Reinforcement learning in autonomous driving

Dantong Xiang

National Elite Institute of Engineering, Chongqing University, Chongqing, 400044, China

20203631@cqu.edu.cn

Abstract. Automatic driving technology has become a highly researched field in recent years, aiming to achieve vehicle driving without human intervention. In this regard, reinforcement learning techniques have played a crucial role. This study discusses and analyses the use of reinforcement learning in automatic driving methods. The research begins with the process of reinforcement learning. In the architectural framework, there is a special emphasis on designing innovative reward functions to encourage safe and socially acceptable driving behaviour, while considering uncertainty factors through advanced Bayesian neural networks. This paper primarily focuses on aspects such as scene understanding, localization and mapping, planning and driving strategies, and control. Furthermore, the paper analyses the key elements of automatic driving and delves into the specific complexities associated with each element. It highlights the utilization of reinforcement learning within the realm of autonomous driving. Reinforcement learning assists autonomous vehicles in understanding the surrounding environment, accurately identifying paths, making intelligent driving decisions, and safely controlling the vehicle. Reinforcement learning especially working with deep learning plays a crucial role in realizing and continuously improving automatic driving. Finally, the paper concludes with a summary and outlook on the entire study.

Keywords: Reinforcement Learning, Autonomous Driving, Deep Reinforcement Learning.

1. Introduction

Autonomous driving (AD) technology holds immense potential to revolutionize the transportation industry by enhancing safety, reducing traffic congestion, and providing increased mobility for all. Reinforcement Learning (RL) has surfaced as a prospective method to train autonomous vehicles to make complex decisions in real-world environments. This research proposal outlines a comprehensive study aimed at leveraging reinforcement learning techniques to enhance the decision-making capabilities of autonomous vehicles for safer and more efficient driving. Autonomous driving technology encompasses several key areas, including perception, decision-making and planning, control systems, localization and mapping and human-machine interaction interface.

RL has made remarkable strides in the field of autonomous driving research. One significant application area is in path planning and decision-making. Researchers have applied reinforcement learning algorithms to autonomous vehicles, enabling them to make intelligent decisions in complex and dynamic traffic environments, such as collision avoidance, optimizing vehicle speed, and making lane changes. A notable example of this approach is the utilization of deep reinforcement learning (DRL) to train autonomous vehicles for lane changes on highways, aimed at enhancing road traffic efficiency and

safety. Furthermore, RL has also made a significant impact on simulation training. Researchers leverage virtual environments to train autonomous driving algorithms, simulating various road conditions and driving scenarios. This approach can substantially reduce the number of trials required on actual roads, mitigating potential safety risks. Additionally, it aids in enhancing algorithm performance and robustness. Another application area is traffic flow optimization. By coordinating the driving of multiple autonomous vehicles, RL algorithms can assist in reducing traffic congestion and enhancing road utilization efficiency. Research indicates that in urban environments, the implementation of intelligent traffic signal control and vehicle-to-vehicle coordination through RL can significantly reduce traffic delays and emissions.

However, despite the extensive potential utilization of RL in the domain of autonomous driving, there are still challenges to address. For instance, RL algorithms often require a substantial amount of training data. Yet, collecting a large volume of real-world data in autonomous driving scenarios can pose challenges related to safety and cost. Additionally, the interpretability of RL models remains a significant concern. In the context of AD, understanding why a model makes specific decisions is crucial, highlighting the need for transparent and interpretable models.

This study discusses and analyses RL-based methods for AD. The focus is primarily on the application of RL in scene understanding, localization and mapping, planning and driving strategies, and control. Additionally, the paper analyses key elements of AD and delves into the specific complexities associated with each element.

2. Methodology

In this section, the main focus is on introducing the methodology of RL in autonomous vehicle technologies. The methodology of this paper involves a systematic and comprehensive approach targeted for enhancing the effectiveness of AD through the strategic integration of cutting-edge technologies. Starting with a meticulous definition of the state space and problem formulation as a Markov Decision Process (MDP), the research primarily centres on utilizing scene understanding, localization and mapping, planning and driving policy, and control. Within the architectural framework, special emphasis is placed on designing innovative reward functions to encourage safe and socially acceptable driving behaviour, while also considering uncertainties through advanced Bayesian neural networks. This framework undergoes meticulous optimization through rigorous simulations, emulating a range of real-world scenarios as a pre-training testing ground for autonomous agents. Subsequently, the agents transition to real-world experiments in controlled environments, ensuring that AI-driven decisions based on core domains effectively lead to tangible enhancements in autonomous driving performance.

2.1. Problem Formulation and State Space Representation

In this phase, the research will define the state space representation of the AD problem. This will involve selecting relevant sensor inputs such as camera images, LiDAR data, GPS coordinates, and vehicle dynamics information. The state space will also encompass contextual information like traffic signals, road signs, and the behaviour of surrounding vehicles and pedestrians. The problem will be formulated as an MDP, with the state, action, reward, and transition functions defined accordingly.

2.2. Reinforcement Learning Architecture

The research will design a hierarchical RL architecture to enable autonomous vehicles to make decisions at different levels of abstraction. The high-level component will focus on strategic maneuver planning, such as lane changes, overtaking, and merging. The low-level component will involve vehicle control actions like acceleration, braking, and steering. Various RL algorithms will be evaluated for each component, considering factors such as stability, convergence speed, and sample efficiency. The chosen algorithms might include Deep Q-Networks (DQN) for maneuver planning and Proximal Policy Optimization (PPO) for control.

2.3. Reward Function Design

This phase will focus on creating reward functions that effectively guide the RL agent's learning process. The research will explore novel reward designs that promote safe and efficient driving behavior. For instance, rewards could be based on adhering to traffic rules, maintaining a safe following distance, and minimizing jerk or sudden maneuvers. The rewards will also incorporate considerations for pedestrian interactions, yielding to other vehicles, and adapting to various road conditions. Socially acceptable driving behaviors will be encouraged through the reward structure.

2.4. Uncertainty Estimation and Risk-awareness

To handle real-world uncertainties, the research will integrate uncertainty estimation techniques into the RL framework. Bayesian neural networks or Monte Carlo dropout will be investigated to model sensor noise and estimation errors. The research will develop strategies to propagate uncertainty through the decision-making process, enabling the agent to make cautious and risk-aware decisions in situations where confidence is low. This will involve balancing exploration and exploitation while considering the level of uncertainty.

2.5. Simulation and Real-world Experiments

Simulation environments will be used extensively throughout the research. Initially, the RL agents will undergo pre-training in simulated environments to acquire basic driving skills and maneuver planning. High-fidelity simulators will replicate various driving scenarios, including urban, highway, and adverse weather conditions. The pre-trained agents will then undergo fine-tuning using reinforcement learning techniques. This phase will also involve creating realistic traffic scenarios, pedestrian behaviors, and dynamic road conditions to ensure robustness.

Once the agents exhibit satisfactory performance in simulations, they will be deployed onto a real-world autonomous vehicle platform. Controlled on-road experiments will be conducted in a safe and controlled environment. The vehicle will be equipped with the necessary sensors, computing hardware, and safety mechanisms to ensure compliance with regulations and ensure safe operation. The real-world experiments will provide validation and benchmarking of the RL agents' performance in actual driving scenarios.

3. Main Link of Autonomous Driving and Application of Reinforcement Learning

The chapter provides an overview of the key elements of AD and delves into the specific intricacies associated with each element. In this segment, it primarily describes the application of RL in the area of AD. RL helps autonomous vehicles understand their surroundings, find their way accurately, make smart driving decisions, and control the car safely. RL takes a vital important role that makes autonomous driving possible and keeps it improving.

RL finds applications in a diverse range of autonomous driving tasks, showcasing its adaptability and effectiveness. These tasks encompass controller optimization, path scheduling, and trajectory optimization, where RL tunes vehicle control for smooth and safe driving. RL also shines in motion planning, dynamic path planning, and complex navigation tasks, enabling vehicles to navigate challenging environments, handle intersections, and adapt to lane changes. Furthermore, RL-driven strategies learn from past experiences, improving performance in areas like merging, splitting, and highway driving. Additionally, RL is employed in reverse RL, fine-tuning reward functions using expert driving data. This versatile approach enhances AD systems' adaptability and decision-making in real-world scenarios [1]. Figure 1 shows the standard blocks of an autonomous driving system [2].
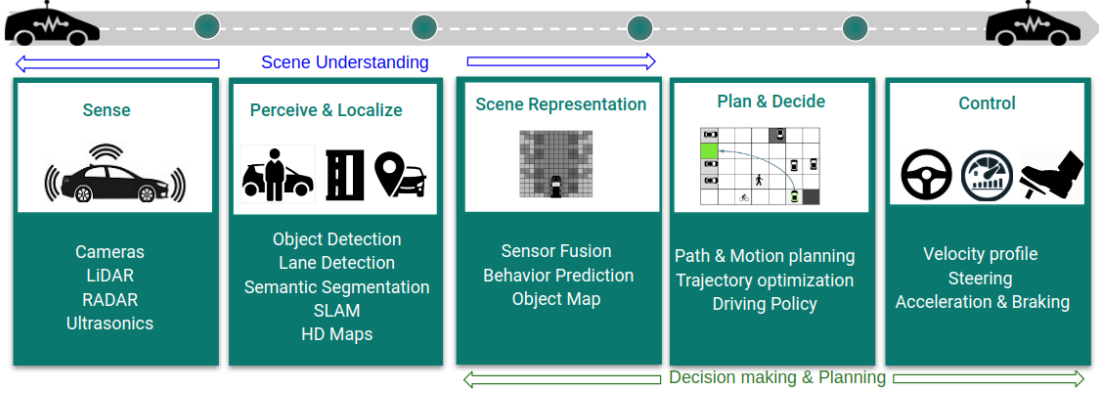
Figure 1. the standard blocks of an autonomous driving system [2].

3.1. Scene Understanding

This module involves the use of different sensors to get ambient data and processing the data with state-of-the-art computer vision and deep learning (DL) methods to identify and interpret objects, road conditions, and the surrounding environment, enabling the vehicle to make informed decisions. RL can be utilized to train agents for object recognition in the environment, enabling the vehicle to detect and monitor objects like pedestrians, vehicles, and obstacles accurately. Agents can learn to understand scenes by assigning rewards based on their ability to correctly identify objects and respond to them appropriately.

3.1.1. Object Detection.
In AD, object detection serves as the "eyes" of the vehicle, allowing it to perceive its environment and make informed decisions based on the presence and behaviour of objects. This information is then used by other components of the autonomous system, such as planning and control, to navigate safely, follow traffic rules, and avoid collisions. Object detection plays a critical role in achieving the perception capabilities required for fully autonomous vehicles.

The primary techniques for object detection in autonomous driving and computer vision are supervised learning and DL like YOLO (You Only Look Once) CNN algorithm [3]. However, RL still play a role in improving or augmenting object detection in specific scenarios. In a study [4], authors proposed two driving models to examine the impact of 3D dynamic object detection in AD. They then formulated an enhanced model that exhibited superior navigation performance and safety standards. They found out that the model called Conditional Imitation Learning Dynamic Objects Low Infractions-Reinforcement Learning (CILDOLI-RL) using Q-Learning and Deep Deterministic Policy Gradient (DDPG) functions better than the other model called Conditional Imitation Learning Dynamic Object (CILD) for safety-critical driving, like in dense traffic driving scenario for autonomous navigation with passenger.

3.1.2. Semantic Segmentation.
Semantic segmentation in AD pertains to a computer vision assignment where every pixel in an image or sensor data is categorized with a semantic label, such as "road," "vehicle," "pedestrian," "building," or "tree." This pixel-level labelling provides a detailed understanding of the scene's composition, allowing autonomous vehicles to discern different road elements and objects in their environment.

RL is not typically directly used semantic segmentation, while deep learning methods now have become the dominant approach for semantic segmentation due to their outstanding performance [5]. However, Reinforcement Learning somehow using as a supplement or improvement. Md. Alimoor Reza and Jana Kosecka proposed a semantic segmentation model for indoor environment using RL, which is a more modular and flexible than traditional monolithic multi-label CRF approach, cutting down the workload for data processing and maximize performance on the benchmark dataset [6].

3.1.3. Sensors Fusion. Autonomous vehicles are equipped with an array of sensors such as cameras, LiDAR (Light Detection and Ranging), radar, ultrasonic sensors, IMUs (Inertial Measurement Units) and GPS (Global Positioning System) (Figure 2). These sensors collect data regarding the vehicle's environment, encompassing other vehicles, pedestrians, traffic signs, and roadway state. Sensor fusion in AD is the process of combining data from multiple sensors, and more, to construct a comprehensive and accurate depiction of the vehicle's environment.
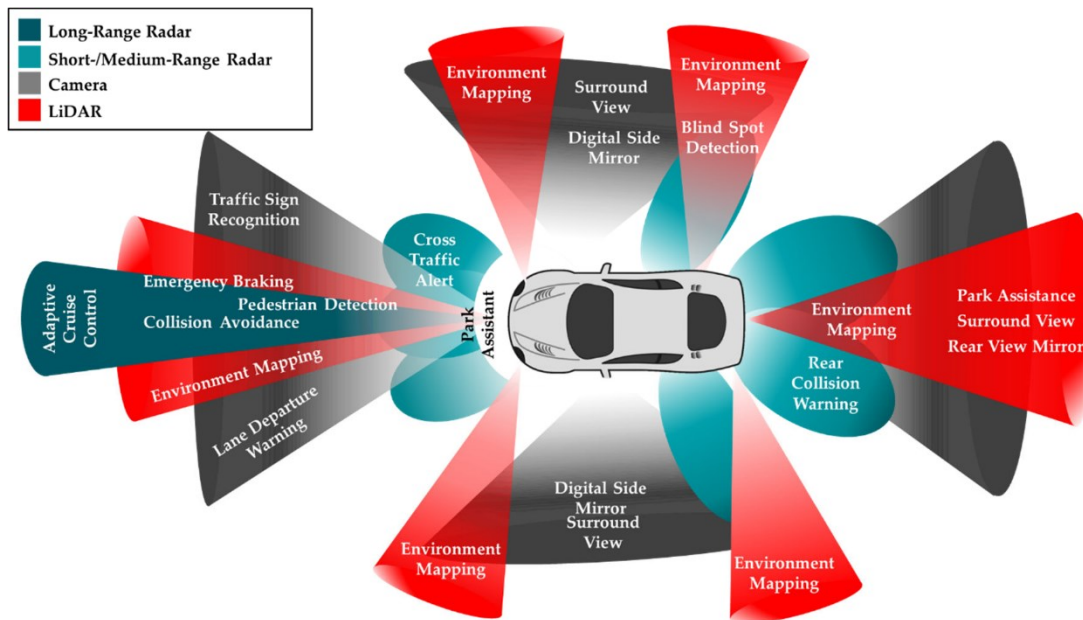
Figure 2. different sensors on the vehicles [7].

3.2. Localization and Mapping

In AD, localization and mapping are fundamental to safe and precise navigation. Localization provides the vehicle with its current position, which is critical for path planning and control. Mapping assists the vehicle in discerning the road layout, identifying the positions of other vehicles, and recognizing the existence of obstacles. Together, they enable the vehicle to make informed decisions, avoid collisions, follow traffic rules, and navigate complex environments. Both localization and mapping are dynamic processes, continually updated as the vehicle moves. The integration of sensor data, machine learning, and advanced algorithms is essential for achieving accurate and reliable results in real-world driving scenarios.

3.3. Planning and Driving Policy

Planning and Driving Policy in AD refer to the processes and strategies that enable a self-driving vehicle to make decisions and control its movements.

There are many examples using reinforcement learning in planning and mapping. For instant, Bojarski et al conducted research on end-to-end learning for self-driving cars, which uses DRL to learn a driving policy directly from raw sensor data [8]. Sallab et al proposed an AD framework using deep reinforcement learning, which use DQN to make planning. The utilization of DQN fasten convergence and improve performance [9]. In a study [10], authors used Q-learning to determine the most effective driving policy that helps the vehicles to change lane. They choose RL for convenience of connecting with perception module and avoiding building the environment model explicitly, taking into account all potential future scenarios. Wang et al [11] proposed a framework for autonomous robot navigation that integrates conventional global planning with DRL-driven local planning. This approach reduces training duration and mitigates the risk of the mobile robot becoming immobilized in the same place. In a study [12], authors proposed a lateral and longitudinal decision-making model based on QDN, Double

DQN(DDQN), and Dueling DQN. In comparative trials, DRL in AD consistently outperforms rule-based methods in safety, efficiency, and generalization. As ADVs increase in mixed traffic, training DRL models gets tougher, highlighting the need for future research in multi-agent RL. In the terms of motion planning, it still have a lot to improve like safety and Sim2Real [13].

3.4. Control
In AD, 'control' pertains to the precise management of a vehicle's movements, including steering, throttle (acceleration), and braking, to navigate the vehicle safely and efficiently. Control systems in autonomous vehicles are responsible for executing the driving policies and paths generated by higher-level planning algorithms.

And several examples of RL applications in control for AD. For instance, in a study [14], authors proposed two approaches (DQN, DDPG (Deep Deterministic Policy Gradient)) to navigate a dynamic urban environment in a simulation, adhering to a predefined route while minimizing collisions and staying on the road, aiming for maximum speed. The result shows that DDRG works better for its continuous control in both speed and steering. Liu et al [15] proposed a new longitudinal motion control method combining DRL with expert demonstrations in order to have better performance in urban driving situations. They compared the results with common RL and Imitation Learning (IL) baseline methods, showing a faster training speed and better safety and efficiency.

In a study [16], authors explored autonomous driving based on vision using DL and RL techniques, which split the vision-based lateral control system into two modules: One relying on neural networks employing multi-task learning is used to analyse driver-view images to predict track features (perception module), and another is rooted in reinforcement learning to utilize these features for control decisions (control module). In a research endeavour [17], the authors introduced a control method that relies on DRL, utilizing the MDP and an associated proximal policy optimization learning algorithm. Their objective was to enable autonomous exploration of a parking lot. The outcomes revealed the attainment of a highly effective controller after just a few hours of training.

In conclusion, the utilization of RL in AD, especially in the planning and control modules, usually work together with deep learning. This synergy harnesses the ability of RL to optimize decision-making through interaction with the environment and pairs it with the capacity of DL to process vast amounts of complex data. This combined approach is driving innovation in autonomous systems, enabling vehicles to adapt, learn, and make intelligent decisions on the road. As AD technology advances, the integration of RL and DL will continue to shape the landscape, promising safer, more efficient, and increasingly sophisticated autonomous transportation solutions. Although DRL had some limitations, such as a lack of continuity in speed and direction control and sim2real processing, it will still continue to advance the progress of self-driving technology in the future.

4. Conclusion
This study has investigated and assessed RL methods applied to AD. Its primary focus has revolved around scene comprehension, positioning and mapping, strategizing for planning and driving, and the aspect of control. The utilization of RL combining with DL in AD has been highlighted, aiding in environmental understanding, accurate path finding, intelligent driving decisions, and safe vehicle control. IT has emphasized the key elements and complexities associated with AD. The limitations like lack of continuity in speed and direction control and sim2real processing are mentioned. Future research can further delve into the reliability, safety, and adaptability aspects of AD, as well as optimize the efficiency and accuracy of driving strategies. Additionally, exploring the handling of more complex traffic scenarios and multi-vehicle cooperative driving can be an interesting direction for further investigation.

References
[1] Udugama, B.. (2023). Review of Deep Reinforcement Learning for Autonomous Driving. 10.48550/arXiv.2302.06370.

[2]   Kiran, Bangalore & Sobh, Ibrahim & Talpaert, Victor & Mannion, Patrick & Sallab, Ahmad & Yogamani, Senthil & Perez, Patrick. (2021). Deep Reinforcement Learning for Autonomous Driving: A Survey. IEEE Transactions on Intelligent Transportation Systems. pp. 1-18.

[3]   Balasubramaniam, Abhishek & Pasricha, Sudeep. (2022). Object Detection in Autonomous Vehicles: Status and Open Challenges. 12.21280/arXiv.2202.02729.

[4]   Wijesekara, P.A.D. Shehan. (2022). Deep 3D Dynamic Object Detection towards Successful and Safe Navigation for Full Autonomous Driving. The Open Transportation Journal. 16. 1-15.

[5]   Tao Shu. (2022). Research on Semantic Segmentation Algorithms in Autonomous Driving. 10.27379/d.cnki.gwhdu.2022.000834

[6]   Reza, Md & Košecká, Jana. (2016). Reinforcement Learning for Semantic Segmentation in Indoor Scenes. 10.2174/18744478-v16-e2208191.

[7]   Yeong, De Jong & Velasco-Hernandez, Gustavo & Barry, John & Walsh, Joseph. (2021). Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review. Sensors. 21. 2140. 10.3390/s21062140.

[8]   Bojarski, Mariusz & Testa, Davide & Dworakowski, Daniel & Firner, Bernhard & Flepp, Beat & Goyal, Prasoon & Jackel. (2016). End to End Learning for Self-Driving Cars. 10.2352/ISSN.2470-1173.2017.19.AVM-023.

[9]   Sallab, Ahmad & Abdou, Mohammed & Perot, Etienne & Yogamani, Senthil. (2017). Deep Reinforcement Learning framework for Autonomous Driving. Electronic Imaging. 2017. 70-76.

[10]   Wang, Pin & Chan, Ching-Yao & de La Fortelle, Arnaud. (2018). A Reinforcement Learning Based Approach for Automated Lane Change Maneuvers. 10.1109/IVS.2018.8500556.

[11]   Wang, Xuanzhi & Sun, Yankang & Xie, Yuyang & Bin, Jiang & Xiao, Jian. (2023). Deep reinforcement learning-aided autonomous navigation with landmark generators. Frontiers in Neurorobotics. 17. 10.3389/fnbot.2023.1200214.

[12]   Cui, Jianxun & Zhao, Boyuan & Qu, Mingcheng. (2023). An Integrated Lateral and Longitudinal Decision-Making Model for Autonomous Driving Based on Deep Reinforcement Learning. Journal of Advanced Transportation. 2023. 1-13.

[13]   Aradi, Szilárd. (2022). Survey of Deep Reinforcement Learning for Motion Planning of Autonomous Vehicles. IEEE Transactions on Intelligent Transportation Systems. 23. 740-759. 10.1109/TITS.2020.3024655.

[14]   Pérez-Gil, Ó., Barea, R., López-Guillén, E. et al. (2022) Deep reinforcement learning based control for Autonomous Vehicles in CARLA. Multimed Tools Appl 81, 3553–3576.

[15]   Liu, Haochen & Huang, Zhiyu & Wu, Jingda & Lv, Chen. (2022). Improved Deep Reinforcement Learning with Expert Demonstrations for Urban Autonomous Driving. 921-928. 10.1109/IV51971.2022.9827073.

[16]   Dong, Li & Zhao, Dongbin & Zhang, Qichao & Yaran, Chen. (2019). Reinforcement Learning and Deep Learning Based Lateral Control for Autonomous Driving. IEEE Computational Intelligence Magazine. 14. 83-98.

[17]   A. Folkers, M. Rick and C. Büskens, (2019) Controlling an Autonomous Vehicle with Deep Reinforcement Learning, 2019 IEEE Intelligent Vehicles Symposium (IV), 2025-2031.