Dean&Francis
ISSN 2959-6157

Application Analysis of Artificial Intelligence Big Data Models in Autonomous Driving

Xingkai Hao[1,*]

[1]Hebei Anxin High School, Baoding, Hebei, China
*Corresponding author: John_xk1211@outlook.com

Abstract:

In recent years, driving assistance systems have gradually gained a foothold in the automotive industry's technology. Although the concept was proposed over half a century ago, it has only entered the public eye in recent years due to limitations in technology, regulations, and other aspects. Autonomous driving technology mainly consists of perception, prediction, decision-making and planning, and control execution. The development of AI in recent years has greatly promoted the progress of the autonomous driving field. Large AI models have significantly improved the accuracy of trajectory prediction and the effectiveness of decision-making and planning in autonomous driving. This paper systematically analyzes how autonomous driving perceives, analyzes, makes decisions and plans like humans, the limitations of different sensors, the necessity of multi-sensor fusion, and the advantages and disadvantages of different AI models in autonomous driving. Finally, it analyzes the future development direction of autonomous driving and the advantages of data engines.

Keywords: Object recognition; Autonomous driving; Artificial intelligence; Trajectory prediction; Decision planning.

1. Introduction

The earliest roots of autonomous driving technology can be traced back to the 1950s when General Motors in the United States proposed the concept of the "Automated Highway System (AHS)". AHS aimed to transform the roads themselves by laying cables beneath them and guiding vehicles along fixed paths through magnetic fields. It also featured a central control system to manage the speed, spacing, and routes of the traffic flow. In 1990, the US Department of Transportation launched the AHS project for testing on California's I-15 highway, where 20 specially modified vehicles traveled in a convoy at 130 km/h with only a few meters between them. However, the AHS project was not put into use due to its high infrastructure transformation costs and lack of flexibility, and was replaced by "single-vehicle intelligence". In the 2010s, autonomous driving entered the commercialization era. In 2010, Google began formal testing of its self-driving cars, which used technologies such as lidar, cameras, and high-precision maps. Since then, more Internet companies like Baidu, Tencent, and Alibaba have entered the autonomous driving field in various ways, bringing a large amount of capital and talent to the research of autonomous driving technology. In 2014, Tesla released the Autopilot assisted driving function, marking the first time advanced assisted driving technology was installed in mass-produced vehicles. Since then, L2+ and L3 level assisted driving have gradually become popular in high-end mass-produced models. The technical routes have gradually diverged into "multi-sensor fusion" and "pure vision". Since 2021, many cities around the world have set up L4 level assisted driving pilot projects, allowing driverless operation. However, as of now, due to technological and regulatory limitations, most cars' autonomous driving is only equipped with L2 level assisted driving.

Autonomous driving is transforming the way people travel in the future. Its core modules include environmental perception, motion prediction, decision planning, and control execution. The emergence of basic models, especially in the fields of natural language processing and computer vision, has taken autonomous driving research to a new level. The uniqueness of these models lies in their training on extensive network-scale datasets, with large dataset sizes. Given the huge amount of data generated by autonomous vehicle services and the advancements in artificial intelligence including natural language processing and AI-generated content (AIGC), the role of artificial intelligence in autonomous driving is becoming increasingly significant. These models play a crucial role in object detection, scene understanding, and decision-making in autonomous driving, even approaching the level of human drivers [1-3].

2. The Application of AI in Autonomous Driving

When humans operate vehicles, they usually rely on the coordination of various senses to perceive the road environment, including vision, hearing, and the vestibular system, and then make judgments and control actions. Such judgments are usually based on experience and rational thinking. The subconscious control of the accelerator, brake, and steering wheel by humans is something that the drivers themselves cannot explain why they do it so precisely. In autonomous driving, the perception capabilities of various sensors in most cases exceed those of humans because these sensors can obtain environmental data more accurately. However, machines cannot drive like humans because the processing ability of vehicles for environmental information can never reach the same level as that of humans. Therefore, the analytical and decision-making capabilities of vehicles need to be closer to or even surpass the level of humans, which requires vehicles to have the same "thinking" and "acting" abilities as human drivers, as shown in Figure 1. Based on the achievements of AI in different fields, deep learning and reinforcement learning are increasingly being applied to autonomous driving, such as identifying the type of obstacles, predicting the trajectories of vehicles and pedestrians ahead, planning the vehicle's movement routes, and performing corresponding vehicle control. The recent advancements in AI technology have significantly improved the performance of these tools, and such components are increasingly implemented using deep learning tools. Deep learning methods have replaced previous methods in a series of tasks and have become the de facto preferred tool in many applications.

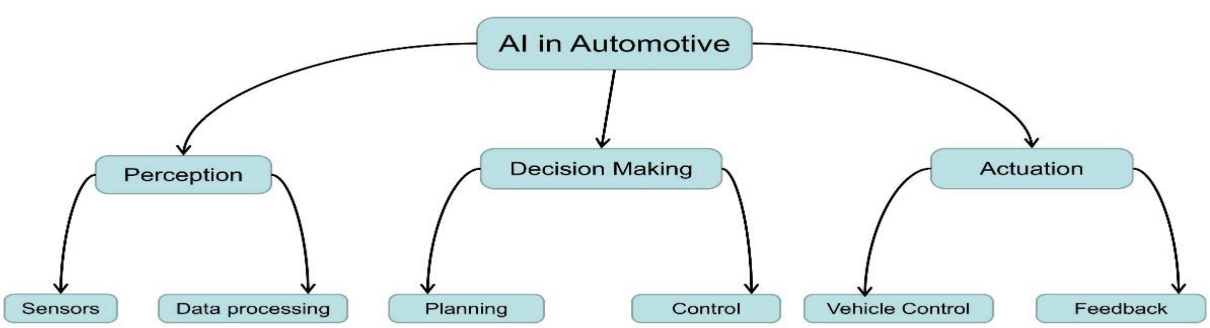

Fig. 1. The application of AI in various modules of autonomous driving

2.1 Perception

Autonomous vehicles obtain information about the surrounding environment through sensors such as lidar, cameras, and ultrasonic waves. AI uses advanced algorithms to process this information to identify information in the surrounding environment, such as the type of obstacles, the speed of surrounding vehicles, and the type of traffic signs. Currently, the main sensor types of autonomous vehicles include cameras, radars, lidars, ultrasonic waves,

ISSN 2959-6157

GPS, and IMU (Inertial Measurement Unit), etc. [4]. Cameras can capture visual information of the surrounding environment, which is helpful for object detection and recognition; radars use radio waves to detect the distance and movement speed of targets from the vehicle, providing key data support for the vehicle's obstacle avoidance function; lidars can create high-resolution 3D maps of the environment, which is helpful for the vehicle to identify the type of obstacles; the application range of ultrasonic waves is relatively small, and it is usually used in parking scenarios, where it can identify obstacles within a short distance; GPS provides the vehicle's location data, allowing the vehicle to locate and plan routes based on this; IMU provides the vehicle's acceleration and vehicle attitude data. The data from these sensors is input into the perception module.

2.1.1 Computer Vision

The research on image analysis for autonomous vehicles is a broad and rapidly developing field that intersects with multiple interdisciplinary areas such as vision, machine learning, sensor fusion, and federated learning. The main advancements in recent years include the application of cutting-edge methods such as convolutional neural networks (CNNs), transformer models, generative adversarial networks (GANs), and visual language models (VLMs). Object detection is the most fundamental task in autonomous driving, and the progress in the AI field is also advancing the accuracy and efficiency of object detection, such as the YOLO model, SSD model, and the DETR model based on transformer. The YOLO model, with its efficiency and accuracy in real-time object detection, has provided reliable technical support for object detection in autonomous driving. Additionally, Ross Girshick and his colleagues developed R-CNN (Region-Based Convolutional Neural Network), which combines region proposal networks with deep learning, significantly improving detection performance [5]. In situations where there are vehicles or pedestrians crossing the road or animals appearing suddenly ahead, the use of AI models can better identify, enabling more efficient and accurate subsequent processing.

2.1.2 Multi-sensor fusion

The significance of multi-sensor fusion lies in complementing each other's strengths and improving the accuracy of perception. By integrating different data sources, it enables precise and real-time analysis of the surrounding environment. A single sensor inevitably has limitations. For instance, cameras may increase noise in poor lighting conditions or overexpose in strong light, and even lose functionality. Moreover, cameras generate a large amount of data per second, which requires certain storage and computing capabilities. The integration of cameras and radars is a key technology in the field of autonomous driving. It involves combining the information from cameras and radar sensors to enhance the perception and decision-making capabilities of autonomous vehicles. The working principle and importance of camera and radar fusion lie in sensor complementarity, redundancy and reliability, object detection and tracking, and filling sensor blind areas. In terms of sensor complementarity, multi-sensor fusion provides the possibility for other sensors to fill the information gaps when one sensor fails, enabling vehicles to still perceive the surrounding environment in complex weather conditions. Moreover, when one sensor provides inaccurate information, other sensors can verify or correct the data, this redundancy reduces the risk of misjudgment and improves the overall safety of the vehicle. In object detection and tracking, cameras can more accurately detect the type of obstacles ahead, while radar can determine the speed and distance of the obstacles, so that the algorithm can predict the trajectory of the obstacles [6].

2.2 Behavioral prediction

In behavior prediction, various methods have been proposed, including physical-based models, rule-based models, and deep learning-based models. Physical-based models use physical formulas to predict vehicle trajectories in the short term, including constant speed models (CV), constant acceleration models (CA), constant steering rate and acceleration models (CTRA), constant speed and velocity models (CTRV), bicycle models, intelligent driver model (IDM), and path tracking models. These models use simple modeling to directly calculate future motion based on the current state of the object. Clearly, this type of model has the advantages of simple modeling and easy calculation compared to other models. On the other hand, these models ignore environmental influencing factors and interactions with other entities. Although they can calculate quickly and accurately in short-term predictions, they become difficult to adapt to complex road conditions.

Rule-based models use known traffic rules and human prior knowledge to predict the future trajectory of vehicles in a structured manner. The simplicity and intuitiveness of this model stem from the use of predefined rules. Based on this, it has the advantages of simple calculation and adjustment. However, similar to physical models, it lacks flexibility in complex scenarios. And since the rules are programmed, adapting to new scenarios becomes difficult.

The models based on deep learning have significantly advanced the progress of long-term motion prediction, demonstrating outstanding performance in integrating various factors, including physics, road conditions, and interactions among multiple vehicles. They have compensated for the deficiencies of physical models and rule-based models in complex road conditions. Among various methods, networks based on RNN, graph-based networks, and Transformer-based networks stand out, each having unique advantages in analyzing the complexity of spatio-temporal data. These methods are all based on supervised learning, but the lack of data has prompted researchers to turn to new solutions.

The existence of self-supervised learning has solved the problem of scarce high-quality data. This method aims to enhance and diversify the available data set through strategies such as data augmentation to alleviate the scarcity of high-quality data [7].

The algorithm for predicting surrounding vehicles at multi-lane turning intersections using recursive neural networks (RNN) with 484 real vehicle trajectory sample data from the road, trained and evaluated 11,662 and 4,998 samples of the network [8]. And using a motion planner based on model predictive control (MPC), it issues instructions to adjust the longitudinal acceleration of the vehicle based on the predicted state of surrounding vehicles. The improved prediction accuracy has significantly enhanced safety by limiting the prediction error within the safety boundary. The application results of the proposed predictor show that the timing of identifying the front vehicle has been improved, and the similarity of longitudinal acceleration with the driver has been improved.

3. Challenges

3.1 Sensors

Firstly, the heterogeneity of cross-modal data makes it difficult to achieve a unified feature space. Secondly, the temporal and spatial asynchrony between different sensors may lead to fusion errors. Moreover, sensor failures (such as lens contamination and signal blockage) may result in the dynamic loss of multimodal information [9]. As a passive sensor, cameras are highly sensitive to light conditions, and their image quality significantly deteriorates in the presence of night and adverse weather conditions. Lidar calculates the distance of objects by measuring the time difference between the transmission and reception of laser signals, directly outputting high-precision point clouds containing spatial geometric information, which has unique advantages in 3D perception. Adverse weather conditions can cause the scattering of laser radar or physical obstruction of the probe, resulting in a decrease in measurement accuracy. Additionally, due to the sparsity and non-uniformity of point clouds, there are difficulties in representing and understanding the point cloud data of lidar. Compared with lidar, radar point clouds are sparser and cannot precisely depict the contours of objects.

3.2 Algorithm efficiency

In practical application scenarios, the extremely short computing time in constantly changing situations may cause deviations in the actual vehicle control, especially during high-speed movement. At the lower control layer of the vehicle system, many algorithms ignore the delays caused by various factors, such as computing, actuator command processing, sensor delays, and actuator dynamics. The delay in computation cannot simply be regarded as a constant, which will reduce the responsiveness of the controller or make it overly conservative.

4. Future Trends

In order to enhance the adaptability to marginalized scenarios, a large amount of high-quality data is an important part in deep learning. It is essential to find and collect high-quality data. “Data Engine” is an integrated system that continuously collects, processes, and continuously improves. It does not rely on human intervention. It uses automatic annotation - through AI perception models to automatically label objects, pedestrians, etc., which significantly shortens the training cycle of the model. Thus, it strengthens the generalization ability and prediction accuracy of the AI model [10].

5. Conclusion

This paper begins by introducing the development history of autonomous driving, the perception of autonomous driving and the AI-involved computer vision and behavior prediction modules. It also summarizes the current development trends and technologies. Autonomous driving is a mainstream trend in future transportation. In recent years, the development of autonomous driving has advanced rapidly and has begun to attract more and more public attention. The emergence of AI has facilitated the development of autonomous driving, especially the appearance of deep learning and self-supervised learning. However, it still faces the challenges of insufficient data and uneven quality. In terms of sensors, the current single sensors all have their own undeniable flaws, and the fusion of multiple sensors has become a necessary condition for the perception function of autonomous vehicles.

ISSN 2959-6157

References

[1] Gao H, Wang Z, Li Y, et al. A survey for foundation models in autonomous driving. arXiv preprint arXiv:2402.01105, 2024.
[2] Mao L M, Cheng J, and Feng B. Research and application of artificial intelligence based vehicle road collaborative autonomous driving research platform Modern Transportation Technology, 2025, 22 (04), 71-75.
[3] Wang T E &Wang J W. The highlighting of the grounding of artificial intelligence models: context context context correlation of understanding Journal of Guizhou University (Social Sciences Edition), 2025, 43 (04), 1-9.
[4] Waghmare A A, Ganesan S, Chen J. Role of Artificial Intelligence in Autonomous Vehicles. Preprint. 2024.
[5] Kozłowski M, Racewicz S, Wierzbicki S. Image analysis in autonomous vehicles: A review of the latest AI solutions and their comparison. Applied Sciences, 2024, 14(18): 8150.
[6] Singh A. Transformer-based sensor fusion for autonomous driving: A survey//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 3312-3317.
[7] Shi J, Chen J, Wang Y, et al. Motion forecasting for autonomous vehicles: a survey. arXiv preprint arXiv:2502.08664, 2025.
[8] Chen L, Wu P, Chitta K, et al. End-to-end autonomous driving: Challenges and frontiers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
[9] Ruan S, Wang R, Shen X, et al. A Survey of Multi-sensor Fusion Perception for Embodied AI: Background, Methods, Challenges and Prospects. arXiv preprint arXiv:2506.19769, 2025.
[10] Jeong Y, Kim S, Yi K. Surround vehicle motion prediction using LSTM-RNN for motion planning of autonomous vehicles at multi-lane turn intersections. IEEE Open Journal of Intelligent Transportation Systems, 2020, 1: 2-14.