Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.38488

Large Language Models for Industrial Robots: Applications, Challenges, and Future Directions

Qianen Zhou

Faculty of Science and Engineering, University of Nottingham Ningbo China, Ningbo, China
sayqz3@nottingham.edu.cn

Abstract. With the accelerating intelligent transformation of global manufacturing, industrial robots have become critical equipment in modern smart factories. However, the growing prevalence of multi-variety, small-batch production has made the high programming threshold, tedious reconfiguration, and limited environmental adaptability of traditional industrial robots increasingly prominent. Over the past few years, large language models (LLMs) have proven to have strong semantic comprehension, knowledge reasoning, and task planning abilities, a new avenue in the intelligent upgrading of industrial robots. The systematic review of the literature on the topic of the industrial robot driven by LLM is presented in this paper, including the mainstream technical directions, and the representative studies in five areas of application: industrial inspection, assembly and collaboration, maintenance and fault recovery, manufacturing operation planning, and robot navigation. Based on this, the paper examines the key challenges to be addressed at present, including model output reliability and safety, the simulation-to-reality gap and real-time constraints, and data scarcity and distribution shift. It also discusses future research opportunities, including knowledge enrichment and reliable reasoning, lightweight deployment and edge computing, and life-long learning and domain adaptation. The purpose of this paper is to serve as a reference in future studies and implementation of LLMs in industrial robots.

Keywords: Large Language Models, Industrial Robots, Task Planning, Reliability and Safety, Domain Adaptation

1. Introduction

With the continuous deployment of automation technologies, global manufacturing is accelerating its intelligent transformation. In the past decades, industrial robots have found wide usage in industrial settings and have become an imperative component in contemporary intelligent factories, as they perform tasks like handling and assembling [1]. But the world of factory production has been transformed. The consumer needs are more diversified, product iteration is faster and multi-variety, small-lot production is the new standard of operation of many manufacturing businesses. However, the most current robots in use in industry are still based on traditional programming paradigms. These robotic systems tend to be environmentally inflexible, are unable to adapt to dynamically changing working environments and cannot learn new tasks on their own [2]. This

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.38488

implies that manual teaching and reconfiguring of robots through manual teaching and trajectory reconfiguring is needed whenever workpieces or production processes are adjusted. The entire process is human resource intensive and is likely to be affected by human introduced operational errors [3], and production systems are hard pressed to match the fast changing production orders.

To overcome the limitations of high-threshold programming and manual reconfiguration when changing production to a new product with conventional industrial robots, scientists have started to consider the implementation of large language models in robot control systems. Over the past few years, as natural language processing advances continuously, large language models have proven to be capable of strong performance in semantic understanding, knowledge reasoning, and task planning, and have been generalized outside of text-only tasks, to embodied intelligence applications like robotics. It is against this context that the advent of large language models has led to revolutionary advances in robotics by giving them the ability to engage in interactive communication, possess semantic understanding, and reason logically like humans [4]. With these capabilities, many studies have used the high-level environmental perception and knowledge-driven planning capabilities of large language models to address the intrinsic constraints of conventional robots in planning algorithms and low-level control [5]. As an illustration, SayCan uses large language models along with robotic affordances to break down instructions written in natural language into executable low-level sequences of actions [6]; VoxPoser uses the code-writing ability of large language models to generate 3D value maps by interacting with vision-language models [7], and PaLM-E incorporates vision, language, and robotic control into a single model [8], allowing robots to solve complex manipulation tasks using multimodal inputs. Based on this, giant language models can be used to help reduce the barriers to robot operation and make the workflows of equipment re-commissioning and task deployment simpler in industrial environments.

Despite the fact that current research has significantly enhanced the intelligent interaction and task-planning ability of robots and removed the pain points of bulky deployment and high professional standards of conventional industrial robots, a set of technical and practical difficulties remains on the way to full implementation of such technologies in real industrial settings [9]. The available review literature shows that large language models can produce erroneous, logically incoherent or even harmful text when generating texts [10]. Given that industrial production is characterized by high demands both in terms of operational accuracy and safety, these erroneous outputs can cause serious risks. Moreover, the majority of current validation experiments remain limited to simulation settings, which poses challenges to the transfer of the technology to real production settings and imposes real-time constraints on practical deployment. Moreover, industrial cases are highly specialized, and quality annotated data to train or fine-tune large language models is often limited, while industrial expertise is usually distributed across technical manuals, maintenance records, and experience. The fast pace of industrial process development also demands models to keep their knowledge up to date in continuous mode, which makes it challenging to stably adapt general-purpose large language models to real-life industrial conditions. It is on this backdrop that the current paper critically examines the literature on industrial robots operated by large language models, discusses the current directions in the mainstream technical literature, and identifies current bottlenecks and future research directions, as well as provides references to further studies in the domain.

DOI: 10.54254/2755-2721/2026.38488

2. LLM-based industrial robots

2.1. Preliminaries of LLM

Large language models (LLMs) are large-scale pre-trained models built upon the Transformer architecture. They are comprised of several stacked Transformer decoder layers which are usually filled with multi-head self attention, feed forward networks, residual networks and layer normalization. By intuition, the model is able to build long term dependencies in a sequence and get semantics of context and intricate correlations. The Transformer architecture has a higher parallel computation capability and scalability than recurrent neural networks and convolutional neural networks, and the number of model parameters can be scaled to billions or even hundreds of billions [11].

The paradigm of pre-training and fine-tuning/alignment is typically two-stage with LLM. The input in the pre-training phase is large-scale unlabeled text and general language representations and world knowledge are trained using autoregressive language modeling tasks. During fine-tuning and alignment phase, model outputs are improved, with instruction tuning and human reinforcement learning, to achieve closer alignment of model outputs with human instructions and task requirements. Through this structural design and training paradigm, LLMs have high levels of semantic understanding and knowledge reasoning, code generation, and task planning [12].

2.2. Industrial inspection robots

2.2.1. Inspection task planning

In inspecting scenarios within industries, traditional inspection robots are usually limited to specific paths and straightforward operations and do not comprehend and decide in complicated scenarios. Z. Wang et al. have created the LLM-MTMP framework that integrates LLMs, multi-agent systems, and retrieval-augmented generation to encode and decode natural language into sequence of operations that can be understood by robots and assist in autonomous planning of tasks and movements in power inspection tasks [13]. The framework uses RAG technology to access information in a knowledge base of a power inspection expert, generates executable task-action sequences in the form of Python scripts and detects execution failures through fault reasoning and feedback, thus dramatically reducing the dependence on human operators of an inspection system [13]. Moreover, a remote human-robot communication system constructed on VLMs allows the operators to remotely execute inspection tasks commands using natural language. The VLM collaboratively analyses the instructions, image data obtained by the visual sensors of the robot, and environmental perception data provided by the LiDAR, thus eliminating ambiguity in instructions, and generating accurate control commands [14].

2.2.2. Zero-shot anomaly detection

In terms of anomaly detection during inspection, VLMs enable robots to perform environmental context reasoning and can give interpretable explanations of anomalies and work with greater intelligence and flexibility [15]. Aydogmus and Boztas combined a quadruped robot, a lightweight edge-side VLM, and a cloud-side VLM, which allowed the detection of anomalies in a zero-shot in both indoor and outdoor conditions. They include an environment conscious class mapping, anomaly deduplication, consistency checking, and fallback reasoning and experiment with several

DOI: 10.54254/2755-2721/2026.38488

state of the art VLMs, such as Gemini-2.5-Pro, GPT-5, and Claude Sonnet 4.5 [15]. In contrast to classical visual anomaly detection techniques, LLMs are capable of processing multimodal sensor data in the form of images, videos, and textual metadata and can give semantically informative insights about anomalous events [16].

2.3. Industrial assembly robots

2.3.1. Human-robot collaborative assembly

The use of LLMs in human-robot collaboration is a shift towards the cognitive manufacturing paradigm of Industry 5.0. L. Wang et al. conducted a systematic literature review of 1,278 studies and categorized uses of LLM in collaborative robotics into three levels of thinking: environmental perception and meaning, task planning and decision-making, and execution and interaction [17]. Their results indicate that LLMs can read natural language instructions and generate assembly code, divide tasks and assign dialogue, and update tasks dynamically, as well as promote human-robot interaction significantly in perceiving the environment, long-horizon task planning, and embodied interaction [17].

2.3.2. Autonomous decision-making and execution

In semi-structured industrial scenarios, where assembling tasks are performed, modern systems are increasingly depending on LLMs and VLMs to substitute code-based inputs with more human-friendly knowledge, such as demonstration videos and technical documents [18]. An inspection system based on VLM, e.g., can learn to detect defective welds based on training videos. Unlike the traditional machine vision that relies on fragile hand-created templates, these models demonstrate generalized understanding and retain accuracy with the variations in lighting conditions, orientation, or the surface. LLMs also have one more feature: the ability to semantically interpret unstructured process documents. A robot with the capability to read maintenance manuals can differentiate between safety-critical defect categories and cosmetic ones, and update its operating procedures with the availability of updated manuals without necessarily having to be reprogrammed. When an autonomous decision is overridden by a human operator, the correction is added to the model in the form of a labelled training case, slowly bringing the workforce tacit knowledge to bear on the decisions made by the system [18].

2.4. Industrial maintenance and fault recovery

2.4.1. Fault diagnosis and prediction

In the field of industrial robot fault diagnosis, LLM applications are evolving from single-modal to multimodal fusion. C. Chen et al. proposed the STF-MLLM framework, a signal-text fusion multimodal large language model that simultaneously captures time-domain transient features and frequency-domain periodic patterns through a parallel time-frequency embedding module, and adaptively combines these features using a lightweight gated fusion mechanism, achieving 96.27% fault diagnosis accuracy in a few-shot scenario with only 10 samples [19]. At the system level of operations and maintenance, J. Chen et al. introduced the DPDC-ILKM framework, which builds a highly dependable industrial large knowledge model combining diagnosis, prediction, decision-making, and control functions to enable intelligent maintenance of industrial swarm robots, and

DOI: 10.54254/2755-2721/2026.38488

further develops a multi-level evolutionary learning mechanism for adaptive learning in complex industrial environments [20].

2.4.2. Fault recovery

In terms of robot fault recovery, VLMs demonstrate remarkable commonsense generalization and reasoning capabilities. Current robot autonomy struggles to operate outside its pre-designed operational design domain, while the real world is full of uncertainties that may lead to failures. H. Chen et al. investigated how to enhance the spatial reasoning capabilities of VLMs by optimizing visual and textual prompts, enabling them to serve as black-box controllers for motion-level position correction and task-level unknown fault recovery [21]. Experiments showed that the optimized prompts improved the accuracy of motion-level position error correction by 65.78% over unoptimized VLM prompts, and increased the success rates of task-level fault detection, problem analysis, and recovery plan generation by 5.8%, 5.8%, and 7.5%, respectively [21].

2.5. Industrial manufacturing operation planning

2.5.1. Shopfloor operation planning

In the area of manufacturing operation planning, the combination of agentic AI and collaborative robots is reshaping conventional manufacturing paradigms. Traditional industrial robots remain bound by rigid programming, so that even a minor change in part geometry necessitates a full rewrite of the software. Contemporary systems employ LLMs and VLMs to replace code-based inputs with human-oriented knowledge, including demonstration videos and technical documents [18]. A human-centered agentic framework driven by LLMs has been proposed to generate and adapt manufacturing programs and their execution. Gerdes et al. carried out a case study on laser-scanned object reconstruction, showing that agent-based interaction can effectively connect high-level user intent with precise machine operations [22].

2.5.2. Welding and casting inspection

In welding inspection, robotic systems equipped with VLMs are capable of identifying surface discontinuities such as porosity, undercut, overlap, and lack of fusion, categorizing them under relevant acceptance standards such as AWS D1.1 or ASME Section IX, and assigning severity ratings [18]. When a defect exceeds the permissible threshold, the system automatically issues a repair work order in the factory MES or ERP system, specifying the defect type, location coordinates, part ID, and recommended repair procedure. In casting inspection, agentic AI operates together with CT scanning or ultrasonic testing equipment to integrate multimodal sensor data, including surface images, volumetric scans, and dimensional measurements, into a unified defect model. An LLM reasoning layer subsequently applies acceptance criteria from design specifications to mark components for review or rework [18].

2.6. Robot navigation

In robot navigation, LLMs are promoting a transition from conventional geometric planning toward semantic navigation. Pan et al. carried out a systematic review that groups LLM-driven robot navigation research into five areas: perception, planning, control, interaction, and coordination, and examines the principal datasets, simulation platforms, and evaluation metrics [23]. This is according

DOI: 10.54254/2755-2721/2026.38488

to their analysis that shows excellent generalization of LLMs in environmental perception and decision reasoning, instruction comprehension, and human-robot interaction [23]. Moreover, Li et al. reviewed the literature on the use of LLM in multi-robot systems in a systematic manner, encompassing several levels such as the high-level level of task allocation, middle level of motion planning, low-level of action generation, and human intervention, reflecting transformative potential in application in areas such as home robots, construction, the control of formations, and in pursuits like target tracking [24].

3. Discussion

3.1. Challenges

3.1.1. Model output reliability and safety

Large language models are inherently probabilistic and their text generation can result in inaccurate, logically inconsistent or even harmful content. Since industrial production requires a high degree of accuracy and safety, such errors may lead to equipment damage, disruption to production or even possible harm to the staff. Unlike traditional industrial control systems based on deterministic algorithms and formal verification, pipelines of controls motivated by LLMs lack interpretable reasoning traces and formal safety guarantees. This is acute especially in situations that are dangerous like chemical manipulations, welding and assembly where one misplaced order could lead to a major disaster. Recent studies have sought to limit the model output with known facts by applying knowledge augmentation, retrieval-augmented generation and cache-augmented generation, therefore, minimizing hallucination and enhancing traceability. But the problem of obtaining a safety assurance at an industrial level and at the same time retain flexibility is a basic challenge. In addition, stochasticity of LLMs implies that the same input can lead to varying outputs, which leads to some inherent conflict with repeatability and consistency found in the industrial world.

3.1.2. Simulation to reality gap and real time constraints

Most of the current validation experiments of the industrial robots driven by LLM have been done in simulated environments. Simulation is a limited method of recreating the complexity of a real production situation (simulation can give controlled conditions and cheaply repeat but it cannot fully simulate the changes in lighting, mechanical wear, sensor noise, and other unpredictable human behavior). Models which perform well during simulation can fail or even fail to perform well when implemented in real-life factories. Bridging this simulation-to-reality gap will require more realistic simulation platforms, and powerful domain-adaptation methods and comprehensive on-site validation. In the meantime, industrial production processes tend to require real-time or near-real-time response, but large language models usually do not scale to small computational resources, and have non-negligible latency of inference. The LLM can be very flexible and adaptable in high-level reasoning, but it is slow; low-level embedded execution is fast and has no semantic insight into complex problems. Hybrid methods seek to compromise this trade-off, but the problem of having both high-level reasoning and low-level real-time control in a single system is still to be solved. Moreover, the deployment of large models on edge devices in factories is limited to the small memory, power and computing capacity, which again makes it more challenging to implement in practice.

DOI: 10.54254/2755-2721/2026.38488

3.1.3. Data scarcity and distribution shift domain

Industrial problems are extremely specialized, and high-quality annotated data with which to train or fine-tune large language models can be hard to find. In contrast with general-purpose domains where large text corpora can be easily found, industrial knowledge can be dispersed in technical manuals, maintenance logs, as well as in expert experience, most of which is unstructured or tacit. Training general-purpose LLMs on particular industry, then, can be costly in terms of domain-specific data and fine-tuning and might not be at all possible in small or medium-sized companies. Moreover, industrial processes are fast changing their models, thus requiring models to keep updating their knowledge, demanding extra data collection and model maintenance. The question of how to stabilize the adaptation of LLMs to industrial settings in the case of limited domain data and constantly changing domain distributions is one of the central issues.

3.2. Future prospects

3.2.1. Knowledge enrichment and reliable reasoning

To tackle the issue of reliability and safety of model outputs, upcoming studies should be aimed at grounding the model-generating process of LLMs on industrial knowledge that is verified. Retrieval-augmented generation, knowledge graphs, and cache-augmented generation can restrict model outputs to trusted knowledge limits, as well as curb hallucination and enhance traceability. Formal verification and uncertainty quantification allow systems to raise warning signals or reject low-confidence output, thus meeting the safety standards of an industrial setting, but still retaining flexibility and offsetting the uncertainties of formally-guaranteed safety in control pipes powered by LLM. Moreover, to address the conflict between the stochastic nature of LLM outputs and the needs of the industry to generate repeating and consistent results, the research in the future should focus on deterministic reasoning, consistency of outputs, and reproducible generation processes so that the models could generate consistent and predictable outputs when presented with similar inputs. Combining knowledge augmentation, formal verification, and consistency constraints, one can enhance the reliability, safety, and repeatability of the LLM outputs in industrial settings without restricting their flexibility.

3.2.2. Lightweight deployment and edge computing

Future studies should facilitate lightweight and efficient deployment of large language models in order to overcome the simulation-to-reality gap and real-time constraints. The use of techniques like parameter-efficient fine-tuning, model compression, quantization, and distillation can significantly decrease the computational and memory footprint of LLMs, enabling them to be more readily deployed to edges in factories and facilitating the constraints of limited memory, power, and computing capacity of edge devices. With edge computing systems, new industrial robots can potentially do high-level reasoning at the edge, eliminating the need to communicate with the cloud to do so, thus enhancing response time, data confidentiality, and system reliability and minimizing the effect of inference latency on production in real time. In the meantime, more realistic simulation platforms, and powerful domain adaptation methods will be used to reduce the simulation-to-reality gap, so that models can continue to perform stably after being transferred to physical factories. Edge deployment may also minimize reliance on network bandwidth enabling robots to retain basic bandwidth features during unstable or disrupted network conditions, which is especially crucial in

DOI: 10.54254/2755-2721/2026.38488

continuous production in industrial environments. The high-level reasoning and low-level real-time control trade-off can be more conveniently balanced in a single framework by the synergy of lightweight models, edge computing, and domain adaptation techniques.

3.2.3. Life-long learning and domain adaptation

To overcome the problems of domain data scarcity and distribution shift, further research needs to facilitate the process of continual learning and domain adaptation of robots, which would allow them to constantly evolve their knowledge and abilities basing on the new data and feedback without losing the previously developed ones. Methods like constant learning, meta-learning as well as memory-enhanced architectures should have significant roles in making robots be able to adjust to the changing production conditions across time, thus lessening their reliance on huge quantities of annotated data. Besides this, future studies ought to explore how stability and generalization of LLMs can be enhanced in an industrial context under constantly changing distributions of domains by updating online, few-shot learning, and cross-domain transfer, to address the knowledge updating requirements of the continuous cycling of industrial processes. With the creation of domain adaptation mechanisms that can be constantly trained, one can reduce the adaptation pressure due to the limited industrial data and high rate of domain exploration, supporting the long-term application of LLMs to industrial robots.

4. Conclusion

Major areas of application of LLM-driven industrial robots are covered in this paper, including industrial inspection, industrial assembly and collaboration, industrial maintenance and fault recovery, manufacturing operation planning, and robot navigation. The analyzed literature indicates that semantic understanding, knowledge reasoning, and task planning are significant benefits of LLMs. They serve well in reducing the barriers to robot operation, streamlining the re-integration of equipment and task placement processes, and initiating the shift of industrial robots out of traditional programming paradigms and into cognitive manufacturing paradigms.

However, numerous challenges continue to occur when it comes to the complete implementation of LLMs in a real industrial setting. To begin with, model output reliability and safety are hard to balance with the high demands of industrial production in precision and consistency. Second, the simulation-to-reality gap and real-time constraints restrict the possibility of migrating these technologies to physical factories. Third, data scarcity and distribution shift, together with scattered industrial knowledge, make it difficult for general-purpose LLMs to adapt stably to concrete industrial cases.

To overcome these difficulties, this paper discusses some of the future lines of research, such as knowledge enrichment and reliable reasoning, lightweight deployment and edge computing, and life-long learning and domain adaptation. With the combination of LLMs and validated industrial knowledge, edge computing architectures, and life-long learning algorithms, the reliability of the latter can be enhanced, their real-time performance and adaptability can be improved, and their flexibility should be retained. It is also advisable in future studies to enhance real industrial validation, by facilitating the shift of LLM-controlled industrial robots from simulation validation into large-scale application.

DOI: 10.54254/2755-2721/2026.38488

References

[1] Bilancia, P., Ferrara, F., Pisu, P., & Sotelo, M. A. (2023). An overview of industrial robots control and programming approaches. *Applied Sciences, 13*(4), 2582.

[2] Arents, J., & Greitans, M. (2022). Smart industrial robot control trends, challenges and opportunities within manufacturing. *Applied Sciences, 12*(2), 937.

[3] Pan, Z., Polden, J., Larkin, N., Van Duin, S., & Norrish, J. (2012). Recent progress on programming methods for industrial robots. *Robotics and Computer-Integrated Manufacturing, 28*(2), 87–94.

[4] Kim, Y., Kim, D., Choi, S., & Kim, J. (2024). A survey on integration of large language models with intelligent robots. *Intelligent Service Robotics, 17*(5), 1091–1107.

[5] Jeong, H., Lee, S., Park, J., & Kim, H. (2024). A survey of robot intelligence with large language models. *Applied Sciences, 14*(19), 8868.

[6] Brohan, A., Chebotar, Y., Finn, C., Hausman, K., Herzog, A., Ho, D., Ibarz, J., Irpan, A., Jang, E., Julian, R., Kalashnikov, D., Levine, S., Lu, Y., Parada, C., Rao, K., Sermanet, P., Toshev, A., Vanhoucke, V., Xia, F., … Zeng, A. (2023). Do as I can, not as I say: Grounding language in robotic affordances. In *Proceedings of the 6th Conference on Robot Learning* (pp. 287–318). PMLR.

[7] Huang, W., Xia, F., Shah, D., Driess, D., Liu, A., Lu, Y., Florence, P., Driess, D., Zeng, A., & Ichnowski, J. (2023). VoxPoser: Composable 3D value maps for robotic manipulation with language models. *arXiv preprint*.

[8] Driess, D., Xia, F., Sajjadi, M. S. M., Lynch, C., Chowdhery, A., Ichter, B., Wahid, A., Tompson, J., Vuong, Q., Yu, T., Huang, W., Chebotar, Y., Sermanet, P., Duckworth, D., Levine, S., Vanhoucke, V., Hausman, K., Toussaint, M., Greff, K., … Florence, P. (2023). PaLM-E: An embodied multimodal language model. *arXiv preprint*.

[9] Helberger, N., & Diakopoulos, N. (2023). ChatGPT and the AI Act. *Internet Policy Review, 12*(1).

[10] Zeng, F., Dong, S., Wang, Y., & Li, X. (2023). Large language models for robotics: A survey. *arXiv preprint*.

[11] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems (Vol. 30,* pp. 5998–6008). Curran Associates.

[12] Firoozi, R., Tucker, J., Tian, S., Majumdar, A., Sun, J., Liu, W., Zhu, Y., Song, S., Kapoor, A., Hausman, K., Ichter, B., & Schwager, M. (2025). Foundation models in robotics: Applications, challenges, and the future. *The International Journal of Robotics Research, 44*(5), 701–739.

[13] Wang, Z., Liu, Y., Zhang, H., & Li, J. (2026). LLM-MTMP: A large language model-based multi-agent task and motion planning framework for power inspection robots. *Journal of Industrial Information Integration, 49*, 101014.

[14] Lan, X., Zhang, L., Xu, Z., & Wang, Y. (2025). Remote human-robot interaction in industrial inspection system based on vision-language models. In *2025 40th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE.

[15] Aydogmus, O., & Boztas, G. (2026). Zero-shot visual anomaly detection on a quadruped robot using state of the art visual language models. *IEEE Access, 14*, 17842–17852.

[16] Cardoso, M., Ferreira, J., Silva, A., & Costa, P. (2026). Automatic fault detection and diagnosis in ROS-based robotic systems using generative AI: A systematic literature review. *Applied Sciences, 16*(11), 5545.

[17] Wang, L., Gao, R. X., Váncza, J., Krüger, J., Wang, X. V., Makris, S., & Chryssolouris, G. (2026). Large language models in human-robot collaboration: A systematic review, trends, and challenges. *Journal of Manufacturing Systems, 85*, 249–268.

[18] Panigrahi, D. (2026). Smarter together: How agentic AI and cobots are transforming industrial operations. *Power Systems Design*.

[19] Chen, C., Zhao, K., Chen, Z., Wang, T., & Shao, H. (2027). Signal–text fusion multimodal large language model for few-shot fault diagnosis of industrial robots. *Robotics and Computer-Integrated Manufacturing, 103*, 103366.

[20] Chen, J., Xu, Y., Tang, J., Xu, X., & Yan, R. (2026). DPDC-ILKM: A multi-agent integrated large knowledge model for intelligent maintenance of industrial swarm robotics. *Journal of Industrial Information Integration, 50*, 101044.

[21] Chen, H., Yao, Y., Liu, R., Liu, C., & Ichnowski, J. (2025). Robot failure recovery using vision-language models with optimized prompts. In *2025 American Control Conference (ACC)* (pp. 1983–1988). IEEE.

[22] Gerdes, H., Reisch, R. T., & Burschka, D. (2025). Shopfloor-integrated manufacturing operation planning through agents. In *2025 IEEE 30th International Conference on Emerging Technologies and Factory Automation (ETFA)* (pp. 1–8). IEEE.

DOI: 10.54254/2755-2721/2026.38488

[23] Pan, H., Huang, S., Yang, J., & Wang, Y. (2026). Robot navigation via foundation language models: A review. *ACM Computing Surveys, 58*(11), 1–38.
[24] Li, P., Wang, Y., Zhang, H., & Liu, J. (2026). Large language models for multi-robot systems: A survey. *Autonomous Robots, 50*(3), 1–25.