# Zero-day GPS Attack Detection and Classification in UAV Networks

Seyyedeh Maryam Mazloom*, Wei-Ping Zhu*, Wessam Ajib†
* Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada,
† Département d'informatique, Université du Québec à Montréal, Montréal, Canada.
Email: seyyedehmaryam.mazloomghalehbala@mail.concordia.ca, weiping@ece.concordia.ca, ajib.wessam@uqam.ca.

***Abstract*—Unmanned aerial vehicles (UAVs) face escalating cybersecurity threats, particularly from GPS spoofing and jamming attacks, which endanger flight safety and mission integrity. This paper introduces a novel UAV cybersecurity framework that integrates statistical extreme value meta-learning (EVML) with a dual-path classifier to detect and classify GPS attacks. The method overcomes the key limitations of some existing approaches, such as excessive training data needs, zero-day threat vulnerability, and disjointed detection/classification. In particular, the proposed framework includes two stages. The first one enables zero-day attack detection through few-shot meta-learning with prototype-based anomaly detection where support sets contain only benign flight data while query sets include both benign samples and synthetically generated attack patterns. Moreover, a prototypical OpenMax layer and extreme value theory are exploited to identify suspicious patterns in GPS telemetry data. The second stage utilizes a novel dual-path architecture that independently processes position-related and signal-related features to classify detected attacks as spoofing or jamming. Our results on real-world attack data demonstrates that the proposed method has exceptional performance with 97.33% detection accuracy and 0% false alarm rate, which significantly outperform the state-of-the-art. Furthermore, attack classification achieves 82.33% accuracy for spoofing and 94.31% for jamming attacks, with an overall F1 score of 0.88.**

***Index Terms*—Unmanned aerial vehicles, GPS spoofing, GPS jamming, extreme value meta-learning, dual-path classification, zero-day attacks, few-shot learning**

## I. Introduction

### A. Context

Unmanned aerial vehicles (UAVs) have become essential assets in mission-critical applications [1]. UAV navigation systems have demonstrated fundamental vulnerabilities, especially in GPS-based positioning, which is critical for autonomous flight operation. Recent surveys also reveal that UAVs face a variety of security threats within the cyber-physical domains [2], [3].

Among the various cyber threats to UAVs, GPS spoofing and jamming are particularly critical as they are easy to execute and can severely compromise flight safety. GPS spoofing attacks manipulate UAV navigation by transmitting falsified positioning signals that cause GPS receivers to calculate incorrect position and time information [4], [5]. GPS jamming attacks exploit the inherent weakness of GPS signals by broadcasting interference within the GPS frequency bands, leading to possible mission failure or vehicle loss [6], [7]. To combat aforementioned attacks, various intrusion detection systems (IDS) have been developed for UAVs, including signature-based, specification-based, anomaly-based, and hybrid detection approaches [8]. Furthermore, recent advances in machine learning have shown great promise. For instance, supervised methods such as tree-based models [3], XGBoost [9], achieve high precision in GPS spoof detection, but they struggle with labeled data scarcity and zero-day detection. Meanwhile, specialized deep learning approaches such as ConvLSTM [10] and DeepSpoofNet [11] demonstrate high accuracy in general intrusion detection, but require extensive labeled attack datasets and large computational resources which are unsuitable for resource-constrained UAV platforms. In addition, the micro air vehicle intrusion detection system (MAVIDS) [1] represents a significant advancement by integrating IDS and machine learning techniques, while allowing on-board deployment.

MAVIDS is considered a leading on-board detection system using one-class classification, although it requires extensive training data and lacks unified threat assessment capabilities [1]. Based on conventional threshold-based attack detection, [12] provides robust attack detection across UAV networks. Recent advances in meta-learning offer potential solutions by enabling rapid adaptation with minimal training samples, as demonstrated in IoT networks [13], and prototype networks reaching 92.86% accuracy for 5G DoS attacks [14]. In [15], meta-WF was adopted for Wi-Fi networks, employing multiple auxiliary networks for rapid knowledge transfer. However, these solutions face critical limitations in UAV contexts. They process network traffic patterns rather than analyzing multidimensional sensor telemetry data from flight systems, require labeled training samples for all attack types, and lack open-set recognition capabilities for zero-day attacks. Authors in [16] introduced extreme value meta-learning (EVML) for hyperspectral image classification, integrating statistical EVT with prototypical networks to improve accuracy by 7-14%. By modeling the tails of feature distributions with the Weibull function and separation anomalies using the P-OpenMax layer, EVML enables accurate zero-day (open-set) attack detection even when limited training samples are available, thus overcoming the limitations of closed-set meta-learning methods. This makes EVML especially promising for UAVs cybersecurity.

### B. Contribution

To address the aforementioned limitations of the state-of-the-art methods, we propose a novel framework that adapts

979-8-3503-9281-4/25/$31.00 ©2025 IEEE

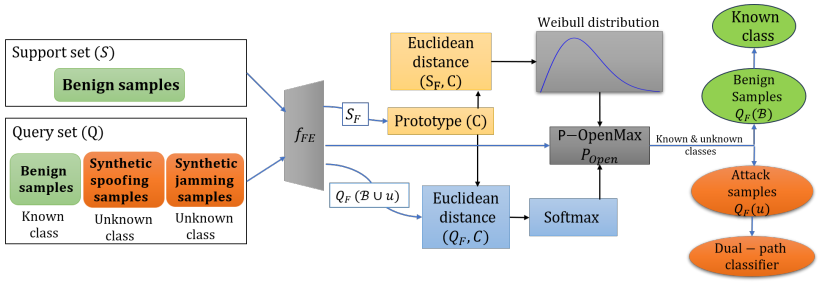


Fig. 1. Architecture of our adapted EVML framework for UAV attack detection with dual-path classifier. It shows the 1D CNN feature extraction network with CBAM attention mechanisms ($f_{\mathrm{FE}}$), prototype computation, and calibrated distance scoring using extreme value theory (EVT) for GPS telemetry data.

EVML for GPS attack detection in UAVs. The contributions of this paper can be summarized as follows:

1) We develop a novel two-stage framework comprising an EVML-based anomaly detector, followed by a specialized dual-path classifier that distinguishes between spoofing and jamming attacks through separate analysis.
2) We apply the EVML to UAV security by redesigning the feature extraction network for one-dimensional (1D) GPS telemetry data and employing a convolutional block attention module (CBAM) to focus on attack-relevant patterns. Additionally, a customized quadruplet loss function is used to maintain topological separation between normal UAV operation and attack patterns in the feature embedding space.
3) We introduce a domain knowledge-driven synthetic attack generation approach to create realistic training data without requiring exposure to real attack patterns, enabling zero-day detection capabilities.
4) Simulations demonstrate superior detection and classification performance compared to traditional methods.

The remainder of this paper is as follows. Section II details the proposed solution. Section III presents an experimental setup and an overview of the end-to-end workflow. Section IV presents the results and discussion. Finally, Section V concludes the paper.

## II. Detection and Classification Solution

In this section, we describe our proposed hierarchical detection-classification framework for UAV GPS attacks. It consists of EVML-based anomaly detection, which leverages extreme value meta-learning to identify suspicious GPS behavior patterns based on minimal training data, and dual-path attack classification employing specialized neural paths to distinguish between spoofing and jamming attacks. Fig. 1 illustrates the adapted EVML-based anomaly detection and dual-path classifier.

### A. EVML-based Anomaly Detection

This step adopts an episodic meta-learning strategy similar to that proposed in [16]. In our framework, the support set consists of the benign class (known class), while the query set includes both benign and attack classes (unknown classes). To enable zero-day attack detection while addressing unknown classes in query sets during training, we employ synthetic attack generation techniques that systematically modify benign flight data to simulate realistic attack patterns (detailed in Section III-A). In particular, the proposed EVML consists of four components to achieve robust anomaly detection.

*1) Feature Extraction:* As can be seen in Fig.1, the feature extractor function ($f_{\mathrm{FE}}$) employs a convolutional neural networks (CNN) architecture enhanced with 1D CBAM layers that focus on attack-relevant temporal patterns in GPS data. The network processes the pre-processed GPS features through convolutional blocks with ReLU activation, batch normalization, and max pooling, followed by attention mechanisms that emphasize critical pattern characteristics. The extracted features are then projected through fully connected layers to a low-dimensional embedding space optimized for prototype-based learning.

*2) Prototype and Distance Modeling:* Similarly to [16], we first compute the prototype of the benign class, then calculate the distances between the support features and this prototype, and finally fit a Weibull distribution to these distances. For the query feature set, we measure the distances from the benign class prototype, and, differently from [16], apply a *softmax* function to these distances to find proximity probabilities, where query samples closer to the prototype are assigned higher probabilities.

*3) P-OpenMax Layer:* The P-OpenMax layer introduces an additional "unknown" class (class 0) to model the open space where outliers (GPS spoofing/jamming attacks) reside. As shown in Algorithm 1, this layer has been adapted to align with our proposed framework. As a result, a query sample is detected as an attack when the unknown class ($y = 0$) receives the highest probability. Once a sample is identified as

**Algorithm 1** P-OpenMax Calibration for UAV Spoofing And Jamming Attack Rejection

**Require:** Query feature $q_j$, prototype $C$, Weibull parameters $\theta, \tau, \lambda$
**Ensure:** Open-set probability $P_{\text{open}}(y = 0)$
1: Compute distance: $d_j \leftarrow \|C - q_j\|_2^2$
2: Calculate Weibull weight: $w_j \leftarrow 1 - \exp\left(-\left(\frac{d_j - \theta}{\tau}\right)^{\lambda}\right)$
3: Calibrate distance: $\tilde{d}_j \leftarrow d_j \cdot w_j$
4: Compute open-space distance: $d_0 \leftarrow \sum_{j=1}^{M(D+1)} d_j \cdot (1 - w_j)$
5: Calculate $P_{\text{open}}$: $P_{\text{open}}(y = 0) \leftarrow \frac{e^{-\tilde{d}_j}}{\sum_{j=1}^{M(D+1)} e^{-\tilde{d}_j} + e^{-d_0}}$
6: **return** $P_{\text{open}}(y = 0)$

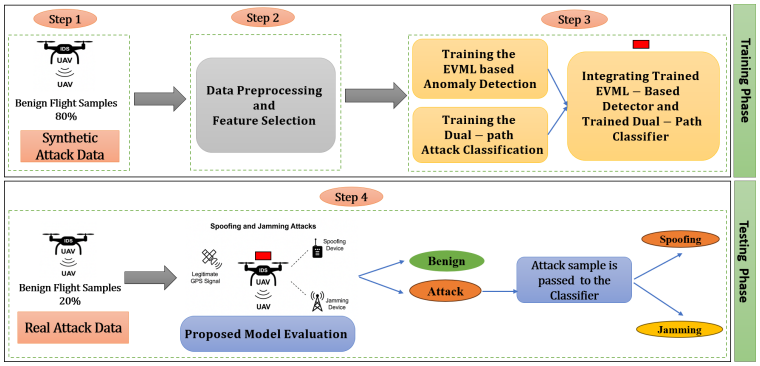


Fig. 2. Overview of the experimental workflow, from dataset to evaluation of the proposed model

an attack, it passes immediately to the dual-path classifier for the attack type identification.

*4) Loss Function:* For our one known class model, the loss function is adapted to enforce two primary constraints where benign samples remain close to their prototype, and attack samples maintain a separation from the benign prototype by margin $\alpha$. This margin controls the minimum distance that attack samples must stay away from the benign prototype, while [16] considers three margins to enforce separations between multiple known classes and the unknown class.

### B. Dual-Path Attack Classification

The proposed dual-path classifier processes exclusively the samples identified as anomalous by the EVML detector. In contrast to the coarse classifier in [16], which employs a single-path neural architecture with a softmax activation function, the dual-path classifier is designed for the fine-grained classification of GPS-based attacks by distinguishing between spoofing and jamming samples, which is necessary to select appropriate countermeasures. Architecturally, the dual-path classifier leverages two parallel neural paths, each customized to the unique characteristics of GPS spoofing and jamming attacks. The position path processes position-related features from Table I (detailed in III-B), such as *lat_y*, *lat_x*, *heading*, *vel_m_s*, and *vel_e_m_s* through dense layers, each followed by *ReLU* activation functions. This path captures patterns in positional feature that are indicative of GPS spoofing attacks, such as subtle coordinate inconsistencies that gradually drift from the true position. In parallel, the signal path focuses on signal characteristics more relevant to jamming detection and processes the signal quality features as listed in Table I including *evh*, *hdop*, *jamming_indicator*, and *noise_per_ms*, through dense layers, all of which distinguish jamming from other types of GPS anomalies. It also maintains layer configurations identical to the position path. The output of each specialized paths is combined with all pre-processed features, and then processed through additional dense layers with dropout regularization to prevent overfitting. The last step in classification is to separate the output layers for spoofing and jamming probabilities with sigmoid activation functions enhanced with regularization $L2$ to improve generalization.

## III. Experimental Setup

Our experimental evaluation of the proposed framework follows a systematic four-step workflow, as illustrated in Fig. 2. They are: step (1) dataset preparation and partitioning with strategic separation of training and testing data, as well as synthetic attack generation; step (2) parallel preprocessing pipelines employed separately for detection and classification components; step (3) integrated training and validation; step (4) combining separately trained models into a unified detection-classification system for real attack data testing.

### A. Dataset

As shown in Step 1 of Fig. 2, the dataset supports both the training and testing requirements of the EVML detector and the dual path classifier. In our experiments, we use the live GPS spoofing and jamming attacks of *"UAV attack dataset"* from the IEEE DataPort repository [17]. The dataset comprises 6,078 benign flight samples, 498 GPS spoofing attack samples (conducted using HackRF software-defined radio), and 1,460 GPS jamming attack samples (white Gaussian noise interference). To ensure a rigorous zero-day attack evaluation, all real attack data was reserved exclusively for testing, while benign data was split into 80% for training and 20% for validation testing. As detailed in Table I, 14 key GPS telemetry features are selected.

To strengthen the training process while preserving real attack data for testing, we developed a synthetic attack generation methodology that systematically modifies benign flight telemetry to simulate realistic attack patterns. For spoofing attacks, we generated subtle position-related deviations while keeping regular signal characteristics, mimicking the stealthy nature of actual GPS spoofing. For jamming attacks, we applied aggressive modifications to *hdop*, *noise_per_ms*, and *jamming_indicator* to reflect signal degradation patterns.

### B. Parallel Preprocessing Pipelines

Our preprocessing approach (step 2 of Fig. 2) employs two parallel pipelines separately to meet the distinct requirements of the EVML detector and dual-path classifier. The EVML detector preprocessing focuses on dimensionality reduction and standardization for prototype-based learning. GPS-related features undergo preparation through a three-stage process.

TABLE I
FEATURES BY IMPORTANCE [1]

| PC | PX4 Feature Name | Description |
|---|---|---|
| 1 | *evh* | Horizontal velocity error |
| 2 | *time_utc_usec* | UTC time in microseconds |
| 3 | *lat_y* | Latitude 2 (from preprocessing) |
| 4 | *lat_x* | Latitude 1 (from preprocessing) |
| 5 | *heading* | Vehicle Heading (from preprocessing) |
| 6 | *z_deriv* | Down position time derivative (m/s) |
| 7 | *v_z* | Z axis velocity |
| 8 | *ax* | North velocity derivative |
| 9 | *hdop* | Horizontal dilution of precision |
| 10 | *vel_m_s* | GPS ground speed (m/s) |
| 11 | *q* [2] | Quaternion rotation from the FRD body frame to the NED earth frame |
| 12 | *jamming_indicator* | PX4 built-in jamming detection |
| 13 | *vel_e_m_s* | GPS east velocity |
| 14 | *noise_per_ms* | GPS RF noise per millisecond |

First, a feature selection is performed to extract from raw GPS signals the 14 most informative features as given in Table I. Next, the standard scaler normalization transforms all features to have zero mean and unit variance. Finally, principal component analysis (PCA) [18] is applied to reduce dimensionality while preserving 85% of the cumulative explained variance.

The dual-path classifier requires specialized preprocessing that incorporates domain-specific engineering features to exploit the differences between spoofing and jamming attack mechanisms. The preprocessing begins with the features from Table I, followed by the engineering features. Domain-specific engineering features include (i) position magnitude computed as the Euclidean norm of coordinate vectors to capture spatial displacement patterns, (ii) velocity magnitude derived from velocity component vectors to identify movement inconsistencies, (iii) position-velocity ratio serving as a key indicator for detecting inconsistencies between reported position and velocity that characterize spoofing attacks, and (iv) hdop-noise product which multiplies horizontal dilution of precision and noise metrics to amplify signal degradation patterns characteristic of jamming attacks. Then, the entire set of features including both base and engineered features is normalized using Standard Scaler to achieve zero mean and unit variance.

### C. Training and validation

The EVML detector and dual-path classifier are trained independently using their respective data streams from Step 2, and then integrated into a sequential detection-classification system optimized for real-time operational deployment.

*1) EVML-based Detector Training:* The EVML detector employs episodic meta-learning with 8 episodes per epoch over 15 epochs, in which each episode includes a support set of 5 benign samples and a query set of 15 samples per class (benign, synthetic spoofing, synthetic jamming), for a total of 45 query samples per episode. The feature extraction network employs a 1D CNN architecture with two convolutional blocks (32 and 64 filters respectively) enhanced with CBAM attention mechanisms for temporal pattern recognition. The features are then projected to a 3D embedding space optimized for prototype-based learning. To fit the Weibull distribution, we consider a tail size parameter $\eta = 5$ that selects the $\eta$ largest support set distances to model the extreme value distribution. In addition, the Weibull parameters for the benign class averaged $\theta = 0$, $\tau = 1.1036 \pm 0.43$, and $\lambda = 0.0006 \pm 0.002$ over episodes. The open quadruplet loss incorporates the margin parameter $\alpha = 1.0$. Regarding the size of this model, the EVML detector contains 10,967 trainable parameters, which is approximately 42.8 kilobyte (KB).

*2) Dual-Path Classifier Training:* The position path processes five position-related features through a 16-neuron dense layer, focusing on spatial inconsistencies characteristic of spoofing attacks. Through an identical 16-neuron dense layer, the signal path analyzes four signal quality factors focused on jamming attacks. The outputs from both paths are combined with all other features and sent through a 32-neuron dense layer with dropout regularization set to 0.3. The last step includes sigmoid activation functions with regularization of $L2$ (0.001), which produce independent probability scores for spoofing and jamming attacks. An Adam optimizer with a learning rate of 0.005 is used, which implements binary cross-entropy loss functions with early stopping mechanisms (patience=10) to ensure optimal generalization. The Dual-path classifier model has 4,114 trainable parameters (16.07 KB).

*3) Hyperparameter Optimization and Model Integration:* Following independent training, the hyperparameters are optimized using a 5-fold cross-validation grid search. The optimization explores parameters including tail size for Weibull fitting (3, 5, 7), learning rate (0.001, 0.005, 0.01), embedding dimension (3, 5), and margin parameter for open quadruplet loss $\alpha$ (0.5, 0.8, 1.0, 1.2), dropout rate (0.2, 0.3, 0.4) and $L2$ regularization strength (0.001, 0.01, 0.1). Parameter selection is based on maximizing the balanced F1 score in all validation folds.

### D. Testing and Evaluation with MAVIDS Integration

The evaluation phase (Step 4 in Fig. 2), extensively accesses the real-world performance and the deployment feasibility of the proposed framework by adopting the evaluation methodology of MAVIDS [1]. The test is carried out on completely unseen real attack data, including 35 randomized test episodes. The evaluation follows the same experimental protocol as MAVIDS, using the live UAV attack dataset in an offline setting to ensure direct comparability with established UAV security solutions. During evaluation, the proposed framework processes the GPS telemetry logs. Although practical onboard deployment and real-time testing are beyond the scope of this study, the proposed framework is specifically designed for integration with onboard IDS agents, supporting future real-time detection and classification within the resource constraints of UAV platforms. All experiments were carried out on a laptop (Intel Core i5-8265U, 1.8 GHz; 16 GB RAM; 64 bit OS) with TensorFlow 2.19.0.

To evaluate both detection and classification performance, we employ a suite of complementary metrics spanning both operational stages. The attack detection metrics are detection accuracy, attack detection rate (ADR), false alarm rate (FAR), and F1 score. Detection accuracy represents the percentage

of correctly classified samples across both benign and attack types. ADR quantifies the percentage of actual attacks correctly identified. FAR measures the percentage of benign samples incorrectly classified as attacks, and the F1 score provides a balanced measure of detection performance through the harmonic mean of *precision* and *recall* metric.

For the attack classification, we evaluate attack-specific classification accuracy (ASCA), F1 score, balanced accuracy, precision and recall metrics. ASCA measures the percentage of detected spoofing and jamming attacks that are classified into their corresponding categories, appropriately. The class-specific F1 score provides individual balanced performance measures for the classification of spoofing and jamming. Balanced accuracy computes the average of spoofing and jamming classification accuracies. Further, *precision* and *recall* are calculated separately for each attack type to assess the classifier's ability to correctly identify categories while minimizing inter-attack misclassifications.

For both the detection phase (benign vs. attack) and the classification phase (spoofing vs. jamming), the *confusion matrices* are computed, providing a detailed breakdown of the detection and classification outcomes. These matrices enable fine-grained analysis of error patterns, which is essential for system optimization and reliability assessment. In addition, we report atomic prediction latency as the overall prediction time divided by the number of processed samples (ms/sample). In this study, an input sample denotes a single preprocessed vector of GPS telemetry features at one time step; see Section III-B.

## IV. Results and Discussion

The proposed unified EVML-based detection framework with dual-path classification is evaluated through the aforementioned metrics on the real-world UAV attack dataset, as detailed below.

### A. Attack Detection Performance

As shown in Table II, the proposed framework achieves an overall detection accuracy of 97.33% (±1.97%) with atomic latency of 1.8539 ms/sample, indicating its effectiveness in distinguishing benign samples from GPS-based attacks on UAVs. The most notable outcome is the perfect false alarm rate of 0.00% (±0.00%) observed across all 35 test episodes. This result is particularly significant in UAV security scenarios, where false alarms can result in unnecessary emergency protocols, mission aborts, or autonomous landing procedures. In addition, the framework maintains a high attack detection rate of 96.00% (±2.96%), successfully identifying 1008 out of 1050 attack samples. The presented method reaches the F1 score of 96.23% (±0.0270) for benign samples and 97.94% (±0.0156) for attack samples, resulting in a macro-averaged F1-score of 97.09% (±0.0156), which highlights the balanced performance of the model in both classes. On the other hand, the detection confusion matrix in Table III illustrates the detection performance, in which all 525 benign samples are correctly identified (100%) while 1008 attack samples are successfully detected and forwarded to the classification stage. The 42 misclassified instances, incorrectly detected as benign, include 27 spoofing samples (5.14% miss rate) and 15 jamming samples (2.86% miss rate), which shows stronger detection performance against jamming attacks compared to spoofing attacks.

TABLE II
Detection and Classification Performance Results

| Metric | Value (%) | Std. Dev. |
|---|---|---|
| **Attack Detection Performance** | | |
| Detection Accuracy | 97.33 | ±1.97 |
| Attack Detection Rate | 96.00 | ±2.96 |
| False Alarm Rate | 0.00 | ±0.00 |
| F1 Score (Macro-Averaged) | 97.09 | ±0.016 |
| F1 Score (Benign) | 96.23 | ±2.70 |
| F1 Score (Attack) | 97.94 | ±1.56 |
| **Attack Classification Performance** | | |
| Balanced Accuracy | 88.00 | — |
| **Spoofing Attack** | | |
| Classification Accuracy | 82.33 | — |
| Precision | 93.00 | — |
| Recall | 82.00 | — |
| F1 Score | 88.00 | — |
| **Jamming Attack** | | |
| Classification Accuracy | 94.31 | — |
| Precision | 85.00 | — |
| Recall | 94.00 | — |
| F1 Score | 89.00 | — |

TABLE III
Detection Phase Confusion Matrix

| | Predicted Benign | Predicted Attack |
|---|---|---|
| **Actual Benign** | 525 (100%) | 0 (0%) |
| **Actual Attack** | 42 (4.00%) | 1008 (96.00%) |

TABLE IV
Attack Type Classification Confusion Matrix

| | Predicted Spoofing | Predicted Jamming |
|---|---|---|
| **True Spoofing** | 410 (82.33%) | 88 (17.67%) |
| **True Jamming** | 29 (5.69%) | 481 (94.31%) |

### B. Classification Performance

The dual-path classifier performance is also presented in Table II. It achieves 82.33% accuracy for spoofing attacks and 94.31% accuracy for jamming attacks, resulting in a balanced accuracy of 88.00%, with atomic latency of 3.4752 ms/sample.

The model exhibited superior performance in jamming attack identification, achieving 94.00% recall and 85.00% precision, yielding an F1-score of 89.00%. This performance is consistent with our design hypothesis that jamming attacks cause more different signal deterioration patterns, which are effectively captured through the signal path's emphasis on *hdop*, noise-level metrics, and built-in jamming indicators.

The spoofing attacks are more challenging to classify correctly, achieving 82.00% recall and 93.00% precision, resulting in an F1-score of 88.00%. The higher precision (93.00%) indicates that when the model predicts a spoofing attack, it is highly reliable. However, the lower recall (82.00%) corresponds to the misclassification pattern shown in Table

TABLE V
MACRO-AVERAGED F1 SCORE COMPARISON FOR GPS ATTACK DETECTION METHODS

| Method | Spoofing | Jamming | Combined Attacks | Atomic Latency (ms/sample) |
|---|---|---|---|---|
| OC-SVM [1] | 88.78% | 56.64% | – | 2.8677 |
| LOF [1] | 90.73% | 78.09% | – | 109.0102 |
| Autoencoder [1] | 90.57% | 94.29% | – | 0.4174 |
| **EVML (Ours)** | – | – | **97.09%** | **1.8539** |

IV, where 88 spoofing samples (17.67%) are misclassified as jamming attacks, which suggests that certain sophisticated spoofing techniques may exhibit signal quality degradation patterns that resemble jamming effects.

### C. Comparative Analysis with MAVIDS Architecture

To assess the EVML framework, we do comparisons with the MAVIDS approach [1], which has used one-class classifiers to detect GPS attacks. Table V presents the performance comparison focusing on macro-averaged F1 score which is the mean of the F1 score for the attack class and the F1 score for the benign class. As shown in Table V, one-class classifiers have strong single-attack detection but lack unified threat assessment capabilities. As a result of the proposed EVML, a unified macro-averaged F1 score of 97.09% is achieved for simultaneous spoofing-jamming detection, which largely outperforms the one-class classifiers. This outperformance comes from our meta-learning architecture's ability to model cross-attack patterns through extreme value distributions, unlike conventional methods that optimize detection thresholds separately for each attack type. As such, EVML is able to achieve this by using less benign samples (maximum 600), than MAVIDS' requirement of 6,078 benign samples. For latency, baseline one-class atomic latencies were measured on an agent onboard (Raspberry Pi 4 Model B companion) [1], while EVML latency was evaluated in our laptop testbed; thus, the absolute values reflect different device capabilities, but both sets of numbers indicate real-time feasibility.

### D. Discussion

Experimental results confirm that the EVML proposed framework successfully addresses key UAV attack detection challenges. It transforms training requirements from extensive labeled attack datasets to minimal benign flight data. Through episodic training with balanced mini-datasets, it also mitigates class imbalance, while the integration of EVT with meta-learning enables robust zero-day attack detection. The proposed dual-path classifier addresses the limitation of existing approaches that only conduct detection without identifying the specific GPS attack type. By leveraging the differences between spoofing and jamming attacks, this classifier enables deployment of specific countermeasures for each attack type, which improves response effectiveness compared to the conventional mitigation strategies. The proposed two-stage architecture (evaluated using the MAVIDS methodology) offers high detection accuracy while meeting the computational constraints of real-time UAV operations. This would result in realistic deployment on resource-constrained devices and provides an effective defense against evolving UAV threats.

## V. CONCLUSION

In this paper, we presented a statistical extreme value meta-learning (EVML)-based framework for enhancing UAV cybersecurity, with a particular focus on zero-day GPS attack detection. By integrating EVML with a dual-path classifier, the proposed approach not only enables accurate identification of previously unseen attacks but also provides fine-grained classification of different GPS attack types. The unified detection-classification architecture achieves superior performance across multiple evaluation metrics, effectively addressing critical operational challenges in UAV security. Moreover, the low data requirements and the computational efficiency of the proposed approach highlight its suitability for deployment in real-world, resource-constrained UAV environments. These results demonstrate the promise of advanced meta-learning strategies for practical UAV threat mitigation.

## REFERENCES


[1] J. Whelan, A. Almehmadi, and K. El-Khatib, "Artificial intelligence for intrusion detection systems in unmanned aerial vehicles," *Comput. Electr. Eng.*, vol. 99, p. 107784, 2022.

[2] Z. Yu *et al.*, "Cybersecurity of Unmanned Aerial Vehicles: A Survey," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 39, no. 9, pp. 182–215, 2024.

[3] T. T. Khoei *et al.*, "A comparative analysis of supervised and unsupervised models for detecting GPS spoofing attack on UAVs," in *2022 IEEE Int. Conf. Electro Inf. Technol. (EIT)*, pp. 279–284, 2022.

[4] G. Panice *et al.*, "A SVM-based detection approach for GPS spoofing attacks to UAV," in *2017 23rd Int. Conf. Autom. Comput.(ICAC)*, pp. 1–11, 2017.

[5] S. Wang, J. Wang, C. Su, and X. Ma, "Intelligent detection algorithm against UAVs' GPS spoofing attack," in *IEEE Int. Conf. Parallel Distrib. Syst. (iCPADS)*, pp. 382–389, 2020.

[6] F. Abdullayeva and O. Valikhanli, "Multimodal deep neural network for UAV GPS jamming attack detection," *Cyber Secur. Appl.*, vol. 3, p. 100094, 2025.

[7] H. O. Slimane *et al.*, "A light boosting-based ml model for detecting deceptive jamming attacks on UAVs," in *IEEE Annu. Comput. Commun. Workshop Conf. (CCWC)*, pp. 0328–0333, 2022.

[8] L. M. Da Silva *et al.*, "Anomaly-based intrusion detection system for in-flight and network security in UAV swarm," in *Int. Conf. Unmanned Aircraft Syst. (ICUAS)*, pp. 812–819, 2023.

[9] G. Aissou *et al.*, "Tree-based supervised ML models for detecting GPS spoofing attacks on UAS," in *IEEE Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, pp. 0649–0653, 2021.

[10] R. Alharthi, "Enhancing unmanned aerial vehicle and smart grid communication security using a ConvLSTM model for intrusion detection," *Frontiers in Energy Research*, vol. 12, p. 1491332, 2024.

[11] A. U. R. Badar *et al.*, "DeepSpoofNet: a framework for securing UAVs against GPS spoofing attacks," *PeerJ Computer Science*, vol. 11, p. e2714, 2025.

[12] 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, p. 103878, 2024.

[13] C. Lu *et al.*, "A few-shot-based model-agnostic meta-learning for intrusion detection in security of internet of things," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 21309–21321, 2023.

[14] Y. Yan *et al.*, "Meta learning-based few-shot intrusion detection for 5G-enabled industrial internet," *Complex Intell. Syst.*, vol. 10, no. 3, pp. 4589–4608, 2024.

[15] T. Li *et al.*, "Meta-WF: Meta-learning-based few-shot wireless impersonation detection for Wi-Fi networks," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3585–3589, 2021.

[16] D. Pal, S. Bose, B. Banerjee, and Y. Jeppu, "Extreme value meta-learning for few-shot open-set recognition of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–16, 2023.

[17] J. Whelan, T. Sangarapillai, O. Minawi, A. Almehmadi, and K. El-Khatib, "UAV Attack Dataset," 2020.

[18] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, no. 1-3, pp. 37–52, 1987.