

심층강화학습 기반 자율주행차량의 차로변경 방법론

Lane Change Methodology for Autonomous Vehicles Based on Deep Reinforcement Learning

박 다 윤* · 배 상 훈** · Trinh Tuan Hung*** · 박 부 기**** · 정 보 경*****

* 주저자 : 부경대학교 지구환경시스템과학부 공간정보시스템공학전공 석사과정

** 교신저자 : 부경대학교 공간정보시스템공학과 교수

*** 공저자 : 부경대학교 지구환경시스템과학부 공간정보시스템공학전공 박사과정

**** 공저자 : 부경대학교 지구환경시스템과학부 공간정보시스템공학전공 박사과정

***** 공저자 : 부경대학교 지구환경시스템과학부 공간정보시스템공학전공 석사과정

DaYoon Park* · SangHoon Bae** · Trinh Tuan Hung*** · Boogi Park*** · Bokyoung Jung*

* Master's Student, Dept. of Spatial Information Eng., Pukyong National University

** Professor, Dept. of Spatial Information Eng., Pukyong National University

*** PhD Student, Dept. of Spatial Information Eng., Pukyong National University

† Corresponding author : Sang Hoon Bae, sbae@pknu.ac.kr

Vol. 22 No.1(2023)
February, 2023
pp.276~290pISSN 1738-0774
eISSN 2384-1729
<https://doi.org/10.12815/kits.2023.22.1.276>Received 30 November 2022
Revised 25 December 2022
Accepted 6 February 2023© 2023. The Korea Institute of
Intelligent Transport Systems. All
rights reserved.

요 약

현재 국내에서는 자율주행차량의 상용화를 목표로 다양한 노력을 기울이고 있으며 자율주행차량이 운영 가이드라인에 따라 안전하고 신속하게 주행할 수 있는 연구들이 대두되고 있다. 본 연구는 자율주행차량의 경로탐색을 미시적인 관점으로 바라보며 Deep Q-Learning을 통해 자율주행차량의 차로변경을 학습시켜 효율성을 입증하고자 한다. 이를 위해 SUMO를 사용하였으며, 시나리오로 출발지에서 랜덤 차로로 출발하여 목적지의 3차로까지 차로변경을 통해 우회전하는 것으로 설정하였다. 연구 결과 시뮬레이션 기반의 차로변경과 Deep Q-Learning을 적용한 시뮬레이션 기반의 차로변경으로 구분하여 분석하였다. 평균 통행 속도는 Deep Q-Learning을 적용한 시뮬레이션의 경우가 적용하지 않은 경우에 비해 약 40% 향상되었으며 평균 대기 시간은 약 2초, 평균 대기 행렬 길이는 약 2.3대 감소하였다.

핵심어 : 자율주행, 경로계획, 차로변경, Deep Q-Learning

ABSTRACT

Several efforts in Korea are currently underway with the goal of commercializing autonomous vehicles. Hence, various studies are emerging on autonomous vehicles that drive safely and quickly according to operating guidelines. The current study examines the path search of an autonomous vehicle from a microscopic viewpoint and tries to prove the efficiency required by learning the lane change of an autonomous vehicle through Deep Q-Learning. A SUMO was used to achieve this purpose. The scenario was set to start with a random lane at the starting point and make a right turn through a lane change to the third lane at the destination. As a result of the study, the analysis was divided into simulation-based lane change and simulation-based lane change applied with Deep Q-Learning. The average traffic speed was improved by about 40% in the case of simulation with Deep Q-Learning applied, compared to the case without application, and the average waiting time was reduced by about 2 seconds and the average queue length by about 2.3 vehicles.

Key words : Autonomous driving, Path planning, Lane-Change, Deep Q-Learning

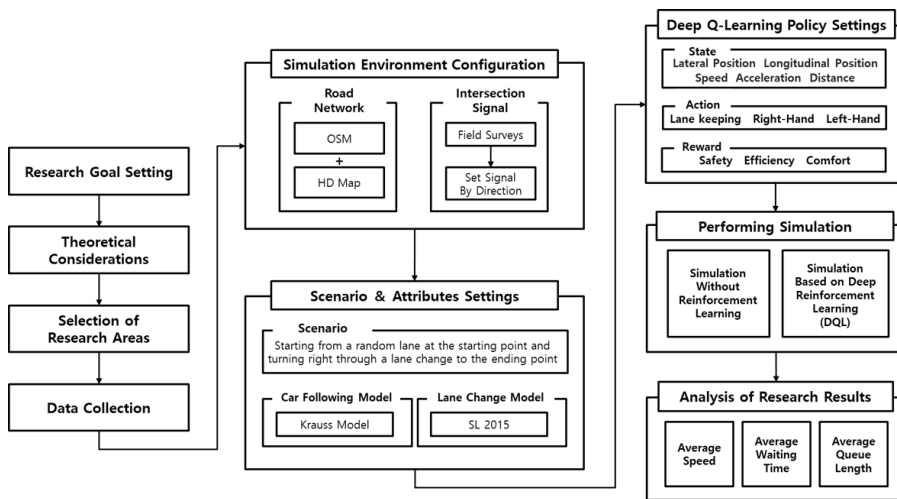
I. 서론

1. 연구 배경 및 목적

2022년 9월 국토교통부에서 발표한 모빌리티 혁신 로드맵에 따르면 정부는 2025년까지 운전자가 개입하지 않는 Lv4단계의 완전자율주행 대중교통을, 2027년까지 완전자율주행 승용차를 출시하겠다는 목표를 세웠다. 자율주행 규제 혁신으로는 2024년까지 Lv4에 대한 운행 제도를 선제적으로 마련할 예정이며 자율주행 친화적 인프라 또한 속도감 있게 구축을 추진할 예정이라고 밝혔다(MOLIT, 2022). 현재 자율주행차량의 상용화를 목표로 지속적인 기술 개발 및 규제 개선이 추진되고 있으며 상용화를 위해서는 자율주행차량이 운영 가이드라인에 따라 안전하고 신속하게 주행할 수 있는 방안이 요구된다. 자율주행기술의 동작 제어는 종방향과 횡방향으로 구분되며 종방향 동작 제어는 차량추종 모델을, 횡방향 동작 제어는 차로변경 모델을 기반으로 한다. 따라서 본 연구는 심층강화학습을 통해 차로변경 모델을 학습시켜 자율주행차량이 주변 교통 상황을 고려하여 선두 차량과 충돌하지 않고 적절한 시점에 차로변경을 수행하여 목적지까지 최적경로로 도달할 수 있도록 하고자 한다.

2. 연구 수행 과정

<Fig. 1>은 본 연구의 수행 과정을 나타낸 것이다. 연구 목적을 설정하여 그에 맞는 이론적 고찰 수행 및 연구 대상지를 선정하였다. 선정된 연구 대상지의 도로 기하구조, 신호 체계 등 관련 데이터를 수집하여 시뮬레이션 환경 구축 시 기초 자료로 사용하였다. 도로 네트워크는 OSM(Open Street Map)과 HD MAP(High Definition Map)을 활용하였으며 교차로 신호 체계는 현장 조사를 기반으로 구성하였다. 시뮬레이션의 시나리오 오는 출발지에서 랜덤 차로로 출발하여 목적지까지 차로변경을 통해 우회전하도록 설정하였다. 이때 차량추종 모델은 Krauss모델을, 차로변경 모델은 SL2015 모델을 사용하였다. 심층강화학습 모델은 Deep Q-Learning을 사용하였으며 연구 결과는 평균 통행 속도, 평균 대기 시간, 평균 대기행렬 길이를 시뮬레이션 기반의 차로변경과 Deep Q-Learning을 적용한 시뮬레이션 기반의 차로변경으로 구분하여 비교 분석하였다.



<Fig. 1> Research Flow chart

II. 이론적 고찰

1. 자율주행차량의 운영 가이드라인

1) ODD(Operational Design Domain)

미국 도로교통 안전국 NHTSA(National Highway Traffic Safety Administration)에서 제시한 ODD는 자율주행 시스템의 특정 작동 조건을 정의하기 위한 운행설계 범위로 <Table 1>과 같이 물리적 인프라, 운영 제약, 연결성, 환경 조건, 객체, 구역으로 구성되어 있다(NHTSA, 2018).

<Table 1> ODD Category Descriptions

ODD elements	Detailed Components	ODD elements	Detailed Components
Physical infrastructure	Roadway type	Environmental conditions	Weather
	Roadway Surfaces		Weather-induced roadway conditions
	Roadway edges & markings		Particulate matter
	Roadway geometry		Illumination
Operation constraints	Speed limits	Objects	Signage
	Traffic conditions		Roadway users
Connectivity	Vehicles		Zone
	Traffic density information	Geo-fencing	
	Remote fleet management system	Traffic management zones	
	Infrastructure sensors and communications	School/construction zones	
Regions/State			
		Interference zones	

NHTSA의 보고서에는 ODD 분류법에 기초하여 자율주행시스템의 기능을 식별하기 위한 체크리스트를 제시하였으며 <Table 2>는 Lv4 Highly Automated Vehicle/TNC Drive에 대한 체크리스트의 일부를 나타낸 것이다.

<Table 2> ODD Checklist : Lv4 Highly Automated Vehicle/TNC

Physical Infrastructure		
Roadway types	Divided/Undivided highway, Arterial, Urban, Rural, Parking, Multi-lane/Single lane etc	Y
Roadway surfaces	Asphalt, Concrete, Mixed	Y
Roadway Geometry	Straightaways, Curves, Hills, Lateral crests, Corners, Negative obstacles, Land width	Y
Operation Constraints		
Speed limits	Minimum speed limit, Maximum speed limit(At least 35 mph is likely to be needed to traverse a city)	Y
Traffic conditions	Altered(accident, emergency vehicle, construction, closed road, special event)	Y
Objects		
Signage	Signs, Traffic Signals, Crosswalks, Railroad crossing, Stopped buses etc	Y
Roadway users	Vehicle types, Stopped vehicles, Other automated vehicles, Pedestrians, Cyclists	Y

2) OEDR(Object and Event Detection and Response)

자율주행시스템이 주행 환경 모니터링을 통해 객체 및 이벤트를 감지, 인식, 분류하고 필요에 따라 대응 준비를 하는 DDT(Dynamic Driving Task)의 하위 작업을 의미한다. 자율주행시스템은 자율주행차량의 운행에 영향을 줄 수 있는 동적 및 정적 객체를 정확하게 인지하고 이에 발생 가능한 상황에 대응할 수 있어야 한다.

OEDR과 관련하여 자율주행시스템의 기능 요소에는 다음을 지원하는 하드웨어 및 소프트웨어 구성요소가 포함된다.

- 감지 (예 : 레이더, 레이저 스캐너, 카메라 등)
- 지각 (예 : 도로 지형 분류, 객체 분할 및 분류 등)
- 모델링 (예 : 영속적인 데이터 매핑, 동적 장애물 추적 및 예측 등)
- 경로 안내 및 계획 (예 : 응답을 구현하기 위한 경로 계획 및 동작 제어 명령)

2. 미시적인 관점에서의 자율주행차량 경로계획

경로계획은 자율주행기술의 중요한 구성 요소 중 하나이며 행동 결정 모듈과 동작 제어 모듈 사이의 중요한 연결 고리이다(Li et al., 2020). 기존의 경로계획은 최단경로탐색 알고리즘인 Dijkstra Algorithm, A* Algorithm, Bellman-Ford Algorithm 등을 활용하여 전체 도로 네트워크를 거시적인 관점으로 바라보며 링크 단위의 경로탐색을 수행하였다. 그러나 자율주행기술의 상용화 시대가 다가오며 최근 연구에서는 경로계획을 미시적인 관점으로 바라보며 차량추종, 차로변경, 교차로 교차 등 다양한 기동 모드를 포함한 탐색을 수행하고 있다. 복잡한 도심환경에서 자율주행을 하는 경우 최종 목적지까지 도달하기 위해 다수의 교차로를 마주하게 되며 차로변경을 통한 좌회전 또는 우회전 이동이 필수적이다. 자율주행 기술의 필수적인 요소 중 하나인 정밀도로지도는 자율주행차량의 위치 결정, 경로 설정, 도로 및 표지 시설 정보 인식 등을 가능하게 하며 3차원으로 표현한 정밀전자지도로 차로단위의 정보로 구축된다(Shim et al., 2020). 따라서 자율주행 상용화 시대에 맞춰 미시적인 관점에서의 정확하고 세밀한 차로단위의 경로계획이 반드시 필요하다.

3. 차로변경을 통한 자율주행차량 경로계획

역동적으로 변화하는 구조 환경에서 자율주행차량은 차로변경을 통해 단시간에 안전하고 원활한 경로를 생성해야 한다. 인간운전자에 의한 차로변경은 선행하는 차량과의 간격을 고려하여 방향지시등을 작동한 후 대기하였다가 공간을 확보한 후 차로변경을 수행한다. 이 과정에서 운전자의 운전 실력, 사회적 이해, 양보 등 다양한 요소들에 의해 영향을 받는다(Chang and Yoo, 2021). 반면 자율주행차량은 정밀도로지도를 통해 주행 중인 차로에 대한 정보를 제공받고 센서 및 통신을 통해 주변 교통 상황을 파악하여 차로변경을 수행하게 된다. 따라서 인적 요인에 영향을 받지 않기 때문에 지체 없이 원활한 차로변경이 가능하다. Choi et al.(2015)은 자율주행차량이 GIS 정보를 이용하여 차로변경 및 경로계획에 대해 올바른 결정을 내릴 수 있는지 논의하고 검증하였다. 실차 테스트를 통해 총 4가지 시나리오를 실험한 결과 차선, 과속 방지턱, 교통 표지판 등 도로 네트워크 정보를 포함한 GIS 데이터를 활용하여 자율주행차량이 장애물을 감지할 경우 차로변경을 통해 우회경로 생성이 가능함을 확인하였다(Choi et al., 2015).

Kwon et al.(2022)은 정밀도로지도를 통해 획득한 레이어 정보를 기반으로 자율주행차량이 장애물 회피를 위해 차로변경을 수행한 후 새로운 경로를 생성하는 알고리즘을 개발하였다. 시뮬레이션 시나리오는 자율주행차량이 직선구간을 통과하여 좌회전하는 경로 내 정적인 장애물이 있는 경우 이를 회피한 후 최종 목적지

까지 도달할 수 있도록 하였다. 연구 결과 자율주행차량이 장애물을 인지한 후 정밀도로지도의 데이터를 기반으로 새로운 지역경로를 생성하였으며 회피 이후에는 기존의 전역경로를 따라가는 모습을 볼 수 있었다(Kwon et al., 2022). 관련 문헌고찰을 통해 차로변경을 통한 자율주행차량 경로계획을 위해 차로단위의 정보를 포함한 정밀도로지도가 필요하며 이를 통해 본 연구의 시뮬레이션 속 도로 네트워크를 구성하였다. 시나리오는 장애물 인지와 같은 특정한 상황이 아닌 회전이동이 필수적인 도심환경에서 차로변경을 통해 우회전하여 최종 목적지까지 최적의 경로로 주행할 수 있도록 설정하였다.

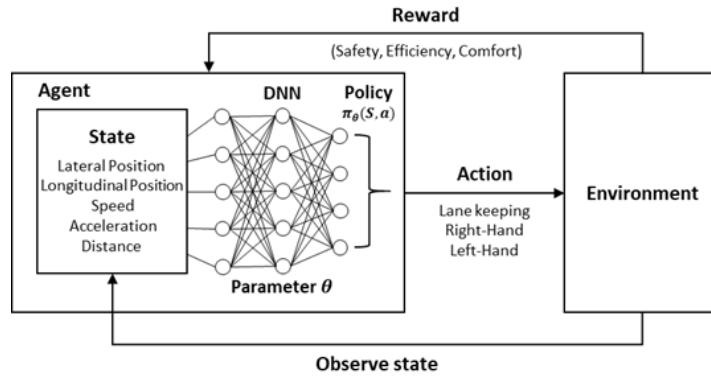
4. 강화학습 기반의 자율주행차량 경로계획

규칙 기반의 경로계획은 미리 정의된 운영 조건에서는 잘 수행될 수 있지만 불확실성과 운전 환경의 복잡성으로 인해 제한된 성능을 보이는 경향이 있다(Wang et al., 2018). 또한 훈련을 위해 가능한 모든 조건(교통량, 도로 표지판, 신호등 등)에서 상당한 양의 데이터를 수집해야 하며, 이는 비용이 많이 발생하며 비현실적이다. 반면 강화학습 기반의 경로계획은 복잡하고 불확실한 환경에서 상호 작용에 의한 학습을 위한 접근법으로 에이전트는 환경으로부터 받은 보상의 총량을 최대화하려고 한다. 따라서 에이전트가 예상치 못한 시나리오일지라도 지능적으로 경로계획을 수행할 수 있는 강화학습 모델을 활용하고자 한다. Kang et al.(2021)은 강화학습 기법인 DQN(Deep Q Network)과 DDPG(Deep Deterministic Policy Gradient)를 동시에 활용하여 자율주행차량의 최적의 경로를 찾는 모델을 제안하였다. DQN은 지역경로를 판단하기 위해 사용하였으며 DDPG는 속도제어를 위해 사용하였다. 강화학습의 State는 총 9개(현재 속도, 우회차선변경 가능 여부, 좌회차선변경 가능 여부, 선행자동차와의 거리, 후행자동차와의 거리, 현재 차선, 경로 내 현재위치, 이동방향, 현재교통신호)로 구분하였으며 Action은 우측 차로변경, 직진, 좌측 차로변경을 각각 [0], [1], [2]로 구분하였다. Reward는 각 Time step의 속도의 합으로 정하였으며 급정지, 충돌 등의 돌발 상황이 발생할 경우에는 -1의 penalty를 부과하였다(Kang et al., 2021). Ye et al.(2020)은 PPO(proximal policy optimization)기반의 강화학습을 사용하여 자동화된 차로변경 전략을 통해 안전하고 효율적인 경로계획을 제안하였다. State는 자아차량을 기준으로 종방향 위치 및 속도, 가속도, 횡방향 위치 및 속도를 포함한 5가지의 상태변수로, 주변 차량을 기준으로 자아차량과의 상대적 거리, 종방향 속도, 가속도, 횡방향 위치를 포함한 4가지 상태변수로 정하였다. Action은 횡방향 동작과 종방향 동작으로 구분하였으며 횡방향 동작은 차로를 유지하는 경우 [0], 목표 차로로 변경하는 경우 [1], 차로변경 조작을 중단하는 경우 [2]로 설정하였다. 종방향 동작은 기존 차로의 선행차량을 따르는 경우 [0], 목표차로의 선행차량을 따르는 경우 [1]로 설정하였다. Reward는 안전성, 효율성, 편안함을 중심으로 설계하였다. 안전성은 충돌 및 근접 충돌 위험에 대한 평가, 효율성은 목표 차로까지의 이동 시간 및 상대거리 평가, 편안함은 저크의 평가로 구성하였다(Ye et al., 2020). 강화학습 기반의 자율주행차량 경로계획과 관련한 논문고찰을 통해 본 논문의 정책 결정 시 적합한 State, Action, Reward를 적용하였다. 상세한 정책 설정에 대한 내용은 III. 연구 내용의 3) Deep Q-Learning의 정책 설정에서 제시하였다.

5. 심층강화학습 모델

본 연구에서 사용된 강화학습 모델은 Deep Q-Learning이며 <Fig. 2>는 모델의 구조를 나타낸 것이다. Deep Q-Learning은 Q-Learning에 DNN(Deep Neural Network)를 결합한 심층강화학습 모델로 Experience Replay Buffer 훈련이 필수적이다. Experience Replay Buffer란 강화학습을 진행하면서 바로 가중치(w)를 학습시키는 것이 아니라 Time-step마다 [S(Current State), A(Action), R(Reward), S'(Next State)] 데이터 셋을 모아 Tuple 형

태로 Replay Memory라는 Buffer Pool에 저장해두고 랜덤하게 Mini-batch를 구성하여 업데이트하는 방법이다. 데이터를 랜덤하게 추출하게 되면 시간에 따른 각 데이터 간의 상관관계가 줄어들게 되며 고르게 분포하게 된다. 이 방법은 일반적인 강화학습 셋업의 훈련 샘플은 높은 상관관계가 있고 데이터 효율성이 낮기 때문에 네트워크의 수렴이 어려워진다는 문제를 해결 할 수 있다(Mnih et al., 2013).



<Fig. 2> Structure of Deep Q-Learning

6. 차량추종 모델

차량추종 모델은 동일한 차로에서 선두 차량의 주행 동작에 반응하여 가속 및 감속과 같은 차량의 종방향 동작을 제어한다. SUMO에 적용이 가능한 차량추종 모델은 Krauss, IDM, ACC 등이 있으며 본 연구는 안전 속도 유지를 중점으로 둔 미시적, 공간 연속적 차량추종 모델인 Krauss 모델을 사용하였다. Krauss 모델은 1998년 Krauß에 의해 개발되었으며 운전자가 선행 차량과의 원하는 거리가 항상 유지되도록 속도를 조정한다는 가정 하에 안전 속도 기반의 차량추종 모델이다. 교통 시뮬레이션에서 선행 차량과 후행 차량이 상호 작용하는 경우 두 차량이 서로 충돌하지 않도록 노력하기 때문에 운전자 중 한 명 이상은 최대 안전 속도보다 높지 않은 속도로 줄여야 한다(DLR, 1998).

$$v \leq v_{safe} \dots\dots\dots (1)$$

v_{safe} (안전 속도)는 식 (2)과 같이 계산되며 차량이 동일한 차로에서 선행 차량과 충돌하지 않고 유지할 수 있는 최대 속도로 정의된다. $v_i(t)$ 은 시간 t 에서 선행 차량의 현재 속도이며, $g(t)$ 은 시간 t 에서 선행 차량과의 거리 간격을 나타낸다. t 은 시뮬레이션 기간, t_r 은 운전자의 반응 시간(약 1초)이며 b 은 차량의 최대 감속도이다(Song et al., 2014).

$$v_{safe} = v_i(t) + \frac{g(t) - v_i(t)t_r}{\frac{v_i(t) + v_f(t)}{2b} + t_r} \dots\dots\dots (2)$$

v_{safe} 가 계산되면 차량의 원하는 속도인 v_{des} (희망 속도)를 계산할 수 있다. 현실에서 차량의 가속도는 물리적 능력, 공기 저항과 같은 효과에 따라 달라지며 가속 기능으로 인해 도로에서 허용되는 최대 속도보다 클

수 있기 때문에 이를 방지하고자 최솟값을 속도로 계산한 것이 식 (3)의 v_{des} 이다.

$$v_{des} = \min[v_{max}, v + at, v_{safe}] \dots\dots\dots (3)$$

7. 차로변경 모델

차로변경 모델은 차량의 횡방향 동작을 제어하며 차량의 경로와 주변의 교통 상황을 바탕으로 단일 시뮬레이션 단계에 대한 차량의 변경 결정을 계산한다. SUMO에서 구현 가능한 차로변경 모델은 3가지(DK2008, LC2013, SL2015)로 본 연구는 차로변경을 연속적인 프로세스로 간주하며 차선의 각도, 횡방향 속도, 가속도 및 기타 정보를 고려한 최신 모델인 SL2015 모델을 사용하였다(Wang et al., 2021). SL2015 모델은 LC2013 모델을 기반으로 한 차로변경 모델로 차량추종을 위한 차로 변경, 협력적 차로 변경, 의무적 차로 변경, 속도 향상을 위한 차로 변경 이외에도 일정한 횡방향 정렬을 유지하기 위한 차로 변경이 가능하다. SL2015 모델은 <Table 3>과 같이 LC2013 모델에서 지원하는 매개 변수 외에도 lcSublane, lcPushy 등 지원하는 주요 매개 변수들이 있다(Eclipse SUMO, 2022a).

<Table 3> SL2015 Parameters used in SUMO

Attribute	Description
lcSublane	The eagerness for using the configured lateral alignment within the lane
lcPushy	Willingness to encroach laterally on other drivers
lcAssertive	Willingness to accept lower front and rear gaps on the target lane
lcImpatience	Dynamic factor for modifying lcAssertive and lcPushy
lcTimeToImpatience	Time to reach maximum impatience
lcAccelLat	Maximum lateral acceleration per second
latAlignment	preferred lateral alignment within a lane

Ⅲ. 연구 내용

1. 연구 범위

본 연구는 부산광역시 남구 대연동 용소삼거리·부경대학교 인근 교차로를 중심으로 수행하였다. 용소삼거리는 명칭은 삼거리이나 경성대학교 정문을 출입하는 차량들로 인하여 실질적으로는 사거리 교차로로 운영된다. 부경대 정문에서 용소삼거리 방향으로 왕복 4차로 구간(486m)과 왕복 5차로로 분리되는 구간(145m)을 선정하였고 교차로를 중심으로 가로 구간(215m)과 직선 구간(60m)을 연구 범위로 선정하였다.

<Table 4>과 같이 NHTSA에서 제시한 Lv4단계의 ODD 체크리스트에 따라 도로 인프라, 운영 제약사항, 환경 조건 등이 부합하다고 판단되어 공간적 범위를 선정하였다.



<Fig. 3> The spatial scope of research

<Table 4> Spatial Coverage Suitable for ODD Checklist

ODD Elements		
Physical infrastructure	Roadway types	Urban, Multi-lane, Bicycle Lanes, Crosswalks
	Intersection types	Signaled, 4-way, Stop sign, Left/Right Turn
	Roadway surfaces	Asphalt
	Roadway Edges and Markings	Lane Markers, Curbs
	Roadway Geometry	Straightaways, Curves
Operational Constraints	Minimum/Maximum Speed Limit	0~30km/h
	Traffic Density	Normal
Objects	Signage	Signs, Traffic Signals, Crosswalks, Roadway User Signals
	Roadway users	Car, Pedestrians, Motorcycles, Cyclists etc
Zones	Geofencing	School Campuses

시간적 범위는 국토교통부가 제시한 모빌리티 혁신 로드맵에 따라 Lv4단계의 승용차가 도입되는 시점인 2027년으로 설정하였다(MOLIT, 2022).

내용적 범위는 자율주행차량이 ODD 체크리스트에 부합하는 환경에서 주변 환경 인지 및 차량 제어를 통해 출발지부터 목적지까지 신속하고 안전하게 가는 것이다. 복잡한 도심환경에서는 반드시 적절한 시점에서의 차로변경이 요구되며 본 연구는 시물레이션 기반의 차로변경과 Deep Q-Learning을 적용한 시물레이션 기반의 차로변경의 평균 통행 속도, 평균 대기 시간, 평균 대기행렬 길이를 각각 비교 분석하고자 하였다.

2. 데이터 수집 및 분석

연구 대상지 내 교차로 신호운영을 파악하기 위해 현장조사를 실시하였다. 현장조사는 평일 점심 시간대에 진행하였으며 용소 삼거리 내 위치한 신호등 4곳의 방향별 신호 데이터를 수집하였다. 용소 삼거리의 신호 현시는 1 Cycle당 160초로 나타났으며 이는 시뮬레이션 환경 구성 시 입력 값으로 사용되었다.

3. 시뮬레이션 환경 구축

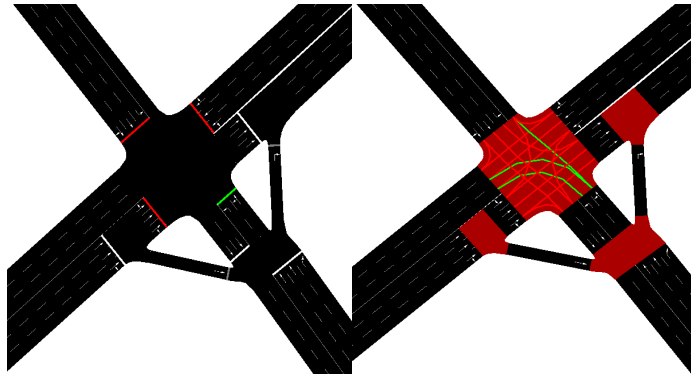
시뮬레이션을 위해 오픈소스 기반의 미시적 교통 시뮬레이터인 SUMO(Simulation of Urban Mobility)를 활용하였다. SUMO는 2001년 독일 항공우주 센터에서 교통 예측, 경로 선택 등 연구 목적으로 개발한 것으로 실제 세계와 가까운 도로 교통 시스템을 모델링 할 수 있다. 또한, TraCI(Traffic Control Interface)를 통해 외부 어플리케이션과 SUMO를 연결하여 객체의 값을 검색하고 조작하는 것이 가능하다(Eclipse SUMO, 2022b).

1) 도로 네트워크 구성

오픈 소스 지도 서비스인 OSM을 통해 기초 환경을 구성한 후 차로 수, 차로 길이 등 세부 네트워크 정보는 2020년 항공영상을 기반으로 제작된 10cm급의 고해상도 정밀도로지도를 활용하였다.

2) 교차로 신호 구성

현장조사를 통해 수집한 방향별 신호주기 데이터를 8단계로 구분하였으며 적용한 Cycle 시간은 총 160초이다. <Fig. 4>는 SUMO의 그래픽 네트워크 에디터인 Netedit를 통해 시뮬레이션 속 교차로의 방향별 신호를 설정하는 과정을 나타낸 것이다.



<Fig. 4> Intersection Signal Setting

4. 심층강화학습 기반 시뮬레이션 수행

1) 환경 변수 설정

시뮬레이션에 적용된 교통량은 실제 대기행렬과 유사하게 대기 행렬이 생기는 교통량을 계산하여 설정하였다. 부경대학교 정문에서 용소삼거리 방향으로 좌회전 차량 520대, 직진 차량 52대, 우회전 차량 390대로 총 962대를 적용하였다.

차량추종 모델과 차로변경 모델은 SUMO에 적용이 가능한 Krauss 모델과 SL2015 모델을 사용하였다. 차량추종 모델의 변수(Min gap, Sigma, Tau)는 <Table 5>와 같다(Lu and Tettamanti, 2018).

<Table 5> Variables in SUMO car following model

Parameter	Regular Vehicle	Connected Automated Vehicle
Min gap(m) : the offset to the leading vehicle when standing in a jam.	2.5	0.5
Sigma : the driver imperfection(between 0 and 1)	0.5	0
Tau(s) : the driver's desired (minimum) time headway(reaction time)	0.9	0.6

2) 시나리오 설정

시나리오는 3600초 동안 출발지에서 랜덤 차로로 출발하여 목적지의 3차로까지 차로변경을 통해 우회전 하는 것으로 설정하였으며, 이는 주변의 교통 상황과 선두 차량과의 간격에 영향을 받는다.

3) Deep Q-Learning의 정책 설정

본 연구는 심층강화학습 기법으로 Deep Q-Learning을 사용하였으며 학습의 State(상태), Action(행동), Reward(보상)을 다음과 같이 정의하였다(Universit  Gustave Eiffel, 2021).

State(상태)는 에이전트와 주변 차량의 연속 상태 변수로 구성하였다. 에이전트 상태 변수는 횡방향 위치, 종방향 위치, 속도, 가속도, 거리 5가지 변수로 구성하였고 주변 차량의 상태 변수는 횡방향 위치, 속도, 가속도, 에이전트와의 상대적 거리 4가지 변수로 구성하였다.

$$S_{agent} = \{lateral\ position, longitudinal\ position, velocity, acceleration, distance\} \dots\dots\dots (4)$$

$$S_{surrounding\ vehicle} = \{lateral\ position, velocity, acceleration, relative\ distance\} \dots\dots\dots (5)$$

Action(행동)은 에이전트가 차로변경을 언제, 어떻게 수행할지 배울 수 있도록 모델을 횡방향과 종방향으로 훈련하였다. 횡방향의 경우 세 가지 동작으로 차로를 유지하는 경우는 [0], 우측으로 통행하는 경우는 [1], 좌측으로 통행하는 경우는 [2]로 구분하였다. 종방향의 경우 두 가지 동작으로 현재 차량의 선행 차량을 따를 경우 [0], 목표 차량의 선행 차량을 따를 경우 [1]로 구분하였다.

$$A_{lateral} = \{keep, right, left\} \dots\dots\dots (6)$$

$$A_{longitudinal} = \{follow\ the\ current\ lane, follow\ the\ target\ lane\} \dots\dots\dots (7)$$

Reward(보상)은 Safety(안전성), Efficiency(효율성), Comfort(편안함)이라는 세 가지 목표를 통해 계산하였다. 먼저 안전성은 충돌 위험에 대한 평가로 에이전트와 주변 차량과의 상대적인 거리가 사전 정의된 임계값보다 작아서 충돌이 임박했음을 나타내는 경우 차로변경을 중단하는 방법을 배울 수 있다. 식 (8)에서 D는 에이전트와 주변 차량에 대한 상대적인 거리를 나타낸 것이다.

$$R_{safety} = \begin{cases} R_{near_collision} & \text{if } d < 20m \\ -200 & \text{if } collision \end{cases} \dots\dots\dots (8)$$

효율성은 식 (10)과 같이 이동시간 및 속도에 대한 평가로 $R_{time}(t)$ 은 시간과 관련된 하위 보상이고, $R_{lane}(t)$ 은 에이전트와 주변 차량의 횡방향 위치의 차이를, $R_{speed}(t)$ 는 속도의 차이를 의미한다.

$$\begin{cases} R_{time}(t) = -\delta t \\ R_{lane}(t) = -|P_x - P_y| \\ R_{speed}(t) = -|V_x - V_y| \end{cases} \dots\dots\dots (9)$$

$$R_{eff}(t) = w_r \cdot R_{time}(t) + w_l \cdot R_{lane}(t) + w_s \cdot R_{speed}(t) \dots\dots\dots (10)$$

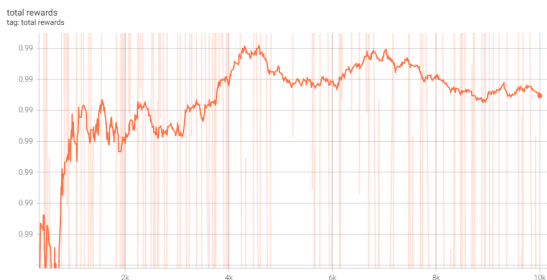
편안함은 횡방향과 종방향의 저크 값의 평가를 통해 계산된다. 저크란 운전자가 브레이크 페달에서 밟을 때면서 기구적 유격으로 인하여 브레이크의 제동력이 발생하는 시점까지 지연이 발생하게 되는데 이때 생기는 차량의 울켜거림을 의미한다. 식 (11)의 a_x 는 횡방향 저크, a_y 는 종방향 저크이며, α 와 β 는 횡방향과 종방향의 편안함에 해당하는 가중치이다. 이 보상 기능은 차량 탑승자의 승차감 향상을 위해 차량의 급가속 또는 감속을 방지하기 위해 도입되었다.

$$R_{comf}(t) = -\alpha \cdot \dot{a}_x(t)^2 - \beta \cdot \dot{a}_y(t)^2 \dots\dots\dots (11)$$

위 내용을 바탕으로 심층강화학습을 수행하였으며 모델 구축을 위해 구글의 전산 프레임워크인 TensorFlow 2.11의 Keras API를 사용하였다. 컴퓨터 환경은 11th Gen Intel(R) Core(TM) i9-11900K @ 3.50GHz 3.50 GHz, Windows 10이며, 언어는 python 3.9 버전을 사용하였다.

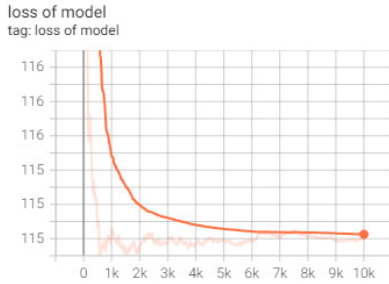
4) 심층강화학습 실행 결과

심층강화학습 훈련 과정은 10,000 epoch 동안 진행되었으며 Step in Epoch= 218, learning_rate=0.00001, discount rate=0.99, batch_size=32, Replay memory start size:33, layer1=64, layer2=128, layer3=64로 실행되었다. 안전성, 효율성, 편안함을 포함한 총 평균 누적 보상에 대한 결과는 <Fig. 5>과 같이 나타났다. 시뮬레이션 중 에이전트는 Reward를 최대화하기 위해 노력하였으며 일부 변동은 있었으나 전반적인 추세는 증가하는 것을 볼 수 있었다. 즉 총 평균 누적 보상의 곡선은 에이전트가 보상을 최대화하기 위한 방법을 성공적으로 학습하였음을 나타냈다.

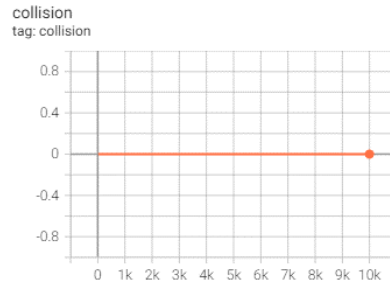


<Fig. 5> Learning Curve of Cumulative Reward

학습 과정에서 총 평균 누적 손실은 <Fig. 6>과 같이 나타났다. epoch가 증가함에 따라 손실 함수 값이 지속적으로 감소함을 보여주었으며 특정 작은 값에 도달하면 모델이 수렴하는 것을 볼 수 있었다. <Fig. 7>의 충돌률은 보상의 안전성 측면에서 선행차량과 후행차량 간의 충돌을 나타낸다. 학습 과정 동안 차로변경 시 선행차량과의 적정 거리를 유지하였으며 높은 차로변경 성능으로 충돌이 발생하지 않아 0%로 나타났다.



<Fig. 6> Total loss



<Fig. 7> Collision rate

IV. 연구 결과

자율주행차량이 주변의 교통 상황과 선두 차량과의 간격을 고려하여 설정된 시나리오에 따라 시뮬레이션 기반의 차로변경을 수행한 경우와 Deep Q-Learning을 적용한 시뮬레이션 기반의 차로변경의 경우로 구분하여 결과를 분석하였다. 연구 결과는 교통류의 이동성 분석지표인 평균 통행 속도, 평균 대기 시간, 평균 대기 행렬 길이로 구분하여 비교를 수행하였다.

1. 평균 통행 속도

효과 척도의 기본 요소로 주행하는 모든 차량들의 통행속도를 평균한 속도이다. 교차로의 신호 체계에 영향을 받으며 적신호가 시작되면 차량의 속도가 떨어지고 녹색 신호가 시작되면 다시 빨라지게 되어 두 경우 모두 유사한 패턴을 보였다. 출발 속도는 동일하게 5m/s로 설정하였으며 <Fig. 8>과 같이 시뮬레이션 기반 차로변경의 경우 평균 통행 속도는 4.94m/s이고 Deep Q-Learning을 적용한 시뮬레이션 기반 차로변경 경우의 평균 통행 속도는 6.95m/s로 약 40% 향상되었다는 것을 알 수 있었다.

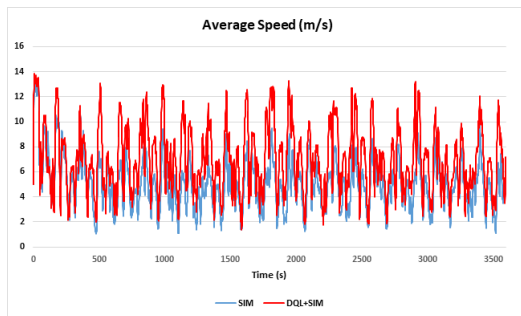
2. 평균 대기 시간

차로변경의 효율성을 입증하기 위하여 SUMO에서 제시한 측정기준에 따라 주행하는 차량의 속도가 0.1m/s 이하일 때 기록이 되며 대기 시간이 길수록 차량 정체 및 통행 시간에 영향을 주게 된다(Erdmann, 2015). 시뮬레이션 기반 차로변경의 경우 46초부터 기록되었고 Deep Q-Learning을 적용한 시뮬레이션 기반 차로변경의 경우는 45초부터 기록되었다. 초기에는 두 경우 모두 녹색 신호가 시작되면 정체가 완화되어 대기 시간이 발생하지 않았으나 시뮬레이션을 진행할수록 시뮬레이션 기반의 차로변경의 경우에는 녹색 신호가 시작되어도 일정 시간 동안 대기 시간이 발생하였다. 즉, 녹색 신호를 부여받아도 신호 시간이 30초 정도

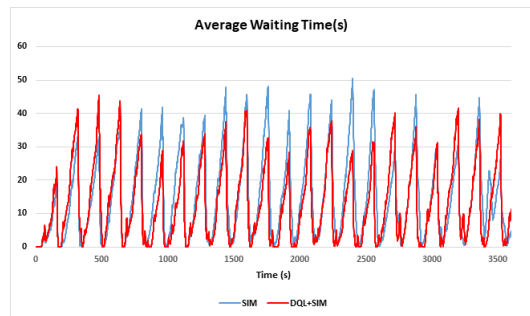
로 짧기 때문에 정체가 모두 완화되지 않았다는 것을 의미한다. <Fig. 9>과 같이 시뮬레이션 기반 차로변경의 경우 평균 대기 시간은 14.68초이며 Deep Q-Learning을 적용한 시뮬레이션 기반 차로변경의 경우 평균 대기 시간은 12.63초로 약 2초 정도 감소하였다.

3. 평균 대기행렬 길이

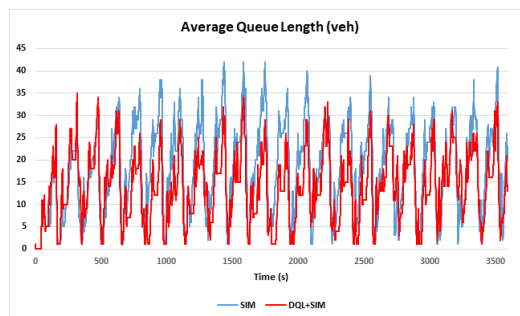
정차된 차량의 시작점부터 마지막 지점까지의 총 차량 대수를 평균한 것을 의미하며 대기행렬 길이에 따라 교차로 신호 운영 및 해당 도로의 서비스 수준에 영향을 준다. 평균 대기 시간이 기록되는 시점부터 기록 되었으며 초기에는 20대 이하로 나타나다가 녹색 신호가 시작되면 1대로 기록되었다. 시뮬레이션을 진행할 수록 평균 대기 시간과 마찬가지로 녹색 신호가 시작되었으나 선두 차량이 다 빠져나가지 못해 대기행렬을 생성하였다. 신호 운영에 따라 시뮬레이션 기반 차로변경의 경우 최대 대기행렬 길이는 42대이며 Deep Q-Learning을 적용한 시뮬레이션 기반 차로변경의 경우 최대 대기행렬 길이는 35대로 나타났다. 평균 대기행렬 길이는 <Fig. 10>과 같이 시뮬레이션 기반 차로변경의 경우가 16.2대이며 Deep Q-Learning을 적용한 시뮬레이션 기반 차로변경의 경우 13.91대로 약 2.3대 감소한 것을 볼 수 있었다.



<Fig. 8> Average Speed(SIM vs DQL+SIM)



<Fig. 9> Average Waiting Time(SIM vs DQL+SIM)



<Fig. 10> Average Queue Length(SIM vs DQL+SIM)

V. 결 론

본 연구는 심층강화학습을 통해 자율주행차량의 최적경로탐색을 링크 단위의 거시적인 관점이 아닌 차량

추종과 차로변경을 기반으로 한 미시적인 관점에서 경로탐색을 수행하였다. 동적으로 변화하는 환경에서 자율주행차량의 횡방향·종방향 동작제어를 통해 주변 교통 상황과 선행하는 차량과의 거리를 고려하여 충돌이나 지체 없이 주행할 수 있도록 하였다. Deep Q-Learning기반의 심층강화학습 실행 결과 총 평균 누적 보상은 일부 변동은 있었으나 전반적으로 증가하는 추세를 보였으며 에이전트가 보상을 최대화하기 위해 성공적으로 학습하였다는 것을 보여주었다. 또한 총 평균 누적 손실은 특정 값에 도달하면 수렴하였으며 안전성을 나타내는 충돌률은 지속적으로 0%로 나타났다. 즉 학습 과정 동안 선행차량과 후행차량 간의 충돌은 발생하지 않았다. 학습된 모델을 통해 본 연구에서는 시뮬레이션 기반 차로변경의 경우와 Deep Q-Learning을 적용한 시뮬레이션 기반 차로변경의 경우로 구분하여 평균 통행 속도, 평균 대기 시간, 평균 대기행렬 길이와 같은 교통류의 이동성에 대한 결과 분석을 진행하였다. 평균 통행 속도는 교차로 신호 체계에 영향을 받아 두 경우 모두 유사한 패턴을 보였으며 Deep Q-Learning을 적용한 경우가 약 40% 향상된 것을 볼 수 있었으며 평균 대기 시간은 시뮬레이션 시간 45-46초부터 기록되었으며 Deep Q-Learning을 적용한 경우가 약 2초 감소하였다. 평균 대기행렬 길이는 평균 대기 시간과 비슷한 패턴을 보이며 Deep Q-Learning을 적용한 경우의 평균 대기행렬 길이가 약 2.3배 감소한 것을 볼 수 있었다. 본 연구는 자율주행차량의 상용화 시대에 맞춰 심층강화학습 기법을 적용하여 차로변경을 통한 경로탐색의 효율성을 입증하였다. 특히 ODD 체크리스트에 부합하는 공간적 범위 선정, 네트워크 구축 시 정밀도로지도 사용 등 실제 자율주행차량의 주행환경에 필수적인 요소들을 연구 내용에 포함시켰다는 점에서 다른 연구들과의 차별성을 가진다. 그러나 실제 교통 상황에서는 이동성 분석 뿐만 아니라 안전성에 대한 고려도 필요하다. 향후 연구에서는 본 연구의 내용을 바탕으로 다양한 경우의 시나리오와 환경에서의 학습이 필요하며 안전성에 대한 평가지표도 추가할 예정이다.

REFERENCES

- Chang, K. J. and Yoo, S. M.(2021), “A Study on Autonomous Vehicle Lane Change Method Using Cooperative Maneuver”, *The Korea Contents Association*, vol. 21, no. 1, pp.139-146.
- Choi, Y. G., Lim, K. I. and Kim, J. H.(2015), “Lane Change and Path Planning of Autonomous Vehicles using GIS”, *12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp.163-166. doi: 10.1109/URAI.2015.7358855
- Deutsches Zentrum für Luft- und Raumfahrt e.V.(DLR)(1998), “Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics”, vol. 55, no. 5, p.5597.
- Eclipse SUMO, <https://sumo.dlr.de/docs/Simulation/SublaneModel.html>, 2022.10.18.
- Eclipse SUMO, <https://sumo.dlr.de/docs/TraCI.html>, 2022.11.16.
- Erdmann, J.(2015), *Modeling Mobility with Open Data, SUMO's lane-changing model*, Springer, pp.105-123.
- Kang, M. S., Yi, I. K., Cho, Y. S. and Shin, O. S.(2021), “Route Selection and Speed Control of Autonomous Vehicles Based on Decentralized Deep Reinforcement Learning”, *The Korean Institute of Communications and Information Sciences*, pp.1501-1502.
- Kwon, J. U., Chae, M. S., Cho, E. Y. and Cho, S. Y.(2022), “Path Planning Algorithm based on Dijkstra Using HD-map”, *The Korean Society of Automotive Engineers*, pp.455-457.
- Li, Z., Liang, H., Zhao, P., Wang, S. and Zhu, H.(2020), “Efficient Lane Change Path Planning based on Quintic spline for Autonomous Vehicles”, *2020 IEEE International Conference on Mechatronics*

- and Automation (ICMA)*, pp.338-344. doi: 10.1109/ICMA49215.2020.9233841
- Lu, Q. and Tettamanti, T.(2018), “Impacts of autonomous vehicles on the urban fundamental diagram”, *5th International Conference on Road and Rail Infrastructure, CETRA 2018*, pp.1265-1271. doi: 10.5592/co/cetra.2018.714
- Ministry of Land, Infrastructure and Transport(MOLIT), http://www.molit.go.kr/USR/NEWS/m_71/dtl.jsp?lcmspage=1&id=95087208, 2022.10.21.
- Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M.(2013), *Playing atari with deep reinforcement learning*, arXiv preprint arXiv:1312.5602.
- National Highway Traffic Safety Administration(NHTSA)(2018), *A Framework for Automated Driving System Testable Cases and Scenarios*, pp.1-162.
- Shim, Y. B., Kim, Y. J., Min, K. W., Lee, S. Y. and Son, H. S.(2020), “A Study on the HD Map Editor for the HD Map based Path Planning”, *The Korean Society of Automotive Engineers*, pp.513-517.
- Song, J., Wu, Y., Xu, Z. and Lin, X.(2014), “Research on car-following model based on SUMO”, *The 7th IEEE/International Conference on Advanced Infocomm Technology, IEEE*, pp.47-55.
- Université Gustave Eiffel(2021), *Reinforcement Learning Project*, pp.1-12.
- Wang, F. J., Lu, Y., Dai, H. I. and Han, H. H.(2021), “Evaluation of Freeway Traffic Management and Control Measures Based on SUMO”, *Journal of Physics: Conference Series*, vol. 1910, 012044. doi: 10.1088/1742-6596/1910/1/012044
- Wang, P., Chan, C. Y. and De La Fortelle, A.(2018), “A Reinforcement Learning Based Approach for Automated Lane Change Maneuvers”, *Proc. IEEE Intelligent Vehicles Symposium*, vol. 2018, no. 2, pp.1379-1384.
- Ye, F., Cheng, X., Wang, P., Chan, C. Y. and Zhang, J.(2020.10), “Automated Lane Change Strategy using Proximal Policy Optimization-based Deep Reinforcement Learning”, *Proc. IEEE Intelligent Vehicles Symposium*, pp.1746-1752.