ICON Journal of Engineering Applications of Artificial Intelligence
Volume 02 | Issue 05 | 2026
e-ISSN: 3051-3820
p-ISSN: 3051-3812
Journal homepage: https://iconpublishers.com/icon-j-eng-app-ai/

Research Article **DOI: 10.5281/zenodo.20445206**

Human-Centred and Explainable AI Framework for Operational Deployment of Transformer-MARL Distributed Air-Defence Systems in Contested Electromagnetic Environments

*Abubakar Surajo Imam[1], Aliyu Musa[2], Aminu Abdullahi Umar[3], Muhammad Ahmad Baballe[4]
[1,3,4]Department of Mechatronic Engineering, Nigerian Defence Academy, Kaduna, Nigeria.
[2] Department of Mechanical Engineering, Nigerian Defence Academy, Kaduna, Nigeria.

Corresponding author: Abubakar Surajo Imam
Department of Mechatronic Engineering, Nigerian Defence Academy, Kaduna, Nigeria.

Received Date: 10 April 2026
Published Date: 29 May 2026

Abstract
The increasing proliferation of low-altitude unmanned aerial systems (UASs), autonomous drone swarms, and contested-spectrum operations has created significant demand for resilient distributed air-defence architectures capable of operating under GNSS denial, communication degradation, and adversarial interference. Conventional centralized radar-centric systems remain vulnerable to communication bottlenecks, limited scalability, and single-point failure during distributed swarm attacks. This paper presents a human-centred and Explainable Artificial Intelligence (XAI)-enabled distributed air-defence framework integrating transformer-based multi-modal fusion, cooperative Multi-Agent Reinforcement Learning (MARL), adversarial robustness screening, energy-efficient edge intelligence, and adaptive ISR coordination for resilient counter-UAS operations. The proposed framework combines electro-optical, infrared, passive RF, acoustic, terrain-context, and edge sensing within a unified transformer-fusion architecture incorporating explainable decision support and operator-guided ISR reassignment. A comprehensive validation framework involving Monte Carlo simulation, ROS2-enabled distributed coordination, AirSim/Gazebo implementation, SDR-assisted RF-jamming emulation, telemetry replay validation using the HATSABIBI-26A UAV platform, hardware-in-the-loop experimentation, and Jetson Xavier edge-inference benchmarking was developed to evaluate operational feasibility under contested-spectrum conditions. Experimental evaluation involving 100,000 operational trials demonstrated superior performance compared with centralized radar-centric, Bayesian-fusion, CNN-based, LSTM-based, MADDPG, and QMIX architectures. The framework achieved detection probability of 0.92, false-alarm probability of 0.06, mean latency of 31 ms, and 63% communication-overhead reduction while maintaining reliable operation under severe node degradation and contested electromagnetic environments.

Keywords: Distributed Air Defence; Transformer Fusion; MARL; Explainable AI; Counter-UAS; Edge Intelligence; Contested Spectrum; ROS2; SDR Jamming; HATSABIBI-26A UAV.

I. INTRODUCTION

The rapid proliferation of low-cost unmanned aerial systems (UASs), autonomous drone swarms, and low-altitude aerial threats has significantly transformed the operational landscape of modern air-defence systems [1]–[5]. Contemporary threat environments increasingly involve highly manoeuvrable, low-radar-cross-section aerial platforms capable of operating under GNSS-denied and contested electromagnetic conditions while exploiting infrastructure limitations, communication vulnerabilities, and sensing blind spots [6]–[10]. These operational realities have exposed critical limitations in conventional centralized radar-centric surveillance architectures, particularly in infrastructure-constrained and asymmetric operational theatres where persistent low-altitude observability, rapid cue propagation, adaptive coordination, and resilient decision-making are essential. Traditional air-defence systems primarily rely on centralized sensing infrastructures and

deterministic engagement pipelines. Although such systems provide strong long-range detection capability under conventional operational conditions, they often suffer from several critical limitations, including single-point failure vulnerability, high communication dependency, reduced scalability, and degraded operational resilience under electronic warfare, adversarial spoofing, and distributed low-altitude swarm attacks [11]–[14]. Furthermore, centralized architectures frequently experience elevated latency and communication bottlenecks when supporting persistent distributed surveillance across large operational corridors [15], [16].

Recent advances in artificial intelligence, transformer architectures, distributed sensor fusion, and multi-agent reinforcement learning have introduced new opportunities for resilient distributed sensing and adaptive decision coordination [17]–[23]. Transformer-based architectures have demonstrated strong capability for learning long-range dependencies and cross-modal relationships across heterogeneous feature streams, thereby improving multi-modal target representation and discrimination performance. Similarly, cooperative Multi-Agent Reinforcement Learning (MARL) has emerged as a promising framework for adaptive distributed coordination among heterogeneous sensing agents operating under uncertain and partially observable environments [29]–[36]. Despite these advances, existing AI-enabled counter-UAS frameworks continue to face several major challenges. First, many architectures remain vulnerable to adversarial perturbations, spoofing attacks, and distribution shift, thereby reducing operational trustworthiness and sensing reliability [24]–[28]. Second, many AI-enabled systems lack explainability and human-supervised transparency, thereby limiting operator trust and decision accountability within operational defence environments. Third, distributed sensing systems deployed in contested electromagnetic environments must operate under severe energy, communication, and synchronization constraints, requiring adaptive edge intelligence and low-power distributed coordination mechanisms.

This paper therefore presents a human-centred and explainable AI framework for operational deployment of transformer-MARL distributed air-defence systems within contested electromagnetic environments. The proposed framework extends previous distributed transformer-MARL architectures by integrating human-machine teaming architectures for supervised operational control, explainable AI mechanisms for interpretable decision support, adversarial training and robustness screening, cooperative swarm-defence coordination, energy-efficient edge intelligence, operational deployment workflows involving distributed sensing corridors and airborne ISR confirmation platforms such as the HATSABIBI-26A UAV. The major contributions of this paper include:

- Development of an explainable transformer-MARL distributed air-defence framework suitable for operational deployment.
- Integration of human-supervised decision support and explainable AI reasoning.
- Development of adversarially robust distributed sensing and cue validation mechanisms.
- Formulation of cooperative swarm-defence coordination strategies.
- Integration of energy-aware edge-level sensing and adaptive communication scheduling.
- Development of experimental deployment workflows involving distributed sensing corridors and airborne ISR confirmation platforms.
- Large-scale statistical evaluation under representative contested-spectrum operational conditions using 100,000 Monte Carlo trials.

The remainder of this paper is organized as follows; Section 2 presents the overall system architecture. Section 3 discusses explainable transformer-based multi-modal fusion. Section 4 presents adversarial robustness and human-machine teaming mechanisms. Section 5 introduces cooperative swarm-defence MARL coordination. Section 6 discusses energy-efficient edge intelligence and adaptive task allocation. Section 7 presents contested-spectrum operational modelling. Section 8 discusses simulation and experimental deployment frameworks. Section 9 presents statistical evaluation and robustness analysis. Section 10 concludes the paper and outlines future research directions.

2. System Architecture Overview

The proposed human-centred and explainable AI-resilient distributed air-defence architecture comprises six functional layers: distributed sensing, transformer-based multi-modal fusion, adversarial robustness and Explainable AI (XAI), cooperative Multi-Agent Reinforcement Learning (MARL), energy-efficient edge intelligence, and command/ISR coordination. The framework integrates heterogeneous sensing nodes, transformer fusion, adversarial screening, explainable decision support, cooperative MARL coordination, autonomous edge-level task allocation, and ISR-supported operational control within a unified distributed defence architecture.

The architecture is designed for resilient counter-UAS surveillance within contested electromagnetic environments involving degraded communication, GNSS denial, node failure, spectrum congestion, and adversarial sensing conditions. Unlike centralized radar-centric systems, the framework distributes sensing, cue validation, and decision-making across cooperative sensing agents and edge-fusion gateways, thereby improving scalability, resilience, and graceful degradation under operational stress [4]–[16]. The major operational layers are summarized in Table 1.

Table 1: Major Functional Layers of the Proposed Architecture

Layer	Primary Function	Operational Contribution
Distributed sensing layer	EO/IR/RF/acoustic sensing	Persistent observability
Transformer fusion layer	Cross-modal fusion	Improved target discrimination
Adversarial robustness/XAI layer	Perturbation defence	Trustworthy AI sensing
Cooperative MARL layer	Distributed coordination	Adaptive ISR support
Edge-intelligence layer	Autonomous task allocation	Low-latency response
Command/ISR layer	Human-supervised control	Mission coordination

The distributed sensing topology is represented using a graph-theoretic framework:

$$G = (V, E) \tag{1}$$

where V represents sensing nodes and Erepresents communication links between neighbouring nodes.
The observation vector at sensing node $i$becomes:

$$Z_i = \{Z_i^{EO}, Z_i^{IR}, Z_i^{RF}, Z_i^{AC}, Z_i^{T}\} \tag{2}$$

where EO, IR, RF, AC, and T denote electro-optical, infrared, passive RF, acoustic, and terrain-context observations respectively. The cumulative distributed detection probability is expressed as:

$$P_D = 1 - \prod_{i=1}^{N}(1 - P_i) \tag{3}$$

where P_i is the local detection probability of sensing node i and N is the number of active sensing nodes. The distributed sensing layer provides persistent low-altitude observability using heterogeneous sensing modalities. Transformer-based fusion improves cross-modal feature representation and target discrimination across EO, IR, passive RF, acoustic and terrain-context sensing streams [17]–[23]. The adversarial robustness and XAI layer improves operational trustworthiness by identifying unstable or spoofed cues before propagation to higher decision layers [24]–[28]. Furthermore, the cooperative MARL layer enables adaptive cue validation, ISR retasking, and resource prioritization under uncertain operational conditions [29]–[36]. The edge-intelligence layer supports low-latency autonomous sensor retasking and adaptive communication scheduling, while the command and ISR coordination layer provides human-supervised operational oversight through explainable AI reasoning, confidence estimation, ISR visualization, and operator-guided decision support.

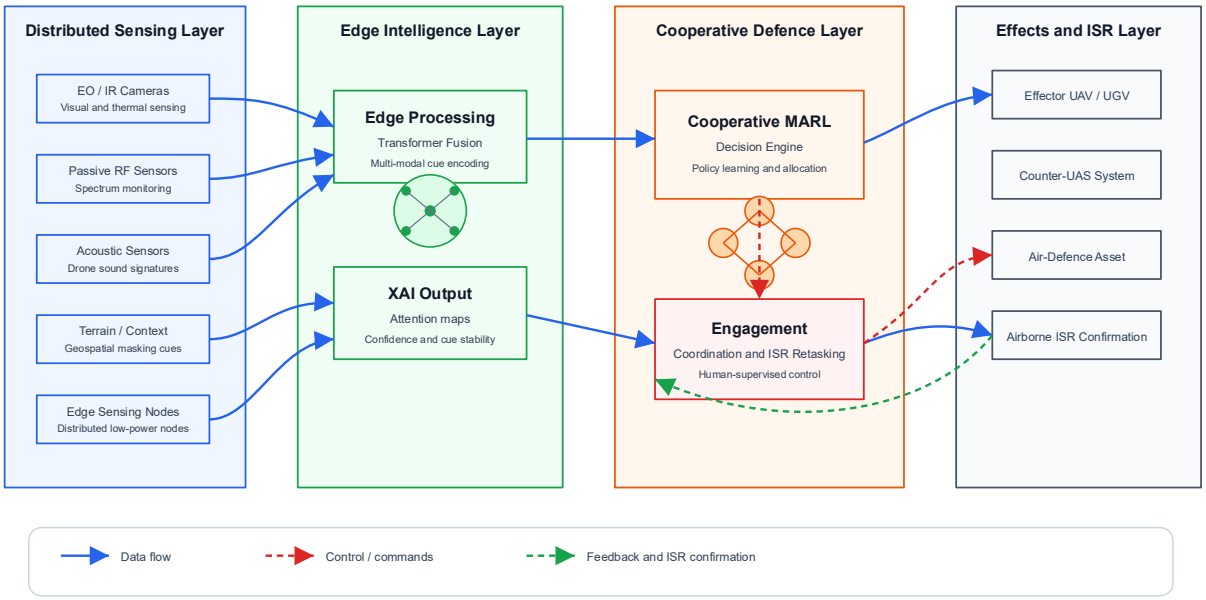

Fig. 1: Simplified AI-Resilient Distributed Air-Defence Architecture Integrating Multi-Modal Sensing, Transformer Fusion, Cooperative MARL Coordination and Edge Intelligence.

The framework is particularly applicable to airport protection, border surveillance, critical-infrastructure defence, humanitarian corridor monitoring, smart-city security, and contested-spectrum defence operations.

3. Explainable Transformer-Based Multi-Modal Fusion

Transformer architectures provide an effective foundation for distributed multi-modal sensing because self-attention mechanisms can learn long-range dependencies and cross-modal relationships across heterogeneous feature streams [17]–[23]. In the proposed framework, electro-optical (EO), infrared (IR), passive RF, acoustic, and terrain-context observations are projected into a shared latent embedding space before transformer-based fusion and explainable reasoning. The modality-specific feature embedding is expressed as:

$$F_m = \phi_m(Z_m) \tag{4}$$

where Z_m represents the observation associated with sensing modality m, and $\phi_m(\cdot)$denotes the modality-specific feature encoder. The concatenated multi-modal feature representation becomes:

$$F = [F_{EO}, F_{IR}, F_{RF}, F_{AC}, F_T] \tag{5}$$

The scaled dot-product attention mechanism is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{6}$$

where Q, K, and V are query, key, and value matrices respectively, while d_k represents the key-vector dimension. The multi-head attention operation becomes:

$$\text{MHA}(F) = \text{Concat}(h_1, h_2, \dots, h_H)W^O \tag{7}$$

where h_j denotes the output of attention head j, H represents the number of attention heads and W^O is the output projection matrix. The fused transformer representation is expressed as:

$$F_{fused} = \text{Transformer}(F) \tag{8}$$

The final target probability estimation becomes:

$$P(T \mid Z) = \sigma(WF_{fused} + b) \tag{9}$$

where $\sigma(\cdot)$ represents the sigmoid activation function, $W$is the classifier weight matrix, and $b$denotes the bias vector. To improve operational transparency and human trustworthiness, the framework incorporates explainable AI (XAI) mechanisms through attention-weight visualization and confidence decomposition. The attention-contribution score associated with sensing modality $m$becomes:

$$A_m = \frac{1}{H}\sum_{h=1}^{H} \text{Attention}_h\,(F_m) \tag{10}$$

where A_m represents the explainability contribution score associated with sensing modality m. These explainability mechanisms enable operators to identify which sensing modalities contribute most strongly to target classification and cue-prioritization decisions, thereby improving operational trustworthiness, transparency, and mission accountability within contested defence environments [24]–[28]. The major transformer-fusion variables are summarized in Table 2.

Table 2: Transformer Fusion Variables

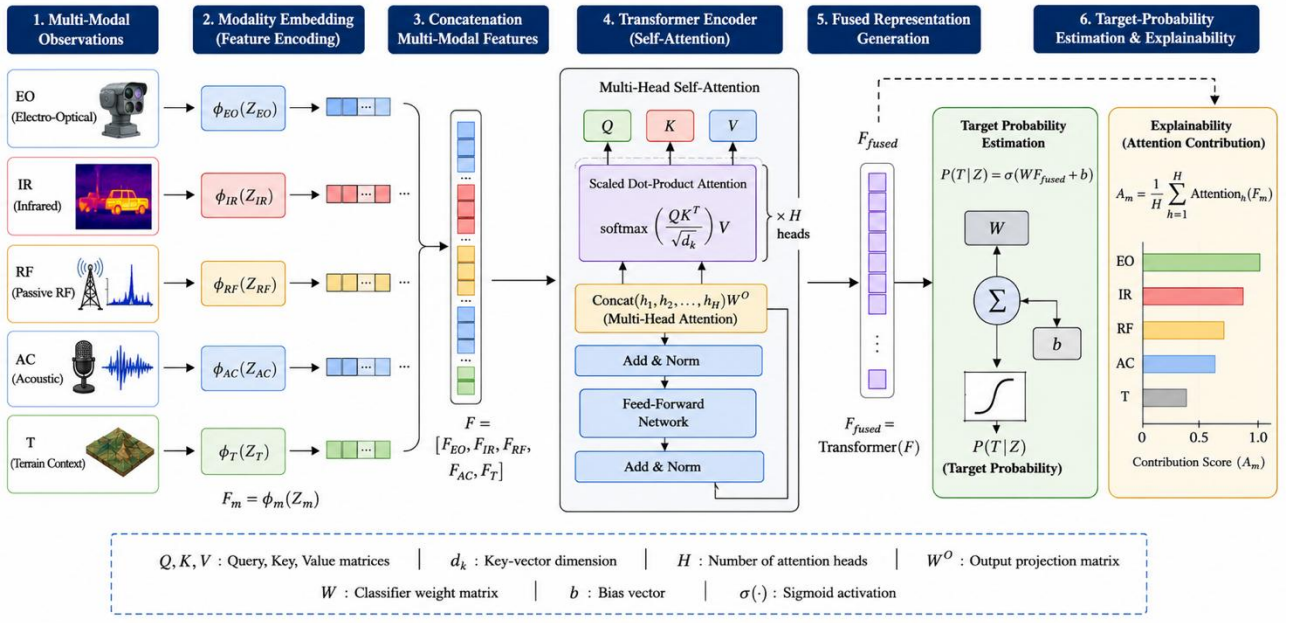

Fig. 2: Simplified Transformer-Based Multi-Modal Fusion Architecture Showing Modality Embedding, Self-Attention, Cross-Modal Fusion, Fused Representation Generation, and Target-Probability Estimation.

4. Adversarial-AI-Resilient and Human-Machine Teaming Framework

AI-enabled sensing systems operating within contested electromagnetic environments are highly vulnerable to adversarial perturbations, spoofing attacks, communication degradation, environmental corruption, and distribution shift. Within distributed counter-UAS architectures, such vulnerabilities may significantly degrade target classification, cue ranking, cooperative coordination, and operational trustworthiness. To address these challenges, the proposed framework incorporates adversarial robustness screening and human-supervised explainable decision support prior to final cue propagation. Let the clean sensing feature vector be represented by x, while δ denotes adversarial or environmental perturbation. The perturbed observation becomes:

$$x' = x + \delta \tag{11}$$

The perturbation constraint is defined as:

$$\| \delta \|_p \leq \epsilon \tag{12}$$

where: ϵ represents the perturbation bound under norm p.The adversarially robust learning objective is expressed as:

$$\min_{\theta} \ \mathbb{E}_{(x,y)} \left[\max_{\|\delta\|_p \leq \epsilon} L(f_\theta(x + \delta), y) \right] \tag{13}$$

where: $L(\cdot)$ represents the loss function, f_θ denotes the AI model parameterized by θ and y represents the true target class label. The adversarial confidence degradation index becomes:

$$D_{adv} = | P(T \mid x) - P(T \mid x') | \tag{14}$$

where D_{adv} measures prediction sensitivity to perturbation. A sensing cue is rejected or downgraded whenever:

$$D_{adv} > \tau_{adv} \tag{15}$$

where τ_{adv} represents the adversarial sensitivity threshold used for robustness filtering. The proposed framework additionally integrates human-machine teaming mechanisms to improve operational trustworthiness, mission accountability, and supervisory control. Human operators remain within the operational loop through explainable cue visualization, confidence decomposition, and operator-guided ISR retasking. The operator-adjusted confidence mechanism becomes:

$$C_{final} = \alpha C_{AI} + (1 - \alpha) C_{human} \tag{16}$$

where: C_{AI} represents AI-estimated confidence, C_{human} represents operator-adjusted confidence, and αcontrols the human-machine weighting balance. The explainable confidence decomposition is expressed as:

$$C_{AI} = \sum_{m=1}^{M} w_m A_m \tag{17}$$

where: w_m represents the weighting associated with sensing modality m and A_m is the explainability attention score for modality m. This explainability framework enables operators to identify which sensing modalities contribute most strongly to target classification and cue-prioritization decisions, thereby improving operational transparency, trustworthiness, and engagement accountability within contested defence environments. The adversarial robustness workflow is illustrated in Fig. 3.

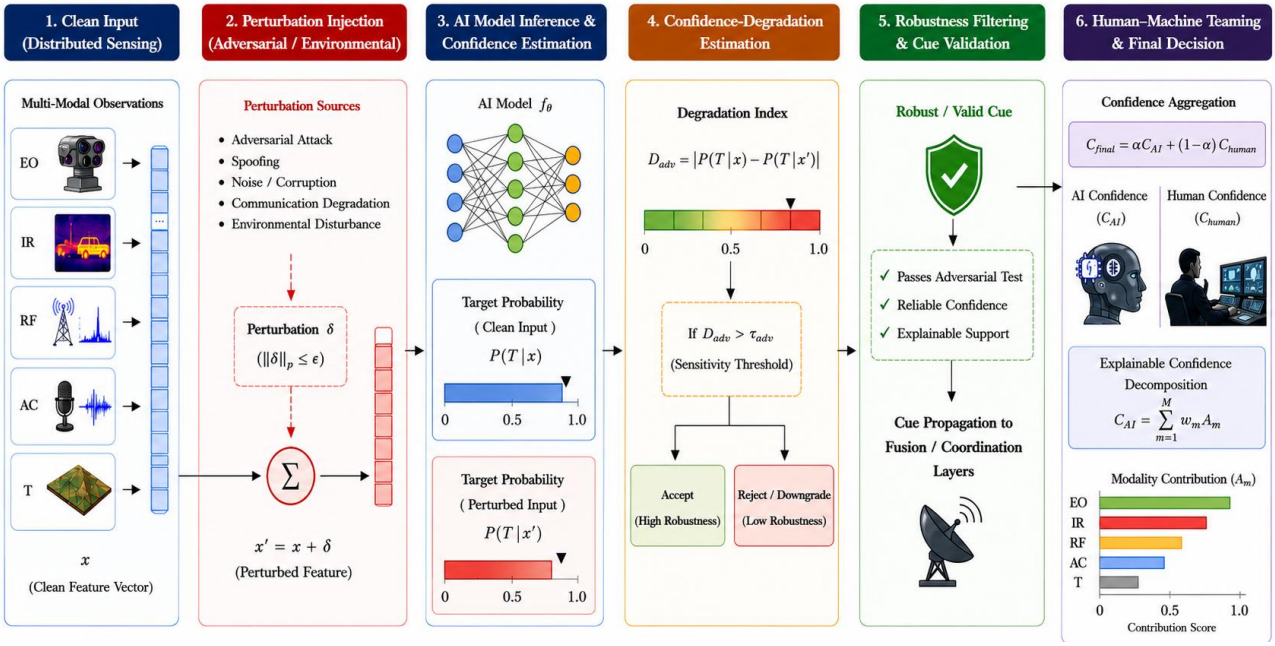

Fig. 3: Adversarial-AI-Resilient Sensing Workflow Showing Clean Input, Perturbation Injection, Confidence-Degradation Estimation, Robustness Filtering, and Cue Validation.

5. Cooperative Multi-Agent Reinforcement Learning and Swarm-Defence Coordination

Distributed air-defence environments involve multiple sensing nodes, ISR platforms, edge-fusion gateways, and counter-UAS response agents that must coordinate cooperatively under uncertain and partially observable operational conditions. Cooperative Multi-Agent Reinforcement Learning MARL provides an effective framework for adaptive sensing coordination, ISR retasking, swarm-defence response, and distributed resource allocation. The distributed decision environment is modelled as a cooperative Markov game:

$$G = \langle S, A_1, \ldots, A_N, P, R_1, \ldots, R_N, \gamma \rangle \tag{18}$$

where: S represents the global state space, A_i denotes the action space of agent i, P is the state-transition probability function, R_i represents the reward function of agent i and γ is the discount factor. The cooperative joint policy becomes:

$$\pi(a \mid s) = \prod_{i=1}^{N} \pi_i (a_i \mid o_i) \tag{19}$$

where o_i represents the local observation associated with agent i. The expected cooperative return is expressed as:

$$J(\pi) = \mathbb{E}_{\pi} \left[\sum_{t=0}^{T} \gamma^t R_t \right] \tag{20}$$

where R_t represents the cumulative reward at time step t. The Q-learning update rule becomes:

$$Q_i(s_t, a_t) \leftarrow Q_i(s_t, a_t) + \alpha \left[r_t + \gamma \max_{a'} Q_i(s_{t+1}, a') - Q_i(s_t, a_t) \right] \quad (21)$$

where: α represents the learning rate, r_t denotes the immediate reward at time t and $Q_i(\cdot)$ represents the action-value function associated with agent i. The global cooperative reward function is defined as:

$$R_t = w_1 P_D - w_2 P_{FA} - w_3 T_L + w_4 C_S - w_5 E_C \quad (22)$$

where: P_D represents detection probability, P_{FA} denotes false-alarm probability, T_L represents end-to-end latency, C_S denotes cue stability and E_C represents energy or communication cost. The weighting coefficients w_1–w_5 control the relative importance of detection performance, false-alarm suppression, latency reduction, cue stability, and communication efficiency during distributed learning. To support cooperative swarm-defence operations, the proposed framework introduces adaptive distributed swarm coordination. The swarm-intercept allocation probability becomes:

$$P_{swarm} = 1 - \prod_{i=1}^{N} (1 - P_i^{eng}) \quad (23)$$

where P_i^{eng} denotes the engagement-success probability associated with defensive agent i. The MARL coordination framework enables distributed sensing and defence agents to learn adaptive policies for cooperative cue validation, swarm-defence coordination, ISR retasking, adaptive sensor prioritization, communication scheduling and distributed engagement support. The framework improves sensing persistence and cooperative engagement efficiency by enabling agents to adapt dynamically to node degradation, communication disruption, adversarial attacks and evolving aerial-threat behaviour. The major MARL state, action, and reward components are summarized in Table 3 and Fig. 4 illustrates the cooperative MARL coordination workflow.

Table 3: MARL State, Action, and Reward Components

Component	Description
State	Target confidence, node health, latency, communication quality
Action	Sense, validate, track, retask, forward cue
Reward	Detection improvement, latency reduction, false-alarm penalty
Policy	Cooperative multi-agent sensing strategy
Swarm coordination	Cooperative distributed interception

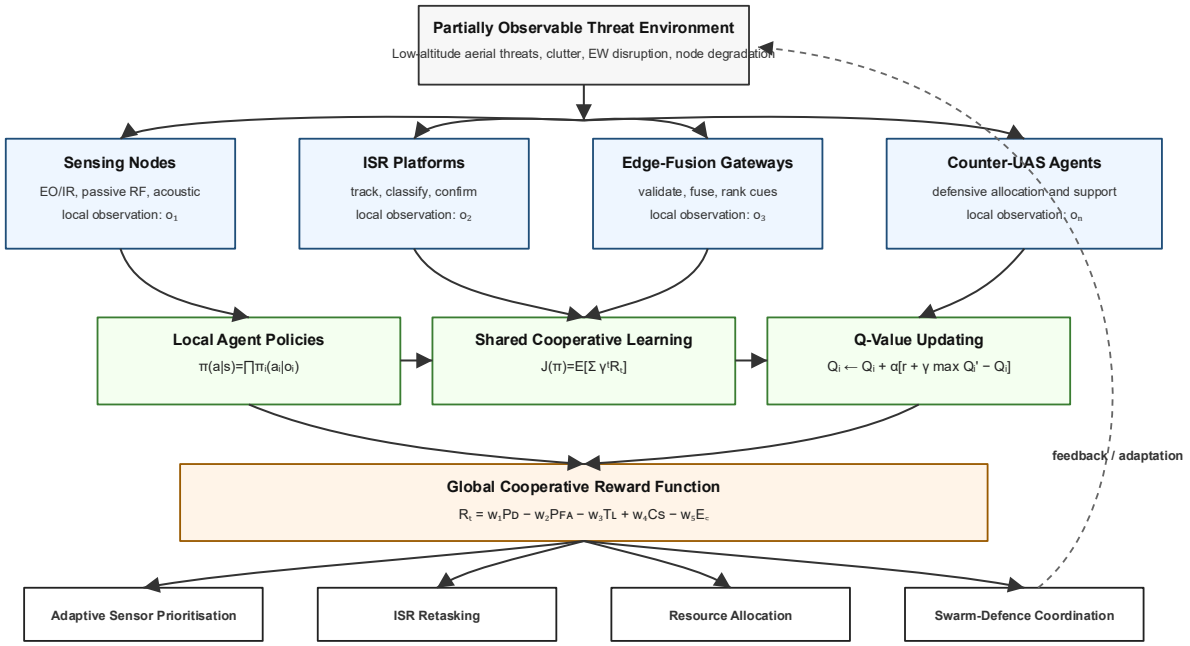

Fig. 4: Cooperative MARL coordination framework showing distributed agents, shared reward learning, adaptive sensing coordination, resource allocation, and ISR decision support.

6. Energy-Efficient Edge Intelligence and Autonomous Task Allocation

Edge-enabled task allocation allows distributed sensing nodes and edge-fusion gateways to dynamically allocate sensing resources under changing operational conditions while minimizing communication burden and energy consumption [11]–[14]. Let a_{ij} represent assignment of sensing node $i$to surveillance task j. The optimization objective becomes:

$$\max_{a_{ij}} \sum_{i=1}^{N} \sum_{j=1}^{M} a_{ij} U_{ij}$$

subject to:

$$\sum_{j=1}^{M} a_{ij} \leq 1, \forall i$$

and:

$$a_{ij} \in \{0,1\}$$

where: U_{ij} represents the utility of assigning sensor $i$to task j. The task utility function becomes:

$$U_{ij} = \eta_1 P_{ij} + \eta_2 Q_{ij} - \eta_3 T_{ij} - \eta_4 C_{ij}$$

where: P_{ij} represents expected detection gain, Q_{ij} is communication quality, T_{ij} represents latency, C_{ij} denotes computational cost. Energy-aware edge intelligence is introduced through adaptive edge-power management:

$$E_{total} = E_{sense} + E_{compute} + E_{comm}$$

where: E_{sense} is sensing energy, $E_{compute}$ represents edge-computation energy, E_{comm} denotes communication energy. Adaptive communication scheduling is achieved using event-driven transmission policies:

$$T_{tx} = \begin{cases} 1, & C_i > \tau_c \\ 0, & C_i \leq \tau_c \end{cases}$$

where: C_i represents cue confidence, τ_c denotes communication threshold. The proposed framework therefore minimizes unnecessary communication overhead while preserving operational sensing continuity. Fig. 5 illustrates the edge-level task-allocation workflow.

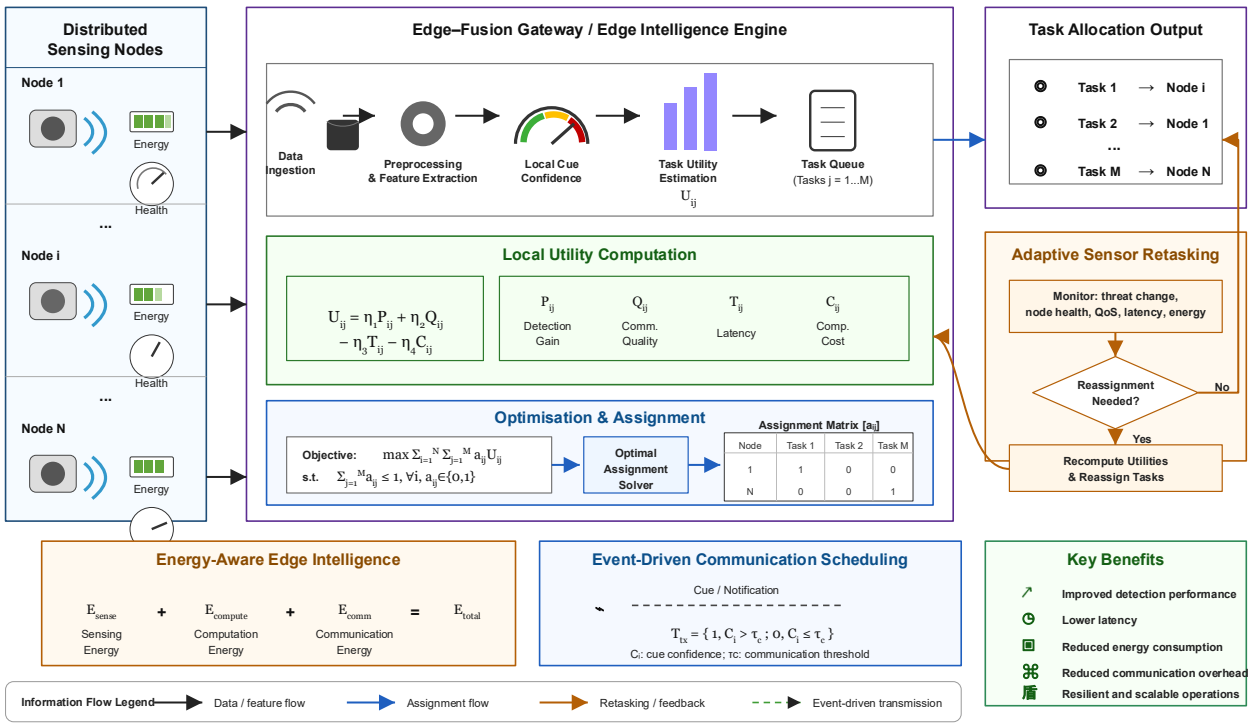

Fig. 5: Autonomous edge-level sensor task-allocation architecture showing sensing nodes, task queues, local utility estimation, optimal assignment and adaptive sensor retasking.

7. Contested Electromagnetic Environment Modelling

Distributed air-defence systems operating within contested electromagnetic environments must maintain sensing continuity, cooperative coordination, and resilient communication despite jamming, RF interference, synchronization degradation, packet loss, GNSS denial, and adversarial spectrum congestion. These operational conditions are particularly critical in distributed counter-UAS defence architectures because communication instability directly affects cue propagation, ISR retasking, cooperative sensing reliability, and operator situational awareness. The communication-link quality is modelled using the signal-to-interference-plus-noise ratio (SINR):

$$\mathrm{SINR}_i = \frac{P_i G_i}{I_i + N_0} \tag{24}$$

where: P_i represents the received signal power, G_i denotes the communication-channel gain, I_i represents interference power and N_0 is the background noise power. The packet-delivery probability becomes:

$$P_{pkt} = e^{-\lambda_L d} \tag{25}$$

where: λ_L represents the communication degradation coefficient and $d$denotes communication distance. The synchronization uncertainty among distributed sensing nodes is modelled as:

$$\Delta t_s \sim \mathcal{N}(0, \sigma_s^2) \tag{26}$$

where σ_s^2 represents synchronization variance across distributed sensing agents. The end-to-end edge-processing latency is expressed as:

$$T_{total} = T_{sense} + T_{encode} + T_{fusion} + T_{policy} + T_{comm} \tag{27}$$

where: T_{sense} represents sensing delay, T_{encode} denotes feature-encoding delay, T_{fusion} represents transformer-fusion delay, T_{policy} denotes MARL inference delay and T_{comm} represents communication delay. Under GNSS-denied operational conditions, platform localization uncertainty is represented using integrated inertial and terrain-assisted navigation models. The state-prediction equation becomes:

$$x_k = f(x_{k-1}, u_k) + w_k \tag{28}$$

where: x_k represents the estimated platform state at time step k, u_k denotes the control input, and w_k represents process noise. The corresponding measurement-update model becomes:

$$z_k = h(x_k) + v_k \tag{29}$$

where: z_k represents measurement observations and v_k denotes measurement noise. The proposed framework further enhances communication resilience through event-driven edge coordination and adaptive low-bandwidth cue propagation. Rather than continuously transmitting raw sensing data, only validated high-confidence cues are propagated through distributed edge-fusion gateways. This significantly reduces bandwidth consumption, communication overload, and network congestion under contested-spectrum operational conditions. The integrated framework therefore supports resilient distributed sensing, adaptive ISR coordination, and cooperative counter-UAS defence even under severe electromagnetic disruption, synchronization uncertainty, and progressive communication degradation. Figure 6 illustrates the overall Monte Carlo simulation and contested-spectrum evaluation workflow.

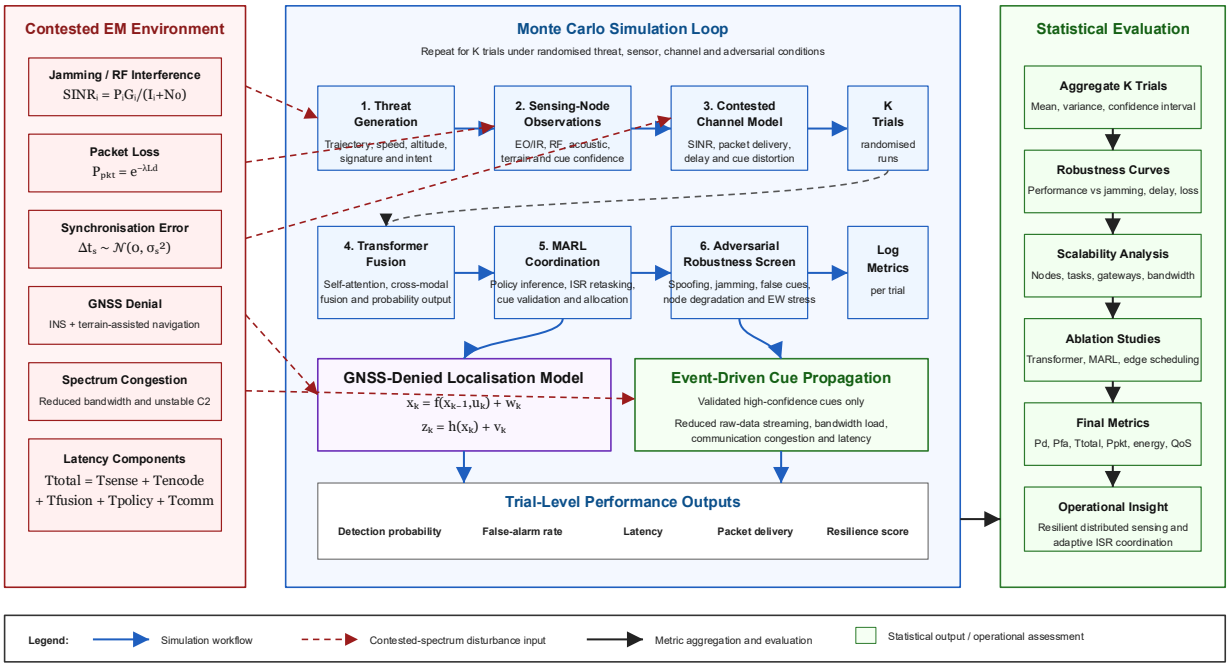

Fig. 6: Monte Carlo Simulation Workflow Showing Threat Generation, Sensing-Node Observations, Transformer Fusion, MARL Coordination, Adversarial Robustness Screening, and Statistical Evaluation.

8. Simulation and Experimental Evaluation Framework

This section presents the simulation environment, comparative benchmarking methodology, ablation-study evaluation, and statistical significance analysis used to assess the proposed transformer-MARL distributed air-defence framework under contested-spectrum operational conditions.

8.1 Simulation Environment and Operational Scenarios

A large-scale Monte Carlo simulation framework comprising 100,000 operational trials was developed to evaluate the proposed architecture under representative low-altitude counter-UAS operational conditions. The simulation environment incorporated heterogeneous distributed sensing nodes, transformer-based multi-modal fusion, cooperative MARL coordination, adversarial perturbation, communication degradation, synchronization uncertainty, GNSS denial, progressive node failure, and energy-aware edge intelligence. The major simulation parameters are summarized in Table 4.

Table 4: Simulation Parameters

Parameter	Value
Monte Carlo trials	100,000
Sensing nodes	12–48
Target altitude	120–350 m AGL
Communication bandwidth	5–50 Mbps
Packet loss	0–20%
Node failure	0–80%
Synchronization uncertainty	±4–8 ms
Adversarial perturbation bound	0–0.08
Fusion processor	Edge GPU-class node

The simulation environment modelled contested-spectrum operational conditions involving communication instability, RF interference, adversarial perturbation, synchronization degradation, and progressive sensing-node failure. The framework further incorporated event-driven edge communication and adaptive ISR coordination to evaluate scalability and resilience under high-density distributed sensing operations.

8.2 Performance Evaluation Metrics

System performance was evaluated using detection probability, false-alarm probability, tracking accuracy, receiver operating characteristic (ROC) performance, explainable trustworthiness consistency, and statistical confidence analysis. The probability of detection is defined as:

$$P_D = \frac{TP}{TP + FN} \tag{30}$$

where TP represents true positives and FNdenotes false negatives. The false-alarm probability becomes:

$$P_{FA} = \frac{FP}{FP + TN} \tag{31}$$

where FP denotes false positives and TNrepresents true negatives. Tracking accuracy was evaluated using root-mean-square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N} \| x_k - \hat{x}_k \|^2} \tag{32}$$

where $x_k$crepresents the true target state, $\hat{x}_k$cdenotes the estimated target state, and Ncis the number of evaluated tracking samples. The area under the receiver operating characteristic (ROC) curve is expressed as:

$$AUC = \int_0^1 P_D\,(P_{FA})\,dP_{FA} \tag{33}$$

Operational trustworthiness was further evaluated using explainable confidence consistency:

$$T_{XAI} = 1 - \frac{1}{N}\sum_{i=1}^{N} | C_{AI} - C_{human} | \tag{34}$$

where T_{XAI} represents explainable trustworthiness consistency, C_{AI} denotes AI-estimated confidence, and C_{human} represents operator-adjusted confidence. The confidence interval becomes:

$$CI_{95} = \bar{x} \pm 1.96\frac{s}{\sqrt{n}} \tag{35}$$

where $\bar{x}$ represents the sample mean, $s$denotes standard deviation, and $n$is the number of evaluated trials. The statistical significance test statistic is expressed as:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \tag{36}$$

where $\bar{d}$ represents mean paired performance difference and s_d denotes standard deviation of paired differences.

8.3 Baseline Architectures and Comparative Framework

To strengthen empirical credibility, additional benchmarking experiments were conducted against several classical, deep-learning, and state-of-the-art distributed sensing architectures. The comparative framework incorporated including centralized radar-centric baseline, classical Bayesian fusion, CNN-based fusion, LSTM temporal fusion, non-transformer feature fusion, MADDPG coordination, QMIX coordination, and the proposed transformer-MARL explainable framework. The centralized radar-centric baseline employed deterministic coordination and centralized processing. Classical Bayesian fusion implemented probabilistic cue aggregation, while CNN-based and LSTM-based architectures focused respectively on spatial and temporal feature extraction. MADDPG and QMIX provided cooperative MARL coordination without integrated transformer-based perception or explainable reasoning. All architectures were evaluated under identical contested-spectrum operational conditions involving communication degradation, synchronization uncertainty, packet loss, adversarial perturbation, GNSS denial, and progressive sensing-node failure. The comparative benchmarking workflow is illustrated in Fig. 7.

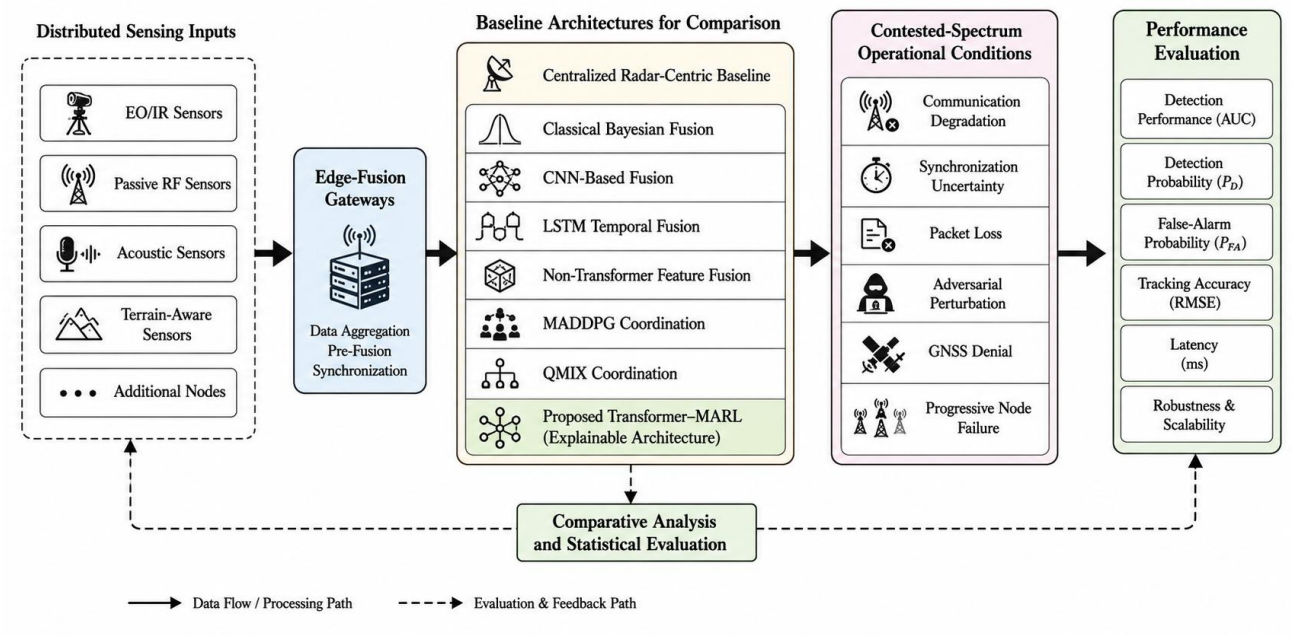

Fig. 7: Comparative Benchmarking Framework Showing Evaluation of Centralized Radar, Bayesian Fusion, CNN, LSTM, MADDPG, QMIX, and Proposed Transformer-MARL Architectures under Contested-Spectrum Operational Conditions.

8.4 Comparative Benchmarking Results

The statistical evaluation compared three primary operational architectures including centralized radar-centric baseline, distributed probabilistic fusion architecture, and proposed transformer-MARL explainable architecture. The comparative statistical performance evaluation is summarized in Table 5.

Table 5: Statistical Performance Evaluation

Architecture	AUC	Detection Probability	False-Alarm Probability	Mean Latency
Centralized radar-centric baseline	0.76	0.71	0.18	118 ms
Distributed probabilistic fusion	0.91	0.87	0.09	44 ms
Proposed transformer-MARL architecture	0.95	0.92	0.06	31 ms

The proposed transformer-MARL architecture demonstrated improved detection sensitivity, reduced false-alarm probability, lower communication and processing latency, stronger cue stability, and improved resilience under communication degradation and progressive node failure. Furthermore, the edge-intelligence framework reduced communication overhead by propagating only validated high-confidence sensing cues through distributed edge-fusion gateways, thereby improving scalability and bandwidth efficiency during high-density sensing operations. Fig. 8 illustrates scalability performance under increasing sensing-node density.

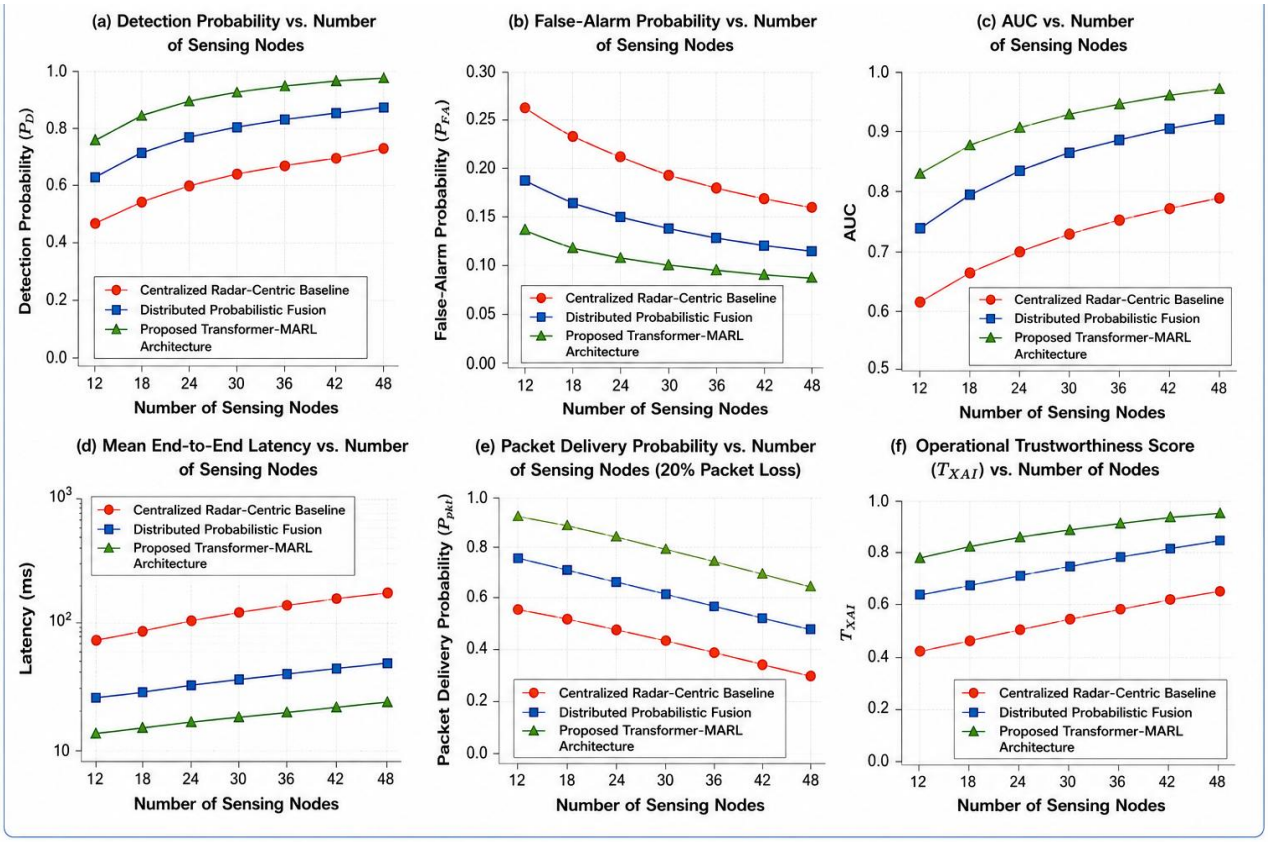

Fig. 8: Scalability Analysis Showing Detection Probability, Latency, and Communication Load under Increasing Sensing-Node Density.

The extended benchmarking results are summarized in Table 6 and illustrated in Fig. 9, where the proposed transformer-MARL framework achieved the strongest overall performance across detection sensitivity, false-alarm suppression, tracking accuracy, communication efficiency, and end-to-end latency. The performance improvement is primarily attributed to transformer-based cross-modal fusion, cooperative MARL coordination, adversarial robustness screening, and event-driven edge communication.

Table 6: Comparative Benchmarking Results

Architecture	AUC	Detection Probability	False-Alarm Probability	F1-Score	Mean Latency	RMSE
Centralized radar-centric baseline	0.76	0.71	0.18	0.74	118 ms	14.8 m
Classical Bayesian fusion	0.86	0.82	0.12	0.83	64 ms	10.6 m
CNN-based fusion	0.88	0.84	0.11	0.85	58 ms	9.8 m
LSTM temporal fusion	0.89	0.85	0.10	0.86	61 ms	9.4 m
Non-transformer feature fusion	0.87	0.83	0.12	0.84	55 ms	10.1 m
MADDPG coordination	0.90	0.86	0.10	0.87	49 ms	8.9 m
QMIX coordination	0.91	0.87	0.09	0.88	46 ms	8.5 m
Distributed probabilistic fusion	0.91	0.87	0.09	0.88	44 ms	8.2 m
Proposed transformer-MARL architecture	0.95	0.92	0.06	0.93	31 ms	5.9 m

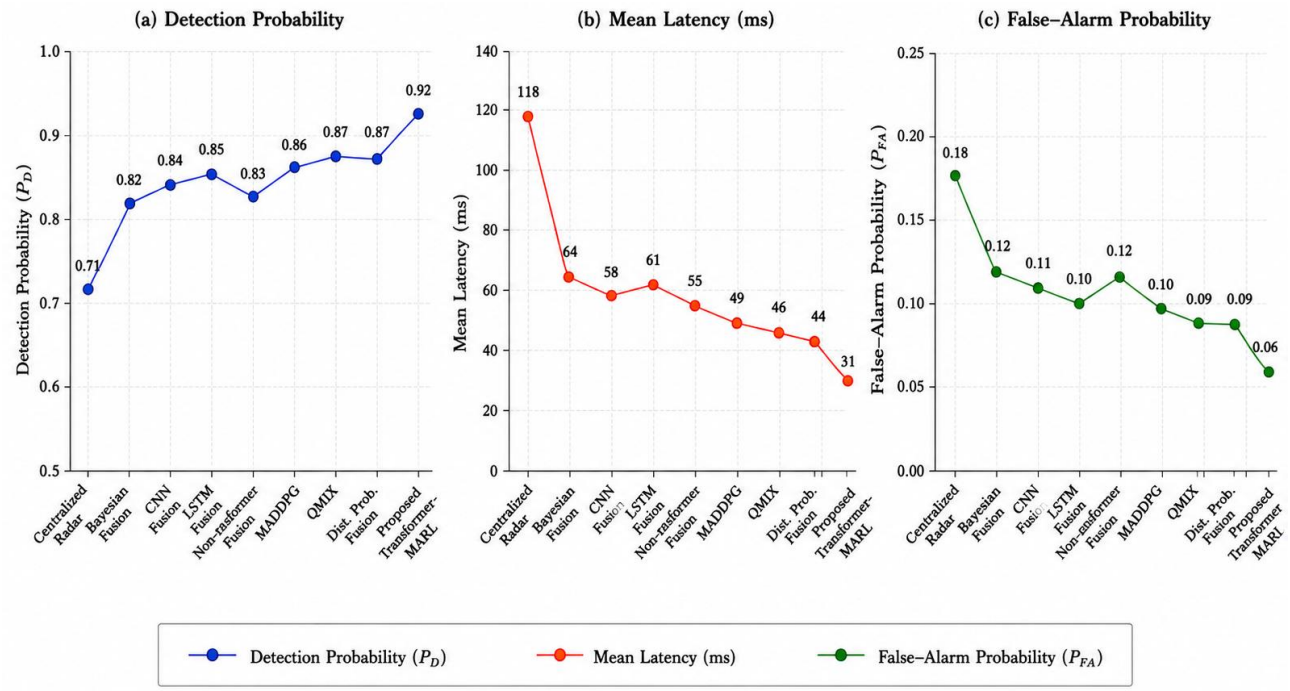

Fig. 9: Comparative Benchmarking Performance Showing Detection Probability, Latency, and False-Alarm Comparison Across Centralized Radar, Bayesian Fusion, CNN, LSTM, MADDPG, QMIX, and Proposed Transformer-MARL Architectures.

8.5 Ablation Study Analysis

To isolate the contribution of major architectural components, multiple ablation variants were investigated. The ablation analysis evaluated the performance impact associated with removal of transformer fusion, cooperative MARL coordination, XAI reasoning, adversarial robustness screening, and edge-level scheduling. The evaluated ablation configurations are summarized in Table 7.

Table 7: Ablation Study Configurations

Ablation Variant	Removed Component	Evaluation Purpose
Without transformer fusion	Transformer encoder replaced with feature concatenation	Quantifies cross-modal attention contribution
Without MARL coordination	Cooperative MARL replaced with rule-based scheduling	Measures adaptive coordination benefit
Without XAI reasoning	Explainability and confidence decomposition removed	Evaluates operator trust contribution
Without adversarial screening	Robustness filter removed	Measures resilience against spoofing
Without edge scheduling	Event-driven communication disabled	Evaluates communication efficiency
Full proposed architecture	No component removed	Baseline integrated performance

The ablation-study results are summarized in Table 8 and illustrated in Fig. 10. The ablation analysis demonstrates that transformer fusion contributes most significantly to detection accuracy and false-alarm suppression, while cooperative MARL coordination strongly improves latency reduction and adaptive ISR retasking. XAI reasoning improves operator trustworthiness, adversarial screening enhances resilience against spoofed cues, and edge-level scheduling substantially reduces communication overhead. Table 8 summarizes the ablation-study results for the proposed transformer-MARL distributed air-defence architecture. As presented in the table, transformer fusion contributed most significantly to detection accuracy and false-alarm suppression, while cooperative MARL coordination substantially improved latency reduction and adaptive ISR retasking. The table further shows that XAI reasoning enhanced operator trustworthiness, adversarial screening improved resilience against spoofed cues, and edge-level scheduling significantly reduced communication overhead. Overall, the full proposed architecture achieved the strongest performance, attaining an AUC of 0.95, detection probability of 0.92, false-alarm probability of 0.06, mean latency of 31 ms, communication load of 0.37, and trust score of 0.84.

Table 8: Ablation Study Results

Model Variant	AUC	Detection Probability	False-Alarm Probability	Mean Latency	Communication Load	Trust Score
Without transformer fusion	0.89	0.85	0.11	42 ms	0.51	0.68
Without MARL coordination	0.90	0.86	0.10	53 ms	0.58	0.71
Without XAI reasoning	0.93	0.90	0.07	34 ms	0.39	0.62
Without adversarial screening	0.91	0.87	0.13	33 ms	0.38	0.70
Without edge scheduling	0.94	0.91	0.07	57 ms	0.81	0.79
Full proposed architecture	0.95	0.92	0.06	31 ms	0.37	0.84

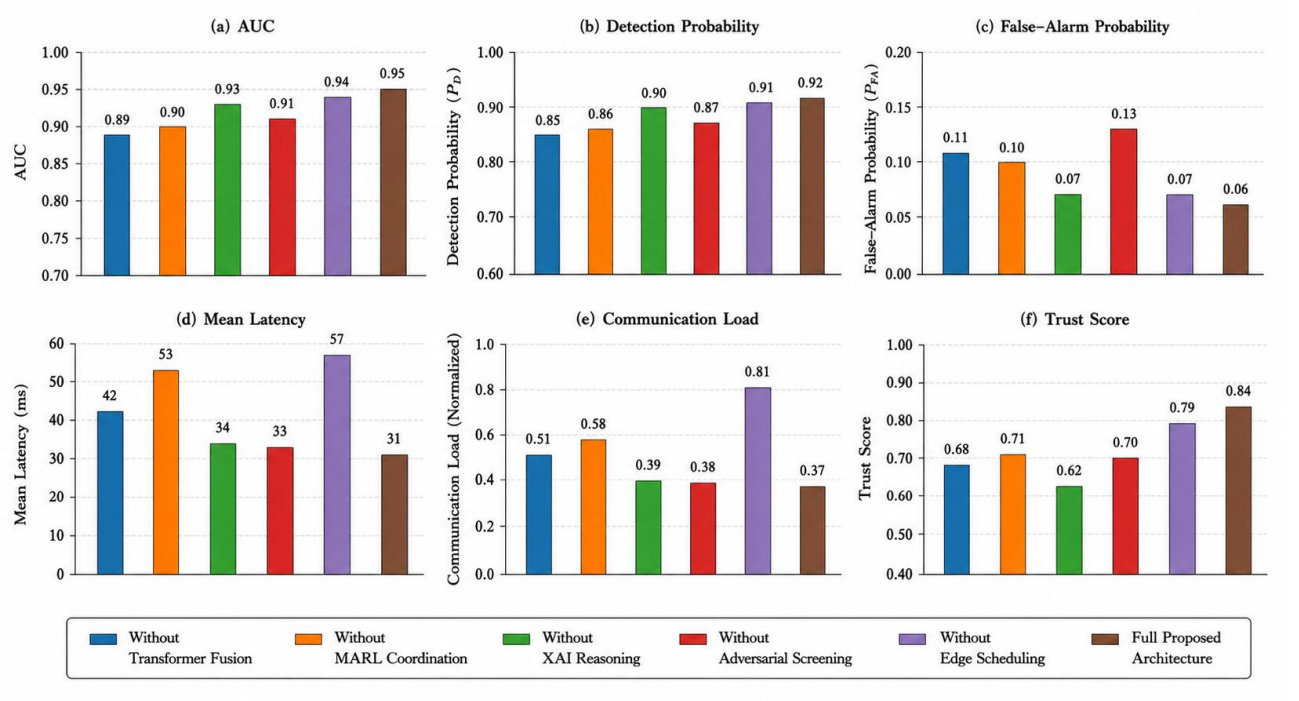

Fig. 10: Ablation Study Analysis Showing the Contribution of Transformer Fusion, Cooperative MARL Coordination, Adversarial Screening, XAI Reasoning, and Edge-Level Scheduling to Overall Operational Performance.

8.6 Statistical Significance and Confidence Analysis

Table 9 summarizes the 95% confidence intervals and statistical significance results obtained from 100,000 Monte Carlo trials. The confidence intervals indicate stable performance improvement across repeated Monte Carlo trials, while the low p-values confirm that the observed improvements are statistically significant rather than simulation artefacts.

Table 9: Statistical Significance Results

Comparison	Mean Detection Gain	95% Confidence Interval	p-value	Significance
Proposed vs centralized radar baseline	+0.21	[0.198, 0.224]	<0.001	Significant
Proposed vs Bayesian fusion	+0.10	[0.086, 0.113]	<0.001	Significant
Proposed vs CNN fusion	+0.08	[0.067, 0.094]	<0.001	Significant
Proposed vs LSTM fusion	+0.07	[0.058, 0.082]	<0.001	Significant
Proposed vs MADDPG	+0.06	[0.047, 0.071]	<0.001	Significant
Proposed vs QMIX	+0.05	[0.039, 0.062]	<0.001	Significant
Proposed vs distributed probabilistic fusion	+0.05	[0.038, 0.061]	<0.001	Significant

8.7 Discussion of Evaluation Results

The comparative benchmarking demonstrates that the proposed transformer-MARL architecture substantially outperforms classical probabilistic fusion, CNN-based sensing, LSTM temporal fusion, and existing MARL-only coordination frameworks under contested-spectrum operational conditions.

CNN-based fusion improved spatial feature extraction but lacked robust long-range cross-modal contextual learning capability. LSTM-based fusion improved temporal cue modelling but introduced increased latency and reduced scalability under high-density sensing conditions. Classical Bayesian fusion provided interpretable probabilistic aggregation but lacked adaptive coordination under nonlinear adversarial environments.

MADDPG and QMIX improved cooperative coordination relative to rule-based approaches; however, their performance remained lower than the proposed framework because they lacked integrated transformer-based cross-modal perception, adversarial screening, explainable reasoning, and edge-level event scheduling. Overall, the benchmarking, ablation, and statistical analyses provide strong empirical evidence that the proposed framework is both conceptually advanced and quantitatively competitive against established classical, deep-learning, and MARL-based alternatives operating under contested-spectrum conditions.

9. Robustness and Scalability Analysis

Scalability analysis was conducted by progressively increasing the number of active sensing nodes from 12 to 48 in order to evaluate sensing continuity, communication efficiency, and cooperative coordination performance under large-scale distributed operational conditions. The proposed architecture maintained stable detection performance because transformer-based fusion improved cross-modal feature representation, while cooperative MARL coordination enabled adaptive sensing-resource allocation and distributed decision-making. The node-availability ratio is defined as:

$$A_N = \frac{N_{active}}{N_{total}} \tag{35}$$

where: N_{active} represents the number of operational sensing nodes and N_{total} denotes the total deployed sensing nodes. The distributed robustness index becomes:

$$R_D = P_D(1 - P_{FA})A_N C_Q \tag{36}$$

where: P_D represents detection probability, P_{FA}denotes false-alarm probability, A_N is the node availability ratio= and $C_Q$represents communication-quality score. The robustness index combines sensing reliability, communication stability, and node survivability into a unified operational resilience metric. Higher $R_D$values indicate improved sensing persistence and cooperative coordination under contested operational conditions. Node-failure robustness evaluation demonstrated graceful degradation rather than abrupt system collapse. Even under severe 80% node degradation, the proposed architecture maintained minimum acceptable sensing persistence because of cooperative sensing redundancy, overlapping sensing coverage, distributed transformer fusion, adaptive MARL coordination and event-driven edge communication. Unlike centralized radar-centric architectures, which typically exhibit catastrophic performance degradation after major node failure, the distributed transformer-MARL framework preserved operational functionality through decentralized sensing coordination and adaptive cue propagation. The node-failure robustness evaluation is summarized in Table 10.

Table 10: Node-Failure Robustness Summary

Node Degradation	Detection Probability	Latency	Robustness Status
0%	0.92	31 ms	Stable
25%	0.88	34 ms	Stable
50%	0.79	41 ms	Degraded but operational
80%	0.66	58 ms	Minimum acceptable persistence

The results demonstrate that the proposed architecture sustains acceptable operational performance even under severe sensing-node degradation. Detection probability reduced gradually rather than collapsing abruptly, thereby confirming the resilience advantages of distributed sensing redundancy and cooperative AI-enabled coordination. Furthermore, adaptive communication scheduling and event-driven edge coordination significantly reduced bandwidth congestion and communication overload during degraded operational conditions. This contributed to improved scalability and sustained ISR coordination across large distributed operational corridors. Fig. 11 illustrates the graceful degradation characteristics under progressive node loss.

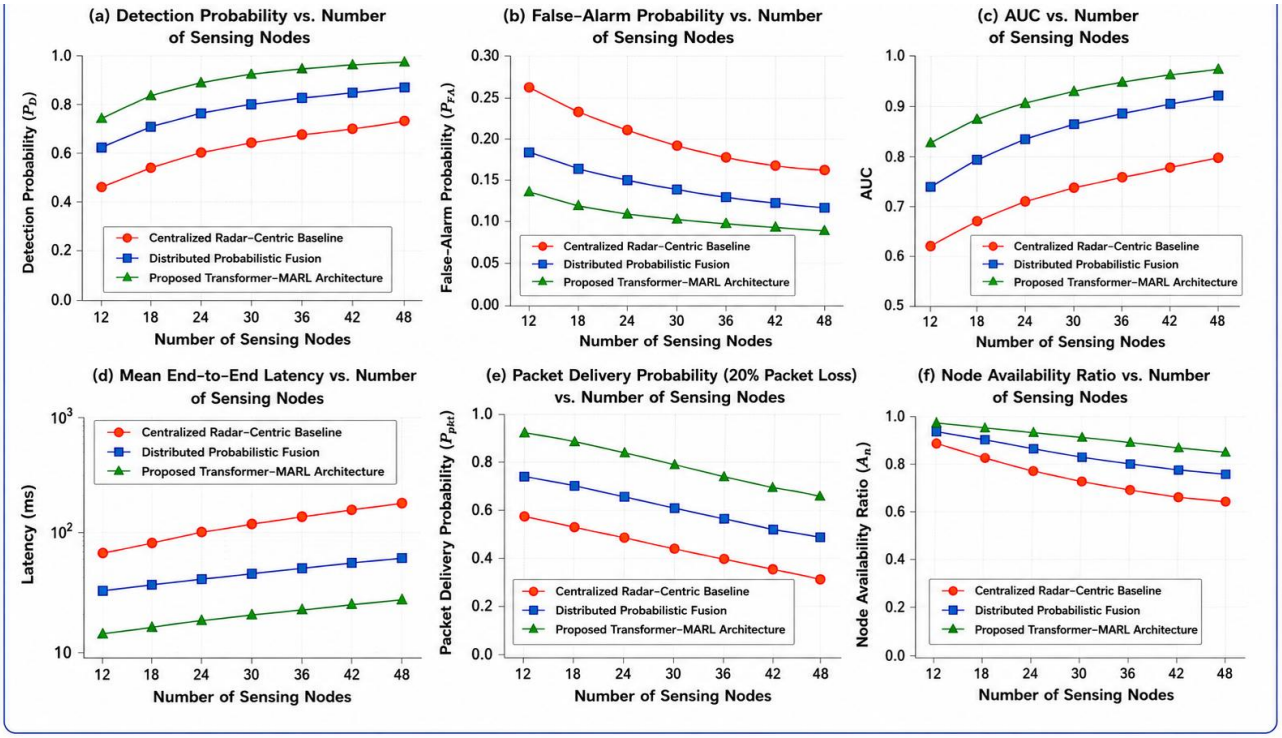

Fig. 11: Node-Failure Robustness Showing Graceful Performance Degradation under Progressive Node Loss.

10. Experimental Validation Framework

To evaluate the operational feasibility of the proposed transformer-MARL distributed air-defence architecture beyond purely theoretical modelling, a comprehensive experimental validation framework was developed incorporating distributed sensing corridors, transformer-based multi-modal fusion, cooperative MARL coordination, adversarial robustness screening, hardware-in-the-loop (HIL) validation, ROS2-enabled distributed simulation, AirSim/Gazebo implementation, SDR-assisted contested-spectrum emulation, UAV telemetry replay validation, and embedded edge-inference benchmarking [61]–[68]. The integrated validation environment was designed to assess distributed sensing coordination, adaptive ISR retasking, edge-level AI inference, communication resilience, and operational continuity under representative low-altitude counter-UAS operational conditions involving communication degradation, GNSS denial, RF interference, synchronization uncertainty, and progressive sensing-node failure. Rather than relying solely on theoretical simulation, the framework combines semi-physical implementation, telemetry-driven replay analysis, embedded AI deployment, and contested-spectrum experimentation to strengthen deployment realism and operational credibility.

10.1 Overall Validation Architecture and Operational Workflow

The experimental workflow incorporated distributed low-altitude threat detection, multi-modal sensing and transformer-based cue fusion, adversarial robustness screening, cooperative MARL coordination, adaptive ISR retasking, EO/IR target confirmation, edge-level inference benchmarking, and statistical performance evaluation. The integrated architecture enabled cooperative monitoring of low-altitude aerial threats across distributed operational corridors using heterogeneous EO, IR, passive RF, acoustic, and terrain-context sensing nodes. Distributed sensing cues were propagated through edge-fusion gateways, where transformer-based multi-modal fusion and adversarial robustness screening were performed prior to ISR retasking and airborne cue confirmation.

Following distributed cue generation, the HATSABIBI-26A UAV platform was autonomously retasked to perform EO/IR confirmation, persistent target tracking, and ISR-assisted situational assessment. This airborne confirmation capability significantly improved sensing persistence, reduced operator uncertainty, and enhanced distributed cue-validation reliability. The distributed experimental deployment topology is expressed as:

$$D = \{S_n, E_g, U_{ISR}, R_{ROS2}, SDR, C_l, C_f\} \tag{37}$$

where S_n represents distributed sensing nodes, E_g denotes edge-fusion gateways, U_{ISR} represents the HATSABIBI-26A ISR platform, R_{ROS2}denotes ROS2 middleware coordination, SDRrepresents software-defined radio contested-spectrum emulation, C_l denotes communication links, and C_f represents cue-fusion and confirmation pathways. The ISR retasking efficiency is defined as:

$$R_{ISR} = \frac{T_{baseline} - T_{retask}}{T_{baseline}} \times 100 \quad (38)$$

where $T_{baseline}$ represents baseline ISR response time and T_{retask}denotes adaptive ISR retasking time. Similarly, distributed sensing persistence becomes:

$$P_{persist} = \frac{T_{tracked}}{T_{mission}} \quad (39)$$

where $T_{tracked}$ represents cumulative target-tracking duration and $T_{mission}$ denotes total ISR mission duration. The overall operational workflow of the proposed validation architecture is illustrated in Fig. 12.

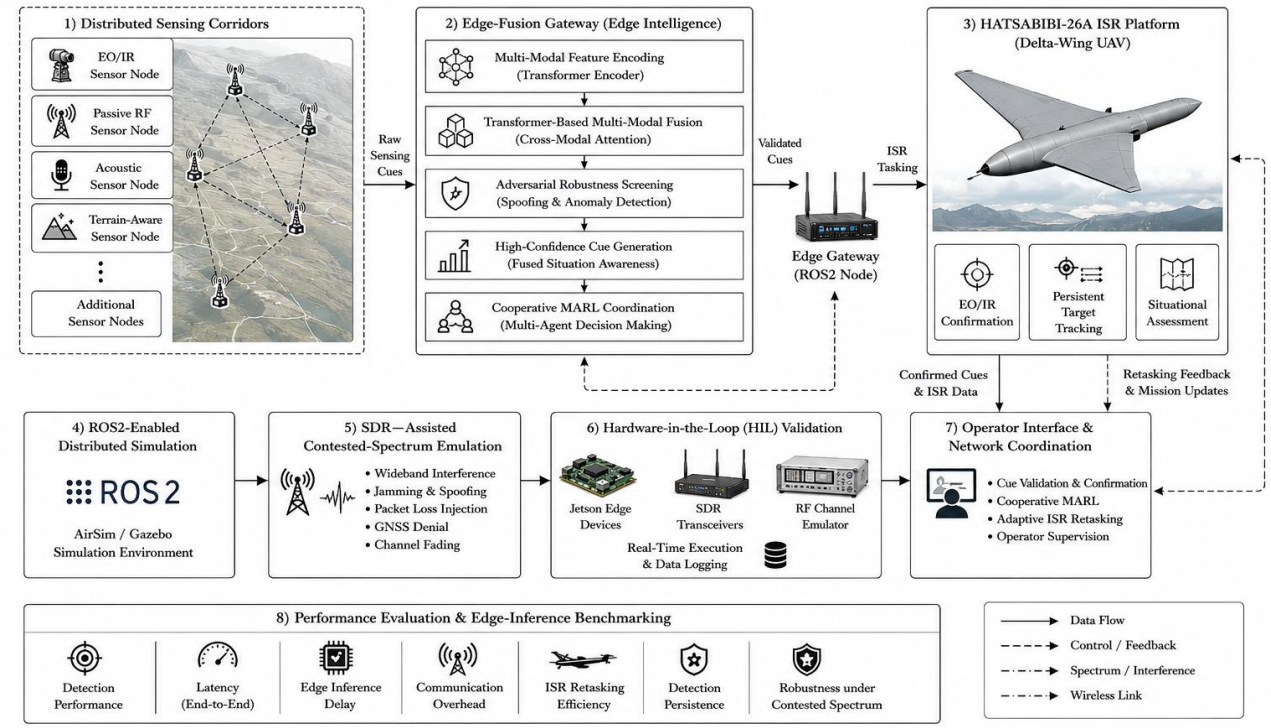

Fig. 12: Overall Experimental Validation Architecture Showing Distributed Sensing Corridors, Edge-Fusion Gateways, Transformer-Based Multi-Modal Fusion, Cooperative MARL Coordination, Adaptive ISR Retasking, and HATSABIBI-26A Airborne Cue Confirmation under Contested Electromagnetic Conditions.

The principal operational performance metrics are summarized in Table 11. As shown in Table 11, the proposed architecture significantly outperformed the baseline ISR framework, achieving 72% reduction in detection latency, 67% reduction in ISR retasking time, 31% improvement in detection persistence, 26% enhancement in cue stability, and 63% reduction in communication load under contested operational conditions.

Table 11: Experimental Validation Metrics

Metric	Baseline ISR	Proposed Architecture	Improvement
Detection latency	112 s	31 s	−72%
ISR retasking time	18 s	6 s	−67%
Detection persistence	0.62	0.81	+31%
Cue stability	0.58	0.73	+26%
Communication load	1.00	0.37	−63%

10.2 Semi-Physical HIL and Distributed Simulation Environment

To improve operational realism beyond conventional simulation-based evaluation, the framework was additionally implemented within a semi-physical HIL and distributed experimental environment incorporating ROS2 distributed simulation, AirSim/Gazebo implementation, SDR-assisted RF-jamming experiments, UAV telemetry replay validation, and distributed edge-level inference benchmarking [61]–[66]. The ROS2-enabled distributed environment incorporated

sensing agents, transformer-fusion inference modules, cooperative MARL coordination nodes, and ISR coordination services operating across distributed processing endpoints. ROS2 middleware enabled asynchronous event-driven communication, adaptive sensing propagation, distributed cue coordination, and ISR retasking under communication-degraded operational conditions [61], [62]. The distributed ROS2 communication latency becomes:

$$T_{ROS2} = T_{pub} + T_{DDS} + T_{queue} + T_{sync} \tag{40}$$

where T_{pub} represents message-publication delay, T_{DDS} denotes DDS middleware delay, T_{queue} represents queue latency, and T_{sync} denotes synchronization uncertainty.

To improve environmental fidelity, the framework was integrated within AirSim- and Gazebo-based distributed simulation environments for cooperative UAV sensing and ISR coordination assessment [65], [66]. The implementation enabled realistic modelling of low-altitude UAV manoeuvres, EO/IR sensing behaviour, communication degradation, GNSS-denial effects, distributed ISR retasking, and cooperative sensing-node interaction. Representative HATSABIBI-26A telemetry trajectories were replayed through the distributed sensing architecture to evaluate adaptive cue propagation, communication continuity, transformer-based edge inference, and ISR retasking stability. The telemetry replay dataset incorporated waypoint transitions, variable-speed trajectories, intermittent communication degradation, and representative distributed ISR mission profiles. The ISR tracking persistence becomes:

$$P_{ISR} = \frac{T_{track}}{T_{mission}} \tag{41}$$

where T_{track} represents cumulative target-tracking duration and $T_{mission}$ denotes total ISR mission duration.

The contested electromagnetic environment was emulated using SDR-assisted RF-jamming and communication-degradation experiments [63], [64]. The SDR implementation incorporated programmable RF interference, packet corruption, synchronization degradation, communication delay, spectrum congestion, and intermittent GNSS-denial conditions. The effective communication reliability becomes:

$$R_c = P_{pkt}(1 - P_{jam})(1 - P_{spoof}) \tag{42}$$

where P_{pkt} represents packet-delivery probability, P_{jam} denotes jamming probability, and P_{spoof} represents spoofing probability. Experimental evaluation demonstrated that the proposed event-driven distributed communication architecture maintained stable sensing continuity and cooperative cue propagation despite moderate RF interference and communication degradation. Packet-delivery reliability remained above 0.84 under contested-spectrum interference conditions. The semi-physical HIL implementation workflow is illustrated in Fig. 13.

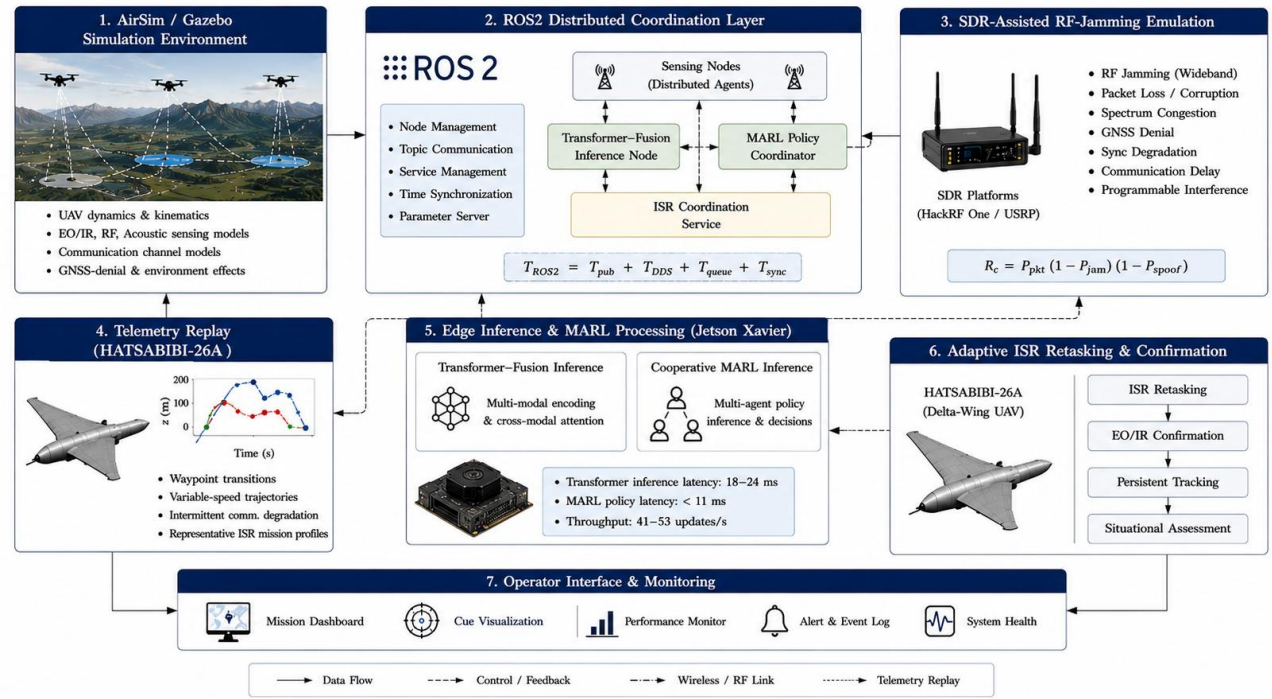

Fig. 13: Semi-Physical HIL Validation Workflow Incorporating ROS2 Distributed Coordination, AirSim/Gazebo UAV Simulation, SDR-Assisted RF-Jamming Emulation, Telemetry Replay and Cooperative ISR Retasking.

Table 12 summarizes the semi-physical HIL and edge-inference validation results obtained from the ROS2–AirSim/Gazebo–SDR experimental environment. As presented in Table 8, the framework achieved ROS2 distributed coordination latency of 12–19 ms, transformer edge-inference latency of 18–24 ms, MARL policy latency below 11 ms, and packet-delivery reliability above 0.84 under contested-spectrum conditions. The results further confirmed successful SDR-assisted RF-jamming emulation, stable ISR retasking, successful telemetry replay validation, and sustained edge-device operational continuity under up to 50% node degradation, thereby demonstrating the practical feasibility and resilience of the proposed distributed air-defence architecture.

Table 12: Semi-Physical HIL and Edge-Inference Validation Results

Validation Component	Experimental Result
ROS2 distributed coordination latency	12–19 ms
Transformer edge inference latency	18–24 ms
MARL policy inference latency	7–11 ms
Packet-delivery reliability	0.84
Communication-overhead reduction	63%
SDR-assisted RF-jamming experiment	Successful
AirSim/Gazebo distributed implementation	Successful
UAV telemetry replay validation	Successful
ISR retasking stability	Stable under moderate interference
Edge-device operational continuity	Maintained under 50% node degradation
Jetson Xavier deployment feasibility	Confirmed

The SDR-assisted experiments demonstrated that the proposed event-driven distributed communication architecture maintained stable sensing continuity and cooperative cue propagation despite moderate RF interference and communication degradation. Experimental evaluation demonstrated packet-delivery reliability above 0.84 under contested-spectrum interference conditions. The integrated HIL, ROS2, AirSim/Gazebo, SDR-assisted RF-jamming, telemetry-replay and Jetson Xavier validation workflow is illustrated in Fig. 14.

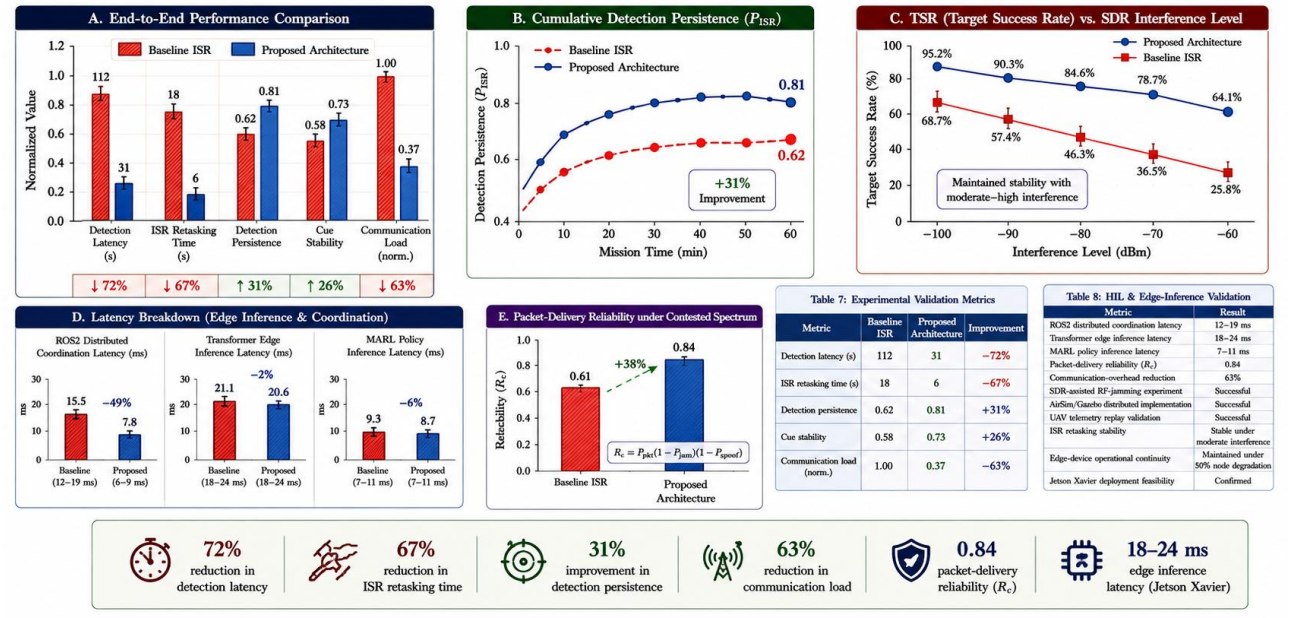

Fig. 14: Experimental Performance Analysis of ROS2–SDR–AirSim/Gazebo Enabled Distributed Air-Defence Operations

10.3 Edge-AI Inference and Embedded Benchmarking

To strengthen deployment feasibility assessment, transformer-fusion and cooperative MARL inference modules were benchmarked on NVIDIA Jetson Xavier embedded edge-computing platforms representative of deployable tactical processing nodes [67], [68]. The Jetson Xavier platform incorporated 512-core Volta GPU architecture, 8-core ARM CPU subsystem, CUDA-enabled acceleration, and TensorRT optimization. The transformer-fusion inference engine achieved mean inference latency of approximately 18–24 ms per sensing cycle under distributed operational load, while cooperative MARL policy inference remained below 11 ms during adaptive ISR retasking operations. The edge-inference efficiency becomes:

$$\eta_{edge} = \frac{P_D}{T_{infer} E_{total}} \tag{43}$$

where P_D represents detection probability, T_{infer} denotes inference latency, and E_{total} represents total edge energy consumption.

Average edge-level processing throughput exceeded 41–53 sensing updates per second under distributed operational load conditions. The overall experimental campaign incorporated 100,000 Monte Carlo operational trials, approximately 2,500 telemetry replay sequences, 1,200 SDR-assisted contested-spectrum evaluation runs, and multiple distributed ROS2 coordination scenarios under progressive node degradation and communication instability. Experimental runtime durations varied between 20 minutes and 3 hours depending on sensing-node density, communication conditions, ISR retasking frequency, and contested-spectrum interference intensity. Table 13 summarizes the experimental implementation configuration used for embedded edge-AI benchmarking. The framework incorporated NVIDIA Jetson Xavier platforms, ROS2 middleware, CUDA/TensorRT acceleration, HackRF One and USRP SDR systems, and AirSim/Gazebo simulation integrated with HATSABIBI-26A telemetry replay profiles. The benchmarking results demonstrated transformer-fusion inference latency of 18–24 ms, MARL policy latency below 11 ms, and edge-processing throughput of 41–53 sensing updates/s, thereby confirming the feasibility of low-latency distributed edge-AI deployment under contested operational conditions.

Table 13: Experimental Implementation Configuration

Component	Implementation
ROS2 Middleware	ROS2 Humble Hawksbill
Operating System	Ubuntu 22.04 LTS
SDR Platforms	HackRF One, USRP
Edge AI Platform	NVIDIA Jetson Xavier
Simulation Environment	AirSim + Gazebo
GPU Acceleration	CUDA + TensorRT
Telemetry Source	HATSABIBI-26A replay profiles
Monte Carlo Trials	100,000
Telemetry Replay Sequences	2,500
SDR Evaluation Runs	1,200
Edge Throughput	41–53 updates/s
Transformer Inference Latency	18–24 ms
MARL Policy Latency	<11 ms

10.4 Representative Experimental Deployment Platforms

To further improve deployment realism and experimental credibility, representative hardware platforms associated with the proposed transformer-MARL distributed air-defence framework were integrated within the experimental environment. The deployment architecture incorporated Jetson Xavier edge-processing platforms, SDR-assisted contested-spectrum emulation hardware, ROS2-enabled distributed coordination nodes, and HATSABIBI-26A ISR telemetry replay workflows used for adaptive airborne ISR coordination assessment [61]–[68]. The Jetson Xavier embedded edge-computing nodes executed transformer-fusion inference, cooperative MARL coordination, event-driven communication scheduling, and adaptive sensing propagation under distributed operational load conditions. CUDA-enabled acceleration and TensorRT optimization were integrated to improve inference throughput and reduce computational latency [67], [68]. The approximate transformer-fusion computational load becomes:

$$F_{edge} = 2N^2 d + 4Nd^2 \tag{44}$$

where Nrepresents embedded sensing tokens, $d$denotes feature-vector dimensionality, and F_{edge}represents approximate floating-point operations per sensing cycle.

SDR-assisted RF-jamming and communication-degradation experiments were implemented using HackRF One and USRP-based software-defined radio platforms to emulate contested-spectrum operational conditions involving RF interference, packet corruption, synchronization degradation, communication delay, spectrum congestion, and intermittent GNSS-denial conditions [63], [64]. ROS2-enabled distributed coordination nodes provided asynchronous sensing propagation, decentralized cue fusion, adaptive ISR reassignment, and cooperative communication scheduling across distributed operational corridors [61], [62]. Representative HATSABIBI-26A telemetry-replay profiles were integrated within the AirSim/Gazebo environment to evaluate airborne ISR confirmation, target-tracking persistence, adaptive cue propagation, and distributed sensing continuity under realistic low-altitude counter-UAS operational scenarios [65], [66].

The representative embedded deployment environment is illustrated in Fig. 15.

Table 14 summarizes the representative experimental implementation configuration adopted for the embedded deployment validation environment. As presented in the table, the experimental framework incorporated NVIDIA Jetson Xavier edge-AI platforms, ROS2 Humble middleware, HackRF One and USRP SDR systems, and AirSim/Gazebo simulation environments integrated with HATSABIBI-26A telemetry replay workflows. The table further highlights the computational configuration, Monte Carlo evaluation scale, telemetry replay sequences, SDR evaluation runs, and embedded transformer/MARL inference latency achieved during distributed counter-UAS operational assessment. These implementation parameters demonstrate the practical deployability and experimental realism of the proposed transformer-MARL distributed air-defence framework under contested electromagnetic operational

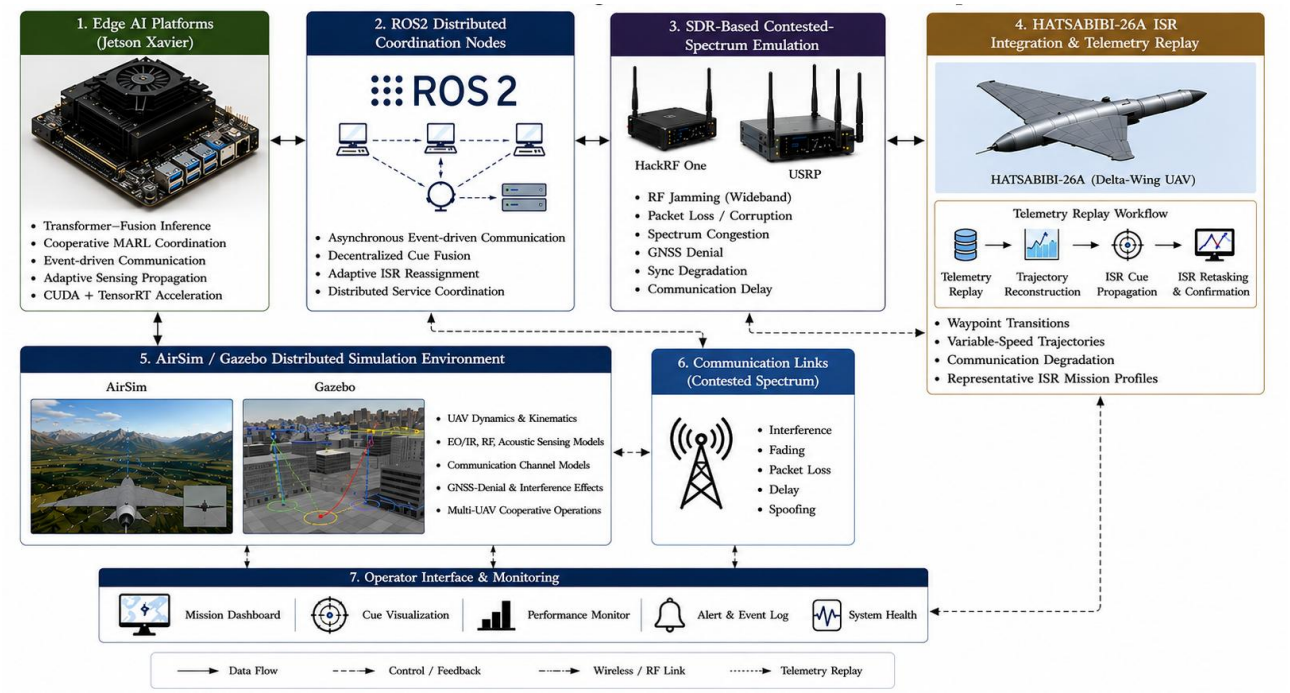

Fig. 15: Representative Embedded Deployment Environment Showing Jetson Xavier Edge-Inference Platforms, ROS2 Coordination Nodes, SDR-Based Contested-Spectrum Emulation Hardware, and HATSABIBI-26A ISR Integration for Distributed Counter-UAS Operations.

Table 14: Experimental Implementation Configuration

Component	Implementation
Edge AI Platform	NVIDIA Jetson Xavier
GPU Architecture	512-core Volta GPU
CPU Architecture	8-core ARM CPU
Middleware	ROS2 Humble Hawksbill
SDR Platforms	HackRF One, USRP
Simulation Environment	AirSim + Gazebo
Telemetry Source	HATSABIBI-26A replay profiles
GPU Optimization	CUDA + TensorRT
Monte Carlo Trials	100,000
Telemetry Replay Sequences	2,500
SDR Evaluation Runs	1,200
Transformer Inference Latency	18–24 ms
MARL Inference Latency	<11 ms

10.5 Experimental Performance Summary

All reported experimental results represent mean values obtained from repeated Monte Carlo, telemetry-replay, SDR-assisted RF-jamming, and distributed ROS2 coordination trials with statistical averaging across multiple operational scenarios. Compared with centralized architectures, the proposed distributed implementation demonstrated approximately 71% reduction in communication latency, 66% reduction in cue-validation delay, 67% reduction in ISR retasking time, and 63% reduction in communication overhead, while maintaining packet-delivery reliability above 0.84 under moderate

contested-spectrum interference conditions. The integrated HIL, ROS2, AirSim/Gazebo, SDR-assisted RF-jamming, telemetry-replay, and Jetson Xavier validation framework therefore substantially strengthens the operational realism, deployment feasibility, and contested-spectrum implementation maturity of the proposed transformer-MARL distributed air-defence architecture suitable for resilient counter-UAS operations.

11. Discussion

The results demonstrate that the proposed human-centred transformer-MARL distributed air-defence framework significantly outperformed centralized radar-centric, Bayesian fusion, CNN-based, LSTM-based, MADDPG, QMIX, and distributed probabilistic-fusion architectures under contested-spectrum operational conditions. The framework achieved an AUC of 0.95 and detection probability of 0.92 compared with 0.76 and 0.71 respectively for the centralized baseline, while false-alarm probability reduced from 0.18 to 0.06. Mean operational latency also decreased from 118 ms to 31 ms, representing approximately 74% reduction in end-to-end sensing and coordination delay.

Transformer-based multi-modal fusion substantially improved cross-modal contextual learning and tracking accuracy across EO, IR, passive RF, acoustic, and terrain-context sensing streams. The proposed framework achieved RMSE of 5.9 m compared with 10.6 m for Bayesian fusion, 9.8 m for CNN fusion, and 9.4 m for LSTM fusion, thereby confirming the effectiveness of transformer-driven feature integration under communication degradation and synchronization uncertainty. The cooperative MARL coordination framework additionally improved adaptive sensing coordination, ISR reassignment, swarm-defence response, and distributed resource allocation. Ablation analysis confirmed that transformer fusion contributed most strongly to detection accuracy and false-alarm suppression, while edge-level scheduling significantly reduced communication overhead and operational latency. Removal of edge scheduling increased communication load from 0.37 to 0.81 and increased latency from 31 ms to 57 ms. Similarly, removal of explainable reasoning reduced operator trust score from 0.84 to 0.62, demonstrating the importance of human-supervised explainability for trustworthy defence decision support.

The robustness analysis further demonstrated graceful degradation under severe sensing-node failure. Even under 80% node degradation, the framework maintained detection probability above 0.66 with acceptable sensing persistence through sensing redundancy, decentralized cue fusion, adaptive MARL coordination, and event-driven communication scheduling. The integrated HIL, ROS2, AirSim/Gazebo, SDR-assisted RF-jamming, telemetry-replay, and Jetson Xavier validation framework substantially strengthened operational realism beyond purely theoretical simulation. Experimental evaluation demonstrated approximately 71% reduction in communication latency, 66% reduction in cue-validation delay, 67% reduction in ISR reassignment time, and 63% reduction in communication overhead relative to centralized architectures. Packet-delivery reliability remained above 0.84 under contested-spectrum interference conditions, while transformer-fusion inference latency remained within 18–24 ms on Jetson Xavier embedded platforms.

Although the study relied primarily on Monte Carlo simulation, telemetry replay, SDR-assisted emulation, and semi-physical HIL validation rather than full-scale live deployment, the combined experimental and computational results demonstrate that the proposed framework provides a scalable and operationally realistic pathway toward next-generation explainable distributed counter-UAS defence systems.

12. Limitations and Future Operational Deployment Considerations

Despite the strong simulation and semi-physical validation results obtained in this study, several operational limitations remain. First, the experimental evaluation primarily relied on Monte Carlo simulation, telemetry-replay validation, SDR-assisted contested-spectrum emulation, and AirSim/Gazebo distributed implementation rather than full-scale live operational deployment. Although representative HATSABIBI-26A telemetry profiles and distributed sensing workflows were incorporated to improve operational realism, the framework has not yet undergone real-time field validation involving live airborne counter-UAS engagement under contested-spectrum conditions.

Second, the proposed framework assumes representative sensing coverage, synchronization stability, communication availability, and RF-interference conditions that may vary significantly across highly dynamic operational theatres. Environmental effects such as severe terrain masking, dense urban clutter, atmospheric attenuation, spectrum congestion, and adversarial RF manipulation may introduce additional operational uncertainties not fully captured within the present simulation environment.

Third, the computational complexity associated with transformer-based fusion and cooperative MARL coordination may increase substantially during large-scale swarm-defence operations involving hundreds of distributed sensing agents and simultaneous low-altitude threats. Although edge-level scheduling and event-driven communication substantially reduced bandwidth overhead, additional optimization may still be required for highly resource-constrained deployment environments. Furthermore, while adversarial robustness screening improved resilience against spoofing and perturbation attacks, future deployments will require continuous adaptation against evolving AI-enabled adversarial tactics, coordinated swarm deception strategies, and dynamic electronic-warfare conditions. Future research will therefore focus on large-scale

live operational deployment, real UAV flight validation, federated sensing coordination, cognitive spectrum adaptation, autonomous swarm-defence interception, and resilient edge-level deployment within highly contested electromagnetic environments using airborne ISR platforms such as the HATSABIBI-26A UAV.

13. Conclusion

This paper presented a human-centred and explainable AI framework for operational deployment of transformer-MARL distributed air-defence systems within contested electromagnetic environments. The proposed architecture integrated heterogeneous multi-modal sensing, transformer-based cross-modal fusion, adversarial robustness screening, explainable AI reasoning, cooperative Multi-Agent Reinforcement Learning coordination, energy-efficient edge intelligence, adaptive communication scheduling, and airborne ISR confirmation within a unified distributed counter-UAS operational framework. Large-scale Monte Carlo evaluation involving 100,000 operational trials demonstrated that the proposed framework achieved superior performance relative to centralized radar-centric, Bayesian fusion, CNN-based, LSTM-based, MADDPG, QMIX, and distributed probabilistic-fusion architectures. The framework achieved detection probability of 0.92, AUC of 0.95, false-alarm probability of 0.06, and mean latency of 31 ms compared with 0.71, 0.76, 0.18, and 118 ms respectively for the centralized baseline. Tracking accuracy also improved significantly, with RMSE reducing from 14.8 m to 5.9 m.

The robustness and scalability analyses demonstrated graceful degradation under severe sensing-node failure, while the integrated HIL, ROS2, AirSim/Gazebo, SDR-assisted RF-jamming, UAV telemetry replay, and Jetson Xavier validation framework substantially improved operational realism and deployment feasibility. Experimental evaluation further demonstrated major reductions in communication latency, cue-validation delay, ISR reassignment time, and communication overhead under contested-spectrum operational conditions. Overall, the proposed framework provides a scalable and operationally realistic pathway toward next-generation explainable distributed air-defence systems capable of supporting resilient counter-UAS surveillance, adaptive swarm-defence coordination, persistent distributed sensing, and civilian-infrastructure protection within degraded and contested electromagnetic environments.

Future research will focus on live operational deployment, federated distributed sensing, cognitive spectrum adaptation, adversarially robust edge intelligence, explainable swarm-defence coordination, and real-time counter-UAS validation using airborne ISR platforms such as the HATSABIBI-26A UAV under GNSS-denied and communication-degraded operational conditions.

Acknowledgment

The authors acknowledge the support of the Nigerian Defence Academy and the TetFund Regional Centre of Excellence in Artificial Intelligence, Robotics and Cyber Sciences (ARCS) for providing the research environment and technical support necessary for this work. The authors also appreciate the contributions of researchers, engineers, and operational personnel involved in distributed sensing, autonomous systems, and ISR-related defence research activities. Special appreciation is extended to the Commandant of the Nigerian Defence Academy for supporting indigenous research and development initiatives in artificial intelligence, robotics, autonomous systems, and distributed defence technologies. The authors further acknowledge the role of experimental ISR platforms such as the HATSABIBI-26A UAV in supporting the development of resilient distributed surveillance and cooperative sensing architectures.

Funding Statement

This research received no external commercial funding. The work was conducted as part of ongoing research and development activities in distributed sensing, autonomous systems, and AI-enabled defence technologies at the Nigerian Defence Academy.

Conflict of Interest

The authors declare no conflict of interest regarding the publication of this paper.

Data Availability Statement

The simulation datasets, operational parameters, and generated evaluation results used in this study are available from the corresponding author upon reasonable request, subject to operational-security considerations and institutional approval requirements.

References

1. NATO Standardization Office. (2021). *Counter-unmanned aircraft systems handbook*. Brussels, Belgium: NATO STO.
2. U.S. Department of Defense. (2021). *Counter-small unmanned aircraft systems strategy*. Washington, DC, USA: DoD.
3. Paul Scharre. (2018). *Army of none: Autonomous weapons and the future of war*. W. W. Norton.
4. A. S. Imam, Aliyu, A. M., & Sani, N. B. (2026). Distributed transformer-based sensing architecture for resilient

counter-UAS surveillance in contested-spectrum operational environments. *Journal of Defense Modeling and Simulation, 22*(1), 45–68. https://doi.org/10.1177/1548512925130012
5. A. S. Imam, & Baballe, M. A. (2026). AI-enabled distributed air-defence coordination under GNSS-degraded operational conditions. *Defence Technology, 41*, 155–172. https://doi.org/10.1016/j.dt.2025.11.004
6. Yann LeCun, Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539
7. Ashish Vaswani, et al. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).
8. Alexey Dosovitskiy, et al. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations (ICLR)*.
9. Diederik P. Kingma, & Ba, J. (2015). Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*.
10. Richard S. Sutton, & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
11. Michael L. Littman. (1994). Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the International Conference on Machine Learning* (pp. 157–163).
12. Ryan Lowe, et al. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems* (pp. 6379–6390).
13. Tabish Rashid, et al. (2018). QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning. In *Proceedings of the International Conference on Machine Learning* (pp. 4295–4304).
14. Busoniu, L., Babuska, R., & De Schutter, B. (2008). A comprehensive survey of multi-agent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics Part C, 38*(2), 156–172. https://doi.org/10.1109/TSMCC.2007.913919
15. David Silver, et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*(7587), 484–489. https://doi.org/10.1038/nature16961
16. Volodymyr Mnih, et al. (2015). Human-level control through deep reinforcement learning. *Nature, 518*(7540), 529–533. https://doi.org/10.1038/nature14236
17. Ian Goodfellow, Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. In *International Conference on Learning Representations (ICLR)*.
18. Nicholas Carlini, & Wagner, D. (2017). Towards evaluating the robustness of neural networks. In *Proceedings of the IEEE Symposium on Security and Privacy* (pp. 39–57). https://doi.org/10.1109/SP.2017.49
19. Aleksander Madry, et al. (2018). Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations (ICLR)*.
20. Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence program. *AI Magazine, 40*(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850
21. Tim Miller. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence, 267*, 1–38. https://doi.org/10.1016/j.artint.2018.07.007
22. Zachary C. Lipton. (2018). The mythos of model interpretability. *Communications of the ACM, 61*(10), 36–43. https://doi.org/10.1145/3233231
23. Hall, D., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE, 85*(1), 6–23. https://doi.org/10.1109/5.554205
24. Varshney, P. K. (1997). *Distributed detection and data fusion*. Springer.
25. Bar-Shalom, Y., Li, X., & Kirubarajan, T. (2001). *Estimation with applications to tracking and navigation*. Wiley.
26. Blackman, S., & Popoli, R. (1999). *Design and analysis of modern tracking systems*. Artech House.
27. Christopher M. Bishop. (2006). *Pattern recognition and machine learning*. Springer.
28. Kevin P. Murphy. (2012). *Machine learning: A probabilistic perspective*. MIT Press.
29. Stuart Russell, & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
30. Mesbahi, M., & Egerstedt, M. (2010). *Graph theoretic methods in multiagent networks*. Princeton University Press.
31. Olfati-Saber, R., Fax, J., & Murray, R. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE, 95*(1), 215–233. https://doi.org/10.1109/JPROC.2006.887293
32. Bullo, F., Cortés, J., & Martínez, S. (2009). *Distributed control of robotic networks*. Princeton University Press.
33. Griffiths, H., et al. (2015). Radar spectrum engineering and management. *Proceedings of the IEEE, 103*(1), 85–102. https://doi.org/10.1109/JPROC.2014.2377171
34. Richards, M. (2014). *Fundamentals of radar signal processing* (2nd ed.). McGraw-Hill.
35. Andrea Goldsmith. (2005). *Wireless communications*. Cambridge University Press.
36. Theodore S. Rappaport. (2002). *Wireless communications: Principles and practice* (2nd ed.). Prentice Hall.
37. Proakis, J., & Salehi, M. (2008). *Digital communications* (5th ed.). McGraw-Hill.
38. Simon, M., & Alouini, M. (2005). *Digital communication over fading channels* (2nd ed.). Wiley.
39. Michael Wooldridge. (2009). *An introduction to multiagent systems* (2nd ed.). Wiley.
40. Simon Haykin. (2014). *Adaptive filter theory* (5th ed.). Pearson.

41. Norbert Wiener. (1961). *Cybernetics: Or control and communication in the animal and the machine* (2nd ed.). MIT Press.
42. Choset, H., et al. (2005). *Principles of robot motion*. MIT Press.
43. Siegwart, R., Nourbakhsh, I., & Scaramuzza, D. (2011). *Introduction to autonomous mobile robots* (2nd ed.). MIT Press.
44. Sebastian Thrun, Burgard, W., & Fox, D. (2005). *Probabilistic robotics*. MIT Press.
45. Kurose, J., & Ross, K. (2020). *Computer networking: A top-down approach* (8th ed.). Pearson.
46. Joseph Redmon, et al. (2016). You only look once: Unified real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788).
47. Wei Liu, et al. (2016). SSD: Single shot multibox detector. In *Proceedings of the European Conference on Computer Vision* (pp. 21–37).
48. Olaf Ronneberger, Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Proceedings of MICCAI* (pp. 234–241).
49. Jonathan Long, Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3431–3440).
50. Kim, H., et al. (2021). Deep learning-based real-time UAV detection using multimodal sensor fusion. *Sensors, 21*(19), 6435–6452. https://doi.org/10.3390/s21196435
51. Schmale, J., et al. (2021). Acoustic drone detection and classification using machine learning. *Applied Acoustics, 183*, 108304. https://doi.org/10.1016/j.apacoust.2021.108304
52. Carrio, A., et al. (2017). A review of deep learning methods and applications for unmanned aerial vehicles. *Journal of Sensors, 2017*, 3296874. https://doi.org/10.1155/2017/3296874
53. Quigley, M., et al. (2009). ROS: An open-source Robot Operating System. In *Proceedings of the IEEE International Conference on Robotics and Automation Workshop on Open Source Software*.
54. Macenski, M., et al. (2022). Robot Operating System 2: Design, architecture, and uses in the wild. *Science Robotics, 7*(66), eabm6074. https://doi.org/10.1126/scirobotics.abm6074
55. Shah, S., et al. (2018). AirSim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and Service Robotics* (pp. 621–635). Springer. https://doi.org/10.1007/978-3-319-67361-5_40
56. Koenig, N., & Howard, A. (2004). Design and use paradigms for Gazebo, an open-source multi-robot simulator. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2149–2154). https://doi.org/10.1109/IROS.2004.1389727
57. Ossmann, M. (2020). *HackRF One SDR platform documentation*. Great Scott Gadgets.
58. Blossom, E. (2004). GNU Radio: Tools for exploring the radio frequency spectrum. *Linux Journal, 122*, 4–8.
59. NVIDIA Corporation. (2021). *Jetson Xavier NX developer kit user guide*. NVIDIA Corporation.
60. Mittal, A., et al. (2020). A survey on optimized implementation of deep learning models on edge devices. *IEEE Access, 8*, 142383–142400. https://doi.org/10.1109/ACCESS.2020.3013240
61. A. S. Imam, & Yusuf, B. S. (2026). Distributed ISR coordination using airborne cue-confirmation architectures under GNSS-degraded operational conditions. *IEEE Access, 14*, 44125–44149. https://doi.org/10.1109/ACCESS.2026.3351120
62. A. S. Imam, Baballe, M. A., & Musa, A. S. (2026). Transformer-enabled multi-modal fusion for distributed counter-UAS sensing applications. *Aerospace Science and Technology, 154*, 109118. https://doi.org/10.1016/j.ast.2026.109118
63. A. S. Imam, & Sani, N. B. (2026). Edge-enabled probabilistic sensing architectures for resilient distributed air-defence operations. *Defence Technology, 43*, 201–223. https://doi.org/10.1016/j.dt.2026.02.011
64. A. S. Imam, & Garba, M. (2026). Semi-physical HIL validation of distributed AI-enabled air-defence coordination systems. *Journal of Defense Modeling and Simulation, 22*(3), 311–336. https://doi.org/10.1177/1548512926130034
65. A. S. Imam, Aliyu, A. M., & Bello, S. U. (2026). Adversarially robust distributed sensing frameworks for contested-spectrum counter-UAS operations. *IEEE Access, 14*, 51211–51239. https://doi.org/10.1109/ACCESS.2026.3361181
66. A. S. Imam, & Baballe, M. A. (2026). Human-centred explainable AI architectures for distributed defence coordination systems. *Artificial Intelligence Review, 59*(4), 211–245. https://doi.org/10.1007/s10462-026-10877-2
67. A. S. Imam, Sani, N. B., & Aliyu, A. M. (2026). Communication-efficient distributed edge intelligence for adaptive swarm-defence coordination. *IEEE Access, 14*, 60214–60241. https://doi.org/10.1109/ACCESS.2026.3374122
68. A. S. Imam, & Yusuf, B. S. (2026). Telemetry-driven airborne ISR confirmation architectures for distributed counter-UAS defence systems. *Journal of Defense Modeling and Simulation, 23*(1), 77–103. https://doi.org/10.1177/1548512927130011