CMPE 260 Final Project, Milestone 1

Name of Leader: Tomy Tran
Other team members: Huang Teng-Wei, Suh Young Joon

One sentence summary

An RL system to perform speed efficient lane changes in a 2D driving world

Proposed project and its novelty, 30%

We propose to build an RL system that will perform speed efficient lane changes in a 2D driving world. The goal is to create a system that can control a simulated driver that will perform safe lane changes while minimizing time braking or slowing down. We will apply Policy Gradients into two formats, one standard implementation and one with exploration additions.

In our simulation, the agent is a vehicle with the goal of reaching a designated destination. TorchDriveSim provides various objects to simulate a realistic road environment, including other vehicles, pedestrians, cyclists, and traffic control elements. Additionally, we can customize these objects with different kinematic models to increase road diversity. For example, a randomly crossing dog could be introduced to add unpredictability, or a work zone could be set up to merge traffic from two lanes into one.

Many researchers have applied reinforcement learning to autonomous driving, in 2D and 3D with basic and more advanced algorithms. Some examples include, enhanced deep Q learning for basic autonomous 2D driving [1], A Decision-Making Strategy for Vehicle Autonomous Braking in Emergency via Deep Reinforcement Learning [2]. More similarly, one group has published Lane-changing system based on deep Q-learning with a request–respond Mechanism [3]. However, to our knowledge no publications exist that consider lane changing to minimize speed decreases and with policy gradient algorithms

Possible datasets/simulators, 25% of grade

Dataset/simulator name	**Contents with examples**	**Pros**	**Cons**
TorchDriveSim	See Figure 1 below	Written to be compatible with pytorch	Doesn’t consider 3D or complex driving models

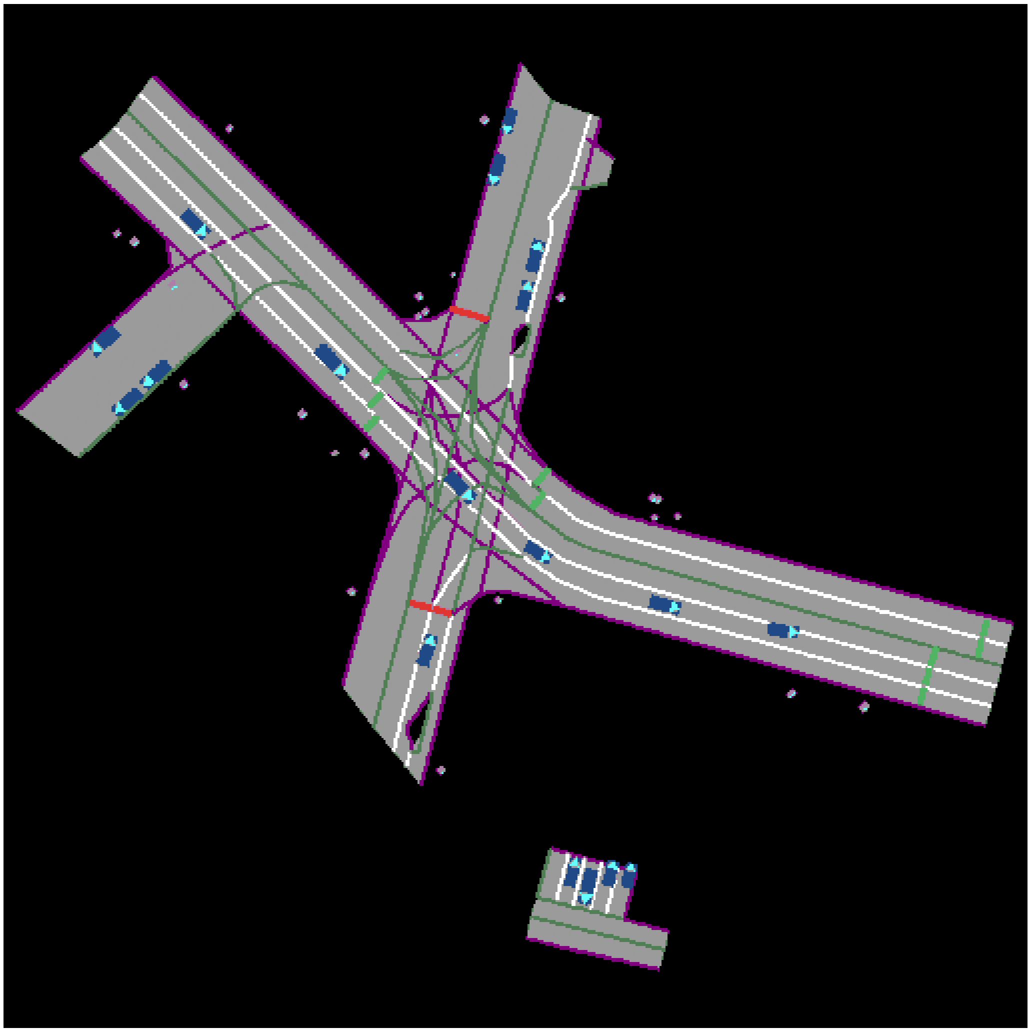

Figure 1. Example image of the torchdrivesim environment

Possible algorithms or methods, 25% of grade

Model name	Description	Reward	Pros	Cons	Notes
Policy Gradient	The model will create different trajectories on a simulated driving environment and then try	Positive for staying on track Positive if arriving time is lower than prediction time	Can deal with continuous action spaces like steering wheel angle and speed	Exploration may be an issue Very noisy reward seeking	N/A

	to perform gradient ascent on those data	Positive if lane changed Positive is driving smooth Negative if speed slows down before goal Negative if collision Negative if breaks laws			
Policy Gradient with exploration boost	The model will create different trajectories on a simulated driving environment with an exploration boost and then try to perform gradient ascent on those data	Positive for staying on track Positive if arriving time is lower than prediction time Positive if lane changed Positive is driving smooth Negative if speed slows down before goal Negative if collision Negative if breaks laws	Can deal with continuous action spaces like steering wheel angle and speed Will help to explore lane changes well	Might take a long time to converge Very noisy reward seeking	N/A

Summary of work plan, 20% of grade

We will be training our agent on the TorchDriveSim 2D simulated driving environment with one car model. Later work can focus on applications to other simulators or other car models.

The algorithms we will try are standard Policy Gradient and Policy Gradient with an exploration boost. Later work can focus on implementing something like A2C or PPO. We will compare the models in terms of reward and convergence time.

We will consider NN hyperparameter tuning as well as exploration factors. We will also consider the qualitative aspect of the agent's driving behavior.

References

[1] https://arxiv.org/html/2402.08780v1

[2]
https://ieeexplore.ieee.org/abstract/document/9067008?casa_token=YPD8wSslKWoAAAAA:g32hoyddvbNAufd1Css1ubsvV9ygwplaJMikbsSWyMo33MF7G2pwjQGZdBcHHUK7xl5CzM5iwbE

[3] https://www.sciencedirect.com/science/article/pii/S095741742301744X