

Autonomous Vehicle Navigation in Highway with Deep Q-Network (DQN) using Reinforcement Learning Approach

Sumiya Tamanna Fujita

Faculty of Manufacturing and Mechatronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Malaysia

Sarah Atifah Binti Saruchi

Faculty of Manufacturing and Mechatronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Malaysia

Ammar A.M. Al-Talib

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

Nurbaiti Wahid

School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 23000 Dungun, Terengganu, Malaysia

Alvi Khan Chowdhury

Monash University Malaysia, Jalan Lagoon Selatan, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

Siti Nurhafizza Maidin

Universiti Teknologi MARA, 23000 Dungun, Terengganu, Malaysia

Email: MIZ24006@adab.umpsa.edu.my, Sarahatifah@umpsa.edu.my, ammart@ucsiuniversity.edu.my, nurbaiti@uitm.edu.my, alvi.khanchowdhury@monash.edu, 2022987123@student.uitm.edu.my

Abstract

Autonomous vehicles (AVs) operate in highly dynamic environments, making them essential in modern transportation systems. Their success depends on real-time decision-making in unpredictable traffic scenarios, which are often beyond the scope of initial design assumptions. This unpredictability limits the effectiveness of traditional rule-based decision-making systems and predefined cost functions for real-time optimization. In critical applications like autonomous driving, reinforcement learning (RL) agents without safety mechanisms often struggle to converge or require extensive training data to produce reliable policies, leading to challenges in achieving safe and efficient operation. To address these challenges, this paper proposes a reinforcement learning (RL)-based framework. The ego vehicle refines its decision-making abilities by interacting with a simulated traffic environment. A short-horizon safety mechanism (SM) is integrated to ensure safer training by providing alternative safe actions during critical scenarios. The RL agent employs a deep neural network to map system states to optimal actions. The SM generalizes risky states, such as near-misses or collisions, rainy environment during night while creating a stable learning environment that enhances learning efficiency and enables meaningful exploration for optimal policy development. The method was validated in a highway driving scenario with varying traffic densities using the DQN algorithm and the CARLA simulator. Results demonstrated that the integration of the safety mechanism significantly improved learning efficiency and enabled the AV to make safe, reliable decisions even in complex and unpredictable traffic conditions.

Keywords: Autonomous Vehicles (AVs), Reinforcement Learning (RL), Deep Q-Network (DQN) Safety Mechanism (SM), CARLA Simulator, Highway Driving Simulation

1. Introduction

This research explores the application of the Deep Q-Network (DQN) algorithm for autonomous vehicle (AV) agents to navigate in highways in adverse weather conditions such as heavy rain. It is extremely challenging because of the reduced visibility, sensor noise, and unpredictable road dynamics. By training the DQN [1], [2], [4], [18] algorithm in simulated environment that replicates such harsh weather conditions, the AV agent

learns to make real-time decisions for lane-keeping, obstacle avoidance, and speed optimization. In order to address a significant gap in the state of autonomous driving technology, the contribution consists of proving how reinforcement learning can effectively increase the resilience and adaptability of AV systems in the face of severe environmental uncertainties.

Many businesses and scientific research organizations are very interested in autonomous driving because of its enormous potential to change current mobility and, most

importantly, reduce the frequency of traffic accidents [1] [6]. In recent years, autonomous driving has attracted a lot of interest from both the general public and researchers. Apart from their commercial potential, autonomous vehicles (AVs) have enormous potential to increase the number of available transportation options and potentially reduce carbon emissions [12].

Autonomous driving, like any other robotic system, involves decision-making. Implementing AV decision making usually involves resolving a number of path planning issues. A lot of calculation is needed for this method, and due to the solver's real-time limitations, the reference path that is generated may not result in a comfortable vehicle motion. In contrast, human drivers make decisions about changing lanes, maintaining speed, braking, and other maneuvers based on their perception of the traffic around them. Once this decision is made, the driver moves the steering and throttle/brake pedal to carry out the maneuver smoothly. The RL framework is a good fit for this approach [13].

The field of autonomous vehicle driving has advanced significantly because of advancements in computational tools like reinforcement learning (RL). One of the primary strategies for the creation of innovative transportation systems is RL's capacity to facilitate decision-making process adaptation through secure collaboration with the environment [5], [17]. The focus of basic RL research has been on low-dimensional problems. The algorithms work well in such predictable and transparent worlds. However, the shift to high-dimensional, complex scenarios, like those faced by self-driving cars, has prompted the creation of increasingly advanced techniques [8]. DQN, which was first presented as an expansion of conventional Q-learning, has since transformed the approach to resolving these difficulties [3].

In order to handle raw, high-dimensional sensory data, DQN uses deep learning technology. This allows for the direct learning of efficient policies from the complex input space without the need for expert features engineering. This ability introduced a new level of practicality concerning RL in autonomous driving as the network's ability to smartly convert continuous sensory streams into useful information is at the core of this sector [7]. DQN's success in a wide range of domains showed not only its versatility and tenacity but also its capacity to successfully negotiate the uncharted territory of the real driving world. In short, DQN has been shown to be able to learn a range of scenarios that can arise, from highway driving to urban congestion, and optimize both routes and driving styles [10].

A significant amount of training data may be needed for any RL methodology that attempts to solve both decision making and actuation at the same time. By clearly defining a hierarchy between high-level decision-making

[1] and low-level actuation, we streamline this process in this work. For low-level actuation, traditional feedback control techniques like PID are employed, while RL is used to solve the decision-making problem. The approach described here somewhat resembles hierarchical reinforcement learning, in which control and decision-making are learned concurrently under a set of related but distinct cost functions [16]. We employ a modified version of the DQN algorithm from [13] to train the decision-making agent.

The trade-off between exploration and exploitation is a crucial feature that sets reinforcement learning apart from other types of machine learning [11]. To get a better reward, an agent requires to try events it hasn't tried before. When learning on a physical platform, like a robotic platform, the agent will therefore attempt every possible action during the initial learning phase. Unfortunately, this curiosity can be expensive and hazardous. Unguided exploration, for instance, may frequently result in collision or near-miss scenarios when training a highway driving agent. This could cause simulation resets, which would slow down the learning process. Furthermore, the trained agent may choose a non-safe maneuver even after convergence because of the function approximation made by the Q-network. It addresses these problems by adding an explicit short-horizon safety check to the DQN decision maker, which is utilized during the implementation phase as well as during learning.

The remainder of the paper is organized as follows. Section 1 presents related works. Section 2 describes the entire methodology of the proposed RL model. The experimental simulation and results are provided in Section 3. Finally, section 4 concludes the whole work.

2. Methodology

This section establishes an RL model for highway autonomous driving. Fig. 1 illustrates the RL process. An agent that interacts with the environment based on observations makes up an RL model. Through interaction with the traffic environment, the agent learns how to maximize the reward function by selecting the best course of action. In the sections that follow, the formulations of the main components of the RL model are displayed in detail. Autonomous driving at highway is regarded as an MDP [15]. MDP is a mathematical decision-making framework, which is described by (S, A, T, R, γ) are as Table 1.

Table 1: Symbol and their description

| Symbol | Description |
|----------|---|
| S | The state space. S_t is the state at time t ($S_t \in S$); |
| A | The action space. a_t is the action at time t ($a_t \in A$); |
| T | The transition model. It describes the transition probability from one state to another. |
| R | The reward function. R_t is the reward for deploying action a_t at state S_t . |
| γ | The discount factor, $\gamma \in [0,1]$. It is used to calculate the cumulative expected reward. |

In RL model, the environment returns a numerical reward from a given reward function R based on the current state and the action that the agent takes. The goal of RL model is to learn an optimal strategy by maximizing the sum of discounted future rewards [1].

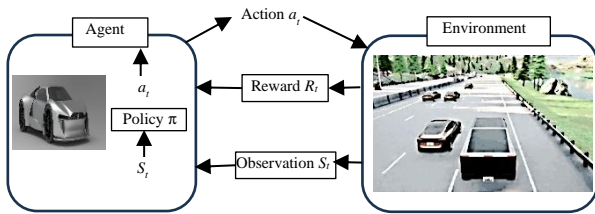


Fig.1 RL model for autonomous driving at highway

2.1 State Space

In our reinforcement learning simulation using CARLA, defining observations plays a critical role in creating a well-defined and precise state space. A single RGB front-facing camera was installed on the autonomous vehicle to capture the front road environment [14]. The images collected by this camera provide essential information about the surroundings. The action space is discrete, consisting of three actions: stop, drive slowly, and drive quickly.

2.2 Action Space

When driving straight across the highway, a vehicle can adjust throttle, steer and brake operations to accelerate, decelerate, or keep a constant velocity to ensure safety according to the surrounding traffic environment [9]. Therefore, the action space for intersection driving policy learning was defined as three velocity-related operations which has been given in the Table 2 below,

Table 2: Action and their description and values

| Action | Throttle | Steer | Brake |
|--------|----------|-------|-------|
| 0 | 0.0 | 0.0 | 0.5 |
| 1 | 0.3 | 0.0 | 0.0 |
| 2 | 0.5 | 0.0 | 0.0 |

2.3 Reward

Once an action is selected, an AV will get a resulting reward. The aim of the RL model about highway driving during rainy weather is to learn an optimal driving policy by maximizing the expectation of the discounted future reward, which indicates that different reward functions will lead to different driving policies. As a result, designing the reward function properly is crucial to directing AVs toward improved learning outcomes. The decision-making process should consider two requirements: (1) avoid collision with other vehicle, (2) do not be too conservative and travel as fast as possible [1].

The simulation's reward system is intended to motivate the agent to behave in a safe and effective manner. If a disruptive action occurs, like a collision, the reward is reduced by 0.7 and the episode is instantly terminated. The agent also receives a penalty of 0.2 for being too close to barriers (less than 1 meter) and 0.2 for keeping a safe distance (1 meter or more). The duration of the episode is also limited, and it ends once the predefined time threshold is exceeded. Furthermore, the reward is influenced by the distance traveled by the agent: traveling less than 30 meters incurs a small penalty of 0.1, while covering a distance between 30 and 100 meters earns a significant reward of 0.7 and ends the episode. The camera input is normalized (divided by 255) for further processing. This structured reward system helps the agent learn safe navigation and optimal behavior within the environment.

3. Results and Discussion

These days, creating secure AV systems requires testing AV models in incredibly lifelike virtual environments. The intricacy of autonomous driving scenarios necessitates testing the suggested methods in simulators that are similar to the real world in order to confirm their effectiveness. In light of this, we test our model in CARLA Simulator [2], which is regarded as the most potent simulator due to its realistic environments and features and suitability for AV technology applications. This section presents the results obtained using our proposed approach. As we can see in Fig.2 and Fig.3 shows the trained DQN model average reward over 100,000 timesteps. It took 1 hour of time to train the model. Four-hundred other vehicles were used in the CARLA environment to complicate the environment for the AV agent. CARLA town-04 highway was utilised for DQN model training purpose.

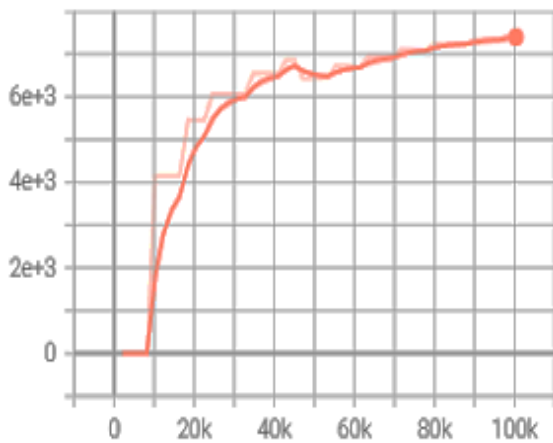


Fig.2 Mean Reward (Y-axis) with Timesteps Trained (X-axis)

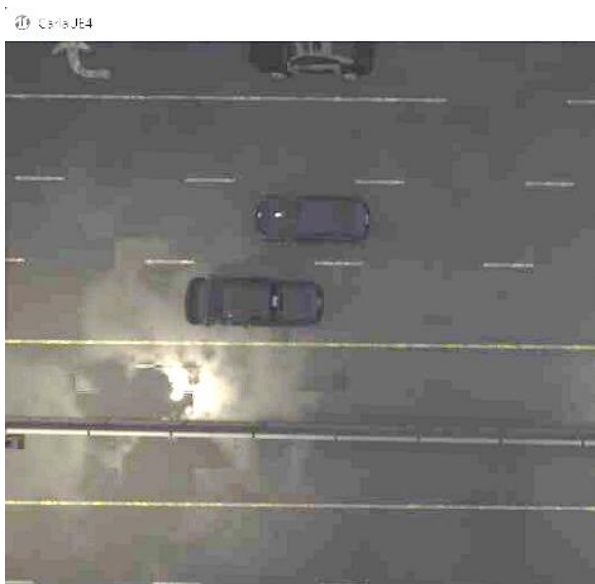


Fig.3. AV agent during rainy weather

To further complicate the training environment, soft rain during nighttime was used to train the model. The rise and consistency of the trained model reward with respect to training timesteps shows that the AV agent was able to drive well in this complicated scenario.

4. Conclusion

This paper presented a control architecture based deep RL framework for safe decision making in autonomous driving. By using well known feedback controller, high level decision by DQN is converted into a low-level actuation throttle, brake, and steering to control an AV in a complex scenario with dense traffic. We have evaluated the learned controller under varying traffic density, the results demonstrate the superior capabilities of the learned AV agent. Lastly, we have illustrated the continuous adaptation framework that has been proved to decrease the frequency of safety triggers

by altering the learning method. The CARLA simulator was used to validate the suggested approach, and the outcomes showed how effective our model is at learning, maintaining stability, and achieving rapid convergence. Given that DQN is among the data sets used in methods that are particularly impacted by noisy or insufficient data [10], we intend to expand this approach in the future to include required lane changes and on-ramp to off-ramp highway navigation.

Acknowledgements

This work was supported in part by Universiti Malaysia Pahang Al-Sultan Abdullah, Faculty of Manufacturing and Mechatronics Engineering Technology. The authors gratefully acknowledge the assistance and support provided throughout this research. Additionally, we extend our sincere gratitude to our co-authors for their valuable contributions to the successful formulation and development of this study.

References

1. Deep Reinforcement Learning Enabled Decision-Making for Autonomous Driving at Intersections Guofa Li^{1,2} · Shenglong Li¹ · Shen Li³ · Yechen Qin⁴ · Dongpu Cao² · Xingda Qu¹ · Bo Cheng⁵
2. A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in Conference on robot learning. PMLR, 2017, pp. 1–16.
3. J. Schulman et al., "Proximal Policy Optimization Algorithms," arXiv preprint arXiv:1707.06347, 2017.
4. M. Alshiekh, R. Bloem, R. Ehlers, B. K'önighofer, S. Niekum, and U. Topcu, "Safe reinforcement learning via shielding," arXiv preprint J. Garcia and F. Fern'andez, "A comprehensive survey on safe reinforcement learning," Journal of Machine Learning Research, vol. 16, no. 1, pp. 1437–1480, 2015. arXiv:1708.08611, 2017.
5. J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," Neural Networks, vol. 61, pp. 85–117, 2015.
6. K. Arulkumar, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 26–38, 2017.
7. L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 38, no. 2, pp. 156–172, 2008.
8. V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529–533, 2015.
9. J. Garcia and F. Fern'andez, "Safe exploration of state and action spaces in reinforcement learning," Journal of Artificial Intelligence Research, vol. 45, pp. 515–564, 2012.
10. D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," Nature, vol. 529, no. 7587, pp. 484–489, 2016.
11. R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction. MIT press Cambridge, 1998, vol. 1, no. 1.
12. B. Paden, M. Ć'ap, S. Z. Yong, D. Yershov, and E. Frazzoli, "A survey of motion planning and control

- techniques for self-driving urban vehicles,” *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 1, pp. 33–55, 2016.
13. Autonomous Highway Driving using Deep Reinforcement Learning, Subramanya Nageshrao¹ and Eric Tseng² and Dimitar Filev²
 14. H. Xu, Y. Gao, F. Yu, and T. Darrell, “End-to-end learning of driving models from large-scale video datasets,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2174–2182.
 15. Bellman, R.: A Markovian decision process. *J. Math. Mech.* 6(5), 679–684 (1957)
 16. K. Frans, J. Ho, X. Chen, P. Abbeel, and J. Schulman, “Meta learning shared hierarchies,” *arXiv preprint arXiv:1710.09767*, 2017.
 17. C. Chen, A. Seff, A. Kornhauser, and J. Xiao, “Deepdriving: Learning affordance for direct perception in autonomous driving,” in *Computer Vision (ICCV), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2722–2730.
 18. Arulkumaran, K., Deisenroth, M.P., Brundage, M., et al.: Deep reinforcement learning: a brief survey. *IEEE Signal Process. Mag.* 34(6), 26–38 (2017)

Authors Introduction

Sumiya Tamanna Fujita



She completed her Bachelor's Degree in Computer and Communication Systems Engineering in 2015 from Universiti Putra Malaysia. Presently, she is pursuing her Master's degree in the Faculty of Manufacturing and Mechatronics Engineering Technology at Universiti

Malaysia Pahang Al-Sultan Abdullah. Her research interests include planning and control, vehicle dynamics, autonomous vehicle systems, artificial intelligence, and robotics.

Sarah Atifah Binti Saruchi



She received her B.Eng. in Mechanical and Aerospace Engineering from Nagoya University, Japan. She received her Master and Doctoral degrees from Malaysia-Japan International Institute of Technology (MJIT), Universiti Teknologi Malaysia. Currently, she is working as senior lecturer at

Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia. Her research interests include machatronics and artificial intelligence.

Ammar Abdulazizf Al Talib



He has finished his B.Sc. and M.Sc. degrees in Mechanical Engineering from the University of Mosul/Iraq. He has finished his Ph.D. degree from UPM University/ Malaysia. He is also a chartered Engineer and Member of the Institute of Mechanical Engineers/UK.(CEng. MIMechE).

He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia, and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Nurbaiti Wahid



She received the M.Eng. in electrical engineering from Universiti Teknologi Malaysia in 2009 and Ph.D. degrees in advanced vehicle system from Universiti Teknologi Malaysia. Currently, she is a Senior Lecturer and Researcher at the College of Engineering, Universiti Teknologi MARA, Malaysia since 2009. Her research interests include planning and control, vehicle dynamics, autonomous vehicle systems, artificial intelligence, and robotics.

Alvi Khan Chowdhury



He received his Bachelor's Degree of Mechatronic Engineering in 2023 from UCSI University, Kuala Lumpur Campus. Currently, he is pursuing his Master's Degree of Engineering Research in Monash University Malaysia Branch.

Siti Nurhafizza Binti Maidin



She received her bachelor's degree in engineering in 2016 from Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia. She also had received her Master's degree in 2018 from Faculty of Advance Informatic School. She is currently a PhD student in Universiti Teknologi

MARA, Malaysia.
