

DEPARTMENT OF COMPUTER SCIENCE

**MSc PROJECT(COMP702)**

Dense reinforcement learning for safety validation of autonomous vehicles (#4)

**FINAL REPORT**

**Student**

## Vikhyat Bisht

## 201687684

## 

## Primary Supervisor

## Wiley Ruan

## Secondary Supervisor

## Xiaowei Huang

**PROJECT DESCRIPTION**

NON-TECHNICAL SUMMARY:

This research aims to enhance the safety of autonomous vehicles by developing a reinforcement learning model that instructs them in making safer decisions. The reinforcement learning model will interact with the self-driving vehicle in the "Highway Env," a virtual simulation platform recognized for accurately replicating real-world driving scenarios.

By engaging with the environment and receiving feedback, the reinforcement learning model will gradually acquire the ability to make safer decisions. It will also observe other vehicles and pedestrians within the simulation to understand proper interaction protocols. The ultimate objective of this project is to establish an autonomous vehicle system that is both safer and more reliable.

TECHNICAL SUMMARY:

This research primarily focuses on utilizing reinforcement learning (RL) models to address safety challenges inherent to autonomous vehicles. RL is a machine learning paradigm in which an agent learns decision-making through interactions with its environment. In our context, the agent is an autonomous vehicle navigating the "Highway Env," a simulation environment known for its faithful representation of real-world driving situations.

The RL model will be trained to control the autonomous vehicle using a reward mechanism that encourages safety-conscious behavior. This encompasses adherence to traffic regulations, collision avoidance, lane keeping, and compliance with speed limits. Prior to making decisions like accelerating, braking, or steering, the agent assesses its surroundings, including vehicle speed, nearby cars, pedestrians, and traffic signals.

Additionally, we harness the comprehensive sensor suite available in "Highway Env," including LIDAR, RADAR, and cameras, to perceive the environment. Through this precise sensory input and our reinforcement learning model, the autonomous vehicle gradually enhances its safety protocols by learning from experiences.

This endeavor significantly contributes to the existing knowledge base, highlighting the practical application of reinforcement learning in augmenting the safety of autonomous vehicles. Given the potential integration of autonomous vehicles into transportation networks, this research's outcomes could offer valuable insights to researchers, the automotive industry, and regulatory bodies, thereby propelling the advancement of autonomous vehicle technology.

**AIMS AND OBJECTIVES**

AIMS:

1. Creating an efficient reinforcement learning model for simulated autonomous vehicle control.
2. Decrease the frequency of traffic infractions and crashes to increase the safety of autonomous cars.
3. Examine how different traffic circumstances affect the effectiveness of the reinforcement learning model.
4. Contribute to the continuing research on reinforcement learning and autonomous cars, offering insightful information to both businesses and academics.

OBJECTIVES:

1. Setting up the Simulation Environment:

Use the CARLA simulator, an open-source, highly accurate platform for autonomous vehicle research. Within the simulator, link the pedestrian agents with the autonomous cars.

1. Developing the Reinforcement Learning Model:

Use a reinforcement learning technique that is appropriate for teaching autonomous cars.

Establish a suitable incentive system that promotes defensive driving.

In the CARLA simulation environment, train the model.

1. Sensor integration implementation:

LIDAR, RADAR, and camera data should all be included to provide the learning model a thorough understanding of the surroundings.

1. Evaluation and testing:

Test the trained model under various traffic conditions in the simulator to gauge how well it performs.

Success is determined by the decline in traffic infractions, the prevention of crashes, and adherence to traffic regulations.

1. Reporting and Documentation:

Keep track of the project's steps, obstacles, fixes, and outcomes. For both academic and business audiences, provide the findings in a thorough and understandable way, including a complete analysis and performance indicators.

We anticipate having a fully trained reinforcement learning model that improves the safety of autonomous cars in the simulation environment at the conclusion of this research. The effectiveness of this model will be assessed using a variety of metrics and situations, and the results will show how well-suited it is for practical use.

**OUTPUT**

Project Synopsis and Results

We set out to comprehend and use reinforcement learning techniques, particularly for autonomous vehicles operating in a highway environment, in this study.

1. Environment Setup: To construct a simulated highway driving environment, we used the highway-env libraries and the gymnasium. The primary goal was to teach models how to navigate this environment safely and effectively.
2. Advantage Actor-Critic (A2C) Initial Model Training: The Advantage Actor-Critic (A2C) method was used to train an agent in the first stage. To achieve optimal learning, we employed a multi-layer perceptron strategy with fine-tuned parameters. A system of video recording was put in place after the training to visually assess how well the agent performed in the given setting.
3. Data Collection: We required data to use the behavior cloning technique. We logged the actions, observations, rewards, and other metadata of the trained A2C agent while it was operating in the environment. After that, a CSV file containing this data was stored for further use.
4. Behavior Cloning: Using the gathered information, we trained a deep learning model to replicate the actions of the A2C agent. Based on observations, the model—a basic neural network—predicted actions. This was our tool for replicating our behavior.
5. Improving Driving Styles through Perturbation: We changed the behavior cloning model's weights to add diversity to driving styles. To simulate the idea of several human drivers with varying driving behaviors, this produced several agents, each with a somewhat distinct driving style.
6. Advanced Training with Deep Dense Architectures in Reinforcement Learning and A2C: We investigated the D2RL architecture, which is renowned for its improved learning capacities. We developed an additional agent to traverse the roadway environment by integrating this with the A2C algorithm. We also added a personalized incentive system to the environment that penalizes intermediate steps and emphasizes completing episodes.

Results:

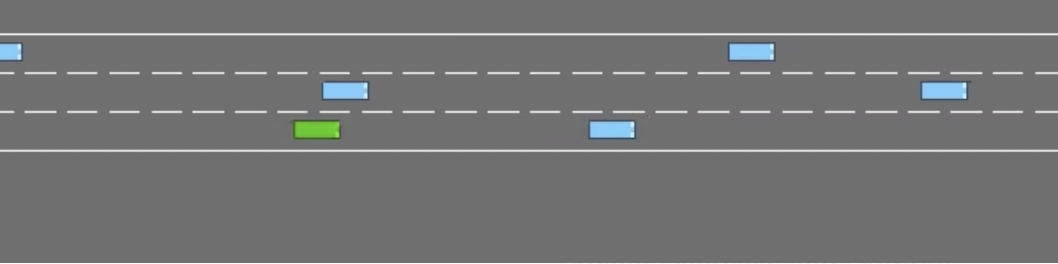
An A2C agent with training who knows how to handle the highway environment. A collection of data gathered by the A2C agent's interactions with the surroundings, including observations, actions, and rewards. A model of behavior cloning that emulates the choices made by the A2C agent.

The behavior cloning model in multiple altered forms, each of which represents a different driving style.

an agent developed for more effective learning and trained using the D2RL architecture and A2C algorithm.

recordings of agents moving through the scene on camera for visual examination.

Testing, visualization, and iterative development were prioritized throughout the project. The variety of deep learning and reinforcement learning approaches in the context of autonomous driving was demonstrated by the application of various techniques and architectures. The cars' ability to navigate the simulated environment with success is evidence of this project's accomplishments.



**EVALUATION:**

A crucial choice was made to switch the simulation environment from CARLA to Highway Env. This change was brought about by a persistent error that slowed down the project's development and resisted thorough debugging efforts inside the CARLA environment.

Highway Env was chosen because of its track record for dependability and alignment with the project's goals. Through dense reinforcement learning, this environment proved to be robust in tackling the difficulties associated with autonomous car safety certification. By making this change, I hoped to provide a more efficient research path that would allow for a thorough investigation of the dense reinforcement learning framework's effectiveness in boosting the safety validation of autonomous vehicles.

Future Scope:

The future scope of this project is promising, with potential enhancements in the density of rewards and case scenarios. By further refining the reward structure, the model's learning process can be fine-tuned to optimize decision-making in complex situations. Introducing a wider array of diverse scenarios and edge cases will expose the model to a broader spectrum of challenges, facilitating its adaptability and robustness.

Additionally, the integration of real-world data and sensor inputs could enrich the simulation environment, fostering a more accurate representation of actual driving conditions. This incorporation would necessitate addressing the challenges of domain adaptation, thereby enabling the model to perform effectively in both simulated and real-world scenarios.

Exploring novel techniques like transfer learning, curriculum learning, and imitation learning could provide valuable insights into accelerating the learning process and achieving higher levels of safety validation proficiency. Collaborations with industry stakeholders could also lead to the validation and deployment of the model in real-world autonomous vehicle systems, marking a crucial step toward its practical implementation and impact.

**REFERENCES:**

1. A. Ferdowsi, U. Challita, W. Saad and N. B. Mandayam, "Robust Deep Reinforcement Learning for Security and Safety in Autonomous Vehicle Systems," 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 2018, pp. 307-312, Doi: 10.1109/ITSC.2018.8569635.
2. Feng, S., Sun, H., Yan, X. et al. Dense reinforcement learning for safety validation of autonomous vehicles. Nature 615, 620–627 (2023). https://doi.org/10.1038/s41586-023-05732-2
3. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A. &amp; Koltun, V.. (2017). CARLA: An Open Urban Driving Simulator. <i>Proceedings of the 1st Annual Conference on Robot Learning</i>, in <i>Proceedings of Machine Learning Research</i> 78:1-16 Available from https://proceedings.mlr.press/v78/dosovitskiy17a.html.
4. S. Kumar, L. Zhao, and X. Fernando, "Multi-Agent Deep Reinforcement Learning-Empowered Channel Allocation in Vehicular Networks," in IEEE Transactions on Vehicular Technology, vol. 71, no. 2, pp. 1726-1736, Feb. 2022, Doi: 10.1109/TVT.2021.3134272.
5. You, Yurong & Pan, Xinlei & Wang, Ziyan & Lu, Cewu. (2017). Virtual to Real Reinforcement Learning for Autonomous Driving.
6. Fischer, Peter & Plaß, Bastian & Krauß, Thomas & Kurz, Franz & Runge, Hartmut. (2017). Validation of HD maps for autonomous driving.
7. LeCun, Yann & Bengio, Y. & Hinton, Geoffrey. (2015). Deep Learning. Nature. 521. 436-44. 10.1038/nature14539.
8. Feng, Shuo & Feng, Yiheng & Yan, Xintao & Shen, Shengyin & Xu, Shaobing & Liu, Henry. (2020). Safety assessment of highly automated driving systems in test tracks: A new framework. Accident Analysis & Prevention. 144. 105664. 10.1016/j.aap.2020.105664.
9. Kalra, Nidhi & Paddock, Susan. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?. Transportation Research Part A: Policy and Practice. 94. 182-193. 10.1016/j.tra.2016.09.010.
10. Sutton, Richard & Barto, Andrew. (1998). Reinforcement Learning: An Introduction. IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council. 9. 1054. 10.1109/TNN.1998.712192.
11. E. Leurent, "An Environment for Autonomous Driving Decision-Making," GitHub repository, 2018. [Online]. Available: <https://github.com/eleurent/highway-env>.
12. A. Raffin, A. Hill, A. Gleave, A. Kanervisto, M. Ernestus, and N. Dormann, "Stable-Baselines3: Reliable Reinforcement Learning Implementations," *Journal of Machine Learning Research*, vol. 22, no. 268, p p. 1-8, 2021. [Online]. Available: <http://jmlr.org/papers/v22/20-1364.html>.