

자율주행을 위한 VAE 기반의 미래 도로 이미지 생성 모델

김동현, 권재락*, 남해운

한양대학교, *미시간대학교

{kissw, hnam}@hanyang.ac.kr, *jrkwon@umich.edu

VAE-based Future Road Image Generation for Autonomous Driving

Donghyun Kim, Jaerock Kwon*, Haewoon Nam

Hanyang Univ., *University of Michigan-Dearborn

요약

최근 딥러닝 모델을 활용한 자율주행 시스템에 대한 연구가 적극적으로 이루어지고 있다. 딥러닝 모델의 연산으로 인해 발생하는 지연시간을 줄이는 것은 자율주행의 실시간 제어를 하는데 있어 중요한 과제이다. 본 논문에서는 이 문제를 해결하기 위한 새로운 방법인 Variational Auto-Encoder (VAE) 구조를 사용하여 미래의 도로 이미지를 생성하는 방법을 제안한다. 이 아이디어는 생성된 미래 이미지를 통해 미래에 사용될 제어 값을 미리 계산하여 처리 시간을 줄이는 방법이다. 미래 이미지를 생성하기 위해 도로 이미지, 차량의 속도, 조향 각, 그리고 시간을 제안하는 모델에 입력하게 되면 입력 정보에 대한 시공간적 특성을 추출하는 잠재공간으로부터 미래 이미지 생성이 가능하다. 제안된 방법의 가능성을 보이기 위해 생성된 이미지를 정답 이미지와 비교하고, 자율주행을 위한 조향 각 예측 모델에 입력한 뒤 출력 값을 분석을 통해 제안하는 방법의 가능성을 보였다.

I. 서론

최근에는 딥러닝 모델을 이용한 자율주행 시스템 연구가 활발하게 진행되고 있다. 자율주행 차량의 제어를 위한 딥러닝 모델은 도로 이미지와 주변 환경 정보를 입력으로 사용하며, 출력으로 차량의 제어 값을 생성하는 구조를 가지고 있다. 이러한 딥러닝 모델은 다양한 패턴과 상황을 학습하여 복잡하고 변화하는 도로 상황에 적용할 수 있으며, 적절한 학습 데이터가 제공되면 강인한 시스템을 구축할 수 있다고 알려져 있다[1]. 또한, end-to-end 학습 방식을 사용하여 센서 입력부터 출력까지의 모든 과정을 하나의 모델로 학습함으로써, 데이터 처리와 제어 알고리즘 설계를 단순화할 수 있다는 장점이 있다[2].

그러나 딥러닝 모델의 네트워크 구조가 복잡하고 파라미터가 많아질수록 처리 시간이 증가한다. 이 때문에 발생하는 지연 시간은 실시간 제어에 큰 영향을 미칠 수 있으며, 고속 주행의 상황에서 지연된 제어 값을 사용하여 차량을 제어하는 것은 치명적인 사고로 이어질 수 있기 때문에 이 문제는 근본적인 해결책이 필요하다. 지연 시간은 주로 하드웨어의 성능, 딥러닝 모델의 복잡도, 입력 데이터의 크기에 영향을 받는데, 하드웨어의 성능 향상을 위한 비용과 딥러닝 모델 경량화 과정에서 발생하는 성능저하로 인해 지연 시간 개선이 어려운 경우가 많다[3].

따라서 본 논문에서는 지연 시간을 줄이기 위한 새로운 방법으로 Variational Auto-Encoder (VAE) 구조를 활용하여 미래의 도로 이미지를 생성하는 방법을 제안한다. 이 방법은 현재의 도로 이미지, 차량의 속도, 조향 각, 그리고 시간 정보들로부터 잠재변수를 추출하여 미래에 다가올 도로의 이미지를 예측하고 생성한다. 제안된 모델의 성능을 검증하기 위해, 생성된 이미지를 ground truth 이미지와 비교하고, end-to-end 기반 조향 각 예측 모델에 입력했을 때의 예측 조향 각과 ground truth 조향 각의 차이와 분석하였다.

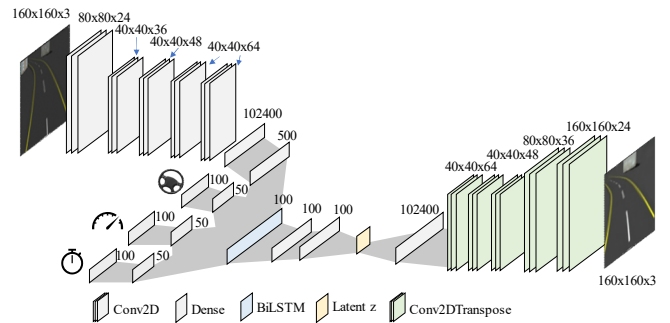


그림 1. 미래 도로 이미지 생성을 위한 VAE 모델 구조

II. 본론

1. 시스템 구조

제안된 모델을 사용하여 미래의 이미지를 생성하고자 할 때, 모델의 입력으로 현재 도로의 이미지 I 와 차량의 속도 v , 조향각 s , 그리고 시간 x 를 사용하며 출력으로 미래 이미지가 생성되게 된다. 여기서 입력과 출력의 의미는 현재 도로 이미지가 I 와 같을 때, 차량을 조향각 s , 속도 v 에서 시간 x 동안 제어했을 때 나타날 미래 이미지를 생성하라는 의미로 쉽게 설명할 수 있다. 이를 위해 모델은 입력 이미지와 조향 각, 속도, 시간의 정보를 압축하는 encoder와 미래의 입력 이미지를 생성하기 위한 decoder로 구성되어 있다.

Encoder에서 입력 이미지로부터 5개의 convolution 레이어를 통해 공간적 특징을 학습하고, fully connected 레이어를 통해 속도, 조향 각, 시간 정보와 결합되어 Bidirectional LSTM (BiLSTM) 레이어에 입력된다. BiLSTM 레이어를 통해 시간적 특징을 학습한 뒤, 잠재공간을 생성하기 위해 sampling layer를 통과하게 되며 이렇게 생성된 잠재공간에는 입력 정보로부터 추출된 시공간적 특징을 나타내는 잠재변수로 구성되게 된다. Decoder는 잠재공간을 입력으로 하여 미래의 이미지를 생성하기 위한 역할을 하며, 잠재 변수는 dense 레이어와 Conv2DTranspose 레이어를 통

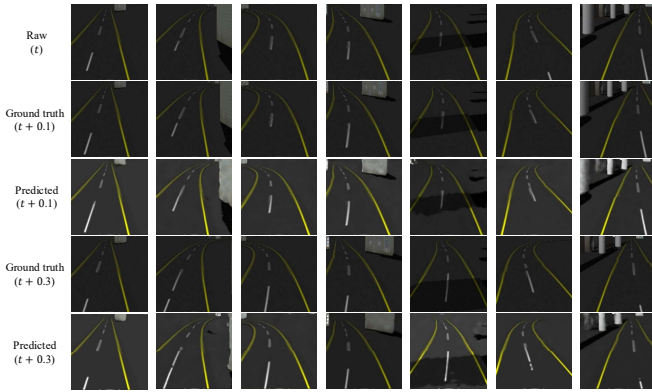


그림 2. 제안하는 모델을 사용하여 생성된 이미지

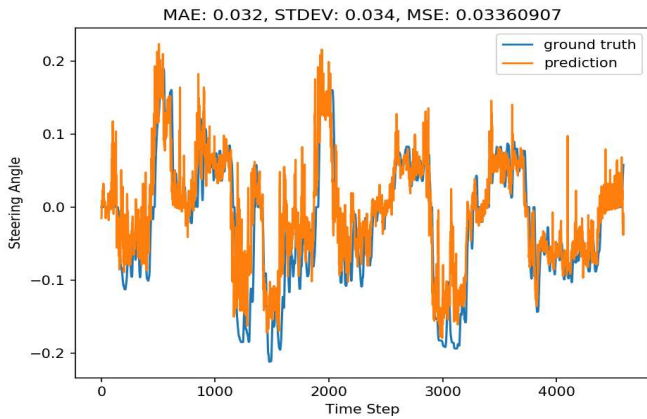


그림 3. 생성된 이미지를 조향 각 예측 모델에 입력한 결과
해 적절한 미래 이미지를 생성하게 된다.

2. 학습 및 데이터 수집

이미지 생성 모델을 학습시키기 위해선 많은 주행 데이터를 필요로 하기 때문에 ROS Melodic 및 Gazebo 9 시뮬레이터를 사용한 가상환경을 구축하였다. 이후 가상 환경의 차량을 이용하여 주행 시 이미지, 조향 각, 시간, 속력을 한 쌍으로 하는 데이터를 수집하였다.

학습에서는 현재 시점 t 에서의 이미지 I_t 와 시간 x 후에 나타날 미래 이미지 I_{t+x} , 그리고 시간 x 동안 수집된 조향 각과 속력의 평균을 사용하였다. 학습에 사용된 총 데이터의 수는 744,180개이며 이 중 85%를 학습에, 15%는 모델을 검증하는데 사용하였다. VAE 모델을 학습하기 위해 2가지 손실 함수를 설정해야하며 reconstruction error를 줄이기 위해 mean squared error (MSE) 를 사용하였고 regularization error를 줄이기 위해 KL-divergence를 사용하여 학습을 진행 하였다.

3. 실험 결과

제안하는 딥러닝 모델의 이미지 생성 성능을 확인하기 위해 미래 시점을 0.1, 0.3초로 설정하여 생성되는 미래 이미지를 확인하였다. 그림 2에서 볼 수 있듯이, 생성된 0.1초 뒤의 이미지와 0.3초 뒤의 이미지는 ground truth 이미지와 매우 유사하게 생성되었으며, 이는 제안된 모델이 입력 정보로부터 시공간적 특징을 효과적으로 추출하여, 미래의 도로 이미지를 생성하였음을 보여준다. 또한 그림 3은 생성된 미래 이미지를 end-to-end 기반 조향 각 예측 모델의 입력으로 사용하였을 때와 ground truth 를 비교한 그래프이다. 검증에 사용된 end-to-end 기반 조향 각 예측 모델은 PilotNet 구조를 사용하였으며, 이 모델은 MSE를 사용하여 검증데이터의 조향 각 오차

를 측정하였을 때 0.0392 ($\approx 3.5^\circ$)의 오차를 보였다. 해당 조향 각 예측 모델에 생성된 미래 도로 이미지를 입력하여 조향 각을 예측했을 때, 생성된 이미지를 모델의 학습에 사용하지 않았음에도 불구하고 MSE가 0.0336($\approx 3.02^\circ$)으로 낮게 나온 것은 미래 도로 이미지가 실제 주행에서 사용될 수 있을 만큼 충분히 실제 이미지와 유사하게 생성되었음을 시사한다.

III. 결론

본 논문에서는 지연 시간을 줄이기 위해 VAE 구조를 활용한 미래 이미지 생성 모델을 제안하였다. 제안하는 모델을 통해 생성된 이미지는 ground truth 이미지와 매우 유사하게 생성되었으며, 자율주행 네트워크의 입력으로 사용 하였을 때도 오차가 적어 실제 주행에서도 사용 할 수 있을 것으로 보인다. 그러나 제안된 모델은 단기간 미래 이미지만 생성 가능하다는 한계가 있다. 따라서 후속 연구에서는 인간의 사고 과정을 모사한 내부 시뮬레이션 방식을 적용하여 먼 미래의 이미지를 효과적으로 생성하는 연구를 진행 할 예정이다.

ACKNOWLEDGMENT

본 연구는 2022년도 중소벤처기업부의 기술개발사업 지원에 의한 연구 임 (No. RS-2022-00164803)

참 고 문 헌

[1] 황병훈, 황효석. “자율주행 인식 알고리즘 학습을 위한 데이터셋 동향,” *전자공학회지*, vol. 50, no. 3, pp. 35-41, 2023.
 [2] 장인식, 정은지, 이현철. “영상 기반 자율경로주행을 위한 관심 영역 기반 End-to-End 심층 학습,” *한국통신학회 학술대회논문집*, pp. 1258-1259, 2022.
 [3] 지민구, 김종찬. “자율주행을 위한 실시간 객체 검출기의 신경망 전환,” *한국정보과학회 학술발표논문집*, pp. 1486-1488, 2021.