

패턴인식 보고서

120220210 고재현

2022년 12월 19일

1 개요

1.1 목적

- 최근 출시되는 자율주행 자동차에 포함된 pattern recognition/computer vision 관련 요소기술을 조사한다.
- 해당 요소기술과 수업에서 다룬 주제들 간의 연관성을 파악한다.

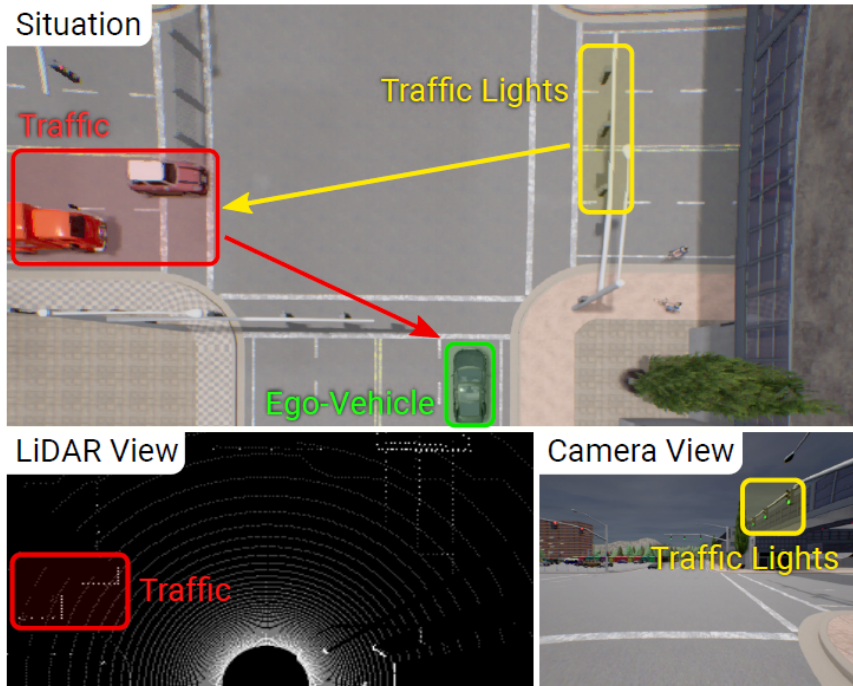
1.2 선정 모델 및 브랜드

선정한 모델은 **Tesla Model S** 이다. 테슬라가 오토파일럿 기능을 앞세워 자율주행 시장을 선도하고 있기 때문이다. 해당 모델에 적용된 pattern recognition/computer vision 관련 요소 기술은 다음과 같다.

- **Autopilot**: 자율주행 기능
- **Autopark**: 주차 자동화 기능
- **Autosteer**: 자동 조향 기능

1.3 선정 논문

[1, Multi-modal fusion transformer for end-to-end autonomous driving] 및 [2, YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors] 논문을 선정하였다. 첫 번째 논문을 선정하게 된 이유는 자율주행 시스템의 구성요소 중 하나인 perception에 대한 연구이면서, 필자의 연구 분야인 multi-modal fusion에 대한 연구이기 때문이다. 두 번째 논문을 선정하게 된 이유는 자율주행 시스템의 구성요소 중 하나인 perception에 대한 연구이면서, 영상처리의 주요한 분야 중 하나인 object detection에 대한 연구이기 때문이다.

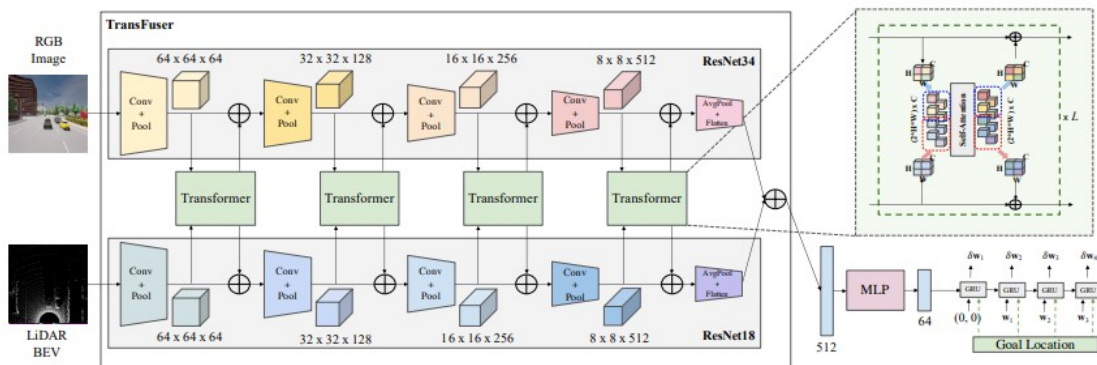


〈그림 1〉 논문에서 해결하려는 문제 상황

2 논문 요약

2.1 Multi-modal fusion transformer for end-to-end autonomous driving

이 논문은 그림 1의 상황처럼 라이다(LiDAR : Light Detection And Ranging) 센서로 얻을 수 있는 주변의 차량의 위치에 따른 교통정보와 카메라로 얻을 수 있는 신호기에 따른 교통정보가 다른 경우, 두 센서로부터 얻을 수 있는 정보를 결합하여 차량의 주행을 제어하는 것을 목적으로 한다. 그림 2는 Transfuser의 구조를 보여준다. 두 센서의 출력으로부터 Resnet 구조



〈그림 2〉 Transfuser 구조

[3]를 이용하여 정보를 추출하는 과정에서, 각 layer의 출력단으로부터 추출된 정보를 Transformer[4]를 이용하여 결합하는 것을 확인할 수 있다.

2.1.1 method

논문에서 제시한 목표는 시내 도로 주행에서의 point-to-point navigation이다. point-to-point navigation은 차량이 목표지점까지 waypoint를 따라 교통법규를 지키면서 다른 차량과의 상호작용을 하며 완주하는 것을 의미한다.

이를 달성하기 위한 방법으로 강화학습 기법 중 하나인 Imitation Learning을 채용하였다. Imitation Learning은 전문가가 직접 주행한 데이터를 따라하도록 agent의 policy를 학습하는 것을 의미한다. 데이터셋은 자율주행 오픈소스 시뮬레이터 CARLA[5]에 있는 urban 가상환경에서 수집했다. 데이터의 왜곡을 줄이기 위해 이미지 입력의 중앙을 잘라내어 256X256X3 크기로 사용했다. LiDAR 센서의 출력 또한 주변부분의 정보를 기반으로 256X256X2 사이즈로 잘라내어 사용하였다. 채널의 한쪽은 지면 위, 한쪽은 지면 아래를 의미한다. 출력은 PID controller로 차량을 제어하기 위해 4개의 waypoint $\{w_t = (x_t, y_t)\}_{t=1}^T$ 로 설정했다.

2.2 YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors

asdf

참고 문헌

- [1] A. Prakash, K. Chitta, and A. Geiger, “Multi-modal fusion transformer for end-to-end autonomous driving,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7073–7083, 2021.
- [2] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, “Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” 2022.
- [3] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *CoRR*, vol. abs/1512.03385, 2015.
- [4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” *CoRR*, vol. abs/1706.03762, 2017.
- [5] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, “CARLA: An open urban driving simulator,” in *Proceedings of the 1st Annual Conference on Robot Learning*, pp. 1–16, 2017.