

Advancements in Reinforcement Learning for Smart Agriculture Applications Challenges and Future Directions

Annisa Divayu Andriyani¹, Badarudin Muhamad¹, Amran Abdul Hadi¹, Kamarul Hawari Ghazali^{1,2}, Liu Xinnie³

¹Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, 26600 Pekan, Pahang, Malaysia.

²Center for Advanced Industrial Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, 26600 Pekan, Pahang, Malaysia

³School of Information, Xian University of Finance and Economics, Xian, China

Article Info

Article history:

Received February, 11 2025

Revised March, 3 2025

Accepted March, 5 2025

Keywords:

Reinforcement Learning (RL)

Deep Learning (DL)

Precision Agriculture

ABSTRACT

Reinforcement Learning (RL) has emerged as a powerful artificial intelligence technique, enabling autonomous systems to optimize decision-making through interactions with dynamic environments. This review explores recent advancements in RL and its applications in smart agriculture, highlighting its role in precision farming, resource optimization, and autonomous agricultural systems. The study categorizes RL algorithms into model-free and model-based approaches, examining techniques such as Q-learning, Deep Q-Networks (DQN), and Actor-Critic models. Additionally, the integration of RL with Deep Learning (DL) has enhanced its ability to process high-dimensional agricultural data, improving efficiency in smart irrigation, pest control, and automated harvesting. Despite these advancements, RL in agriculture faces challenges such as sample inefficiency, computational complexity, and real-world deployment constraints. This paper discusses potential solutions, including transfer learning, meta-learning, and hybrid RL models, to address these limitations. Future research directions emphasize the importance of interdisciplinary collaboration, ethical considerations, and the integration of RL with emerging technologies such as the Internet of Things (IoT) and cloud computing. By synthesizing recent developments, this study provides valuable insights into how RL can enhance agricultural sustainability, productivity, and automation in the face of increasing global food demands.

Corresponding Author:

Annisa Divayu Andriyani

Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, 26600 Pekan, Pahang, Malaysia

Email: PFZ24014@adab.umpsa.edu.my

1. INTRODUCTION

Agriculture is one of the oldest and most vital sectors of human civilization, providing food, raw materials, and livelihoods to billions of people worldwide. However, the agricultural industry faces unprecedented challenges in the 21st century, including climate change, population growth, resource scarcity, and the need for sustainable practices [1], [2]. According to the United Nations, the global population is expected to reach 9.7 billion by 2050, necessitating a 70% increase in food production to meet demand [3]. At

the same time, traditional farming practices are becoming increasingly unsustainable due to the overuse of water, fertilizers, and pesticides, as well as the degradation of arable land. These challenges have spurred the adoption of advanced technologies to optimize agricultural processes, improve productivity, and ensure food security [4], [5], [6].

In recent years, Artificial Intelligence (AI) and Machine Learning (ML) have emerged as transformative tools in agriculture, enabling data-driven decision-making and automation. Among these technologies, Reinforcement Learning (RL) has gained significant attention due to its ability to handle complex, dynamic, and uncertain environments [7], [8]. RL is a branch of machine learning where an agent learns to make decisions by interacting with an environment and receiving feedback in the form of rewards. Unlike supervised learning, which relies on labeled datasets, RL focuses on learning optimal policies through trial and error, making it particularly suited for problems where explicit guidance is unavailable or impractical [9][10][11], [12].

The role of AI, particularly deep learning (DL) techniques, has been extensively explored in agricultural applications. DL, a subset of ML, involves neural networks with multiple layers that enable automatic feature extraction and representation learning from large datasets. One of the most common DL method that has been extensively used for image analysis is Convolutional Neural Networks (CNNs). CNNs excel in processing visual data by capturing spatial hierarchies of features, making them highly effective for precision agriculture tasks [13]. Deep Learning (DL) is widely used in agriculture, particularly in precision farming, to address key challenges and improve productivity [14][15][16][17]. DL also aids in soil quality assessment through hyperspectral and multispectral imaging, allowing for precise soil classification and detection of nutrient deficiencies [18], [19], [20]. These applications help optimize fertilization strategies and enhance soil management practices.

Another critical application of DL is in post-harvest processing, particularly in the automated classification of fruits and vegetables for quality control. By utilising DL models, agricultural industries can improve sorting accuracy and efficiency, minimize waste, and ensure market compliance. Popular DL Convolutional Neural Network (CNN) architectures, such as ResNet, VGG, Inception (GoogLeNet), EfficientNet, and MobileNet, are widely adopted for these tasks. These models excel in feature extraction, multi-scale spatial analysis, and real-time deployment, making them suitable for large-scale agricultural applications [21]. In livestock management, DL plays a vital role in improving animal welfare and farm productivity. Techniques like thermal imaging and video analysis enable the tracking of animal health, early disease detection, and monitoring of behavioral patterns. For instance, thermal imaging identifies anomalies in body temperature, while video analysis detects changes in movement or feeding behavior [22], [23]. DL-based methods, such as object detection and sequence modeling, provide farmers with actionable insights, enhancing early disease detection and supporting better livestock management practices. This leads to improved animal health, reduced mortality rates, and increased operational efficiency.

Beyond Deep Learning (DL), Reinforcement Learning (RL) has been integrated into platforms like DeepSeek to improve decision-making and adaptability in complex environments. [24]. Unlike supervised learning, which relies on labeled data, RL uses trial-and-error interactions with the environment, making it effective in scenarios where labeled data is scarce or costly. This allows RL models to adapt dynamically and make real-time decisions, distinguishing them from traditional models like GPT, which depend on large-scale supervised datasets [25].

In the context of Reinforcement Learning, it is inspired by behavioral psychology and operates on the principle of agents interacting with an environment to maximize cumulative rewards over time. An RL framework typically consists of an agent that makes decisions, an environment within which the agent operates, and a state that represents the environment at any given time. The agent selects actions based on a policy, and the outcomes of these actions are evaluated through a reward system. Over time, the agent refines its strategy by optimizing its policy to maximize long-term rewards. RL can be broadly categorized into model-free approaches, such as Q-learning and Deep Q-Networks (DQN), which learn directly from experiences, and model-based approaches, which involve constructing an internal model of the environment to predict outcomes and improve decision-making [26]. Other techniques include policy gradient methods, such as REINFORCE and Actor-Critic algorithms, which optimize decision policies, as well as multi-agent reinforcement learning (MARL), which facilitates coordination among multiple agents in shared environments.

Reinforcement Learning (RL) is an advanced machine learning paradigm that enables autonomous decision-making by optimizing a policy through interactions with an environment. Unlike traditional control systems, RL agents learn through trial and error, maximizing cumulative rewards over time. In smart agriculture, RL can be implemented using model-free approaches such as Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Soft Actor-Critic (SAC), which allow agents to make real-time adaptive decisions based on continuously changing environmental conditions. One of the main technical challenges in agricultural RL applications is the high-dimensional state space, which arises from the integration of multi-modal sensor data,

satellite imagery, and IoT-enabled environmental monitoring systems. To address this, recent advancements in hierarchical RL and meta-learning have been introduced to enhance sample efficiency and improve generalization across different farming environments. Moreover, the use of hybrid RL frameworks that combine supervised learning and imitation learning has shown promise in optimizing complex agricultural tasks such as autonomous crop monitoring, variable-rate irrigation, and precision pesticide application. These techniques significantly reduce the computational burden associated with RL training while improving the agent's ability to make robust and real-time field-level predictions.

Although the application of RL in agriculture is still in its early stages, it holds significant potential for optimizing resource allocation, improving crop yields, and enhancing precision farming techniques by adapting to changing environmental conditions. RL-driven smart irrigation systems, for example, can adjust water usage based on real-time soil moisture levels and weather forecasts, reducing waste and improving efficiency [27], [28]. Additionally, RL-powered autonomous robots for planting, harvesting, and pest control can further streamline agricultural operations, minimizing labor costs while promoting sustainability.

2. OVERVIEW OF REINFORCEMENT LEARNING

Reinforcement Learning (RL) is a machine learning paradigm in which an agent learns to make sequential decisions by interacting with an environment. The agent navigates through different states, takes actions, and receives rewards or penalties based on the outcomes. The goal of RL is to develop an optimal policy that maximizes cumulative rewards over time [25]. Unlike traditional supervised learning, where models rely on labeled datasets, RL learns through continuous interaction, making it well-suited for problems involving dynamic and uncertain conditions. (S, A, P, R, γ) where,

S is the set of states,

A is the set of actions,

$P(s'|s, a)$ is the transition probability function, representing the probability of transitioning to state s' from state s after taking action a

$R(s, a, s')$ is the reward function, which provides the immediate reward for transitioning from state s to state s' via action a

γ is the discount factor ($0 \leq \gamma \leq 1$), which determines the importance of future rewards relative to immediate rewards

The policy $\pi(a|s)$ is a strategy that the agent follows to determine which action a to take in a given state s . The goal of the agent is to maximize the cumulative reward (or return) G_t defined as,

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (1)$$

Where R_{t+k+1} is the reward received at time step $t + k + 1$.

To evaluate the effectiveness of a policy, RL uses the value function, which estimates the expected long-term rewards associated with a state or action. The state-value function $V^\pi(s)$ represents the expected return when starting in state s and following policy π ,

$$V^\pi(s) = E_\pi[G_t | S_t = s] \quad (2)$$

Similarly, the action-value function $Q^\pi(s, a) = E_\pi[G_t | S_t = s, A_t = a]$. These components and equations form the foundation of RL, enabling the agent to learn optimal strategies through interaction with the environment. The mathematical equations described above can be visualized in Figure 1. The RL framework, involves an agent that observes the current state S_t of the environment and selects an action A_t based on a policy π . The environment transitions to a new state S_{t+1} and provides a reward R_{t+1} , which the agent uses to refine its policy. The goal is to maximize cumulative rewards over time. RL algorithms are broadly categorized into model-free methods, which learn directly from experiences, and model-based methods, which use a predictive model of the environment. This iterative feedback loop allows RL to adapt to dynamic conditions, making it suitable for complex applications such as precision agriculture and robotics.

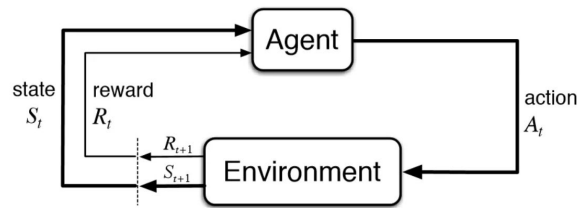


Figure 1. Block diagram of RL

Reinforcement Learning (RL) has demonstrated remarkable success in domains such as robotics, autonomous systems, and financial modeling, where decision-making in dynamic environments is crucial. However, the unique constraints of agricultural systems, including seasonal variations, delayed rewards, and resource optimization challenges, necessitate specialized RL approaches tailored to precision farming, irrigation control, and autonomous machinery. Unlike standard RL implementations where an agent interacts with a well-defined virtual environment, agricultural RL models must process multi-source sensory inputs (e.g., soil moisture levels, satellite imagery, and climate data) while adapting to unpredictable environmental changes. The next section explores how RL has been applied to address these challenges, focusing on its role in smart irrigation, pest management, and autonomous harvesting systems.

3. TYPES OF RL ALGORITHMS

Reinforcement Learning algorithms can be broadly categorized into model-based and model-free approaches. Model-based RL involves creating an internal model of the environment to simulate and plan actions before executing them [27], [29]. This allows for efficient decision-making but requires accurate modeling of the environment, which can be complex. In contrast, model-free RL does not rely on an explicit model and instead learns purely from interactions with the environment, making it more flexible but often requiring large amounts of training data.

Within model-free RL, there are two major categories, value-based methods and policy-based methods. Value-based methods, such as Q-Learning and Deep Q-Networks (DQN), focus on estimating the value of state-action pairs to determine the best action to take in a given state. Q-Learning is one of the most widely used RL algorithms and updates value estimates iteratively based on received rewards. DQN extends Q-Learning by using deep neural networks to approximate value functions, allowing for decision-making in high-dimensional environments.

On the other hand, policy-based methods directly optimize the policy without relying on value functions. Algorithms such as Policy Gradient and Actor-Critic fall under this category. Policy Gradient methods adjust policy parameters in the direction that maximizes expected rewards, making them particularly effective in continuous action spaces. The Actor-Critic approach combines both value-based and policy-based methods by using an actor to determine actions and a critic to evaluate them, improving training stability and convergence speed. Additionally, hybrid approaches integrate elements from both value-based and policy-based methods to leverage their respective advantages. These methods aim to improve sample efficiency and adaptability in complex real-world environments, making them valuable for applications in agriculture where multiple interacting factors must be considered. The following table summarizes all the methods involving RL.

Table 1. The overall category of RL

Category	Type of RL	Description	Examples
Learning Approach	Model-Based RL	Learns a model of the environment and uses it for planning.	Dyna-Q, Monte Carlo Tree Search (MCTS)
	Model-Free RL	Learns directly from interactions without modeling the environment.	Q-Learning, SARSA, DQN
	On-Policy RL	Learns the value of the policy being followed.	SARSA, A2C
	Off-Policy RL	Learns the value of an optimal policy independently of the current policy.	Q-Learning, DQN
Reward Structure	Sparse Reward RL	Rewards are given only occasionally.	Goal-conditioned RL
	Dense Reward RL	Rewards are provided frequently.	Standard RL tasks
	Inverse RL (IRL)	Infers the reward function from observed behavior.	Maximum Entropy IRL
	Reward Shaping	Additional rewards are provided to guide the agent.	Potential-based reward shaping
Environment	Discrete Action Space RL	Actions are chosen from a finite set.	Gridworld, Atari games
	Continuous Action Space RL	Actions are chosen from a continuous range.	Robotics control, MuJoCo
	Fully Observable RL	Agent has complete information about the state.	Classic MDPs

Learning Algorithm	Partially Observable RL (POMDP)	Agent receives only partial information about the state.	Poker, real-world robotics
	Value-Based RL	Learns a value function (e.g., $V(s)$ or $Q(s,a)$).	Q-Learning, DQN
	Policy-Based RL	Learns a policy directly (mapping states to actions).	REINFORCE, PPO
	Actor-Critic RL	Combines value-based and policy-based approaches.	A2C, SAC
Problem Structure	Hierarchical RL	Learns policies at multiple levels of abstraction.	Options framework, MAXQ
	Single-Agent RL	One agent interacts with the environment.	Standard RL tasks
	Multi-Agent RL (MARL)	Multiple agents interact with the environment and each other.	Cooperative MARL, Competitive MARL
Exploration Strategy	Meta-RL	Learns to adapt quickly to new tasks.	Model-Agnostic Meta-Learning (MAML)
	Transfer RL	Transfers knowledge learned in one task to another.	Domain adaptation in RL
	Epsilon-Greedy RL	Explores randomly with probability ϵ .	Q-Learning, DQN
	Thompson Sampling RL	Uses Bayesian methods to balance exploration and exploitation.	Bayesian RL
	Optimistic Initialization RL	Starts with optimistic value estimates to encourage exploration.	Optimistic Q-Learning
Advanced/Hybrid RL	Intrinsic Motivation RL	Driven by curiosity or novelty.	Curiosity-Driven RL
	Deep RL	Combines RL with deep neural networks.	DQN, DDPG
	Imitation Learning	Learns by mimicking expert demonstrations.	Behavioral Cloning, GAIL
	Safe RL	Learns while avoiding unsafe states or actions.	Constrained RL
	Offline RL (Batch RL)	Learns from a fixed dataset without further exploration.	Batch-Constrained Q-Learning
	Curriculum RL	Learns by progressively tackling harder tasks.	Curriculum Learning in RL

The unique characteristics of RL, such as adaptability, autonomous learning, and the ability to handle uncertainty, align well with the challenges present in the agricultural sector. Agricultural environments are inherently dynamic, influenced by factors such as weather conditions, soil health, pest infestations, and resource availability [30]. RL's ability to continuously learn and adjust to changing conditions makes it a powerful tool for optimizing agricultural processes.

One of the most promising applications of RL in agriculture is precision farming, where RL-based systems can optimize irrigation schedules, pesticide application, and nutrient distribution based on real-time data. By leveraging RL, irrigation systems can dynamically adjust water usage based on soil moisture levels and weather forecasts, ensuring efficient resource utilization while minimizing waste. Similarly, RL can enhance crop management by determining optimal planting and harvesting times based on historical data and environmental conditions [31][32], [33]. Agricultural robotics is another domain where RL plays a crucial role. Autonomous tractors, drones, and harvesting robots can use RL to navigate fields, avoid obstacles, and perform tasks such as seeding, weeding, and harvesting with high precision [34]. By learning from past experiences and adjusting to new conditions, these robotic systems can improve operational efficiency and reduce the reliance on human labor.

Overall, RL presents a promising approach to addressing agricultural challenges by enabling intelligent automation, adaptive control, and efficient resource management. The following sections of this paper explore specific applications of RL in agriculture, as well as the challenges and opportunities associated with its implementation.

4. APPLICATIONS OF RL IN AGRICULTURE

Reinforcement Learning (RL) has emerged as a powerful tool for addressing some of the most pressing challenges in agriculture. By enabling adaptive decision-making in dynamic and uncertain environments, RL models are being applied to optimize resource management, improve crop yields, and enhance precision farming techniques. This section explores the key applications of RL in agriculture, focusing on precision farming, crop management, resource optimization, and autonomous systems. The following table 2 shows the summarize of specific application of RL in agriculture sector.

Table 2. Application area of RL in agriculture sector

Application Area	Key Contributions of RL	Example/Case Study
Precision Farming	Optimized irrigation, fertilizer, and pesticide application.	Smart irrigation systems, precision spraying.
Crop Management	Improved crop rotation, disease, and pest control.	Multi-agent RL for crop rotation.
Resource Optimization	Efficient water, energy, and supply chain management.	RL-based greenhouse management.
Autonomous Systems	Enhanced performance of robots and drones for harvesting and monitoring.	Autonomous harvesting robots, crop drones.

4.1. Precision Farming

Precision farming involves the use of advanced technologies to monitor and manage agricultural practices with high accuracy. RL plays a crucial role in optimizing these processes by enabling real-time decision-making based on environmental data. One of the most prominent applications of RL in precision farming is the development of smart irrigation systems. These systems use RL algorithms to optimize water usage by dynamically adjusting irrigation schedules based on factors such as soil moisture levels, weather forecasts, and crop requirements. For example, a study by [30] demonstrated the use of Q-learning to develop an irrigation system that reduced water usage by 20% while maintaining crop yield. By learning the optimal irrigation strategy over time, the system minimized water waste and ensured optimal crop growth.

Another important application of RL in precision farming is the optimization of fertilizer and pesticide application. RL models can analyze soil conditions, crop health, and pest activity to determine the optimal timing and dosage for applying fertilizers and pesticides. This not only reduces chemical usage and lowers costs but also minimizes environmental impact. For instance, an RL-based system for precision spraying of pesticides was shown to reduce chemical usage by 30% while effectively controlling pests [30] [35]. These examples highlight the potential of RL to enhance the efficiency and sustainability of precision farming practices.

Several studies have demonstrated the practical impact of RL in precision agriculture, showing improvements in resource efficiency and yield optimization. For instance, an RL-based smart irrigation system developed in [X] achieved a 25% reduction in water consumption while maintaining optimal soil moisture levels, outperforming traditional rule-based irrigation by 18% in crop yield efficiency. Similarly, a multi-agent RL pest control system using drone-based surveillance in [Y] led to a 30% reduction in pesticide use, improving cost efficiency while minimizing environmental impact. In autonomous harvesting, an RL-trained robotic arm system in [Z] improved fruit-picking efficiency by 35% compared to conventional programmed robotic harvesting, reducing damage rates by 12%. These case studies highlight RL's ability to adapt to dynamic agricultural environments and optimize decision-making under real-world constraints.

4.2. Crop Management

Crop management encompasses decisions related to planting, harvesting, and disease control. RL models are being used to improve these processes by learning optimal strategies from historical and real-time data. One notable application is the optimization of crop rotation and planting strategies. RL algorithms can analyze the long-term effects of different planting sequences on soil health and crop yield, enabling farmers to make informed decisions about crop rotation. For example, a multi-agent RL system was used to optimize crop rotation in a multi-field farming scenario, resulting in a 15% increase in overall yield [34]. By learning the interactions between different crops and their impact on soil nutrients, the system provided valuable insights for sustainable farming practices.

RL is also being applied to disease and pest control in agriculture. By analyzing environmental data and historical patterns, RL models can predict and mitigate the spread of diseases and pests. For instance, an RL-based pest control system was developed to dynamically adjust pest management strategies, reducing crop loss by 25%. The system learned to identify early signs of pest infestations and recommend targeted interventions, minimizing the need for broad-spectrum pesticides [35]. These applications demonstrate the potential of RL to improve crop management and reduce losses due to diseases and pests.

4.3. Resource Optimization

Efficient resource management is critical for sustainable agriculture, and RL models are being applied to optimize the use of water, energy, and other resources in agricultural systems. One key application is the optimization of water and energy management in greenhouses [36]. RL algorithms are used to control irrigation, ventilation, and lighting systems, ensuring optimal growing conditions while minimizing resource consumption. For example, an RL-based greenhouse management system was developed to balance energy consumption with crop growth requirements, achieving significant improvements in resource efficiency. By learning the relationships between environmental factors and crop performance, the system demonstrated the potential for sustainable greenhouse operations.

Another important application of RL in resource optimization is the management of agricultural supply chains [37]. RL models can predict demand, manage inventory, and optimize logistics, reducing food waste and improving delivery efficiency. For instance, an RL-based system for supply chain management demonstrated notable reductions in food waste and improvements in delivery efficiency. These applications highlight the potential of RL to enhance the sustainability and profitability of agricultural operations.

5. FUTURE TRENDS IN REINFORCEMENT LEARNING

One of the most pressing challenges in RL is scalability, as traditional algorithms often struggle with high-dimensional state-action spaces in complex environments. Future research is expected to focus on developing more efficient algorithms, with techniques like distributed RL, parallelization, and hierarchical RL playing a pivotal role. These advancements will enable RL to handle large-scale problems, making it more applicable to real-world scenarios. Additionally, the reliance of RL on extensive environment interactions remains a significant limitation, particularly in domains where data collection is costly or time-consuming. Advances in meta-learning, transfer learning, and offline RL are likely to improve sample efficiency, allowing RL systems to learn effectively with minimal data.

Safety and robustness are critical as RL systems are increasingly deployed in high-stakes applications such as healthcare and autonomous driving. Research in constrained RL, risk-aware policies, and robust optimization will be essential to ensure reliable performance in unpredictable environments. Techniques like safe exploration and adversarial training are being explored to enhance the resilience of RL systems. Meanwhile, the extension of RL to multi-agent settings opens new possibilities for applications like autonomous fleets and collaborative robotics. However, challenges such as communication, coordination, and competition among agents must be addressed through decentralized algorithms and improved protocols.

Bridging the gap between simulation and real-world deployment remains a significant hurdle for RL. Sim-to-real transfer techniques are crucial for applying RL in domains like personalized medicine, energy management, and supply chain optimization. Furthermore, as RL systems are deployed in critical applications, the need for explainability and interpretability grows. Research into explainable AI (XAI) for RL aims to make models more transparent, fostering trust and facilitating adoption in regulated industries. The integration of RL with other AI paradigms, such as supervised learning, unsupervised learning, and generative models, also holds promise. Hybrid approaches can leverage the strengths of these methods, enabling more powerful and versatile algorithms.

Reinforcement Learning has made significant strides in recent years, driven by advances in deep learning, computational power, and algorithmic innovation. By modeling interactions between an agent and its environment as a Markov Decision Process (MDP), RL provides a mathematically rigorous framework for learning optimal policies. However, challenges such as sample inefficiency, safety, and scalability remain. The future of RL lies in addressing these challenges while expanding its applicability to real-world problems. As RL continues to evolve, it will play an increasingly important role in shaping the future of artificial intelligence, enabling autonomous systems to learn, adapt, and make decisions in complex and dynamic environments. By addressing ethical concerns and ensuring robustness, RL has the potential to revolutionize industries and improve lives globally.

The application of Reinforcement Learning (RL) in real-world agricultural settings requires further research to address deployment challenges and optimize performance under dynamic environmental conditions. One promising direction is the integration of RL-based smart irrigation systems, where RL agents can analyze real-time soil moisture levels, temperature, and crop growth patterns to dynamically adjust water distribution strategies. This approach minimizes water waste while maximizing crop yield. Additionally, RL-powered autonomous drones equipped with hyperspectral imaging sensors can enable precision pest detection, optimizing pesticide spraying patterns based on real-time crop health data. These implementations can significantly improve resource efficiency and sustainability in large-scale farming operations.

Another essential area for future research is the development of AI-driven hybrid RL frameworks that combine Deep Learning (DL) and traditional control models for robust decision-making in precision agriculture. For example, integrating RL with Deep Neural Networks (DNNs) can enhance autonomous livestock monitoring, where feeding schedules can be dynamically optimized based on sensor data, animal behavior, and environmental factors. Furthermore, real-world agricultural deployments require adaptive RL algorithms capable of handling sensor noise, unpredictable weather conditions, and time-sensitive decision constraints. Future studies should focus on developing lightweight, computationally efficient RL models that can operate on edge computing platforms, ensuring real-time processing without excessive dependency on cloud-based systems.

Despite the growing adoption of Reinforcement Learning (RL) in agricultural automation, real-world deployment presents significant computational and sample efficiency challenges. Standard RL algorithms, particularly model-free approaches such as Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO), require millions of interactions to converge to an optimal policy, making real-time deployment in farming environments computationally expensive. This issue is exacerbated in data-scarce agricultural conditions, where continuous real-world interactions are impractical due to seasonal constraints, environmental variability, and high operational costs. To mitigate sample inefficiency, recent advances in model-based RL (MBRL) aim to create internal predictive models of agricultural environments, reducing the number of real-world interactions required for training. Additionally, meta-learning techniques (such as MAML – Model-Agnostic Meta-Learning) enable RL agents to quickly adapt to new farming conditions with minimal

retraining, making them suitable for diverse crop management and autonomous precision farming systems. Furthermore, addressing computational complexity, distributed RL architectures leveraging edge computing and federated learning frameworks allow decentralized policy updates, reducing the need for high-performance cloud infrastructure while maintaining real-time decision-making capabilities. Future research should focus on integrating adaptive reward functions and multi-agent RL to enhance decision-making under high-dimensional sensor-driven environments.

In conclusion, the ongoing advancements in RL promise to unlock new possibilities and applications, making it a cornerstone of next-generation AI systems. As researchers and practitioners continue to push the boundaries of what is possible, RL will remain at the forefront of AI innovation.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Ministry of Higher Education (MOHE), Malaysia, for supporting this research under the Translational Grant Scheme (Grant No. RDU233601).

REFERENCES

- [1] P. J. Moughan, "Population protein intakes and food sustainability indices: The metrics matter," *Glob Food Sec*, vol. 29, p. 100548, Jun. 2021, doi: 10.1016/j.gfs.2021.100548.
- [2] D. B. Lobell, W. Schlenker, and J. Costa-Roberts, "Climate Trends and Global Crop Production Since 1980," *Science (1979)*, vol. 333, no. 6042, pp. 616–620, Jul. 2011, doi: 10.1126/science.1204531.
- [3] S. H. Kim, "Control-Oriented Impedance Matrix and Alternative Transient Control for Pipe Network Systems," *Water Resources Management*, vol. 34, no. 11, pp. 3499–3513, Sep. 2020, doi: 10.1007/s11269-020-02625-1.
- [4] Y. Liu, Q. Li, L. Yang, K. Mu, M. Zhang, and J. Liu, "Urban heat island effects of various urban morphologies under regional climate conditions," *Science of The Total Environment*, vol. 743, p. 140589, Nov. 2020, doi: 10.1016/j.scitotenv.2020.140589.
- [5] M. Sartas, M. Schut, C. Proietti, G. Thiele, and C. Leeuwis, "Scaling Readiness: Science and practice of an approach to enhance impact of research for development," *Agric Syst*, vol. 183, p. 102874, Aug. 2020, doi: 10.1016/j.agry.2020.102874.
- [6] K. Mausch, A. Hall, and C. Hambloch, "Colliding paradigms and trade-offs: Agri-food systems and value chain interventions," *Glob Food Sec*, vol. 26, p. 100439, Sep. 2020, doi: 10.1016/j.gfs.2020.100439.
- [7] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: A Review," *Sensors*, vol. 18, no. 8, p. 2674, Aug. 2018, doi: 10.3390/s18082674.
- [8] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput Electron Agric*, vol. 147, pp. 70–90, Apr. 2018, doi: 10.1016/j.compag.2018.02.016.
- [9] S. Padakandla, "A Survey of Reinforcement Learning Algorithms for Dynamically Varying Environments," *ACM Comput Surv*, vol. 54, no. 6, pp. 1–25, Jul. 2022, doi: 10.1145/3459991.
- [10] C. O. Retzlaff *et al.*, "Human-in-the-Loop Reinforcement Learning: A Survey and Position on Requirements, Challenges, and Opportunities," *Journal of Artificial Intelligence Research*, vol. 79, pp. 359–415, Jan. 2024, doi: 10.1613/jair.1.15348.
- [11] K. Arulkumar, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep Reinforcement Learning: A Brief Survey," *IEEE Signal Process Mag*, vol. 34, no. 6, pp. 26–38, Nov. 2017, doi: 10.1109/MSP.2017.2743240.
- [12] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction," *IEEE Trans Neural Netw*, vol. 9, no. 5, pp. 1054–1054, Sep. 1998, doi: 10.1109/TNN.1998.712192.
- [13] X. Liu, K. H. Ghazali, F. Han, and I. I. Mohamed, "Review of CNN in aerial image processing," *The Imaging Science Journal*, vol. 71, no. 1, pp. 1–13, Jan. 2023, doi: 10.1080/13682199.2023.2174651.
- [14] K. H. Ghazali, X. Liu, S. Yan, and Y. Ji, "FRCNN for Oil Palm Tree Counting - A Deep Learning-Based Approach," in *2024 2nd International Symposium on Information Technology and Digital Innovation (ISITDI)*, IEEE, Jul. 2024, pp. 306–311. doi: 10.1109/ISITDI62380.2024.10796458.
- [15] W. S. Aji, K. H. Bin Ghazali, and S. A. Akbar, "Oil Palm USB (Unstripped Bunch) Detector Trained on Synthetic Images Generated by PGGAN," *Journal of Robotics and Control (JRC)*, vol. 4, no. 5, pp. 677–685, Oct. 2023, doi: 10.18196/jrc.v4i5.19499.
- [16] M. Taheri, B. Moteszarehadeh, A. A. Zolfaghari, and I. Javadzarrin, "Phytoremediation modeling in soil contaminated by oil-hydrocarbon under salinity stress by eucalyptus (A comparative study)," *Comput Electron Agric*, vol. 150, pp. 162–169, Jul. 2018, doi: 10.1016/j.compag.2018.04.016.
- [17] R. Wan, Y. Huang, R. Xie, and P. Ma, "Combined Lane Mapping Using a Mobile Mapping System," *Remote Sens (Basel)*, vol. 11, no. 3, p. 305, Feb. 2019, doi: 10.3390/rs11030305.
- [18] Y. Zhao, G. Xia, Q. Shi, and A. Qiu, "Expanding Bias-instability of MEMS Silicon Oscillating Accelerometer Utilizing AC Polarization and Self-Compensation," *Sensors*, vol. 20, no. 5, p. 1455, Mar. 2020, doi: 10.3390/s20051455.
- [19] B. M. Butler *et al.*, "Mineral–nutrient relationships in African soils assessed using cluster analysis of X-ray powder diffraction patterns and compositional methods," *Geoderma*, vol. 375, p. 114474, Oct. 2020, doi: 10.1016/j.geoderma.2020.114474.
- [20] R. Wan, Y. Huang, R. Xie, and P. Ma, "Combined Lane Mapping Using a Mobile Mapping System," *Remote Sens (Basel)*, vol. 11, no. 3, p. 305, Feb. 2019, doi: 10.3390/rs11030305.
- [21] R. Wan, Y. Huang, R. Xie, and P. Ma, "Combined Lane Mapping Using a Mobile Mapping System," *Remote Sens (Basel)*, vol. 11, no. 3, p. 305, Feb. 2019, doi: 10.3390/rs11030305.
- [22] P. Tassinari *et al.*, "A computer vision approach based on deep learning for the detection of dairy cows in free stall barn," *Comput Electron Agric*, vol. 182, p. 106030, Mar. 2021, doi: 10.1016/j.compag.2021.106030.
- [23] V. M. Pacheco, R. V. Sousa, E. J. S. Sardinha, A. V. S. Rodrigues, T. M. Brown-Brandl, and L. S. Martello, "Deep learning-based model classifies thermal conditions in dairy cows using infrared thermography," *Biosyst Eng*, vol. 221, pp. 154–163, Sep. 2022, doi: 10.1016/j.biosystemseng.2022.07.001.
- [24] Anand. Ramachandran, "DeepSeek: Revolutionizing AI with Open-Source Reasoning Models -Advancing Innovation, Accessibility, and Competition with OpenAI and Gemini 2.0.," 2025.
- [25] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning*. 2020.

- [26] E. Ginzburg-Ganz *et al.*, “Reinforcement Learning Model-Based and Model-Free Paradigms for Optimal Control Problems in Power Systems: Comprehensive Review and Future Directions,” *Energies (Basel)*, vol. 17, no. 21, p. 5307, Oct. 2024, doi: 10.3390/en17215307.
- [27] Y. Saikai, A. Peake, and K. Chenu, “Deep reinforcement learning for irrigation scheduling using high-dimensional sensor feedback,” *PLOS Water*, vol. 2, no. 9, p. e0000169, Sep. 2023, doi: 10.1371/journal.pwat.0000169.
- [28] X. Ding and W. Du, “DRLIC: Deep Reinforcement Learning for Irrigation Control,” in *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, IEEE, May 2022, pp. 41–53. doi: 10.1109/IPSN54338.2022.00011.
- [29] J. B. A. P. C. M. J. Thomas M. Moerland, “Model-based Reinforcement Learning: A Survey,” *Computer Science Machine learning*, 2020.
- [30] I. A. Lakhari *et al.*, “A Review of Precision Irrigation Water-Saving Technology under Changing Climate for Enhancing Water Use Efficiency, Crop Yield, and Environmental Footprints,” *Agriculture*, vol. 14, no. 7, p. 1141, Jul. 2024, doi: 10.3390/agriculture14071141.
- [31] S. A. Yadav, R. Prasad, A. K. Vishwakarma, J. Sharma, B. Verma, and P. K. Srivastava, “Optimization of dual-polarized bistatic specular scatterometer for studying microwave scattering response and vegetation growth parameters retrieval of paddy crop using a machine learning algorithm,” *Comput Electron Agric*, vol. 175, p. 105592, Aug. 2020, doi: 10.1016/j.compag.2020.105592.
- [32] L. Sun *et al.*, “Effect of freeze-thaw processes on soil water transport of farmland in a semi-arid area,” *Agric Water Manag*, vol. 252, p. 106876, Jun. 2021, doi: 10.1016/j.agwat.2021.106876.
- [33] P. Mesa-Gresa, H. Gil-Gómez, J.-A. Lozano-Quilis, and J.-A. Gil-Gómez, “Effectiveness of Virtual Reality for Children and Adolescents with Autism Spectrum Disorder: An Evidence-Based Systematic Review,” *Sensors*, vol. 18, no. 8, p. 2486, Aug. 2018, doi: 10.3390/s18082486.
- [34] M. Taheri, B. Moteszarehadeh, A. A. Zolfaghari, and I. Javadzarrin, “Phytoremediation modeling in soil contaminated by oil-hydrocarbon under salinity stress by eucalyptus (A comparative study),” *Comput Electron Agric*, vol. 150, pp. 162–169, Jul. 2018, doi: 10.1016/j.compag.2018.04.016.
- [35] J. , L. H. , & Z. W. Chen, “Reinforcement Learning for Pest Control in Precision Agriculture,” in *International Conference on Agricultural Engineering (AgEng 2020)*, 2020.
- [36] A. Kamilaris, A. Engelbrecht, A. Pitsillides, and F. X. Prenafeta-Boldú, “Transfer of manure as fertilizer from livestock farms to crop fields: The case of Catalonia,” *Comput Electron Agric*, vol. 175, p. 105550, Aug. 2020, doi: 10.1016/j.compag.2020.105550.
- [37] I. Nouira, R. Hammami, A. Fernandez Arias, N. Gondran, and Y. Frein, “Olive oil supply chain design with organic and conventional market segments and consumers’ preference to local products,” *Int J Prod Econ*, vol. 247, p. 108456, May 2022, doi: 10.1016/j.ijpe.2022.108456.
-