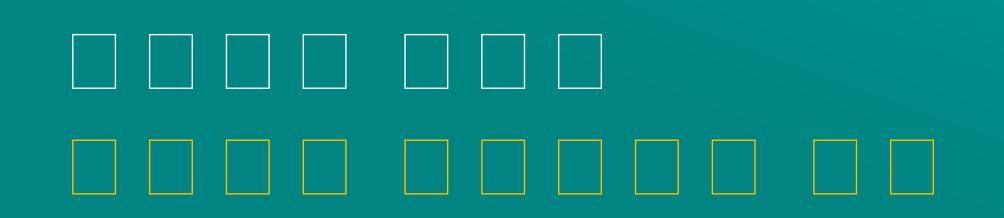
2023학년도 1학기 소프트웨어캡스톤디자인 경진대회





пппп

본 과제의 목적은 유니티를 활용한 자율주행 시뮬레이터 개발로, 크게 시뮬레이션 환경 구축, 알고리즘 개발 및 검증이란 두 가지 단계로 나누어 과제를 진행한다. 시뮬레이션 환경 구축에서는 차선, 교차로, 차량 물리 등과 같은 환경적인 요소를 구현해 다양한 도로 상황을 재현하며, 알고리즘 개발 및 검증에서는 자율주행 연구에서 많이 사용되는 강화학습을 이용해 알고리즘을 개발한다. 시뮬레이션 환경 구현을 위한 개발 플랫폼으로 유니티를 선정한 이유는 실제 물리 법칙을 반영할 수 있으며 강화학습이 진행되는 Python 환경과 쉽게 통신할 수 있다는 장점을 가지기 때문이다. 따라서, 강화학습에 필요한 데이터는 유니티 환경을 통해 수집되며, 파이썬 환경으로 전달되어 모델이 적절한 결정을 내릴 수 있도록 학습된다. 결과적으로, 본 과제의 자율주행 환경이 강화학습 알고리즘을 잘 학습할 수 있는 환경인지를 검증함과 동시에, 알고리즘 자체의 성능도 평가하는 과정을 거치며 시뮬레이터 개발을 진행한다.

학습에 사용되는 도로 환경은 도로의 커브 구간이 90도, S자 커브로 구분되는 2종의 도로가 있으며, 정체나 사고환경을 가정한 장애물(차량)을 추가해 총 4종의 도로로 제작되었다. 학습의 주체인 차량에는 총 5개의 Raycast 센서가 있으며 전방 2개, 좌/우 각 1개, 후방 1개로 구성된다. 이 센서들은 현실에서의 LiDAR 센서와 카메라를 대체하며, 연산량을 줄여 학습 속도를 개선할 수 있도록 도와준다. 또한 차량에는 타이어, 서스펜션 등의 물리 모델이 구현되어 있어 실제 차량과 유사한 움직임을 내도록 구현되었다.

강화학습의 알고리즘으로는 PPO(Proximal Policy Optimization)을 선정하였다. PPO는 에피소드를 수집하는 Policy와 업데이트되는 Policy가 동일한 On-Policy 알고리즘으로, 실시간 학습이 가능해 많은 양의 데이터를 저장하지 않아도 된다는 장점이 있다. 또한, PPO는 기존 강화학습 알고리즘에 비해 알고리즘의 연산 및 구현이 쉽고, Policy의 급격한 변화를 방지하는 안정적인 업데이트 방식을 통해 다양한 강화학습 문제에 대해서 전반적으로 좋은 성능을 가진다는 특징이 있다.

따라서, 본 과제에서는 PPO의 하이퍼파라미터와 차량의 센서 성능 값을 수정하며 자율주행 모델의 성능을 높이는 단계를 거쳤다. 자율주행 모델은 움직이는 물체를 탐지하는 모델(PoliceAgent)과 그렇지 않은 모델(CarAgent)로 나누어 총 2가지의 모델을 학습하였으며, 각각 150,000,000 steps, 50,000,000 steps의 학습 시간을 두어 복잡한 환경일수록 더 많은 시간동안 학습할 수 있게 설정하였다. 결과적으로 4종의 트랙에서 전반적으로 높은 보상을 받을 수 있는 모델을 개발할 수 있었으며, 특히 움직이는 물체를 탐지하는 PoliceAgent 모델의 경우, 주행 중 또 다른 차량이 나타나면 추월을 시도하거나 서행하는 등의 움직임을 보였다.

본 과제는 자율주행 솔루션을 제공하는 방식의 사업화를 기대할 수 있다. 먼저, 개인이나 연구 단체의 경우에는 고가의 차량, 센서 등을 구매하는 초기 비용이 부담스럽고, 기업에 비해 개발에 참여하는 인원이 적어 장비의 유지/보수 과정에서 많은 시간이 소요될 수 있으므로 적은 초기 비용과 빠른 시제품 개발을 위한 활용 방안을 고려할 수 있다. 다음으로 기업의 경우, 더 현실적인 물리 법칙을 반영한 환경을 제공하는 자율주행 SW개발 플랫폼으로의 활용을 고려할 수 있다. 본 과제가 지속해서 발전한다면 유니티의 높은 개발 접근성이란 장점을 살려 자동차 시뮬레이션 환경에 국한되지 않고 비행기, 기차, 선박 등 다양한 교통수단의 자율주행 분야로 확장된 통합 시뮬레이션 환경으로의 발전을 기대할 수 있다.