

# LAB MEETING

국민대학교 지능형 차량 신호 처리 연구실 학부연구생 김지원

2024.04.09(수)



국민대학교  
KOOKMIN UNIVERSITY

# DDPG 알고리즘을 활용한 차선 유지 제어 강화학습



국민대학교  
KOOKMIN UNIVERSITY

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

## Task 변경 (Traffic 없는 차선 유지 제어)

## Task 변경 이유

1. 많은 Traffic으로 인해 광범위한 고려사항 및 환경 설정 발생.
2. Lidar Sensor을 활용하지 않고 UAQ를 통한 데이터 추출.
3. 크기가 크지만 다양성이 없는 Observation 설정.

## 주요 변경 사항

- Traffic  
17대 -> 0대
- Observation  
35data -> 9data
- Observation 다양성  
1개(traffic 정보) -> 4개(왼쪽 차선 정보, 오른쪽 차선 정보, 차량 정보, 도로-차량 정보)
- Sensor  
Lidar(실질적으로 활용 x) & Collision -> Load & Line

## Observation 시각화

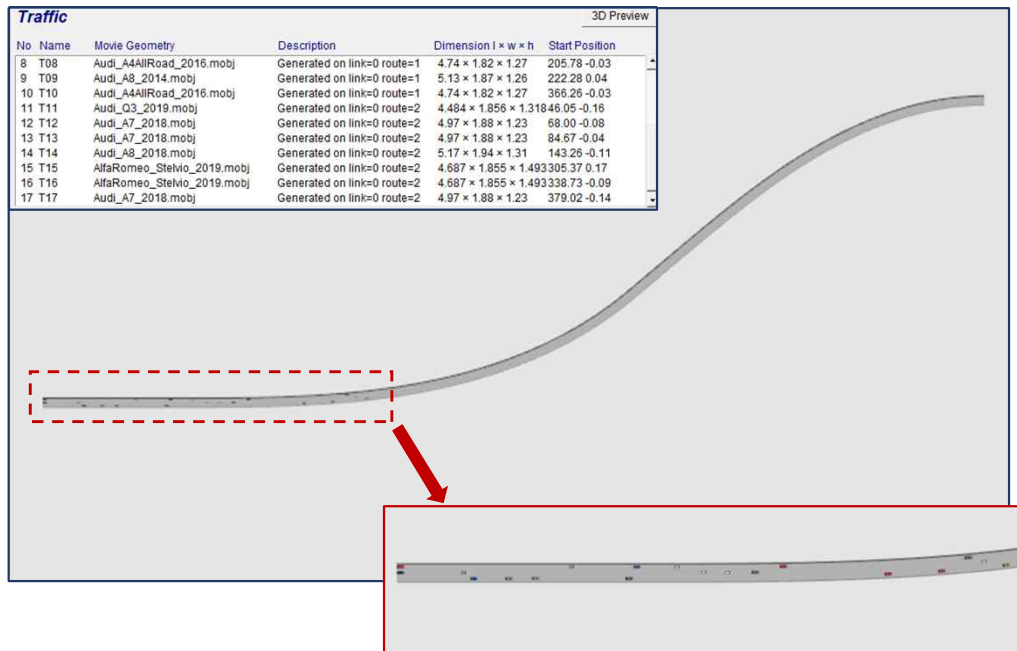
$$\begin{bmatrix} X_1 & Y_1 & Vx_1 & Vy_1 & Ax_1 & Ay_1 & Yaw_1 \\ X_2 & Y_2 & Vx_2 & Vy_2 & Ax_2 & Ay_2 & Yaw_2 \\ X_3 & Y_3 & Vx_3 & Vy_3 & Ax_3 & Ay_3 & Yaw_3 \\ X_4 & Y_4 & Vx_4 & Vy_4 & Ax_4 & Ay_4 & Yaw_4 \\ X_5 & Y_5 & Vx_5 & Vy_5 & Ax_5 & Ay_5 & Yaw_5 \end{bmatrix}$$



$$[Lx \quad Ly \quad Rx \quad Ry \quad X_{ego} \quad Y_{ego} \quad Dev_{avg} \quad Dev_{dist} \quad Curve]$$

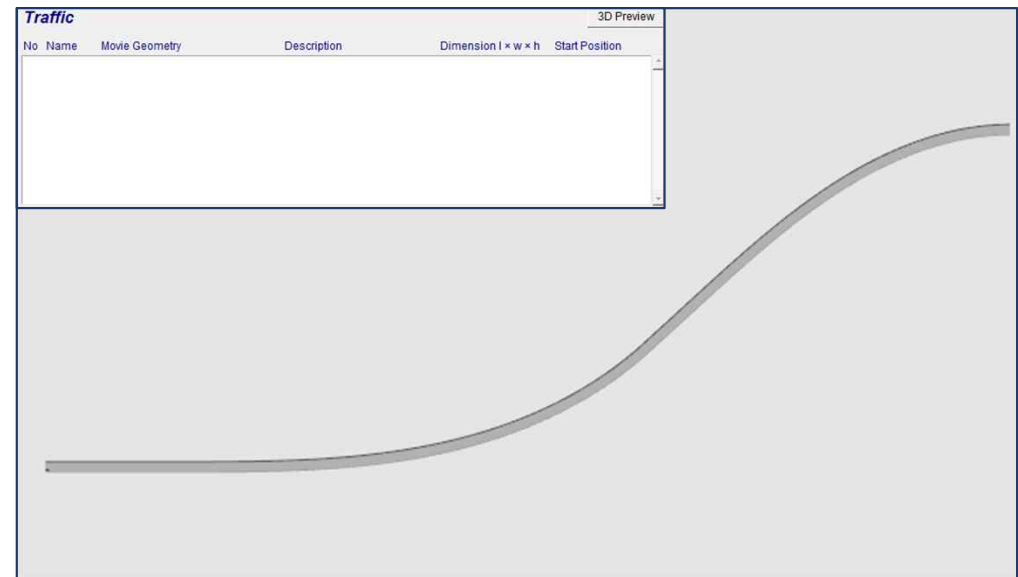
## DDPG 알고리즘 환경(Scenario)

## 변경 전



1. 총 17대의 Traffic 존재.
2. Traffic들의 정보는 RL Agent의 Observation data의 대부분을 차지.

## 변경 후



1. 기존 Traffic을 모두 제거함으로써 시나리오 간단화.  
-> 차선 유지 제어에 집중.
2. Observation 크기 간소화(35 -> 9).  
-> Sampling 효율 향상.

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

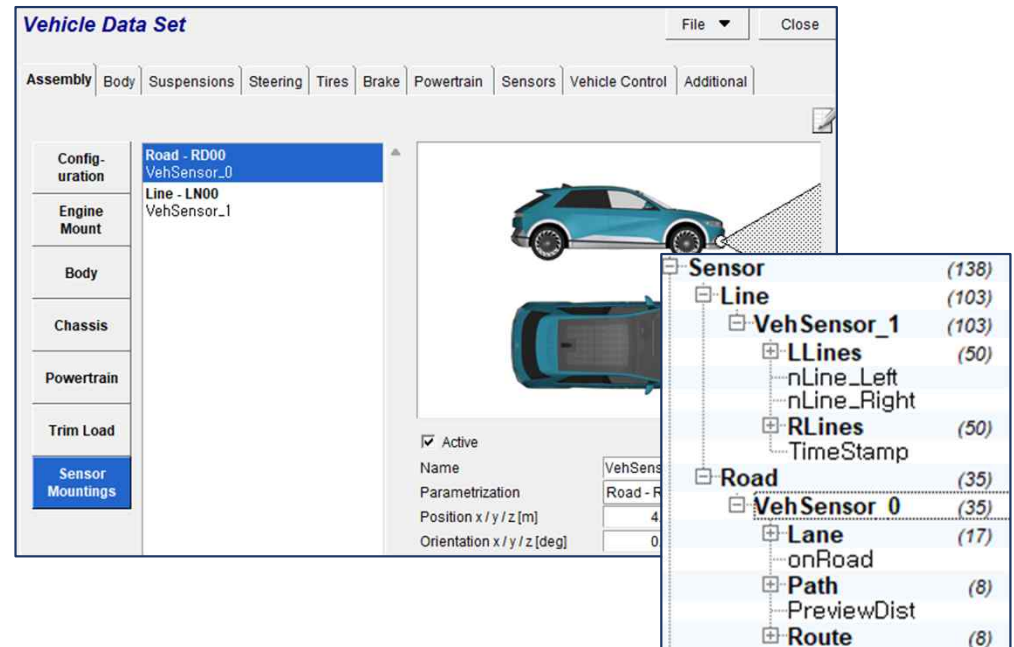
## DDPG 알고리즘 환경(Sensor)

## 변경 전



1. Lidar Sensor와 Collision Sensor 활용.
2. 실질적으로 Lidar Sensor 미사용.
3. Reward function과 Check if done에서 Collision Sensor parameter 활용.

## 변경 후



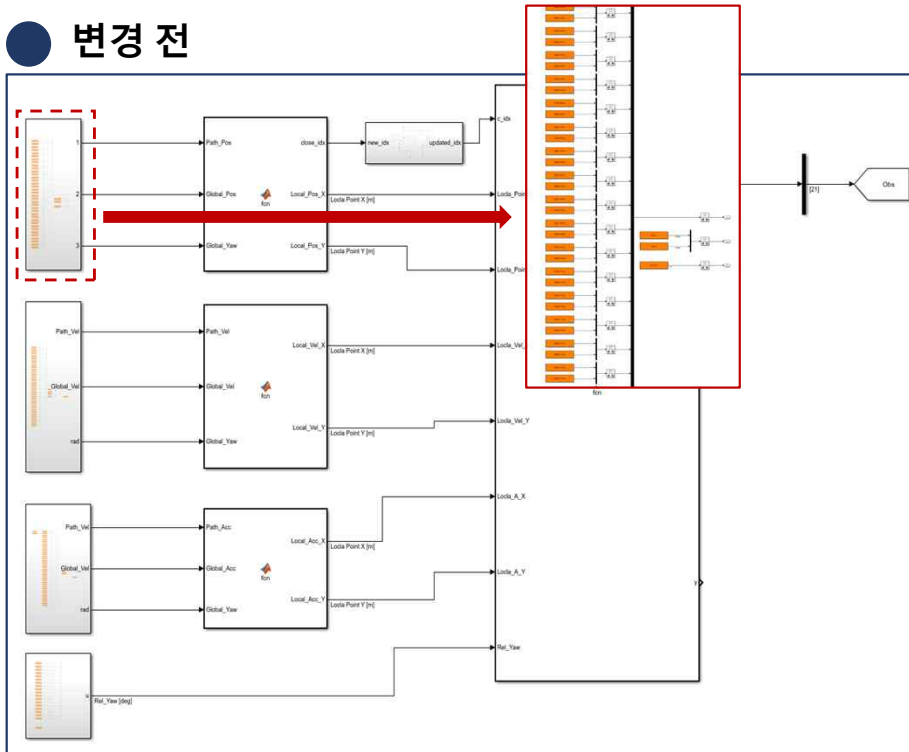
1. Road Sensor와 Line Sensor 활용.
2. 차선 위치, 도로 곡률, 도로-차량 이탈거리, 이탈각도 등 다양한 정보 추출 가능.
3. Observation 다양성 보장.

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

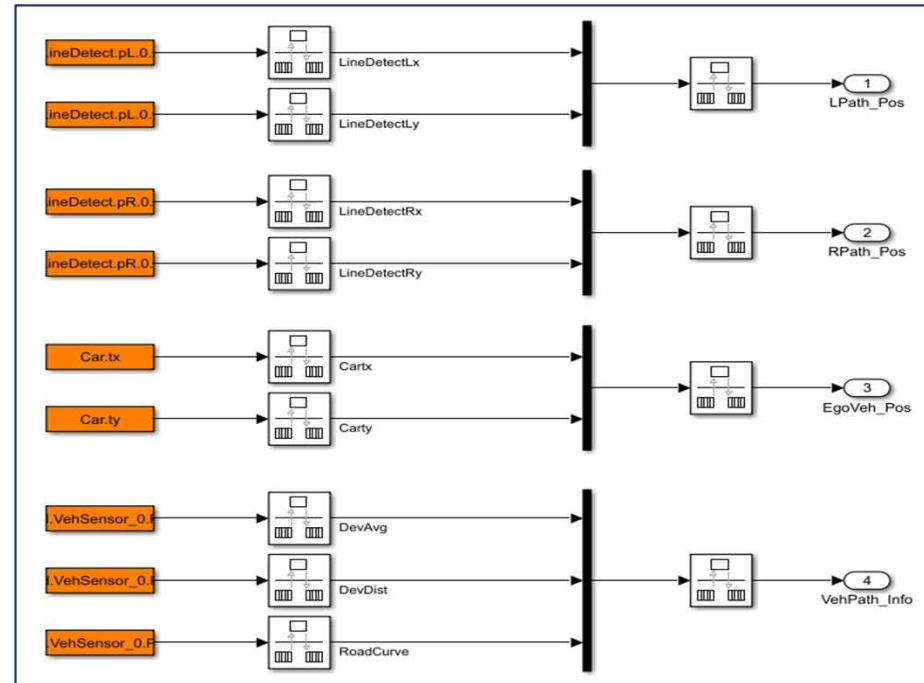
## DDPG 알고리즘 환경(Observation)

## 변경 전



1. Lidar Sensor로 추출된 파라미터가 아닌 Carmaker UAQ를 사용하여 Traffic들의 정보를 추출.
2. Observation 크기를 줄이고자 ego 차량과 가까운 5개의 traffic 정보만 Observation으로 활용.
3. 총 5 x 7(x, y, vx, vy, ax, ay, yaw), 35개의 observation을 사용.

## 변경 후



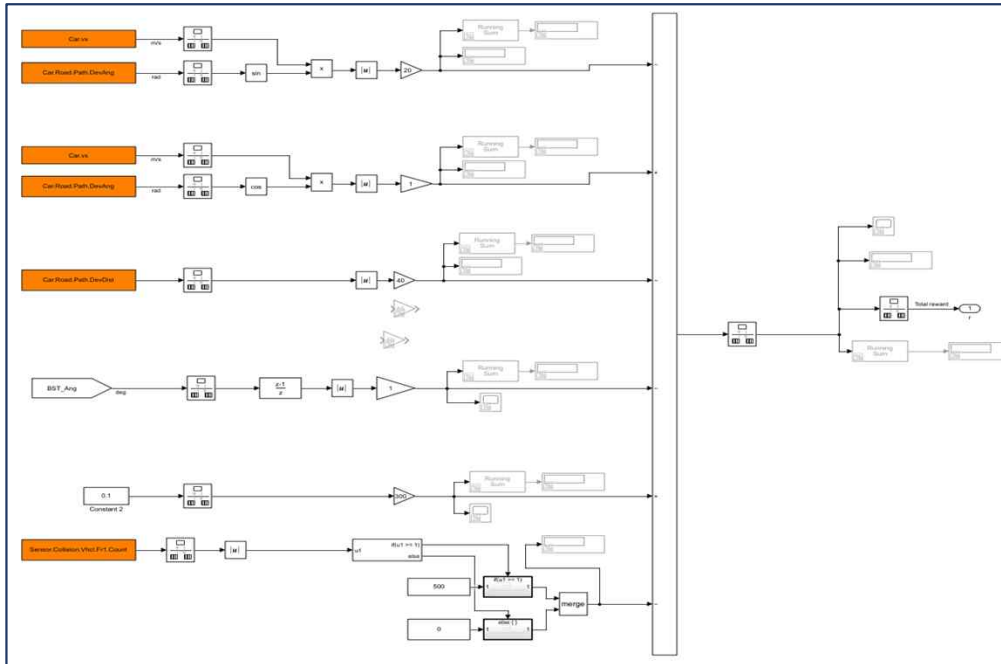
1. Line Sensor를 통해 왼쪽 차선, 오른쪽 차선 위치 정보를 추출.
2. Road Sensor를 통해 이탈 각도, 이탈 거리, 도로 곡률 정보를 추출.
3. 총 9(Lx, Ly, Rx, Ry, x, y, DevAng, DevDist, Curve)개의 Observation 활용.

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

## DDPG 알고리즘 환경(Reward function)

## 변경 전



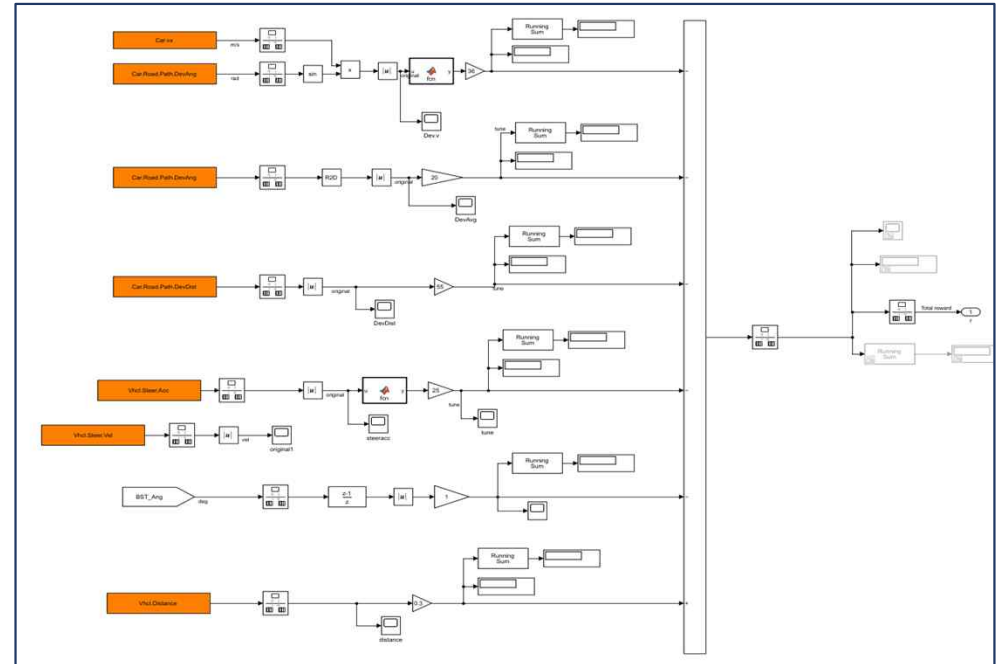
## 1. 보상함수 마이너스 인자

- 차량 이탈 속도
- 차량 이탈 각도
- 차량 이탈 거리
- 차량 충돌 여부

## 2. 보상함수 플러스 인자

- 상수

## 변경 후



## 1. 보상함수 마이너스 인자

- 차량 이탈 속도
- 차량 이탈 각도
- 차량 이탈 거리
- 차량 스티어링 휠 각가속도 (횡방향 제어 시, 승차감 및 안정성 보장)

## 2. 보상함수 플러스 인자

- 주행거리 (시나리오와 상관관계를 갖는 파라미터)

## Matlab function block in Reward function



차량 이탈 속도 연산 시, Car.tx를 시간에 대해 미분한 값인 Car.vx 사용.  
->도로 곡률과 차량 주행 곡률 사이에 delay 발생.

차량 이탈 속도 최대 허용 범위를 0.4m/s ~ 0.6m/s로 정의하는 차량 이탈 경고 시스템(LDWS) 성능 요구 사항 검증 논문을 참고하여 범위 설정.

## 차량 스티어링 휠 각가속도 인자

조향 시스템 반응성과 승차감 사이의 균형을 분석하는 논문에서 자율주행 차량의 조향 각가속도를  $500 \sim 700 \text{ deg/s}^2$  이하로 설정하는 것을 권장.

500[deg/s^2]는 Simulink 환경에서 단위 환산 시 0.87[-]임을 참고하여 범위 설정.

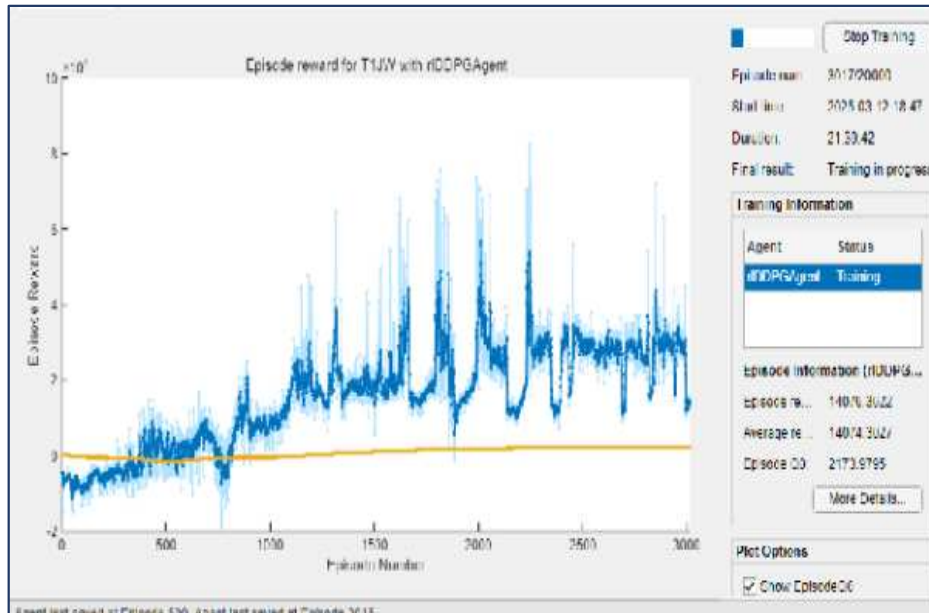


## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

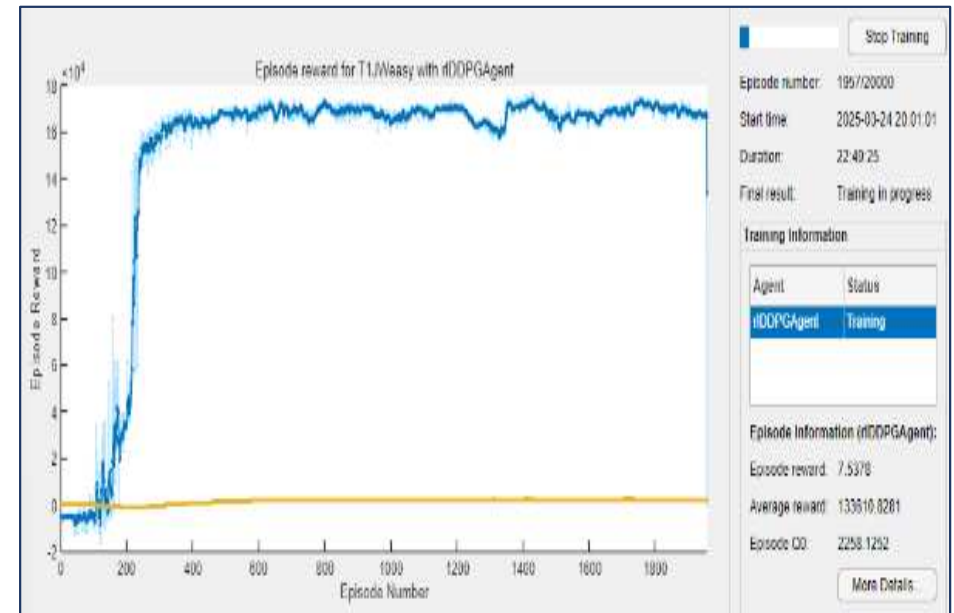
## DDPG 알고리즘을 활용한 차선 유지 제어 강화학습

## ■ 변경 전, 후 보상 그래프 비교

## ● 변경 전



## ● 변경 후



## ● 변경 전 -&gt; 변경 후

1. 에피소드 진행 속도: 1.7x 배속 -> 2.8x배속
2. 약 18,000회 에피소드 동안 최적 경로 도달 x -> 약 200회 만에 최적경로 도달 및 1.2km 시나리오 완주
3. 충분한 에피소드가 진행되어도 주기적인 차선 이탈 발생 -> 최적 경로 도달 후, 차선 이탈 완전 소멸

## 4월 개인연구 계획

### 강화학습

1. 횡방향 오차 최소화 (Pure Pursuit, Stanley)
  - Normalization 적용 (Reward function, Network)
  - Reward function, Observation 고도화
2. SAC, TD3 활용 및 DDPG 알고리즘과 성능 비교

감사합니다.

국민대학교 지능형 차량 신호 처리 연구실 학부연구생 김지원

2024.04.09(목)



국민대학교  
KOOKMIN UNIVERSITY