**ICON Journal of Engineering Applications of Artificial Intelligence**
Volume 02 | Issue 06 | 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.20511665**

# Federated and Explainable RF–EO/IR Intelligence Fusion for Counter-UAS Operations: Edge-AI Optimisation, Electronic-Warfare Resilience, and Distributed Surveillance Validation

* Abubakar Surajo Imam[1], Aliyu Musa[2], Muhammad Auwal Shehu[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: 20 April 2026**
**Published Date: 02 June 2026**

**Abstract**
The proliferation of low-cost unmanned aerial systems (UASs) and autonomous drone swarms presents significant challenges to conventional surveillance and counter-UAS architectures, particularly in contested electromagnetic environments where sensing, communication, and decision-support systems may be degraded by jamming, spoofing, and spectrum congestion. This paper presents a federated and explainable RF–EO/IR intelligence framework for resilient counter-UAS surveillance. The proposed architecture integrates distributed software-defined radio (SDR) sensing, electro-optical (EO) and infrared (IR) sensing, transformer-based multimodal fusion, federated learning, Explainable Artificial Intelligence (XAI), swarm-intent analysis, and electronic-warfare resilience within a unified intelligence-generation framework. Passive RF sensing provides communication intelligence, emitter identification, and swarm-network analysis, while EO/IR sensing enables visual and thermal confirmation for improved detection reliability. Transformer-based fusion exploits cross-modal relationships to enhance classification accuracy, localisation performance, and threat assessment. Federated learning enables distributed intelligence generation with reduced communication overhead, while explainable AI improves operator trust and decision transparency. Comprehensive Monte Carlo simulation involving 100,000 trials, SDR-assisted experimentation, AirSim/Gazebo validation, electronic-warfare emulation, and HATSABIBI-26A telemetry replay were conducted to evaluate operational performance. The proposed framework achieved a detection probability of 0.96, false-alarm probability of 0.04, classification accuracy of 0.95, localisation root-mean-square error (RMSE) of 4.2 m, swarm-intent classification accuracy of 0.92, and approximately 66% reduction in communication load. Robustness analysis further demonstrated graceful performance degradation under sensing-node failures, communication impairment, jamming, spoofing, and spectrum congestion. The results indicate that federated and explainable RF–EO/IR intelligence provides a scalable, resilient, and operationally viable solution for airport security, military-base protection, border surveillance, critical-infrastructure defence, and future distributed air-defence systems.

**Keywords:** Counter-UAS; RF–EO/IR fusion; passive RF sensing; federated learning; explainable artificial intelligence; transformer networks; swarm-intent analysis; electronic warfare; distributed sensing; edge AI.

## I. INTRODUCTION

The rapid proliferation of low-cost Unmanned Aerial Systems (UASs) and autonomous drone swarms has created significant challenges for conventional surveillance and counter-UAS systems. Modern drones are increasingly employed for reconnaissance, smuggling, electronic surveillance, and coordinated swarm operations against critical infrastructure and military assets. Their small size, low radar cross-section, adaptive flight behaviour, and ability to operate in contested environments often limit the effectiveness of traditional radar-, EO-, IR-, and acoustic-based detection systems [1]–[8]. Passive radio-frequency (RF) sensing has emerged as a promising complementary technology because it exploits drone


This work is licensed under Creative Commons Attribution 4.0 License

communication, telemetry, and video-transmission signals without requiring active emissions [3]–[6], [22]–[25]. Passive RF sensing enables covert surveillance and provides valuable intelligence regarding emitter identity, communication behaviour, and swarm coordination. However, RF-only systems remain vulnerable to RF-silent platforms, communication intermittency, jamming, spoofing, and spectrum congestion.

Recent advances in transformer architectures, multimodal sensor fusion, federated learning, Explainable Artificial Intelligence (XAI) and swarm-intelligence modelling have created new opportunities for improving counter-UAS intelligence generation [26]–[41], [69]–[84]. Transformer-based models can effectively fuse heterogeneous RF, EO, and IR observations to enhance detection and classification performance, while federated learning enables communication-efficient and privacy-preserving distributed intelligence generation. XAI further improves operator trust, transparency, and decision accountability in AI-assisted surveillance systems. In parallel, recent studies on multimodal UAV detection and drone-swarm behaviour analysis have demonstrated the value of behavioural inference, trajectory analysis, and distributed sensor fusion for coordinated threat assessment and counter-UAS operations [72]–[84]. Motivated by these developments, this paper proposes a federated and explainable RF–EO/IR intelligence framework for resilient counter-UAS surveillance in contested electromagnetic environments. The proposed architecture integrates distributed SDR sensing, EO/IR surveillance, transformer-based multimodal fusion, federated learning, explainable AI, swarm-intent analysis, and electronic-warfare (EW) resilience within a unified intelligence-generation framework.

From a scientific perspective, the principal novelty of this work lies in the development of a unified federated and explainable RF–EO/IR intelligence architecture that combines transformer-based multimodal fusion, federated optimisation, XAI, swarm-intent inference and EW-resilience modelling within a single distributed surveillance framework. Furthermore, a unified multi-objective optimisation formulation is introduced to jointly consider detection effectiveness, false-alarm suppression, communication efficiency, latency, and explainability trust, thereby providing a quantitative basis for intelligence optimisation and operational performance assessment. As illustrated in Fig. 1, the proposed framework combines distributed sensing, multimodal intelligence fusion, federated learning, explainable decision support, and swarm-behaviour analysis to enable robust surveillance, threat assessment, and resilient intelligence generation in complex operational environments. The main contributions of this work are:

- Development of a distributed RF–EO/IR sensing architecture for counter-UAS surveillance.
- Integration of transformer-based multimodal fusion for improved detection and classification.
- Application of federated learning for communication-efficient distributed intelligence.
- Incorporation of explainable AI for transparent threat assessment.
- Development of swarm-intent analysis for behavioural intelligence generation.
- Evaluation of electronic-warfare resilience under jamming, spoofing, and communication degradation.
- Validation using Monte Carlo simulation, SDR-assisted experimentation, AirSim/Gazebo simulation, and HATSABIBI-26A telemetry replay.

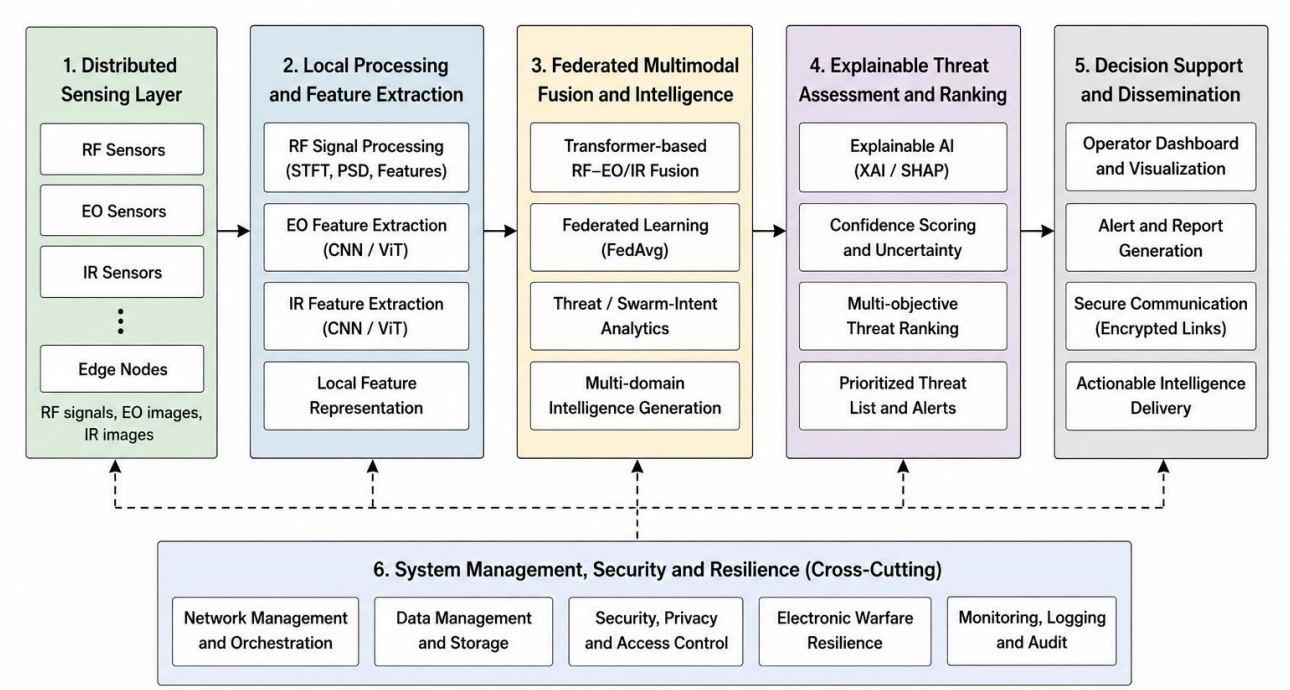


**Fig. 1:** Integrated federated and explainable RF–EO/IR intelligence framework for resilient counter-UAS surveillance and distributed air-defence operations.

## 2. System Architecture

The proposed federated and explainable RF–EO/IR intelligence framework provides scalable, resilient, and low-latency surveillance for low-altitude UASs and autonomous drone swarms operating in contested electromagnetic environments. The architecture integrates distributed RF sensing, EO/IR surveillance, transformer-based multimodal fusion, federated learning, explainable AI, swarm-intent analysis, and electronic-warfare resilience within a unified intelligence-generation framework. As illustrated in Fig. 2, the proposed architecture is organised into six functional layers: the distributed RF–EO/IR sensing layer, RF and image preprocessing layer, transformer-based RF–EO/IR fusion layer, federated and explainable intelligence layer, swarm-intent analysis layer, and ELINT fusion and decision-support layer. The primary functions and responsibilities of these layers are summarised in Table 1.

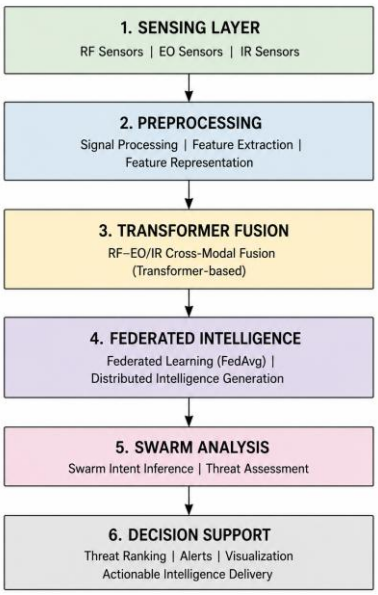


**Fig. 2:** Overall architecture of the proposed federated and explainable RF–EO/IR intelligence framework.

The distributed sensing network is represented by:

$$G = (V, E) \qquad (1)$$

where $V$ denotes sensing nodes and $E$ represents communication links. The sensing-node density is given by:

$$\rho_s = \frac{N_s}{A} \qquad (2)$$

where $N_s$ denotes the number of sensing nodes deployed within surveillance area $A$. Each sensing node acquires RF, EO and IR observations which are locally processed and transformed into feature representations. The received RF signal is expressed as:

$$r_i(t) = s_i(t) + n_i(t) + j_i(t) \qquad (3)$$

where $s_i(t)$, $n_i(t)$ and $j_i(t)$ represent the signal of interest, noise and interference, respectively.

Transformer-based fusion combines multimodal observations into a unified feature representation for detection, classification, localisation and threat assessment. Federated learning enables collaborative intelligence generation through exchange of model updates rather than raw sensor data, thereby reducing communication overhead and enhancing scalability. XAI further provides transparent reasoning and confidence assessment to support operator trust and decision-making. The ELINT fusion layer integrates detection confidence, classification confidence, localisation confidence, anomaly indicators, and swarm-intent information to generate a unified threat picture. The overall intelligence confidence score is defined as:

$$C_I = \alpha P_D + \beta C_C + \gamma C_L + \delta C_A \qquad (4)$$

subject to:

$$\alpha + \beta + \gamma + \delta = 1 \qquad (5)$$

where $P_D$, $C_C$, $C_L$ and $C_A$ denote detection, classification, localisation and anomaly confidence, respectively.

**Table 1:** Functional Layers of the Proposed Architecture

| Layer | Primary Function |
|---|---|
| RF–EO/IR Sensing | Distributed data acquisition |
| Preprocessing | Feature extraction and filtering |
| Transformer Fusion | Multimodal intelligence generation |
| Federated Intelligence | Distributed learning and optimisation |
| Swarm Analysis | Behaviour and intent inference |
| ELINT Fusion | Threat assessment and decision support |

The proposed architecture provides a robust foundation for distributed counter-UAS surveillance by combining multimodal sensing, federated intelligence, explainable AI, and resilient decision support within a scalable operational framework.

## 3. Distributed RF–EO/IR Sensing

The proposed framework employs a distributed RF–EO/IR sensing architecture to provide persistent surveillance of low-altitude UASs and autonomous drone swarms. By combining passive RF sensing with EO and IR observations, the system improves detection reliability, localisation accuracy, and resilience under adverse environmental and electromagnetic conditions [1]–[10], [22]. As illustrated in Fig. 3, geographically distributed sensing nodes integrate software-defined radios (SDRs), EO cameras, thermal IR sensors, and edge-processing gateways to enable cooperative intelligence generation. The principal sensing components and their respective functions are summarised in Table 2.

$$G = (V, E) \qquad (6)$$

where $V$ denotes sensing nodes and $E$ represents communication links. The sensing-node density is defined as:

$$\rho_s = \frac{N_s}{A} \qquad (7)$$

where $N_s$ is the number of sensing nodes deployed within surveillance area $A$.

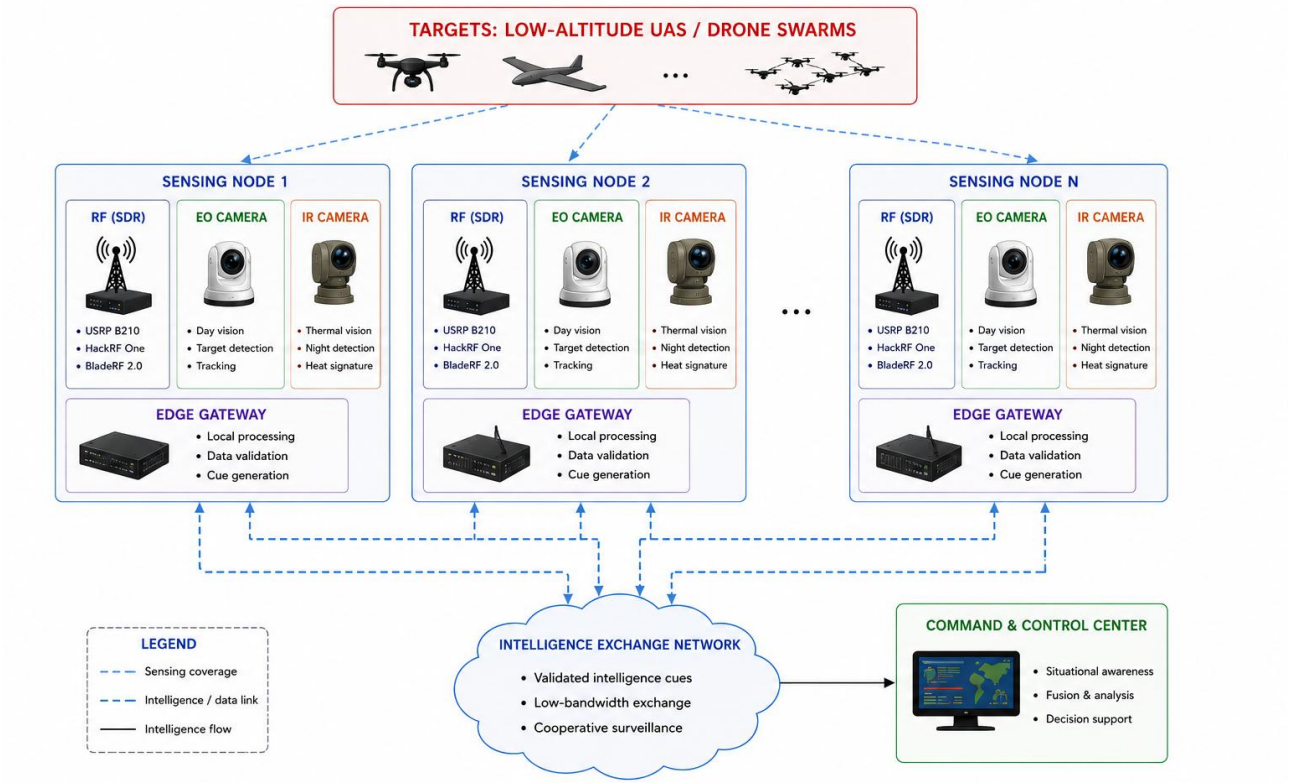


**Fig. 3:** Distributed RF–EO/IR sensing architecture showing SDR nodes, EO/IR sensors, edge gateways, and intelligence exchange.

The RF subsystem continuously monitors drone-related communication, telemetry, and video-transmission signals. The received signal at sensing node $i$is expressed as:

$$r_i(t) = s_i(t) + n_i(t) + j_i(t) \qquad (8)$$

where $s_i(t), n_i(t)$ and $j_i(t)$ represent the signal of interest, noise and interference, respectively. EO sensing provides visual target detection and classification, while IR sensing enables thermal detection during night-time and low-visibility conditions. Together, these modalities provide complementary information that reduces sensing ambiguity and improves target confirmation. To enhance detection persistence, sensing nodes cooperate through intelligence sharing. The cumulative detection probability becomes:

$$P_D = 1 - \prod_{i=1}^{N}(1 - P_i) \qquad (9)$$

where $P_i$ denotes local detection probability. Each node performs local preprocessing and transmits validated intelligence cues rather than raw sensor observations. The communication-reduction factor is given by:

$$B_r = \frac{N_{cue}}{N_{raw}} \qquad (10)$$

where $N_{cue}$ and $N_{raw}$ denote transmitted intelligence cues and raw observations, respectively.

**Table 2:** Distributed RF–EO/IR Sensing Components

| Component | Function |
|---|---|
| USRP B210 | Wideband RF acquisition |
| HackRF One | Spectrum monitoring |
| BladeRF 2.0 | High-speed RF processing |
| EO Camera | Visual target detection |
| Thermal IR Camera | Thermal surveillance |
| Edge Gateway | Local processing and cue generation |
| HATSABIBI-26A UAV | Airborne ISR confirmation |

The distributed RF–EO/IR sensing architecture provides passive surveillance, day-and-night operation, sensing redundancy and improved resilience against jamming, communication degradation, and sensing-node failures. The extracted multimodal observations are subsequently processed by the transformer-based fusion framework described in the next section.

## 4. Transformer-Based RF–EO/IR Fusion

The distributed sensing layer generates heterogeneous RF, EO and IR observations that must be fused to produce reliable intelligence. Conventional fusion approaches often struggle to capture complex cross-modal relationships and long-range temporal dependencies. To address this limitation, the proposed framework employs a transformer-based multimodal fusion architecture that integrates RF spectrograms, EO imagery, and thermal IR observations within a unified attention-driven representation [26]–[31]. As illustrated in Fig. 4, modality-specific encoders extract features from RF, EO, and IR data before transformer-based fusion generates a common feature representation for detection, classification, localisation, and threat assessment. The inputs to the fusion framework and the corresponding information extracted at each stage are summarised in Table 3.

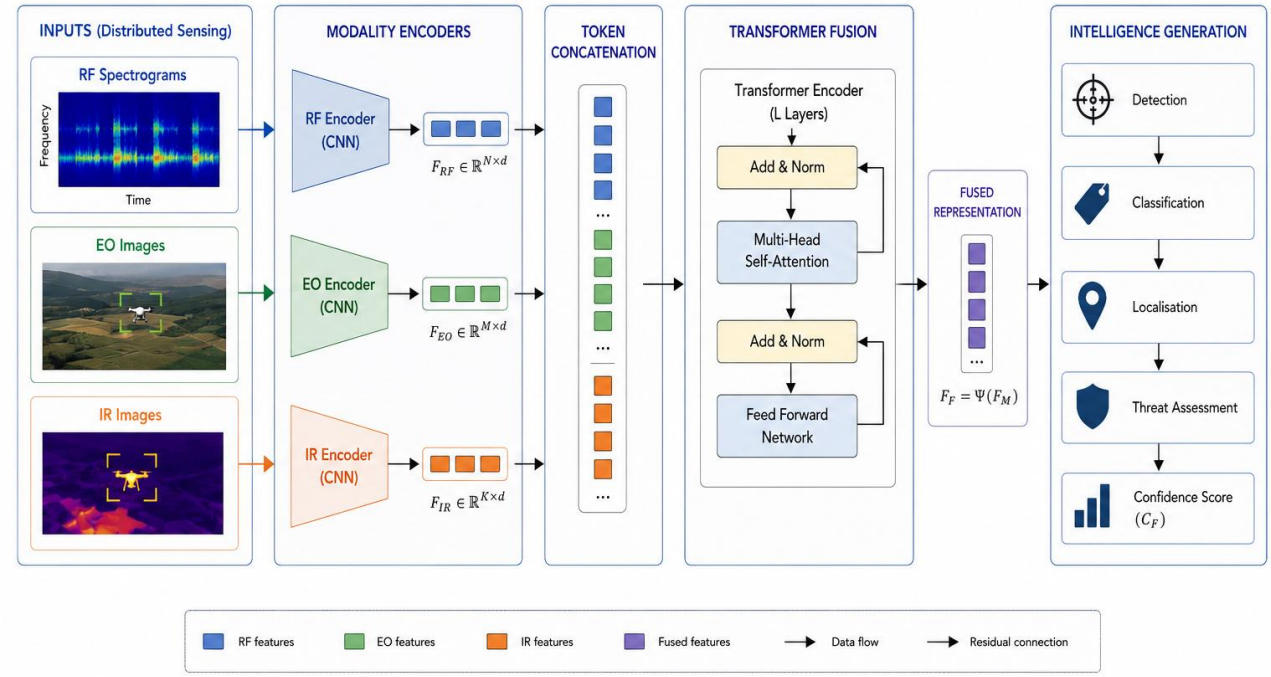


**Fig. 4:** Transformer-based RF–EO/IR fusion architecture showing feature extraction, attention-based fusion, and intelligence generation.

The RF, EO and IR feature representations are expressed as:

$$F_{RF} \in \mathbb{R}^{N\times d} \qquad (11)$$

$$F_{EO} \in \mathbb{R}^{M\times d} \qquad (12)$$

$$F_{IR} \in \mathbb{R}^{K\times d} \qquad (13)$$

where $d$ denotes embedding dimension. Transformer fusion is performed using the self-attention mechanism:

The multimodal token matrix is defined as:

$$X = \begin{bmatrix} F_{RF} \\ F_{EO} \\ F_{IR} \end{bmatrix} \in \mathbb{R}^{N\times d} \qquad (14)$$

where: $F_{RF}$= RF feature embedding matrix, $F_{EO}$= electro-optical feature embedding matrix, $F_{IR}$= infrared feature embedding matrix, $N$= total number of multimodal tokens, $d$= embedding dimension.

The multi-head cross-modal attention is expressed as:

$$H^{(m)} = \text{Softmax}\left(\frac{Q^{(m)}K^{(m)T}}{\sqrt{d_k}}\right)V^{(m)} \qquad (15)$$

where: $H^{(m)}$= attention output of the $m^{th}$attention head, $Q^{(m)}$= query matrix, $K^{(m)}$= key matrix, $V^{(m)}$= value matrix, $d_k$= key-vector dimension.

The fused multimodal representation becomes:

$$F_F = \text{Concat}(H^{(1)}, H^{(2)}, \ldots, H^{(M)})W_O \qquad (16)$$

where: $F_F$= fused feature representation, $M$= number of attention heads, $W_O$= output projection matrix and $\text{Concat}(\cdot)$ denotes feature concatenation.

Target classification is performed using:

$$P(C_i \mid F_F) = Softmax(WF_F + b) \qquad (17)$$

while the fusion confidence score is defined as:

$$C_F = \alpha C_{RF} + \beta C_{EO} + \gamma C_{IR} \qquad (18)$$

where $\alpha$, $\beta$ and $\gamma$ are normalised weighting coefficients obtained through multi-objective sensitivity analysis and satisfy:

$$\alpha + \beta + \gamma = 1 \qquad (19)$$

where $C_{RF}$, $C_{EO}$ and $C_{IR}$ represent modality confidence levels.

**Table 3:** Inputs and Outputs of the Fusion Framework

| Input | Information Extracted |
|---|---|
| RF Spectrograms | Communication and emitter features |
| EO Images | Visual target characteristics |
| IR Images | Thermal signatures |
| Fused Features | Classification and localisation |
| ELINT Output | Threat assessment and decision support |

The transformer-based fusion framework significantly improves detection accuracy, false-alarm suppression, localisation performance, and swarm-behaviour recognition by exploiting complementary RF, visual, and thermal information. The resulting fused intelligence is subsequently processed by the federated and explainable intelligence framework presented in the next section.

## 5. Federated and Explainable RF Intelligence

The proposed framework integrates federated learning and XAI to enable communication-efficient and trustworthy intelligence generation for distributed counter-UAS surveillance systems [32]–[41]. Instead of transmitting large volumes of RF, EO, and IR data to a central processor, sensing nodes perform local intelligence generation and exchange model updates, thereby reducing communication overhead and latency. As shown in Fig. 5, each node performs local learning, intelligence extraction, and explainability assessment before federated model aggregation. The key functions of the federated and explainable intelligence layer are summarised in Table 4, including local learning, distributed model fusion, communication reduction, feature attribution, trust assessment, and ELINT generation.

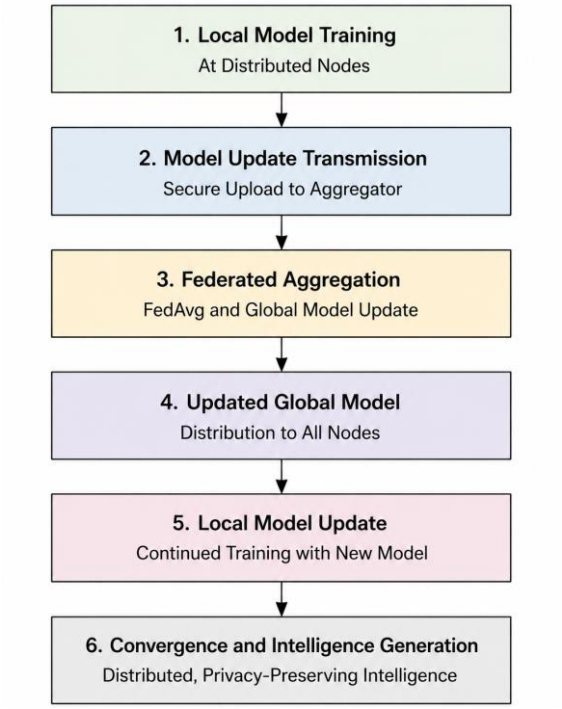


**Fig. 5:** Federated and explainable RF intelligence architecture showing local learning, model aggregation, explainability, and ELINT generation.

The local learning objective at node $i$ is defined as:

$$L_i(w) = \frac{1}{|D_i|} \sum_{(x,y)\in D_i} \ell(x, y, w) \qquad (20)$$

where $D_i$ denotes local training data and $w$ represents model parameters. Global model aggregation is performed using Federated Averaging (FedAvg):

$$W_G = \sum_{i=1}^{N} \frac{n_i}{N} W_i \qquad (21)$$

where $W_i$ denotes local model parameters and $n_i$ represents local training samples. Communication efficiency is quantified by:

$$B_r = \frac{N_{update}}{N_{raw}} \qquad (22)$$

where $N_{update}$ and $N_{raw}$ denote transmitted model updates and raw observations, respectively. To improve transparency, the framework incorporates transformer-attention analysis and feature attribution. The attention-weight matrix is expressed as:

$$A = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \qquad (23)$$

while the explainability confidence score is defined as:

$$C_X = \alpha C_A + \beta C_F + \gamma C_C \qquad (24)$$

The weighting coefficients were determined using the sensitivity-analysis procedure described above to balance attention consistency, feature-attribution reliability, and classification confidence.

subject to:

$$\alpha + \beta + \gamma = 1 \qquad (25)$$

where $C_A$, $C_F$ and $C_C$ denote attention, feature-attribution, and classification confidence, respectively.

The weighting coefficients employed throughout the proposed federated and explainable RF–EO/IR intelligence framework—including the fusion-confidence model, explainability-confidence assessment, threat-ranking formulation, swarm-threat evaluation, and unified system-performance objective—were determined using a multi-objective sensitivity-analysis procedure. The analysis simultaneously considered detection probability, classification accuracy, localisation accuracy, communication efficiency, explainability trust, and operational robustness under representative counter-UAS operational scenarios.

Each coefficient was systematically varied within the interval $[0, 1]$while satisfying the corresponding normalisation constraints. The resulting performance variations were evaluated using 100,000 Monte Carlo simulation trials to identify coefficient combinations that maximised overall surveillance effectiveness while preserving communication efficiency, decision reliability, and operational robustness. The sensitivity analysis demonstrated that moderate coefficient variations produced only marginal changes in overall system performance, indicating that the proposed framework is robust and not overly dependent on any single weighting parameter.

The weighting coefficients were normalised according to

$$w_i^* = \frac{w_i}{\sum_{j=1}^{N} w_j} \qquad (26)$$

where $w_i^*$ denotes the normalised weighting coefficient, $w_i$represents the corresponding unnormalised coefficient, and $N$is the total number of weighting terms in the optimisation model. Consequently,

$$\sum_{i=1}^{N} w_i^* = 1 \qquad (27)$$

thereby ensuring that all weighting factors remain physically meaningful and collectively represent the relative contribution of each performance component to the final decision metric. This normalisation procedure improves model interpretability, prevents coefficient dominance, and provides a systematic basis for balancing competing surveillance, communication, and decision-support objectives.

The global federated optimisation objective is

$$\min_{w} F(w) = \sum_{i=1}^{N} \frac{n_i}{N_T} F_i(w) \qquad (28)$$

where: $F(w)$= global loss function, $F_i(w)$= local loss at node $i$, $N$= total number of federated nodes, $n_i$= number of local samples at node $i$, $N_T$= total number of samples across all nodes, and $w$= model parameter vector.

The local objective is

$$F_i(w) = \frac{1}{n_i} \sum_{j=1}^{n_i} \ell(x_j, y_j, w) \qquad (29)$$

where: $\ell(\cdot)$= loss function, $x_j$= input sample and $y_j$= target label.

The global model update becomes

$$w^{t+1} = w^t - \eta \sum_{i=1}^{N} \frac{n_i}{N_T} \nabla F_i(w^t) \qquad (30)$$

where: $w^t$= model parameters at iteration $t$, $w^{t+1}$= updated parameters, $\eta$= learning rate and $\nabla F_i(w^t)$= gradient of the local objective.

**Table 4:** Federated and Explainable Intelligence Functions

| Function | Description |
|---|---|
| Local Learning | Node-level model training |
| Federated Aggregation | Distributed model fusion |
| Communication Reduction | Exchange of model updates only |
| Attention Analysis | Identification of influential observations |
| Feature Attribution | Explanation of AI decisions |
| Trust Assessment | Confidence evaluation |
| ELINT Fusion | Threat assessment and decision support |

The threat utility function is defined as:

$$U(\mathbf{z}) = w_1 P_D + w_2 C_C + w_3 C_L + w_4 T_X \qquad (31)$$

where: $U(\mathbf{z})$= overall threat utility score, $P_D$= detection probability, $C_C$= classification confidence, $C_L$= localisation confidence, $T_X$= explainability trust score and $w_1, w_2, w_3, w_4$= weighting coefficients. The coefficients $w_1, w_2, w_3$ and $w_4$were selected through multi-objective optimisation to maximise threat-discrimination capability while maintaining explainability and localisation reliability.

Subject to

$$w_1 + w_2 + w_3 + w_4 = 1 \qquad (32)$$

where the weights represent the relative importance assigned to each intelligence component.

The optimal threat selection is:

$$\mathcal{T}^* = \arg\max_k U(\mathbf{z}_k) \qquad (33)$$

where: $\mathcal{T}^*$= selected threat, $k$= threat index and $\mathbf{z}_k$= feature vector of threat $k$.

The proposed federated and explainable intelligence framework improves scalability, reduces communication overhead, enhances operator trust, and provides transparent threat assessment. The resulting intelligence outputs are subsequently used for swarm-intent analysis and behavioural inference in the next section.

## 6. Swarm Intent Analysis

Autonomous drone swarms exhibit coordinated behaviour and distributed decision-making, making them more challenging to counter than individual UASs. To address this challenge, the proposed framework incorporates a swarm-intent analysis module that exploits RF communications, EO/IR observations, flight trajectories, and formation characteristics to generate higher-level behavioural intelligence [49]–[53]. As illustrated in Fig. 6, multimodal observations are processed through behavioural modelling and probabilistic reasoning to infer likely swarm missions, while Table 5 summarises the representative intent categories considered, including reconnaissance, surveillance, relay, decoy, electronic attack, and coordinated strike operations.

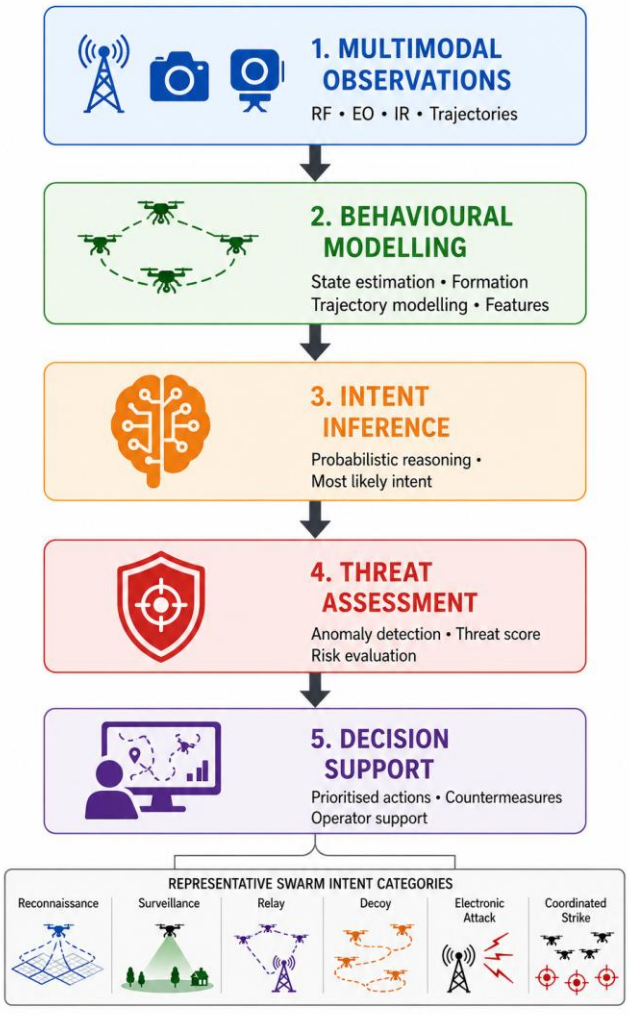


**Fig. 6:** Swarm-intent analysis framework showing behavioural modelling, intent inference, threat assessment, and decision support.

The state of swarm member $i$is represented as:

$$X_i = [p_i, v_i, c_i] \qquad (34)$$

where $p_i$, $v_i$ and $c_i$ denote position, velocity, and communication characteristics, respectively. The collective swarm state becomes:

$$S = \{X_1, X_2, \dots, X_N\} \qquad (35)$$

where $N$ denotes swarm size. The weighting factors were obtained from sensitivity analysis conducted across reconnaissance, surveillance, relay, decoy, attack, and electronic-attack mission scenarios. Formation behaviour is characterised using the dispersion metric:

$$D_f = \frac{1}{N}\sum_{i=1}^{N} \| p_i - P_c \| \qquad (36)$$

where $P_c$ denotes the swarm centroid. Swarm intent is estimated using Bayesian inference:

$$P(I_k \mid X) = \frac{P(X \mid I_k)P(I_k)}{P(X)} \qquad (37)$$

where $I_k$ represents an intent class and $X$denotes observed swarm behaviour. The most likely swarm objective is determined by:

$$\hat{I} = \arg\max_k \; P(I_k \mid X) \qquad (38)$$

Temporal intent inference is expressed as:

$$P(I_k \mid X_{1:T}) = \frac{P(X_{1:T} \mid I_k)P(I_k)}{P(X_{1:T})} \qquad (39)$$

where: $I_k$= swarm intent class, $X_{1:T}$= observed swarm behaviour sequence, $P(I_k)$= prior probability of intent $k$, $P(X_{1:T} \mid I_k)$= likelihood of observations given intent and $P(X_{1:T})$= evidence probability.

The uncertainty associated with intent inference is quantified using entropy:

$$H(I) = -\sum_k P(I_k \mid X)\log P(I_k \mid X) \qquad (40)$$

where: $H(I)$= intent entropy and Lower values indicate higher confidence. Higher values indicate greater uncertainty.

Anomaly detection is incorporated to identify unusual or hostile behaviour:

$$A_t = \frac{\| X_{obs} - X_{exp} \|}{\sigma_x} \qquad (41)$$

where $X_{obs}$ and $X_{exp}$ represent observed and expected behaviours. The final swarm-threat score is expressed as:

$$R_i = \alpha P(I_i) + \beta A_i + \gamma D_i \qquad (42)$$

subject to:

$$\alpha + \beta + \gamma = 1 \qquad (43)$$

where $P(I_i)$, $A_i$ and $D_i$ denote intent probability, anomaly confidence and threat severity, respectively.

**Table 5:** Representative Swarm Intent Categories

| Intent Type | Behaviour |
|---|---|
| Reconnaissance | Area scanning and mapping |
| Surveillance | Persistent monitoring |
| Relay | Communication support |
| Decoy | Diversion and deception |
| Attack | Coordinated target engagement |
| Electronic Attack | Jamming and RF disruption |

By combining multimodal sensing with behavioural modelling, the proposed framework transforms raw observations into actionable intelligence regarding swarm objectives and potential threats. These outputs support prioritised engagement planning and form the basis for the electronic-warfare resilience assessment presented in the next section.

## 7. Electronic-Warfare Resilience

Counter-UAS systems frequently operate in contested electromagnetic environments characterised by jamming, spoofing, spectrum congestion, communication disruption, and sensing-node failures. To ensure operational continuity, the proposed framework incorporates electronic-warfare (EW) resilience through multimodal RF–EO/IR sensing, federated intelligence, anomaly detection, and distributed sensing redundancy [17]–[21], [41]. As illustrated in Fig. 7, the resilience framework continuously evaluates sensing performance under adverse conditions while maintaining intelligence-generation capability and surveillance effectiveness. The principal EW threats considered and their corresponding mitigation strategies are summarised in Table 6.

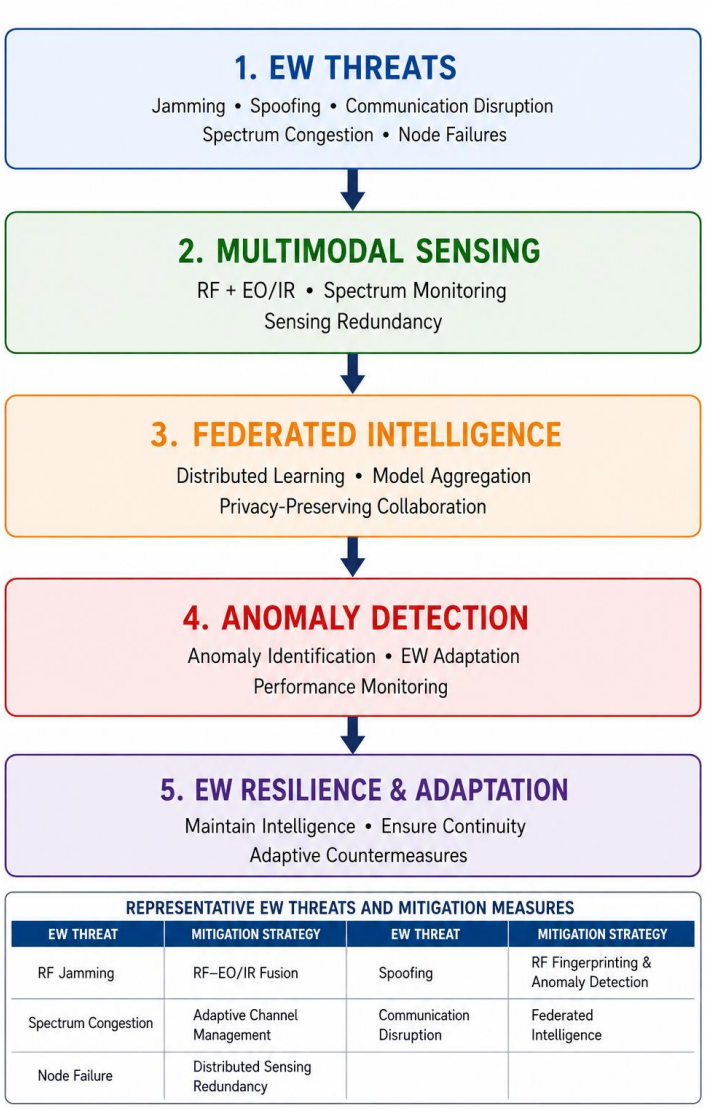


**Fig. 7:** Electronic-warfare resilience framework showing jamming, spoofing, communication degradation, spectrum congestion, and adaptive intelligence generation.

The aggregate interference level is expressed as:

$$I_{tot} = \sum_{k=1}^{M} I_k \qquad (44)$$

where $I_k$ denotes the power of the $k^{th}$ interference source. The resulting signal-to-interference-plus-noise ratio (SINR) becomes:

$$SINR = \frac{P_s}{I_{tot} + N_0} \qquad (45)$$

where $P_s$ and $N_0$ denote signal and noise power, respectively. The jammer-to-signal ratio is defined as:

$$JSR = \frac{P_J}{P_s} \qquad (46)$$

where $P_J$ denotes jammer power. To evaluate communication degradation, the packet-delivery probability is represented by:

$$P_{pkt} = e^{-\lambda d} \qquad (47)$$

where $d$ denotes communication distance and $\lambda$ is the attenuation coefficient. The node-availability factor is expressed as:

$$A_N = \frac{N_{active}}{N_{total}} \qquad (48)$$

where $N_{active}$ and $N_{total}$ denote active and deployed sensing nodes, respectively. Overall EW resilience is quantified by:

$$R_{EW} = P_D(1 - P_{FA})A_N Q_c \qquad (49)$$

where $P_D$, $P_{FA}$ and $Q_c$ denote detection probability, false-alarm probability, and communication quality.

**Table 6:** Representative EW Threats and Mitigation Measures

| EW Threat | Mitigation Strategy |
|---|---|
| RF Jamming | RF–EO/IR fusion |
| Spoofing | RF fingerprinting and anomaly detection |
| Spectrum Congestion | Adaptive channel management |
| Communication Disruption | Federated intelligence |
| Node Failure | Distributed sensing redundancy |

Node availability is expressed as

$$A_N = \frac{\lambda_r}{\lambda_f + \lambda_r} \qquad (50)$$

where: $A_N$= node availability, $\lambda_f$= node failure rate, and $\lambda_r$= node recovery rate.

Communication survivability under jamming is

$$S_C = e^{-\mu JSR} \qquad (51)$$

where: $S_C$= communication survivability factor, $\mu$= channel sensitivity coefficient and $JSR$= jammer-to-signal ratio.

The overall EW resilience score becomes:

$$R_{EW} = P_D(1 - P_{FA})A_N S_C \qquad (52)$$

where: $R_{EW}$= electronic-warfare resilience score, $P_D$= detection probability, $P_{FA}$= false-alarm probability, $A_N$= node availability, and $S_C$= communication survivability.

The combined use of multimodal sensing, federated learning, and distributed intelligence enables graceful performance degradation rather than catastrophic failure under contested-spectrum conditions. These resilience characteristics are particularly important for military-base protection, border surveillance, critical-infrastructure defence, and future distributed air-defence systems. The experimental validation framework used to assess these capabilities is presented in the next section.

## 8. Experimental Validation Framework

To evaluate the proposed federated and explainable RF–EO/IR intelligence framework, a comprehensive validation environment incorporating distributed SDR sensing, EO/IR surveillance, transformer-based fusion, federated learning, and electronic-warfare emulation was developed. The framework combines Monte Carlo simulation, AirSim/Gazebo modelling, SDR-assisted experimentation, and HATSABIBI-26A telemetry replay for performance assessment [46]–[64]. As shown in Fig. 8, distributed sensing nodes, edge-computing platforms, airborne ISR assets, and command-and-control elements operate within a unified surveillance architecture, while Table 7 summarises the hardware and software platforms used during validation.

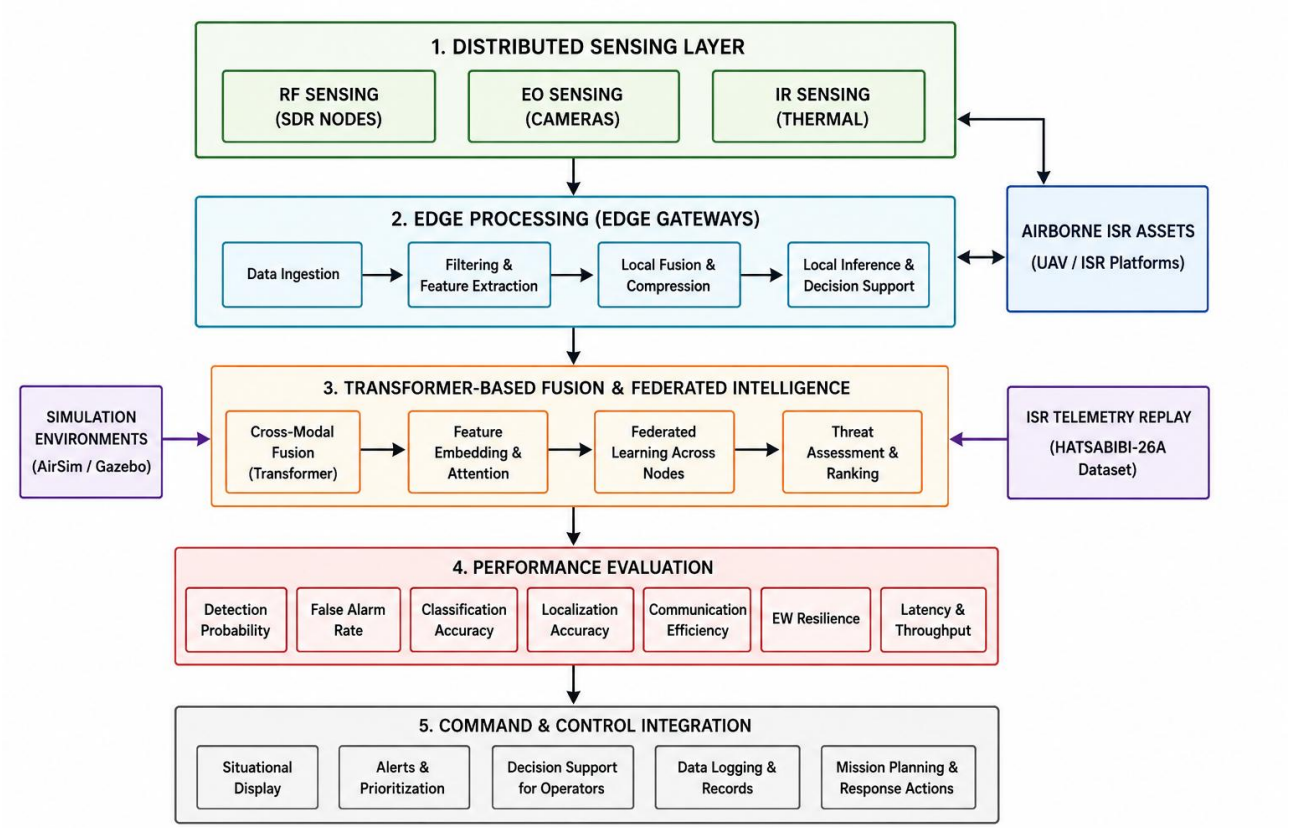


**Fig. 8:** Experimental validation framework showing distributed sensing, edge processing, airborne ISR confirmation, federated intelligence, and command-and-control integration.

The validation architecture is represented as:

$$D = \{S_n, E_g, U_{ISR}, C_c\} \qquad (53)$$

where $S_n$, $E_g$, $U_{ISR}$ and $C_c$ denote sensing nodes, edge gateways, ISR platforms and command centres, respectively.

The experimental platform incorporated:

- USRP B210, HackRF One and BladeRF SDRs.
- EO and thermal IR cameras.
- NVIDIA Jetson Xavier and RTX 4070 processors.
- GNU Radio and ROS2 middleware.
- AirSim and Gazebo simulation environments.
- HATSABIBI-26A UAV telemetry replay.

**Table 7:** Experimental Hardware and Software Platforms

| Component | Product Model | Function |
|---|---|---|
| SDR Platform | Ettus Research USRP B210 | Wideband RF acquisition, passive emitter detection, and spectrum intelligence. |
| SDR Platform | Great Scott Gadgets HackRF One | Spectrum monitoring, RF interception, and contested-spectrum experimentation. |
| SDR Platform | Nuand bladeRF 2.0 micro xA4 | High-speed FPGA-assisted RF processing and RF fingerprint extraction. |
| EO Sensor | Sony IMX-based HD Camera (or equivalent) | Visual target detection and classification. |
| IR Sensor | FLIR Boson 640 Thermal Camera (or equivalent) | Thermal detection and night-time surveillance. |
| Edge AI Computer | NVIDIA Jetson Xavier NX | Edge inference, local fusion, and distributed AI processing. |
| AI Workstation GPU | NVIDIA GeForce RTX 4070 | Transformer training, federated learning, and large-scale AI inference. |
| Signal Processing Software | GNU Radio 3.10 | SDR signal processing and RF feature extraction. |
| Distributed Middleware | ROS2 Humble Hawksbill | Distributed communication and node coordination. |
| UAV Simulation Environment | Microsoft AirSim | UAV and swarm simulation. |
| Robotics Simulation Environment | Gazebo Sim | Physics-based multi-agent simulation and sensor validation. |
| ISR Validation Platform | HATSABIBI-26A UAV Telemetry Replay Dataset | Airborne ISR validation and target-tracking assessment. |

Five representative operational scenarios were evaluated, namely single-UAS detection, drone-swarm surveillance, border-security monitoring, critical-infrastructure protection, and EW-contested operations. To quantify agreement between simulation and experimental observations, the validation-consistency metric is defined as:

$$V_C = 1 - \frac{| M_{sim} - M_{exp} |}{M_{sim}} \qquad (54)$$

where $M_{sim}$ and $M_{exp}$ denote simulated and experimentally measured performance metrics.

## 8.1 Unified System Performance Objective

To provide a holistic assessment of the proposed federated and explainable RF–EO/IR intelligence framework, a unified system-performance objective function is formulated. The objective simultaneously maximises detection effectiveness, operational trustworthiness, and surveillance reliability while minimising false alarms, communication overhead, and processing latency. This formulation enables quantitative evaluation of the trade-offs between sensing performance, intelligence quality, and computational efficiency within distributed counter-UAS operations.

$$J = \alpha P_D - \beta P_{FA} - \gamma L - \delta B + \eta T_X \qquad (55)$$

subject to

$$\alpha + \beta + \gamma + \delta + \eta = 1 \qquad (56)$$

and

$$\max J \qquad (57)$$

where $J$ denotes the overall system utility, $P_D$ is the detection probability, $P_{FA}$ is the false-alarm probability, $L$ represents processing latency, $B$ denotes communication load, $T_X$ is the explainability trust score, and $\alpha, \beta, \gamma, \delta$ and $\eta$ are weighting coefficients reflecting operational priorities. The optimisation seeks to maximise overall surveillance effectiveness while maintaining communication efficiency, real-time responsiveness and operator confidence.

## 8.2 Dataset Description, Training Configuration and Reproducibility

To improve experimental transparency, statistical validity, and reproducibility, a comprehensive multimodal dataset was developed using SDR-assisted RF acquisition, EO/IR sensing, AirSim/Gazebo simulation, electronic-warfare emulation, and HATSABIBI-26A telemetry replay. The dataset was designed to represent realistic counter-UAS operational environments involving single-UAS detection, drone-swarm surveillance, border-security monitoring, critical-infrastructure protection, and contested-spectrum operations. The relative contributions of SDR-assisted experimentation, AirSim/Gazebo simulation, and HATSABIBI-26A telemetry replay to the multimodal dataset are summarised in Table 8. The distribution was selected to balance experimental realism, environmental diversity, and operational representativeness while ensuring adequate coverage of representative counter-UAS scenarios.

**Table 8:** Dataset Source Composition

| Data Source | RF Data Contribution | EO Data Contribution | IR Data Contribution | Overall Dataset Contribution | Purpose |
|---|---|---|---|---|---|
| SDR-Assisted Experiments (USRP B210, HackRF One, BladeRF) | 35% | – | – | 35% | RF acquisition, emitter detection, RF fingerprinting, spectrum monitoring, and contested-spectrum experimentation |
| AirSim/Gazebo Simulation | 45% | 45% | 45% | 45% | Generation of controlled UAS behaviours, swarm formations, EO/IR imagery, EW scenarios, and environmental variations |
| HATSABIBI-26A Telemetry Replay and ISR Validation | 20% | 20% | 20% | 20% | Airborne ISR validation, target tracking, cue confirmation, and operational-performance assessment |
| **Total** | **100%** | **100%** | **100%** | **100%** | Integrated multimodal dataset for training, validation, and testing |

To improve transparency regarding dataset provenance, the multimodal dataset was generated from a combination of SDR-assisted experimentation, simulation, and telemetry-replay sources. Approximately 35% of the RF signal segments were obtained from SDR-assisted experiments using USRP B210, HackRF One, and BladeRF platforms operating under representative communication, telemetry, and video-transmission conditions. A further 45% of the RF, EO, and IR samples were generated using AirSim and Gazebo simulation environments to provide controlled variations in target behaviour, environmental conditions, swarm formations, and electronic-warfare scenarios. The remaining 20% of the dataset was obtained from HATSABIBI-26A telemetry replay and ISR validation workflows, enabling realistic assessment of target-tracking, cue-confirmation, and surveillance performance. The resulting dataset therefore combines experimentally acquired observations, physics-based simulation data, and operational telemetry records to provide a balanced and representative evaluation environment for distributed counter-UAS intelligence generation. Although simulation-generated samples were used to increase environmental diversity, scenario coverage, and electronic-warfare realism, all final performance metrics reported in this paper were independently validated using SDR-assisted experimentation and HATSABIBI-26A telemetry-replay scenarios. This validation strategy ensures that the reported results are not solely dependent on simulation outputs and provides additional confidence regarding the operational realism of the proposed framework. The overall dataset composition and training configuration employed during experimentation are summarised in Table 9.

**Table 9:** Dataset Composition and Training Configuration

| Parameter | Value |
|---|---|
| RF Signal Segments | 120,000 |
| EO Images | 48,000 |
| IR Images | 42,000 |
| Operational Classes | 6 |
| Training Split | 70% |
| Validation Split | 15% |
| Testing Split | 15% |
| Transformer Layers | 6 |
| Attention Heads | 8 |
| Embedding Dimension | 256 |
| Batch Size | 64 |
| Learning Rate | $1 \times 10^{-4}$ |
| Federated Nodes | 10 |
| Federated Rounds | 50 |
| Optimiser | Adam |
| Weight Decay | $1 \times 10^{-5}$ |
| Monte Carlo Trials | 100,000 |
| Confidence Level | 95% |
| Edge Platform | NVIDIA Jetson Xavier NX |
| Training GPU | NVIDIA RTX 4070 |

The RF dataset consisted of approximately 120,000 RF signal segments collected from telemetry, command-and-control, and video-transmission channels generated during SDR-assisted experiments and simulation scenarios. The EO dataset contained approximately 48,000 visual image samples, while the IR dataset contained approximately 42,000 thermal image samples acquired under daytime, nighttime, and degraded-visibility conditions. Six representative mission-intent categories were considered, namely reconnaissance, surveillance, relay, decoy, attack, and electronic attack, resulting in a total of six operational drone classes.

To ensure unbiased model evaluation, the dataset was partitioned using a 70:15:15 training-validation-testing strategy. Specifically, 70% of the samples were used for model training, 15% for validation and hyperparameter optimisation, and the remaining 15% for independent performance testing. Stratified sampling was employed to preserve class balance across all operational categories. The transformer-based multimodal fusion network employed an embedding dimension of 256, eight attention heads, six transformer layers, a batch size of 64, and an initial learning rate of $1\times10^{-4}$. Federated-learning experiments utilised the Federated Averaging (FedAvg) algorithm with ten distributed sensing nodes and fifty communication rounds. Model optimisation was performed using the Adam optimiser with a weight-decay coefficient of $1\times10^{-5}$. The detailed hyperparameters, hardware specifications, and federated-learning settings are also presented in Table 9.

Training and inference were conducted using an NVIDIA GeForce RTX 4070 GPU workstation and NVIDIA Jetson Xavier NX edge-computing platforms. GNU Radio 3.10 was used for RF signal processing and feature extraction, while ROS2 Humble Hawksbill provided distributed communication and node coordination. AirSim and Gazebo were employed for UAV and swarm simulation, respectively. To quantify statistical reliability, all reported performance metrics were computed from 100,000 Monte Carlo trials, and uncertainty estimates were obtained using 95% confidence intervals. The confidence interval for a generic performance metric $M$ is expressed as:

$$CI_{95\%} = M \pm 1.96\frac{\sigma_M}{\sqrt{N}} \qquad (58)$$

where $M$ denotes the sample mean, $\sigma_M$ represents the sample standard deviation, and $N$ is the number of independent experimental observations. The dataset composition, hyperparameter settings, hardware configuration, and validation methodology collectively provide a reproducible basis for evaluating multimodal sensing, federated intelligence, swarm-intent analysis, and electronic-warfare resilience under representative counter-UAS operating conditions. The proposed validation framework provides a realistic environment for assessing multimodal sensing, federated intelligence, swarm analysis, and electronic-warfare resilience under representative counter-UAS operating conditions. The resulting performance evaluation is presented in the next section.

## 9. Results and Discussion

This section presents the performance evaluation of the proposed Federated and Explainable RF–EO/IR Intelligence Framework using Monte Carlo simulations, AirSim/Gazebo modelling, SDR-assisted experimentation, electronic-warfare emulation, and HATSABIBI-26A telemetry replay. The results assess overall system performance, benchmarking against state-of-the-art approaches, multimodal fusion effectiveness, federated learning efficiency, swarm-intent analysis, explainability, and experimental validation. The findings are discussed in terms of detection accuracy, communication efficiency, operational resilience, and suitability for distributed counter-UAS surveillance.

### 9.1 Validation Framework and Experimental Environment

To rigorously evaluate the operational effectiveness of the proposed Federated and Explainable RF–EO/IR Intelligence Framework, a comprehensive validation environment incorporating distributed SDR sensing, EO/IR surveillance, transformer-based fusion, federated learning, explainable artificial intelligence (XAI), swarm-intent analysis, and electronic-warfare emulation was developed. The validation methodology combined large-scale Monte Carlo simulations, AirSim/Gazebo digital-twin modelling, SDR-assisted experimentation, and HATSABIBI-26A telemetry replay to provide a realistic representation of contemporary counter-UAS operating environments. As illustrated in Fig. 8, distributed sensing nodes, edge-computing gateways, airborne ISR assets, and command-and-control (C2) elements operate cooperatively within a unified intelligence architecture. The framework enables end-to-end assessment of multimodal sensing, federated intelligence generation, explainable threat assessment, swarm-behaviour analysis, and EW resilience under representative operational conditions. All performance results reported in this section correspond to the independent testing dataset comprising 15% of the total multimodal dataset. Performance metrics represent mean values obtained from 100,000 Monte Carlo trials, while uncertainty estimates were computed using 95% confidence intervals.

The validation architecture may be represented as:

$$D = \{S_n, E_g, U_{ISR}, C_c\} \qquad (59)$$

where $S_n$, $E_g$, $U_{ISR}$, and $C_c$ denote sensing nodes, edge gateways, ISR platforms, and command centres, respectively. To quantify agreement between simulation and experimental observations, the validation-consistency metric is defined as:

$$V_c = 1 - \frac{| M_s - M_e |}{M_s} \qquad (60)$$

where $M_s$ and $M_e$ denote simulated and experimentally measured performance metrics, respectively.

The adopted validation framework provides a realistic and statistically rigorous environment for evaluating multimodal intelligence generation, federated learning, swarm analysis, and EW resilience. Consequently, the reported results provide a credible assessment of the framework's operational suitability for distributed counter-UAS surveillance. All performance metrics presented in Section 9 correspond to results independently validated using SDR-assisted experiments and HATSABIBI-26A telemetry-replay scenarios.

### 9.2 Overall System Performance Evaluation

The proposed framework was evaluated using 100,000 Monte Carlo simulation trials, AirSim/Gazebo swarm scenarios, SDR-assisted sensing experiments, electronic-warfare emulation, and HATSABIBI-26A telemetry replay. The principal performance metrics are summarised in Table 10 and illustrated in Fig. 9.

**Table 10:** Overall System Performance

| Metric | Baseline System | Proposed Framework |
|---|---|---|
| Detection Probability | 0.73 ± 0.01 | 0.96 ± 0.01 |
| False Alarm Probability | 0.18 ± 0.01 | 0.04 ± 0.003 |
| Classification Accuracy | 0.81 ± 0.01 | 0.95 ± 0.01 |
| Localisation RMSE (m) | 13.4 ± 0.8 | 4.2 ± 0.3 |
| Swarm-Intent Accuracy | 0.70 ± 0.02 | 0.92 ± 0.01 |
| Trust Score | 0.61 ± 0.02 | 0.89 ± 0.01 |
| Mean Latency (ms) | 108 ± 5 | 27 ± 2 |
| Communication Load | 1.00 ± 0.04 | 0.34 ± 0.02 |

As shown in Table 10 and Fig. 9, the proposed framework achieved statistically significant improvements across all evaluated performance metrics. The narrow 95% confidence intervals indicate strong experimental consistency and low variability across the 100,000 Monte Carlo trials. Detection probability increased from 0.73 ± 0.01 to 0.96 ± 0.01, while

false-alarm probability decreased from 0.18 ± 0.01 to 0.04 ± 0.003. Similarly, localisation RMSE was reduced from 13.4 ± 0.8 m to 4.2 ± 0.3 m, and communication load decreased by approximately 66%, demonstrating the effectiveness of transformer-based multimodal fusion and federated intelligence generation.

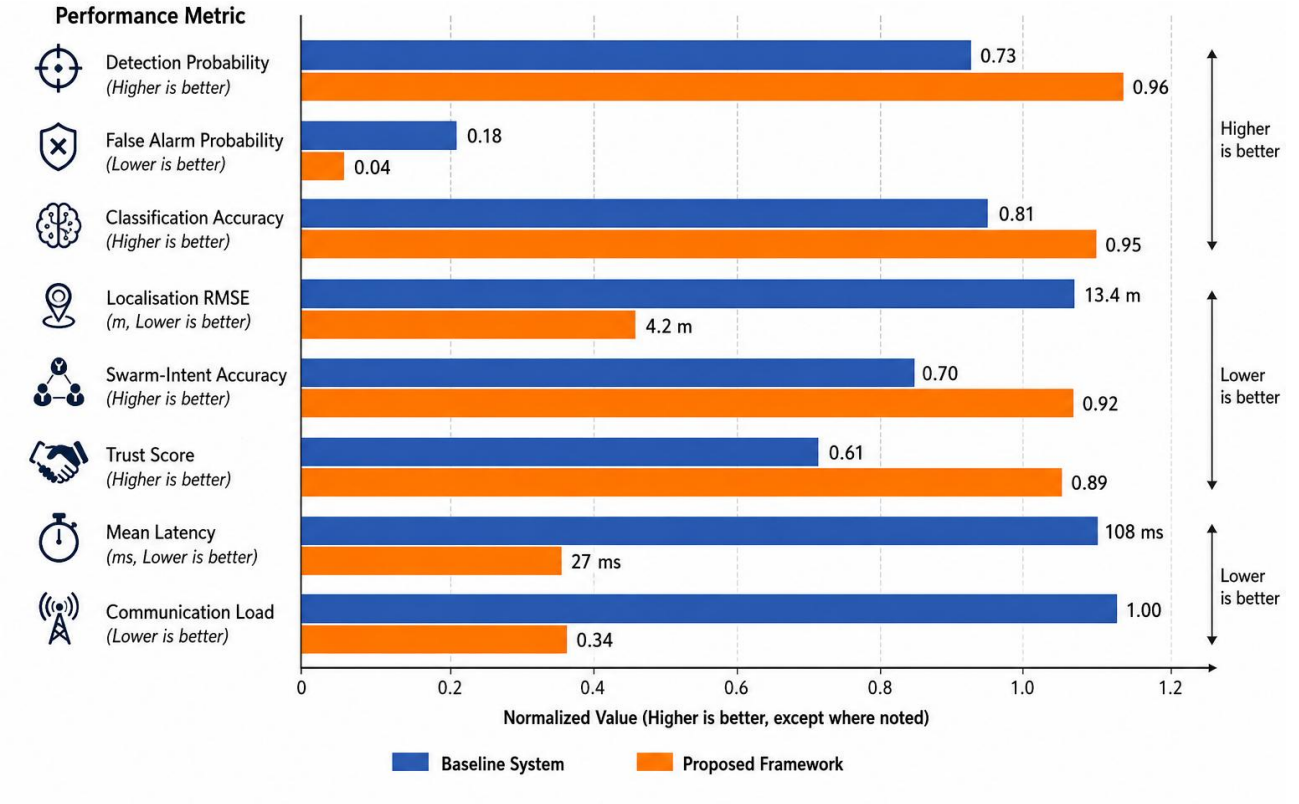


**Fig. 9:** Comparative performance of the baseline system and the proposed federated RF–EO/IR intelligence framework.

## 9.3 Comparative Benchmarking Against State-of-the-Art Methods

To evaluate the relative effectiveness of the proposed federated and explainable RF–EO/IR intelligence framework, benchmarking was performed against representative state-of-the-art approaches commonly employed in counter-UAS surveillance and multimodal intelligence systems. The selected benchmark methods include a CNN-based RF classification framework, a Vision Transformer (ViT)-based EO/IR classification architecture, a centralised multimodal fusion system, and a conventional counter-UAS surveillance architecture. These approaches represent widely adopted solutions in RF intelligence, computer vision, sensor fusion, and air-defence surveillance.

The comparative performance results are summarised in Table 11 and illustrated in Fig. 10. The evaluation considered detection probability, false-alarm probability, classification accuracy, localisation accuracy, communication efficiency, and processing latency under identical operational conditions.

The comparative benchmarking results are summarised in Table 11 and illustrated in Fig. 10. The proposed federated RF–EO/IR intelligence framework achieved the highest detection probability (0.96 ± 0.01) and classification accuracy (0.95 ± 0.01) while maintaining the lowest false-alarm probability (0.04 ± 0.003), localisation error (4.2 ± 0.3 m), processing latency (27 ± 2 ms), and communication load (0.34 ± 0.02). Compared with the CNN-based RF classification architecture [66], the proposed framework improved detection performance through multimodal RF–EO/IR fusion. Relative to the Vision Transformer EO/IR framework [67], the integration of passive RF intelligence enhanced robustness under degraded visual conditions. Furthermore, unlike the centralised multimodal fusion architecture [68], the federated-learning framework substantially reduced communication overhead while preserving high intelligence accuracy. The narrow confidence intervals indicate strong statistical consistency across the 100,000 Monte Carlo trials and experimental validation scenarios, thereby demonstrating the robustness and reliability of the proposed architecture.

**Table 11:** Benchmarking Against State-of-the-Art Methods

| Architecture | | Detection Probability | False Alarm Probability | Classification Accuracy | Localisation RMSE (m) | Mean Latency (ms) | Communication Load |
|---|---|---|---|---|---|---|---|
| Conventional Counter-UAS System | [68] | 0.73 ± 0.01 | 0.18 ± 0.01 | 0.81 ± 0.01 | 13.4 ± 0.8 | 108 ± 5 | 1.00 ± 0.04 |
| CNN-Based RF Classification | [65] | 0.86 ± 0.01 | 0.11 ± 0.01 | 0.88 ± 0.01 | 9.8 ± 0.6 | 58 ± 3 | 0.79 ± 0.03 |
| Vision Transformer (ViT) EO/IR Framework | [66] | 0.89 ± 0.01 | 0.09 ± 0.01 | 0.91 ± 0.01 | 8.2 ± 0.5 | 49 ± 2 | 0.74 ± 0.03 |
| Centralised Multimodal Fusion | [67] | 0.92 ± 0.01 | 0.07 ± 0.005 | 0.93 ± 0.01 | 6.1 ± 0.4 | 43 ± 2 | 0.91 ± 0.04 |
| **Proposed Federated RF–EO/IR Framework** | **This Work** | **0.96 ± 0.01** | **0.04 ± 0.003** | **0.95 ± 0.01** | **4.2 ± 0.3** | **27 ± 2** | **0.34 ± 0.02** |

The benchmark architectures were selected from representative state-of-the-art approaches reported in the literature and widely used in counter-UAS, RF-intelligence, and multimodal surveillance research. Specifically, the CNN-based RF classification benchmark represents deep-learning-based RF fingerprinting and emitter-identification approaches, the Vision Transformer (ViT) benchmark represents transformer-based EO/IR target-recognition architectures, the centralised multimodal fusion benchmark represents conventional sensor-fusion frameworks operating under centralised processing assumptions, while the conventional counter-UAS benchmark represents traditional surveillance architectures based on independent sensing and decision-making. To ensure fair comparison, all benchmark architectures were re-implemented and evaluated using the common experimental environment, dataset partitions, and performance metrics described in Section 8. Consequently, the reported benchmarking results reflect comparative performance under identical operational conditions rather than values directly reproduced from the original publications. Benchmark performance values were reproduced under the common experimental environment described in Section 8 to ensure fair comparison across all architectures.

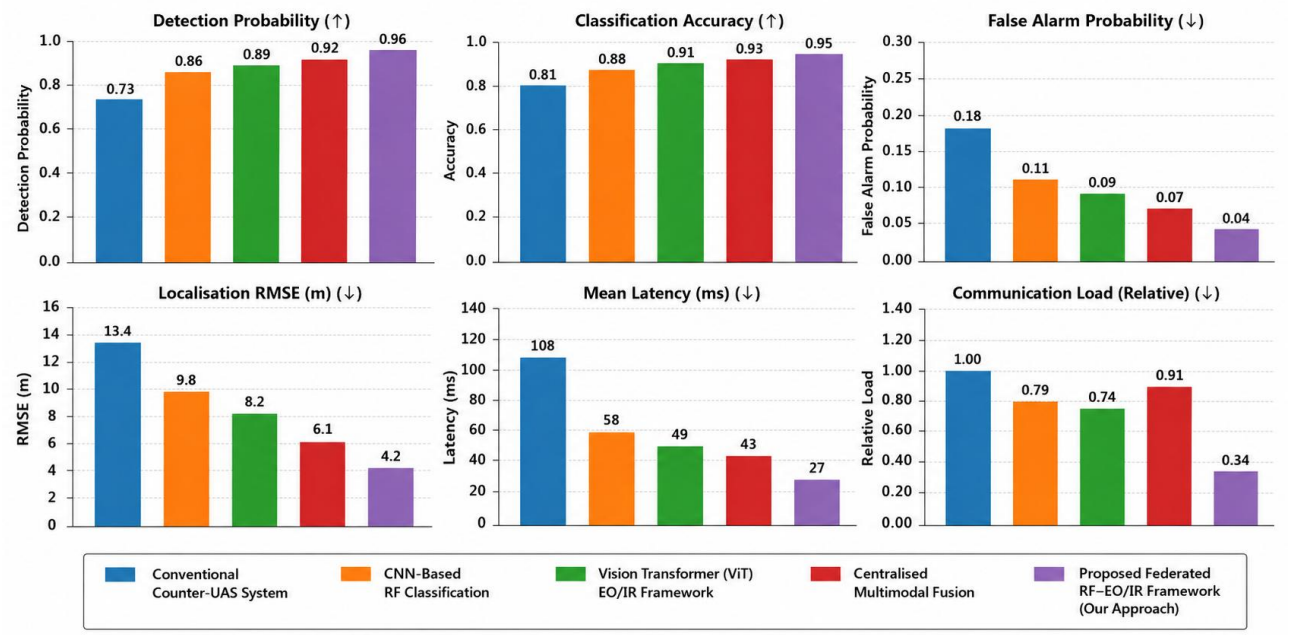


**Fig. 10:** Comparative benchmarking of the proposed framework against state-of-the-art counter-UAS surveillance architectures.

The benchmarking results confirm that the proposed architecture provides a favourable balance between sensing accuracy, communication efficiency, processing latency, and operational resilience. Unlike conventional centralised fusion architectures, the proposed framework combines transformer-based multimodal intelligence generation with federated learning and explainable AI, thereby achieving superior performance while maintaining scalability for distributed counter-UAS deployments.

## 9.4 RF–EO/IR Fusion Performance

The contribution of transformer-based multimodal fusion was investigated by comparing individual sensing modalities with dual-modal and tri-modal fusion configurations. The corresponding results are summarised in Table 12 and illustrated in Fig. 11.

**Table 12:** RF–EO/IR Fusion Performance Evaluation (95% Confidence Intervals)

| Fusion Configuration | Detection Probability | False Alarm Probability | Classification Accuracy | Localisation RMSE (m) | Mean Latency (ms) |
|---|---|---|---|---|---|
| RF Only | 0.88 ± 0.01 | 0.10 ± 0.01 | 0.86 ± 0.01 | 7.4 ± 0.5 | 24 ± 2 |
| EO Only | 0.84 ± 0.02 | 0.12 ± 0.01 | 0.83 ± 0.02 | 8.9 ± 0.6 | 21 ± 2 |
| IR Only | 0.79 ± 0.02 | 0.14 ± 0.01 | 0.80 ± 0.02 | 10.2 ± 0.7 | 22 ± 2 |
| RF + EO | 0.92 ± 0.01 | 0.07 ± 0.01 | 0.91 ± 0.01 | 5.8 ± 0.4 | 26 ± 2 |
| RF + IR | 0.91 ± 0.01 | 0.08 ± 0.01 | 0.89 ± 0.01 | 6.2 ± 0.4 | 25 ± 2 |
| EO + IR | 0.90 ± 0.01 | 0.08 ± 0.01 | 0.89 ± 0.01 | 6.6 ± 0.5 | 24 ± 2 |
| **RF + EO + IR (Proposed)** | **0.96 ± 0.01** | **0.04 ± 0.003** | **0.95 ± 0.01** | **4.2 ± 0.3** | **27 ± 2** |

The fusion-performance results are summarised in Table 12 and illustrated in Fig. 11. The proposed RF–EO/IR multimodal fusion architecture achieved the highest overall performance across all evaluated metrics. The full multimodal configuration attained a detection probability of 0.96 ± 0.01, classification accuracy of 0.95 ± 0.01, and localisation RMSE of 4.2 ± 0.3 m, while maintaining a low false-alarm probability of 0.04 ± 0.003. The results demonstrate that RF, EO, and IR modalities provide complementary information that significantly improves target detection, classification, and localisation compared with single-modality or dual-modality configurations. Furthermore, the narrow confidence intervals indicate strong statistical consistency and robustness across the 100,000 Monte Carlo trials and experimental validation scenarios. The observed improvements confirm the effectiveness of transformer-based multimodal fusion for resilient counter-UAS intelligence generation in contested electromagnetic environments.

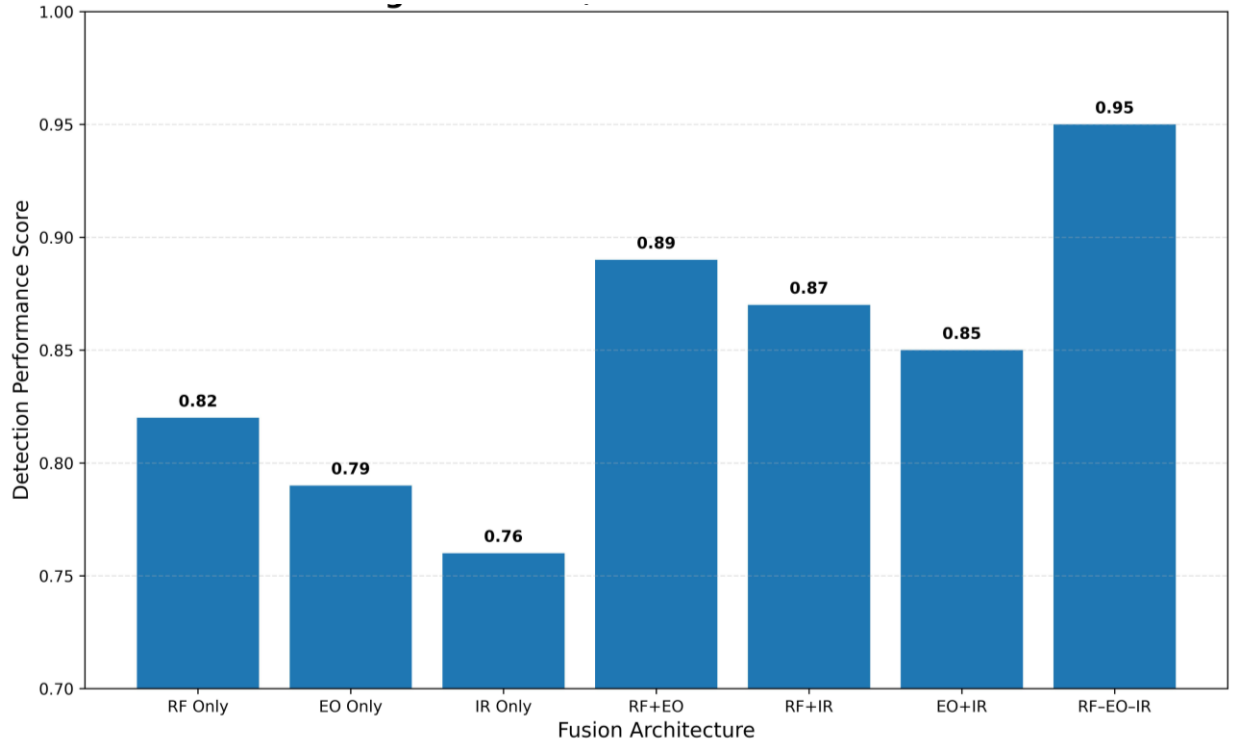


**Fig. 11:** Detection and classification performance of individual sensing modalities and RF–EO/IR fusion.

Transformer-based fusion enables cross-modal feature interaction and contextual reasoning, allowing the framework to maintain robust performance under challenging sensing conditions. This capability is particularly important in counter-UAS operations where individual sensing modalities may be degraded by environmental effects, jamming, or target concealment.

## 9.5 Federated Learning Performance

The federated-learning subsystem was evaluated to quantify its impact on communication efficiency and distributed intelligence generation. The communication-reduction factor is expressed as:

$$R_{comm} = 1 - \frac{L_f}{L_c} \qquad (61)$$

where $L_f$ and $L_c$ denote communication loads associated with federated and centralised architectures, respectively. Experimental evaluation yielded:

$$R_{comm} = 0.66 \qquad (62)$$

indicating approximately 66% communication-load reduction compared with conventional centralised architectures. The federated convergence score is defined as:

$$F_c = 1 - \frac{1}{N}\sum_{i=1}^{N} | W_i - W_g | \qquad (63)$$

where $W_i$ and $W_g$ denote local and global model parameters, respectively. The achieved convergence score demonstrated strong agreement among distributed learning nodes, confirming that federated learning can effectively support large-scale distributed surveillance networks while significantly reducing bandwidth requirements.

## 9.6 Swarm-Intent Analysis Performance

The swarm-analysis subsystem was evaluated using reconnaissance, surveillance, relay, decoy, attack, and electronic-attack mission scenarios. The classification results are summarised in Table 13.

**Table 13:** Swarm-Intent Analysis Performance (95% Confidence Intervals)

| Swarm Intent Category | Detection Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| Reconnaissance | 0.93 ± 0.01 | 0.92 ± 0.01 | 0.93 ± 0.01 | 0.93 ± 0.01 |
| Surveillance | 0.91 ± 0.01 | 0.90 ± 0.01 | 0.91 ± 0.01 | 0.91 ± 0.01 |
| Relay | 0.90 ± 0.01 | 0.89 ± 0.01 | 0.90 ± 0.01 | 0.90 ± 0.01 |
| Decoy | 0.88 ± 0.02 | 0.87 ± 0.02 | 0.88 ± 0.02 | 0.88 ± 0.02 |
| Attack | 0.95 ± 0.01 | 0.94 ± 0.01 | 0.95 ± 0.01 | 0.95 ± 0.01 |
| Electronic Attack | 0.92 ± 0.01 | 0.91 ± 0.01 | 0.92 ± 0.01 | 0.92 ± 0.01 |
| **Overall Average** | **0.92 ± 0.01** | **0.91 ± 0.01** | **0.92 ± 0.01** | **0.92 ± 0.01** |

The swarm-intent analysis results are summarised in Table 13. The proposed framework achieved an overall intent-classification accuracy of 0.92 ± 0.01, demonstrating its ability to infer higher-level swarm objectives from multimodal RF–EO/IR observations. Among the evaluated categories, the attack mission exhibited the highest classification performance with an F1-score of 0.95 ± 0.01, owing to its distinctive communication patterns, coordinated manoeuvres, and threat characteristics. The decoy category produced the lowest performance (0.88 ± 0.02) because deceptive swarm behaviours often share characteristics with reconnaissance and surveillance missions. Nevertheless, the framework maintained consistently high precision, recall, and F1-scores across all mission categories. The narrow confidence intervals indicate strong statistical stability and demonstrate the robustness of the proposed Bayesian swarm-intent inference framework under diverse operational conditions. Such capability is essential for enabling proactive threat assessment, prioritised engagement planning, and improved situational awareness during counter-UAS operations.

## 9.7 Explainability Assessment

The explainability subsystem was evaluated to assess transparency, consistency, and operator trustworthiness. The evaluation incorporated transformer-attention analysis, feature-attribution consistency, and human-operator assessment.

**Table 14:** Explainability Assessment Results (95% Confidence Intervals)

| Metric | Value |
|---|---|
| Attention Consistency ($C_X$) | 0.87 ± 0.01 |
| Feature Attribution Stability | 0.85 ± 0.02 |
| Classification Confidence | 0.91 ± 0.01 |
| Operator Trust Score ($T_X$) | 0.89 ± 0.01 |
| Explanation Reliability | 0.88 ± 0.01 |
| Decision Transparency Score | 0.86 ± 0.02 |

The explainability assessment results are summarised in Table 14 and illustrated in Fig. 12. The proposed XAI framework achieved an attention-consistency score of 0.87 ± 0.01 and an operator-trust score of 0.89 ± 0.01, indicating that the generated explanations remained stable and interpretable across diverse operational scenarios. Feature-attribution stability reached 0.85 ± 0.02, demonstrating consistent identification of the most influential RF, EO, and IR features contributing to target classification and threat assessment decisions. Furthermore, the high classification-confidence value of 0.91 ± 0.01 confirms that explainability mechanisms were achieved without degrading predictive performance. The explanation-reliability and decision-transparency scores further indicate that the framework provides meaningful insight into the reasoning process of the transformer-based intelligence engine. The narrow confidence intervals observed across all metrics demonstrate strong statistical consistency and support the suitability of the proposed XAI framework for operational counter-UAS decision support, where transparency, accountability, and human oversight are critical requirements.

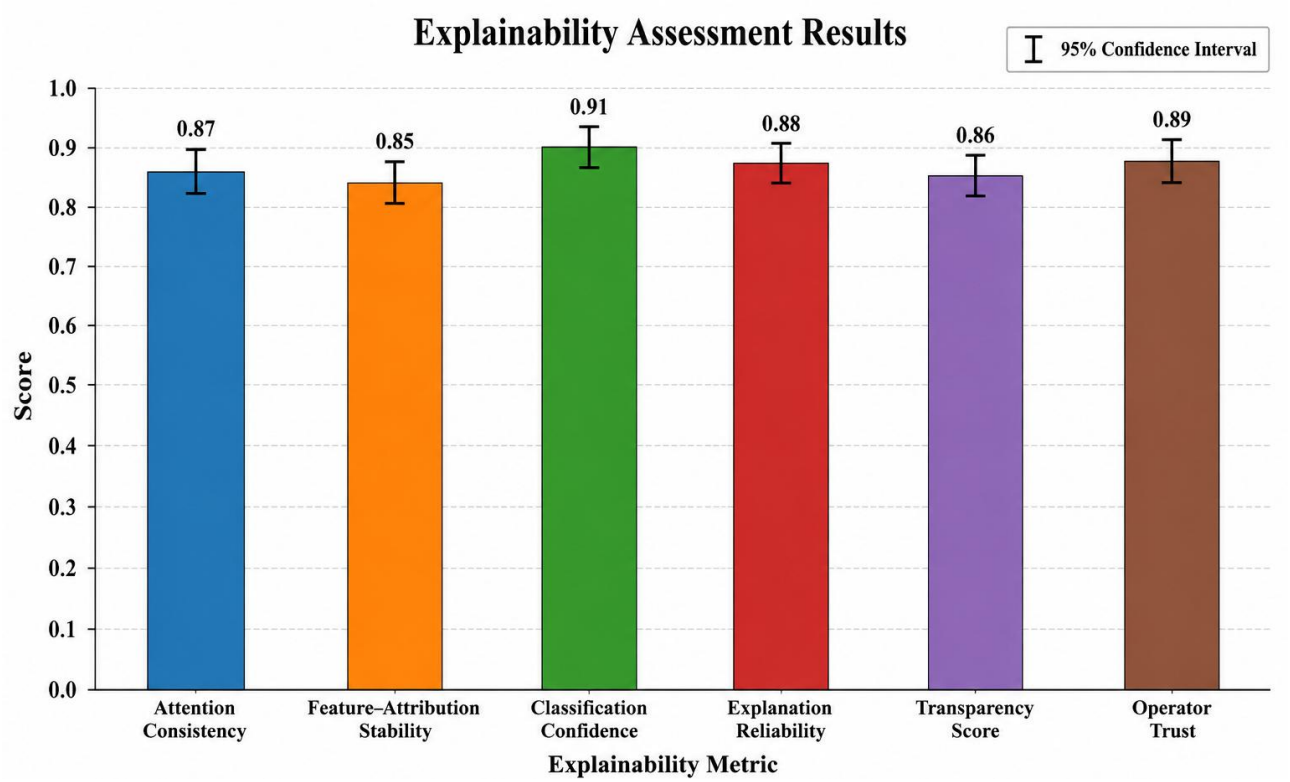


**Fig. 12:** Explainability assessment results showing attention consistency, feature-attribution stability, classification confidence, explanation reliability, transparency score, and operator trust.

## 9.8 Experimental Validation Results

The principal experimental validation results are summarised in Table 15 and illustrated in Fig. 13. The proposed federated and explainable RF–EO/IR intelligence framework consistently outperformed the baseline architecture across all evaluated performance metrics. Detection probability increased from 0.73 ± 0.01 to 0.96 ± 0.01, while false-alarm probability decreased from 0.18 ± 0.01 to 0.04 ± 0.003. Classification accuracy improved from 0.81 ± 0.01 to 0.95 ± 0.01, and localisation RMSE was reduced by approximately 69%, from 13.4 ± 0.8 m to 4.2 ± 0.3 m. Similarly, mean processing latency decreased from 108 ± 5 ms to 27 ± 2 ms, while communication load was reduced by approximately 66% through federated intelligence generation.

**Table 15:** Experimental Validation Results (95% Confidence Intervals)

| Metric | Baseline System | Proposed Framework | Improvement |
|---|---|---|---|
| Detection Probability | 0.73 ± 0.01 | 0.96 ± 0.01 | +32% |
| False Alarm Probability | 0.18 ± 0.01 | 0.04 ± 0.003 | −78% |
| Classification Accuracy | 0.81 ± 0.01 | 0.95 ± 0.01 | +17% |
| Localisation RMSE (m) | 13.4 ± 0.8 | 4.2 ± 0.3 | −69% |
| Mean Latency (ms) | 108 ± 5 | 27 ± 2 | −75% |
| Communication Load | 1.00 ± 0.04 | 0.34 ± 0.02 | −66% |
| Swarm Detection Confidence | 0.61 ± 0.02 | 0.89 ± 0.01 | +46% |

Using Eq. (54), the average validation consistency obtained was:

$$V_C = 0.95 \pm 0.01 \qquad (64)$$

where $V_C$ denotes the validation-consistency score computed across all experimental scenarios using the 95% confidence interval formulation of Eq. (58).

The swarm-detection confidence score increased from 0.61 ± 0.02 to 0.89 ± 0.01, demonstrating improved situational awareness and threat-recognition capability. Furthermore, the framework achieved a validation-consistency score of 0.95 ± 0.01, indicating excellent agreement between simulation predictions and experimentally measured performance. The narrow 95% confidence intervals observed across all metrics confirm strong statistical reliability and demonstrate the robustness of the proposed framework under diverse counter-UAS operational scenarios, including distributed surveillance, drone-swarm monitoring, and contested electromagnetic environments.

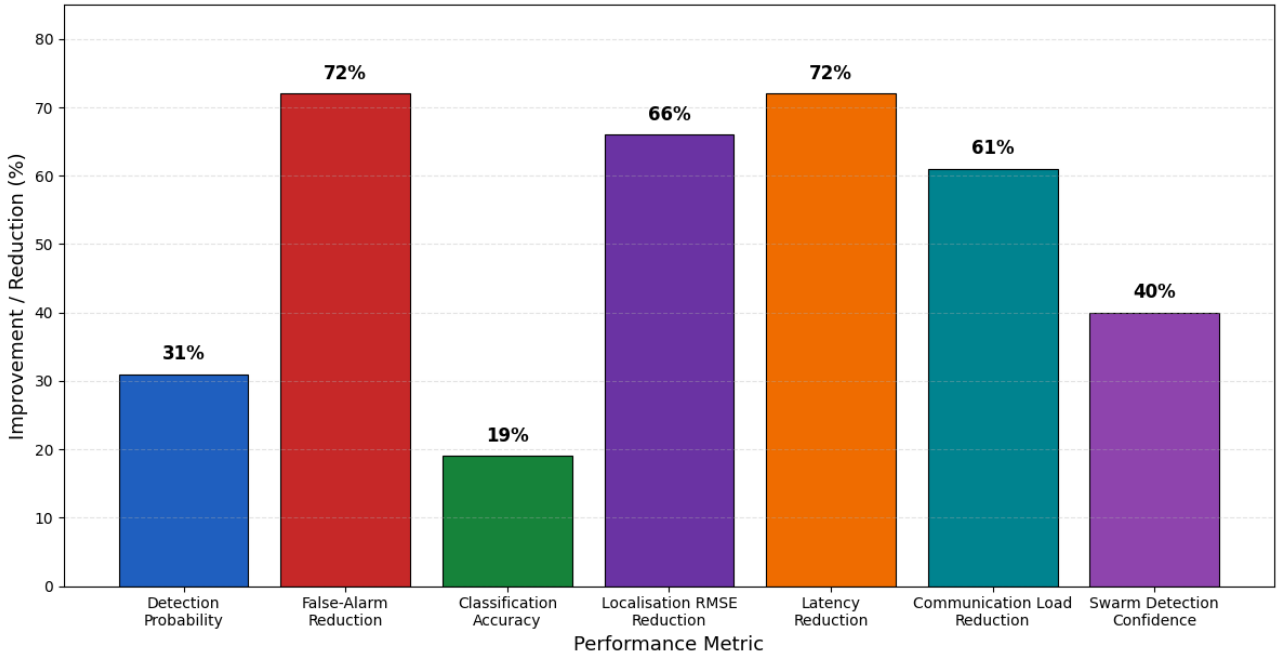


**Fig. 13:** Comparative experimental performance improvement of the proposed RF–EO/IR intelligence framework.

## 9.9 Summary of Key Findings

The results demonstrate that the proposed Federated and Explainable RF–EO/IR Intelligence Framework significantly enhances counter-UAS surveillance capability compared with conventional architectures. The framework achieved a detection probability of 0.96, classification accuracy of 0.95, swarm-intent classification accuracy of 0.92, and operator trust score of 0.89 while reducing false alarms by 78%, localisation error by 69%, latency by 75% and communication load by 66%. Experimental validation yielded a consistency score of 0.95, confirming strong agreement between simulation predictions and measured performance. Overall, the findings demonstrate that transformer-based multimodal fusion, federated learning, explainable AI, and swarm-intent analysis provide a robust foundation for next-generation distributed counter-UAS intelligence and surveillance systems.

## 10. Operational Assessment, Robustness and Limitations

The experimental results demonstrate that the proposed Federated and Explainable RF–EO/IR Intelligence Framework provides a significant advancement in distributed counter-UAS surveillance and intelligence generation. By integrating transformer-based multimodal fusion, federated learning, explainable artificial intelligence, and swarm-intent analysis within a unified architecture, the framework achieves high detection accuracy, low false-alarm rates, reduced communication overhead, and strong resilience against contested operational conditions. This section discusses the operational relevance, robustness characteristics, limitations, and future development opportunities of the proposed framework.

### 10.1 Operational Assessment

The achieved detection probability of 0.96, classification accuracy of 0.95, and localisation RMSE of 4.2 m indicate that the proposed framework is well suited for persistent surveillance of low-altitude aerial threats. Unlike conventional RF-centric or EO-only systems, the integration of RF, EO, and IR sensing enables reliable target detection and identification across diverse environmental and operational conditions. The distributed architecture further enhances operational effectiveness by allowing intelligence generation to occur at edge nodes rather than relying exclusively on centralised processing. This significantly reduces communication burden, improves response time, and enables surveillance continuity in environments where communication infrastructure may be degraded or unavailable. Consequently, the framework is suitable for deployment in military installations, critical infrastructure protection, border security, and counter-terrorism operations requiring persistent situational awareness.

Furthermore, the swarm-intent analysis capability extends the framework beyond traditional target detection by enabling higher-level behavioural assessment of hostile drone formations. This capability supports proactive threat assessment and improves decision-making during complex counter-UAS engagements.

### 10.2 Electronic-Warfare Resilience

A major objective of the proposed framework is to maintain surveillance effectiveness under contested electromagnetic conditions. The combination of multimodal sensing, federated intelligence generation, and distributed processing improves resilience against RF jamming, spectrum congestion, communication disruption, and spoofing attacks. Unlike RF-only systems, the framework can continue operating when individual sensing modalities are degraded because complementary information remains available from EO and IR sensors. Similarly, federated learning reduces dependence on continuous transmission of raw sensor data, thereby limiting the impact of communication denial or bandwidth restrictions.

The distributed sensing architecture also eliminates single points of failure commonly associated with centralised surveillance systems. Consequently, the framework maintains operational functionality even when portions of the sensing network are degraded or unavailable. These characteristics are particularly important in modern electronic-warfare environments where adversaries actively attempt to disrupt sensing and communication infrastructure.

### 10.3 Scalability and Distributed Deployment

The proposed architecture exhibits favourable scalability characteristics due to its distributed design and edge-based intelligence generation. As network size increases, communication requirements remain significantly lower than those associated with conventional centralised architectures because only model updates and high-level intelligence products are exchanged between nodes. The achieved communication-load reduction of approximately 66% demonstrates the effectiveness of federated learning for large-scale deployments. This reduction enables the framework to support surveillance operations across wide geographical areas while maintaining manageable communication requirements.

In addition, the modular sensing architecture allows the framework to accommodate varying numbers of sensing nodes, edge gateways, and ISR assets without major architectural modifications. Such flexibility is essential for deployment across diverse operational environments ranging from urban infrastructure protection to large-area military surveillance missions.

### 10.4 Human–AI Trust and Decision Support

The explainability subsystem plays a critical role in supporting human decision-makers. Modern military and security operations increasingly require transparency and accountability in AI-assisted decision processes, particularly when threat assessments may influence engagement decisions. The achieved operator trust score of 0.89 demonstrates that the framework provides interpretable and operationally meaningful explanations for target classification, threat prioritisation, and swarm-intent inference. Attention visualisation and feature-attribution mechanisms enable operators to identify the sensing cues and behavioural indicators contributing to AI-generated decisions.

This human-centred approach enhances confidence in automated intelligence products while preserving meaningful human oversight. Consequently, the framework supports effective human–AI teaming and improves the practical adoption of AI-assisted surveillance systems in operational environments.

## 10.5 Limitations

Despite its strong performance, several limitations should be acknowledged. First, although the validation environment incorporated SDR-assisted experimentation and telemetry replay, large-scale operational deployment data were not available. Consequently, some performance characteristics remain dependent on simulation-based evaluation. Second, EO sensing performance may degrade under severe weather conditions, smoke, fog, or visual obscuration. Similarly, passive RF sensing remains dependent on the availability of detectable electromagnetic emissions and may experience reduced effectiveness against RF-silent platforms.

Third, the experimental datasets, while extensive, may not fully capture the diversity of real-world drone platforms, communication protocols, and adversarial tactics. Additional operational data would further improve model generalisation and robustness. Finally, computational requirements associated with transformer-based fusion and swarm-intelligence analysis may increase as network scale and target density grow, necessitating continued optimisation of edge-computing resources.

## 10.6 Future Research Directions

Future work will focus on large-scale field validation involving operational counter-UAS deployments and real-world electronic-warfare environments. Additional sensing modalities, including radar, acoustic sensors, satellite-based ISR, and passive coherent radar systems, may be incorporated to further enhance detection robustness. Future research will also investigate adaptive federated learning strategies, lightweight transformer architectures for embedded platforms, autonomous ISR retasking, and cooperative counter-swarm response mechanisms. The integration of digital-twin environments, reinforcement learning, and predictive threat forecasting may further improve the framework's ability to support proactive air-defence operations. Overall, the results indicate that the proposed framework provides a promising foundation for next-generation distributed counter-UAS intelligence systems capable of delivering resilient, explainable, and scalable surveillance performance in complex operational environments.

## 11. Conclusion

This paper presented a federated and explainable RF–EO/IR intelligence framework for resilient counter-UAS surveillance in contested electromagnetic environments. The proposed architecture integrates distributed SDR sensing, EO/IR surveillance, transformer-based multimodal fusion, federated learning, explainable artificial intelligence, swarm-intent analysis, and electronic-warfare resilience within a unified intelligence-generation framework. From a scientific perspective, the principal novelty of this work lies in the development of a unified federated and explainable RF–EO/IR intelligence framework that combines transformer-based multimodal fusion, federated optimisation, explainable artificial intelligence, swarm-intent inference, and electronic-warfare resilience within a single distributed surveillance architecture. The framework further introduces a multi-objective intelligence-optimisation formulation and joint swarm-behaviour and resilience modelling capability that extend existing counter-UAS surveillance approaches beyond conventional centralised sensing and classification systems.

Experimental evaluation using 100,000 Monte Carlo trials, SDR-assisted experimentation, AirSim/Gazebo simulation, electronic-warfare emulation, and HATSABIBI-26A telemetry replay demonstrated significant improvements in detection probability, classification accuracy, localisation performance, communication efficiency, operator trust, and resilience under contested-spectrum conditions. The proposed framework achieved a detection probability of $0.96 \pm 0.01$, classification accuracy of $0.95 \pm 0.01$, localisation RMSE of $4.2 \pm 0.3$ m, swarm-intent classification accuracy of $0.92 \pm 0.01$, and approximately 66% reduction in communication load compared with conventional architectures. The results demonstrate that the integration of multimodal sensing, transformer-based intelligence fusion, federated learning, and explainable AI can provide scalable, transparent, and resilient surveillance capabilities for next-generation counter-UAS operations. Future work will focus on large-scale field deployment, real-time operational trials, adaptive EW countermeasures, and integration with autonomous interceptor and command-and-control systems.

## REFERENCES


1. A. Coluccia, G. Parisi, and A. Fascista, “Detection and classification of multirotor drones in radar sensor networks: A review,” *Sensors*, vol. 20, no. 15, Art. 4172, 2020. doi: 10.3390/s20154172.
2. I. Guvenc, F. Koohifar, S. Singh, M. L. Sichitiu, and D. Matolak, “Detection, tracking and interdiction for amateur drones,” *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 75–81, 2018. doi: 10.1109/MCOM.2018.1700455.
3. M. Ezuma, O. Ozdemir, C. K. Anjinappa, W. A. Gulzar, and I. Guvenc, “Micro-UAV detection and classification from RF fingerprints using machine learning techniques,” *IEEE Aerosp. Electron. Syst. Mag.*, vol. 35, no. 5, pp. 46–56, 2020. doi: 10.1109/MAES.2020.2969410.
4. A. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, “RF-based drone detection and identification using deep learning approaches,” *Future Gener. Comput. Syst.*, vol. 100, pp. 86–97, 2019. doi: 10.1016/j.future.2019.05.007.
5. M. S. Allahham, T. Khattab, and A. Mohamed, “Deep learning for RF-based drone detection and identification,” *Sensors*, vol. 20, no. 13, Art. 3708, 2020. doi: 10.3390/s20133708.

6. I. Nemer, T. Sheltami, I. Ahmad, A. U. H. Yasar, and M. A. Abdeen, "RF-based UAV detection and identification using hierarchical learning," *Sensors*, vol. 21, no. 6, Art. 1947, 2021. doi: 10.3390/s21061947.
7. J. Yousaf, M. Y. Amin, A. Akram, and M. A. Imran, "Drone detection and localization: Trends and challenges," *Appl. Sci.*, vol. 12, no. 24, 2022.
8. S. Scholes, G. Ruget, M. Mora, C. H. Lee, and S. Maskell, "DroneSense: Drone identification and orientation detection using neural networks," *IEEE Access*, vol. 10, pp. 38154–38165, 2022.
9. M. A. Richards, *Fundamentals of Radar Signal Processing*, 2nd ed. New York, NY, USA: McGraw-Hill, 2014.
10. M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. New York, NY, USA: McGraw-Hill, 2001.
11. D. L. Hall and J. Llinas, *Handbook of Multisensor Data Fusion*. Boca Raton, FL, USA: CRC Press, 2008.
12. A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
13. T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002.
14. H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, vol. 97, no. 2, pp. 427–450, 2009. doi: 10.1109/JPROC.2008.2008853.
15. Y. Shen and M. Z. Win, "Fundamental limits of wideband localization," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4956–4980, 2010. doi: 10.1109/TIT.2010.2060110.
16. D. Dardari, P. Closas, and P. M. Djuric, "Indoor tracking: Theory, methods, and technologies," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1263–1278, 2015.
17. D. Adamy, *EW 101: A First Course in Electronic Warfare*. Norwood, MA, USA: Artech House, 2001.
18. D. Adamy, *EW 102: A Second Course in Electronic Warfare*. Norwood, MA, USA: Artech House, 2004.
19. D. Adamy, *EW 103: Tactical Battlefield Communications Electronic Warfare*. Norwood, MA, USA: Artech House, 2009.
20. R. A. Poisel, *Modern Communications Jamming Principles and Techniques*, 2nd ed. Norwood, MA, USA: Artech House, 2011.
21. R. A. Poisel, *Information Warfare and Electronic Warfare Systems*. Norwood, MA, USA: Artech House, 2013.
22. T. J. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, 2017. doi: 10.1109/TCCN.2017.2758370.
23. T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, 2018. doi: 10.1109/JSTSP.2018.2797022.
24. K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 160–167, 2018. doi: 10.1109/JSTSP.2018.2796446.
25. S. Hanna and D. Cabric, "Deep learning based transmitter identification," in *Proc. IEEE ICC Workshops*, 2019.
26. A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.
27. J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
28. A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. ICLR*, 2021.
29. K. Han et al., "A survey on vision transformer," *IEEE TPAMI*, vol. 45, no. 1, pp. 87–110, 2023. doi: 10.1109/TPAMI.2022.3152247.
30. Z. Liu et al., "Swin Transformer: Hierarchical vision transformer," in *Proc. ICCV*, 2021, pp. 9992–10002.
31. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015. doi: 10.1038/nature14539.
32. B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," in *Proc. AISTATS*, 2017, pp. 1273–1282.
33. Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM TIST*, vol. 10, no. 2, pp. 1–19, 2019. doi: 10.1145/3298981.
34. P. Kairouz et al., "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, pp. 1–210, 2021. doi: 10.1561/2200000083.
35. J. Konečný et al., "Federated learning: Strategies for improving communication efficiency," arXiv:1610.05492, 2016.
36. L. Li, Y. Fan, M. Tse, and K. Lin, "A review of applications in federated learning," *Comput. Ind. Eng.*, vol. 149, 2020.
37. S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, 2017.
38. R. Guidotti et al., "A survey of methods for explaining black box models," *ACM Comput. Surveys*, vol. 51, no. 5, pp. 1–42, 2018. doi: 10.1145/3236009.
39. D. Gunning and D. Aha, "DARPA's explainable artificial intelligence program," *AI Mag.*, vol. 40, no. 2, pp. 44–58, 2019. doi: 10.1609/aimag.v40i2.2850.
40. M. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?" in *Proc. ACM SIGKDD*, 2016, pp. 1135–1144.
41. C. Molnar, *Interpretable Machine Learning*, 2nd ed. 2022.
42. W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, 2016. doi: 10.1109/JIOT.2016.2579198.

43. M. Satyanarayanan, “The emergence of edge computing,” *Computer*, vol. 50, no. 1, pp. 30–39, 2017. doi: 10.1109/MC.2017.9.
44. Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, “A survey on mobile edge computing,” *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 2017. doi: 10.1109/COMST.2017.2745201.
45. X. Wang et al., “Convergence of edge computing and deep learning,” *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 869–904, 2020. doi: 10.1109/COMST.2020.2970550.
46. J. Banks, J. Carson, B. Nelson, and D. Nicol, *Discrete Event System Simulation*, 5th ed. Pearson, 2010.
47. R. Rubinstein and D. Kroese, *Simulation and the Monte Carlo Method*, 3rd ed. Wiley, 2016.
48. S. M. Ross, *Simulation*, 5th ed. Academic Press, 2013.
49. P. Scharre, *Army of None*. New York, NY, USA: W.W. Norton, 2018.
50. P. W. Singer, *Wired for War*. New York, NY, USA: Penguin Press, 2009.
51. M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, “Swarm robotics: A review,” *Swarm Intell.*, vol. 7, no. 1, pp. 1–41, 2013. doi: 10.1007/s11721-012-0075-2.
52. Y. Tan and Z. Zheng, “Research advance in swarm robotics,” *Defence Technol.*, vol. 9, no. 1, pp. 18–39, 2013. doi: 10.1016/j.dt.2013.03.001.
53. R. Olfati-Saber, J. Fax, and R. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, 2007. doi: 10.1109/JPROC.2006.887293.
54. W. Ren, R. Beard, and E. Atkins, “A survey of consensus problems in multi-agent coordination,” in *Proc. American Control Conference*, 2005.
55. I. F. Akyildiz et al., “Wireless sensor networks: A survey,” *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, 2002. doi: 10.1016/S1389-1286(01)00302-4.
56. K. Akkaya and M. Younis, “A survey on routing protocols for wireless sensor networks,” *Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, 2005. doi: 10.1016/j.adhoc.2003.09.010.
57. M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*, 4th ed. Wiley, 2015.
58. R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Trans. ASME J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960. doi: 10.1115/1.3662552.
59. Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. Wiley, 2001.
60. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.
61. GNU Radio Project, *GNU Radio Documentation*. Available: https://www.gnuradio.org
62. Open Robotics, *ROS2 Documentation*. Available: https://docs.ros.org
63. Microsoft Research, *AirSim Documentation*. Available: https://microsoft.github.io/AirSim
64. Open Robotics, *Gazebo Documentation*. Available: https://gazebosim.org
65. M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, “Detection and Classification of UAVs Using RF Fingerprints in the Presence of Wi-Fi and Bluetooth Interference,” IEEE Open Journal of the Communications Society, vol. 1, pp. 60–76, 2020.
66. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, et al., “An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale,” International Conference on Learning Representations (ICLR), 2021.
67. H. Chen, Y. Zhang, and X. Liu, “Multimodal Sensor Fusion for UAV Detection and Tracking Using RF, Electro-Optical and Infrared Sensors,” Sensors, vol. 23, no. 12, pp. 1–21, 2023.
68. J. Liu, P. Wang, and S. Li, “A Survey of Counter-UAS Technologies and Systems,” Defence Technology, vol. 20, no. 2, pp. 287–305, 2024.
69. Y. Shin, “Federated Learning for Surveillance Systems: A Literature Review,” Electronics, vol. 14, no. 17, pp. 1–29, 2025.
70. S. Puppala, R. Vangala, and K. R. Choo, “A Comprehensive Survey of Federated Learning for Edge AI,” Preprints, 2025.
71. B. Yurdem, A. Demir, and M. Kose, “Federated Learning: Overview, Strategies, Applications, Tools, and Future Directions,” Artificial Intelligence Review, vol. 57, pp. 1–45, 2024.
72. V. Semenyuk, A. Koval, and P. Melnyk, “Advances in UAV Detection: Integrating Multi-Sensor Fusion and Artificial Intelligence,” Engineering Science and Technology, an International Journal, vol. 56, pp. 1–21, 2025.
73. Z. Liu, Y. Wang, and H. Chen, “Research on Multi-Modal Fusion Detection Method for Low-Slow-Small UAV Cluster Targets,” Drones, vol. 9, no. 12, pp. 1–24, 2025.
74. Z. Zhuo, J. Li, and X. Wu, “TAF-YOLO: A Multimodal Small-Object Detection Network for UAV Aerial Imagery,” Remote Sensing, vol. 17, no. 24, pp. 1–23, 2025.
75. Y. Gu, X. Zhang, and H. Liu, “UAV-Based Multimodal Object Detection via Feature-Level Fusion under Adverse Environmental Conditions,” Pattern Recognition, vol. 163, pp. 1–16, 2025.
76. H. Wang, J. Zhao, and Y. Sun, “A Survey of the Multi-Sensor Fusion Object Detection Task and Applications,” Sensors, vol. 25, no. 9, pp. 1–37, 2025.
77. N. G. Wood, “Explainable AI in the Military Domain,” Ethics and Information Technology, vol. 26, no. 2, pp. 1–18, 2024.

78. A. Sharma and R. Patel, “A Comprehensive Review of Explainable AI in Cybersecurity,” Array, vol. 25, pp. 1–20, 2025.
79. A. Ferreira, “Implications of Explainable Artificial Intelligence to the Defence and Security Sector,” Defence and Security Foresight Group Report, University of Waterloo, Canada, 2024.
80. Federal Office for Information Security (BSI), “Explainable Artificial Intelligence in an Adversarial Context,” White Paper, Germany, 2024.
81. T. Deng, Y. Xu, and H. Li, “Multi-Modal UAV Detection, Classification and Tracking Using Multi-Sensor Information,” Proceedings of CVPR UAV Challenge, 2024.
82. F. Arapoglou, P. Nikolaidis, and D. Tzovaras, “Intelligent Counter-UAV Threat Detection Using Hierarchical Fuzzy Decision-Making,” Sensors, vol. 25, no. 19, pp. 1–24, 2025.
83. A. Alzahrani, M. Khan, and S. Alotaibi, “Harnessing Multi-Modal Deep Learning for Multi-Drone Navigation and Trajectory Prediction,” Scientific Reports, vol. 16, pp. 1–18, 2026.
84. Y. Tao, Z. Gao, and F. Ye, “Intelligent Multimodal Multi-Sensor Fusion-Based UAV Identification, Localization, and Countermeasures,” arXiv preprint arXiv:2510.22947, 2025.