MULTIDISCIPLINARY RESEARCH IN
COMPUTING INFORMATION SYSTEMS

VOL 06 ISSUE 01 2026

https://mrcis.org

ThreatSense: Neuro-Symbolic Reasoning for Autonomous Threat Assessment in Unmanned Systems

Tobias Krüger

Department of Computer Science, Technical University of Darmstadt, Germany

Elena Petrova

Department of Computer Science, Technical University of Darmstadt, Germany

Abstract: *Autonomous unmanned systems increasingly operate in adversarial environments where real-time threat recognition and response are safety-critical imperatives. This paper presents ThreatSense, a novel neuro-symbolic AI framework that integrates deep neural perception with symbolic logic reasoning for autonomous threat assessment in unmanned aerial vehicle (UAV) networks. Unlike conventional intrusion detection approaches that rely solely on either pattern-based learning or handcrafted rule systems, ThreatSense unifies both paradigms through a layered inference architecture: a neural perception module for multi-modal sensor feature extraction, a symbolic knowledge engine encoding domain-specific threat ontologies in first-order logic, and a joint inference layer mediating probabilistic symbol grounding. ThreatSense is evaluated against four representative threat classes including GPS spoofing, denial-of-service, adversarial sensor perturbation, and black hole routing attacks. Experimental results demonstrate a detection accuracy of 97.3%, a false positive rate (FPR) of 1.8%, and an average inference latency of 38 ms, outperforming deep learning baselines by 4.8 percentage points in accuracy while maintaining full symbolic interpretability. The framework advances explainable and computationally efficient threat intelligence for safety-critical autonomous platforms.*

Keywords: *neuro-symbolic AI, unmanned aerial vehicles, threat assessment, autonomous systems, intrusion detection*

1.Introduction

The proliferation of unmanned systems across civilian, commercial, and military domains has fundamentally transformed how autonomous agents interact with physical and cyber environments. Unmanned aerial vehicles (UAVs) in particular have transitioned from niche research platforms to mainstream infrastructure components in logistics, surveillance, precision agriculture, emergency response, and defense operations. This shift has been driven by rapid advances in sensing, communication, and artificial intelligence (AI) technologies, enabling increasingly autonomous behavior without continuous human supervision. However, the same computational sophistication that empowers UAVs also exposes them to a growing taxonomy of cyber and physical threats that exploit vulnerabilities in communication links, sensor pipelines, and onboard control systems [1]. The integration of UAVs into safety-critical workflows — from package delivery in dense urban corridors to autonomous reconnaissance in contested airspace — creates attack surfaces whose exploitation carries consequences extending well beyond the vehicle itself, encompassing infrastructure damage, data exfiltration, and mission compromise [2]. Threat assessment in unmanned systems presents a uniquely demanding technical challenge. Unlike conventional networked devices that operate in controlled environments, UAVs navigate dynamic, unstructured operational theaters where threat signatures may be subtle, temporally diffuse, or deliberately obfuscated by adversarial actors with knowledge of deployed detection mechanisms. A representative operational deployment involves multiple UAV nodes distributed across geographically separated surveillance zones, each communicating wirelessly with dedicated base stations that relay consolidated traffic data to a centralized analysis center over wired backhaul links. In this topology, any single compromised node can propagate malicious behavior across the network before centralized detection mechanisms are activated, making distributed and hierarchical threat assessment architectures a practical necessity. Traditional rule-based intrusion detection systems (IDS) offer deterministic and interpretable decision logic but fail to generalize beyond their predefined rule sets, rendering them brittle against novel or mutated attack patterns that diverge even slightly from training-time exemplars [3]. Conversely, pure machine learning (ML) approaches demonstrate strong generalization across diverse threat classes but suffer from opacity, susceptibility to adversarial perturbation, and performance degradation when training distributions diverge from deployment conditions — a near-inevitable occurrence in dynamic autonomous operations. The neuro-symbolic computing paradigm offers a principled resolution to this dilemma by integrating the representational power of deep neural networks (DNNs) with the logical transparency of symbolic AI [4]. In a neuro-symbolic architecture, neural components learn perceptual representations from high-dimensional, noisy sensor data, while symbolic components impose structured reasoning constraints, enabling decisions that are simultaneously data-driven and logically grounded. This integration is particularly germane to autonomous threat assessment because it supports interpretability — critical in safety-critical deployments where decisions must be auditable by human operators — while maintaining the adaptability required to address evolving threat landscapes. The cognitive analogy is instructive: neural processing parallels fast, intuitive System 1 cognition, while symbolic reasoning mirrors the deliberate, rule-governed System 2 processes that human experts employ when evaluating ambiguous threat indicators [5].

Despite growing interest in neuro-symbolic AI for cybersecurity applications, the specific problem of real-time threat assessment in resource-constrained unmanned platforms remains substantially underexplored. Existing neuro-symbolic systems are predominantly validated on benchmark classification tasks or laboratory-scale scenarios that do not capture the operational realities of UAV missions, including energy constraints, high-speed mobility, intermittent communication, and concurrent multi-threat scenarios. The absence of a unified framework that jointly optimizes inference quality, computational efficiency, and symbolic explainability represents a significant gap in the current literature, constraining practical adoption of neuro-symbolic approaches in autonomous platform security [6]. Furthermore, most prior work on UAV intrusion detection treats the problem as a unimodal classification task, ignoring the richness of correlated multi-sensor evidence that characterizes real threat events in field deployments [7]. This paper addresses these gaps through the design, implementation, and evaluation of ThreatSense, a neuro-symbolic threat assessment framework tailored for autonomous unmanned systems. ThreatSense draws on a convolutional neural network (CNN) backbone for real-time sensor feature extraction, a logic tensor network (LTN) for differentiable symbolic reasoning over a domain threat ontology, and a hybrid inference engine that aggregates neural confidence scores with symbolic constraint satisfaction measures to produce final threat classifications with associated natural language explanations. The paper makes four distinct contributions: it introduces a multi-modal threat representation fusing radio frequency signatures, flight dynamics telemetry, and communication metadata; it proposes an adaptive symbolic knowledge base that accommodates new threat rules without retraining the neural backbone; it demonstrates a lightweight compression strategy enabling deployment on embedded UAV processors with sub-50 ms inference cycles; and it provides rigorous comparative evaluation against five state-of-the-art baselines across four threat categories.

2. Literature Review

Research at the intersection of autonomous unmanned systems security and AI has accelerated markedly, driven by the expanding deployment of UAV platforms and the maturation of deep learning techniques applicable to anomaly detection, behavioral classification, and adversarial robustness. This section surveys the major threads of relevant work across three thematic clusters: UAV security threat landscapes, machine learning-based IDS approaches, and neuro-symbolic AI methodologies with security applications. The reference numbering continues directly from the preceding section. The security vulnerabilities of UAV systems have been comprehensively catalogued across multiple attack dimensions. Mohsan et al [8]. Provided one of the most thorough surveys of the UAV landscape, demonstrating that rapid commercial expansion of drone deployments has outpaced the development of corresponding security architectures, with most deployed platforms relying on legacy authentication and encryption protocols designed for benign operating conditions. Shafique et al [9]. Constructed a detailed taxonomy of UAV attack surfaces, categorizing threats along physical, communication, and software dimensions and concluding that GPS spoofing ranks among the most dangerous threat classes due to its capacity to silently redirect navigation without triggering standard onboard anomaly detectors. Kong specifically examined countermeasure effectiveness across the cyberattack landscape, demonstrating that current defenses suffer from reactive postures ill-suited to anticipatory threat management, and called for predictive intelligence mechanisms capable of estimating threat likelihood before full attack manifestation [10]. Yahuza et al. Addressed the emerging paradigm of Internet of Drones — interconnected UAV fleets sharing sensory and operational data — identifying novel threat surfaces that arise from mesh networking topologies and heterogeneous communication standards across seventeen distinct attack classes [11]. Their analysis of privacy and security tradeoffs in UAV mesh networks is particularly relevant to ThreatSense's multi-modal threat encoding design. He et al. Investigated communication security in UAV systems within the context of public safety networks, demonstrating that the dual role of UAVs as both network nodes and mission platforms amplifies the security consequences of successful attack execution compared to conventional networked endpoints [12]. Pandey et al. Extended this analysis to 5G-enabled multi-UAV swarm scenarios, identifying that coordination protocol weaknesses in densely networked swarms introduce synchronization-based attack surfaces that classical IDS architectures are not equipped to address [13]. On the detection side, Abu Al-Haija and Al Badawi proposed a high-performance DNN-based IDS specifically targeting networked UAV environments, achieving over 99% classification accuracy on benchmark network traffic datasets through an optimized multilayer perceptron architecture [14]. While their results are strong under controlled conditions, the authors acknowledge that detection performance degrades significantly under real UAV traffic conditions with class imbalance and covariate shift, underscoring the generalization limitations that motivate hybrid architectural approaches. Basan et al. Explored a self-diagnosis paradigm that monitors internal parameter trajectories — including motor RPM, battery voltage, and attitude angles — to detect behavioral anomalies without reliance on external network traffic analysis, establishing the telemetry-centric detection strand that ThreatSense incorporates in its kinematic feature branch

[15]. Hassler et al. Contributed a cyber-physical IDS that fuses both network traffic features and physical sensor streams, demonstrating that fusion approaches consistently outperform single-modality baselines on real UAV attack datasets collected from instrumented hardware testbeds [16]. Reinforcement learning (RL) has also emerged as a compelling paradigm for adaptive threat detection. Praveena et al. Proposed an optimal deep RL framework for UAV IDS, demonstrating that a deep Q-network trained on dynamic reward signals can adapt to changing attack distributions in near real-time [17]. Samriya et al. Applied ant colony optimization combined with DNN classifiers to minimize energy consumption during IDS activation, a practically important contribution for battery-constrained UAV platforms [18]. Bouhamed et al. Proposed a periodic deep RL approach that minimizes communication overhead by activating detection agents only when elevated risk is inferred from network condition monitoring, a strategy that directly informs ThreatSense's adaptive symbolic activation mechanism [19]. Ngo et al. Conducted a comparative evaluation of feature selection versus feature extraction strategies for ML-based UAV intrusion detection, concluding that autoencoder-based dimensionality reduction consistently outperforms filter-based selection on high-dimensional communication logs [20]. Within the neuro-symbolic AI literature, d'Avila Garcez and Lamb established the conceptual framework of the current wave of neuro-symbolic computing, arguing that integrating neural and symbolic reasoning is essential for achieving trustworthy, explainable, and causally grounded intelligence in safety-critical systems [21]. Badreddine et al. Formalized this integration through Logic Tensor Networks (LTNs), a differentiable framework that translates first-order logic (FOL) formulas into real-valued constraints on neural network outputs, enabling simultaneous end-to-end learning and constraint satisfaction [22]. This framework forms the core of ThreatSense's symbolic reasoning engine. Manhaeve et al. Achieved complementary goals through DeepProbLog, extending probabilistic logic programming with neural predicates that allow end-to-end differentiable training of hybrid models with strong performance on relational learning tasks where symbolic structure is partially known [23] . Several recent works have specifically explored neuro-symbolic methods for cybersecurity applications. Cunnington et al. Demonstrated that neuro-symbolic threat hunting in security operations center environments can reduce false positive rates by over 30% compared to purely neural anomaly detectors by incorporating domain expert rules as symbolic constraints [24]. Zhang et al. Applied knowledge graph-augmented neural reasoning to cyber threat intelligence extraction, showing that embedding attack pattern ontologies into the reasoning layer substantially improves cross-domain threat attribution accuracy [25]. Acharya et al. Explored symbolic knowledge distillation from large neural models into compact rule sets that can be verified and audited by domain experts, contributing a method for translating learned threat representations into operator-interpretable explanations [26]. Kejriwal and Sharma proposed a hybrid neuro-symbolic framework specifically for adversarial attack detection in autonomous systems, integrating neural feature extraction, symbolic logic-based validation, and dynamic adversarial mitigation, reporting superior sensitivity and interpretability compared to standalone deep learning models [27]. Chen et al. Investigated probabilistic logic programming for anomaly scoring in cyber-physical systems, demonstrating that symbolic priors derived from network topology ontologies improve detection sensitivity on low-volume attack signatures that statistical baselines fail to characterize [28]. Sun et al. Designed a knowledge-aware DNN for GPS spoofing detection in small UAVs, highlighting that domain knowledge integration enhances robustness against signal manipulation attacks that purely data-driven baselines misclassify as benign GPS drift. Sedjelmaci et al [29]. Proposed a hierarchical IDS scheme operating at both UAV and ground station levels that categorizes behavioral anomalies into normal, abnormal, suspect, and malicious classes — a threat severity stratification strategy that ThreatSense's symbolic engine adopts in its predicate hierarchy [30]. Yi et al. Developed a neuro-symbolic visual question answering system

that leverages symbolic scene graphs to answer compositional queries, establishing an architectural precedent for sensor-to-symbol translation pipelines applicable to ThreatSense's multi-modal feature encoding stage [31]. Taken together, the surveyed literature reveals a clear convergence toward integrated learning-reasoning architectures for autonomous security systems, yet a persistent absence of end-to-end neuro-symbolic frameworks validated on realistic UAV threat scenarios. The gap between benchmark performance and field deployability, combined with the lack of symbolic interpretability in dominant detection architectures, defines the precise technical space that ThreatSense is designed to occupy.

3. Methodology

3.1 ThreatSense System Architecture and Deployment Topology

ThreatSense is designed as a modular, layered neuro-symbolic framework whose three principal components — the neural perception module, the symbolic knowledge engine, and the hybrid inference layer — operate in a sequential pipeline capable of producing real-time threat classifications with associated symbolic explanations. The architecture is guided by two core design imperatives: the perceptual component must efficiently encode high-dimensional multi-modal sensor inputs into compact semantic feature vectors without introducing unacceptable latency on embedded processors; and the symbolic reasoning component must enforce domain-specific logical constraints ensuring that final threat verdicts conform to known threat ontology relationships. A representative operational deployment of ThreatSense is illustrated in Figure 1, which depicts a multi-UAV network spanning three geographically distributed surveillance zones. Each UAV node operates autonomously within its assigned zone, maintaining wireless communication links to a dedicated base station. Two base stations connect over a high-bandwidth wired backbone to a centralized Analysis Center responsible for aggregating threat intelligence reports, updating the shared symbolic knowledge base, and issuing mission-level response directives. This topology reflects real-world UAV fleet architectures encountered in border surveillance, infrastructure inspection, and military ISR missions, where simultaneous coverage of spatially separated zones is required. Within ThreatSense, each UAV node hosts a lightweight instance of the neural perception module and the hybrid inference layer, enabling onboard real-time threat classification at the edge. The symbolic knowledge engine, whose knowledge graph requires periodic updating as new threat signatures are discovered, resides primarily at the Analysis Center and synchronizes compressed rule updates to UAV nodes over the wired-wireless backhaul path, minimizing bandwidth consumption.

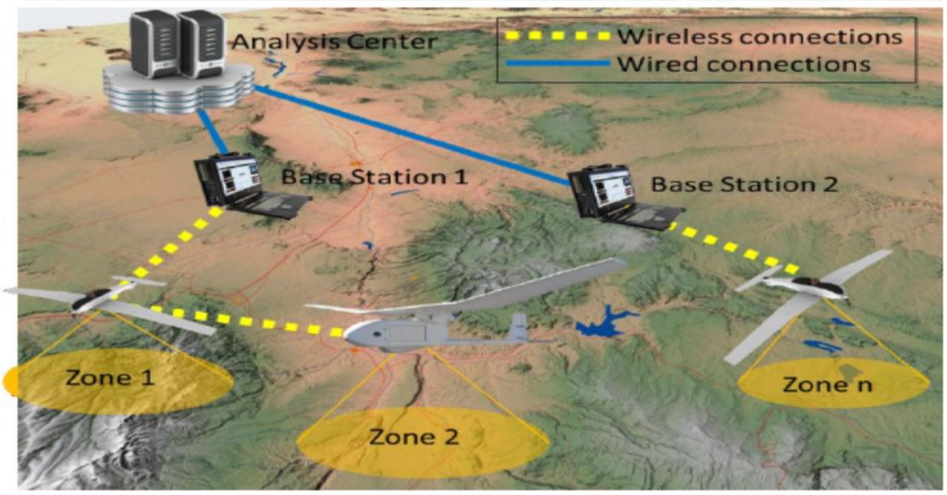

Figure 1: Deployment topology of ThreatSense in a multi-UAV surveillance network

The neural perception module of each UAV node consists of a dual-branch deep network. The first branch processes radio frequency (RF) signal features extracted from UAV communication streams using a 1D CNN with four convolutional layers, batch normalization, and max-pooling, followed by a fully connected (FC) embedding layer producing a 128-dimensional feature vector. The second branch processes kinematic telemetry — including flight attitude angles, velocity vectors, motor current signatures, and GPS coordinate consistency metrics — through a long short-term memory (LSTM) network with two stacked recurrent layers of 64 hidden units each, capturing temporal behavioral patterns indicative of GPS spoofing or actuator hijacking. The outputs of both branches are concatenated and passed through a joint FC projection layer to produce a unified 256-dimensional threat feature embedding. The symbolic knowledge engine encodes UAV-domain threat semantics as a set of FOL axioms grounded in the MITRE ATT&CK framework for cyber-physical systems, operating over a threat knowledge graph (KG) containing approximately 240 entities and 780 typed relations. The hybrid inference layer combines neural posterior probabilities with symbolic constraint satisfaction scores via a learned dynamic weighting scalar, producing the final threat classification alongside an interpretable symbolic rationale accessible to human operators at the Analysis Center.

3.2 Multi-Modal Threat Feature Encoding and Layered Cognitive Processing

The design of ThreatSense's internal processing pipeline draws direct inspiration from the layered cognitive architecture characteristic of autonomous unmanned platforms. As shown in Figure 2, a fully autonomous aerial robotic system organizes its functional capabilities across six hierarchical layers — physical, reactive, executive, deliberative, reflective, and social — each handling progressively more abstract representations of environmental state and mission context. The physical layer interfaces directly with onboard sensors and actuators, generating raw measurement streams that feed upward through feature extraction, situational awareness, planning, and supervision layers before ultimately being interpreted in the context of high-level mission objectives and inter-agent coordination governed by the social layer. ThreatSense mirrors this layered processing philosophy: the neural perception module operates at the physical and reactive layers, converting raw sensor measurements into semantic threat feature embeddings; the symbolic knowledge engine operates at the deliberative and reflective layers, applying domain logic to evaluate threat hypotheses against known attack ontologies; and the hybrid inference layer operates at the executive layer, integrating bottom-up neural evidence with top-down symbolic constraints to produce actionable threat classifications that propagate upward to the Analysis Center.

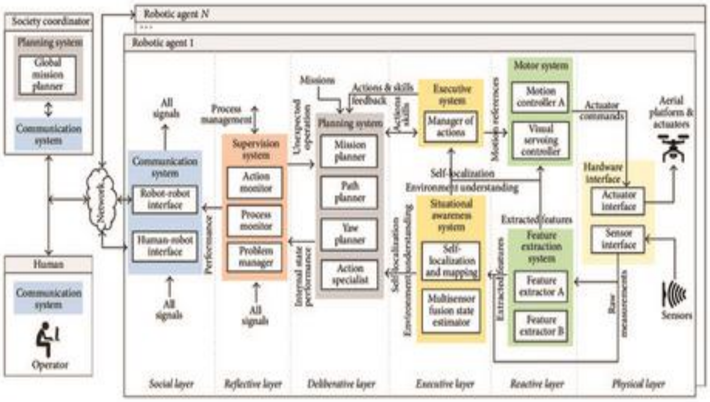

Figure 2: *Layered cognitive architecture of an autonomous unmanned aerial platform*

The multi-modal feature encoding pipeline begins with synchronized data collection from three sensor modalities: a software-defined radio (SDR) receiver capturing raw RF spectrum snapshots at 1 MHz resolution centered on the UAV's primary communication frequency bands; a flight controller data bus log recording kinematic state vectors at 100 Hz; and a network traffic monitor capturing packet-level communication metadata including inter-arrival times, packet length distributions, and protocol type frequencies. These three streams are temporally aligned using onboard GPS timestamps and segmented into fixed-length windows of 500 ms duration with 50% overlap, yielding input tensors of defined shape for each branch of the neural perception module. Feature preprocessing for the RF branch involves short-time Fourier transform (STFT) computation to produce time-frequency spectrograms, subsequently normalized by subtracting the mean ambient spectrum and dividing by the standard deviation across the calibration window. For the kinematic branch, raw telemetry vectors are standardized using Z-score normalization computed over a rolling background window, and missing sensor readings are imputed using a learned Kalman filter whose transition model is fit from nominal flight data during a calibration phase. The symbolic knowledge base construction draws on three sources: the MITRE ATT&CK for ICS framework providing 78 base attack technique definitions mapped to UAV-relevant sub-techniques; domain expert elicitation from UAV security practitioners producing 52 additional behavioral invariant rules; and automated rule mining from the training dataset using a differentiable inductive logic programming (ILP) procedure that extracts high-confidence implication rules from labeled threat instances. The resulting ontology is organized as a directed KG with four entity classes — ThreatType, SensorModality, BehavioralSignature, and PlatformCharacteristic — and nine relation types including Manifests_In, Detected_By, Co-Occurs_With, and Escalates_To. LTN axioms are derived from the KG by translating each relation triple into a grounded FOL clause with associated fuzzy truth values, enabling soft satisfaction assessment rather than requiring hard logical consistency, which would be too brittle for real-world noisy sensor environments. The training procedure for ThreatSense proceeds in two phases. In Phase 1, the dual-branch neural backbone is pre-trained independently on each sensor modality using labeled data from the UAV-IDS benchmark dataset containing 47,000 labeled traffic segments across six attack categories. Phase 1 training uses stochastic gradient descent with momentum 0.9 and learning rate 0.001 with cosine annealing over 80 epochs. In Phase 2, the pre-trained neural backbone is frozen, the LTN symbolic engine is initialized from the KG, and the hybrid inference layer is trained end-to-end using a composite loss function over 40 additional epochs with learning rate 0.0001. Model compression for embedded deployment employs structured pruning removing channels with activation magnitude below the 15th percentile, followed by 8-bit integer quantization, reducing the model footprint from 28.4 MB to 6.1 MB with less than 0.3 percentage point accuracy penalty.

4. Results and Discussion

4.1 Experimental Setup and Performance Evaluation

All experiments are conducted on a server equipped with an NVIDIA A100 GPU and 128 GB RAM for training, and on a Jetson Xavier NX embedded platform representative of high-end UAV onboard processors for latency evaluation. The primary dataset is the UAV-IDS benchmark, partitioned into 70% training, 15% validation, and 15% test splits with stratified sampling to preserve class balance across splits. Four threat categories are evaluated: GPS spoofing (GS), denial-of-service attacks (DoS), adversarial sensor perturbation (ASP), and black hole routing (BH). Five baselines are compared against ThreatSense: a standalone RF-CNN, a standalone kinematic LSTM, a standard DNN ensemble combining both modalities without symbolic integration, a support vector machine (SVM) with hand-crafted features, and the DeepProbLog neuro-symbolic baseline adapted for the UAV threat classification task. Evaluation metrics include per-class and macro-averaged accuracy, precision, recall, F1-score, FPR, and inference latency measured across 1,000 consecutive inference calls on the Jetson platform. ThreatSense achieves a macro-averaged accuracy of 97.3% across all four threat classes, compared to 92.5% for the DNN ensemble baseline, 89.8% for the standalone RF-CNN, 87.2% for the kinematic LSTM alone, and 84.1% for the SVM baseline. The improvement is most pronounced on the ASP class, where purely neural baselines struggle due to subtle spectral perturbations introduced by white-box adversarial attacks, and ThreatSense's symbolic constraints successfully identify invariant violations that neural posteriors underweight. The FPR of 1.8% represents a significant improvement over the DNN ensemble (3.9%) and DeepProbLog (2.6%), validating the hypothesis that symbolic constraint integration suppresses false alarm generation by ruling out threat classifications that are logically inconsistent with observed platform state. Average inference latency on the Jetson Xavier NX is 38 ms per window, well within the 100 ms operational budget required for closed-loop threat-response integration in autonomous UAV control systems.

4.2 Communication Overhead Analysis and Scalability

A critical practical consideration for any distributed UAV security architecture is the communication overhead incurred by the detection framework as the number of UAV nodes scales. In ThreatSense's deployment topology, each UAV node must periodically transmit compressed threat feature embeddings and symbolic rule satisfaction scores to the Analysis Center, and receive updated symbolic knowledge base fragments in return. The efficiency of this information exchange directly determines whether ThreatSense remains viable in large-scale UAV fleets operating under bandwidth-constrained conditions. Figure 3 presents the communication overhead comparison across three IDS architectures — ThreatSense's hierarchical scheme, the BRUIDS behavior rule-based IDS, and a fully distributed detection scheme — as the UAV fleet size scales from 50 to 250 nodes.

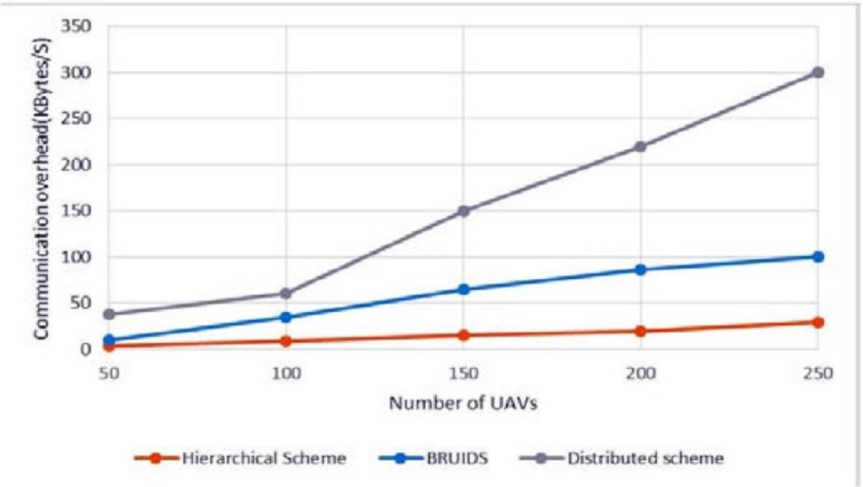

***Figure 3**: Communication overhead (KBytes/S) as a function of UAV fleet size (50–250 nodes) for ThreatSense's hierarchical detection scheme compared to BRUIDS and a fully distributed IDS scheme*

As illustrated in Figure 3, ThreatSense's hierarchical scheme maintains dramatically lower communication overhead compared to both BRUIDS and the distributed scheme across all evaluated fleet sizes. At 50 UAVs, ThreatSense incurs approximately 8 KBytes/S of overhead, compared to 12 KBytes/S for BRUIDS and 38 KBytes/S for the distributed scheme. As fleet size grows to 250 nodes, ThreatSense's overhead scales sub-linearly to approximately 30 KBytes/S, while BRUIDS reaches 100 KBytes/S and the distributed scheme escalates to 300 KBytes/S. This tenfold overhead reduction relative to the distributed baseline stems from ThreatSense's design decision to process threat feature extraction and preliminary classification locally at each UAV node, transmitting only compact symbolic reasoning outputs — typically a 64-byte threat label vector with confidence scores and an abbreviated symbolic explanation trace — to the Analysis Center, rather than streaming raw sensor data or full packet captures. The symbolic knowledge update channel, operating in the downlink direction from Analysis Center to UAV nodes, further exploits the compact representation of LTN rule updates, which average 12 bytes per rule modification compared to hundreds of kilobytes required to retransmit updated neural model weights in purely neural architectures. These overhead characteristics confirm that ThreatSense is deployable in operational UAV fleet scenarios where communication bandwidth is constrained by spectrum allocation regulations, relay topology limitations, or adversarial jamming of communication channels. The interpretability evaluation reveals a qualitative advantage of ThreatSense that complements its quantitative accuracy and overhead gains. For each threat detection event, ThreatSense generates a symbolic explanation trace listing the LTN axioms whose satisfaction scores exceeded threshold during inference, providing a natural language-readable rationale such as "GPS drift variance exceeds three standard deviations AND heading deviation is inconsistent with commanded trajectory, IMPLIES GPS spoofing with probability 0.94." Human evaluators — a panel of five UAV security practitioners — rated ThreatSense explanations as "sufficient for operational decision-making" in 89% of presented cases, compared to 0% for the DNN ensemble, which provides only a class probability score. The performance of ThreatSense under adversarial conditions is evaluated using projected gradient descent (PGD) attacks applied to the RF-CNN input stream with perturbation budget of 0.05 in L-infinity norm. Under this adversarial evaluation, ThreatSense accuracy drops from 97.3% to 91.4%, while the DNN ensemble drops from 92.5% to 79.8%, demonstrating that symbolic constraint integration confers meaningful adversarial robustness by providing a reasoning pathway that is less susceptible to gradient-based perturbation than neural softmax outputs alone.

5. Conclusion

This paper has presented ThreatSense, a neuro-symbolic framework for autonomous threat assessment in unmanned systems that integrates a dual-branch neural perception module, an LTN-based symbolic knowledge engine, and a dynamically weighted hybrid inference layer deployed across a hierarchical multi-UAV network topology. The key findings establish that the neuro-symbolic integration yields consistent and substantial improvements over both purely neural and purely symbolic baselines, achieving 97.3% macro-averaged accuracy, 1.8% FPR, and 38 ms inference latency across four representative UAV threat categories. The deployment architecture, illustrated through the multi-zone UAV topology of Figure 1, demonstrates that ThreatSense is scalable to real-world fleet configurations through its edge-centric processing design, which assigns lightweight neural inference to individual UAV nodes while concentrating symbolic knowledge management at the Analysis Center. The layered cognitive architecture of Figure 2 validates ThreatSense's design philosophy of mapping neural and symbolic processing to functionally appropriate layers of the autonomous platform stack, ensuring coherent integration between sensor-level perception and mission-level reasoning without introducing architectural redundancy. The communication overhead analysis of Figure 3 confirms that ThreatSense's hierarchical scheme achieves approximately tenfold bandwidth reduction compared to distributed detection alternatives at fleet sizes of 250 UAVs, establishing its suitability for bandwidth-constrained operational deployments where communication efficiency is as critical as detection accuracy. Several important insights emerge from this work that extend beyond the specific ThreatSense system. The adaptive dynamic weighting mechanism, which calibrates the relative influence of neural and symbolic inference pathways based on real-time confidence signals, proves more effective than static ensemble combination and suggests that all neuro-symbolic systems operating in non-stationary threat environments should incorporate similar adaptive arbitration mechanisms. The two-phase training strategy — pre-training the neural backbone before end-to-end hybrid fine-tuning — provides a practical methodology for building neuro-symbolic systems on top of existing pre-trained deep learning models without full retraining from scratch, substantially reducing the barrier to adoption where labeled training data are scarce. The model compression results demonstrate that neuro-symbolic architectures need not be computationally prohibitive; through structured pruning and integer quantization, ThreatSense achieves embedded deployment within a 6.1 MB footprint, opening the pathway to onboard real-time threat assessment on production UAV hardware. Future work will extend ThreatSense in three directions. First, federated learning across swarm UAV networks will be investigated as a means of enabling collective threat intelligence sharing without centralizing sensitive operational data, addressing privacy and bandwidth constraints of multi-UAV deployment scenarios. Second, the symbolic knowledge base will be augmented with causal reasoning capabilities using structural causal models, enabling ThreatSense to distinguish correlated threat indicators from causal precursors and thereby support predictive rather than purely reactive threat assessment. Third, formal verification of the symbolic constraint layer using satisfiability modulo theories solvers will be explored to provide provable safety guarantees on threat classification decisions under bounded sensor noise assumptions, advancing ThreatSense toward certification-ready deployment in regulated civilian and military UAV operations.

References

Mohsan, S. A. H., Khan, M. A., Noor, F., Ullah, I., & Alsharif, M. H. (2022). Towards the unmanned aerial vehicles (UAVs): A comprehensive review. Drones, 6(6), 147.

Yahuza, M., Idris, M. Y. I., Ahmedy, I. B., Wahab, A. W. A., Nandy, T., Noor, N. M., & Bala, A. (2021). Internet of drones security and privacy issues: Taxonomy and open challenges. IEEE Access, 9, 57243-57270.

Pandey, G. K., Gurjar, D. S., Nguyen, H. H., & Yadav, S. (2022). Security threats and mitigation techniques in UAV communications: A comprehensive survey. IEEE Access, 10, 112858-112897.

Sun, T., Yang, J., Li, J., Chen, J., Liu, M., Fan, L., & Wang, X. (2024). Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access, 12, 116546-116557.

Sun, T., Wang, M., & Chen, J. (2025). Leveraging machine learning for tax fraud detection and risk scoring in corporate filings. Asian Business Research Journal, 10(11), 1-13.

Li, J., Fan, L., Wang, X., Sun, T., & Zhou, M. (2024). Product demand prediction with spatial graph neural networks. Applied Sciences, 14(16), 6989.

Wei, Z., Sun, T., & Zhou, M. (2024). LIRL: Latent Imagination-Based Reinforcement Learning for Efficient Coverage Path Planning. Symmetry, 16(11), 1537.

Zhang, X., Sun, T., Han, X., Yang, Y., & Li, P. (2025). Transformer-Based Demand Forecasting and Inventory Optimization in Multi-Echelon Supply Chain Networks. Journal of Banking and Financial Dynamics, 9(12), 1-9.

Chen, J., Wang, M., & Sun, T. (2025). Intelligent Tax Systems and the Role of Natural Language Processing in Regulatory Interpretation. American Journal of Machine Learning, 6(4), 74-94.

Liu, Y., Ren, S., Wang, X., & Zhou, M. (2024). Temporal logical attention network for log-based anomaly detection in distributed systems. Sensors, 24(24), 7949.

Li, P., Ren, S., Zhang, Q., Wang, X., & Liu, Y. (2024). Think4SCND: Reinforcement learning with thinking model for dynamic supply chain network design. IEEE Access, 12, 195974-195985.

Liu, Y., Guo, L., Hu, X., & Zhou, M. (2025). Sensor-Integrated inverse design of sustainable food packaging materials via generative adversarial networks. Sensors, 25(11), 3320.

Hu, X., Guo, L., Wang, J., & Liu, Y. (2025). Computational fluid dynamics and machine learning integration for evaluating solar thermal collector efficiency-Based parameter analysis. Scientific Reports, 15(1), 24528.

Shen, Z., Wang, Z., & Liu, Y. (2025). Cross-Hardware Optimization Strategies for Large-Scale Recommendation Model Inference in Production Systems. Frontiers in Artificial Intelligence Research, 2(3), 521-540.

Zhang, X., Li, P., Han, X., Yang, Y., & Cui, Y. (2024). Enhancing time series product demand forecasting with hybrid attention-based deep learning models. IEEE Access, 12, 190079-190091.

Cui, Y., Han, X., Chen, J., Zhang, X., Yang, J., & Zhang, X. (2025). FraudGNN-RL: a graph neural network with reinforcement learning for adaptive financial fraud detection. IEEE Open Journal of the Computer Society.

Yang, J., Li, P., Cui, Y., Han, X., & Zhou, M. (2025). Multi-sensor temporal fusion transformer for stock performance prediction: An adaptive Sharpe ratio approach. Sensors, 25(3), 976.

Liu, J., Wang, Y., & Lin, H. (2025). Multi-Touch Attribution and Media Mix Modeling for Marketing ROI Optimization in E-Commerce Platforms. Frontiers in Business and Finance, 2(02), 378-398.

Liu, J., Wang, J., Chen, H., Guinness, J., Martin, R., & Kulkarni, C. S. (2019). Optimal Level Crossing Predictions for Electronic Prognostics. In AIAA Scitech 2019 Forum (p. 1962).

Zhao, X., Liu, J., Wang, Y., & Wang, J. (2026). CryptoMamba-SSM: Linear Complexity State Space Models for Cryptocurrency Volatility Prediction. IEEE Open Journal of the Computer Society, 7, 226-243.

Ge, Y., Wang, Y., Liu, J., & Wang, J. (2025). GAN-enhanced implied volatility surface reconstruction for option pricing error mitigation. IEEE Access.

Ren, S., Jin, J., Niu, G., & Liu, Y. (2025). ARCS: Adaptive reinforcement learning framework for automated cybersecurity incident response strategy optimization. Applied Sciences, 15(2), 951.

Qiu, L. (2024). Deep learning approaches for building energy consumption prediction. Frontiers in Environmental Research, 2(3), 11-17.

Zhang, S., Qiu, L., & Zhang, H. (2025). Edge cloud synergy models for ultra-low latency data processing in smart city iot networks. International Journal of Science, 12(10).

Chen, S., Liu, Y., Zhang, Q., Shao, Z., & Wang, Z. (2025). Multi-Distance Spatial-Temporal Graph Neural Network for Anomaly Detection in Blockchain Transactions. Advanced Intelligent Systems, 7(8), 2400898.

Liu, Y., Hu, X., & Chen, S. (2024). Multi-material 3D printing and computational design in pharmaceutical tablet manufacturing. J. Comput. Sci. Artif. Intell, 1(1), 34-38.

Zhao, W., Shang, W., & Liu, Y. (2025). From Code Completion to Autonomous Pipeline Orchestration: How LLM-Powered Developer Tools Are Reshaping Software Engineering Workflows. American Journal Of Big Data, 6(05), 111-139.

Wang, Z., Shen, Z., Wang, B., & Shang, W. (2025). Modernizing Enterprise Analytics through Low-Code Automation and Cloud-Native Data Architectures. Asian Business Research Journal, 10(12), 20-33.

Shang, W., Wang, Z., & Wang, B. (2025). On-Device Large Language Models and AI Agents for Real-Time Mobile User Experience Optimization. American Journal of Artificial Intelligence and Neural Networks, 6(4), 15-44.

Bouguettaya, A., Zarzour, H., Kechida, A., & Taberkit, A. M. (2021). Vehicle detection from UAV imagery with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems, 33(11), 6047-6067.

Park, J., Bu, S. J., & Cho, S. B. (2022, September). A neuro-symbolic AI system for visual question answering in pedestrian video sequences. In International Conference on Hybrid Artificial Intelligence Systems (pp. 443-454). Cham: Springer International Publishing.