Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 14
TPCEE 2026

Application value and development trend of machine vision technology in AI era

Jiachang Chen [1, *]

[1] Zhejiang Jinyun High School, Lishui, China

* Corresponding Author Email: smion666@163.com

Abstract. This paper summarizes the application value and development trend of machine vision technology in AI era. With the breakthrough of deep learning and other technologies, machine vision has achieved a leap from "seeing" to "understanding". This technology shows its core value in manufacturing, medical care, autonomous driving and other fields by improving detection efficiency, realizing micron-level precision control and promoting industrial intelligent transformation. However, its large-scale application still faces challenges such as data dependence, poor environmental adaptability and computing power conflict. In the future, technology will develop in the direction of 3D visual perception, edge computing collaboration, multi-modal fusion and generative AI drive, and it is committed to strengthening interpretability and general intelligence to promote the deep landing of technology and industrial innovation.

Keywords: Machine vision; AI; Application value; Development trend; Deep learning.

1. Introduction

The rapid progress of AI has triggered a paradigm shift in many industries, from manufacturing, health care to autonomous driving. When entering the era of Industry 4.0, the demand for intelligent systems that can simulate human cognitive functions has surged. Among these abilities, visual perception is the cornerstone, which enables machines to explain and understand the complex visual world.

As the "eye" of AI, machine vision technology has significantly evolved from traditional rule-based image processing to modern data-driven methods [1]. Unlike human inspectors, machine vision systems provide unparalleled speed, accuracy, and the ability to work continuously without fatigue [2]. However, despite the surge in the application of machine vision, the comprehensive analysis of its specific value proposition in the current AI background is still limited. Many existing researches only focus on algorithm improvement narrowly, and often ignore the practical significance and define the evolution trend of its future deployment. In addition, with the expansion of hardware capabilities and the maturity of deep learning algorithms, the boundaries of machine vision applications are being further pushed up, which makes it necessary to re-examine the development track of this technology.

Therefore, this paper systematically summarizes the application value and development trend of machine vision technology in the AI era. Specifically, this paper examines key technical drivers, analyzes diverse application scenarios, and discusses key challenges such as data dependence and environmental robustness. Finally, this paper discusses the emerging trends including 3D vision and edge computing, and provides insights for the future direction of this revolutionary technology.

2. Technical evolution and key technologies

The development process of machine vision technology is essentially a leap from "seeing" to "understanding", and its evolution path can be clearly divided into two stages: traditional machine vision and a new generation of AI-driven machine vision. The structure of a machine learning platform is shown in Figure 1.

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

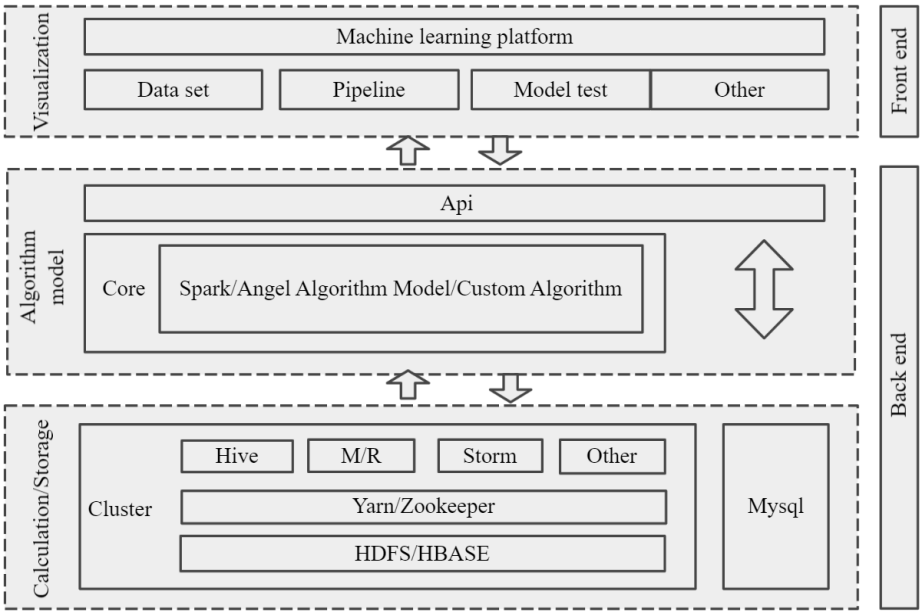

Figure 1. Machine learning platform structure

In terms of technological evolution, the early traditional machine vision mainly relied on hand-designed feature extraction algorithm and rule-based matching technology. Although this method is stable in specific scenes, it has weak generalization ability in the face of illumination changes, background interference and unstructured scenes, and it is highly dependent on expert experience to adjust parameters. With the breakthrough of deep learning, machine vision has entered an intelligent era driven by data. With the introduction of Convolutional Neural Network (CNN), Transformer architecture and large model technology [3-4], machine vision has the ability of end-to-end feature self-learning, which fundamentally solves the problem of difficult feature extraction of traditional algorithms in complex scenes and realizes the paradigm shift from "machine judgment" to "intelligent cognition".

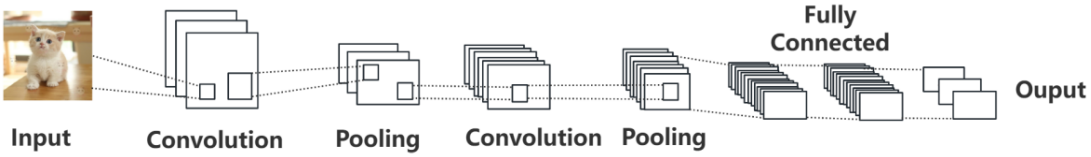

Figure 2. Typical CNN structure schematic

At the key technical level, machine vision in the AI era is mainly supported by three pillars: algorithm model, computing hardware and data processing. Based on CNN, the ability of visual task processing has been continuously improved. A typical CNN structure is shown in figure 2. The target detection algorithm and the emerging visual Transformer (ViT) architecture in recent years not only greatly improve the recognition accuracy, but also enhance the ability to understand the global context information of the image [5]. The popularity of high-performance computing chips provides a physical basis for efficient training and real-time reasoning of deep learning models, making real-time processing of high-resolution images possible. The construction of large-scale annotation data set is the key to improve the performance of the model. At the same time, in view of the difficulty in obtaining scene data such as industry, transfer learning, small sample learning and synthetic data training technology have become important technical supplements, effectively reducing the dependence on massive labeled data.

3. Application value analysis

At the moment when AI is deeply empowered, machine vision technology has broken through the limitations of a single technical tool and become the core force driving multi-industry efficiency innovation, quality upgrading and model reconstruction. Its application value runs through the core

links of production and operation, industrial transformation and social development, showing irreplaceable strategic significance (Figure 3).

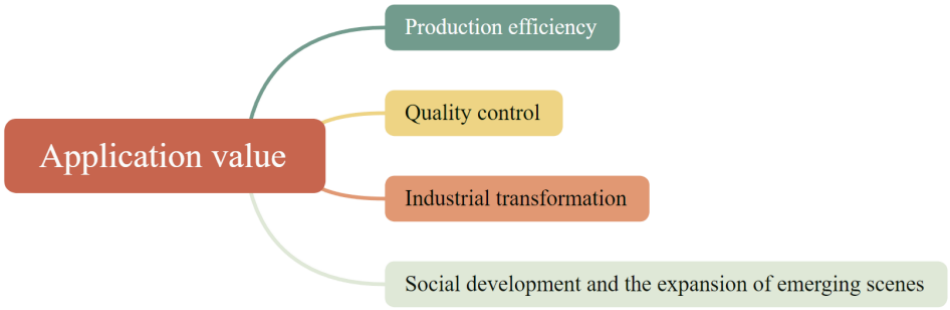

Figure 3. Embodiment of application value of machine vision technology

From the perspective of production efficiency, machine vision technology completely breaks the efficiency bottleneck and physiological limitation of traditional manual operation with millisecond image recognition speed and all-weather uninterrupted working characteristics [6]. In high-intensity and repetitive scenes such as manufacturing assembly lines and logistics sorting centers, it can accurately capture product defects, quickly complete size detection and material positioning, improve the detection efficiency by several times or even dozens of times compared with manual work, and avoid the problems of misjudgment and missed judgment caused by manual fatigue, which not only ensures the continuity of production rhythm, but also greatly reduces the rework cost and capacity loss caused by detection errors, laying a solid foundation for efficiency in large-scale and standardized production.

At the level of quality control, machine vision technology builds a standardized and traceable quality assurance system. Relying on deep learning algorithm and high-precision imaging ability, it can accurately identify the subtle defects and dimensional deviations of products, and the detection accuracy can reach micron level, far exceeding the perceptual limit of human vision [7]. Whether it is solder joint inspection of electronic components, precision verification of automobile parts or packaging integrity verification of pharmaceutical products, this technology can provide objective and consistent inspection results, prevent unqualified products from entering the market from the source, effectively reduce quality risks and after-sales costs, help enterprises establish high-quality reputation and build a quality defense line for industrial core competitiveness.

From the perspective of industrial transformation, machine vision technology is the key engine to promote the upgrading of traditional industries to intelligence. It can not only replace manual work in dangerous and harsh environments, but also ensure personnel safety and reduce labor risks. It can also realize unmanned and flexible transformation of production processes through deep integration with industrial robots and automated production lines. For example, in the scene of high-risk chemical inspection and high-altitude facility inspection, machine vision can be equipped with intelligent equipment to complete remote monitoring and fault identification to avoid personnel safety risks; In customized production scenarios, it can quickly adapt to the testing requirements of different products, help enterprises achieve flexible production in small batches and varieties, accelerate the transformation of traditional industries from labor-intensive to technology-intensive, and inject core kinetic energy into the high-quality development of industries.

At the level of social development and emerging scene expansion, the value of machine vision technology continues to extend, deeply empowering smart life and frontier fields. In the field of smart medical care, it assists doctors to complete accurate image diagnosis and surgical navigation, greatly improving the efficiency and accuracy of diagnosis and treatment and protecting public health; In the field of automatic driving, machine vision, as the "visual center" of vehicles, can perceive the road environment in real time, identify traffic signs and obstacles, and provide the core guarantee for driving safety; In the construction of smart cities, it supports traffic flow monitoring, public safety early warning and other scenarios to optimize the efficiency of urban governance.

4. Current challenges

Although machine vision technology has made great progress driven by AI, its large-scale landing and deep application still face several key challenges, as shown in Table 1.

Table 1. The current challenges of machine vision technology

Challenge dimension	core problem
Data Dependence and Quality Bottleneck	Relying on massive and high-quality annotation data; The high cost and long period of acquisition in the professional field restrict the generalization ability of the model.
Insufficient environmental robustness and adaptability	In the real and complex environment such as illumination, occlusion and cluttered background, the recognition accuracy is significantly reduced and the stability is poor.
The contradiction between computing power and real-time performance	High-performance model needs strong computing power, which conflicts with the requirements of low latency and low power consumption at the edge.
Interpretability and lack of trust	The "black box" feature of deep learning model leads to opaque decision-making process, and it is difficult to establish trust in high-reliability scenarios.
Small sample and long tail problem	It is easy to miss or misjudge the "long tail" event at low frequency, and it is difficult to collect enough samples for training.

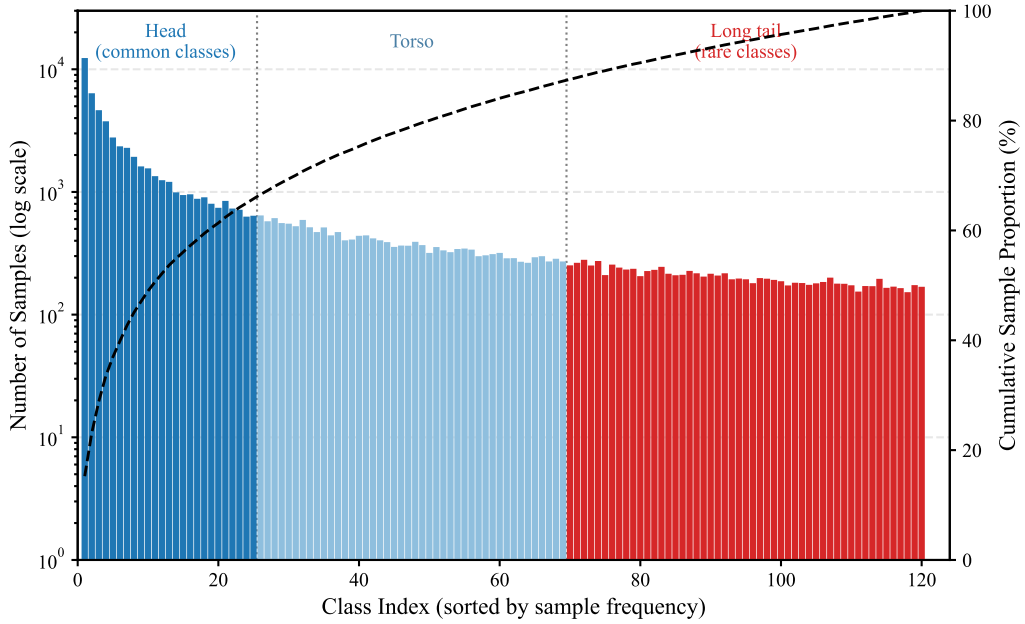

Figure 4. Long tail distribution in visual recognition task

Modern machine vision system based on deep learning is highly dependent on massive, high-quality and accurately labeled training data. In many professional or subdivision fields, it is expensive and long to obtain enough and diverse data sets, which seriously restricts the generalization ability and actual performance of the model. At present, most visual models perform well in the controlled ideal environment, but in the real world, complex factors such as illumination change, visual angle deviation, occlusion, background clutter and weather conditions will significantly reduce the recognition accuracy [8]. How to improve the stability and anti-interference ability of the model in dynamic and unstructured environment is still a major problem. High-performance deep learning models usually need strong computing resources, which conflicts with the limited computing power and power consumption of edge devices and embedded devices. In the scenes that are extremely sensitive to delay, such as automatic driving and industrial online detection, how to realize millisecond real-time reasoning while ensuring accuracy is the core pain point of the project landing. Deep neural network is often regarded as a "black box", and its decision-making process lacks transparency and interpretability. This is particularly prominent in applications with high reliability requirements such as medical diagnosis and safety monitoring. When the system gives wrong judgment, it is difficult to trace the cause and correct it, which further hinders users' trust and adoption of technology [9]. In real scenes, "common" situations are often only a few, and a large number of

"long tail" events occur with extremely low frequency (Figure 4). The traditional model is easy to miss or misjudge these rare categories, and it is almost impossible to collect enough long tail samples, which puts forward higher requirements for the learning ability of the algorithm with few samples.

5. Future development trend

5.1. Deep integration of 3D vision and depth perception

Traditional two-dimensional vision has inherent limitations in spatial understanding and scene reconstruction. In the future, 3D sensing technology based on structured light [10], ToF (time of flight), lidar and binocular stereo vision will accelerate its maturity, and push machine vision from "plane recognition" to "stereo understanding". Especially in industrial sorting, warehousing and logistics, autonomous driving and other scenes, 3D vision will give machines the ability to accurately perceive the relationship between object posture, volume and space, and realize higher-level interaction and decision-making.

5.2. Collaborative promotion of edge computing and end-side intelligence

With the rapid increase in the number of Internet of Things (IoT) devices, it has become an irreversible trend to migrate visual reasoning from the cloud to the edge and even the terminal chip. The combination of dedicated AI chip and lightweight model will enable terminal devices such as cameras and sensors to have real-time and low-power local visual analysis capabilities. It not only reduces bandwidth and delay costs, but also provides better protection in data privacy and security.

5.3. Multi-modal perception and cross-modal learning

It is difficult for a single visual mode to cope with all the information needs in a complex real environment. In the future, machine vision will be deeply integrated with multi-modal signals such as voice, touch, radar and infrared, and a more robust and comprehensive environmental understanding system will be built through cross-modal alignment and joint representation learning. This integration will play a key role in intelligent driving, man-machine cooperative robots and smart cities.

5.4. Generative AI and innovation of training paradigm driven by synthetic data

The scarcity of high-quality labeling data has always been a bottleneck restricting the development of machine vision. Generative AI technology, represented by diffusion model and GAN (Generative Adversarial Network), can efficiently generate realistic synthetic training data and significantly reduce the cost of data collection and labeling. Meanwhile, breakthroughs in paradigms such as "self supervised learning" and "few sample learning" will enable models to achieve rapid generalization with minimal or even no manual annotation, greatly expanding the applicability of machine vision in long tail scenes and small sample domains.

5.5. Explanatory, safety and strengthening of ethical governance

With the deep penetration of machine vision in high-risk fields such as medical diagnosis, security monitoring and automatic driving, the interpretability, fairness and safety of the model will become the prerequisite for technology landing. Future research will pay more attention to the development of interpretable AI method, the defense mechanism against attacks, and the construction of ethical norms for data use and algorithm decision-making, so as to promote the benign interaction between technological development and social governance.

5.6. Exploration of flexibility and universal visual intelligence

At present, machine vision systems are mostly "professionals" and customized for specific tasks. Inspired by the big model, a "general visual agent" with general visual understanding ability is

expected to emerge in the future, which can adapt to various downstream tasks with a few hints and realize the qualitative change from "dedicated detection" to "open understanding", which will fundamentally reshape the technical framework and business model of machine vision.

6. Conclusion

Machine vision technology has become the core engine to drive multi-industry intelligent upgrading, which shows irreplaceable strategic value in production efficiency, quality control, industrial transformation and emerging scene expansion. Through the deep integration of deep learning algorithm and high-precision imaging ability, machine vision breaks through the efficiency bottleneck and physiological limitation of traditional manual work, realizes the paradigm shift from "machine judgment" to "intelligent cognition", and improves the manufacturing defect detection accuracy, medical diagnosis efficiency and automatic driving safety. However, the large-scale implementation of technology still faces challenges such as high data dependence, insufficient environmental adaptability, contradiction between computing power and real-time, lack of model interpretability and long tail problem. The future development trend will focus on the integration of 3D vision and depth perception, the collaboration of edge computing and end-side intelligence, the collaboration of multimodal perception, the innovation of synthetic data training paradigm driven by generative AI, and the strengthening of interpretability and ethical governance. With the penetration of large-scale model technology, the emergence of universal visual agents will promote the transition of machine vision from "dedicated detection" to "open understanding", and inject new kinetic energy into industrial automation, smart medical care and smart city construction.

References

[1] Jaemyung S, Sultan M M, U. T R, et al. Trends and Prospect of Machine Vision Technology for Stresses and Diseases Detection in Precision Agriculture [J]. AgriEngineering, 2022, 5(1): 20-39. DOI:10.3390/AGRIENGINEERING5010003.

[2] Ferreira T R F, Couto D M L, Saporetti M C. Using deep learning applied in computer vision for the inclusion of individuals with tetraplegia through assistive robotics via facial features [J]. Soft Computing, 2025, 29(15-16): 1-29. DOI:10.1007/S00500-025-10906-3.

[3] Dong G, Pan X, Liu S, et al. A review of machine vision technology for defect detection in curved ceramic materials[J]. Nondestructive Testing and Evaluation,2025,40(7):2797-2823.DOI:10.1080/10589759.2024.2404497.

[4] Hu Y, Yan H, Wang P, et al. A study on the application of machine vision technology in defect detection of CFRP composites [J]. Journal of Reinforced Plastics and Composites, 2026, 45(15-16): 3640-3654. DOI:10.1177/07316844251391518.

[5] Shi W, Qi H, Liu Z. HALCON Machine Vision Technology in Automated Inspection Study [J]. International Core Journal of Engineering, 2025, 11(3): 47-57. DOI:10.6919/ICJE.202503_11(3).0006.

[6] Pekias A, Maraslidis S G, Tsipouras G M, et al. Power Supply Technologies for Drones and Machine Vision Applications: A Comparative Analysis and Future Trends [J]. Telecom, 2023, 4(3): 459-476. DOI:10.3390/TELECOM4030024.

[7] Chen H. Application of Automatic Recognition and Positioning Technology of Industrial Robot based on Monocular Vision Technology in Vehicle Manufacturing [J]. International Journal of Vehicle Structures & Systems, 2023, 15(3): 446-453. DOI:10.4273/IJVSS.15.3.27.

[8] Guo S, Xu J, Han C. Application and effect simulation of image recognition technology based on machine vision feature parameters in art teaching [J]. Soft Computing,2023,27(12):8471-8479. DOI:10.1007/S00500-023-08149-1.

[9] Yujie C. Explore the Application of Machine Vision Technology in the Automation of Machinery Manufacturing[J]. Education Reform and Development, 2024, 6(4): 126-132. DOI:10.26689/ERD.V6I4.6819.

[10] Zhang D, Huang S, Sun X, et al. Advances in 2D and 3D machine vision technologies for morphological characterization of granular food products: from laboratory to application [J]. Journal of Food Measurement and Characterization, 2025, 19(12): 1-27. DOI:10.1007/S11694-025-03697-6.