

고정밀 지도를 이용한 모방 학습 기반 종단간 자율 주행

한승태, 조상재, 정명근, 김태선, 공승현*
한국과학기술원

seungtaeh@kaist.ac.kr, *skong@kaist.ac.kr

Imitation Learning-based End-to-end Autonomous Driving with High Definition Map

Seung-Tae Han, Sangjae Cho, Myung-Keun Jeong, Taeseon Kim, Seung-Hyun Kong*
Korea Advanced Institute of Science and Technology

요 약

참조 정책을 모방하도록 인공지능 네트워크를 학습시키는 모방 학습(Imitation Learning) 기반 종단간(End-to-end) 자율 주행 기술은 동일한 상태에서 취득한 관측치에 대해 사람에 따라 다른 행동을 취할 수 있다는 의사 결정 다양성 문제가 있다. 이러한 문제를 해결하기 위해, 본 논문에서는 차량 전방 이미지와 고정밀 지도 이미지로부터 차량의 조향각을 매핑하는 MIMSNet 을 제안한다. 그리고 직접 제작한 데이터셋을 사용하여 MIMSNet 과 기존 연구에서 제시된 네트워크를 학습 및 평가하고, 그 결과를 비교하여 MIMSNet 의 우월한 조향각 추정 성능을 입증한다.

I. 서 론

자율 주행 기술의 대표적 패러다임으로는 모듈러 방식과 종단간(End-to-end) 방식이 있다. 모듈러 방식은 사람이 정의한 논리에 따라 인식, 판단, 그리고 제어 모듈 등으로 구성된다. 이러한 방식은 복잡한 주행 상황에서의 성능에 한계가 있으며, 높은 비용이 요구되어 효율적이지 않다.

반면, 종단간 방식은 인식에서부터 제어까지의 모든 과정을 단일 알고리즘으로 처리하기 때문에, 모듈러 방식의 단점을 극복할 수 있다. 최근 인공지능의 비약적 발전에 따라, 모방 학습(Imitation Learning) 기반 종단간 방식에 대한 연구가 활발히 진행되고 있다. 모방 학습은 사람이 직접 운전하는 상태에서 취득한 관측치와 그때 사람이 결정한 행동을 참조하여, 인공지능이 사람의 주행 정책을 스스로 모방하도록 학습하는 방식이다. 이러한 방식은 대표적으로 [1-4]에서 제시되었다.

[1]는 차량의 전방 이미지만을 조향각으로 매핑하는 기술을 제시하였고, 차선 유지 정도만을 실증하였다. 그러나, 해당 기술은 동일한 상태의 관측치에 대해 사람에 따라 서로 다른 행동을 취할 수 있다는 의사 결정 다양성 문제가 있다. 이를 해결하기 위해, [2, 3]는 방향 지시등 정보를 네트워크에 입력함으로써, 선회 방향 결정이 가능한 기술을 제시하였다. 하지만, 이 기술들은 데이터셋 구성 및 실증 시 방향 지시등을 수동으로 조작해야 한다는 한계가 있다. [4]는 단순화된 지도 이미지를 네트워크에 입력하는 기술들을 제시하였지만, 도로를 하나의 차선으로만 도시한 지도를 활용하기 때문에, 교차로 회전을 위한 차선

변경 등 다양한 환경에서의 적절한 행동을 출력하기에 한계가 있다. 이러한 한계는 네트워크에 도로의 지형 정보 및 전역 경로 정보를 반영하여 해결할 수 있다.

본 논문에서는 차량 전방 이미지와 고정밀 지도 이미지, 그리고 차량의 속도로부터 조향각을 출력하는 MIMSNet(Mapping from Image and Map to Steering Network)을 제안한다. 또한 MIMSNet 및 기존 연구에서 제시한 네트워크를 비교하여 MIMSNet 의 우월성을 입증한다.

II. 본론

A. 데이터셋

본 논문에서는 한국과학기술원 내의 도로에서 사람이 직접 운전하며 측정한 차량 전방 이미지 i_f , 속도 v , GPS 정보, heading 과 그때 사람이 결정한 조향각 δ 을 기록하여 데이터셋을 구성하였다. 이때, 데이터들은 맑은 날, 흐린 날, 비가 오는 날, 그리고 밤의 다양한 주행 환경에서 취득하였다.

i_f 는 해상도 640×480 의 원본 RGB 이미지를 관심 영역에 대해 크로핑(Cropping) 및 흑백 변환하여, 64×200 로 해상도를 조정하여 녹화하였다. 고정밀 지도 이미지 i_m 은 전역 경로 정보를 포함한 고정밀 지도를 차량의 heading 에 따라 회전시키고, GPS 측정치를 기준으로 전방 50m, 측방 각 25m 씩 관심 영역을 설정하여 크로핑한 후, 100×100 RGB 이미지로 변환하여 출력 및 취득하였다. 우리는 각 데이터를 15Hz 로 다운 샘플링 및 시간 동기화하여 데이터셋

$D = \left\{ \langle i_{f,t}, i_{m,t}, v_t, \delta_{r,t} \rangle \right\}_{t=1}^N$ 를 구성하였고, 총 데이터의 개수 N 은 100,000이다. 학습 및 평가 데이터셋은 각각 다른 주행 경로에서 취득한 데이터로 구성하였고, 그 비율은 9:1로 하였다.

B. 학습

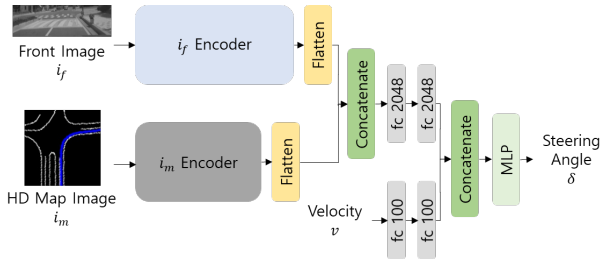


Fig. 1. 제안하는 MIMSNNet 구조

본 논문에서 제시하는 MIMSNNet은 i_f , i_m , v 로부터 $\delta = f(i_f, i_m, v)$ 를 매핑한다. Fig. 1과 같이 MIMSNNet은 i_f 를 수정된(Modified) VGGNet[5]을 사용하여 $128 \times 4 \times 12$ 크기로 인코딩하고, 인코딩된 특징 맵(feature map)을 최종적으로 벡터화한다. i_m 은 i_f 비해 적은 특징들을 포함하기 때문에, 비교적 얇은 인공 신경망을 거쳐 $48 \times 4 \times 4$ 로 인코딩된 뒤, 벡터화된다. v 는 두 개의 fully connected layer를 거쳐 100개의 뉴런으로 인코딩된다. 출력되는 각 벡터들은 병합되어 Multilayer Perceptron(MLP)를 통해 조향각으로 회귀(Regression)된다. 이때, MIMSNNet의 마지막 Fully Connected Layer를 제외한 모든 layer 이후에는 활성화 함수 Rectified Linear Unit(ReLU)를 적용한다.

MIMSNNet은 식 (1)과 같이 δ 와 δ_r 간의 Root Mean Square Error(RMSE)가 최소가 되도록 네트워크 파라미터 θ 를 업데이트하며 학습한다.

$$\underset{\theta}{\text{minimize}} \text{RMSE}(f(i_f, i_m, v; \theta), \delta_r) \quad (1)$$

Minibatch는 16개이고, θ 는 learning rate가 0.0001인 조건에서 Adam을 통해 최적화된다. 본 논문에서는 총 50 에포크의 과정을 NVIDIA GeForce GTX 1080 Ti GPU를 사용하여 수행하였다.

C. 평가

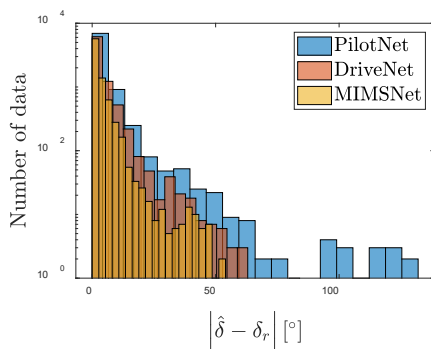


Fig. 2. PilotNet, DriveNet, 그리고 MIMSNNet의 평가 결과 오차 히스토그램

이 절에서는 학습된 MIMSNNet, [1]에서 제시한 PilotNet, 그리고 [2]에서 제시한 DriveNet을 평가 및 그 성능들을 비교한다. 우리는 공정한 비교를 위해 PilotNet을 MIMSNNet의 i_f 인코더에 MLP를 추가한 것으로 수정하였고, DriveNet은 수정된 PilotNet의

두번째 fully connected layer 출력 뉴런에 방향 지시등 정보를 병합하는 구조로 대체하였다. 방향 지시등 정보는 [2]에 제시된 바와 같이 100 배 스케일링하여 입력하였다. 우리는 MIMSNNet과 기존 2 가지 기술에 대해 참조 조향값과 평가 결과 조향값 간 차이의 절대값을 계산하고, 각 값의 범위에 해당하는 데이터의 개수를 나타내는 히스토그램을 Fig. 2에 도시하였다.

Fig. 2의 성능 결과를 통해 MIMSNNet은 고정밀 이미지의 특징을 조향값 매핑에 반영함으로써, PilotNet이 내포하고 있는 의사 결정 다양성 문제를 해결함을 알 수 있다. 또한, 히스토그램 상 MIMSNNet의 출력 오차가 DriveNet에 비해 낮은 구간에 분포해 있다는 점에 대해, MIMSNNet이 고정밀 지도의 자세한 도로 지형 정보 및 전역 경로의 상대적인 곡률로부터 적절한 조향값을 추정할 수 있다는 것을 증명할 수 있다. Fig. 2에는 도시하지 않았지만, MIMSNNet은 교차로 회전을 위한 차선 변경과 같은 다양한 환경에서의 적절한 행동을 출력함을 실증을 통해 입증할 수 있었다.

III. 결론

본 논문에서는 모방 학습 기반 종단간 자율 주행의 의사 결정 다양성 문제를 해결하기 위해 MIMSNNet을 제시하였다. 또한 직접 제작한 데이터셋을 사용하여 MIMSNNet과 기존 네트워크의 조향각 추정 성능을 비교하였고, 이를 통해 MIMSNNet의 우월성을 입증하였다. 본 연구에서는 한국과학기술원 내의 주행 데이터만을 사용하였지만, 미래에는 학습 데이터의 다양성을 제고하여 변화하는 환경에 강인한 자율 주행이 가능할 것으로 예상된다.

ACKNOWLEDGMENT

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

참고 문헌

- [1] Bojarski, M. "End to end learning for self-driving cars," arXiv preprint arXiv:1604.07316, 2016.
- [2] Hubschneider, C. "Adding navigation to the equation: Turning decisions for end-to-end vehicle control," IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017.
- [3] Codevilla, F. "End-to-end driving via conditional imitation learning," IEEE International Conference on Robotics and Automation (ICRA), 2018.
- [4] Amini, A. "Variational end-to-end navigation and localization," International Conference on Robotics and Automation (ICRA), 2019.

- [5] Simonyan, K. "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.