

S-BEV 이미지를 이용한 효율적인 자율 주행

김정훈, 조상재, 공승현*
한국과학기술원

jeonghoon.kim@kaist.ac.kr, *skong@kaist.ac.kr

Efficient Autonomous Driving using S-BEV image

Jeong-Hoon Kim, Sangjae Cho, Seung-Hyun Kong*
Korea Advanced Institute of Science and Technology

요약

종단간(End-to-end) 자율주행 시스템이 입력 받는 센서의 원 측정치는 주행에 불필요한 고차원의 특징(feature)들을 내포하고 있어 도심지와 같은 복잡한 환경에서 네트워크 입력으로써의 효율성이 떨어진다. 이러한 문제를 완화시키기 위해, 우리는 전역 경로를 포함한 고정밀 지도(HDmap), 카메라 이미지와 HDmap 을 기반으로 생성된 Shadow Bird's Eye View (S-BEV), 그리고 속도 정보를 입력 받아 지역 경로를 추정하는 Efficient Autonomous Driving Network(EADNet)을 제안한다. EADNet 은 S-BEV 를 통해 주변 객체를 인지하고 HD map 으로부터 도로 상황을 이해한 후 가장 안전하고 효율적인 경로를 생성한다. 우리는 시뮬레이션 결과를 통해 EADNet 의 우수한 주행 안전성과 효율성을 입증한다.

I. 서론

자율 주행은 하나의 간단한 기술이 아닌, 여러가지 복잡한 모듈이 유기적으로 동작하도록 구현되곤 하는데, 이를 모듈러 방식 자율주행 기술이라고 한다. 해당 기술은 일반적으로 인식, 판단, 제어 모듈로 이루어져 있고, 각 모듈의 역할이 제한적이어서 도심의 도로와 같이 복잡한 상황에서는 안정적인 자율주행에 한계가 있다.

이러한 문제를 해결하고자, 기존 여러 모듈을 하나의 인공지능 네트워크로 대체하는 종단간(End-to-end) 방식의 자율주행 기술이 등장했다[1-3]. 그 중에서도 전문가의 행동데이터를 지도 학습하여 그 행동을 모방하도록 하는 모방학습 기반 종단간 자율주행이 활발히 연구되고 있다. 하지만, 모방학습은 학습 데이터에 매우 의존적이고, 규칙 기반의 모듈러 방식 자율주행 기술에 비해 인공지능의 출력을 예측하기 어렵다는 문제가 있다.

본 논문에서는 종단간 방식과 모듈러 방식을 융합한 EADNet(Efficient Autonomous Driving Network)을 제안한다. 제안하는 기술은 인공지능 기반의 인지-판단 통합 모듈로 구성되며, EADNet 의 출력을 PID 제어기의 입력으로 사용하여 자율주행 시스템을 구성한다. EADNet 은 차량 주변 상황을 표현하는 Shadow Bird's Eye View (S-BEV) 이미지, 전역 경로를 포함하는 고정밀 지도(HDmap)를 표현하는 이미지(HD+G), 그리고 차량의 속도를 입력으로 받아 앞으로 이동할 지역 경로를 출력한다. 이후, PID 제어기는 해당 목적지를 따라갈 수 있도록 적합한 가속값과 조향각을 차량에 전달한다.

본 논문의 구성은 다음과 같다. 2 장에서는 EADNet 에 대한 자세한 소개와 시뮬레이션 실험 결과를 보여준다. 3 장은 본 논문의 결론이다.

II. 본론

A. 실험 환경

우리는 안전한 실험과 다양한 데이터 취득을 위하여 CARLA 시뮬레이터를 사용한다. 데이터셋은 사람이 로지텍 G29 레이싱휠을 사용하여 취득한 가속값 및 조향각, 그리고 Fig. 1 과 같은 다양한 BEV 이미지를 포함한다.

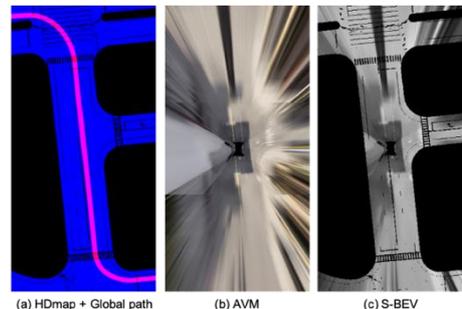


Fig. 1. 다양한 정보를 표현하는 BEV 이미지

사람이 실제 운전을 하면서 다음과 같은 주행 정책을 준수하였다.

1. 전역 경로 위에 장애물이 없을 경우
 - ① 전역 경로를 따라 목표 속도에 맞춰 주행한다.
2. 전역 경로 위에 장애물이 있을 경우
 - I. 차선 변경이 가능한 경우
 - ② 차선을 변경하되 목표 속도에 맞춰 전역 경로와 같은 방향으로 주행한다.

II. 차선 변경이 가능하지 않은 경우

- ③ 전역 경로를 따라 앞차와의 간격을 유지하며 주행한다.

이러한 주행 정책을 EADNet 에 학습시키기 위해서, 우리는 데이터 셋의 각 프레임을 HD+G, Shadow-Bird Eye View (S-BEV), 그리고 차량의 속도로 구성했다. 우리는 자차의 전, 후, 좌, 우를 향하는 카메라의 이미지로부터 생성되는 BEV 이미지를 HDmap 에서 표현되는 도로 영역으로 필터링하여 단채널의 Shadow-Bird Eye View (S-BEV) 이미지를 생성한다 (Fig.1.c 참고). S-BEV 데이터셋은 특정 프레임의 시점을 t , 그때의 위치를 S_t 라고 한다면 $S \in (S_t, \dots, S_{t+T-1})$ 를 정답 라벨로 제공한다. 해당 라벨은 회귀 대상이 되는 지역 경로로써 해당 시점을 포함한 T 개의 목적지로 구성된다.

B. 학습

제안하는 EADNet 은 Fig.2 와 같이 Fig.1.a, c 를 3 채널로 합쳐서 표현한 이미지를 입력으로 받는다. 해당 이미지는 이미지 인코더로 인해 크기가 500 인 벡터로 인코딩 되고 동시에 속도 정보도 fully connected layer 를 거쳐 동일한 크기의 벡터로 변환된다. 이 두 벡터는 크기가 1000 인 벡터로 결합된 후 Multi-layer Perceptron 에 통과되어 크기가 20 인 벡터 \hat{S} 로 출력된다. 활성화 함수는 Gaussain Error Linear Unit(GELU)를 사용하며, 네트워크의 출력인 \hat{S} 과 정답 라벨 S 간의 Mean Square Error(MSE)가 최소가 되도록 학습시킨다. 학습 시 최적화 함수는 Adam 을 사용하였다.

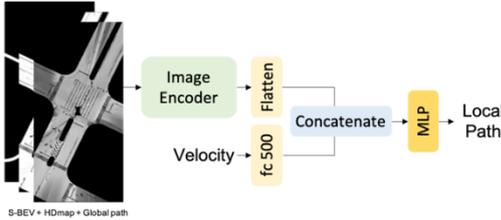


Fig. 2. 제안하는 EADNet 구조

C. 평가

이번 절에서는 시뮬레이션 결과로 EADNet 의 우수성을 보여준다.

Fig.3.a 와 b 는 학습된 EADNet 의 출력인 지역 경로를 S-BEV 와 HD+G 로 시각화 한 결과이다. Fig.3.c 는 시뮬레이터에서 정확히 얻을 수 있는 주변 차량 위치 정보를 HD+G 로 시각화 한 것이다. Fig.3 은 자차 전방에 멈춰 있는 다른 차량이 있는 주행 상황을 보여주고 있다.

Fig.3 에서 전방의 차량을 회피하기 위해 전역경로로부터 벗어나는 지역경로를 생성하는 것을 확인할 수 있다. 이는 S-BEV 로 부터 전방 차량을 인지했고, 네트워크가 해당 상황에서의 최적의 경로를 생성했다고 볼 수 있다.

이러한 결과는 EADNet 이 기존 중단간 주행방식과 다르게 Interpretability 가 우수함을 보여준다. 그리고, EADNet 은 HD+G 에서 보이는 전역 경로만을 따라가지 않고, S-BEV 이미지로부터 장애물을 인식한 후 도로 상황에 적절하게 주행하는 능력이 있음을 보여준다.

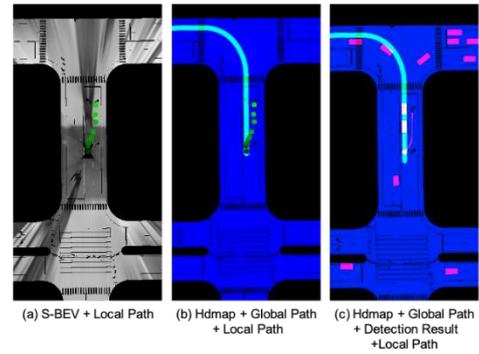


Fig. 3. EADNet 출력 시각화

III. 결론

우리는 모듈형 자율주행과 중단간 자율 주행을 융합한 EADNet 을 제시하였다. EADNet 은 S-BEV 를 통해 주변 객체에 대한 정보를 학습할 수 있고 제어값이 아닌 지역 경로를 출력하여 중단간 주행 방식의 Interpretability 를 제고하였다. 또한, 주변 도로 상황에 따른 최적의 주행 성능을 보여주었다.

추후 연구로 시뮬레이터와 실제 도로환경간 도메인을 일반화 시켜 Simulation-to-real 적용하는 기술을 고려해 볼 수 있다.

ACKNOWLEDGMENT

이 논문은 2021 년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1A2C300837011, 예상 못한 상황에서 빠르게 적응하는 메타-강화학습기반 End-to-end 자율주행 기술 개발).

참고 문헌

- [1] Q. Wang, L. Chen, B. Tian, W. Tian, L. Li, and D. Cao, "End-to-end autonomous driving: An angle branched network approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 11599-11610, Dec. 2019.
- [2] Y. Xiao, F. Codevilla, A. Gurram, O. Urfalioglu, and A. M. López, "Multimodal end-to-end autonomous driving," 2019, arXiv:1906.03199.
- [3] Codevilla, F. "End-to-end driving via conditional imitation learning," *IEEE International Conference on Robotics and Automation (ICRA)*, 2018.
- [4] Tolstikhin, Ilya O., et al. "Mlp-mixer: An all-mlp architecture for vision." *Advances in Neural Information Processing Systems 34* (2021): 24261-24272.