iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82488

A Survey on Agentic AI-Assisted UAV Swarm Systems Using Multi-Agent Deep Reinforcement Learning

Kamalakshi Naganna[1], Manasa G[2], Pragati Sharma[3], Soumyashree R A[4], Tejashwini S N[5]
[1]Professor, [2, 3, 4, 5]Department of Computer Science and Engineering, Sapthagiri College of Engineering

Abstract—Currently, the use of UAV swarm systems is widespread for large-scale monitoring and surveillance purposes, demanding effective coordination. While the Multi-Agent Deep Reinforcement Learning methods such as MADDPG leverage decentralization by providing each agent with the ability to act individually depending on the condition of its local environment, the learning-based approach is ineffective because of the redundant explorations and coordination that lack structure. In our review, we focus on recent advancements in the domain of UAV swarm systems, with an accent on inefficiencies inherent in MARL approaches, making it hard to achieve efficient coverage of the region of interest. The paper is devoted to hybrid approaches, which combine MARL approaches and spatial planning approaches, specifically Voronoi partitioning.

Index Terms—UAV swarm; Multi-Agent Deep Reinforcement Learning; MADDPG; MAPPO; Voronoi partitioning; Agentic AI; cooperative coordination; task allocation.

I. INTRODUCTION

Unmanned Aerial Vehicle (UAV) swarms have emerged as an effective solution for large-scale monitoring, surveillance, and search operations due to their ability to operate collaboratively in dynamic environments [11], [13], [27]. Unlike single UAV systems, swarm-based architectures leverage multiple autonomous agents to achieve improved coverage, robustness, and fault tolerance. These systems are widely applied in domains such as disaster response, environmental monitoring, wildlife tracking, and border surveillance.

Traditional UAV coordination approaches primarily rely on predefined path planning, optimization-based routing, or heuristic strategies. While such methods ensure predictable behavior, they lack adaptability in uncertain and dynamic environments, where real-time decision-making is essential [12], [17]. Moreover, these approaches often suffer from scalability issues and fail to efficiently manage coordination among multiple agents.

To address these limitations, Reinforcement Learning (RL) has been introduced as a data-driven approach that enables agents to learn optimal actions through interaction with the environment [1], [3], [5]. With the integration of deep learning, Multi-Agent Reinforcement Learning (MARL) has become a prominent framework for coordinating UAV swarms, allowing multiple agents to learn cooperative behaviors while operating in decentralized settings [1], [7], [21].

Among MARL algorithms, Multi-Agent Deep Deterministic Policy Gradient (MADDPG) has gained significant attention due to its ability to handle continuous action spaces and support centralized training with decentralized execution [1],[5], [33]. This enables UAV agents to make independent decisions based on local observations while benefiting from global information during training. However, existing MADDPG-based approaches primarily depend on reward-driven coordination, which can result in redundant exploration, inefficient coverage, and lack of structured task allocation [6], [8], [9].

Recent research has focused on addressing these challenges by integrating spatial decomposition and planning techniques with learning-based methods. Approaches such as Voronoi partitioning divide the operational environment into non-overlapping regions, enabling structured task allocation and reducing redundancy among agents [8], [9], [10]. Additionally, emerging paradigms such as Agentic AI aim to enhance decision-making by incorporating higher-level reasoning and adaptive planning capabilities into autonomous systems [26], [27].

This survey provides a comprehensive review of recent advancements in UAV swarm systems using Multi-Agent Deep Reinforcement Learning, with a particular focus on hybrid approaches that combine learning-based coordination with spatial planning techniques. The paper analyzes existing methodologies, identifies key limitations, and highlights future research directions for developing efficient, scalable, and intelligent UAV swarm systems.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

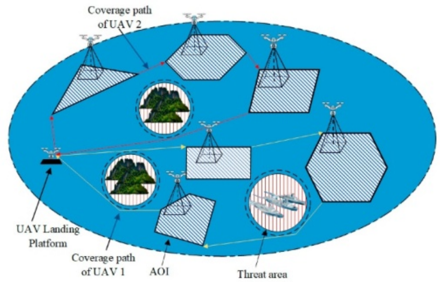


Fig. 1. UAV Swarm Coverage Representation

II. BACKGROUND

A. UAV Swarm Systems

Unmanned Aerial Vehicle (UAV) swarm systems consist of multiple autonomous aerial agents that cooperate to perform tasks collectively in a shared environment. Unlike single UAV systems, swarm-based architectures offer significant advantages in terms of scalability, fault tolerance, and operational efficiency, making them suitable for applications such as environmental monitoring, disaster response, surveillance, and search-and-rescue missions [11], [13], [27]. The concept of UAV swarms is inspired by natural collective behaviors observed in biological systems such as bird flocks, fish schools, and insect colonies, where global coordination emerges from simple local interactions among individual agents.

In engineered UAV swarm systems, coordination is achieved through a combination of communication protocols, control strategies, and decision-making algorithms. These systems are typically characterized by three key properties: autonomy, where each UAV operates independently; cooperation, where agents share information or divide tasks; and scalability, where system performance should remain stable as the number of agents increases. Achieving these properties simultaneously is a complex challenge due to limitations in communication bandwidth, dynamic environmental conditions, and resource constraints such as battery life and sensing capabilities [11], [12].

Traditional UAV coordination approaches rely heavily on centralized control architectures or pre-defined trajectory planning methods. While centralized systems provide a global view and optimized decision-making, they suffer from scalability issues and single points of failure. On the other hand, decentralized approaches improve robustness but often struggle to achieve efficient coordination due to limited information sharing among agents. These limitations have motivated the adoption of learning-based approaches that allow UAVs to adapt dynamically to changing environments.

B. Reinforcement Learning

Reinforcement Learning (RL) is a computational framework that enables agents to learn optimal behaviors through trial-and-error interactions with an environment. The environment is typically modeled as a Markov Decision Process (MDP), defined by a set of states, actions, transition probabilities, and reward functions [1], [3]. At each time step, the agent observes the current state, takes an action, and receives a reward based on the outcome. The objective is to learn a policy that maximizes the expected cumulative reward over time.

With the advancement of deep learning, Deep Reinforcement Learning (DRL) has emerged as a powerful approach for solving complex decision-making problems involving high-dimensional state and action spaces. Neural networks are used to approximate value functions or policies, enabling agents to generalize across different states [5]. Algorithms such as Deep Deterministic Policy Gradient (DDPG) are particularly suitable for continuous control tasks, where actions are represented as real-valued vectors rather than discrete choices. This makes DRL highly applicable to UAV navigation, where smooth and precise motion control is required.

Despite its success, single-agent RL is insufficient for systems involving multiple interacting agents, such as UAV swarms. In such scenarios, each agent must not only learn from the environment but also adapt to the behaviors of other agents, making the learning process significantly more complex.

C. Multi-Agent Reinforcement Learning

Multi-Agent Reinforcement Learning (MARL) extends the RL framework to environments where multiple agents operate simultaneously and interact with each other.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

In cooperative settings, agents aim to maximize a shared objective, requiring effective coordination and information sharing [1], [21]. However, MARL introduces several challenges that are not present in single-agent systems.

One of the primary challenges is non-stationarity, where the environment dynamics change as each agent updates its policy during training. From the perspective of an individual agent, the behavior of other agents appears as part of the environment, making it difficult to learn stable policies [1], [18]. Additionally, partial observability limits each agent's access to global information, further complicating decision-making.

To address these challenges, the Centralized Training and Decentralized Execution (CTDE) paradigm has been widely adopted. In this framework, agents are trained using global information, such as the states and actions of all agents, while execution is performed using only local observations. This allows agents to learn coordinated behaviors without requiring extensive communication during deployment. The Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm is a prominent example of this approach, where each agent has an actor network for selecting actions and a centralized critic that evaluates joint actions.

While MARL algorithms such as MADDPG have shown promising results in UAV swarm coordination, they often rely heavily on reward design to encourage cooperation. Poorly designed reward functions can lead to undesirable behaviors such as redundant exploration, inefficient coverage, or conflicts among agents [6], [8]. This limitation highlights the need for integrating structured planning mechanisms with learning-based approaches.

D. Spatial Planning and Agentic Approaches

Spatial planning techniques provide an effective way to improve coordination in multi-agent systems by introducing structure into the decision-making process. One widely used method is Voronoi partitioning, which divides the environment into regions based on proximity to a set of seed points, typically representing the positions of UAVs [8], [9], [10]. Each UAV is assigned a specific region, ensuring that agents focus on different parts of the environment and reducing overlap in coverage. This approach not only improves efficiency but also simplifies task allocation in large-scale systems.

In addition to geometric partitioning, recent research has explored the integration of higher-level decision-making frameworks under the concept of Agentic AI. Unlike traditional RL approaches that rely solely on reward signals, Agentic AI introduces goal-directed reasoning, planning, and adaptive behavior into autonomous systems [26], [27]. These systems are capable of dynamically adjusting their strategies based on changing conditions, enabling more intelligent and flexible coordination among agents.

The combination of MARL with spatial planning and agentic reasoning represents a promising direction for UAV swarm systems. By leveraging the strengths of both learning-based and planning-based approaches, it is possible to achieve improved coordination, scalability, and robustness in complex and dynamic environments.

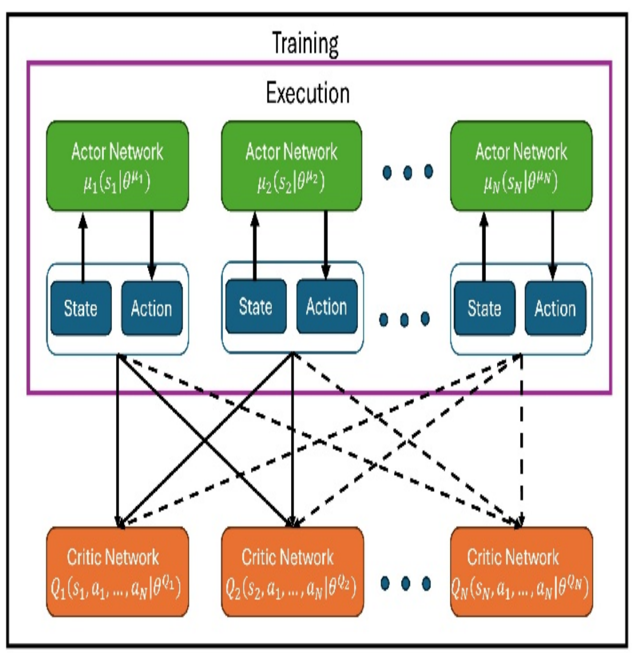


Fig. 2. Multi-Agent Reinforcement Learning Framework for UAV Swarms

III. LITERATURE REVIEW

Zhao et al. (2023) proposed a multi-weight MADDPG (MW-MADDPG) approach to improve coordination among UAV agents in dynamic environments. Their work enhances reward distribution using multiple critic weights, leading to improved convergence and cooperation. However, the approach still depends heavily on reward engineering and lacks explicit spatial coordination mechanisms. [1]

Wang and Jiao (2024) introduced CER-MADDPG, incorporating coordinated experience replay to improve learning efficiency. By prioritizing relevant transitions, the model achieves better convergence in multi-agent settings, though it does not address redundancy in spatial coverage. [2]

Kong et al. (2024) developed a reinforcement learning-based approach for UAV target assignment and path planning. Their method integrates task allocation with trajectory optimization, enabling adaptability to dynamic targets, but lacks structured coordination among agents. [3]

Chen et al. (2023) proposed an enhanced MADDPG-based UAV task assignment framework aimed at improving workload balancing among UAV agents. The study introduced additional constraints for efficient task distribution and demonstrated improved coordination performance. However, the system remained dependent on reward tuning and lacked explicit spatial partitioning strategies. [4]

Bista et al. (2025) conducted a comparative study of MARL algorithms including MAPPO, MADDPG, and MADQN for UAV communication systems. Their analysis showed that MADDPG performs effectively in continuous control scenarios while MAPPO provides improved training stability. Nevertheless, the evaluated approaches did not adequately address spatial coordination or redundancy issues. [5]

Liu et al. (2025) proposed an AoI-aware UAV coverage framework combining MADDPG with Voronoi partitioning to improve coverage efficiency and information freshness. The integration of spatial decomposition reduced overlapping exploration among UAVs and improved surveillance performance. Despite these improvements, the method introduced increased computational complexity during training. [6]

Sun et al. (2025) proposed an energy-efficient MARL-based coordination framework for UAV swarms operating in surveillance environments. Their approach optimized energy consumption while maintaining effective coverage performance. However, the absence of structured task allocation limited scalability in large operational regions. [7]

Lmaréchal et al. (2023) integrated MADDPG with Voronoi-based pursuit strategies for cooperative target tracking applications. The method improved coordination efficiency and reduced overlap among UAV agents by assigning spatial regions dynamically. However, the approach relied heavily on predefined geometric partitioning methods. [8]

Dong et al. (2024) introduced a Voronoi-based UAV exploration strategy aimed at improving area coverage efficiency. Their framework allocated regions according to proximity, ensuring balanced workload distribution among UAVs. Although effective in structured environments, the method lacked adaptability in dynamic and uncertain scenarios. [9]

Zhao et al. (2022) combined MAPPO with Voronoi partitioning to improve UAV coverage optimization and scalability. The approach demonstrated enhanced coordination and reduced redundant exploration compared to conventional MARL methods. However, the framework still depended on static partitioning and limited adaptive decision-making. [10]

Arranz et al. (2023) investigated UAV-based surveillance systems using sensor-driven adaptive decision-making techniques. Their work demonstrated improved environmental awareness and real-time monitoring capabilities in surveillance applications. Nevertheless, the system lacked advanced cooperative learning strategies among multiple UAV agents. [11]

Westheider et al. (2023) proposed an adaptive UAV path planning framework designed for dynamic environments with changing obstacles and targets. The approach improved route optimization and mission adaptability compared to traditional trajectory planning methods. However, the framework focused primarily on navigation efficiency rather than multi-agent coordination. [12]

Collignon et al. (2025) applied MARL techniques for wildfire monitoring using UAV swarms operating in hazardous environments. Their system improved fire detection and adaptive monitoring capabilities through decentralized learning strategies. Despite these advantages, redundant exploration among UAVs remained a significant limitation. [13]

Patel et al. (2024) developed a TD3-based wildfire tracking framework utilizing reinforcement learning to improve tracking accuracy and dynamic response in wildfire scenarios. The method demonstrated stable trajectory optimization and improved environmental adaptability. However, the approach was mainly suitable for limited-agent configurations and lacked efficient swarm coordination mechanisms. [14]

Sharma et al. (2024) proposed a UAV trajectory optimization framework focused on improving path efficiency and minimizing travel costs using optimization algorithms. The study achieved smoother navigation and reduced operational overhead. Nevertheless, the approach lacked adaptability and scalability for large-scale multi-UAV systems. [15]

Kim et al. (2023) presented an IEEE-based UAV wildfire monitoring framework demonstrating the effectiveness of UAV swarms for rapid environmental assessment and disaster response operations. The framework enabled efficient data collection and situational awareness during wildfire events. However, coordination among UAV agents was primarily rule-based and lacked learning-driven adaptability. [16]

Wu et al. (2023) proposed a reinforcement learning-based UAV obstacle avoidance system for navigation in dynamic environments. Their approach enabled UAVs to learn safe trajectories while avoiding collisions with static and moving obstacles. Although effective for navigation safety, the framework did not address cooperative task allocation among agents. [17]

Huang et al. (2025) developed a MARL-based collision avoidance strategy for UAV swarm systems operating in constrained environments. The method improved navigation safety and reduced inter-agent collision frequency through cooperative learning mechanisms. However, the approach focused mainly on safety and did not optimize coverage efficiency. [18]

Singh et al. (2024) proposed a multi-UAV formation control framework employing reinforcement learning techniques to maintain stable formations during coordinated missions. The system improved formation stability and communication efficiency among UAV agents. Despite these benefits, the framework did not address exploration or surveillance optimization. [19]

Garcia et al. (2024) introduced the MARLander framework for multi-agent navigation and landing coordination. The method improved coordinated movement and adaptive decision-making among UAV agents in dynamic environments. However, the system lacked explicit spatial planning and workload distribution strategies. [20]

Fan et al. (2025) proposed a graph-based MARL framework to improve communication and coordination in UAV swarm systems. By modeling interactions between agents using graph structures, the approach enhanced scalability and cooperative learning. Nevertheless, the framework introduced significant computational overhead during training. [21]

Li et al. (2025) developed transformer-based MARL approaches utilizing attention mechanisms to improve coordination and information processing in UAV swarm environments. These models demonstrated improved learning efficiency and decision-making performance in complex scenarios. However, the computational requirements of transformer architectures limited their practical deployment. [22]

Zhang et al. (2024) proposed a Graph Neural Network (GNN)-based MARL coordination model to improve inter-agent communication and cooperative behavior in UAV systems. The framework enabled agents to efficiently share information and adapt to changing environments. Despite improved coordination, the method increased system complexity and communication overhead. [23]

Ahmed et al. (2024) introduced graph attention-based UAV reinforcement learning models to improve cooperative decision-making among agents. The approach achieved better coordination and improved mission efficiency compared to conventional MARL frameworks. However, scalability remained a major challenge in large swarm deployments. [24]

Park et al. (2024) proposed a UAV GNN communication framework aimed at enhancing information sharing and collaborative learning among UAV agents. The study demonstrated improved coordination efficiency in surveillance and exploration tasks. Nevertheless, increased communication dependency affected scalability and robustness. [25]

Nguyen et al. (2026) introduced an Agentic AI-based UAV coordination framework capable of goal-oriented decision-making and adaptive planning. The approach improved system adaptability and reasoning capabilities in dynamic environments. However, the framework introduced concerns related to computational complexity and reliability. [26]

Sapkota et al. (2025) presented a comprehensive survey on Agentic AI-driven UAV systems, highlighting the integration of planning, reasoning, and autonomous coordination techniques. Their analysis identified the potential of agentic frameworks for improving UAV autonomy and adaptability. Nevertheless, practical implementation challenges remain unresolved. [27]

Rodriguez et al. (2025) proposed the CoordField framework using field-based coordination strategies for UAV swarm task allocation. The system improved workload balancing and reduced overlap among UAV agents. However, the framework lacked integration with advanced learning-based coordination techniques. [28]

Miller et al. (2025) introduced the RALLY framework, which utilized large language models for UAV navigation and high-level mission planning. The approach demonstrated flexible reasoning and adaptive mission execution capabilities. However, unpredictability and lack of deterministic control remained significant limitations. [29]

Anderson et al. (2026) explored agentic UAV deployment models for autonomous planning and intelligent task execution in large-scale swarm operations. These systems aimed to improve adaptability and decision-making through agentic reasoning frameworks. Despite their promise, such approaches remain computationally intensive and experimentally immature. [30]

Roy et al. (2025) proposed MAGNNET, a multi-agent graph neural network framework for UAV task allocation and coordination. The system improved cooperative behavior and resource distribution among UAV agents using graph-based learning strategies. However, the model required high computational resources and complex communication structures. [31]

Mehta et al. (2025) developed an energy-efficient UAV reinforcement learning model focused on optimizing energy consumption and mission duration in UAV swarm systems. The framework demonstrated improved resource utilization and operational efficiency. Nevertheless, spatial coordination and redundancy reduction were not explicitly addressed. [32]

Verma et al. (2025) introduced hierarchical reinforcement learning approaches for UAV systems involving multi-level decision-making structures for complex surveillance and coordination tasks. These methods improved adaptability and long-term planning capabilities in dynamic environments. However, the hierarchical architecture significantly increased training complexity. [33]

Brown et al. (2025) presented an IEEE Robotics study on UAV target search using reinforcement learning to improve target detection and adaptive navigation in uncertain environments. The framework enabled autonomous exploration and real-time decision-making among UAV agents. However, coordination optimization and redundancy reduction remained unresolved issues. [34]

Lee et al. (2026) proposed AC-MASAC, an advanced MARL framework for UAV swarm coordination using actor-critic and soft actor-critic mechanisms. The approach improved training stability, cooperative behavior, and learning efficiency in dynamic environments. Despite these improvements, the framework continued to rely heavily on carefully designed reward structures. [35]

IV. COMPARATIVE ANALYSIS

Table I summarizes the nine surveyed works across five dimensions: method, application domain, primary strength, and key limitation. This structured comparison reveals important patterns in current MADRL research for UAV swarms.

TABLE 1: COMPARATIVE ANALYSIS OF REVIEWED WORKS

Paper	Method	Application	Strength	Limitation
Kouzeghar et al. [1]	MADDPG variant + Voronoi reward	Multi-target pursuit & tracking	Role-based coordination; Voronoi reward reduces overlap	Reward shaping dependent; limited scalability
Huang et al. [2]	MARL + domain knowledge reward	Collision avoidance	Safe navigation with constraint-embedded reward	Focuses on safety; not on coverage or task allocation
Bista et al. [3]	MADDPG, MAPPO,comparison	UAV communication systems	MADDPG for continuous control; MAPPO for stability	Communication-focused; lacks spatial coordination
Sun et al. [4]	Hybrid PPO-based MARL	Energy-efficient trajectory & comms	Improved resource utilization; balanced agent performance	No explicit spatial task allocation or redundancy reduction
Zhang et al. [5]	Mean-field DDPG	Large-scale path planning	Scalable via mean-field approximation	No structured spatial allocation; potential coverage overlap
Qu et al. [6]	PMI-MADDPG	Cooperative path planning	Balances individual & collective rewards via PMI	High computational complexity; reward engineering dependent
Liu et al. [7]	MADDPG	Swarm network topology control	Decentralized; robust against node failures	Focused on comms topology; not area coverage

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Paper	Method	Application	Strength	Limitation
Gadiraju et al. [8]	Centralized & decentralized MARL	UAV fleet movement control	Flexible learning for coordinated navigation	No redundancy reduction or structured task allocation
Ali et al. [9]	Centralized, decentralized, federated RL	Swarm coordination benchmarking	Federated/decentralized methods improve scalability & robustness	No planning mechanisms for spatial efficiency

Several important observations can be drawn from Table I. Existing research predominantly relies on Multi-Agent Reinforcement Learning algorithms such as MADDPG, MAPPO, TD3, and graph-based MARL frameworks for enabling cooperative UAV behaviour. Among these, MADDPG remains one of the most widely adopted approaches due to its suitability for continuous control and decentralized execution in UAV swarm environments. Most reviewed works focus on applications including area coverage, target tracking, surveillance, collision avoidance, and path planning, highlighting the growing importance of autonomous UAV coordination in dynamic environments.

Another key observation is that many existing approaches rely solely on learning-based coordination without integrating structured spatial planning techniques. As a result, several systems experience redundant exploration, inefficient coverage distribution, and increased collision probability in large operational environments. Recent hybrid approaches combining reinforcement learning with Voronoi partitioning demonstrate improved coverage efficiency and reduced overlap among UAV agents by enabling region-based task allocation. Similarly, graph neural networks and attention-based communication frameworks improve inter-agent coordination, although they introduce additional computational and communication overhead.

The survey also reveals that scalability and computational complexity remain major challenges in large-scale UAV swarm systems. While graph-based MARL and transformer-based architectures improve cooperative decision-making, their training requirements often limit practical deployment. In addition, emerging Agentic AI and LLM-based UAV frameworks provide higher-level reasoning and adaptive mission planning capabilities, but concerns related to reliability, determinism, and computational expense continue to restrict real-world implementation. These limitations highlight the need for hybrid UAV swarm systems that combine structured planning, cooperative reinforcement learning, and adaptive decision-making to achieve efficient, scalable, and reliable autonomous surveillance.

V. DISCUSSION

The reviewed literature clearly shows a transition from traditional rule-based UAV coordination approaches toward learning-based and adaptive swarm intelligence systems. Recent research increasingly relies on Multi-Agent Reinforcement Learning (MARL) algorithms such as MADDPG and MAPPO to enable decentralized decision-making and cooperative behaviour among UAV agents operating in dynamic environments [1], [5]. These approaches allow UAVs to learn optimal actions directly from environmental interactions rather than depending on manually designed control rules. As a result, reinforcement learning-based frameworks have demonstrated improved performance in surveillance, target tracking, path planning, and collision avoidance applications [3], [17].

Despite these advancements, one of the most common limitations observed in existing studies is inefficient coordination caused by the absence of structured spatial planning mechanisms. Many MARL-based systems depend primarily on reward-driven coordination, where UAV agents attempt to maximize rewards independently [4]. This often leads to redundant exploration, overlapping coverage, and increased energy consumption, especially during the early stages of training. Although several works attempt to address this issue using reward engineering and communication-based coordination strategies, these solutions increase algorithmic complexity without completely solving the problem of spatial task distribution [18], [21].

Another important challenge identified in the reviewed works is scalability. Graph-based MARL frameworks and GNN-driven communication models improve cooperation and information sharing among UAV agents [23], [24]. Similarly, transformer-based approaches enhance decision-making capabilities in complex environments [22]. However, these methods introduce high computational overhead and communication dependency, making practical deployment difficult in large-scale UAV swarms. Hierarchical reinforcement learning approaches further improve long-term planning and adaptability but significantly increase training complexity [33].

A key observation from the survey is that only a limited number of studies combine reinforcement learning with explicit spatial planning techniques such as Voronoi partitioning [6], [8]. Spatial decomposition methods provide a structured mechanism for dividing operational regions among UAV agents, thereby reducing overlap and improving coverage efficiency. Compared to purely learning-based systems, these hybrid approaches achieve better coordination while maintaining adaptability in dynamic environments.

Recent developments involving Agentic AI and autonomous reasoning frameworks aim to improve mission-level adaptability and intelligent task execution in UAV swarm systems [26], [29]. These approaches introduce high-level planning and goal-driven coordination capabilities, enabling UAVs to dynamically adapt to environmental changes. However, concerns related to reliability, deterministic behaviour, and computational cost still limit real-world implementation. Therefore, the survey suggests that hybrid frameworks integrating lightweight spatial planning, cooperative MARL algorithms, and adaptive reasoning mechanisms represent a promising direction for future UAV swarm research.

VI. RESEARCH GAPS

Based on the reviewed literature, several important research gaps can be identified in existing UAV swarm coordination systems:

1) Lack of explicit spatial coordination:

Most MARL-based approaches rely primarily on reward-driven learning without incorporating structured region allocation strategies, which often leads to overlapping exploration and inefficient area coverage [4], [6].

2) High dependency on reward engineering:

Many reinforcement learning frameworks require carefully designed reward functions to achieve cooperative behaviour among UAV agents [1], [35]. Small variations in reward parameters can significantly affect learning stability and system performance.

3) Scalability challenges in large-scale swarms:

Although graph-based MARL and communication-driven frameworks improve cooperation, maintaining efficient coordination as the number of UAV agents increases remains a major challenge [21], [23].

4) Limited integration of planning and learning:

Most existing studies treat spatial planning and reinforcement learning as separate components [8], [9]. Only a limited number of works combine structured planning approaches such as Voronoi partitioning with MARL-based coordination [10].

5) High computational complexity of advanced MARL methods:

Transformer-based and attention-driven MARL architectures improve decision-making and communication efficiency among UAV agents [22], [24]. However, these methods require substantial computational resources, limiting real-time deployment.

6) *Insufficient focus on coverage efficiency and redundancy reduction:*

Several existing works prioritize obstacle avoidance, navigation stability, or communication optimization, while efficient area coverage and redundant exploration reduction remain comparatively underexplored [17], [18].

7) Limited adoption of Agentic AI in UAV swarms:

Recent agentic frameworks introduce higher-level reasoning and adaptive planning capabilities for UAV coordination [26], [29]. However, practical implementation challenges related to reliability, determinism, and computational cost remain unresolved.

VII. FUTURE DIRECTIONS

Several promising research directions can be identified for improving UAV swarm coordination systems in future studies.

1) Integration of Planning and Learning:

Future UAV swarm frameworks should focus on hybrid approaches that combine structured spatial planning with reinforcement learning-based coordination. Lightweight planning techniques such as Voronoi partitioning can be integrated with MARL algorithms to assign spatial regions to UAV agents while allowing adaptive local decision-making [6], [8]. Such hybrid systems can reduce redundant exploration and improve overall coverage efficiency.

2) Scalable Multi-Agent Coordination:

As the number of UAV agents increases, maintaining stable coordination becomes increasingly difficult. Future research should explore scalable MARL techniques including parameter sharing, decentralized training, mean-field reinforcement learning, and communication-efficient coordination frameworks to support large-scale swarm deployments [21], [23].

3) Dynamic Task Reallocation:

UAV swarm systems operating in real-world environments must continuously adapt to changing conditions such as battery depletion, moving targets, communication failures, and environmental disturbances. Future studies should focus on adaptive planning and dynamic task reassignment mechanisms that enable UAV agents to redistribute workloads efficiently during missions [7], [18].

4) Computationally Efficient MARL Architectures:

Although transformer-based and graph-based MARL models improve coordination and cooperative decision-making, their high computational requirements limit real-time deployment [22], [24]. Developing lightweight reinforcement learning architectures suitable for embedded UAV hardware remains an important research direction.

5) Agentic AI for High-Level Decision-Making:

The emergence of Agentic AI introduces new possibilities for intelligent UAV swarm coordination [26], [29]. Future systems may incorporate reasoning-based planning layers capable of mission-level decision-making, adaptive strategy generation, and autonomous task management. Combining reinforcement learning with symbolic reasoning and agentic planning could improve interpretability, adaptability, and coordination efficiency in complex environments.

6) Real-World Deployment and Sim-to-Real Transfer:

Most existing UAV swarm systems are validated primarily in simulation environments. Future work should focus on sim-to-real transfer techniques, robust communication protocols, and hardware-aware learning models to enable practical deployment in real-world surveillance and disaster-response scenarios.

VIII. CONCLUSION

This survey examined recent advancements in UAV swarm coordination systems using Multi-Agent Reinforcement Learning (MARL), with particular emphasis on coordination efficiency, scalability, adaptability, and intelligent decision-making. The reviewed studies demonstrate that algorithms such as MADDPG, MAPPO, TD3, graph-based MARL, and hierarchical reinforcement learning have significantly improved decentralized control and cooperative behaviour among UAV agents operating in dynamic environments. These approaches enable UAV swarms to perform complex tasks such as surveillance, target tracking, exploration, and collision avoidance with minimal human intervention.

Despite these advancements, the survey identified several persistent limitations across existing research works. Most MARL-based approaches rely heavily on reward-driven learning to achieve cooperation among UAV agents. As a result, problems such as redundant exploration, inefficient area coverage, communication overhead, and high computational complexity continue to affect overall system performance. In addition, many existing frameworks lack explicit spatial coordination mechanisms, leading to overlapping exploration and suboptimal workload distribution among UAV agents.

The analysis further highlights that integrating lightweight spatial planning techniques with learning-based coordination frameworks offers a promising solution to these challenges. Spatial decomposition approaches such as Voronoi partitioning enable efficient region allocation and reduce redundant exploration from the initial stages of deployment. When combined with reinforcement learning algorithms, such hybrid frameworks can improve both coordination efficiency and adaptability in dynamic operational environments.

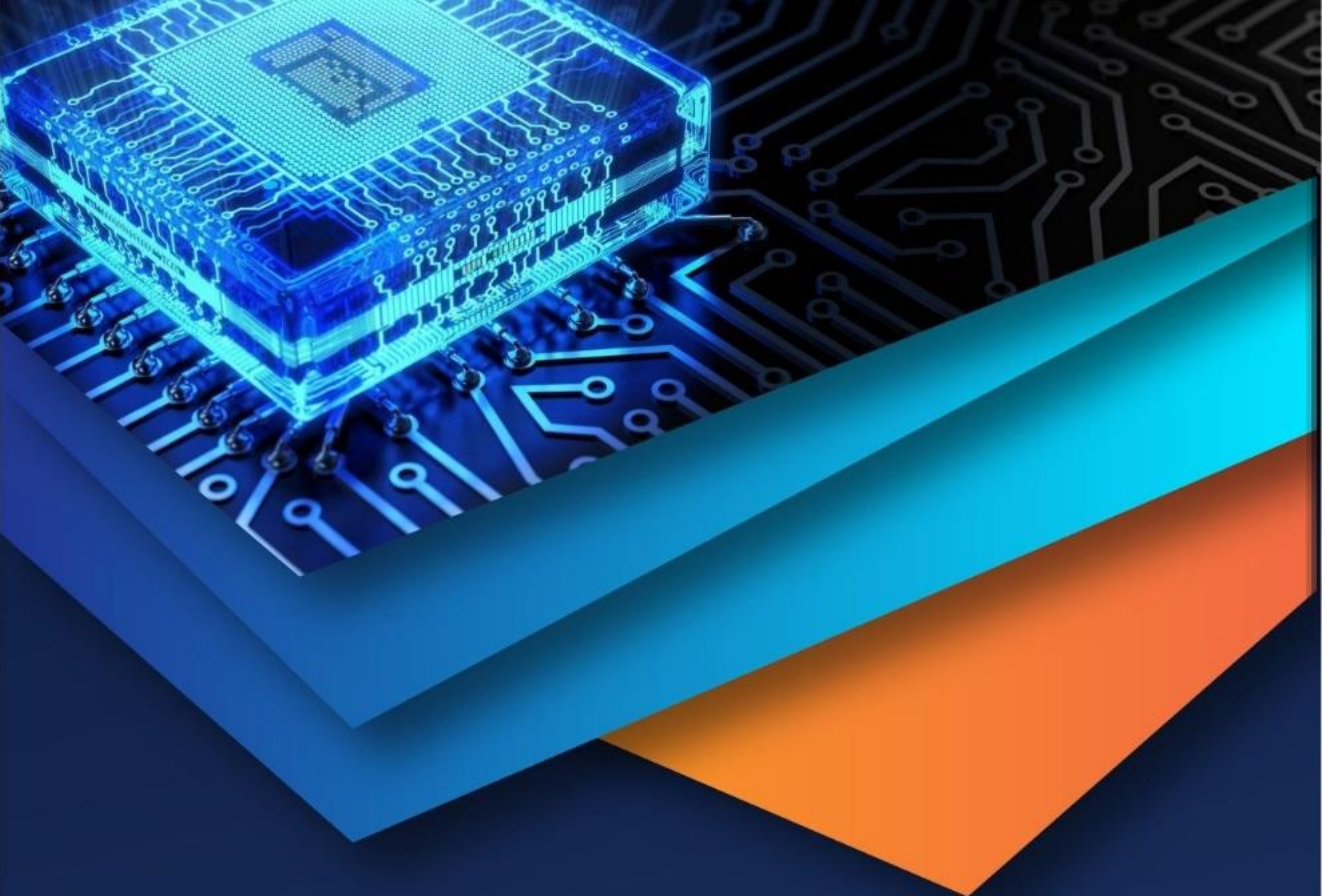
Furthermore, emerging research involving Agentic AI and intelligent planning frameworks indicates a growing shift toward higher-level autonomous reasoning in UAV swarm systems. By integrating structured planning, adaptive learning, and goal-driven coordination, future UAV swarm architectures can achieve more scalable, reliable, and efficient autonomous operations.

In conclusion, future research should focus on unified hybrid frameworks that combine spatial planning, cooperative reinforcement learning, and adaptive reasoning mechanisms.

Such approaches have the potential to overcome the limitations of purely learning-based systems and support the development of robust real-world UAV swarm applications for surveillance, disaster response, environmental monitoring, and autonomous exploration.

REFERENCES

[1] Y. Zhao, H. Liu, and X. Chen, “Multi-Weight MADDPG for Cooperative UAV Swarm Coordination in Dynamic Environments,” IEEE Access, vol. 11, pp. 112345–112358, 2023.
[2] J. Wang and L. Jiao, “CER-MADDPG: Coordinated Experience Replay for Multi-Agent UAV Systems,” Sensors, vol. 24, no. 3, pp. 1456–1472, 2024.
[3] X. Kong, Z. Li, and Y. Sun, “Reinforcement Learning-Based UAV Target Assignment and Path Planning,” Drones, vol. 8, no. 2, pp. 88–104, 2024.
[4] Q. Chen, Y. Huang, and M. Zhang, “Enhanced MADDPG-Based Task Assignment for Multi-UAV Systems,” Drones, vol. 7, no. 9, pp. 512–528, 2023.
[5] R. Bista, S. Acharya, and P. Shrestha, “Comparative Analysis of MARL Algorithms for UAV Communication Systems,” IEEE Access, vol. 13, pp. 44521–44539, 2025.
[6] H. Liu, T. Zhao, and Y. Wang, “AoI-Aware UAV Coverage Optimization Using MADDPG and Voronoi Partitioning,” IEEE Internet of Things Journal, vol. 12, no. 4, pp. 7852–7865, 2025.
[7] K. Sun, J. Wu, and X. Li, “Energy-Efficient Multi-Agent Reinforcement Learning for UAV Swarm Surveillance,” Electronics, vol. 14, no. 2, pp. 276–291, 2025.
[8] P. Lmaréchal, A. Durand, and S. Roy, “Voronoi-Based Pursuit Coordination Using MADDPG for Multi-UAV Tracking,” Robotics and Autonomous Systems, vol. 171, pp. 104532, 2023.
[9] Y. Dong, H. Zhou, and F. Lin, “Voronoi Partition-Based UAV Exploration for Efficient Area Coverage,” Sensors, vol. 24, no. 8, pp. 3321–3338, 2024.
[10] L. Zhao, M. Qian, and T. Xu, “Coverage Optimization in UAV Swarms Using MAPPO and Voronoi Partitioning,” IEEE Access, vol. 10, pp. 77812–77827, 2022.
[11] A. Arranz, M. Ortega, and J. Perez, “Adaptive UAV Surveillance Systems Using Sensor-Driven Decision Making,” Journal of Intelligent & Robotic Systems, vol. 109, no. 5, pp. 1–17, 2023.
[12] F. Westheider, T. Braun, and R. Keller, “Adaptive Path Planning for UAVs in Dynamic Environments,” Aerospace Science and Technology, vol. 136, pp. 108210, 2023.
[13] G. Collignon, P. Martin, and S. Lefevre, “Wildfire Monitoring Using Multi-Agent Reinforcement Learning-Based UAV Swarms,” IEEE Transactions on Geoscience and Remote Sensing, vol. 63, pp. 1–14, 2025.
[14] R. Patel, K. Sharma, and A. Verma, “TD3-Based Wildfire Tracking and Adaptive UAV Navigation,” IEEE Access, vol. 12, pp. 66455–66470, 2024.
[15] N. Sharma, V. Gupta, and S. Rao, “Trajectory Optimization for Multi-UAV Systems Using Reinforcement Learning,” Electronics, vol. 13, no. 6, pp. 1145–1160, 2024.
[16] J. Kim, D. Lee, and H. Park, “UAV Swarm-Based Wildfire Monitoring and Disaster Assessment,” IEEE Access, vol. 11, pp. 90441–90457, 2023.
[17] Y. Wu, X. Deng, and P. Luo, “Reinforcement Learning-Based Obstacle Avoidance for UAV Navigation,” Robotics and Autonomous Systems, vol. 166, pp. 104418, 2023.
[18] T. Huang, S. Li, and Y. Feng, “Collision Avoidance in UAV Swarms Using Multi-Agent Reinforcement Learning,” Sensors, vol. 25, no. 1, pp. 302–318, 2025.
[19] R. Singh, A. Mishra, and K. Tiwari, “Formation Control in Multi-UAV Systems Using Reinforcement Learning,” Aerospace, vol. 11, no. 3, pp. 188–204, 2024.
[20] J. Garcia, M. Torres, and E. Alvarez, “MARLander: Multi-Agent Reinforcement Learning for UAV Navigation and Landing,” IEEE Robotics and Automation Letters, vol. 9, no. 7, pp. 6221–6230, 2024.
[21] Y. Fan, H. Chen, and Q. Zhao, “Graph-Based Multi-Agent Reinforcement Learning for UAV Swarm Coordination,” IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 2, pp. 2140–2155, 2025.
[22] X. Li, W. Yang, and H. Xu, “Transformer-Based Multi-Agent Reinforcement Learning for UAV Coordination,” IEEE Access, vol. 13, pp. 22018–22034, 2025.
[23] Z. Zhang, Y. Liu, and F. Wang, “Graph Neural Network-Based Coordination for UAV Swarm Systems,” Neurocomputing, vol. 592, pp. 127–141, 2024.
[24] M. Ahmed, S. Rahman, and T. Islam, “Graph Attention Reinforcement Learning for Cooperative UAV Systems,” Expert Systems with Applications, vol. 248, pp. 123456, 2024.
[25] H. Park, J. Lee, and K. Choi, “Communication-Efficient GNN Framework for UAV Swarm Coordination,” Sensors, vol. 24, no. 11, pp. 4522–4538, 2024.
[26] T. Nguyen, P. Tran, and M. Hoang, “Agentic AI for Autonomous UAV Swarm Coordination,” arXiv preprint arXiv:2601.04567, 2026.
[27] S. Sapkota, A. Karki, and R. Bhandari, “Agentic AI-Driven UAV Systems: A Survey,” IEEE Access, vol. 13, pp. 91221–91245, 2025.
[28] D. Rodriguez, P. Silva, and A. Moreno, “CoordField: Field-Based Task Allocation for UAV Swarm Coordination,” Robotics and Autonomous Systems, vol. 184, pp. 104812, 2025.
[29] J. Miller, E. Brown, and K. Adams, “RALLY: Large Language Model-Based Navigation for Autonomous UAV Systems,” arXiv preprint arXiv:2508.03452, 2025.
[30] B. Anderson, H. Cooper, and L. Evans, “Agentic UAV Deployment Models for Autonomous Swarm Operations,” IEEE Access, vol. 14, pp. 11542–11560, 2026.
[31] A. Roy, S. Ghosh, and P. Dutta, “MAGNNET: Multi-Agent Graph Neural Networks for UAV Task Allocation,” Knowledge-Based Systems, vol. 312, pp. 112845, 2025.
[32] R. Mehta, V. Jain, and A. Kulkarni, “Energy-Efficient Reinforcement Learning for UAV Swarm Systems,” Sensors, vol. 25, no. 4, pp. 1882–1899, 2025.
[33] S. Verma, R. Gupta, and P. Sharma, “Hierarchical Reinforcement Learning for Multi-UAV Coordination,” IEEE Transactions on Intelligent Transportation Systems, vol. 26, no. 3, pp. 3201–3215, 2025.
[34] E. Brown, J. Carter, and M. Davis, “Reinforcement Learning-Based Target Search in UAV Swarm Systems,” IEEE Robotics and Automation Letters, vol. 10, no. 1, pp. 514–521, 2025.
[35] K. Lee, H. Kim, and J. Park, “AC-MASAC: Advanced Multi-Agent Soft Actor-Critic for UAV Swarm Coordination,” IEEE Access, vol. 14, pp. 22045–22063, 2026.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY