

Autonomous Driving Control Strategy Based on Deep Reinforcement Learning

Liangjun Yang^{1,a,*}

ITrier College of Sustainable Technology, Yantai University, Yantai, China

a. hamlete@nwfsc.edu

**corresponding author*

Abstract: This paper discusses an autonomous driving control strategy based on Deep Reinforcement Learning (DRL), which aims to improve the decision-making ability of autonomous driving system in complex traffic environments. Deep reinforcement learning has a wide range of applications in many fields, such as robotics and medicine. Autonomous driving has emerged as a significant research focus in recent years. By combining deep learning and reinforcement learning, the model is able to autonomously learn and optimize driving behavior under dynamically changing road conditions. The DRL-based control strategy performs well in vehicle obstacle avoidance, pedestrian recognition, and traffic rule compliance in the face of complex environments such as city streets, intersections, and congested road sections, significantly improving the safety and efficiency of autonomous driving. This article will first introduce deep reinforcement learning. Then, the autonomous driving control strategy based on deep reinforcement learning is introduced. This research provides valuable insights for developing and implementing DRL-based autonomous driving systems.

Keywords: Autonomous Driving, Deep Reinforcement Learning, Markov Decision Process.

1. Introduction

In recent years, with the rapid development of artificial intelligence and automation technology, autonomous driving technology has become an important part of intelligent transportation systems. DRL effectively combines deep learning and reinforcement learning. This integration enables high-dimensional data processing and automatic feature extraction, minimizing manual intervention. This approach is particularly suitable for complex decision-making in autonomous driving systems because it is particularly suitable for control problems in dynamic environments by interacting with the environment to learn optimal decision-making strategies. Research indicates that autonomous vehicles can improve road safety while alleviating traffic congestion, reducing environmental pollution, and conserving fuel resources [1-3]. Traffic conditions in cities are often very complex, especially during the morning and evening rush hours. Autonomous driving systems need to face a variety of traffic participants, such as bicycles, motor vehicles, and pedestrians. At the same time, autonomous driving systems analyze and process complex traffic signals and massive amounts of traffic information. Such complexities pose higher demands on autonomous driving systems' real-time decision-making and adaptability. Traditional autonomous driving control strategies often rely on hand-designed rules and models, and the adaptability and flexibility of these strategies are

insufficient in the face of complex and changeable urban traffic environments. Therefore, control strategies based on Deep Reinforcement Learning (DRL) have emerged as the times require, and have become a powerful tool to solve various challenges in autonomous driving.

This article begins with an overview of DRL, covering its operating principles and modes. Markov decision-making process is also highlighted in this paper as an important concept used to describe reinforcement learning tasks. The autonomous driving control strategy based on deep reinforcement learning will be introduced in the second part of this paper. This section explains the fundamentals of autonomous driving systems, particularly sensor operation and data processing mechanisms. This article will focus on the combination of autonomous driving systems and deep reinforcement learning. At the same time, this section will also introduce the lateral motion control algorithm in the autonomous driving system.

2. Autonomous Driving Control Strategy Based on Deep Reinforcement Learning

In recent years, algorithm-based technology has been continuously advancing. As a result, the topic of autonomous driving systems (ADSs) has received widespread attention from researchers. Mature and efficient autonomous driving technology requires strong algorithms. The autonomous driving control strategies are a set of methods designed to ensure the safe and efficient operation of autonomous vehicles. One crucial control target is to minimize the difference between the dynamic position and attitude of the real car and those of the dynamic waypoint (virtual car). This enables the autonomous vehicle to accurately follow the path through the waypoint [4].

Deep learning and reinforcement learning (RL) are important branches in the field of artificial intelligence. While they have different applications and theoretical foundations, these two approaches can be combined to form deep reinforcement learning. For example, in some complex tasks, deep learning can be used to extract features, while reinforcement learning (RL) is responsible for decision-making and policy optimization. The Deep Reinforcement Learning model combines deep learning and reinforcement learning (RL) to solve more complex decision-making problems. The application of deep reinforcement learning to the autonomous driving control strategy can solve the problem of autonomous driving in the face of complex road conditions and dangerous driving environments. The main function of this approach is to utilize deep reinforcement learning methods based on sensing and vision. These methods are used to detect roads, vehicles, and pedestrians. By doing so, the system can avoid dangers and prevent traffic violations.

2.1. Deep reinforcement learning

A fundamental element of reinforcement learning (RL) is the agent's ability to learn effective behaviors. This process involves the gradual modification or acquisition of new skills and actions. Another critical feature of RL is its reliance on trial-and-error experiences, in contrast to approaches like dynamic programming, which operate under the assumption of having full prior knowledge of the environment. Consequently, an RL agent does not need complete information or control over its surroundings; it simply requires the capacity to interact with the environment and gather data. In offline scenarios, experiences are collected beforehand and then utilized as a batch for learning, which is why this approach is also referred to as batch RL [5].

Deep learning models are characterized by their use of multiple layers that represent data in various forms. Essentially, they consist of a series of foundational components, each serving a specific purpose. These components include autoencoders for data compression and feature learning, Restricted Boltzmann Machines (RBMs) for probabilistic modeling, and convolutional layers for processing grid-like data such as images. During the training phase, raw input data is introduced to a multi-layered network. Each layer performs nonlinear transformations on its inputs, generating

outputs that serve as inputs for subsequent layers in the deep architecture. The representation produced by the final layer can be utilized to build classifiers or applications that benefit from a more efficient and high-performing hierarchical abstraction of the data. This hierarchical abstraction allows the model to capture increasingly complex features at each layer, from simple edges to more intricate patterns. By applying nonlinear transformations to their inputs, each layer aims to learn and identify the underlying factors that explain the data. This iterative process ultimately results in the development of a hierarchy of increasingly abstract representations [6].

Markov decision process (MDP) (Figure 1) is an important part of reinforcement learning. It provides a mathematical framework for modeling decision-making in situations where outcomes are partly random and partly under the control of a decision-maker. Reinforcement Learning (RL) is an area of artificial intelligence influenced by behaviorist psychology. It operates on the principle of learning through trial and error while interacting with a stochastic environment. The foundation of RL lies in the concept of a Markov Decision Process (MDP), which addresses sequential decision-making problems. An MDP is formally defined by a 5-tuple. This tuple includes: a set of states and actions (S, A), a reward function (R), a state transition probability matrix (P), and a discount factor γ . The matrix P describes transitions from each states to its successor states s' , while γ falls within the range $[0, 1]$. This discount factor prioritizes more immediate rewards over those expected in the future. An environment qualifies as an MDP when the state S encompasses all the information necessary for the agent to make optimal decisions [7].

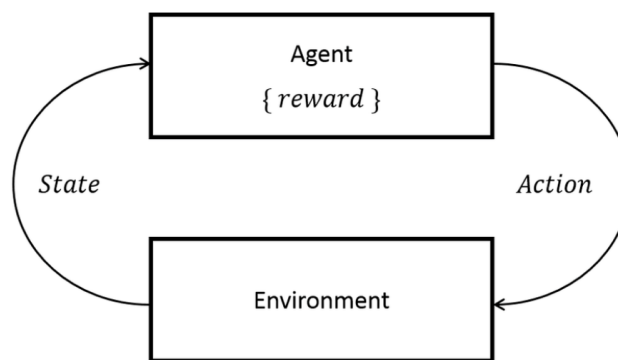


Figure 1: Markov decision process [7]

A reinforcement learning (RL) agent can be described using a Markov Decision Process (MDP). This situation is referred to as a finite MDP when both the state and action spaces are finite. Finite MDPs are critical in RL studies, with a significant portion of the literature assuming that the environment follows this structure. In the finite MDP framework, the RL agent interacts with the environment by executing actions, which lead to observations and rewards. Through these interactions, the agent learns from its experiences and enhances its decision-making strategies over time [6].

The article "A View on Deep Reinforcement Learning in System Optimization" reviews various efforts to apply deep reinforcement learning (DRL) in the field of system optimization and introduces a set of evaluation metrics [8]. It discusses the challenges associated with integrating DRL into systems and provides a detailed description of how to formulate the DRL environment. Additionally, the paper proposes a framework and toolkit for assessing DRL solutions, illustrated with concrete examples like DeepRM. The authors conclude that while DRL holds significant promise for system optimization, it also faces challenges such as low sample efficiency, instability, and limited generalization ability. Importantly, the article presents a novel set of key indicators for evaluating

DRL applications in system optimization, which is expected to enhance the standardization and reproducibility of future research in this area.

2.2. Autonomous driving control strategy based on deep reinforcement learning

Due to the limitations of traditional machine learning methods in implementing key functions of autonomous driving such as scene understanding and motion planning, researchers have shifted their focus to deep learning and reinforcement learning methods. This once again demonstrates the potential of deep reinforcement learning in enabling these functions. Autonomous driving systems rely on various sensors, such as cameras and radar, to collect real-time environmental information for identifying obstacles, traffic signs, and other vehicles. After processing this data, deep reinforcement learning (DRL) models use real-time information to learn through trial and error, optimizing driving strategies and decision-making processes. Sensors provide the essential environmental perception, while DRL intelligently controls the vehicle based on this information, enabling safe and efficient autonomous driving functionality.

Sensors are essential for measuring and detecting various environmental properties and tracking changes over time, providing the crucial input data for deep reinforcement learning models in autonomous driving systems. They can be classified into two main types: (1) Exteroceptive Sensors: These sensors gather information about the surrounding environment of the vehicle. Notable examples include cameras and Light Detection and Ranging (LiDAR), which help identify obstacles, road conditions, and other critical elements. (2) Proprioceptive Sensors: These sensors focus on the vehicle's internal parameters. This category encompasses systems such as the Global Navigation Satellite System (GNSS) for precise location tracking and wheel odometers that measure the distance traveled.

By integrating data from both types of sensors, autonomous vehicles can effectively understand their environment and maintain a comprehensive awareness of their own dynamics [9]. Autonomous driving systems necessitate the integration of data from various sensors. These sensors can be categorized based on the dimensionality of their data output: low-dimensional sensors, such as LiDAR, produce relatively simple data structures, while high-dimensional sensors, like cameras, generate complex, multi-layered data. Interestingly, while raw camera images are high-dimensional, the essential information for autonomous driving tasks is significantly lower in dimensionality. The critical components that influence driving decisions include moving vehicles, available free space on the road ahead, the location of curbs, and other relevant environmental factors. These key elements form the basis for the autonomous vehicle's decision-making process.

In this context, intricate details of vehicles are often irrelevant; only their spatial positions matter. As a result, the memory bandwidth required for pertinent information is considerably reduced. By effectively extracting this relevant data and eliminating non-essential information, this paper can enhance both the accuracy and efficiency of autonomous driving systems. This strategy also minimizes the computational and memory demands, which are vital for the embedded systems that will support the autonomous driving control unit. Overall, focusing on essential information optimizes processing capabilities and improves performance in real-time driving scenarios [10].

The primary objective of the perception module is to create an intermediate-level representation of the environment, such as a bird's-eye view map that illustrates all obstacles and agents. This representation serves as input for a decision-making system that ultimately determines the driving policy. Key elements included in this environmental state are lane positions, drivable areas, locations of agents like vehicles and pedestrians, and the status of traffic lights, among others [11]. Each of these elements plays a crucial role in informing the vehicle's decision-making process, enabling it to navigate safely and efficiently. Uncertainties present in perception can propagate through the information processing chain, potentially leading to incorrect decisions or dangerous situations. This

highlights the necessity for robust sensing to ensure safety. To improve the reliability of detection, it is important to incorporate multiple data sources. This can be achieved by integrating various perception tasks, such as semantic segmentation, motion estimation, depth estimation, and soiling detection, which can be efficiently unified into a multi-task model [11].

Research on autonomous driving control issues based on DL and RL technologies has yielded significant results. However, there are still many challenges to overcome before fully autonomous vehicles can become widespread. In dense traffic conditions, autonomous vehicles often face significant challenges in making appropriate driving decisions due to the high mobility of various road users, including other vehicles, bicycles, and pedestrians. Deep learning (DL) and reinforcement learning (RL) techniques have shown promising results in managing such complexities. Nevertheless, more intricate scenarios, like intersections and crosswalks, still pose significant challenges and require further exploration. Current decision-making methods typically rely on supervised learning to replicate human driving behavior. Unfortunately, these approaches do not encompass all potential driving scenarios, and it is extremely challenging to gather comprehensive data for every situation, especially as these scenarios can vary significantly from one country to another. Additionally, researchers must focus on two critical goals: ensuring the robustness and flexibility of their proposed solutions. These qualities are essential for enabling autonomous vehicles to navigate any scenario without needing prior knowledge, thus making them truly adaptable to real-world conditions [9].

In the article "Research on Strategy and Algorithm of Lateral Motion Control for Autonomous Driving Electric Vehicle", the author establishes a safety evaluation model for curvilinear driving, including slip stability and rollover stability [12]. Through the simulation of vehicle dynamics software, the functional relationship between the safe driving speed and the curve radius (Figure 2) was established. The upper level automatic steering control strategy is innovatively designed. The lower lateral motion control algorithm matched with the upper control strategy was studied, and the optimal control of the lateral motion deviation was realized. Real vehicle experiments were carried out to verify the performance of the designed control strategy and algorithm. The experimental results show that the control strategy and algorithm designed in this paper perform well under the conditions of straight road and large curvature curve, and have good real-time performance and stability.

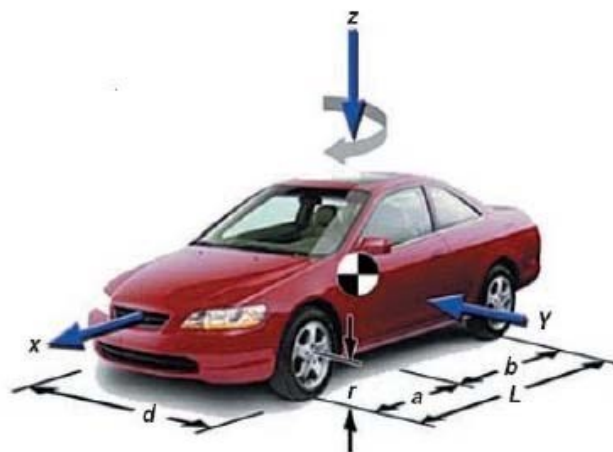


Figure 2: Motor movement diagram [12]

The article "A Survey of Deep RL and IL for Autonomous Driving Policy Learning" constructs a literature classification system from a systems perspective and identifies five patterns for integrating DRL/DIL models into autonomous driving architectures [13]. The authors comprehensively review the design of DRL/DIL models for specific autonomous driving tasks, including the design of state space, action space, and reinforcement learning rewards. Creatively and comprehensively analyzes

the application of DRL/DIL in autonomous driving from multiple perspectives, providing a classification of systems and in-depth discussions such as driving safety, interaction with other traffic participants, and environmental uncertainty. Finally, the authors propose the direction of future research and possible themes for exploration.

In the article "Formulation of Deep Reinforcement Learning Architecture: Toward Autonomous Driving for On-Ramp Merge", the authors focus on the application of Deep Reinforcement Learning (DRL) techniques to achieve autonomous driving during the merging of highway ramps [14]. The authors modeled the interactive environment through a Long Short-Term Memory (LSTM) (Figure 3) to extract the internal state from historical driving information. The internal state generated by the LSTM is then passed as input to the Deep Q-Network (DQN) for action selection to optimize long-term rewards. Then, the DQN parameters are updated through empirical playback and the target Q network to avoid the local optimal and divergence problems. Finally, the reward function is designed to guide the learning process by considering security, smoothness and timeliness. The authors innovatively combine LSTM and DQN to capture the impact of historical information on long-term rewards, and improve the reliability of the strategy.

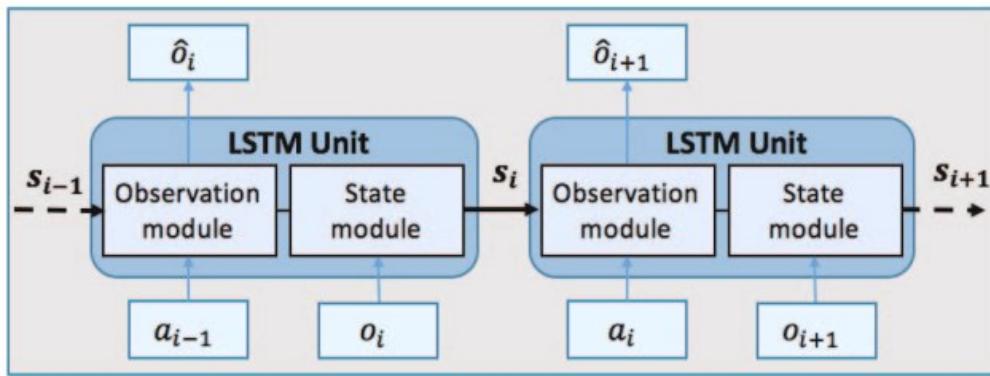


Figure 3: LSTM architecture [14]

3. Conclusion

In this paper, the author provides a comprehensive overview of the advancements and future applications of autonomous driving control strategies leveraging deep reinforcement learning (DRL). As autonomous driving technology continues to progress rapidly, DRL, an emerging artificial intelligence approach, is being increasingly applied across various domains, including vehicle control, path planning, and decision-making. The paper reviews recent literature to analyze key technologies, model architectures, and practical applications, summarizing their benefits and challenges in complex traffic situations.

The authors' comparative analysis of existing research indicates that deep reinforcement learning significantly enhances the autonomous decision-making capabilities of vehicles in dynamic environments, optimizes driving behavior, and improves safety. Among various applications, DRL demonstrates particular effectiveness in managing uncertainties and adapting to complex traffic scenarios. This paper emphasizes the principles and functions of essential algorithmic processes, such as Markov Decision Processes (MDP), within the DRL framework. The paper further explores how autonomous driving integrates various technologies, from GNSS to LSTM networks and DQN. Current research challenges include limited model interpretability, extended training periods, and poor adaptation to extreme conditions. Future research should concentrate on optimizing algorithms, integrating data from various sensors, and validating applications in real-world scenarios. Beyond technical aspects, addressing policy, regulatory, and ethical issues remains crucial for ensuring safety

and public acceptance. In conclusion, autonomous driving control strategies based on deep reinforcement learning offer significant potential for future applications, providing robust support for developing a more intelligent transportation system. This research provides valuable insights to advance autonomous driving technology development.

References

- [1] Xiao Liu (2010). *A literature review on the study of urban traffic congestion*. *ECONOMIC RESEARCH GUIDE*, 4, pp. 102-103.
- [2] Yang Lu (2012). *Environmental Pollution from an Open Macro Perspective: A Review*. *Economic Research*, 2, pp. 146-158.
- [3] Wei Yifan, Han Xuebing, Lu Languang, Wang Hewu, Li Jianqiu & Ouyang Minggao (2022), *Technology Prospects of Carbon Neutrality-oriented New-energy Vehicles and Vehicle-grid Interaction*, *Automotive Engineering*, 44(4), pp. 449-464. <https://doi.org/10.19562/j.chinasae.qcgc.2022.04.001>
- [4] Eugenio Alcalá; Laura Sellart; Vicenç Puig; Joseba Quevedo; Jordi Saludes; David Vázquez & Antonio López (2016). *Comparison of two non-linear model-based control strategies for autonomous vehicles*. In *2016 24th Mediterranean Conference on Control and Automation (MED)* (pp. 846-851).
- [5] Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G. Bellemare & Joelle Pineau (2018). *An Introduction to Deep Reinforcement Learning*. *Foundations and Trends in Machine Learning*, Vol. 11(No. 3-4), pp 219–354. <https://doi.org/10.1561/22000000071>
- [6] Seyed Sajad Mousavi, Michael Schukat & Enda Howley (2018). *Deep Reinforcement Learning: An Overview*. In *Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016* (pp. 426-440).
- [7] Youssef Fenjiro, Houda Benbrahim (2018). *Deep Reinforcement Learning Overview of the State of the Art*. *Journal of Automation Mobile Robotics and Intelligent Systems*, Vol 12(No 3), pp. 20-39. https://doi.org/10.14313/JAMRIS_3-2018/15
- [8] Ameer Haj-Ali, Nesreen K. Ahmed, Ted Willke, Joseph E. Gonzalez, Krste Asanovic & Ion Stoica (2019). *A View on Deep Reinforcement Learning in System Optimization*. *arXiv*. <https://doi.org/10.48550/arXiv.1908.01275>
- [9] Badr Ben Elallid, Nabil Benamar, Abdelhakim Senhaji Hafid, Tajjeeddine Rachidi & Nabil Mrani (2022). *A Comprehensive Survey on the Application of Deep and Reinforcement Learning Approaches in Autonomous Driving*. *Journal of King Saud University - Computer and Information Sciences*, Vol. 34(No. 9), pp. 7366-7390. <https://doi.org/10.1016/j.jksuci.2022.03.013>
- [10] Ahmad El Sallab, Mohammed Abdou, Etienne Perot & Senthil Yogamani (2017). *Deep Reinforcement Learning framework for Autonomous Driving*. *arXiv*, pp. 70-76. <https://doi.org/10.2352/ISSN.2470-1173.2017.19.AVM-023>
- [11] B Ravi Kiran, Ibrahim Sobh, Victor Talpaert, Patrick Mannion, Ahmad A. Al Sallab, Senthil Yogamani & Patrick Pérez (2021). *Deep Reinforcement Learning for Autonomous Driving: A Survey*. *IEEE Transactions on Intelligent Transportation Systems*, 23(6), pp. 4909-4926. <https://doi.org/10.1109/TITS.2021.3054625>
- [12] Jian Zhong, Xinbo Chen (2019). *Research on Strategy and Algorithm of Lateral Motion Control for Autonomous Driving Electric Vehicle*. In *2019 3rd Conference on Vehicle Control and Intelligence (CVCI)* (pp. 1-6).
- [13] Zeyu Zhu, Huijing Zhao (2022). *A Survey of Deep RL and IL for Autonomous Driving Policy Learning*. *IEEE Transactions on Intelligent Transportation Systems*, 23(9), pp. 14043 - 14065. <https://doi.org/10.1109/TITS.2021.3134702>
- [14] Pin Wang, Ching-Yao Chan (2018). *Formulation of Deep Reinforcement Learning Architecture Toward Autonomous Driving for On-Ramp Merge*. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1-6).