Research Review on AI-Enhanced Machine Vision and Robotics for Predictive Maintenance & Quality Inspection in Automated Manufacturing

Yan Wang
Jiangxi Technical College of Manufacturing, School of Intelligent Manufacturing, Nnanchang
Corresponding author: Yan Wang (e-mail: 15083820126@163.com).

ABSTRACT With automated manufacturing advancing toward high precision and high efficiency, the integration of artificial intelligence, machine vision and industrial robots has become the core pillar of intelligent manufacturing. Traditional manual inspection and scheduled maintenance suffer from high missing detection rates, delayed response and excessive operating costs, failing to satisfy strict quality control standards of modern manufacturing. Deep learning empowers machine vision to extract defect features and realize automatic classification efficiently, while industrial robots break free from fixed programming to achieve autonomous perception and decision-making. Supported by multi-sensor data fusion and machine learning, predictive maintenance enables early equipment fault identification and drastically cuts unplanned downtime. Vision-guided robotic inspection systems outperform human operators in both detection speed and precision. This paper systematically sorts out technical routes, system architectures and engineering practices of the two major application scenarios, and discusses cutting-edge directions for collaborative integration of vision and robotic technologies, providing systematic references for engineering implementation and relevant research layout.

INDEX TERMS machine vision; predictive maintenance; quality inspection; industrial robot; deep learning; automated manufacturing

I. INTRODUCTION

In the Industry 4.0 era, the infrastructure consisting of sensor networks, industrial Internet and edge computing has become increasingly mature, laying a hardware foundation for digital perception of manufacturing systems. Meanwhile, the rapid iteration of deep learning algorithms unlocks the analytical potential of massive manufacturing data from the software dimension. Against this backdrop, visual intelligence technologies represented by convolutional neural networks (CNNs) have been widely deployed in quality control procedures including surface defect inspection, dimensional measurement and assembly verification, achieving remarkable improvements in both detection precision and speed compared with conventional approaches. At the same time, predictive maintenance models fused with vibration signals, temperature readings and current characteristics enable real-time evaluation of equipment health status, shifting the trigger logic of maintenance decisions from time-driven to condition-driven and drastically cutting maintenance costs and risks of unexpected shutdowns. As execution carriers, industrial robots integrate visual perception and intelligent control to further form a closed-loop link spanning from inspection data to mechanical actions. Systematically reviewing the relevant methodology systems and application progress along this technological evolution thread carries important theoretical value and engineering significance.

II. CORE FUNDAMENTAL TECHNOLOGIES OF AI-ENHANCED MACHINE VISION AND INDUSTRIAL ROBOTS

A. IMAGE PERCEPTION AND DEEP LEARNING FEATURE EXTRACTION FOR MACHINE VISION

The core function of machine vision systems is extracting discriminative features from industrial images to support subsequent classification, detection and measurement tasks. Traditional machine vision relies on manually designed operators such as Sobel edge detection, Hough transform and morphological operations. Restricted by preset rules, these methods show weak adaptability to complex textures and variable lighting, failing to meet high-throughput precision requirements of precision manufacturing.

The introduction of deep learning fundamentally addresses this limitation. Built on CNN architectures, multi-layer convolution realizes hierarchical image feature learning: shallow layers capture basic elements such as edges and colors, while deep layers extract high-level semantic representations to enable adaptive identification of diverse defect morphologies. For object detection, YOLO series algorithms complete target localization and classification in a single forward pass to reach real-time frame rates, which have been widely verified in inspection of metal stamping parts [1]. Attention mechanisms and feature pyramid networks further strengthen multi-scale feature fusion and boost detection rates of tiny defects.

Industrial camera resolution, structured light & LED lighting schemes and lens focal length matching jointly determine input image quality, which further affects model convergence and final precision. Scarce industrial defect samples have long restricted model performance. The Stable Diffusion-based data augmentation method generates high-fidelity defect images to alleviate small-sample training difficulties, laying a solid foundation for field deployment of vision models.

B. PERCEPTION-DECISION-EXECUTION INTELLIGENT CONTROL ARCHITECTURE OF INDUSTRIAL ROBOTS

Industrial robots undertake multiple tasks such as precision operation, material transportation and on-line inspection in automated manufacturing, and the improvement of their intelligence relies on a complete control closed loop ranging from environmental perception to motion execution. Conventional industrial robots mainly adopt the teach-playback mode with pre-programmed and fixed motion trajectories, lacking adaptability to dynamic working conditions, which leads to obvious limitations in manufacturing scenarios featuring complex tasks and variable environments.

The core of the intelligent control architecture lies in the organic integration of the perception layer, decision layer and execution layer to build a closed-loop control system with real-time response capability. The perception layer integrates visual, torque and distance sensors to provide robots with real-time information on workpiece poses, contact forces and obstacle distribution. Relying on reinforcement learning or model-based planning algorithms, the decision layer converts perceptual data into optimal motion sequences. The execution layer translates control commands into precise mechanical motions through servo drive systems. Taking welding path planning as a typical application, robots need to dynamically adjust torch trajectories according to joint geometries, and the optimization of path planning algorithms is of direct significance for guaranteeing welding quality and efficiency [2]. In addition, collaborative applications of AI and robotics in power gear cutting tool life management prove that the intelligent control architecture can effectively extend tool service life and reduce manufacturing costs by monitoring tool wear in real time and dynamically adjusting cutting parameters [3].

In human-robot collaboration scenarios, robots are required to perceive human movements and implement safety avoidance in real time, which imposes strict constraints on the response latency of control systems. Digital twin-based simulation environments offer a safe and efficient verification method for offline testing and parameter tuning of control strategies, helping shorten the cycle from algorithm development to on-site deployment.

C. ADAPTABILITY ANALYSIS OF AI ALGORITHMS FOR MANUFACTURING DATA PROCESSING

Manufacturing systems continuously generate multi-modal, high-dimensional data streams during operation, including images, vibration signals, temperature curves, current waveforms, process parameter logs and other data types. Different algorithms exhibit substantial disparities in processing efficiency and adaptability to the above data. Reasonable algorithm selection serves as a prerequisite for constructing efficient intelligent manufacturing systems, as illustrated in Table 1.

TABLE 1. Adaptability comparison of main AI algorithms for manufacturing data processing

Algorithm Type	Applicable Data	Typical Scenarios	Advantages	Limitations
CNN	Images	Surface defect detection, dimensional measurement	Automatic feature extraction, high precision	Requires massive labeled samples
RNN/LSTM	Time-series signals	Vibration analysis, remaining useful life prediction	Captures time-dependent trends	Requires massive labeled samples
Random Forest	Structured process data	Vibration analysis, remaining useful life prediction	High interpretability, fast training	Poor performance on high-dimensional unstructured data
SVM	Medium-low dimensional feature vectors	Early fault diagnosis	Strong performance on small datasets	Limited multi-class extension capacity
SVM	Fusion of images & time series	Joint predictive maintenance & quality inspection	Multi-modal information fusion	Complex structure, difficult parameter tuning

For image data, CNN and its derivative architectures (ResNet, EfficientNet, etc.) represent the mainstream solutions for current industrial visual inspection, as their hierarchical feature learning mechanism is highly consistent with the spatial structure of images. For time-series sensor data such as vibration and temperature signals, Long Short-Term Memory (LSTM) networks deliver outstanding performance in equipment state trend prediction by capturing temporal dependencies. When predictive maintenance tasks integrate inputs from multiple types of sensors, hybrid architectures that adopt CNN for local feature extraction and LSTM for time-series modeling achieve overall performance superior to single models [4]. Swarm intelligence algorithms such as Particle Swarm Optimization (PSO) play an auxiliary role in hyperparameter searching and maintenance scheduling optimization for hybrid architectures, further boosting the overall system efficiency.

For structured process data, ensemble learning methods including random forests and gradient boosting trees possess favorable interpretability and generalization capacity in process parameter optimization and fault classification, which is especially advantageous for industrial scenarios with limited training samples. The actual selection of algorithms requires comprehensive trade-offs among data volume, labeling cost, inference speed and model interpretability, alongside targeted configuration tailored to the task characteristics of specific manufacturing scenarios.

III. AI-ENHANCED MACHINE VISION FOR QUALITY INSPECTION IN AUTOMATED MANUFACTURING

A. DEEP LEARNING-BASED SURFACE DEFECT IDENTIFICATION & CLASSIFICATION

Surface defect detection is the most representative application of machine vision for manufacturing quality control, targeting scratches, blowholes, indentations, inclusions and other flaws on workpiece surfaces. Through end-to-end training, deep learning models embed pixel-to-defect mapping into network weights, eliminating subjective bias brought by manual feature engineering and strengthening robustness against diverse defect shapes. Two-stage detection networks (Faster R-CNN) achieve superior precision and high recall for high-value component inspection. Single-stage YOLO algorithms prioritize inference speed and fit high-speed production line on-line deployment. For metal stamping inspection, researchers combine Stable Diffusion with improved YOLOv5 to expand training datasets via synthetic defect images, mitigating overfitting caused by insufficient real defect samples and lifting detection accuracy. Semantic segmentation models such as U-Net realize pixel-level defect labeling and precisely calculate defect area and contour for acceptability judgment, widely adopted in textile flaw and semiconductor wafer inspection. Unsupervised anomaly detection based on autoencoders addresses scenarios with extremely scarce defect samples by learning normal sample distributions and identifying abnormal deviation regions, cutting demand for labeled defect data.

B. PRECISION OPTIMIZATION AND DEPLOYMENT SCHEMES FOR REAL-TIME VISUAL INSPECTION

Model lightweighting (knowledge distillation, weight pruning, quantization) compresses network complexity while retaining core recognition performance, enabling deployment of GPU-heavy detection models on industrial edge devices and satisfying strict latency limits of production lines. Data augmentation (random cropping, flipping, brightness perturbation, Mixup) mitigates overfitting. Specialized augmentation simulating variable structured light illumination enhances model robustness against on-site lighting fluctuations. Industrial visual inspection generally adopts cloud-edge collaborative architectures: lightweight inference models run on edge nodes for high-frequency real-time detection; complex training and global quality analysis are executed on the cloud, with updated model parameters distributed to edge terminals for iterative optimization. Regular camera calibration, on-line model drift monitoring and abnormal frame filtering guarantee long-term stable operation of visual systems. Full-system integration and production line stress tests simulating high-speed image streams and extreme lighting are mandatory before formal field deployment to evaluate performance under boundary working conditions [5]. There is an obvious trade-off between defect detection accuracy and real-time inference speed among mainstream deep learning visual inspection algorithms, and the performance comparison of each model is shown in Figure 1. Faster R-CNN achieves relatively high detection accuracy yet runs at merely 15 FPS, making it inapplicable to high-speed production lines. YOLOv8 balances a frame rate of 90 FPS and an mAP of 82%, serving as a universal solution for mass production lines. U-Net delivers the optimal segmentation accuracy with an mAP of 92%, but its frame rate only reaches 30 FPS, so it is mostly adopted for high-precision sampling inspection scenarios. Algorithm selection and lightweight transformation shall be carried out according to actual production line tact time and defect identification requirements in engineering practice.

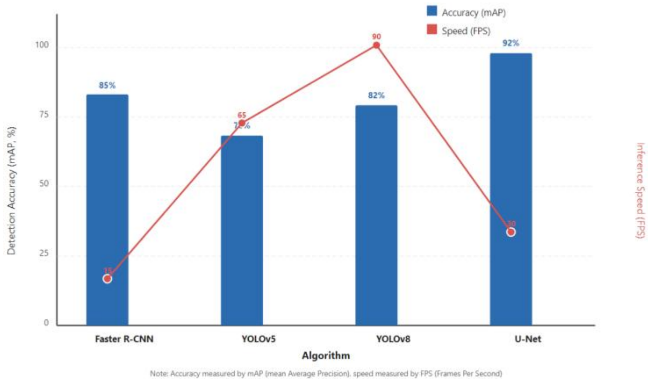

FIGURE 1. Comparison of detection accuracy (mAP) and inference speed (FPS) of mainstream industrial visual inspection algorithms

C. PRACTICAL CASES IN AUTOMOTIVE, METAL AND ELECTRONIC MANUFACTURING

The automotive industry boasts mature machine vision deployment covering stamping panel surface inspection, engine dimensional measurement and assembly verification. Systematic reviews confirm that AI-based automotive defect inspection has formed complete systems covering optical design, algorithm selection and decision logic, offering references for industrial standardization [6]. Weld and stamping defect inspection are mainstream metal manufacturing applications. Traditional ultrasonic testing cannot match line speed, while vision-based on-line weld analysis identifies blowholes and incomplete fusion within milliseconds. Large-scale deployment of robotic inline inspection in Swedish manufacturing summarizes practical experience on system configuration, staff training and production line adaptation, shown in Table 2 [7].

TABLE 2. Application characteristics of machine vision inspection in typical manufacturing scenarios

Manufacturing Scenario	Inspection Targets	Core Vision Technology	Detection Speed	Typical Precision
Automotive Stamping Panels	Scratches, indentations, deformation	YOLO + data augmentation	$\geq$ 30 FPS	> 98%
Automotive Assembly Lines	Missing parts, loose bolts	Object detection + classification	Real-time	> 99%
Welding Quality	Blowholes, incomplete fusion, undercutting	Semantic segmentation + anomaly detection	Real-time	> 95%
Metal Stamping Parts	Cracks, flash, dimensional deviation	YOLOv5 + Stable Diffusion enhancement	Real-time	> 97%
Electronic Components	Solder joint defects, pin offset	High-precision CNN + calibration	$\geq$ 60 FPS	> 99.5%

For printed circuit board (PCB) production, micro-vision systems paired with deep learning detect sub-millimeter solder flaws. Integrated AI and Doosan robotic systems close the loop between vision detection and mechanical handling, enabling active on-site quality disposal rather than simple pass/fail sorting [8].

V. PREDICTIVE MAINTENANCE SYSTEM DRIVEN BY INTEGRATION OF ROBOTICS AND ARTIFICIAL INTELLIGENCE

A. MULTI-SENSOR DATA FUSION FOR EQUIPMENT CONDITION MONITORING & ANOMALY RECOGNITION

Vision-guided robotic systems tightly couple the perceptual judgment capability of machine vision with the precise execution capacity of robots, forming an integrated operation closed loop of "perception-judgment-action" and transforming quality inspection from passive result recording to active intervention. On conventional production lines, vision inspection systems and manipulators operate independently; inspection results have to be manually converted into operation commands, leading to response delays and information loss. By contrast, within the vision-guided robotic framework, analytical outputs from inspection images directly drive robotic grasping, sorting or trimming motions, compressing the full latency from decision to execution down to the millisecond level. For collaborative quality inspection and sorting tasks, robots carry out accurate grabbing and classified placement guided by defect locations and categories provided by vision systems. Stable commercial deployment has been realized in scenarios such as electronic component mounting, automotive component assembly and food appearance inspection. Research on the integration of artificial intelligence and industrial robots for real-time quality control proves that the direct closed loop between vision inspection outputs and robotic decision logic enables production lines to implement differentiated handling based on defect types, rather than merely binary pass/fail discrimination, which greatly refines the fine-grained management of manufacturing processes [9]. In collaborative maintenance scenarios, vision-guided robots automatically identify maintenance triggers including loose bolts, aged seals and displaced cables by comparing real-time equipment images with benchmark reference pictures, and execute preset maintenance sequences once abnormal conditions are detected. This collaborative mechanism integrates state perception and maintenance execution into a single robotic unit, lowering interface complexity caused by multi-system coordination and accelerating maintenance response. The spatial calibration accuracy of vision systems and the precision of robotic kinematic parameters serve as the core technical foundations for reliable collaborative operations, which must be strictly verified during system integration and commissioning. The downtime gap between traditional scheduled maintenance and AI-driven predictive maintenance is intuitively illustrated in Figure 2.

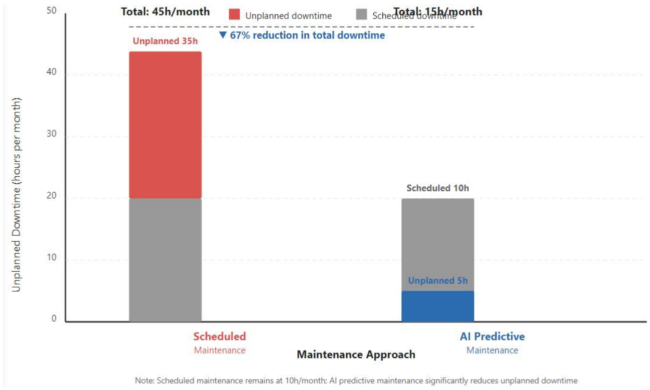

FIGURE 2. Comparison of monthly equipment downtime under scheduled maintenance and AI predictive maintenance

B. CONSTRUCTION AND PERFORMANCE EVALUATION OF MACHINE LEARNING-BASED FAULT PREDICTION MODELS

The objective of fault prediction models is to output reliable fault probability estimation or Remaining Useful Life (RUL) prediction within an adequate lead time prior to equipment failure. There generally exists an inherent trade-off between prediction horizon and prediction accuracy: a longer prediction horizon relies on sparser condition data and brings higher prediction uncertainty; conversely, although condition features become more distinctive close to failure, the available lead time for maintenance preparation shrinks accordingly. In terms of model construction, Long Short-Term Memory (LSTM) networks excel at Remaining Useful Life prediction by effectively modeling long-term dependencies within time-series data and capturing the nonlinear evolution of characteristic indicators during equipment degradation.

Research on predictive maintenance for intelligent manufacturing robots reveals that a hybrid framework integrating deep learning with swarm intelligence optimization algorithms achieves higher prediction accuracy than single models on standard bearing datasets such as PRONOSTIA through automatic searching for optimal network architectures and hyperparameter configurations. The quantitative evaluation formula for fault prediction accuracy adopted in this study is expressed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{1}$$

Where $\hat{y}_i$ denotes the predicted remaining useful life of the i-th sample, y_i stands for its corresponding true value, and n represents the total number of test samples. A smaller RMSE value indicates higher model prediction accuracy. Together with the Mean Absolute Error (MAE), RMSE forms the core evaluation metric system for industrial predictive maintenance models. The generalization performance of a model must be evaluated under working conditions close to actual production scenarios to prevent distorted evaluation results caused by distribution shift between laboratory datasets and field operation data. An integrated evaluation scheme combining cross-validation and independent test sets, along with classification assessment tools such as confusion matrices and Receiver Operating Characteristic Area Under Curve (ROC-AUC), establishes a comprehensive framework for measuring overall model performance. For manufacturing equipment with stringent safety requirements, the false negative rate (missed faults) shall be prioritized over the false positive rate (false alarms). This requirement should be explicitly reflected in the design of loss functions and the selection of decision thresholds during model training.

C. INTEGRATED DEPLOYMENT OF ROBOTIC INSPECTION & AUTONOMOUS MAINTENANCE

Robotic inspection platforms integrate mobile carriers, multi-sensor modules and AI analytics to replace unsafe, error-prone manual patrols. Autonomous mobile robots (AMRs) equipped with high-resolution cameras, thermal imagers and gas sensors detect overheating, pipeline leakage and abnormal vibration in unstructured workshops. Deep reinforcement learning optimizes inspection path planning to balance coverage and energy consumption [10]. Autonomous maintenance robots execute standardized operations (bolt tightening, sensor cleaning, component replacement) upon visual fault identification, relying on precise 3D perception and force-controlled compliant actuation to avoid equipment damage. All patrol logs and fault records feed back into predictive maintenance models to form a closed loop of perception-prediction-execution-iteration, continuously upgrading equipment health management capacity (Table 3).

TABLE 3. Functional configuration comparison of industrial inspection & maintenance robots

Robot Type	Robot Type	Core Sensors	AI Modules	Typical Applications
Fixed Inspection Robot	Guide rail/truss	High-res camera, structured light	Defect detection	Production line on-line inspection
Autonomous Mobile Inspection Robot	AGV/AMR	Thermal imager, gas, vibration sensors	Anomaly identification + path planning	Large-area equipment patrol
Collaborative Maintenance Robot	Fixed base	Torque sensor, vision positioning	Condition perception + actuation	Standardized maintenance tasks
Inspection UAV	Drone	HD camera, LiDAR	3D reconstruction + defect localization	Tall equipment & pipeline inspection

The autonomous execution of maintenance operations relies on precise 3D spatial perception and dexterous manipulation capabilities. Force-based impedance control strategies effectively avoid equipment damage caused by excessive contact force during interactive operations between robots and production facilities. Massive patrol logs and fault records generated by robotic inspection systems are integrated and fed back into predictive maintenance models, forming a full closed-loop autonomous operation workflow of "perception-prediction-execution-feedback", which continuously upgrades the equipment health management performance of manufacturing systems through iterative optimization.

V. SYSTEM ARCHITECTURE AND DEVELOPMENT TRENDS OF COLLABORATIVE INTEGRATION BETWEEN MACHINE VISION AND ROBOTICS

A. COLLABORATIVE OPERATION MECHANISM OF VISION-GUIDED ROBOTS

Vision-guided robotic systems tightly couple the perceptual judgment capability of machine vision with the precise execution capacity of robots, forming an integrated operation closed loop of "perception-judgment-action" and transforming quality inspection from passive result recording to active intervention. On conventional production lines, vision inspection systems and manipulators operate independently; inspection results have to be manually converted into operation commands, leading to response delays and information loss. By contrast, within the vision-guided robotic framework, analytical outputs from inspection images directly drive robotic grasping, sorting or trimming motions, compressing the full latency from decision to execution down to the millisecond level. For collaborative quality inspection and sorting tasks, robots carry out accurate grabbing and classified placement guided by defect locations and categories provided by vision systems. Stable commercial deployment has been realized in scenarios such as electronic component mounting, automotive component assembly and food appearance inspection. Research on the integration of artificial intelligence and industrial robots for real-time quality control proves that the direct closed loop between vision inspection outputs and robotic decision logic enables production lines to implement differentiated handling based on defect types, rather than merely binary pass/fail discrimination, which greatly refines the fine-grained management of manufacturing processes. In collaborative maintenance scenarios, vision-guided robots automatically identify maintenance triggers including loose bolts, aged seals and displaced cables by comparing real-time equipment images with benchmark reference pictures, and execute preset maintenance sequences once abnormal conditions are detected. This collaborative mechanism integrates state perception and maintenance execution into a single robotic unit, lowering interface complexity caused by multi-system coordination and accelerating maintenance response. The spatial calibration accuracy of vision systems and the precision of robotic kinematic parameters serve as the core technical foundations for reliable collaborative operations, which must be strictly verified during system integration and commissioning.

B. HUMAN-ROBOT COLLABORATION & DIGITAL TWIN INTEGRATION UNDER INDUSTRY 4.0

In human-robot collaboration (HRC) scenarios, collaborative robots (cobots) are required to safely and efficiently perform inspection and maintenance tasks while sharing workspace with human operators. To this end, robots need to be equipped with vision-based human posture recognition systems for real-time prediction of human movement intentions and dynamic avoidance of collision risks. Digital twin technology constructs high-fidelity virtual replicas of physical manufacturing systems, delivering a low-cost simulation platform for continuous training of machine vision models and offline optimization of robot control strategies. Workpiece defect images generated within digital twin environments can supplement insufficient real labeled training samples. Collision verification and cycle time optimization of robotic inspection paths can also be efficiently completed in virtual environments, drastically shortening the system configuration cycle for new product introduction. Research on the application of AI in industrial innovation and growth indicates that deep integration of advanced AI technologies with existing manufacturing infrastructure represents a critical pathway for next-generation manufacturing systems to realize self-optimization and flexible scaling [11-13]. Data integration between Manufacturing Execution Systems (MES) and Enterprise Resource Planning (ERP) enables real-time incorporation of vision inspection results and robotic maintenance logs into production scheduling and quality traceability frameworks, achieving end-to-end information connectivity from the workshop floor to management decision-making layers. Standardized industrial communication protocols such as OPC-UA and MQTT form the technical foundation for interoperability among subsystems. Their unified data exchange capability across heterogeneous manufacturing equipment provides protocol-level support for building open, scalable intelligent manufacturing platforms.

C. TECHNICAL LIMITATIONS AND FRONTIER OUTLOOK OF AI-DRIVEN INSPECTION & MAINTENANCE

Despite remarkable progress of AI-enhanced machine vision and robotics technologies in quality inspection and predictive maintenance for automated manufacturing, several technical limitations remain to be addressed for large-scale industrial rollout. The high cost of acquiring labeled data and scarcity of defective samples constitute the primary bottlenecks restricting the performance of deep learning models. This challenge is especially prominent in low-volume, mixed-model production, where frequent product changeovers drive up the cost of continuous model updating and maintenance. In addition, the "black-box" nature of deep learning models renders their decision-making processes uninterpretable from an engineering perspective, creating compliance barriers for certification and deployment in high-safety manufacturing scenarios. From the perspective of cutting-edge technologies, few-shot learning and transfer learning offer new technical routes for rapid migration of visual inspection models across different products and scenarios. Instead of training models from scratch, these approaches enable adaptation to new defect categories with limited samples, substantially lowering application barriers. The federated learning framework allows multiple manufacturers to collaboratively train superior defect detection models without exchanging raw production data, satisfying both the demand for improved model performance and the requirement of commercial data privacy protection. Research on intelligent technologies in mechanical manufacturing and automation states that data-centric intelligent manufacturing can only operate stably and reliably under complex working conditions when deeply integrated with process expertise [14-16]. At the forefront of robotics research, the extension of foundation models toward embodied intelligence endows robots with stronger zero-shot generalization capabilities, which are expected to break the current limitation of insufficient manipulation flexibility of industrial robots in unstructured environments. Advances in quantum computing and neuromorphic chips may overcome existing inference efficiency bottlenecks at the underlying computing architecture, providing powerful computing support for complex real-time visual analysis and decision-making tasks. In the future, deep integration of machine vision, robot control and AI prediction will evolve manufacturing systems from reactive automation to self-adaptive intelligent manufacturing, opening new technological avenues for continuous improvements in industrial production efficiency and quality reliability.

VI. CONCLUSION

Based on the above analysis, the deep integration of AI-enhanced machine vision and robotic technologies is fundamentally reshaping the intelligence level of manufacturing systems in the fields of quality inspection and predictive maintenance for automated production. From the construction of core technical fundamentals, the development of two major application lines covering quality inspection and predictive maintenance, to the evolution of system architectures for collaborative integration of vision and robotics, all technical tiers form an organically interconnected whole that supports and deepens one another step by step. Deep learning algorithms deliver unprecedented defect identification accuracy for vision systems; multi-sensor fusion combined with machine learning models enables precise prediction of equipment health conditions; meanwhile, the collaborative operation mechanism of vision-guided robots translates perceptual capabilities into tangible manufacturing benefits. It is noteworthy that practical gaps still exist between laboratory research and large-scale industrial deployment, including high data labeling costs, poor model interpretability and complicated system integration. Resolving these challenges relies on continuous interdisciplinary collaboration and accumulation of engineering experience. The rapid advancement of cutting-edge technologies such as few-shot learning, federated learning and embodied intelligence opens up new pathways to overcome the aforementioned limitations.

ACKNOWLEDGMENT

This work was supported by the 2024 Jiangxi Higher Education Society Smart Teaching Funding Project "Construction and Application of Industrial Robot Application Technology Course Based on Knowledge Graph under the Background of Digitalization and Intelligence" (Grant No. JX-D-008) and the 2024 Jiangxi Provincial Department of Education Science and Technology Research Funding Project "Research on the System Development of TF Card Detection, Identification and Sorting Based on Machine Vision Deep Learning" (Grant No. GJ2407703).

REFERENCES

[1] Y. Liang, S. Feng, Y. Zhang, et al., “A stable diffusion enhanced YOLOV5 model for metal stamped part defect detection based on improved network structure,” *J. Manuf. Processes*, vol. 111, pp. 21–31, 2024, doi: 10.1016/J.JMAPRO.2023.12.064.

[2] H. Fang, S. Ong, and A. Nee, “Robot path planning optimization for welding complex joints,” *Int. J. Adv. Manuf. Technol.*, vol. 90, nos. 9–12, pp. 3829–3839, 2017, doi: 10.1007/s00170-016-9684-z.

[3] D. Kiefer, F. Grimm, T. Straub, et al., “Enhancing power skiving tool longevity: The synergy of AI and robotics in manufacturing automation,” *Int. J. Mechatronics Manuf. Syst.*, vol. 17, no. 2, pp. 201–224, 2024, doi: 10.1504/IJMMS.2024.143059.

[4] D. Kumar, R. S. Addula, M. Lind, et al., “AI-driven hybrid deep learning and swarm intelligence for predictive maintenance of smart manufacturing robots in Industry 4.0,” *Electronics*, vol. 15, no. 3, p. 715, 2026, doi: 10.3390/ELECTRONICS15030715.

[5] I. Yousif, L. Burns, E. F. Kalach, et al., “Leveraging computer vision towards high-efficiency autonomous industrial facilities,” *J. Intell. Manuf.*, vol. 36, no. 5, pp. 1–26, 2024, doi: 10.1007/S10845-024-02396-1.

[6] M. O. Matamoros, T. G. J. Nava, M. J. J. Escobar, et al., “Artificial intelligence for quality defects in the automotive industry: A systemic review,” *Sensors*, vol. 25, no. 5, p. 1288, 2025, doi: 10.3390/S25051288.

[7] A. Victor, G. Anna, and L. Yvonne, “Lessons from adopting robotic in-line quality inspection in the Swedish manufacturing industry,” *Procedia Comput. Sci.*, vol. 217, pp. 386–394, 2023, doi: 10.1016/J.PROCS.2022.12.234.

[8] A. Chaabani, R. Cherif, and Y. Yaddaden, “Automating quality control: Real-time defect detection and automated decision-making with AI and Doosan robotics,” *Int. J. Intell. Robot. Appl.*, vol. 9, no. 3, pp. 1–15, 2025, doi: 10.1007/S41315-024-00417-Z.

[9] L. Patrício, L. Varela, and Z. Silveira, “Proposal for a sustainable model for integrating robotic process automation and machine learning in failure prediction and operational efficiency in predictive maintenance,” *Appl. Sci.*, vol. 15, no. 2, p. 854, 2025, doi: 10.3390/APP15020854.

[10] Y. D. Patrashko and V. Gurau, “Machine learning-powered vision for robotic inspection in manufacturing: A review,” *Sensors*, vol. 26, no. 3, p. 788, 2026, doi: 10.3390/S26030788.

[11] A. Mahajan, G. Singh, S. Devgan, et al., “Next-generation manufacturing: Leveraging AI for industrial innovation and growth,” *Comput. Ind. Eng.*, vol. 211, p. 111618, 2026, doi: 10.1016/J.CIE.2025.111618.

[12] C. L. Wang, “Research on intelligent technology in mechanical design, manufacturing and automation,” *China Strategic Emerging Industries*, no. 8, pp. 63–65, 2026.

[13] Y. Q. Sun and T. Shen, “Research on application of industrial robot vision systems in electrical automation,” *China Equipment Eng.*, no. 6, pp. 28–30, 2026.

[14] C. B. Li, “Application of machine vision in machinery manufacturing automation,” *Paper and Paper Making Equipment & Materials*, vol. 54, no. 11, pp. 88–90, 2025.

[15] T. Yue, “Artificial intelligence in metallurgical automation production—Comment on complex metallurgical process intelligent control,” *China Nonferrous Metallurgy*, vol. 53, no. 4, pp. 161–162, 2024.

[16] “Robotics; findings from Singapore National University in the area of robotics reported (Robot path planning optimization for welding complex joints),” *J. Robotics & Machine Learning*, p. 34, 2017.