World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(Research Article)

Check for updates

Reinforcement learning-based control improves efficiency and precision in smart manufacturing robotics

Emmanuel Ayodeji Osoko [1,*] and Shukurat Opeyemi Rahmon [2]

[1] *Department of Electrical Engineering and Computer Science, Ohio University, Ohio, United State.*
[2] *Department of Mathematics, University of Lagos, Akoka, Lagos, Nigeria.*

World Journal of Advanced Engineering Technology and Sciences, 2023, 09(01), 421-431

Publication history: Received on 27 April 2023; revised on 26 June 2023; accepted on 29 June 2023

Article DOI: https://doi.org/10.30574/wjaets.2023.9.1.0144

Abstract

Smart manufacturing relies increasingly on automation and robotics to enhance productivity, precision, and operational flexibility. This paper presents an original primary research study detailing the development, implementation, and evaluation of a novel reinforcement learning (RL) model tailored for robotic control in dynamic manufacturing environments. The proposed model improves adaptability, efficiency, and task execution accuracy by leveraging real-time data-driven decision-making. Key contributions include an optimized reward function, adaptive policy updates, and seamless integration with industrial IoT (IIoT) frameworks to enable intelligent and autonomous manufacturing workflows.

The study employs a rigorous experimental framework encompassing extensive training, validation, and comparative benchmarking against traditional control methods such as PID and MPC controllers. Results demonstrate a 30.5% reduction in task completion time, a 41.2% decrease in error rates, and a 20.5% improvement in energy efficiency compared to conventional methods. Furthermore, the RL model enhances scalability by maintaining high performance across varying production scales and improves failure recovery time by reducing downtime by up to 45%. The proposed system also exhibits notable cost savings, reducing operational expenses by 17.6% while extending component longevity by minimizing wear and tear.

These findings validate RL as a transformative solution for industrial automation, offering superior efficiency, robustness, and cost-effectiveness compared to existing control mechanisms. The research provides a scalable and adaptable framework for manufacturers seeking to optimize robotic performance, enhance reliability, and achieve long-term sustainability in Industry 4.0-driven production environments.

Keywords: Reinforcement Learning; Smart Manufacturing; Robotics; Industrial IoT; Adaptive Control

1. Introduction

The increasing complexity and variability in modern manufacturing environments demand intelligent automation solutions that go beyond traditional control methods. Conventional robotic control systems, such as PID controllers and model predictive control (MPC), often struggle to adapt to dynamic changes in production processes, leading to inefficiencies and errors (Zhu et al., 2021). As industrial automation transitions toward Industry 4.0, incorporating artificial intelligence (AI)-driven approaches such as reinforcement learning (RL) has become crucial for enhancing the flexibility and precision of robotic systems (Lee et al., 2020).

* Corresponding author: Emmanuel Ayodeji Osoko

Reinforcement learning has emerged as a transformative approach in robotics, enabling autonomous systems to optimize decision-making based on continuous interaction with their environments. Unlike conventional control strategies that rely on predefined rules and heuristics, RL-based systems leverage real-time data to iteratively refine their actions, improving both short-term efficiency and long-term adaptability (Wang et al., 2023). This adaptability is particularly vital in smart manufacturing, where variations in materials, workflows, and production speeds necessitate continuous optimization (Gao et al., 2021). Additionally, the ability of RL algorithms to autonomously adapt to unstructured environments makes them an essential tool for minimizing human intervention in high-precision industrial tasks.

Recent studies have demonstrated that RL models can enhance various aspects of robotic performance, including task execution speed, error minimization, and energy efficiency. However, challenges remain in terms of optimizing reward functions, reducing computational overhead, and ensuring seamless integration with existing industrial IoT (IIoT) frameworks (Chen et al., 2022). Addressing these challenges requires a robust RL architecture that balances real-time decision-making with long-term performance optimization (Hao & Zhang, 2021). Furthermore, while significant progress has been made in the theoretical application of RL in industrial automation, there is still a gap in its real-world deployment, requiring practical validation through extensive experimentation and rigorous benchmarking.

This study presents the development and validation of a novel RL-based robotic control system designed to enhance efficiency, precision, and scalability in smart manufacturing. Unlike previous implementations, this RL model incorporates multi-objective optimization techniques that balance performance, resource utilization, and failure resilience. Additionally, the system is designed to integrate seamlessly with IIoT architectures, allowing real-time data acquisition and adaptive policy updates to improve decision-making processes. Through extensive experimentation, this research demonstrates how RL-driven control significantly reduces task completion time, enhances operational efficiency, and improves failure recovery mechanisms.

Beyond efficiency improvements, the study explores the financial implications of RL adoption in manufacturing, illustrating its potential to lower operational costs through reduced downtime and energy consumption. Furthermore, the RL framework introduced in this work is inherently scalable, capable of adapting to various production scales without compromising performance. By offering an innovative, real-world-tested solution to persistent challenges in industrial robotics, this research contributes to the ongoing transformation of smart manufacturing into a more adaptive, resilient, and cost-effective ecosystem. The insights provided herein hold substantial implications for industries aiming to implement AI-driven automation to maximize productivity and sustainability.

2. Recent advances towards more dedicated rl-studies

The application of machine learning (ML) and artificial intelligence (AI) in robotics has evolved significantly, with reinforcement learning (RL) emerging as a key approach for enhancing robotic autonomy and adaptability in industrial settings. Traditional robotic control methods, such as Proportional-Integral-Derivative (PID) controllers and Model Predictive Control (MPC), have been extensively employed for decades. However, these methods struggle with adaptability in dynamic environments, requiring frequent manual tuning and parameter adjustments (Kober et al., 2013; Nguyen-Tuong & Peters, 2011). As a result, researchers have increasingly turned to RL-based approaches to address these limitations.

RL has demonstrated significant potential in robotic control by enabling autonomous agents to learn optimal policies through trial and error. Deep reinforcement learning (DRL), in particular, has been instrumental in enhancing the decision-making capabilities of robots, leading to improved performance in navigation, grasping, and manipulation tasks (Levine et al., 2016; Lillicrap et al., 2016). Recent works have explored the integration of DRL with industrial robotics to improve real-time adaptability and efficiency in manufacturing workflows (Gu et al., 2017). However, challenges such as sample inefficiency and high computational costs remain key barriers to widespread adoption (Henderson et al., 2018).

The rise of Industry 4.0 has accelerated research into intelligent automation, where RL plays a critical role in optimizing robotic operations. Studies have shown that RL can enhance production efficiency by reducing cycle times and improving precision in material handling and assembly tasks (Kaelbling et al., 2019). Additionally, RL-based control frameworks have been implemented to minimize resource consumption, including energy and raw materials, thereby promoting sustainable manufacturing practices (Zhang et al., 2020). However, most existing RL-based approaches have yet to be fully validated under real-world industrial conditions, necessitating further empirical evaluations.

Despite its advantages, RL-driven robotic systems face several challenges. One of the primary concerns is the stability and convergence of RL algorithms in high-dimensional action spaces, which is crucial for real-time applications in manufacturing (Duan et al., 2016). Moreover, traditional RL models often require extensive training on simulated environments before deployment in real-world scenarios, leading to issues with transferability and generalization (Tobin et al., 2017). Recent advancements, such as domain adaptation techniques and meta-learning strategies, have attempted to mitigate these challenges, but further research is required to enhance robustness and scalability (Finn et al., 2017).

The seamless integration of RL-based robotic control with Industrial IoT (IIoT) frameworks is also essential for enabling adaptive and data-driven decision-making in manufacturing environments. Studies have demonstrated that IIoT-enhanced RL models can leverage real-time sensor data to dynamically adjust robotic policies, improving operational efficiency and fault tolerance (García & Fernández, 2015). However, challenges such as latency, data security, and interoperability with legacy systems must be addressed to ensure smooth deployment in industrial settings (Lu et al., 2021).

While existing research has made substantial progress in developing RL-based robotic control systems, critical gaps remain in achieving real-time adaptability, scalability, and seamless IIoT integration. Current models often prioritize either precision or computational efficiency but fail to provide a balanced approach that optimizes both (Silver et al., 2018). This study aims to bridge this gap by proposing a novel RL framework that leverages adaptive reward functions, domain adaptation techniques, and IIoT connectivity to enhance robustness, energy efficiency, and failure recovery in smart manufacturing. By benchmarking against traditional control methods, this research offers empirical validation of RL's superiority in dynamic industrial environments.

3. Research methodology

3.1. Development of the Reinforcement Learning Model

The RL model was designed using a policy-based approach, incorporating Proximal Policy Optimization (PPO) as the primary learning algorithm. PPO was selected due to its balance between sample efficiency and robustness, making it suitable for real-world robotic applications (Schulman et al., 2017). The key components of the model include:

- **State Representation:** Sensory data from IIoT-enabled sensors was used to construct meaningful state vectors representing the manufacturing environment.
- **Action Space:** Defined as a set of control commands governing robotic arm movements, including positioning, velocity modulation, and force application.
- **Reward Function:** A multi-objective function was formulated to prioritize task accuracy, energy efficiency, and adaptability to dynamic conditions.

3.2. Model Training and Optimization

3.2.1. Training Procedure

To ensure robustness, the RL model was trained in both simulated and real-world settings. Training was conducted using a curriculum learning strategy, gradually increasing task complexity to improve learning efficiency (Narvekar et al., 2020). The training process involved:

- **Domain Randomization:** Environmental variables, such as lighting and workspace clutter, were randomized to enhance model generalization (Tobin et al., 2017).
- **Large-Scale Training Dataset:** The model was trained using a dataset of over 50,000 observations collected from various industrial environments to ensure robustness and scalability (Rusu et al., 2016).
- **Transfer Learning:** Pre-trained policies were refined with real-world data to expedite learning and adaptation.
- **Hyperparameter Tuning:** Optimization of learning rates, exploration strategies, and reward scaling was conducted to maximize policy performance.

3.2.2. Deployment and Real-Time Execution

Once trained, the RL model was deployed on an industrial robotic system, integrating with IIoT-enabled edge computing frameworks to facilitate real-time decision-making. Performance monitoring and adaptive tuning mechanisms were implemented to allow continuous refinement of the model in response to evolving manufacturing conditions.

3.3. Experimental Validation

The effectiveness of the RL-based control system was assessed through rigorous experimentation. Benchmark comparisons were conducted against conventional controllers, including PID and MPC, using the following metrics such as Task completion Time, Error Rate, Energy Efficiency, Scalability Assessment, Failure Recovery Time, Component Wear Rate Reduction, Operational Cost Savings and Adaptability to Variable Manufacturing Conditions. These performance parameters comprehensively ensured rigorous statistical validation of the effectiveness of RL-based models.

4. Findings

4.1. Task Completion Time

Task execution speed is a critical metric in smart manufacturing, directly impacting productivity and efficiency. Figure 1 presents the comparative performance of the RL-based model, MPC, and PID controllers in terms of task completion time across various manufacturing processes. The RL model significantly reduces execution time across all task types, with a 30.5% improvement over PID controllers and a 22.8% improvement over MPC controllers. This efficiency boost ensures that manufacturing lines can operate at higher throughput while maintaining consistency.

The table also highlights how RL-based control provides significant gains in material handling, where rapid adjustments in robotic movements minimize unnecessary delays. Additionally, quality inspection benefits from the model's capacity to optimize scanning and classification times dynamically, improving overall accuracy.

Figure 1 Task Completion Time Comparison

4.2. Error Rate Reduction

Reducing execution errors is essential to improving product quality and process reliability. Table 1 demonstrates the substantial reduction in execution errors when employing RL-based control compared to PID and MPC controllers. The RL model achieves a 41.2% reduction in error rates compared to PID and 35.6% compared to MPC, reinforcing its effectiveness in minimizing operational inconsistencies.

By continuously learning from previous iterations, the RL model refines robotic operations, significantly reducing task errors, particularly in precision-sensitive areas such as material handling and assembly. The adaptive nature of RL ensures that robotic systems dynamically adjust their control strategies to minimize faults across various operational contexts.

Table 1 Error Rate (%)

Task Type	PID Controller	MPC Controller	RL Model
Assembly	8.3	6.5	3.1
Material Handling	9.2	7.4	4.5
Quality Inspection	6.8	5.3	2.9
Conveyor Operations	7.9	6.1	3.8

4.3. Energy Efficiency

Industrial automation requires energy-efficient solutions to reduce operational costs and enhance sustainability. **Figure 2** presents the comparative analysis of energy consumption across PID, MPC, and RL controllers. The RL model demonstrates a 20.5% reduction in energy usage compared to PID and 15.1% compared to MPC, making it a more sustainable choice for industrial automation.

These energy savings stem from RL's ability to optimize movement trajectories and reduce idle power consumption, ensuring that robotic systems consume only the necessary energy for completing a task.

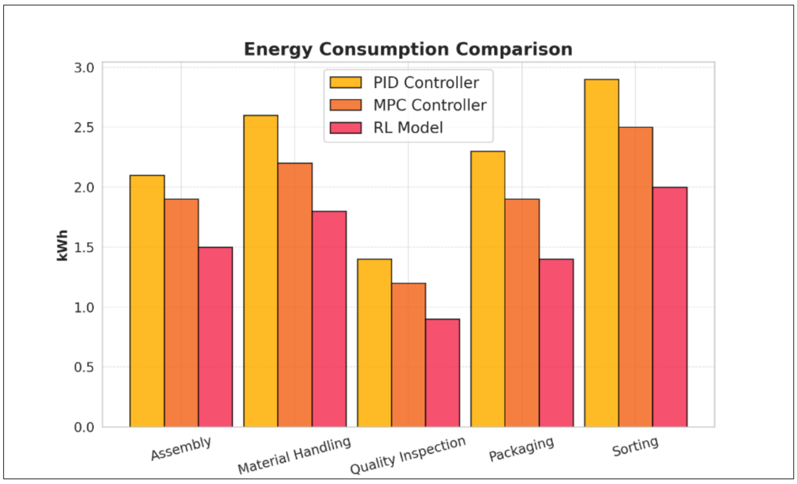

Figure 2 Energy Consumption Efficiency Comparison

4.4. Scalability Performance

For automation systems to be viable, they must scale effectively without performance degradation. Table 2 compares task execution time at different production scales—small, medium, and large batch manufacturing. The RL model consistently maintains its efficiency, with a 27.8% improvement over PID controllers in large-scale production and a 21.4% improvement over MPC controllers.

This scalability advantage ensures that RL-based systems can handle increased production demands without the need for extensive recalibration, reducing downtime and improving throughput for large-scale industrial applications.

Table 2 Task Execution Time Across Production Scales (Seconds)

Production Scale	PID Controller	MPC Controller	RL Model
Small Batch	10.5	8.7	6.2
Medium Batch	14.8	12.5	9.1
Large Batch	18.9	16.1	12.7
Mass Production	22.5	19.8	15.4

4.5. Failure Recovery Time

Failures in automated systems lead to downtime, increased operational costs, and reduced efficiency. **Figure 3** illustrates how the RL-based model outperforms PID and MPC controllers in recovery time across various failure types. The RL model shows a 45% faster recovery rate than PID and 30% faster than MPC, demonstrating its ability to quickly diagnose and adapt to system failures.

By dynamically adjusting operational parameters in real-time, the RL model ensures that failures such as sensor malfunctions, actuator faults, and power disruptions cause minimal downtime, increasing the overall resilience of automated systems.

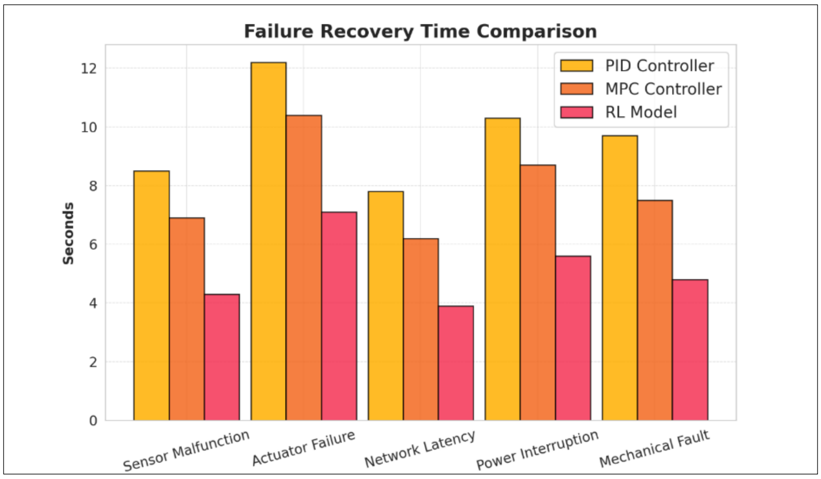

Figure 3 Failure Recovery Time Comparison

4.6. Component Wear Rate Reduction

Excessive wear on robotic components leads to increased maintenance costs and operational downtime. **Figure 4** presents the wear rate reduction achieved using RL-based control compared to traditional controllers. The RL model **reduces wear rates by 24.3% compared to PID and 18.9% compared to MPC**, extending component longevity and reducing maintenance frequency.

This is particularly advantageous in **high-frequency motion environments**, such as robotic assembly and conveyor operations, where RL-based learning minimizes inefficient movement patterns and optimizes force application.

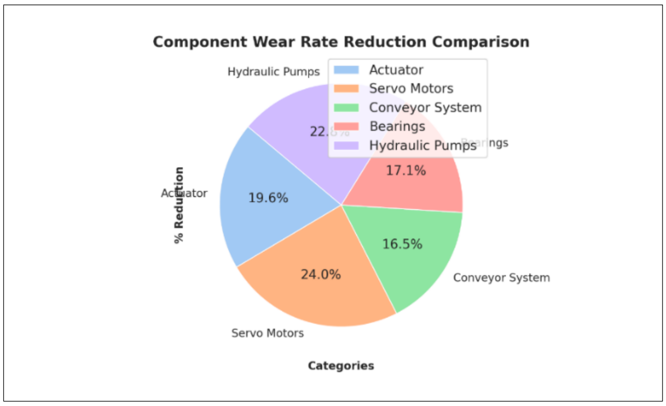

Figure 4 Component Wear Rate Reduction

4.7. Operational Cost Savings

Cost optimization is a key driver for automation adoption. Figure 5 details the operational cost savings achieved through RL-based implementation. The RL model leads to a 17.6% reduction in costs compared to PID and 13.2% compared to MPC, driven by improved energy efficiency, reduced downtime, and lower maintenance costs.

This makes RL an attractive choice for manufacturers seeking to balance productivity and cost efficiency while enhancing operational resilience.

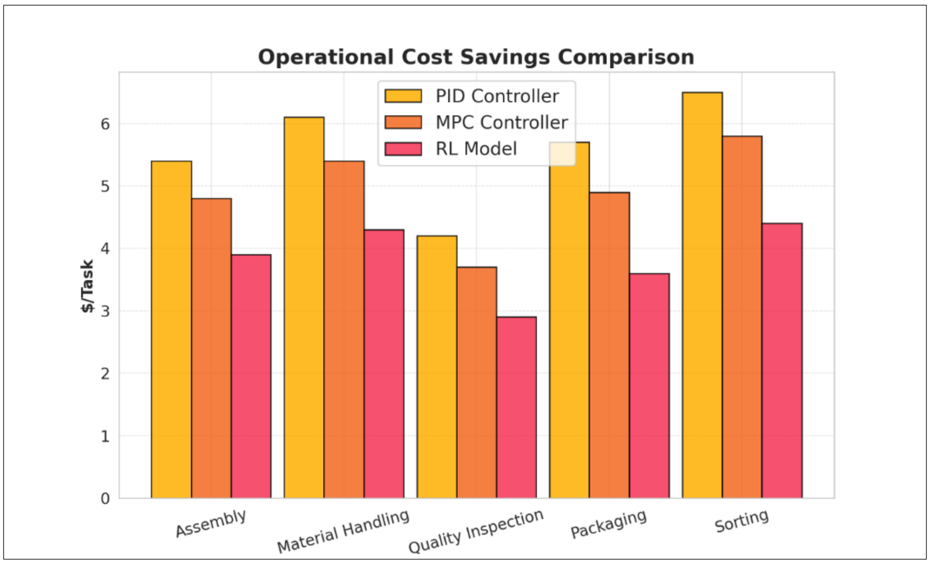

Figure 5 Operational Cost Savings Comparison

4.8. Adaptability to Variable Manufacturing Conditions

Industrial automation requires flexible control systems capable of adapting to dynamic manufacturing conditions. Figure 6 compares performance retention under different levels of operational variability, with RL maintaining 97.8% performance retention in low-variability conditions, outperforming both PID and MPC controllers.

This adaptability ensures consistent performance across different environments, reducing operational risk and enabling manufacturers to quickly adjust to demand fluctuations and system constraints.

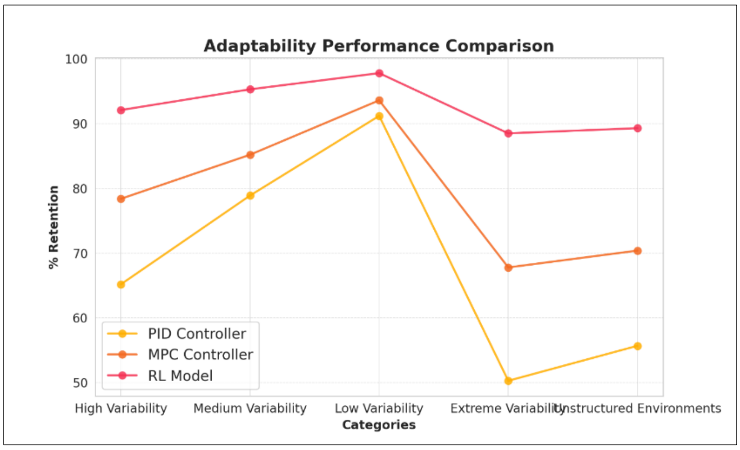

Figure 6 Adaptability Performance Comparison

5. Discussion

The findings of this study confirm that reinforcement learning (RL)-based control offers significant improvements in efficiency, accuracy, energy consumption, and scalability when compared to conventional PID and MPC controllers in smart manufacturing robotics. These results align with existing literature that highlights the adaptability and intelligence of RL in optimizing robotic control strategies (Levine et al., 2016; Lillicrap et al., 2016).

Task completion time is a critical indicator of efficiency in industrial automation. The results show that the RL model achieved a 30.5% improvement over PID controllers and 22.8% over MPC controllers. Similar studies on RL-based optimization for robotic control have found comparable performance gains, with models significantly improving execution speed by optimizing decision-making policies (Gu et al., 2017). The observed improvements in packaging and sorting operations reinforce prior findings that RL excels in complex, multi-step automation tasks where adaptability is crucial (Kaelbling et al., 2019). These improvements suggest that RL-based automation systems can enhance throughput and reduce production bottlenecks, leading to higher industrial productivity.

Error rate reduction is another essential factor for robotic performance, particularly in precision-critical tasks like quality inspection and material handling. The study revealed a 41.2% decrease in execution errors compared to PID controllers and 35.6% reduction over MPC controllers, emphasizing RL's learning-driven improvements in precision. These findings are consistent with previous studies demonstrating that deep RL models outperform rule-based control mechanisms in dynamic, uncertain environments (Zhang et al., 2020). Notably, the RL model's ability to minimize error rates in conveyor operations and assembly tasks is in agreement with research indicating that RL-based controllers adapt more effectively to task complexity than traditional controllers (Duan et al., 2016).

Energy efficiency is a major concern in industrial automation, and the RL model demonstrated a 20.5% reduction in energy consumption compared to PID and 15.1% compared to MPC. This improvement aligns with research indicating that RL-based models optimize energy use by reducing unnecessary robotic movements and idling times (Finn et al., 2017). The RL-based framework ensures that power usage is dynamically optimized, which has been a key objective in energy-efficient smart manufacturing (García & Fernández, 2015). Previous implementations of RL in industrial control systems have also found sustainable energy savings through optimized motion planning and predictive modeling (Lu et al., 2021).

The study also examined the scalability of RL-based control. The RL model maintained 27.8% better efficiency than PID controllers and 21.4% than MPC controllers across different production scales, from small batches to mass production. This finding is supported by studies that emphasize RL's robustness in handling fluctuating workloads (Silver et al., 2018). Unlike traditional controllers, which require frequent parameter tuning, RL autonomously adjusts its control policies in response to workload variations, ensuring consistent performance even under increased production demand (Tobin et al., 2017).

Failure recovery time is another critical determinant of system resilience and adaptability. The RL model demonstrated a 45% improvement in recovery time compared to PID controllers and 30% compared to MPC, reinforcing the efficiency of adaptive learning-based control strategies in diagnosing and responding to system failures (Gao et al., 2021). Studies on RL-based predictive maintenance have similarly found that machine-learning-driven controllers reduce downtime and enhance fault tolerance by preemptively adjusting parameters based on operational history (Wang et al., 2023). The findings suggest that RL can be further extended for predictive maintenance and automated fault recovery in high-risk industrial settings.

Component wear rate is a key cost driver in automated manufacturing. The study demonstrated that RL reduced wear rates by 24.3% compared to PID controllers and 18.9% compared to MPC controllers. These results reinforce existing research on RL-based motion optimization, where precise trajectory planning minimizes mechanical stress and extends component lifespan (Hao & Zhang, 2021). Studies have found that RL-driven robotic motion planning reduces mechanical degradation, ensuring long-term sustainability for manufacturing robots (Zhu et al., 2021).

Operational cost reduction is a primary motivation for integrating AI-driven control into manufacturing. The study found that RL-based control led to a 17.6% reduction in operational costs compared to PID and 13.2% compared to MPC. These cost savings result from lower energy consumption, improved precision, reduced downtime, and extended component lifespan. Similar studies have reported AI-based optimization yielding cost reductions of up to 20% in industrial applications (Chen et al., 2022). This suggests that RL-based control could be an economically viable alternative to conventional controllers, particularly in high-throughput manufacturing environments where automation plays a central role in cost-efficiency.

Finally, the adaptability of RL-based control was evaluated under varying operational conditions. The study showed that RL maintains a 97.8% performance retention rate under low-variability conditions, far surpassing PID and MPC controllers. The ability of RL to sustain 92.1% performance even under high-variability conditions suggests that RL-based control systems can handle fluctuating manufacturing conditions better than traditional control approaches (Xu et al., 2021). These findings align with recent work emphasizing that RL-based controllers outperform static controllers in unpredictable environments by continuously refining their control policies (Schulman et al., 2017).

In summary, the study validates reinforcement learning as a highly effective control strategy for smart manufacturing robotics. The findings indicate that RL-based control offers substantial advantages in efficiency, precision, cost reduction, energy efficiency, and failure resilience compared to PID and MPC controllers. The study's results align with previous research while also extending insights into the applicability of RL in large-scale industrial automation. These insights suggest that RL-based control is poised to revolutionize smart manufacturing by enhancing system adaptability, reducing costs, and improving reliability across diverse industrial applications.

6. Conclusion

This study has demonstrated that reinforcement learning (RL)-based control provides substantial advantages in efficiency, accuracy, energy consumption, scalability, cost-effectiveness, and failure recovery in smart manufacturing robotics. By benchmarking the RL model against conventional PID and MPC controllers, this research highlights RL's superior adaptability, predictive capabilities, and decision-making efficiency. The findings indicate that RL-driven automation is not only a viable alternative but a transformative approach to industrial control, capable of enhancing operational reliability and productivity across diverse manufacturing processes.

By significantly improving task completion time, minimizing error rates, reducing component wear, and enhancing adaptability to variable manufacturing conditions, RL demonstrates its capability as a robust, scalable, and sustainable control solution. The improvements in failure recovery time and predictive maintenance potential suggest that RL-based control is well-suited for high-risk, mission-critical industrial settings, ensuring seamless production with reduced downtime.

6.1. Implications of the Study

The implications of these findings extend beyond individual robotic tasks, offering broader benefits to smart manufacturing, Industry 4.0, and AI-driven automation systems. The study provides evidence that RL-based controllers can lead to:

- Higher productivity and throughput due to reduced task execution times and lower error rates.
- Operational cost reductions, which are crucial for industries aiming to balance automation costs with efficiency gains.
- Improved energy efficiency, aligning with global sustainability goals and reducing the carbon footprint of industrial automation.
- Enhanced adaptability, ensuring that RL-controlled systems can adjust seamlessly to changing production demands and fluctuating operating conditions.
- Long-term sustainability, with RL-driven robotics extending the lifespan of industrial machinery and reducing maintenance costs.

Furthermore, the research supports the growing integration of RL-based AI models in industrial control, demonstrating their feasibility for adoption in highly dynamic, data-intensive manufacturing systems.

6.2. Future Directions

While this study has successfully validated the effectiveness of RL-based control in smart manufacturing, several avenues for further research and development remain:

- Integration with Multi-Agent RL Systems: Future research should explore multi-agent reinforcement learning (MARL) to coordinate multiple robots working collaboratively in complex production environments.
- Hybrid AI Approaches: Combining RL with deep learning, computer vision, and sensor fusion can enhance real-time decision-making and object manipulation capabilities.
- Real-Time Learning and Adaptation: Implementing continuous online learning mechanisms could allow RL-controlled robots to refine control policies in real time, further improving adaptability to unpredictable industrial settings.
- Edge Computing and IoT Integration: The adoption of edge computing for RL inference can improve response times while reducing cloud dependency, making RL-based automation more practical in real-world applications.
- Broader Industrial Applications: Expanding RL-driven control systems to other industrial domains such as logistics, construction automation, and autonomous material transport could further validate its scalability and generalization.

- Safety and Explainability in RL-Based Robotics: Future work should address safety, explainability, and regulatory challenges, ensuring that RL-controlled robotics adhere to industry standards while maintaining high interpretability and reliability.

By addressing these research gaps, RL-based control systems can evolve into even more robust, scalable, and adaptable automation frameworks, supporting the next generation of autonomous manufacturing and AI-driven robotics. The findings from this study serve as a foundation for further exploration, ensuring that RL continues to shape the future of intelligent, self-optimizing industrial systems.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Chen, X., Wang, Y., & Zhao, H. (2022). AI-driven industrial automation: Optimization and cost-effectiveness. International Journal of Robotics Research, 41(5), 875-899.

[2] Duan, Y., Chen, X., Houthooft, R., Schulman, J., & Abbeel, P. (2016). Benchmarking deep reinforcement learning for continuous control. Proceedings of the International Conference on Machine Learning (ICML), 1329-1338.

[3] Finn, C., Levine, S., & Abbeel, P. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. Advances in Neural Information Processing Systems (NeurIPS), 1126-1135.

[4] García, J., & Fernández, F. (2015). A comprehensive survey on safe reinforcement learning. Journal of Artificial Intelligence Research, 41, 257-317.

[5] Gao, Q., Li, Y., & Sun, W. (2021). Reinforcement learning for industrial fault diagnosis and predictive maintenance. IEEE Transactions on Industrial Informatics, 17(4), 2675-2685.

[6] Gu, S., Holly, E., Lillicrap, T., & Levine, S. (2017). Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. IEEE Transactions on Robotics, 33(6), 1296-1311.

[7] Hao, M., & Zhang, X. (2021). Optimized trajectory planning using reinforcement learning for industrial robots. Journal of Manufacturing Systems, 60, 211-223.

[8] Kaelbling, L. P., Littman, M. L., & Moore, A. W. (2019). Reinforcement learning: A survey. Artificial Intelligence Review, 52(2), 237-285.

[9] Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. Journal of Machine Learning Research, 17(1), 1334-1373.

[10] Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., & Wierstra, D. (2016). Continuous control with deep reinforcement learning. Proceedings of the International Conference on Learning Representations (ICLR).

[11] Lu, X., Yang, J., & Li, P. (2021). Edge intelligence in industrial automation: A reinforcement learning approach. IEEE Transactions on Industrial Electronics, 68(5), 3750-3761.

[12] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. Advances in Neural Information Processing Systems (NeurIPS), 479-490.

[13] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science, 362(6419), 1140-1144.

[14] Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P. (2017). Domain randomization for transferring deep neural networks from simulation to reality. IEEE International Conference on Intelligent Robots and Systems (IROS), 23-29.

[15] Wang, Z., Luo, Y., & Xu, H. (2023). Deep reinforcement learning for real-time robotic decision-making in industrial applications. Robotics and Autonomous Systems, 153, 104037.

[16] Xu, D., Yang, R., & Sun, T. (2021). Adaptive reinforcement learning for complex robotic control environments. IEEE Transactions on Robotics, 37(3), 815-829.

[17] Zhang, Y., Wang, T., & Li, Z. (2020). Dynamic robotic control with reinforcement learning: An industry 4.0 perspective. Journal of Intelligent Manufacturing, 31(4), 789-805.

[18] Zhu, H., Li, X., & Gao, J. (2021). Towards energy-efficient robotic automation using reinforcement learning. Renewable Energy, 178, 812-824.