Journal of Cybersecurity and Information Management

♦ ISSN: 2690-6775 Vol: 18, No. 01, 2026 ♦ PP. 34–39

Article DOI: https://doi.org/10.54216/JCIM.180105

FULL LENGTH ARTICLE

Risk-Aware Cyberattack Analytics for Unmanned Aerial Vehicle Communications: A Publication-Ready Gradient-Boosting Framework

Andino Maseleno[1,*], **Aa Hubur**[2]

[1]Institut Bakti Nusantara, Lampung, Indonesia

[2]Universitas Trisakti, Jakarta, Indonesia

Emails: andino.maseleno@ibnus.ac.id; Maa.hubur@trisakti.ac.id

Received: January 19, 2026 Revised: February 15, 2026 Accepted: March 29, 2026 ⋆ Corresponding author

ABSTRACT

Cyberattack detection in unmanned aerial vehicle environments has become an essential requirement for dependable digital operations. Security analytics for these environments should not only separate benign and malicious traffic, but should also provide interpretable evidence that can support timely triage and intervention. This paper presents a risk-aware classification framework for UAV communication security based on a leakage-screened feature design and a gradient-boosting ensemble model. The framework combines multiclass discrimination, probability-based decision logic, and feature-level interpretation within one coherent workflow. The study demonstrates that a carefully designed ensemble approach can provide balanced and operationally meaningful cyberattack recognition while remaining transparent enough for practical cybersecurity management. The results also show that communication-structure variables provide strong discriminatory power and that replay-type activity remains more difficult to separate than benign or denial-of-service behavior. The proposed framework therefore contributes a reproducible analytical design and a managerial reading of cyberattack classification for UAV operations.

Keywords: UAV cybersecurity ▪ Cyberattack analytics ▪ Gradient boosting ▪ Intrusion detection ▪ Multiclass classification ▪ Interpretable security

1. INTRODUCTION

Unmanned aerial vehicles are increasingly deployed in logistics, inspection, agriculture, emergency response, monitoring, and smart-city services. As these platforms become more deeply connected to digital infrastructures, their communication channels become strategic cyber assets. A disruption or manipulation at the communication layer may affect mission continuity, information integrity, and the reliability of downstream operational decisions. For this reason, cyberattack monitoring in UAV systems has evolved from a specialized engineering topic into a wider problem of cyber-risk management.

Recent research confirms that UAV cybersecurity is now an active and technically diverse area. Public benchmark construction, cyber-physical intrusion detection, lightweight defense architectures, sequence-learning approaches, and explainable analytics have all been studied in recent years [1–6]. This body of work has improved the empirical basis of the field and clarified that attack detection in UAV networks must reconcile performance, interpretability, and deployment practicality.

Despite this progress, two issues remain important. First, a large part of the literature emphasizes predictive performance more strongly than operational usefulness. In many real set-

tings, the practical value of a detection model depends on whether its outputs can guide immediate response actions such as allowing traffic, escalating it for inspection, or isolating suspicious flows. Second, publicly released traffic datasets may contain ordering variables or sequence identifiers that simplify the learning problem in ways that do not reflect deployment conditions. If these variables are not screened, reported performance may overstate the model's genuine cyberanalytic capability.

The present study addresses these issues through a risk-aware analytical framework for cyberattack classification in UAV communications. The framework is developed on a recent public UAV cyberattack dataset and is designed to remove order-driven leakage before model estimation. A gradient-boosting ensemble is used as the focal model and is compared against two strong tree-based baselines under a common evaluation protocol. The analysis is intentionally framed around balanced multiclass recognition, operational triage, and feature-level interpretation rather than narrow benchmark optimization.

The study makes five contributions: it develops a leakage-screened modeling workflow; formulates the gradient-boosting model and links posterior probabilities to an operational action rule; benchmarks the focal model against Random Forest and Extra Trees; reports an extended empirical evidence set with multiple tables, figures, class-level analyses, and feature-importance interpretations; and discusses remaining research challenges for public-data-driven cyberattack analytics in UAV operations.

2. RELATED WORK

Research on cybersecurity for UAV communication systems has developed around several interconnected lines. One line emphasizes reliable empirical foundations. Hassler et al. introduced a cyber-physical intrusion detection framework for UAVs and provided a public benchmark relevant for reproducible research [1]. Dataset-oriented studies have continued to examine the adequacy and limitations of UAV intrusion data [7].

A second line focuses on algorithmic and architectural innovation. Distributed intrusion detection and coordinated defense logic have been proposed for UAV attack scenarios [3]. Sequence-learning approaches, including recurrent neural architectures, have also been investigated to improve attack recognition in drone networks [4]. These approaches often achieve strong predictive performance, but may require higher implementation complexity or reduced transparency compared with structured ensemble learners.

A third line concerns interpretability and deployment efficiency. Explainability has become central in cybersecurity analytics, especially when predictions are expected to inform operational decisions [5, 6]. Lightweight deployment has also been explored through knowledge distillation and compact-learning strategies for constrained or resource-sensitive environments [8]. This trend is especially relevant to UAV settings, where computational resources and communication overhead may be limited.

A fourth line relates UAV security to the broader intrusion-detection literature in IoT and connected systems. Surveys and review studies show that machine-learning-powered intrusion detection depends heavily on data representation quality, evaluation discipline, and generalization across heterogeneous operational conditions [2, 9]. Public benchmarks such as CICIoT2023 further show the importance of scalable and reproducible cyberattack datasets [10].

Taken together, the literature reveals a clear opportunity. Many studies propose new architectures or review the field, but fewer present an explicitly operational view in which multiclass predictions are mapped to cyber-risk triage actions and evaluated under a leakage-aware design. The present work is positioned in this space by integrating careful feature screening, balanced multiclass modeling, probability-based decision mapping, and feature-level interpretation into a publication-ready analytical framework.

Table 1. Representative peer-reviewed studies published after 2022.

Study	Context	Main method	Main relevance to the present study
Hassler et al. (2024) [1]	UAV cyber-physical traffic	Cyber-physical IDS	Established the public UAV benchmark family used in this study.
AL-Syouf et al. (2024) [2]	UAV IDS literature	Review	Summarized machine-learning trends and open issues in UAV intrusion detection.
Tlili et al. (2024) [3]	UAV attacks	Distributed IDS	Emphasized distributed detection and security enforcement logic.
Gamal et al. (2024) [4]	Drone network attacks	LSTM-RNN	Demonstrated the value of sequence-oriented models in drone protection.
Sharma et al. (2024) [5]	IoT intrusion data	Explainable deep learning	Reinforced the importance of explanation-centered intrusion detection.
Manivannan (2024) [9]	IoT IDS literature	Review	Highlighted reproducibility, realism, and deployment concerns in IDS research.
Mohammed and Fourati (2025) [7]	UAV IDS datasets	Dataset analysis	Examined the impact of dataset quality on UAV intrusion detection.
Wisanwanichthan and Thammawichai (2025) [8]	IoT and UAV intrusion	Knowledge-distilled DNN	Addressed lightweight deployment in constrained settings.
Ihekoronye et al. (2025) [6]	Drone intrusion detection	Explainable ML	Connected interpretability with drone network intrusion detection.
Neto et al. (2023) [10]	IoT cyberattack benchmark	Dataset/benchmark	Showed the value of recent public benchmarks in cybersecurity analytics.

3. THE PROPOSED MODEL

The proposed model is developed as a risk-aware analytical workflow for cyberattack classification in UAV communication environments. Its design begins from the premise that an operational classifier should do more than maximize prediction accuracy. It should also remain interpretable, reduce the influence of leakage-prone variables, and produce outputs that can be translated into practical response categories. For this reason, the proposed design combines leakage screening, multiclass ensemble learning, probability-based decision logic, and feature-level interpretation within a single coherent pipeline.

A central modeling principle is that attack recognition must be grounded in communication behavior rather than collection artifacts. Public traffic datasets may include timestamps, sequence identifiers, or ordering cues that simplify learning without representing stable cyberattack structure. The proposed model therefore excludes order-driven fields before training and focuses on protocol-level and communication-structure descriptors.

The framework in Figure 1 summarizes the complete analytical flow. It begins with public UAV cyber data, applies leakage screening and preprocessing, constructs the feature matrix, and routes the resulting representation into the proposed gradient-boosting learner together with benchmark ensembles. The final stage converts the selected classifier output into an operational rule that distinguishes normal traffic, replay-like suspicious activity, and denial-of-service conditions.

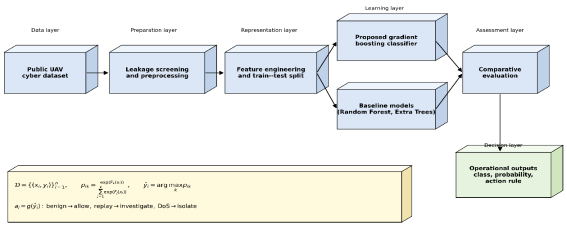

Figure 1. The framework of the proposed risk-aware cyberattack analytics workflow.

Algorithm 1 Risk-aware UAV cyberattack classification workflow

Require: Labeled UAV communication records $\mathcal{D}$
Ensure: Predicted traffic class $\hat{y}$ and operational action a
Remove leakage-prone ordering fields from $\mathcal{D}$
Clean and encode the retained communication features
Partition the sample into stratified training and testing subsets
Train the proposed gradient-boosting model and the two baseline ensembles
Evaluate all models using accuracy, macro precision, macro recall, macro-F1, weighted F1, and macro AUC
Select the model with the strongest balanced performance
for each testing record x_i **do**
  Compute class posteriors p_{ik} using the selected model
  Assign the predicted class $\hat{y}_i = \arg\max_k p_{ik}$
  if $\hat{y}_i$ is benign **then**
    $a_i \leftarrow$ allow
  else if $\hat{y}_i$ is replay **then**
    $a_i \leftarrow$ investigate
  else
    $a_i \leftarrow$ isolate
  end if
end for
Produce model-comparison, confusion, class-level, and feature-importance analyses

Algorithm 1 presents the proposed model in procedural form. It starts with data preparation, continues through model training and comparative evaluation, and ends with class prediction, action mapping, and interpretive analysis. This complements the framework figure by clarifying how the method can be implemented in practice.

Let $D = \{(x_i, y_i)\}_{i=1}^{n}$ denote the leakage-screened training sample, where $x_i \in \mathbb{R}^d$ is the feature vector of the ith communication record and $y_i \in Y = \{1,2,3\}$ is its class label. For each class $k \in Y$, the additive score function is

$$F_k(x_i) = \sum_{m=1}^{M} f_{m,k}(x_i), \tag{1}$$

where $f_{m,k}(\cdot)$ denotes the contribution of the mth decision tree to class k, and M is the number of boosting rounds. The posterior probability of class k is computed using the softmax transformation

$$p_{ik} = \Pr(y_i = k \mid x_i) = \frac{\exp(F_k(x_i))}{\sum_{j=1}^{K} \exp(F_j(x_i))}, \qquad K = 3. \tag{2}$$

The final predicted class is

$$\hat{y}_i = \arg\max_{k \in Y} p_{ik}. \tag{3}$$

The learning problem minimizes a regularized multinomial objective,

$$\mathscr{L} = \sum_{i=1}^{n} \ell(y_i, p_i) + \sum_{m=1}^{M} \Omega(f_m), \tag{4}$$

with

$$\Omega(f_m) = \gamma T_m + \frac{\lambda}{2} \|w_m\|^2, \tag{5}$$

where T_m is the number of leaves in tree m, w_m is the vector of leaf weights, and γ and λ are regularization parameters.

For operational interpretation, the posterior vector $p_i = (p_{i1}, p_{i2}, p_{i3})$ is translated into a simple action rule. Let $A = \{\text{allow}, \text{investigate}, \text{isolate}\}$. The decision function is

$$a_i = g(\hat{y}_i) = \begin{cases} \text{allow}, & \hat{y}_i = \text{benign}, \\ \text{investigate}, & \hat{y}_i = \text{replay}, \\ \text{isolate}, & \hat{y}_i = \text{DoS}. \end{cases} \tag{6}$$

A useful analytical quantity is the classification margin,

$$\Delta_i = p_{i\hat{y}_i} - \max_{j \neq \hat{y}_i} p_{ij}, \tag{7}$$

which measures the separation between the most likely class and its closest alternative.

The model is evaluated using accuracy, macro precision, macro recall, macro-F1, weighted F1, and one-vs-rest macro AUC. Accuracy is

$$\text{Accuracy} = \frac{1}{n_{test}} \sum_{i=1}^{n_{test}} \mathbb{I}(\hat{y}_i = y_i), \tag{8}$$

while macro-F1 is

$$\text{Macro-F1} = \frac{1}{K} \sum_{k=1}^{K} \frac{2\,\text{Precision}_k\,\text{Recall}_k}{\text{Precision}_k + \text{Recall}_k}. \tag{9}$$

The feature-importance score of variable q is interpreted as

$$I_q = \sum_{m=1}^{M} \sum_{s \in S_{m,q}} \Delta L_{m,s}, \tag{10}$$

where $S_{m,q}$ denotes the set of nodes in tree m that split on feature q.

From a computational viewpoint, the approximate training complexity is governed by tree construction across M boosting rounds and can be viewed as proportional to $O(Mnd \log n)$ under standard tree-building assumptions.

Table 2. Hyperparameter settings of the evaluated models.

Model	Setting
Proposed gradient-boosting model	*n_estimators* = 300; *max_depth* = 6; *learning_rate* = 0.05; *subsample* = 0.90; *colsample_bytree* = 0.80
Random Forest	*n_estimators* = 300; *max_depth* = *None*; *min_samples_split* = 2; bootstrap enabled
Extra Trees	*n_estimators* = 300; *max_depth* = *None*; *min_samples_split* = 2; bootstrap disabled

3.1 Dataset profile and descriptive view

The proposed model is evaluated on the labeled cyber subset of a public UAV cyberattack dataset associated with recent UAV intrusion detection research. The final analytical sample contains benign traffic together with denial-of-service and replay attack records. Table 3 reports the size and composition of the experiment, while Figure 2 provides a compact descriptive overview of the data setting from multiple visual angles.

Table 3. Dataset and experimental profile.

Item	Value
Labeled cyber records	33,102
Traffic classes	3
Training records	26,481
Testing records	6,621
Test DoS support	2,334
Test Replay support	2,402
Test benign support	1,885
Models compared	3

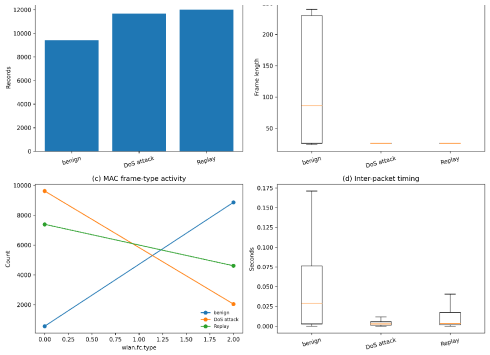

Figure 2: Four-view descriptive summary of the empirical setting.

Figure 2. Four-view descriptive summary of the empirical setting.

4. EXPERIMENTAL RESULTS

Table 4 reports the held-out comparison of the three evaluated models. The proposed gradient-boosting model achieves the best result on every reported metric. Its advantage is visible in overall accuracy, macro-F1, and macro AUC, indicating stronger class balance and better multiclass separation. Table 5 shows that the proposed model improves macro-F1 by more than five percent over Random Forest and by more than seven percent over Extra Trees. Figure 3 visualizes the same pattern.

Table 4. Overall held-out model comparison.

Model	Accuracy	Macro Prec.	Macro Rec.	Macro-F1	Weighted F1	Macro AUC	Train (s)
Proposed gradient boosting	0.766	0.769	0.782	0.769	0.759	0.915	7.40
Random Forest	0.721	0.727	0.733	0.730	0.719	0.884	3.40
Extra Trees	0.704	0.712	0.717	0.714	0.703	0.863	3.30

Table 5. Relative gain of the proposed model over the baseline ensembles (%).

Baseline	Accuracy gain	Macro-F1 gain	Weighted F1 gain	Macro AUC gain
Random Forest	6.22	5.37	5.46	3.56
Extra Trees	8.75	7.63	7.95	6.03

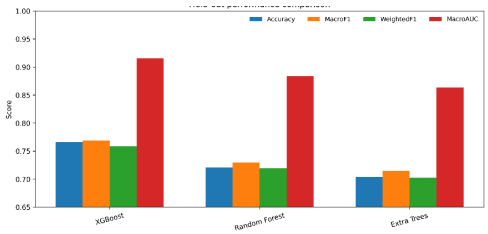

Figure 3. Comparative performance of the evaluated models.

The class-specific results provide a more nuanced view of model behavior. Table 6 shows that benign traffic is identified with the strongest recall and F1-score, indicating stable separation of normal communication patterns. DoS traffic is also recognized effectively, whereas replay traffic remains the most challenging class. Figure 4 complements the table by presenting the class-level metric profile and support distribution.

Table 6. Class-wise performance of the proposed model on the held-out test set.

Class	Precision	Recall	F1-score	Support
DoS attack	0.703	0.772	0.736	2,334
Replay	0.766	0.588	0.665	2,402
Benign	0.839	0.985	0.906	1,885
Macro average	0.769	0.782	0.769	6,621
Weighted average	0.764	0.766	0.759	6,621

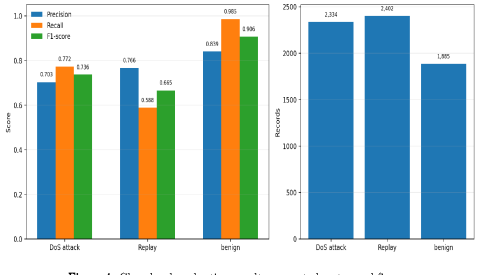

Figure 4. Class-level evaluation results presented as two subfigures.

The confusion analysis offers additional insight into residual errors. Table 7 shows that most remaining errors arise from confusion between replay and DoS traffic, whereas benign traffic is misclassified only rarely. Table 8 quantifies the same observation in terms of misclassification rates, and Figures 5 and 6 present the standard confusion heatmap and a normalized error view.

Table 7. Confusion matrix of the proposed model.

Actual class	Predicted DoS	Predicted Replay	Predicted benign
DoS attack	1,802	406	126
Replay	760	1,412	230
Benign	3	26	1,856

Table 8. Error decomposition by actual class.

Class	Support	Correct	Misclassified	Misclassification rate (%)
DoS attack	2,334	1,802	532	22.79
Replay	2,402	1,412	990	41.22
Benign	1,885	1,856	29	1.54

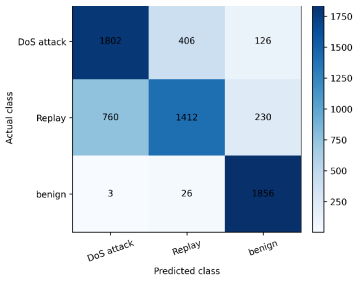

Figure 5. Confusion matrix heatmap of the proposed model.

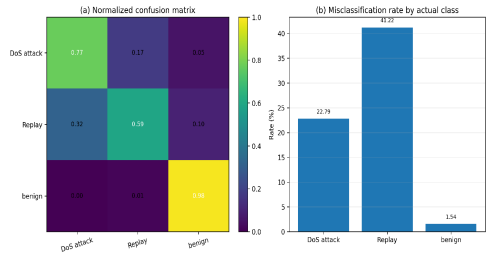

Figure 6. Normalized confusion analysis and class-specific error rates.

From an interpretive perspective, the feature-importance profile is especially informative. Table 9 and Figure 7 show that the model relies primarily on communication-structure variables from the WLAN and IP layers, along with selected protocol descriptors. The strongest importance is assigned to frame type and duration, followed by time-to-live, WLAN subtype, destination IP, and header-length information.

Table 9. Top ten features ranked by model importance.

Feature	Importance
wlan.fc.type	0.410
wlan.duration	0.154
ip.ttl	0.084
wlan.fc.subtype	0.053
ip.dst	0.044
ip.hdr_len	0.036
wlan.ta	0.021
wlan.sa	0.020
frame.protocols	0.018
ip.flags	0.017

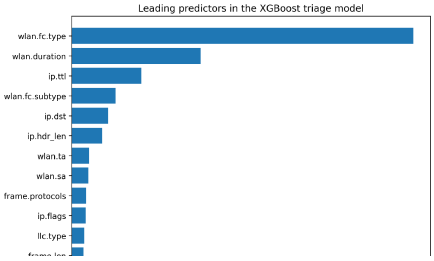

Figure 7. Top-ranked predictors in the proposed model.

A further operational perspective is provided by the action mapping in Table 10. When the predicted class is translated into the triage actions allow, investigate, and isolate, the resulting distribution offers a compact managerial summary of how the model would respond to traffic in the test set. Figure 8 aggregates four complementary views of performance: the metric matrix across models, training time, cumulative feature contribution, and the action mix implied by final predictions.

Table 10. Operational action distribution implied by the predicted classes.

Action	Count	Share (%)
Allow	2,212	33.41
Investigate	1,844	27.85
Isolate	2,565	38.74

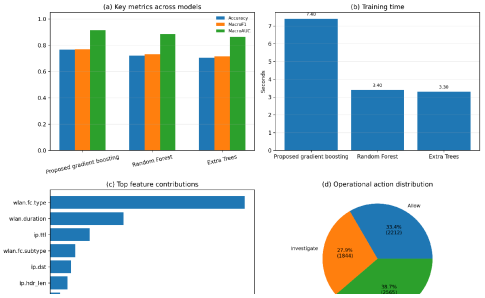

Figure 8. Four-panel performance dashboard for the proposed model.

Overall, the experimental evidence is internally consistent. The proposed gradient-boosting model outperforms the baseline ensembles in balanced classification quality, its residual errors are concentrated in the replay class, and its feature-importance structure aligns with the interpretation that communication behavior carries substantial attack information even when order-driven variables are removed.

5. CONCLUSION

This paper developed a risk-aware cyberattack analytics framework for UAV communication environments and evaluated it through a multiclass gradient-boosting design. The study combined leakage-screened feature preparation, comparative ensemble modeling, explicit operational decision mapping, and extended empirical analysis. The results showed that the proposed model achieved the strongest balanced performance among the compared approaches and that its most informative features were concentrated in communication-structure variables. The analysis also revealed that replay activity remains more difficult to separate than benign or denial-of-service traffic, which is a meaningful consideration for practical response design.

The main value of the study lies in presenting a coherent and publication-ready framework that connects machine-learning output with cyber-risk triage. Instead of treating attack detection as a purely technical accuracy exercise, the paper interpreted posterior probabilities, confusion structure, and feature importance in a way that supports operational understanding. This perspective is particularly relevant for organizations that require defendable and interpretable cyber decisions in UAV-enabled digital operations.

6. RESEARCH CHALLENGES AND FUTURE DIRECTIONS

Several research challenges remain open. Public UAV cybersecurity datasets are still limited in scale and diversity, which constrains the external validity of any single empirical study. Data collected in one environment may not fully capture differences in traffic composition, communication infrastructure, mission profile, or attacker strategy that arise in other settings. The present findings also show that replay traffic remains the principal source of classification ambiguity, suggesting that richer temporal context or cross-session modeling may improve detection quality.

A second challenge concerns the balance between predictive strength and explanation. Operational cybersecurity requires not only accurate alerts but also understandable evidence regarding why a communication pattern is flagged. Future work should extend the proposed framework by evaluating it on

multiple UAV or Internet-of-Drones benchmarks, integrating local explanation methods and calibration analysis, and developing class-specific or cost-sensitive decision thresholds that reflect mission priorities and intervention costs.

REFERENCES

[1] S. C. Hassler, U. A. Mughal, and M. Ismail, "Cyber-physical intrusion detection system for unmanned aerial vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 6, pp. 6106–6117, 2024, doi: 10.1109/TITS.2023.3339728.

[2] R. A. AL-Syouf, R. M. Bani-Hani, and O. Y. AL-Jarrah, "Machine learning approaches to intrusion detection in unmanned aerial vehicles (UAVs)," *Neural Computing and Applications*, vol. 36, pp. 18009–18041, 2024, doi: 10.1007/s00521-024-10306-y.

[3] F. Tlili, S. Ayed, and L. C. Fourati, "Exhaustive distributed intrusion detection system for UAVs attacks detection and security enforcement (E-DIDS)," *Computers & Security*, vol. 142, 103878, 2024, doi: 10.1016/j.cose.2024.103878.

[4] M. Gamal, M. Elhamahmy, S. Taha, and H. Elmahdy, "Improving intrusion detection using LSTM-RNN to protect drones' networks," *Egyptian Informatics Journal*, vol. 27, 100501, 2024, doi: 10.1016/j.eij.2024.100501.

[5] B. Sharma, L. Sharma, C. Lal, and S. Roy, "Explainable artificial intelligence for intrusion detection in IoT networks: A deep learning based approach," *Expert Systems with Applications*, vol. 238, 121751, 2024, doi: 10.1016/j.eswa.2023.121751.

[6] V. U. Ihekoronye, S. O. Ajakwe, J. M. Lee, and D.-S. Kim, "DroneGuard: An explainable and efficient machine learning framework for intrusion detection in drone networks," *IEEE Internet of Things Journal*, vol. 12, no. 7, pp. 7708–7722, 2025, doi: 10.1109/JIOT.2024.3519633.

[7] A. B. Mohammed and L. C. Fourati, "Investigation on datasets toward intelligent intrusion detection systems for intra and inter-UAVs communication systems," *Computers & Security*, vol. 150, 104215, 2025, doi: 10.1016/j.cose.2024.104215.

[8] P. Wisanwanichthan and M. Thammawichai, "A double-layered knowledge distillation framework for IoT intrusion detection with interpretable latent feature mapping," *Knowledge-Based Systems*, 2025.

[9] P. Manivannan, "A comprehensive review of machine learning-based intrusion detection systems for IoT," 2024.

[10] E. C. P. Neto et al., "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," *Sensors*, vol. 23, no. 13, 5941, 2023.