A survey on the use of artificial intelligence in autonomous driving

Boyuan Kong *

Beijing Institude of technology University, Beijing, China

* Corresponding Author Email: 1120211912@bit.edu.cn

Abstract. Autonomous driving and artificial intelligence are the most popular research projects in the field of technology today. As the high technology, autonomous driving relies on perception, decision-making, and control systems, and the performance of these systems largely depends on the application of artificial intelligence nowadays. Fortunately, there are plenty of applications of artificial intelligence in several aspects of autonomous driving. This paper aims to introduce the relationship between autonomous driving and artificial intelligence by reviewing several literatures and analyzing the applications of deep learning (DL), reinforcement learning (RL), and graph neural networks (GNN) in autonomous driving.

Keywords: Autonomous driving, artificial intelligence, deep learning, reinforcement learning.

1. Introduction

1.1. Motivation

Autonomous driving is surging and has already highly impacted the transportation. However, in order to ensure the safety, efficiency and reliability, it is still needed that we should solve many problems in perception, decision-making, and control systems, especially in complex environments. Artificial intelligence, especially emerging technologies in AI such as deep learning, reinforcement learning, and graph neural networks, have gradually become the key factor in solving these challenges. Therefore, studying the relationship between artificial intelligence and autonomous driving has special and important significance.

1.2. Fundamental concepts

Autonomous driving relies on the collaborative efforts of artificial intelligence, radar and other monitoring equipment and image processing. Autonomous driving is typically divided into 6 levels, ranging from L0 to L5 [1]. At present, many car manufacturers and technology companies have achieved L2 and L3 level autonomous driving, but there is still a long way to go before fully automated L5 level. It is a complex system that needs to simultaneously meet the goals of environmental awareness, path planning, and real-time driving control. There are 3 important parts of the autonomous driving: perception, planning, and control.

Perception is fusing data from multiple sensors, using computer vision and deep learning algorithms to build a 3D map. Planning is planning the vehicle's high-level behaviors, Predicting the behaviors and trajectories of other road users to generate a safe and efficient driving path. Control is controlling the vehicle's actual movements, ensuring the vehicle accurately follows the path provided by decision making part [2].

1.3. Contribution

This paper reviewed several literatures and found out the relationship between autonomous driving and artificial intelligence, analyzing developments and applications of deep learning (DL), reinforcement learning (RL), and graph neural networks (GNN) in autonomous driving.

2. Application of Deep Learning in Autonomous Driving

2.1. Fundamental concepts

Deep learning, as a machine learning method, has unique characteristics, and its characteristic is to extract complex features from data and learn through a multi-layer network structure. It is a half-theoretical and half- empirical modeling approach which uses mathematical knowledge and computer algorithms to construct a comprehensive architecture. Deep learning can extract useful information from large amounts of data, greatly improving the performance of models. Nowadays, Deep learning has achieved a lot in many fields, and has also been widely applied in autonomous driving [3].

2.1.1. Deep Neural Networks

Deep neural networks construct models by simulating the human brain and learning its structure and function. DNN is the fundamental architecture of deep learning, and DNN includes an input layer, multiple hidden layers, and an output layer. By transmitting information layer by layer and applying nonlinear activation functions, DNN can automatically extract and learn complex features of data.

2.1.2. Convolutional Neural Networks

Convolutional neural networks, inspired by natural organisms, simulate the way organisms perceive the world through visual perception and construct feedforward neural networks. The first layer inputs the image and convolves it, resulting in a feature map with a depth of 3 for the second layer. Continuously repeat, unfold and connect the feature map, and pass it to the fully connected layer.

2.1.3. Recurrent Neural Networks

The essence of recurrent neural networks is the ability to remember and make inferences based on the content of these memories. Therefore, its output depends on the current input and memory. The nodes between hidden layers are connected, which gives the network strong memory and allows it to process input sequences of any time sequence using its internal memory.

2.2. Application

2.2.1. Perception

It is good for deep learning methods to detecting and recognizing objects in 2D images and 3D point clouds. Deep learning achieves comprehensive perception of the environment by processing multi-source sensor data. For example, CNN plays an important role in object detection, object recognition, and scene segmentation in camera images, enabling vehicles to recognize roads, traffic signs, pedestrians, and other vehicles. CNN extracts spatial features from images through convolution operations and combines them with other techniques such as fully convolutional networks (FCN) and generative adversarial networks (GAN) to achieve high-precision image segmentation and object detection. Now deep learning models such as YOLO, SSD, and Faster R-CNN have been widely applied in these tasks and have shown excellent performance in terms of accuracy and real-time performance [3].

2.2.2. Decision-Making

The path planning method based on deep learning can calculate the optimal driving path in real time by learning environment and rules. Models combining RNN and LSTM can predict the vehicle trajectories, thereby assisting vehicles in path planning in complex urban environments [4].

2.2.3. Control

Deep learning has been used in an end-to-end autonomous driving, which is a way to directly input sensors to control outputs, which avoids the separation of 3 important parts in the traditional auto drive system. The end-to-end method based on deep learning maps camera input directly to control commands for steering wheel, throttle, and brake through neural networks. This method simplifies the system structure, reduces the accumulation of errors in intermediate links, and has potential application prospects. Studies such as NVIDIA's DAVE-2 system have demonstrated the feasibility of it, although challenges such as data dependency, interpretability, and safety still need to be addressed [5].

3. Application of Reinforcement Learning in Autonomous Driving

3.1. Fundamental concepts

Reinforcement learning learns optimal strategies by interacting with the environment. It guides intelligent agents to gradually improve their decisions through reward and punishment mechanisms, ultimately learning to adopt optimal behavior in specific environments. The framework of reinforcement learning mainly consists of the following cores, state, action, reward, policy. It aims to interact with the surrounding environment and obtain the optimal solution. Through error process, the agent selects actions in each state and adjusts its strategy based on the rewards to maximize long-term cumulative rewards [6]. The application of reinforcement learning in autonomous driving is increasing, especially in decision-making and control tasks in complex and uncertain environments, demonstrating great potential.

3.1.1. Markov Decision Process

Markov Decision Process (MDP) is widely used in reinforcement learning. Specifically, it includes states, possible actions, transition probabilities between states, and a reward function. The goal of decision-making is to select appropriate actions in each state that maximize the total reward obtained in the long term.

3.1.2. Deep Q-Networks

Deep Q-Network (DQN) is used to solve the problem of intelligent agents learning how to make decisions and obtain maximum cumulative rewards while interacting with the environment. Unlike traditional Q-learning, DQN uses deep neural networks to estimate Q-value functions and improves decision strategies by optimizing the weights of the network, enabling it to handle problems with high-dimensional state spaces.

3.1.3. Policy Gradient Method

The strategy gradient method directly optimizes the strategy function and can work in a continuous action space. This means skipping the evaluation stage of action value and directly moving from the input state to the output strategy. Common strategy gradient methods include REINFORCE, Near End Policy Optimization (PPO), and Policy Gradient Reinforcement Learning (A3C).

3.2. Application

3.2.1. Perception

It is good for reinforcement learning methods in multi-modal perception data fusion. Automatic driving perception systems usually need to integrate data from different sensors. RL can be used to optimize the fusion process of these multimodal data by learning how to weight and integrate various sensor data in different contexts, thereby providing more accurate perception results in complex environments [7]. However, reinforcement learning is mainly used for decision-making and control.

3.2.2. Decision-Making

Reinforcement learning trains vehicles to learn optimal driving strategies in different environments by simulating complex traffic scenarios. Imitation learning is a reinforcement learning method that accelerates strategy learning by imitating expert behavior. It helps autonomous vehicles quickly acquire effective driving strategies in the initial stage by learning the behavior of human drivers, thereby reducing training time and data requirements.

3.2.3. Control

Reinforcement learning is used to develop Adaptive Cruise Control (ACC) systems that enable vehicles to automatically adjust their speed to maintain a safe distance under different speeds and traffic conditions, dynamically plan paths to avoid obstacles, and select the optimal route. Reinforcement learning algorithms such as DQN and PPO have achieved efficient response to dynamic traffic conditions through online learning and real-time policy adjustment. The DDPG algorithm can learn how to drive autonomously in complex urban environments by training in a simulation environment. SAC and PPO algorithms are used to optimize the path selection of vehicles on congested roads, thereby reducing driving time and energy consumption [8].

4. Application of Graph Neural Networks Learning in Autonomous Driving

4.1. Fundamental concepts

Graph Neural Networks (GNNs) are a kind of deep learning model to solve the problems form graph data. A graph structure is a universal data structure where nodes represent entities and edges represent relationships between entities [9]. In autonomous driving, many problems can be naturally represented as graphs, such as traffic networks, scene modeling, and the interaction of multi-agent systems. Graph neural networks can effectively process and analyze complex relationships while maintaining the topological structure information of the graph by learning the features of nodes and edges.

4.1.1. Graph Convolutional Network

Graph Convolutional Network (GCN): GCN is a neural network that extends convolution operations to graph structured data. It learns node feature representations by propagating and aggregating information within the neighborhood of the graph. GCN is commonly used for scene understanding and traffic flow prediction in autonomous driving.

4.1.2. Graph Attention Network

Graph Attention Network (GAT): GAT introduces an attention mechanism that flexibly captures relationships between nodes by assigning different weights to neighboring nodes in the graph. GAT performs well in handling complex traffic scenarios and collaborative problems in multi-agent systems.

4.1.3. Dynamic Graph Networks

DGN is a neural network that processes dynamically changing graph structures, suitable for situations where nodes and edges in transportation networks are dynamically changing. It can capture time-varying relationships and has wide applications in traffic prediction and path planning in autonomous driving.

4.2. Application

4.2.1. Scene Modeling

Autonomous vehicles require modeling and understanding of complex driving scenarios. By modeling the driving scene as a graph structure, GNN can effectively capture entities and their interrelationships in the scene. For example, GCN can be used for scene segmentation to identify and classify different areas in the driving scene, such as lanes, road boundaries, obstacles, etc. The transportation network can be naturally represented as a graph structure, where nodes represent road intersections or segments, and edges represent the connectivity relationships of roads. Graph neural networks can effectively predict traffic flow, estimate traffic status, and detect congestion by learning the characteristics of nodes and edges in transportation networks. For example, models such as GCN and GAT can capture spatiotemporal dependencies in transportation networks by performing convolution or attention operations on graph structures, thereby achieving more accurate traffic predictions [10].

4.2.2. Traffic Prediction

GNN can utilize the graph structure information of transportation networks to achieve more accurate spatiotemporal prediction. For example, combining GCN and LSTM models can simultaneously capture spatial and temporal dependencies in transportation networks, thereby achieving dynamic prediction of future traffic flow [11]. This method demonstrates higher prediction accuracy when dealing with sudden traffic incidents or complex changes in traffic patterns. Meanwhile, path planning is a crucial task in autonomous driving, involving the calculation of the optimal route from the starting point to the endpoint. Graph neural networks can quickly estimate the optimal path by learning the graph structure of road networks. The DGN model is particularly suitable for real-time path planning in dynamic traffic environments, as it can handle dynamic changes in traffic conditions and quickly adjust paths.

5. Conclusion

5.1. Summary

This article provides a literature review of the application of artificial intelligence technology in autonomous driving in recent years, and analyzes the application of AI technologies such as deep learning, reinforcement learning, and graph neural networks in the core aspects of autonomous driving: perception, decision-making, and control. Deep learning has achieved a lot in many fields such as image processing, object detection, and path planning. Reinforcement learning demonstrates unique advantages through interactive learning with the environment, especially in dynamic decision-making and complex control tasks. Graph neural networks does well in areas such as transportation networks, scene modeling, path planning, and multi-agent collaboration by effectively modeling graph structured data.

5.2. Future Prospects

The performance of AI models in autonomous driving is dependent on large-scale, high-quality data, and will increasingly rely on multimodal data. Self-supervised learning, and generative adversarial networks (GANs) may show the importance in reducing data annotation requirements and improving model training efficiency. And GNN can play an important role in multimodal field. However, the ethical and legal issues arising from autonomous driving, for example, the division of decision-making responsibilities in emergency situations, also need to be fully studied and resolved while technology develops.

The future development of autonomous driving technology cannot separate from the innovation of artificial intelligence. With the continuous advancement of technology, AI will provide more powerful perception, decision-making, and control capabilities for autonomous driving. However, while pursuing technological breakthroughs, it is necessary to pay attention to the safety, ethical, and legal challenges, ensuring that autonomous driving technology can safely and reliably serve society.

References

[1] Parekh, Darsh, et al. ‘A Review on Autonomous Vehicles: Progress, Methods and Challenges’. Electronics, vol. 11, no. 14, MDPI AG, July 2022, p. 2162.

[2] Ma, Yifang, et al. "Artificial intelligence applications in the development of autonomous vehicles: A survey." IEEE/CAA Journal of Automatica Sinica 7.2 (2020): 315-329.

[3] Grigorescu, Sorin, et al. "A survey of deep learning techniques for autonomous driving." Journal of field robotics 37.3 (2020): 362-386.

[4] Niu, Zhaoyang, Guoqiang Zhong, and Hui Yu. "A review on the attention mechanism of deep learning." Neurocomputing 452 (2021): 48-62.

[5] Chib, Pranav Singh, and Pravendra Singh. "Recent advancements in end-to-end autonomous driving using deep learning: A survey." IEEE Transactions on Intelligent Vehicles (2023).

[6] Kiran, B. Ravi, et al. "Deep reinforcement learning for autonomous driving: A survey." IEEE Transactions on Intelligent Transportation Systems 23.6 (2021): 4909-4926.

[7] Chen, Guoxi, Ya Zhang, and Xinde Li. "Attention-based highway safety planner for autonomous driving via deep reinforcement learning." IEEE Transactions on Vehicular Technology (2023).

[8] Ge, Lun, et al. "Deep reinforcement learning navigation via decision transformer in autonomous driving." Frontiers in Neurorobotics 18 (2024): 1338189.

[9] Wu, Zonghan, et al. "A comprehensive survey on graph neural networks." IEEE transactions on neural networks and learning systems 32.1 (2020): 4-24.

[10] Xi, Zerong, and Gita Sukthankar. "A Graph Representation for Autonomous Driving." The 36th Conference on Neural Information Processing Systems Workshop. Vol. 7. No. 8. 2022.

[11] CAI, Peide, et al. "DQ-GAT: Towards safe and efficient autonomous driving with deep Q-learning and graph attention networks." IEEE Transactions on Intelligent Transportation Systems 23.11 (2022): 21102-21112.