Key Technologies, Challenges and Future Prospects of Neural Network in Automatic Driving System

Jiale Liu *

School of Springboard International Bilingual School, Beijing, 694107, China

* Corresponding Author Email: calebliujiale123@gmail.com

Abstract. Automatic driving technology is a disruptive force to reshape the future travel and transportation system. Its development has undergone a fundamental transformation from a rule-based modular approach to a data-driven deep learning paradigm. This paper systematically summarizes the key role, frontier progress and core challenges of neural network as the core technology engine in the whole technology stack of automatic driving perception, decision-making, planning and control. Firstly, this paper summarizes the application foundation of convolutional neural network, cyclic neural network, transformer and graph neural network in the field of automatic driving; Then the representative work and implementation path in the core tasks of multi-sensor fusion 3D target detection, interactive behavior prediction, end-to-end driving are analyzed in detail. The analysis shows that although the neural network has greatly improved the performance ceiling of the system, its inherent black box characteristics, vulnerability to adversarial attacks, difficult to deal with the "long tail problem" and high computing costs are still the fundamental obstacles restricting its safe landing. Finally, this paper looks forward to the future research directions of multimodal large-scale model, causal reasoning, neural radiation field simulation and vehicle road coordination, aiming to provide a comprehensive and profound perspective of technology development for researchers in the field.

Keywords: Automatic driving; Deep learning; Convolutional neural network; Transformer; End to end learning.

1. Introduction

Nowadays, automatic driving is becoming increasingly important. Most of cars in today already have automatic drive ability, and individual uses is function more and more frequent. With the rapid development of artificial intelligence, sensor technology and computing platforms, autonomous driving, as a disruptive technology, is moving from the realm of science fiction to reality, and is expected to fundamentally reshape the form of future transportation. Its core value lies in significantly enhancing road safety levels by eliminating human errors. Moreover, autonomous driving technology has profound social and economic significance in improving traffic efficiency, providing mobility solutions for the disabled and the elderly, and giving rise to new business models.

So, people can know that improving automatic driving technology is currently a heavy task. The most important part in part for achieving autonomous driving is neural network. Because neural network take responsibility for perceiving the environment, making decisions and maintain safe driving as a result. Better ability on perceiving the environment allows car perceive more potential risk and make a more accurate distinction. Making a more sensible decision allows car running on a safer route. Both of improvements are crucial for car’s safety.

Therefore, the development of neural network is of great significance for autonomous driving. We have many researchers studying this problem, and they have their own ideas. For example, the deeptest test framework developed by Tian et al. systematically adopts neuron coverage as the guiding standard for the first time [1]. It can automatically generate test cases to simulate the changes of rain and fog in the original road image, so as to find thousands of fragile scenarios that may lead to vehicle departure from the lane and other fatal errors in multiple high-performance automatic driving models. This reveals that the data-driven model has serious reliability risk in the face of rare "inflection point", which opens up an important direction for the test and verification of autonomous driving. The end-to-end pilotnet model proposed by bojarski et al. successfully proved for the first time that a single convolutional neural network can map directly from the pixel image of the front camera to the steering wheel control command [2], challenging the necessity of traditional modular design and causing extensive discussion in the industry. In terms of environmental perception, Chen et al. proposed a multi view 3D network [3], which realized Pseudo 3D target detection without lidar by fusing multi view camera images. This proves the potential of CNN in complex stereo vision tasks and provides the possibility of low-cost autonomous driving solutions.

The structure of this paper is as follows: the second chapter briefly introduces the architecture of autonomous driving system and the theoretical basis of neural network. As the core part, the third chapter reviews the application status of neural network in key modules such as environmental perception, high-precision positioning, decision planning and so on. The fourth chapter will focus on the end-to-end autonomous driving paradigm. The fifth chapter will analyze the main challenges faced by the technology, such as security, interpretability and data dependency. Chapter 6 looks forward to the future development of the integration of neural network and autonomous driving. Finally, the seventh chapter summarizes the full text.

This paper aims to make a comprehensive and systematic review of the application of neural network in automatic driving. The main contribution of this paper is to provide a panoramic perspective, systematically explain the key role of neural network in the whole technology stack of automatic driving, such as perception, positioning, decision-making, planning and control, and deeply analyze the principles, advantages and limitations behind it.

2. Fundamentals of Autonomous Driving and Neural Networks

2.1. Overview of Autonomous Driving System Architecture

The design of autonomous driving system mainly follows two modes: modular architecture and end-to-end architecture.

Modular architecture is the current mainstream commercial solution. It decomposes the complex autonomous driving task into a series of serial or parallel sub modules, mainly including perception, prediction, decision planning and control. The advantage of this architecture lies in the clear division of labor between modules, which is convenient for debugging and verification. However, its disadvantage is that errors can be propagated and accumulated between modules.

End to end architecture represents a more radical approach. Its purpose is to use complex neural network to directly map the original data obtained from the sensor to the final control command. The pioneering work of bojarski et al. pilotnet verified the architecture [2]. Its advantages include simple system structure and reducing the possibility of error accumulation, but its core challenge stems from the "black box" nature, resulting in poor interpretability, difficult debugging and highly challenging security verification.

2.2. Introduction to Core Neural Network Models

As the cornerstone of image data processing, convolutional neural network (CNN) is widely used in target detection, semantic segmentation and other sensing tasks.

Recurrent neural network (RNN) and long-term and short-term memory network (LSTM) are designed to process sequential data and are suitable for temporal modeling tasks such as behavior prediction and trajectory tracking.

Using its self-attention mechanism, transformer can effectively capture the dependencies between distance elements in the sequence, and apply it to multi-sensor fusion and scene understanding.

Gan and AE generation are used for data expansion, simulation environment generation and anomaly detection.

Graphical neural network (GNN) can effectively infer complex social interaction behavior, which is the core technology of the next generation behavior prediction system.

3. Application of Neural Networks in Core Modules of Autonomous Driving

3.1. Environmental Perception

Perception is the "eye" of autonomous driving system, and neural network significantly enhances its visual ability.

Object detection has developed from 2D to 3D. In two-dimensional detection, single-stage model (such as yolo) and two-stage model (such as fast r-cnn) have been very mature [4]. Three-dimensional target detection can provide the exact size and location of the target, which is very important for security. The method based on LIDAR point cloud, such as point column [5], voxels the irregular point cloud before using 3D CNN to process it; Vision based methods, such as the mv3d network proposed by Chen et al. [3], realize 3D detection by fusing multiple views from camera images and point clouds.

Semantic/instance segmentation classifies each pixel in the image. The complete convolutional network (FCN) architecture lays the foundation for such tasks [6], and can accurately depict the outline of the driveable area, lane line and single object, providing rich environmental information for path planning.

Sensor fusion is the key to improve the perceptual robustness. With its powerful multimodal information integration capability, transformer architecture shows great potential in feature level fusion [7], effectively combining high-resolution texture information from the camera and accurate depth information from lidar.

3.2. Decision-Making and Planning

The "brain" of decision-making and planning as a tool has evolved from a rule driven approach to a learning driven approach.

Behavior prediction is the premise of safety decision. At present, graph neural network has become the mainstream; For example, vectornet accurately predicts the future trajectory of interaction by modeling the topological relationship between road elements and traffic participants [8].

In the aspect of motion planning, planners based on imitation learning learn the behavior data of drivers; The driving strategy based on reinforcement learning interacts with the simulated environment and learns the optimal driving strategy through repeated experiments. These methods can generate more human like and smoother trajectories.

3.3. Motion control

At the control level, the neural network can be used as a controller to directly learn the complex nonlinear mapping from the system state to the actuator instructions, as shown in pilotnet in autonomous driving [2]. This method has the potential to surpass the traditional PID controller, especially when dealing with nonlinear systems and adapting to dynamic environment, without accurate parameter tuning.

However, due to its black box characteristics, neural network controllers face challenges in stability verification and robustness assurance. The current research explores the hybrid structure, which combines the neural network with the classical control method to take advantage of the advantages of the two methods and ensure the reliability of the system in safety critical applications.

4. Challenges and Limitations

Although neural network has achieved remarkable success, its application in the field of autonomous driving with high safety requirements still faces severe challenges.

4.1. Safety and robustness

The decision-making boundary of neural network is very fragile and vulnerable to hostile attacks: adding imperceptible small disturbances to the input image may cause the model to make completely wrong judgments [9]. In addition, the long tail problem is the main obstacle to actual deployment, which means that the training data cannot cover all possible rare scenarios in the real world.

4.2. Interpretability and Trustworthy AI

The "black box" characteristic of neural network makes it difficult for people to understand its internal decision logic. When an accident occurs, we cannot trace the cause of the fault as we do with traditional procedures. Developing interpretable AI (Xai) technology to visualize and attribute model decisions is a key step to build user trust and obtain regulatory approval.

4.3. Computational Efficiency and Real-Time Performance

Large scale neural network model needs a lot of calculation, and autonomous driving requires high real-time performance. How to run these models efficiently on an embedded platform with limited resources is a major engineering challenge.

4.4. Data dependence and simulation

The performance of neural networks largely depends on large-scale, high-quality and diverse annotation data. Therefore, using simulation platform and generative adaptive networks (GAN) to generate a large number of real synthetic data and conduct large-scale virtual test (as shown in Tian et al. deeptest) has become the key [6].

5. Future Outlook

5.1. Algorithmic Level: From Perceptual Intelligence to Cognitive Intelligence

Future research will no longer be limited to a single perceptual task, but to pursue higher cognitive ability.

Based on the successful experience of chatgpt, the multimodal large-scale model will establish a basic visual language model in the field of autonomous driving (such as drivegpt) [10], so as to understand more abstract traffic concepts and extend it to unknown scenes.

Causal reasoning enables the model not only to learn the statistical correlation in the data, but also to understand the causal relationship between various elements in the scene, so as to carry out causal reasoning and make more rational decisions in the face of new situations.

Continuous learning/lifelong learning develop models that can continuously learn and develop from new data and environments without forgetting existing knowledge, so that the automatic driving system can continuously adapt and improve.

5.2. System Architecture Level: From Simulation Testing to Vehicle-Infrastructure cooperation

Neural radiation field (nerf), an emerging technology, can generate surreal three-dimensional scenes from sparse images, and conduct security verification by creating virtual worlds that are almost indistinguishable from reality, which will completely change the simulation test of automatic driving.

Vehicle to everything (v2x) combines on-board intelligence with roadside intelligent facilities (cameras, radars), and the neural network will process more macro traffic flow information, realize over the horizon perception and collaborative decision-making, and fundamentally improve safety and efficiency.

5.3. Testing, Verification, and Commercialization

Reliable simulation test and formal verification combine the traditional program verification method with deep learning to provide mathematical security for the behavior of neural networks.

The technology will gradually expand from restricted scenes (such as robot axis, port, mining area) to open roads. By continuously collecting "long tail data" and iterating the model in actual operation, we can finally realize fully autonomous driving in all scenarios.

6. Conclusion

In this paper, the application of neural network in autonomous driving system is comprehensively and systematically reviewed. By investigating the technical evolution of neural networks in the core modules of environmental perception, decision planning and motion control, it is clear that neural networks, especially the deep learning model, have evolved from auxiliary tools to the technical cornerstone to promote the development of the whole system. From the end-to-end drive initiated by bojarski et al., to the multi view 3D detection explored by Chen et al., and then to the automated test framework to solve security challenges, these studies jointly demonstrate the great potential of data-driven methods in understanding and processing complex, open and dynamic environments.

However, there are serious challenges behind the prosperity of this technology. The unexplainability of neural networks, their sensitivity to hostile examples and the thorny "long tail problem" have led to a crisis of trust in the application of neural networks in autonomous driving with strict safety requirements. Future breakthroughs may not be limited to improving the performance of a single model, but rely more on innovation at the architecture level: for example, use multimodal large-scale models to achieve in-depth scene understanding and causal reasoning, use neural radiation field (nerf) to build a high-fidelity simulation environment for thorough verification, and overcome the perceived limitations of single vehicle intelligence through vehicle to everything (v2x) cooperation.

In short, the neural network has outlined an exciting technical blueprint for the realization of fully autonomous driving. However, turning this blueprint into a safe and reliable reality requires continuous and in-depth exploration by academia and industry in the fields of algorithm robustness, system architecture and verification standards.

References

[1] Tian Y, Pei K, Jana S, et al. DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars. Proceedings of the 27th ACM SIGSOFT International Symposium on Software Testing and Analysis, 2018.

[2] Bojarski M, et al. End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316, 2016.

[3] Chen X, Ma H, Wan J, Li B, & Xia T. Multi-view 3d object detection network for autonomous driving. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 1907-1915), 2017.

[4] Redmon J, Divvala S, Girshick R, & Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

[5] Lang A H, Vora S, Caesar H, Zhou L, Yang J, & Beijbom O. PointPillars: Fast Encoders for Object Detection from Point Clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.

[6] Long J, Shelhamer E, & Darrell T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.

[7] Prakash A, Chitta K, & Geiger A. Multi-Modal Fusion Transformer for End-to-End Autonomous Driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.

[8] Gao J, Sun C, Zhao H, Shen Y, Anguelov D, Li C, & Schmid C. VectorNet: Encoding HD Maps and Agent Dynamics from Vectorized Representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.

[9] Szegedy C, Zaremba W, Sutskever I, Bruna J, Erhan D, Goodfellow I, & Fergus R. Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199, 2013.

[10] Wang Z, et al. DriveGPT: Towards Scalable Autonomous Driving with Large-Scale Video Pre-Training. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.