Development of Machine Vision and AI for Intelligent Driving Technology

Jinlin Yang

College of Life Sciences and Technology, University of Electronic Science and Technology of China, Chengdu, 611731, China
2024140902025@std.uestc.edu.cn

Abstract. Intelligent driving has been popular research in automotive high-tech in recent years, and some researchers have found that machine vision plays an important role in environment detection and other contents of intelligent driving technology, however, there is a lack of unified cognition for the specific application scenarios and corresponding use characteristics of machine vision. Therefore, this study further discusses the development and application of machine vision in intelligent driving technology. It is found that machine vision has been widely applied in intelligent driving technology in self-perception, environment detection, behavioral decision-making of a single vehicle and multi-vehicle cooperative scenarios, etc. In decision-making, artificial intelligence, deep learning, and other technologies are also applied, and the establishment of large models and so on makes the intelligent driving technology better adaptability to complex road conditions and real road conditions. Therefore, specific applications have improved the adaptability, intelligence, and autonomy of intelligent driving. Therefore, this research is aimed at promoting the application of artificial intelligence and machine vision in smart driving technology where new technologies may find a way to enhance and apply.

Keywords: Artificial intelligence, autonomous driving, machine vision.

1 Introduction

At present, the development of intelligent driving is booming, relying on machine vision, large models, and other technologies to develop rapidly and start pilot experiments, many companies' intelligent driving products in environment perception, path planning and intelligent decision-making, and other core areas of continuous innovation, and some of the companies' products have even completed the initial commercialization. A Report on the Development of Automotive Intelligent Driving points out that it has become an industry consensus to promote the development of intelligent functions with intelligent driving as the core, and intelligent driving can be said to have become the core competition point in the development stage of automotive intelligentization [1, 2].

In these popular development areas, vehicle identification, environmental detection technology has been developed for a long time, there has been a certain establishment,

J. Iqbal (ed.), *Proceedings of the 2025 2nd International Conference on Mechanics, Electronics Engineering and Automation (ICMEEA 2025)*, Advances in Engineering Research 272,
https://doi.org/10.2991/978-94-6463-821-9_59

and for behavioral decision-making and other aspects of the participation of artificial intelligence large models, but also need a large amount of data support for the training of algorithms, but due to limitations in the scope and scenarios, it is difficult to guarantee the accuracy and consistency, the lack of effective standards and systems, as well as data security and privacy protection, etc. At present, China is not in line with the international leading level. However, due to limitations in scope and scenarios, difficulty in guaranteeing accuracy and consistency, lack of effective standards and systems, as well as data security and privacy protection, etc., there is a gap between China and the international leading level. It is these gaps partly exist, the current research field is still deficient: the application of large models and other applications have not yet been fully integrated into the development of intelligent driving technology; on the basis of recognizing the existing gaps, there is a need to further develop behavioral decision-making based on algorithmic training, and to promote the experience enhancement of intelligent driving technology.

This paper carries out a study on the advantages and shortcomings of the application of machine vision and artificial intelligence in the field of intelligent driving, with the aim of clarifying that the current research can further use machine vision and artificial intelligence in depth to improve intelligent driving technology and enhance its experience, so as to clarify the developmental areas of the current intelligent driving technology and conduct in-depth research.

2 Historical Background of Intelligent Driving Technologies

2.1 Origins and Early Explorations of Intelligent Driving Technologies

The concept of intelligent driving technology can be traced back to the 1970s at the earliest, and the early development can be demarcated by the year 2000. Developed countries such as the United States, the United Kingdom, and Germany took the lead in opening the theoretical research and experimental verification of autonomous driving technology. In 1977, the self-driving car developed by Japan's Tsukuba Engineering Research Laboratory achieved autonomous driving at 20 km/hour through a camera and an analog computer, which became an early landmark achievement [3]. In the 1980s, the U.S. Defense Advanced Research Projects Agency (DARPA) promoted the military application of autonomous driving technology by funding several projects to promote the military application of autonomous driving technology, such as the ALV (Autonomous Land Vehicle) project in 1984, which utilized LiDAR and computer vision to achieve navigation in complex terrain.

China's research started a little later, but since the late 1980s, the National University of Defense Technology (NUDT) successfully developed the first self-driving car, ATB-1 (Autonomous Test Bed-1), in 1992, marking the entry of domestic self-driving technology into the experimental stage. The commonality of this stage is that the technology is highly dependent on sensors and regularized algorithms, is mostly confined to laboratory environments, and cannot yet be commercially applied.

2.2 Technological Breakthroughs and Budding Industrialization

This period, roughly the 2000s-2010s, marked a watershed in technological development when DARPA organized the Autopilot Challenge in the Mojave Desert in 2004. Vehicles such as Carnegie Mellon's Sandstorm and Stanford's Stanley demonstrated the potential of combining LIDAR, GPS, and SLAM (Simultaneous Localization and Map Architecture) technologies, spurring a deeper convergence between academia and industry. In 2009, Google launched the Waymo project, which for the first time applied deep learning to an autonomous driving perception system, realizing the processing of complex scenarios on urban roads [3].

In the same period, Chinese technology enterprises began to make efforts. In 2013, Baidu set up the automatic driving division, and in 2017, it launched the Apollo open platform, which integrates high-precision maps, simulation testing, and vehicle-circuit coordination technology, accelerating the ecological construction of the domestic industrial chain [4]. This stage is characterized by the differentiation of technology routes: the “vision-dominant faction” represented by Tesla and the “multi-sensor fusion faction” represented by Waymo form a competitive pattern [5].

2.3 Standard Setting and Commercialization Pilot

This period is roughly 2010-early 2020s. The J3016 standard issued by the Society of Automotive Engineers International (SAE) in 2014 defined the L0-L5 automated driving classification system, which became the reference framework for global technology development. China released the national standard for Automated Vehicle Driving Classification (GB/T 40429-2021) in 2021, which is in line with the SAE standard and further standardizes industrial development.

In terms of commercialization, Tesla launched its Autopilot system in 2015, the first time to mass-produce L2-level assisted driving function on board; Waymo launched its Robotaxi service in Phoenix in 2018, opening public testing of L4-level autonomous driving. China, on the other hand, has promoted the technology through policy pilots, such as Beijing and Shanghai, which allow self-driving vehicles to carry out road tests in limited areas.

3 Status of Development of Intelligent Driving Technology

3.1 Technology Architecture and Core Progress

The core of intelligent driving technology lies in the synergistic optimization of perception, decision-making, and execution, and the breakthroughs in machine vision, artificial intelligence (AI), and big model technologies have injected new vitality into this field. Next, this part analyzes in detail the aspects of machine vision at the perception layer, AI algorithms at the decision-making layer, and the application of big models in intelligent driving.

Machine vision technology for the perceptual layer. Machine vision is one of the core technologies in the perception layer of intelligent driving, and its task is to realize environment perception, target detection, and semantic segmentation through image data captured by cameras. In recent years, the combination of deep learning and computer vision has driven the rapid development of machine vision technology.

Target Detection and Semantic Segmentation. Target detection is a fundamental task in machine vision, aiming at recognizing targets such as vehicles, pedestrians, traffic signs, etc. in images. The YOLO (You Only Look Once) series of algorithms has become the mainstream choice due to its high efficiency. YOLOv8 improves the inference speed to 0.2ms/frame while maintaining high accuracy, which is suitable for in-vehicle scenarios with high real-time requirements. Semantic segmentation, on the other hand, realizes scene understanding through pixel-level classification, for example, the DeepLabv3+ model achieves 82.1% mIoU (mean intersection and merger ratio) on the Cityscapes dataset [6]. Baidu Apollo uses an improved Mask R-CNN model combined with multi-scale feature fusion technology to significantly improve target detection accuracy in nighttime scenes [7]. In addition, the Huawei ADS 2.0 system introduces the GOD (General Obstacle Detection) network, which realizes the recognition of 1400 types of shaped obstacles, including long-tailed targets such as construction fences and animals, through self-supervised learning [8].

Multimodal Fusion and BEV Sensing. Single sensors have limitations, such as creating problems such as camera susceptibility to light and LiDAR performance degradation in rainy and foggy weather. Multi-modal fusion technology improves the comprehensiveness of the sensing system by integrating data from cameras, LiDAR, and millimeter-wave radar. The BEV (Bird's Eye View) sensing model achieves cross-modal feature fusion by mapping multi-sensor data into a unified bird's eye view space. For example, Tesla's Occupancy Network model generates dynamic 3D occupancy raster maps through an 8-camera system to recognize complex scenes without relying on high-precision maps [9]. Xiaopeng Auto proposed XNet 2.0 architecture, which utilizes the Transformer model to achieve cross-camera feature fusion, and the target detection accuracy is improved to 99.3% [8]. In addition, the Horizon Journey 5 chip supports BEV model deployment with 128 TOPS arithmetic, which significantly reduces the hardware cost [8].

Visual Enhancement in Low Light and Bad Weather. Low light and bad weather are technical difficulties in machine vision. The research team at Tsinghua University proposed a low-light enhancement algorithm based on GAN (Generative Adversarial Network), which improves target detection accuracy by generating high-quality images; Huawei ADS 2.0 system introduces multi-spectral fusion technology, which combines the visible light and infrared camera data and improves the target recognition rate up to 95% in rainy and foggy weather [8].

AI algorithms at the decision-making level. The decision-making layer is the judgmental core of intelligent driving, whose task is to generate driving strategies and drive vehicles based on sensed data. Traditional rule-based algorithms find it difficult to cope with complex scenarios, while AI algorithms are data-driven to improve the degree of adaptation to complex road conditions that are mostly seen in daily life.

Reinforcement Learning and Game Decision Making. Reinforcement Learning (Reinforcement Learning) optimizes the decision-making strategy through a trial and error mechanism and has outstanding performance in gaming scenarios. For example, Waymo's ChauffeurNet model achieves decision optimization in complex intersection scenarios through simulation training, and the success rate of intersection passing is increased to 98%; DDT's autonomous driving adopts Multi-intelligence Reinforcement Learning (MARL) technology to achieve optimization of the interaction strategy between vehicles and pedestrians and non-motorized vehicles [9].

End-to-end decision modelling. The end-to-end model maps sensor inputs to control commands directly through neural networks, reducing the loss of information transfer between modules. Mobileye's SuperVision system adopts an end-to-end architecture, and the decision latency in complex scenarios is reduced to 80ms from 500ms in traditional architectures [9]. Tesla's FSD Beta system realizes full scene coverage “from parking space to parking space” through an end-to-end model [8].

Brain-like computing and impulse neural networks. Neuromorphic Computing (NC) improves decision-making efficiency by simulating the spatio-temporal information processing mechanism of the human brain. A research team from Tsinghua University has proposed a decision-making model based on a pulsed neural network (SNN), which reduces the misjudgment rate by 37% in unexpected pedestrian avoidance scenarios. Huawei ADS 3.0 system introduces brain-like computing technology to achieve real-time optimization of decision logic through dynamic weight adjustment [8].

Application of large models to intelligent driving. Large models are trained to achieve generalized capabilities through massive amounts of data, opening new possibilities for intelligent driving technology.

Big Language Modeling and Scene Understanding. Large Language Model (LLM) realizes scene understanding and interaction through natural language processing technology. For example, Baidu Apollo introduced the Wenxin large model to achieve dynamic adjustment of navigation paths through voice commands; Tesla's Dojo supercomputer achieves improved generalization ability for complex scenes through large model training [8].

Integration of Multimodal Large Models and Perceptual Decision Making. The multimodal big model realizes the integration of perception and decision-making by fusing multi-source data such as vision, speech, and text. For example, Huawei's Pangu grand model improves the decision-making accuracy to 99% in complex intersection scenarios through multimodal fusion technology [8]; Xiaopeng Automobile proposes the XGPT model, which achieves real-time optimization of the driving strategy through a multimodal grand model [8].

Large model-driven simulation and testing. Large models accelerate the testing and validation of intelligent driving systems by generating massive simulation scenarios. Waymo's Carcraft platform generates 10 billion kilometers of simulation data through large models, covering typical traffic scenarios around the world [9]. The Chinese company Tencent autonomous driving uses the TAD Sim simulation platform to generate traffic scenarios with Chinese characteristics through large models, such as the scenario of electric vehicles jaywalking, which significantly improves the testing efficiency [9].

3.2 Industrial Application and Market Pattern

The commercialization of intelligent driving technology presents a differentiated path of “popularizing L2+ for passenger cars and breaking through L4 for commercial vehicles”, and reconfigures the value distribution of the automotive industry chain.

Passenger vehicle market: scaled penetration of L2+ functions. In 2024, the penetration rate of new cars with L2-assisted driving will reach 34% globally, with China leading the way with a penetration rate of 42% [8]. The technology generation differentiation is significant: the first echelon contains Huawei, Xiaopeng, and Ideal, these companies have realized the urban area NOA (Navigate on Autopilot) full coverage, and support for high-precision maps without the “ point-to-point ” navigation; the second echelon contains BYD, Geely, these companies are mainly high-speed NOA, relying on high-precision map coverage; the third echelon contains BYD, Geely, these companies are mainly high-speed NOA, rely on The second echelon includes BYD and Geely, which focus on high-speed NOA and rely on high-precision map coverage; the third echelon, i.e., traditional joint venture brands, still focus on the basic functions of ACC (Adaptive Cruise Control) + LCC (Lane Centering).

Breakthroughs in commercial vehicles and specific scenarios. L4 technology for closed scenarios has entered the commercial operation stage. For example, Xijing Technology has deployed 50 unmanned collector trucks at Xiamen Far Sea Terminal, realizing all-weather operation through 5G+V2X; Treadmill's unmanned mining trucks have accumulated more than 1 million kilometers of operation in Inner Mongolia's mines; and Waymo has operated more than 500 unmanned cabs in San Francisco, with a one-way cost reduced to $0.3/km.

Supply chain reconstruction and localization substitution. The localization rate of chips, operating systems, and other core links has increased significantly, including the Horizon Journey 5 chip with an arithmetic power of 128 TOPS, which supports the deployment of BEV sensing models, Huawei Hongmeng OS realizing real-time optimization under microkernel architecture, and Four-Dimensional New obtaining the country's first L3-level high-precision map review number.

3.3 Policy and Standardization Challenges

The global policy environment shows a balance between "technology promotion" and "risk regulation", but legal and ethical issues still need to be broken through. In China, the "Dual Intelligence Pilot" promotes the construction of vehicle-road cooperation, but the restriction of cross-border data flow leads to increased compliance costs for foreign enterprises. In California, the DMV allows Robotaxi to operate without a safety officer but requires companies to disclose disengagement rate data.

In the EU, the Artificial Intelligence Act classifies systems above L3 as "high-risk AI" and requires third-party certification.

4 Directions for the development of intelligent driving technologies

4.1 Analysis of Key Challenges in Current Research

In the field of machine vision and artificial intelligence, big model, and intelligent driving technology for their application have yet to be developed. Currently, because the proposal and development of machine vision and artificial intelligence lag behind the research of intelligent driving, the compatibility of intelligent driving for these fields is still insufficient and has room for development. In addition, due to the breakthrough of three-dimensional vision and other technologies, more technologies can be applied to the development of intelligent driving, to improve the intelligence and autonomy of intelligent driving technology. Therefore, the next section mainly discusses these directions to study the feasibility of the application of machine vision, deep learning algorithms, and artificial intelligence in intelligent driving technology [10].

4.2 Application of Deep Learning Algorithms

The deep learning algorithm is the key technology to make breakthroughs in the field of machine vision. By building deep neural networks, machine vision systems can automatically learn and extract image features, significantly improving the accuracy and efficiency of target detection, image classification, and other tasks. There are two main reasons for the rapid application of deep learning: easier access to a large number of manually labeled datasets, such as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC); and deep learning algorithms that can process graphics in parallel on GPUs, which improves learning efficiency and prediction capability. By utilizing the autonomous learning characteristics of deep neural networks, the huge and complex neural network models are first trained through high-performance GPUs, and then ported to embedded development platforms, real-time and efficient processing of image and video information can be realized. In recent years, from automatic driving startups, and Internet companies to major OEMs, they are actively exploring the use of GPUs to build neural networks to realize the ultimate driverless [10,11].

4.3 Breakthroughs in 3D vision technology

Three-dimensional vision technology realizes accurate perception of complex scenes by acquiring the depth information of objects. In recent years, the advancement of structured light, LIDAR, and other technologies has made 3D vision widely used in robot navigation, virtual reality, and other fields. In the field of unmanned driving, 3D vision technology can generate a 3D model of the road environment in real time, providing accurate navigation information for vehicles [12].

4.4 Integration of cross-modal recognition technology

Cross-modal recognition technology improves the comprehensive cognitive ability of the machine vision system by integrating multiple data forms such as images, speech, and text. For example, in the field of smart homes, cross-modal recognition technology is able to analyze the user's voice commands and gesture movements at the same time, providing a more natural human-computer interaction experience [12].

4.5 Increase in Intelligence and Autonomy

In the future, the machine vision system will be more intelligent and autonomous. By combining reinforcement learning and adaptive algorithms, machine vision systems are able to make decisions and perform tasks autonomously in complex environments. In the field of intelligent manufacturing, intelligent machine vision systems can automatically adjust the inspection parameters and processes according to the real-time status of the production line, so as to realize the overall optimization of the production process. Similarly, in the field of intelligent driving, the intelligent machine vision system can also be formed through the formation of machine vision-assisted decision-making, adaptive to more complicated road conditions, and realizing how to make decisions through the existing model training to correctly arrive at the decision and guide the vehicle driving in the unpredicted road conditions [13].

4.6 Technical Evolution of Big Model

Big models are based on the Transformer architecture, and through massive data training, they show strong generalization and reusability. For example, OpenAI's GPT series of models can achieve high performance in multiple fields such as natural language processing and computer vision through large-scale pre-training. Multi-modal large models improve the comprehensive cognitive ability of AI systems by integrating text, image, speech, and other data forms. Among them, OpenAI's Sora model can simultaneously process video and text data to generate high-quality multimodal content, providing new tools for film and television production, advertising design, and other fields. Such multimodal content can provide ideas for data exchange, communication and information sharing between multi-vehicle collaboration [13].

5 Conclusion

Intelligent driving technology has been a breakthrough in the automotive industry and technology field in recent years. It is profoundly affecting the shape of modern transportation systems and future travel modes. The results of this paper show that the realization of intelligent driving relies on the support of advanced machine vision and artificial intelligence technologies as well as large models. With the continuous progress of these technologies, intelligent driving is developing towards higher safety, better performance, and wider adaptability. This paper examines the lack of application and existing results of machine vision and artificial intelligence technologies in intelligent driving technology and concludes that intelligent driving technology can utilize these two technologies to further improve the results.

To use these technologies to continue to develop intelligent driving technology, first of all, it is necessary to apply the superior ability of deep learning in image recognition and other aspects to gradually form a pattern of multi-vehicle networking and informationization. Advances in deep learning technology have greatly enhanced the ability of machine vision to be applied in automatic driving, and companies can utilize the adaptive characteristics of deep learning to enable machine vision systems to be updated and adjusted in real time according to new road conditions or new environmental conditions, thereby effectively improving the intelligence of the system. However, at the same time, the effectiveness of cameras and sensors will be affected due to the poor performance of machine vision in low-light and bad weather, such as rain, snow, haze, etc., and it is difficult for the vision system to deal with complex dynamic environments, especially in the process of high-speed driving, and the recognition and processing of long-distance objects is still a difficult problem. Therefore, self-driving cars need to combine other sensors, such as LiDAR, millimeter wave radar, and other sensors that have become more mature to enhance the system's multi-dimensional perception ability.

Secondly, the application of artificial intelligence and large models of massive data processing and fusion capabilities, so that the self-driving vehicle can mimic the human driver's decision-making process and make a reasonable response in different road scenarios. The core of the automatic driving system lies in the ability to make decisions and plans in the rapidly changing road environment. Enterprises can strengthen the use of AI technologies such as deep learning and reinforcement learning to train the automatic driving system, so that it can learn from a large amount of driving data to optimize path selection, and obstacle avoidance strategies, and make its decisions to comply with traffic laws and regulations and adapt to complex road conditions. At the same time, enterprises should also note that AI still faces many risks and challenges. AI systems still have certain risks when dealing with extreme and special situations, and AI may not be able to make optimal decisions when encountering traffic accidents, sudden weather, or other uncommon situations. Second, AI's decision-making process often lacks sufficient interpretability, and in certain high-risk scenarios, such difficult interpretability may lead to trust issues and ethical controversies. Therefore, how to improve the transparency and security of AI systems remains an important challenge for future smart driving technology.

The development of intelligent driving technology relies on continuous innovation and breakthroughs in machine vision, artificial intelligence, and large model technology. Although the current level of technology has made significant progress, it still faces many challenges in terms of environmental adaptability, safety, and regulations. In the future, intelligent driving technology will develop in the direction of greater maturity, safety, and popularization, bringing a more efficient, safe, and environmentally friendly travel experience.

References

1. Zhang D.: 'Application and development trend of artificial intelligence in automobile manufacturing'. Robotics Technology and Application, **06**, 5-11 (2024)
2. Tencent Automotive, 'Hundred People's Association: Intelligent Driving Becomes the Competitive Focus of Automotive Intelligence, China's Higher-order Intelligent Driving Solution Leads the Global Automotive Intelligence Transformation', 2025/1/22, cited in 2025/3/16. cited in https://news.qq.com/rain/a/20250122A04WGE00?suid=&media_ id= (2025)
3. New Financial News, I'ntelligent autonomous driving: technology status and future development prospects', 10/7/2024. Retrieved on Feb/2/2025. Retrieved from: https://www.sohu.com/a/817323618_267471 (2025)
4. Li X.: 'Development status, challenges and response of automatic driving'. People's Forum, **18**, 68-72 (2023)
5. Slender, 'A Survey of Autonomous Driving: Common Practices and Emerging Technologies', 9/9/2020.Retrieved on 2/9/2025. https://blog.csdn.net/slender_1031/article/details/108504466 (2025)
6. Netease Technology Express, 'Application and Development Prospects of Machine Vision Technology Discussion', 2024/10/28, Retrieved on 2025/3/16.Retrieved from: https://m.163.com/dy/article/JFKC0J850556A727.html (2024)
7. Yu T, Chai H, Meng R.: 'Research on the application of machine vision in intelligent driving'. Digital Communication World, **07**, 115-117 (2024)
8. Discovery Report, 'An Article on Smart Driving Update and Trends', 2025/1/23, cited in 2025/3/16, quoted from https://news.qq.com/rain/a/20250123A01UL000 (2025)
9. Smart Driving Cutting Edge, 'An article to talk about the application and impact of big modeling technology in autonomous driving', 2024/1/8, cited in 2025/3/16, cited in https://so.html5.qq.com/page/real/search_news?docid=70000021_702659b436e50452& faker=1 (2025)
10. Wang C, Li Z, Zhao X, et al.: 'A review of driver state monitoring algorithms based on image and deep learning'. China Highway Journal, **1**, 1-26 (2024)
11. Fan J, Shan S, Lin X.: 'Analysis of key technology of machine vision for intelligent driving vehicles'. Automotive Practical Technology, **48**(21):173-178 (2023)
12. Ziad A, Duong V P, Ali A M, et al. 'Intelligent Driver Assistance and Energy Management Systems of Hybrid Electric Autonomous Vehicles'. Sustainability, **14**(15):9378-9378 (2022)

13. Huang H, Teng H, Liu Y, et al.: 'Research on the application of machine vision technology in automobile manufacturing industry'. Automotive Process and Materials, **06**, 8-15 (2022)