BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

AI-Driven Multi-Drone Swarms for Defence Surveillance and Reconnaissance: A Literature Study and Conceptual Design Direction

[1]Sakshi Sunil Pawar, [2]Prof. Varsha Kolhatkar

[1]Student, [2]Professor
Department of Artificial Intelligence and Data Science Engineering
Ajeenkya DY Patil School of Engineering, Pune, India

Abstract

With a single unmanned aerial vehicle (UAV) monitoring just a small area, once that one gets jammed, damaged or shot down, the surveillance mission ends with it. When the work is distributed across multiple drones, this is different: when one drone fails, the mission continues. This paper examines eight published papers from the last five years (2021–2025) that explore how such groups can be made to cooperate, and the transition from hand-designed formation and communication rules to coordination learned using Multi-Agent Reinforcement Learning (MARL). When studied together, the studies demonstrate that learned coordination performs well in simulation for pursuit, task allocation and layered command, but very little experimental work has been done with large physical swarms, degraded communications or missions beyond a single scenario. To illustrate a conceptual direction for design, we sketch a four-layer design direction (perception, shared awareness, learned coordination, and resilience with evaluation) and provide a list of measurements that would be reported in a subsequent experimental study. This paper is an analysis and design agenda based on literature; no swarm was constructed, flown or simulated.

Keywords: Drone swarm, multi-agent reinforcement learning, UAV, defence surveillance, reconnaissance, distributed coordination, situational awareness, autonomous systems.

Introduction

Armed forces rely on aerial platforms to gather intelligence in areas that would be expensive and risky to patrol on foot. A single UAV is useful but limited: it has a small footprint and limited battery life, and it can be jammed, can malfunction, or can be attacked by an adversary. Once it is lost, no one is watching.

A swarm reorganises the problem. Instead of one aircraft responsible for everything, several UAVs share area scanning, target following and data relay. Their coverage overlaps to provide redundancy, so the mission does not depend on any single vehicle. The challenge is that a fleet of drones is not necessarily a swarm: getting the drones to cooperate is itself a control and learning problem.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Early answers were rule-based, defining what each drone had to do to maintain formation and to communicate with the others. With the advent of reinforcement learning, researchers began handing the development of cooperative behaviour to the drones themselves, distributing work among multiple targets, performing multiple tasks concurrently and reacting to contexts that could not be described beforehand [1–8]. Shared sensing and computer vision are further additions that let each drone contribute to a "big picture" that helps people make faster decisions.

The contributions of this paper are:

1. an organised reading of eight recent studies, grouped according to the coordination approach each employs (Related Work);
2. five gaps between current achievements and defence surveillance requirements (Research Gaps); and
3. a Conceptual Design Direction (CDD) of four layers that address these gaps, with an evaluation plan for future work (Conceptual Design Direction and Proposed Evaluation Plan).

The scope is purposely limited to literature and concept. The studies are not intended to be comprehensive, but to represent the evolution of the field and its open problems.

Problem Statement and Objectives

A. Problem Statement

Using isolated drones for defence surveillance has three implications: the observed area remains very limited, there is no backup if a unit is lost or corrupted, and there is no way to coordinate what each unit observes or does. Together these reduce reliance on real-time situational awareness of the target area during reconnaissance over extensive or contested regions, making it advantageous to use multiple drones coordinated by a learning approach.

B. Objectives

- To follow the evolution of coordination in AI-powered UAV swarms from rule-based formation control to multi-agent reinforcement learning.
- To explore how MARL enables members of a swarm to share tasks, monitor targets and chase them together.
- To evaluate the capability of computer vision and shared sensing to augment wide-area surveillance and real-time awareness.
- To recognise the shortcomings of current swarm research in terms of coordination, communication and reliability.
- To describe the vision for an approach to defence surveillance based on learning swarms.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Related Work

The eight studies fall into four groups, depending on which part of the swarm problem they tackle. The grouping and the stated weaknesses are this paper's own reading; the original authors may frame their work differently.

A. Hand-Designed Coordination and Resilient Management

The proposal in [7] by Wubben et al. is a swarm management scheme designed to withstand change. It maintains formation, coordination and communication even when members leave, and enables the group to rearrange itself. Such a scheme is hard to carry over to large real-world swarms with high autonomy. Brown and Raj [8] propose a guidance law for a surveillance swarm to follow a faster, capable hostile UAV, keeping the drones positioned so that they never lose sight of the target. The result has been validated primarily through simulation and assumes a working network.

B. Learning-Based Coordination with MARL

To track a target, Xia et al. [5] allowed the UAVs to share observations during tracking and learn by multi-agent reinforcement learning. The group performed more cohesively than individual drones, although results depend on training factors and environmental assumptions. Kouzeghar et al. [3] present a decentralised MADDPG, where UAVs of diverse roles chase multiple targets without any

central control; the work includes simulations and a few physical flight tests. Learning is structured in layers in [6] and [1]. In [6], hierarchical MARL is used to make joint decisions for a suppression-of-enemy-air-defence (SEAD) mission, but the method is restricted to the SEAD mission and requires extensive training time. In [1], top-level agents assign tasks to individual drones, which then perform them independently, improving cooperative exploration; validation on larger, defence-scale swarms is still awaited.

C. Training Platforms and Scaling

Lou et al. [4] present Air-M, a many-agent RL environment that allows training of large aerial swarms and provides a transfer path for moving learned behaviour to real vehicles. The quality of the transfer depends on how well the simulated models and test conditions represent reality.

D. Surveys

Javed et al. [2] provide an overview of recent progress in UAV swarms, including formation control, communication, path planning, autonomy, coordination and security. They find that scalability, security, autonomy and system integration remain unresolved.

E. Observed Progression

Arranged chronologically, the studies clearly illustrate a progression from rule-based coordination to swarm management, to single-agent reinforcement learning, to multi-agent reinforcement learning, and finally to hierarchical and scalable swarm intelligence. Earlier studies focused on formation, inter-drone communication, tracking and survivability after losses [7, 8]. Later studies use deep and multi-agent RL for pursuit, allocation and exploration [1, 3, 5, 6]. Current topics are reliability of communication, autonomous decision making, learning at large scale, and simulation-to-real transfer [1, 2, 4].

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

TABLE 1

SUMMARY OF THE REVIEWED STUDIES

Ref. (Year)	Approach	Main Contribution	Weakness Noted
[1] Salt (2025)	Hierarchical Agent Deep RL	Upper-level task assignment with autonomous lower-level actions	Needs validation at defence scale
[2] Javed (2024)	State-of-the-Art Review	Survey of swarm control, networking, and security	Scalability, security, and integration remain open challenges
[3] Kouzeghar (2023)	Decentralised MADDPG	Multi-target pursuit with mixed UAV roles	Mostly simulation-based, with few physical trials
[4] Lou (2023)	Many-Agent RL Platform (Air-M)	Large-swarm training with a real-world transfer pathway	Transfer depends on model realism
[5] Xia (2022)	MARL for Target Tracking	Shared observations enable coherent tracking	Sensitive to training assumptions
[6] Yue (2022)	Hierarchical Multi-Agent RL	Layered joint decisions in SEAD missions	Single scenario and computationally intensive training
[7] Wubben (2021)	Resilient, Reconfigurable Management	Operations continue despite the loss of individual members	Large-scale real-world deployments remain challenging
[8] Brown (2021)	Surveillance Guidance Law	Continuous tracking of a faster target	Simulation-based; requires reliable networking

Research Gaps

Comparing the studies produces five recurring shortcomings.

G1. Simulation dominates. Most coordination methods are shown inside simulators. Trials with large physical swarms under defence-like conditions are rare [1, 3, 4, 8].

G2. Scale worsens known problems. Scalability, security and system integration appear again and again as unresolved, and become harder as swarm size grows [2, 7].

G3. Single-mission focus. Work built around one mission type, such as pursuit or SEAD, does not show that the strategy carries over to general surveillance and reconnaissance [3, 6].

G4. Degraded communication is neglected. Methods that rely on a dependable network [5, 8] say little about operating over jammed or intermittent links, a normal condition in defence use.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

G5. Fragmented frameworks. Few studies combine vision-based sensing, multi-agent learning and defence-grade reliability in one design.

Taken together, no reviewed work offers a single framework in which a swarm senses through vision, learns to coordinate, and keeps working under loss and interference at realistic scale.

Conceptual Design Direction

This section describes what a more complete framework would need to contain. It is a proposal on paper, derived from the gaps above, and is not an implemented system.

A. Layers

1. **Perception layer.** Each drone uses computer vision algorithms to analyse camera and sensor data, and to identify and locate objects of interest, generating a small local observation rather than a large video feed.
2. **Shared-awareness layer.** Local observations are shared with neighbouring drones and integrated into a shared operating picture. The layer must withstand delay, packet loss and temporary isolation of members (G4).
3. **Lear ned-coordination layer.** A hierarchical MARL scheme, derived from and inspired by [1, 3, 6], assigns tasks at a higher level and lets the drones perform them independently at a lower level. Training for multiple mission types is designed to address G3.
4. **Resilience and evaluation layer.** This layer watches swarm health, triggers reconfiguration when members fail [7], and scores behaviour against the measures in the Proposed Evaluation Plan, including tests with degraded links and reduced swarm size.

B. Mapping Gaps to Layers

Table II pairs each gap from the Research Gaps section with the design response and the layer that carries it.

C. Information Flow

In one operating cycle, each drone maps its sensor input to a “local observation”. Adjacent drones share observations and keep the common picture up to date. Low-level actions follow, with each drone picking its action from the tasks selected for it by the coordination layer. The resilience layer tracks the results of the coverage (what was achieved) and the losses of members (who was lost), and feeds these into the next cycle for re-assignment of tasks. Sending compact observations rather than raw imagery also decreases bandwidth requirements, which is important in low-bandwidth environments.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

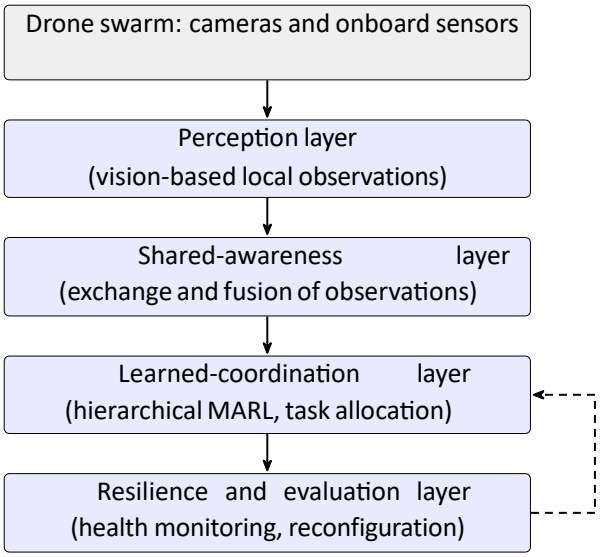

Figure 1. Proposed four-layer conceptual framework. Outcomes of each cycle feed back to the coordination layer.

TABLE 2

HOW EACH GAP IS ADDRESSED

Gap	Design Response	Layer
G1: Simulation-Dominated Research	A staged validation plan involving simulation, small-scale physical trials, and progressively larger swarms	Resilience and Evaluation
G2: Scalability Limitations	Hierarchical task assignment, with each drone reasoning within a limited neighbourhood	Learned Coordination
G3: Single-Mission Focus	Training and testing across multiple mission types and operational scenarios	Learned Coordination
G4: Degraded Communication Links	Local-first decision-making and information fusion tolerant of missing messages	Shared Awareness
G5: Fragmented Frameworks	Integrating perception, awareness, learning, and resilience into a unified pipeline	All Layers

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Proposed Evaluation Plan

Because nothing has been built, Table III lists the questions a future experimental study should answer and the measurements that would answer them. These measures are suggestions made by this paper, not figures reported in the reviewed studies.

TABLE 3

QUESTIONS AND MEASURES FOR FUTURE STUDY

Evaluation Question	Suggested Measure
Does the swarm cover more ground than a single UAV?	Fraction of the target area observed within a fixed time
Is tracking continuous?	Proportion of time a target remains under observation
Does the mission survive losses?	Mission success rate as swarm members are progressively removed
Does it work under poor communication links?	Performance under increasing packet loss and communication delay
Does it scale?	Performance and computational cost as swarm size increases
Does simulation transfer to real-world conditions?	Performance gap between simulation and physical trial results
Is the approach generalisable?	Performance evaluation across multiple mission types

Operational and Ethical Considerations

For defence swarms, technical solutions alone are not enough. First, the system should not replace decision makers but assist them: it is intended as a surveillance and situational-awareness tool, and decision makers remain in control of any action that has consequences. Second, security is an ongoing open issue in the literature [2], and a swarm that learns and communicates should be secured against three types of attack: spoofing, interception and takeover. Third, learned behaviour can break down in unfamiliar situations; conservative testing is therefore advocated, along with clear fallback behaviour and honest reporting of failure cases as well as average results. Lastly, appropriate permissions for airspace and safety would be required for any physical trial.

Conclusion

All the studies discussed here illustrate a shift from prescribed coordination rules to learned cooperation. Each of resilience, pursuit, task allocation and layered decision making has been shown, typically in simulation. The literature lacks a single design that combines vision-based sensing, multi-agent learning and reliable operation under loss and jamming at a realistic scale. The four-layer direction suggested above is a way of structuring that missing combination, and the evaluation plan suggests how it might be tested. The natural next step is a prototype, tested first in simulation and then in realistic surveillance conditions. At present this paper is only a conceptual analysis; the proposed framework has not been implemented or executed.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

References

[1] N. M. Salt, F. Arvin, and J. Hu, "Hierarchical multi-agent deep reinforcement learning for cooperative exploration of UAV swarms," in *Proc. 2025 IEEE 31st Int. Conf. Parallel and Distributed Systems (ICPADS)*, 2025, pp. 1–8, doi: 10.1109/ICPADS67057.2025.11323196.

[2] S. Javed, A. Hassan Akbar, R. Ahmad, W. Ahmed, R. Ahmed, A. Saadat, and M. Guizani, "State-of-the-art and future research challenges in UAV swarms," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19023–19045, 2024, doi: 10.1109/JIOT.2024.3364230.

[3] M. Kouzeghar, Y. Song, M. Meghjani, and R. Bouffanais, "Multitarget pursuit by a decentralized heterogeneous UAV swarm using deep multi-agent reinforcement learning," in *Proc. 2023 IEEE Int. Conf. Robotics and Automation (ICRA)*, 2023, pp. 3289–3295, doi: 10.1109/ICRA48891.2023.10160919.

[4] J. Lou, W. Wu, S. Liao, and R. Shi, "Air-M: A visual reality many-agent reinforcement learning platform for large-scale aerial unmanned system," in *Proc. 2023 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2023, pp. 5598–5605, doi: 10.1109/IROS55552.2023.10341405.

[5] Z. Xia, J. Du, J. Wang, C. Jiang, Y. Ren, G. Li, and Z. Han, "Multi-agent reinforcement learning aided intelligent UAV swarm for target tracking," *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 931–945, 2022, doi: 10.1109/TVT.2021.3129504.

[6] L. Yue, R. Yang, J. Zuo, Y. Zhang, Q. Li, and Y. Zhang, "Unmanned aerial vehicle swarm cooperative decision-making for SEAD mission: A hierarchical multiagent reinforcement learning approach," *IEEE Access*, vol. 10, pp. 92177–92191, 2022, doi: 10.1109/ACCESS.2022.3202938.

[7] J. Wubben, F. Fabra, C. T. Calafate, J.-C. Cano, and P. Manzoni, "A novel resilient and reconfigurable swarm management scheme," *Computer Networks*, vol. 194, p. 108119, 2021, doi: 10.1016/j.comnet.2021.108119.

[8] J. Brown and N. Raj, "Guidance law for a surveillance UAV swarm tracking a high capability malicious UAV," in *Proc. 2021 IEEE Asia Pacific Conf. Wireless and Mobile (APWiMob)*, 2021, doi: 10.1109/APWiMob51111.2021.9435240.