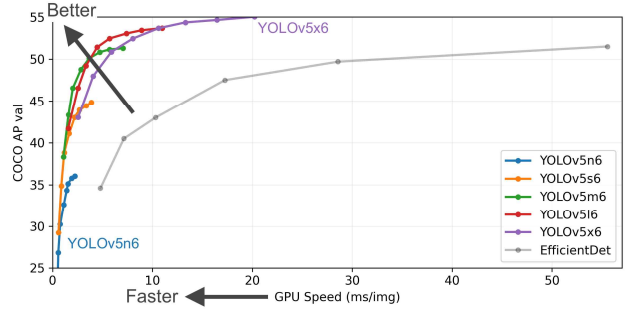
2. 본론

2-1. 데이터 세트 구성

딥러닝을 기반한 객체 인식 및 거리 계산은 데이터를 기반한 학습으로 진행되기 때문에 학습 데이터가 매우 중요하다. 본 논문에서는 깊이 데이터만 사용하는 객체 감지 성능을 확인하기 위해 ZED 카메라 이용하여 색상 데이터와 깊이 데이터를 동시에 획득한 데이터 세트를 구축하여 성능을 비교했다.

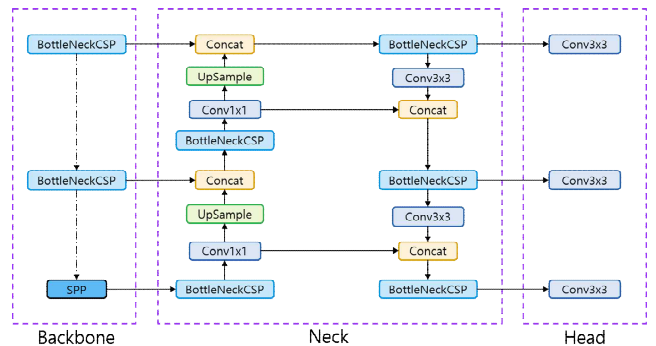
직접 획득한 학습 데이터 세트는 표 1과 같이 구성된다. 촬영 데이터 중 객체 감지에 의미 있다고 판단되는 깊이 데이터 18729장을 선별하고, YOLOv5 학습에 필요한 train, validation 데이터 비율을 9:1로 나누어 학습에 사용하였다. 데이터의 해상도와 포맷은 ZED 카메라에서 지원하는 1080p 해상도와 PNG파일로 RGB 24bit(색상 데이터), GrayScale 16bit(깊이 데이터) 형태로 구성되고, 자율 주행 시 인식되는 객체의 클래스(Class)은 사람, 자전거, 자동차, 오토바이 등 총 8개로 구성했다.

2-2. 객체 감지 모델 학습



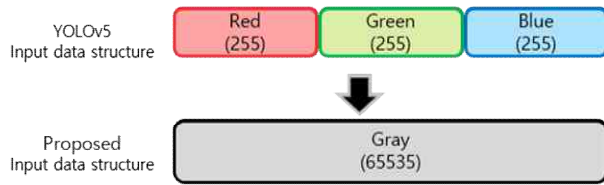
<그림 2 - 객체 감지 모델 성능 비교>

차량 운행 및 자율주행 시나리오에서 객체 감지 기술은 안전과 직결되기 때문에 처리 시간이 중요하다. 객체 감지 모델 중 최근 주목 받고 있는 모델은 YOLO의 최근 버전인 YOLOv5[4]와 EfficientDet[3]가 있다. 그림 2는 COCO 데이터 세트[5]를 이용하여 시간 축(X축) 대비 성능이 좋은(Y축) 두가지 객체 감지 모델의 성능을 보여준다. 두가지 모델 중 본 논문의 객체 감지 모델을 선정하기 위해 객관적인 성능 비교 자료로 그림 2를 참고하였다. 그림 2에서 YOLOv5의 처리 속도가 EfficientDet보다 빠르기 때문에 본 논문 시나리오의 학습 모델로 YOLOv5가 적합하다고 판단하여 사용했다.



<그림 3 - YOLOv5 구조>

YOLOv5의 구조는 그림 3과 같다. 본 논문에서 제안하는 객체 감지 기술은 YOLOv5의 네트워크 구조를 동일하게 이용하면서 입력 데이터를 깊이 데이터에 맞는 구조로 변경하여 학습한다.



<그림 4 - YOLOv5 입력 데이터 구조>

YOLOv5 모델에서 입력으로 사용하는 색상 데이터 구조는 기본적으로 RGB 3채널 이미지를 사용하지만, 본 논문에서 구축한 깊이 데이터를 입력 데이터로 사용하기 위해 그림 4와 같이 Gray Scale 1채널로 변경하여 학습한다.

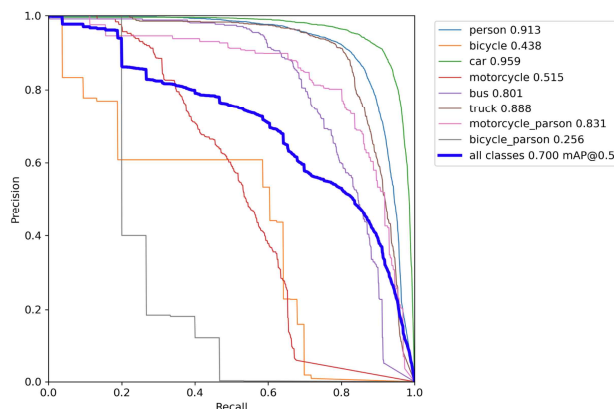
2-3. 학습 환경

깊이 데이터를 사용하는 객체 감지 학습을 위한 하드웨어 구성은 표 2와 같다. OS는 장기지원 버전인 Ubuntu 20.04 LTS를 사용하고, GPU는 학습의 속도를 높이기 위해 Nvidia TITAN RTX를 2개로 사용하여 병렬 학습을 진행했다.

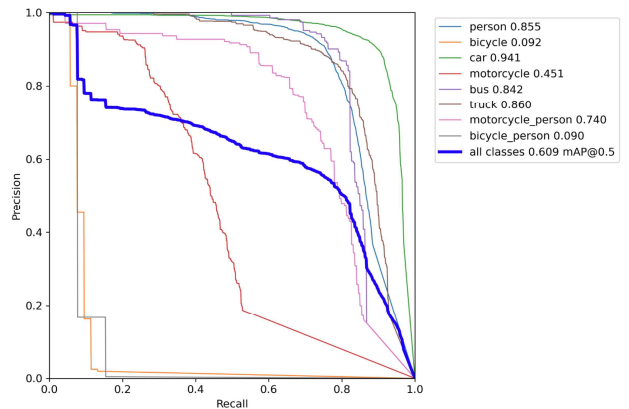
OS	Ubuntu 20.04 LTS
CPU	Intel(R) Core(TM) i9-7960X CPU @ 2.80GHz
RAM	64GB
GPU	NVIDIA TITAN RTX (x2)

<표 2 - 하드웨어 구성>

3. 결과



<그림 5 - 색상(Color) 영상 학습 결과 그래프>



<그림 6 - 깊이(Depth) 영상 학습 결과 그래프>

그림 5, 6는 본 논문에서 구축한 영상으로 학습된 정량적 결과를 정리한 그래프다. 그래프를 통해 객체 감지의 평가지표가 되는 mAP의 면적을 확인할 수 있다. 깊이 영상으로 학습한 전체 클래스의 mAP는 0.609로 색상 영상으로 학습한 mAP 0.700과 0.091의 근소한 차이를 확인할 수 있다. 공통적으로 클래스 중 bicycle_person, bicycle, motorcycle은 mAP의 면적보다 적은 값으로 객체 감지 학습이 어려운 것을 확인할 수 있다.

4. 결론

본 논문에서 제안한 깊이 데이터만 이용하는 객체 감지 기술의 학습 성능은 색상 데이터의 객체 감지 성능과 비교하여 근소한 차이를 확인할 수 있었다. 그래서 깊이 데이터만 사용하여 객체 감지 기술을 자율주행에 적용하더라도 색상 데이터를 사용한 객체 감지 성능과 비슷할 것으로 예상된다. 하지만 3절에서 설명한 것처럼 본 논문의 색상과 깊이 데이터 모두 특정 클래스는 객체 감지가 어려운 점을 확인했고, 객체 감지 모델의 성능이 발전되면서 해당 클래스도 잘 검출되는 객체 감지 기술을 개발한다면 좀더 안정적인 단일 깊이 센서를 이용한 객체 감지 기술이 완성될 것으로 전망한다.

Acknowledgement

본 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2021-0-00802, 속성을 유지하는 지능적 미디어 화면비 변환 기술 개발, 기여율)

참고 문헌

- [1] YOLO (You only look once): Real-Time Object Detection, <https://pjreddie.com/darknet/yolo/>
- [2] ZED 2 camera: <https://www.stereolabs.com/zed-2/>
- [3] Mingxing Tan Ruoming Pang Quoc V. Le, "EfficientDet: Scalable and Efficient Object Detection", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 10781-10790
- [4] YOLOv5: <https://github.com/ultralytics/yolov5>
- [5] COCO dataset: <https://cocodataset.org/#download>