IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 December 2023, accepted 24 January 2024, date of publication 30 January 2024, date of current version 26 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3359700

RESEARCH ARTICLE

Enhanced Machine Learning Ensemble Approach for Securing Small Unmanned Aerial Vehicles From GPS Spoofing Attacks

ALA' ABDULMAJID ESHMAWI[1], MUHAMMAD UMER[2], IMRAN ASHRAF[3], AND YONGWAN PARK[3]

[1]Department of Cybersecurity, College of Computer Science and Engineering, University of Jeddah, Jeddah 23218, Saudi Arabia
[2]Department of Computer Science and Information Technology, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan
[3]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, Gyeongsangbuk 38541, Republic of Korea

Corresponding authors: Yongwan Park (ywpark@yu.ac.kr) and Imran Ashraf (ashrafimran@live.com)

This work was supported in part by the 2021 Yeungnam University Research Grant under Grant 221A380146, and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2021R1A6A1A03039493.

ABSTRACT Unmanned aerial vehicles (UAVs) substantially rely on the utilization of global positioning systems (GPS) to navigate. A simulator for commercial GPS applications with false GPS signals can lead to the deviation of a GPS-guided drone from its planned path. As a result, an anti-spoofing technology is required to assure UAV operating safety. Several approaches have been developed to detect GPS spoofing, however, predominantly such methods rely on additional hardware. Using additional hardware might not be an ideal solution for small and low-capacity UAVs. Detecting signal spoofing attacks in small UAVS has significant importance. This study presents a stacked ensemble approach to detect GPS signal spoofing within the context of small UAVs. The initial phase involves outlining the sequential procedures for obtaining and preparing the GPS signal dataset, including details about the UAV hardware, blocker, data collection timing, environmental factors, and the utilization of z-score normalization for preprocessing. Then controlled simulation tests with varying experimental conditions are conducted and the model is built using a support vector machine and convolutional neural network. Additionally, a comprehensive comparative assessment is conducted to analyze the efficacy of the proposed model against traditional machine learning models. Experimental results demonstrate notably good performance by the proposed model with a 99.74% accuracy, showing its superior performance in the context of GPS signal spoofing in small UAVs.

INDEX TERMS Unmanned aerial vehicles, autonomous vehicles, GPS spoofing, cybersecurity, ensemble machine learning.

I. INTRODUCTION

Small unmanned aerial vehicles (UAVs) commonly integrate a variety of sensor types, where global positioning system (GPS) receivers hold paramount significance. These receivers play a pivotal role in establishing the precise position, spatial coordinates, and altitude of the UAV by receiving signals transmitted from satellites. The acquired GPS signals contribute to enhancing the UAV's navigation accuracy, consequently elevating its efficacy in executing missions [1]. In summation, GPS technology stands as a pivotal component within the realm of small UAVs. Nonetheless, the vulnerability of GPS signals to spoofing introduces a substantial risk, particularly in sectors like aviation, military, navigation, and civil safety, where disruptions to synchronization and navigation can result in severe consequences. Several countries have reported GPS signal spoofing security vulnerabilities [2].

UAVs were initially developed for military goals like practicing anti-aircraft techniques, gathering intelligence, killing opponents, destroying hostile objects, and so on.

The associate editor coordinating the review of this manuscript and approving it for publication was Juan Liu.

With fast technological improvement over the last two to three decades, the usage of UAVs has expanded beyond military uses to various civilian and commercial applications. In Germany, DHL's logistics business employed UAVs to carry medicine twice a day over a 12-kilometre journey to the car-free island of Juist [3]. The United States Federal Aviation Administration has granted Alphabet, the parent company of Google, permission to use UAVs to transport meals in 2019 [4]. UAVs are utilized for a variety of other tasks, including animal monitoring, search and rescue operations, community surveillance, ambulance service, firefighting, journalism, aerial filming, and panoramic photography [5]. Utilizing UAVs in conjunction with the Internet of Things (IoT) sensors on the ground offers a range of applications [6]. These include assisting agricultural companies in surveying land and crops, aiding energy companies in monitoring power infrastructure and operational equipment, and supporting insurance companies in property and asset inspections.

The use of GPS services has dramatically expanded recently. The market for GPS tracking devices is projected to witness growth, escalating from its existing value of 1.57 billion to an estimated 3.38 billion by the year 2025 [7]. Ensuring safety often revolves around monitoring the present location of a mobile entity. In the case of autonomous vehicles, their navigation system relies on GPS signals to determine the current latitude, longitude, acceleration, and orientation, aiding the vehicle in reaching its intended destination. The widespread availability of GPS-equipped devices and affordable spoofing equipment has given rise to an increased risk of malicious GPS attacks. These attacks are facilitated by the prevalence of unencrypted GPS signals and the ease with which attackers can manipulate standard GPS signal structures using programmable radio devices such as HackRF or USRP [8]. This enables attackers to launch GPS spoofing attacks from a distance, disrupting genuine GPS signals and leading target vehicles' navigation systems astray. Researchers have demonstrated the potential to control autonomous vehicles' paths, even causing them to deviate off-road using tools like HackRF [9]. Besides navigation, numerous applications and services also heavily rely on GPS data to enhance their functionalities and user interfaces [10].

Small UAVs have been investigated for their vulnerability to GPS spoofing attacks in recent years [11], [12]. To tackle the issue of susceptibility to GPS signal spoofing in small UAVs, numerous researchers have employed techniques such as fingerprint and multipath detection [13]. These approaches aim to identify and counteract the manipulation of GPS signals in small UAVs. The employed methods face significant challenges. Multipath detection is ineffective against subtle GPS signal spoofing and demands extra hardware and precise clock synchronization. Noisy and Weak signals falter energy detection, risking the misclassification of genuine signals as spoofed. Fingerprinting needs extensive data for training, and substantial receiver hardware, and is susceptible to environmental shifts.

Numerous studies have introduced traditional machine learning (ML) methods aimed at categorizing and identifying GPS spoofing. These models offer proficient frameworks for identifying GPS spoofing incidents [14]. Over the last decade, ensemble learning methods have emerged as a prominent advancement in the machine learning field because of their superior performance when compared to conventional machine learning approaches [15].

The challenge of GPS spoofing in small UAVs is complex. New, tricky spoofing techniques can trick traditional detection methods, needing extra tools and precise timing to counter. Weak GPS signals can confuse detection systems, risking mistakes in identifying real and fake signals. Protecting small UAVs from spoofing is crucial as they are used in vital roles like security and delivery. Solving these challenges means making reliable detection methods to keep UAV operations safe and secure in important tasks. The motivation for this study is to recognize and classify GPS spoofing incidents in UAVs. Numerous research investigations have concentrated on categorizing and identifying UAV-targeted spoofing attacks with the help of machine learning methods. This study relies on machine learning models and ensemble classifiers. The models undergo training and testing utilizing a dataset comprising 13 GPS signal characteristics derived from real-time experiments. This study makes the following contributions

- A stacked ensemble model is introduced which leverages the strength of both machine learning and deep learning models to identify GPS signal spoofing in small UAVs.
- A systematic approach is designed for dataset acquisition, preparation, and controlled simulation tests.
- For better results from the proposed approach, the data is preprocessed using z-score normalization techniques before being used to train machine learning models.
- Performance comparison to analyze the efficacy of the proposed approach in comparison to well-known machine learning models in the context of detecting GPS signal spoofing.

Section II provides a succinct overview of contemporary literature and significant progress within the domain of IoT-based solutions. Section III presents the data and methodologies utilized in the conducted experiments. Section IV elaborates on the machine learning and deep learning models. Section V offers a comprehensive insight into the experimental outcomes, accompanied by a detailed analysis. Lastly, Section VI serves as the conclusion, wrapping up the article.

II. RELATED WORK

This section explores the advancements, methodologies, and findings from various studies that have addressed the challenges and intricacies of identifying and countering GPS signal manipulation. By examining the range of approaches, techniques, and outcomes reported in prior research, this

section sets the stage for the novel contributions and insights presented in this study. Numerous investigations have been conducted concerning the Categories of GPS Spoofing Signals and the detection of GPS spoofing.

A. CATEGORIES OF GPS SPOOFING SIGNALS

GPS spoofing signals are classified as meaconing or generative spoofing. In meaconing, an authentic GPS signal is captured and rebroadcast, resulting in a timing offset when compared to real GPS signals. Spoofing a UAV using meaconing necessitates jamming to reacquire signals after entering signal-tracking mode [16]. In contrast, generative spoofing requires synchronizing GPS time and modifying navigation signals with a spoofing simulator [17]. The signal's strength steadily increases to alter the target receiver's tracking loops and move it to a counterfeit place. Detection avoidance is achievable using generative spoofing, even when GPS tracking is active. Meaconing attacks may be distinguished by examining the time offset, but generative spoofing offers a more predictable danger to tiny UAVs due to its complexity.

B. GPS SPOOFING ATTACKS DETECTION

Numerous studies have focused on detecting and mitigating GPS spoofing incidents. One approach relies on acceleration error analysis derived from GPS receiver and inertial measurement unit (IMU) measurements [18]. Another strategy utilizes GPS data and IMU, applying the XGBoost model and Genetic Algorithm in a two-step process to detect spoofing attacks [14]. Artificial neural networks were employed to classify GPS signals based on features like pseudo-range, Doppler shift, and SNR, yielding promising detection efficiency [19]. These efforts collectively highlight effective methods for identifying GPS spoofing with various data-driven techniques.

Another method for detecting GPS spoofing was outlined in [20], utilizing vision sensors, monocular cameras, and IMUs in conjunction with UAV sensors. This method employs the fusion of vision sensor and IMU data for GPS spoofing detection. Additionally, a vision-based UAV spoofing identification technique utilizing visual odometry was proposed in [21], leveraging UAV cameras that remain unaffected by fabricated GPS signals. The relative trajectory acquired from images through visual odometry is cross-referenced with GPS-derived flight trajectory data to detect spoofed signals. while another [22] utilizes vision sensors and IMUs for detection. An additional vision-based method [23] exploits Visual Odometry to detect spoofing by comparing UAV camera-derived trajectory with GPS-based flight trajectory information. These techniques collectively address GPS spoofing concerns through predictive modeling, sensor fusion, and image-based analysis, contributing to improved UAV flight security and resistance to deception signals [24].

In [2], a novel framework for GPS spoofing detection was introduced, requiring minimal initial setup and emphasizing information fusion. This real-time approach utilizes IMU data to determine the UAV's present location, cross-referencing it with GPS-derived location information to flag potential GPS spoofing attacks. In [25], an innovative algorithm was presented to handle abrupt system state changes caused by GPS spoofing attacks. Employing a particle filter algorithm, this method counteracts the effects of GPS spoofing by manipulating prediction discrepancies, resulting in improved UAV position estimation and reduced errors.

C. SPOOFING ATTACKS CLASSIFICATION WITH MACHINE LEARNING

In [26], a counter-spoofing model was introduced, employing linear regression for optimal UAV route prediction and long short-term memory (LSTM) for trajectory prediction. The model incorporates multiple identification schemes for GPS spoofing signals, enhancing UAV flight sensitivity and safety to deceptive signal detection. Simulation experiments indicate its efficacy in countering GPS spoofing without escalating hardware expenses.

In various studies, innovative methods for countering GPS spoofing have been proposed. One approach [26] combines linear regression and LSTM to predict optimal UAV routes and enhance sensitivity to GPS spoofing signals. A similar approach for spoofing detection and classification was put out by [27] based on the least absolute shrinkage and selection operator. This technique employs signal processing methods to differentiate authentic and spoofed signals using code-phase values and incorporates a threshold to minimize false alarms. These studies collectively offer diverse strategies for combatting GPS spoofing, encompassing real-time detection, particle filter compensation, and signal processing techniques for accurate identification and classification.

In [28], a methodology was introduced involving multiple models with varying K-fold values, integrating voting techniques to select the best model. Authors [29] explored a resilient framework for detecting and estimating GPS spoofing attacks. The authors addressed sensor drift concerns by managing estimation errors. Machine learning methods offer promise for GPS spoofing detection in small civilian UAVs by eliminating the need for extra hardware. In [30], a support vector machine (SVM)-based approach was proposed to identify UAV GPS spoofing attacks through state estimation analysis. Alternatively, [31] introduced a method relying on received signal strength measurements to establish a credible residence area, effectively discerning between authentic and spoofed GPS positions. In [32], machine learning models were employed for spoofing detection, compared to a path-based approach. Notably, these techniques offer the potential for robust detection, considering varied attack durations and flight patterns.

D. GPS SPOOFING ATTACKS CLASSIFICATION USING DEEP LEARNING MODELS

The recent developments in deep neural networks (DNN) offer a potential resource for filtering out data anomalies.

TABLE 1. Summary of related work on GPS spoofing detection on UAVs.

Ref.	Year	Dataset	Approach/Classifiers	Findings	Limitations
[27]	2021	Real-time GPS-related data	LR anti-spoofing model	Increased capability to resist GPS spoofing, no additional hardware cost, and easy implementation.	Limited evaluation of diverse attack scenarios.
[23]	2023	MPU9250	Motion Processing Units (MPUs)	The method uses data from all three axes to identify GPS spoofing and retrieve correct GPS locations.	Dependency on specific IMU hardware; effectiveness in complex spoofing scenarios not addressed.
[28]	2020	Texas Spoofing Test Battery (TEXBAT) data	LASSO	Correlation profiles, analysis of the contribution of individual components from desired and spoofed signals.	Evaluation limited to specific spoofing test data; generalizability concerns.
[31]	2021	Spoofing data set TEXBAT	Support Vector Machines	After evaluating the effectiveness of various kernel functions and comparing their results with those of earlier detection techniques, a precise and effective automatic detection method using a coarse Gaussian function is created.	Lack of evaluation on diverse spoofing scenarios; scalability concerns.
[33]	2022	A real-time dataset (13 signals features)	Machine learning models	The approaches suggested dynamically choose the model that produces the best results for identifying attacks.	Generalizability concerns; lack of extensive real-world testing.
[34]	2021	UAV flight logs and telemetry data	LSTM	Authors applied LSTM classifier and autoencoder for the GPS spoofing attacks classification.	Performance on complex spoofing scenarios not discussed; scalability concerns.
[35]	2021	Real-time dataset	MultiLayer Perceptron (MLP)	Authors tested three statistical models of MLP under different base stations.	Lack of evaluation under diverse environmental conditions; scalability concerns.
[36]	2023	Real-time dataset	IR-UWB distance measurement	The authors offer a method for detecting GPS spoofing attacks in UAV swarms using an illustration of an IR-UWB supported UAV swarm.	Scalability concerns for larger UAV swarms; real-world validation required.
[37]	2023	GPS signal dataset	PCA-CNN-LSTM	Comparing the proposed model to existing deep learning and machine learning models, it fared better.	Resource-intensive methods; real-time feasibility not extensively discussed.
[38]	2019	Real-time dataset	The ensemble model using the Salp Swarm Algorithm	The weight optimization technique improved the results.	Scalability concerns for varying attack types; generalizability not thoroughly addressed.
[39]	2022	Real-time dataset	1D convolutional neural network	The proposed model enabled detection on mobile platforms.	Performance under varying environmental conditions not extensively evaluated.
[40]	2023	Real-time UAV sensor data	CNN-BiLSTM-Attention (CBA)	The proposed method is evaluated on actual attack scenarios, such as denial-of-service (DoS) attacks that spoof the GPS, and it shows both efficacy and interpretability.	Limited evaluation on diverse spoofing attacks; scalability concerns in complex scenarios.
[41]	2022	Real-time UAV cellular data	Deep ensemble learning methods	The formulated spoofing detection problem as non-linear optimization problem.	Scalability concerns for complex spoofing scenarios; real-world validation required.
[42]	2022	TEXBAT dataset and MAVLINK dataset	MLP	The proposed system generate alarm on spoofing attack detection.	Real-time response in dynamic environments not extensively addressed; scalability concerns.

A study employed 1D CNN for GPS spoofing detection for small UAVs. The authors applied an LSTM autoencoder and classifier for GPS spoofing [33]. Authors introduced a multilayer perceptron (MLP) model in [34]. It was trained using statistical characteristics extracted from path loss measurements collected from adjacent base stations. This trained model is then utilized to assess the credibility of the GPS position. The study introduced an innovative application of deep learning techniques to counteract the impacts of spoofing attacks aimed at one or multiple PMUs concurrently [42].

Through MLP techniques, there are also various ways to identify GPS signal faking. A novel method for spoofing identification is to utilize a neural network (NN), as presented by [43]. The technique collected information on the early-late phase and signal intensity from the tracking loop's correlation output to assess whether the signal was phony. Based on the results, a 99.3% true detection rate is reported. When two deep learning models were evaluated, the researchers discovered that MLP performed better than LSTM. Their method accurately identified GPS spoofing attacks with accuracies of 83.2%(TEXBAT dataset) and 99.9% (MAVLINK dataset) [41]. Dang et al. [40] investigated the effectiveness of statistics from the base stations for spoofing attack detection on cellular UAVs. The MLP model has a simple structure and performs well with some less complex datasets. To avoid overfitting, MLP needs a lot of training in spoofing data because it is not robust enough to handle complicated spoofing datasets.

In addition to the aforementioned deep learning techniques, spoofing attack detection techniques based on CNN have also been used. Using a simple model, the study [38]

a novel antispoofing technique. The suggested solution used the ResNet architecture, which made it more effective than SVM at detecting most spoof signals. Flight experiments were used to gauge the algorithm's efficacy. Wu et al. [39] suggested a methodology for cyber attack detection in real-time. The strategy makes use of an attention model based on CNN and bidirectional LSTM (BiLSTM). The model was successfully used to detect spoofing attacks in a simulation setting. Results revealed a 99.1% spoofing detection accuracy.

E. ROLE OF ENSEMBLE MODELS FOR GPS SPOOFING ATTACKS CLASSIFICATION

Spoofers are constantly designing and developing innovative techniques to attack UAVs, making the security of UAVs challenging. Stand-alone models may not be an appropriate choice to detect such attacks, so ensemble models can be leveraged in this regard. As an illustration, in [44], an ensemble model was introduced to categorize and identify attacks in wireless networks. The proposed technique integrates multiple base learners to get a strong meta-learner. The proposed stacking strategy exhibited superior performance compared to its constituent base learners. Similarly, [37] undertook a comparison of distinct ensemble models to predict RSS power for UAVs. The outcomes showcased the supremacy of stacking over standalone models. The authors [36] developed a framework for GPS signal spoofing detection in small UAVs utilizing the PCA-CNN-LSTM approach, as well as a method for detecting signal spoofing attacks in small UAVs. Many studies are confined to small datasets or scenarios, restricting the applicability of their proposed methods to a wider range of spoofing contexts. The diversity and complexity of spoofing attacks are not fully addressed or evaluated in some works, potentially leaving gaps in understanding and defense against various types of spoofing tactics. Table 1 summarizes the summary of existing studies and their limitations.

In brief, diverse techniques are proposed for detecting GPS spoofing attacks utilizing machine learning algorithms. These methods encompass feature selection and extraction strategies; however, their effectiveness remains limited. To address the challenge of identifying GPS signal spoofing in autonomous vehicles, this study recommends an efficacious approach by employing machine and deep learning algorithms, along with ensemble techniques.

This study stands out by introducing a stacked ensemble model, merging the strengths of machine learning and deep learning models to detect GPS signal spoofing in small UAVs. Unlike previous research, it adopts a structured methodology, systematically acquiring and preparing datasets, and conducting controlled simulation tests. Notably, the study employs z-score normalization techniques for data preprocessing, a step often overlooked in prior works, aiming to optimize model performance. Additionally, it shows a comprehensive performance comparison, analyzing the effectiveness of the proposed approach against established machine learning models specifically for GPS signal spoofing detection. These aspects, from the novel model architecture to the systematic approach and detailed performance analysis, make this study important, potentially offering enhanced detection capabilities for safeguarding small UAVs against spoofing threats.

TABLE 2. Attributes of the dataset.

Attributes	Detailed description
PRN	Satellite Vehicle Number
DO	Carrier Doppler inHz
PD	Pseudo-range in meter
RX	Receiver Time
TOW	Time of the Week in seconds
CP	Carrier Phase Cycles
EC	Magnitude of the EarlyCorrelator
LC	Magnitude of the LateCorrelator
PC	The Magnitude of the PromptCorrelator
PIP	Prompt in phase correlator
PQP	Prompt Quadrature Component
TCD	Carrier Doppler in Trackingloop in Hz
CN0	Carrier to Noise Ratio in dB-Hz

III. DATA COLLECTION AND PREPROCESSING

This section encompasses the crucial stages of GPS spoofing attack detection, starting with the collection of GPS signal data and the subsequent data preprocessing.

A. DATASET

This study makes use of a collection of GPS spoofing incidents. This dataset [45] comprises data from genuine GPS signals gathered from various places to simulate a moving and stationary autonomous automobile using a universal software radio peripheral device configured as a GPS receiver. During the data collection process, 13 features are obtained from eight parallel mediums at various receiver stages (i.e., tracking, navigation decoding, and acquisition). In addition to the gathered legitimate GPS signals, three types of GPS spoofing attempts were simulated: basic, intermediate, and complicated. The generated dataset includes 158,170 samples with a balanced distribution reflecting three types of simulated GPS spoofing attempts and 55% genuine occurrences. Table 1 describes each feature.

B. DATA PREPROCESSING

Data preparation is critical since it enhances model performance and results in more accurate features. In this part, data preparation is accomplished by data analysis, cleansing, and outlier removal using Z-score normalization. To begin preprocessing the GPS spoofing dataset, its size, and general information are examined using the .shape property. To obtain the data types and the quantity of non-null values for each variable, use the.info() function. The variables PD, TCD, and CNO are then investigated as well as their distribution. Following that, we examine the dataset for missing values. We infer from our data exploration that the dataset contains all values. Then we look for outliers in the data.

C. Z-SCORE NORMALIZATION

Following analysis, the data are normalized using the z-score method. The dataset used in this study contains an 8-channel GPS receiver. It indicates that each feature in the dataset contains values of 8 different channels. Every feature in each channel contains 158,170 continuous numeric values ranging from -3368 to 491783. Due to the higher diversity in features and channel values, Z-score normalization (standardization) is one of the most important data preprocessing techniques used for machine learning. Algorithms that compute the distance between the features are biased towards numerically larger values if the data is not scaled. Tree-based algorithms are fairly insensitive to the scale of the features. Also, feature scaling helps machine learning and deep learning algorithms train and converge faster. For these reasons, this study adopted the z-score normalization using Scikit-Learn provided transformer, StandardScaler, for standardization. It translates the data to the mean vector of the original data to the origin and squishes or expands. The z-score normalization is applied to the entire dataset to scale up the dataset and avoid any kind of favoritism to the higher numeric values.

The statistical process of standardization, also known as Z-score normalization, transforms a value distribution into a mean of 0 and a standard deviation of 1. Both feature scaling and data preparation frequently involve its utilization. The mean is subtracted, and the standard deviation is divided, to normalize the Z-score. The Z-score of a data point may be calculated using the following equation: The Z-score can be calculated using

$$z = \frac{(x - \mu)}{\sigma} \tag{1}$$

where z denotes the data point's Z-score, x denotes its value, *mu* is the dataset's mean (average), and *sigma* denotes its standard deviation.

The Z-score displays the number of standard deviations of a data point from the mean value. Positive and negative Z-scores indicate above the mean, and below the mean values, respectively. Using z-score normalization, the data distribution is altered which facilitates the comparison of different features of the dataset.

IV. MATERIALS AND METHODS

Within this section, the methods employed for GPS spoofing detection are discussed. The supervised machine learning models, deep learning models, and ensemble models are used in experiments. The hyperparameter details of the models are given in Table 3.

A. RANDOM FOREST

Random forest (RF) [46] generates multiple trees and employs randomness to mitigate variability. It has gained substantial attention in research for addressing classification and regression tasks involving groups of data. RF adopts a bagging strategy, combining predictions through majority voting, and operates on bootstrap samples from the initial dataset. The operational concept of RF can be described as follows:

$$p = mode\{T_1(y), T_2(y), \ldots, T_m(y)\} \tag{2}$$

$$p = mode\left\{\sum_{m=1}^{m} T_m(y)\right\} \tag{3}$$

The final output, designated as 'p,' is determined by employing majority voting among the predictions generated by the individual trees, denoted as T_1, T_2, and T_m."

TABLE 3. Hyperparameter details of all classifiers.

Classifier	Hyperparameter
LR	C = 10, class_weight='balanced', l1_ratio = 0.7, max_iter = 3000, penalty = 'elasticnet', solver = 'saga'
SVM	C = 300, class_weight = 'balanced'
RF	n_estimators = 300, criterion='entropy', max_depth = 30,
DT	criterion='entropy' , max_depth = 30,
ETC	n_estimators = 300, max_depth = 30, criterion='entropy'
KNN	n_neighbors = 5, leaf_size = 35
VC	criteria='soft', n_jobs = -1
CNN	Stride = (1 × 1), pool size= (@ 2), filter= (@ 256), Dense neuron (60), activation ='Relu'
LSTM	return_sequences= True, Dense neurons/Units (60), activation ='Relu'
RNN	Dense neurons/Units (60), activation ='Relu'

B. SUPPORT VECTOR MACHINE

SVM [47] stands as a potent machine learning methodology, proficient in tackling both classification and regression challenges. It achieves this by employing the kernel trick to transform data and establish optimal boundary lines, referred to as hyperplanes, that demarcate different outputs effectively. These hyperplanes serve to segregate data points of distinct types. The foundational principle of data classification revolves around crafting a function that consistently assigns labels to data points, all the while minimizing errors or maximizing margins. A wider margin surrounding the separating function translates to fewer errors. By constructing this function, labels are more distinctly segregated. In this specific instance, the linear kernel is utilized, which offers high accuracy.

C. LOGISTIC REGRESSION

Logistic regression (LR), as documented in Wright's work [48], presents a prominent approach for addressing challenges in classification tasks. Rooted in statistical principles, this method operates by leveraging the concept of probabilities. It demonstrates particular efficacy when dealing with binary data, aiming to predict outcomes using one or more explanatory variables. Logistic regression employs a sigmoid function, also termed a logistic function, to establish relationships within categorical data. This sigmoid function transforms input values into a range between 0 and 1, giving rise to an S-shaped curve. This enables logistic regression to gauge the probability of a specific class or event occurrence.

The numerical value indicated within logistic regression involves the translation of real numbers, where ''e'' signifies

the natural logarithmic base. For model optimization, logistic regression is executed with 100 iterations (max_iter). The parameter ''penalty'' is configured as ''12,'' thereby determining the penalty norm applied to the model.

D. DECISION TREE

Decision trees (DT), as outlined in the work by Breiman [49], constitute a fundamental yet potent supervised machine-learning methodology, adept at accommodating both numerical and categorical input. Its exceptional adaptability has led to its widespread utilization across diverse fields. The root node is selected based on the Gini Index which is calculated as:

$$Gini = 1 - \sum_{i=1}^{classes} p(i|t)^2 \tag{4}$$

A key advantage of decision trees is their straightforward implementation. These trees employ decision rules and subsets of features at varying classification levels. They comprise branches featuring internal and leaf nodes. Internal nodes represent individual features, while branches leading to groupings represent combinations of features. Each leaf node corresponds to a class, embodying an example. The efficacy of a decision tree's construction profoundly influences its performance on training datasets.

E. K NEAREST NEIGHBOUR

The K-nearest neighbors (KNN) classifier [50] is a widely recognized algorithm. It operates as a non-parametric, instance-driven learning technique, thereby avoiding any assumptions about the underlying data distribution. Employing estimated distances, the KNN algorithm selects K examples from the training dataset that exhibit the closest resemblance to the new instance. Once K's nearest neighbors are identified, the algorithm conducts a majority voting process among their associated class labels. Consequently, the projected class for a novel instance is determined by the class label that attains the highest frequency within the set of K neighbors.

F. EXTRA TREE CLASSIFIER

The extra tree classifier (ETC) algorithm, as introduced by Sharaff [51], shares similarities with the Random Forest (RF) technique, albeit with distinct tree-building approaches. Unlike RF, ETC opts for using the original dataset to construct trees, foregoing the utilization of bootstrap samples. ETC's decision-making process is rooted in random data sampling from the top k-best features. The selection of the optimal feature for tree partitioning employs the Gini index. While both ETC and RF serve as ensemble learning models for classification tasks, their divergence emerges in the manner of tree assembly within their respective forests. In the case of ETC, K features are randomly sampled from the feature pool and subsequently allocated to the test nodes of each tree.

G. LONG SHORT-TERM MEMORY

LSTM represents an advanced paradigm within the realm of deep learning and stands as an evolutionary development of Recurrent Neural Networks (RNNs), as referenced by [52]. Within the architecture of an LSTM, crucial elements such as the forget gate (f_k), the input gate (i_k), and the output gate (o_k) are embedded. These gates serve to facilitate the controlled passage of data, enabling the retention of salient information while filtering out extraneous data, contingent on the predetermined dropout threshold. Central to the LSTM model is the inclusion of a dedicated memory component denoted as C_k, functioning as a repository for pivotal information. Notably, various iterations of LSTM configurations exist.

The LSTM model corresponds to the associated weights involving matrix components. The accumulated hidden state until the $(k-1)$ time step is denoted as h, while s_k signifies the input at that specific time step. The bias term is symbolized by b. During the $(k-1)$ time step, adjustments are made to the memory cell block represented as c. Every neuron within the output layer of the LSTM maintains connections with all neurons within the dense layer, indicating a fully interconnected structure.

H. CONVOLUTIONAL NEURAL NETWORK

CNN is designed to capture intricate patterns through the utilization of convolutional and pooling layers, as highlighted by Yamashita [53]. CNNs find widespread application in tasks such as image segmentation and classification. The robustness of layered CNN models is bolstered by end-to-end training, ensuring their adaptability.

Functioning as a feed-forward network model, CNN's convolutional layers process input data by applying filters to the output of preceding levels. CNN also contains pooling, dropout, and fully connected layers. Pooling contributes to feature selection by reducing feature dimensions, and it can be implemented as either average or max-pooling. The outputs of preceding layers are directed to fully connected layers, which ultimately determine the final outcome. Dropout layers are strategically employed to mitigate overfitting risks. The selection of an appropriate activation function is crucial in discerning the significance of input information.

I. PROPOSED ENSEMBLE MODEL

This ensemble model combines machine learning and a deep learning algorithm, creating a synergistic union. Ensembling stands as a potent strategy involving the aggregation of predictions from diverse models to enhance accuracy and resilience. Each model within an ensemble brings forth its distinct merits and limitations, and their amalgamation yields superior overall performance. The proposed architecture is presented in Figure 1.

The ensemble model's functionality revolves around harmonizing the predictions from two distinct learning algorithms. The conventional strategy for constructing an

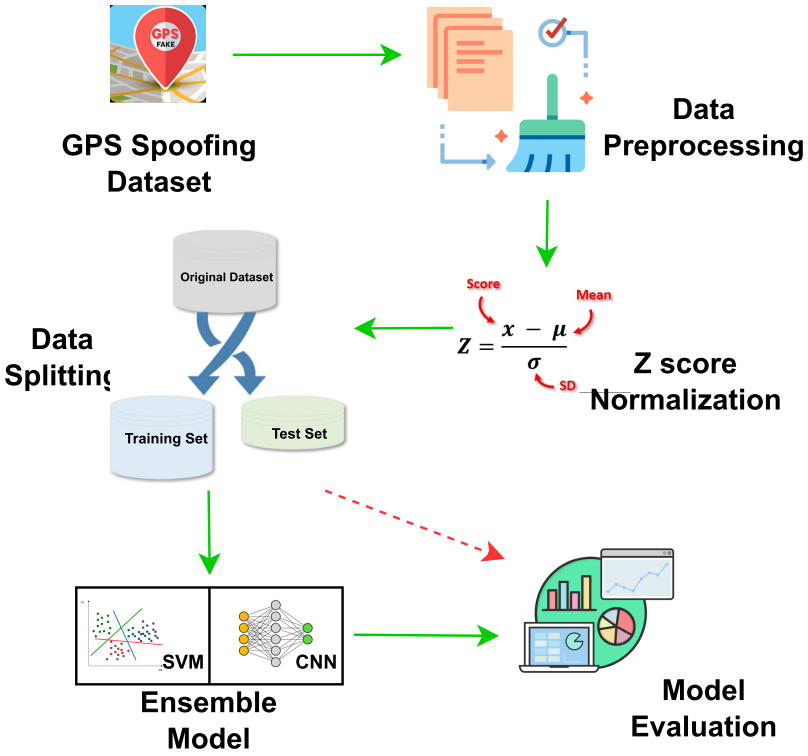

FIGURE 1. Architecture diagram of the proposed approach.

ensemble involves training multiple models on an identical dataset, followed by fusing their predictions. The SVM-CNN ensemble model adheres to this methodology, training SVM and CNN models separately on the same dataset. These models individually generate predictive probabilities for each class of the target variable. This assortment of predicted probabilities can then be combined to render a conclusive prediction for every observation within the dataset. A prevalent method to aggregate these predictions is through a weighted average of predicted probabilities, wherein the weights are determined based on each model's performance on a validation set.

The proposed ensemble model leverages the strengths of both machine learning and deep learning algorithms, leading to enhanced accuracy and robustness in predictions. By training diverse models on the air quality dataset and merging their predictions, we elevate the model's capability for generalization and curtail overfitting. The operational dynamics of the envisaged ensemble model are encapsulated by

$$\widehat{p} = argmax\{\sum_{i}^{n} SVM_i, \sum_{i}^{n} CNN_i\}. \tag{5}$$

where $\sum_{i}^{n} SVM_i$ and $\sum_{i}^{n} CNN_i$ yield prediction probabilities for each test sample. Subsequently, the probabilities generated by SVM and CNN for each test case undergo assessment using the soft voting criterion.

The ensemble model arrives at its final class designation by considering the highest average probability across classes and by aggregating the projected probabilities from both classifiers. The conclusive prediction corresponds to the class with the most substantial probability score, as evidenced by

$$VC(SVM + CNN) = argmax(g(x)) \tag{6}$$

V. EXPERIMENTS AND RESULTS

Within this section, an in-depth analysis of the performance exhibited by the proposed model is conducted. This model employs a variety of machine learning classifiers and is deployed on a GPS spoofing dataset. A comprehensive evaluation is undertaken, encompassing multiple metrics such as accuracy, recall, precision, and F1 score. These metrics collectively serve as evaluative benchmarks to ascertain how effectively the proposed model performs in comparison to established approaches. The spoofing data has been partitioned into training and test sets, following a 70:30 ratio.

A. IMPLEMENTATION DETAILS

The experimental setup involved training the model. The prescribed model was developed within the Python 3.8 programming environment. For performance evaluation, accuracy, precision, recall, and F1 score are used with the following equations

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

TABLE 4. Experimental results of machine learning models.

Classifiers	Accuracy	Precision	Recall	F1 score
RF	95.89%	95.62%	95.25%	95.45%
ETC	91.47%	85.17%	87.74%	86.29%
LR	93.09%	91.29%	93.44%	92.54%
KNN	89.68%	88.50%	89.74%	89.18%
DT	94.49%	94.22%	94.58%	94.37%
SVM	96.50%	97.34%	98.18%	97.87%

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively.

B. RESULTS OF MACHINE LEARNING MODELS

The experiments incorporate an array of machine learning classifiers including RF, ETC, LR, KNN, DT, and SVM. Table 4 illustrates the results of machine learning models and offers an overview of their overall efficacy. The SVM model exhibited a commendable score with 96.50% accuracy in accurately identifying instances within the test set. Its precision, recall, and F1-score are recorded as 97.34%, 98.18%, and 97.87% respectively. Conversely, the KNN model demonstrated a notably low score on all metrics with 89.68% accuracy, 88.50% precision, 89.74% recall, and 89.18% F1 score. ETC, LR, and DT have shown moderate levels of performance in identifying spoofing attacks from GPS-based datasets.

The experimental findings demonstrate that both RF and SVM exhibit better results when compared to other models in precisely categorizing GPS spoofing attacks. By analyzing these results, it's evident that the SVM classifier outperformed the other models in terms of accuracy, precision, recall, and F1 score. This indicates that the SVM model exhibited a high level of accuracy in identifying GPS spoofing attacks and was particularly effective in achieving both high precision and recall rates. These results suggest that the SVM model holds promise as a robust tool among machine learning models for detecting and classifying GPS spoofing attacks.

C. RESULTS OF DEEP LEARNING MODELS

Deep learning models used in experiments include LSTM, RNN, and CNN. The performance of these models is presented in Table 5. Results show the performance evaluation of different classifiers, LSTM has shown better results than RNN in accurately identifying spoofing attacks. The CNN achieves an accuracy of 97.37% and 98.47% score of precision. The recall metric measures the classifier's ability to detect actual positive instances, with CNN achieving 98.88%. Lastly, the F1-score, harmonizing precision and recall, is 98.69% for CNN. These results collectively underscore CNN's superior performance among the deep learning models, highlighting its potential as a robust tool for accurate GPS spoofing attack classification.

TABLE 5. Experimental results of deep learning models.

Classifiers	Accuracy	Precision	Recall	F1 score
LSTM	90.29%	87.16%	87.37%	87.24%
RNN	88.34%	85.52%	85.19%	85.32%
CNN	97.37%	98.47%	98.88%	98.69%

Collectively, these findings accentuate the superior performance of the CNN architecture among the evaluated deep learning models. CNN's ability to accurately identify GPS spoofing attacks, as evidenced by its high precision, recall, and F1 score, signifies its potential as a reliable and robust tool for combating such security threats. The CNN's capacity to extract and discern significant spatial and temporal features from the data contributes substantially to its efficacy in classification tasks, thereby establishing it as a promising choice for accurate GPS spoofing attack detection and classification. These results underscore the viability of CNN as a pivotal component in the arsenal against GPS spoofing threats, paving the way for enhanced security measures and reliable identification of spoofing attempts.

D. RESULTS OF ENSEMBLE MODELS

Table 6 presents the performance of ensemble models in the context of classifying GPS spoofing attacks. These ensemble models combine the capabilities of multiple individual classifiers to improve predictive accuracy. The ensemble configuration of models is labeled as "Ensemble (RF-CNN)," "Ensemble (RF-LSTM)," "Ensemble (SVM-LSTM)," and "Proposed Ensemble (SVM-CNN)." It can be observed that an ensemble of machine learning and deep learning models has shown improved results. Notably, the "Proposed Ensemble (SVM-CNN)" stands out with exceptional results, achieving an accuracy of 99.72%, a precision of 99.65%, a recall of 99.77%, and an F1-score of 99.72%. Overall, the table demonstrates that ensemble models, such as "Proposed Ensemble (SVM-CNN)," yield remarkable results in accurately classifying GPS spoofing attacks. These ensemble models leverage the complementary strengths of their constituent classifiers, leading to enhanced predictive accuracy, precision, recall, and F1-score. These findings suggest that ensemble techniques hold significant promise for the effective detection and categorization of GPS spoofing attacks.

The results from the ensemble models, notably the proposed Ensemble (SVM-CNN), demonstrate a significant leap in accurately identifying GPS spoofing attacks. By combining the strengths of various classifiers, these ensembles showcase exceptional performance, achieving robust accuracy, precision, recall, and F1 scores. This robustness indicates the potential of ensemble techniques as a cornerstone in bolstering security measures against GPS spoofing threats. The visual representation in Figure 2 further solidifies the ensemble models' dominance in classifier accuracy, underlining their promise for effective detection

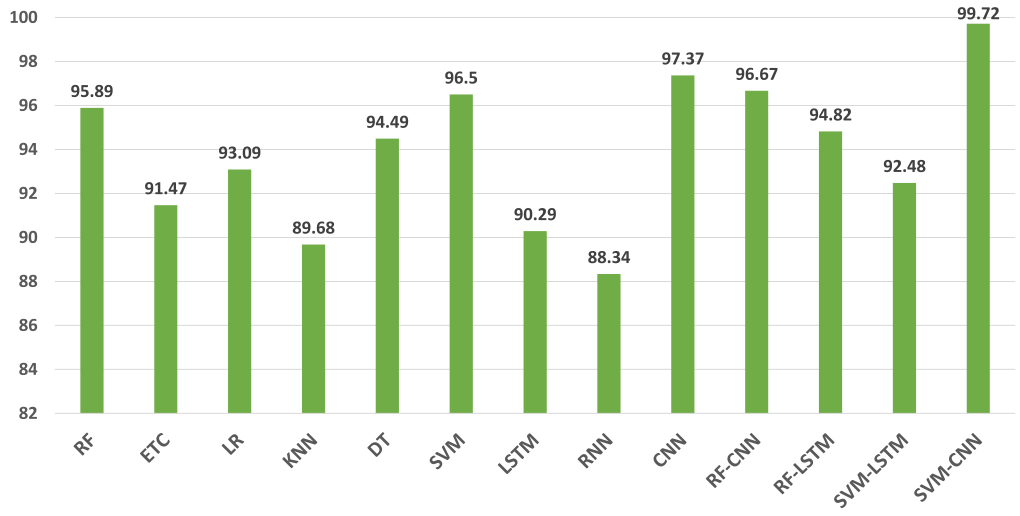

FIGURE 2. Accuracy comparison of all employed models.

TABLE 6. Experimental results of ensemble models.

Classifiers	Performance in %			
	Accuracy	Precision	Recall	F1 score
RF-CNN	97.67	97.20	97.78	97.49
RF-LSTM	94.82	91.37	92.08	91.89
SVM-LSTM	92.48	94.52	94.18	94.29
Proposed (SVM-CNN)	99.72	99.65	99.77	99.72

and classification of GPS spoofing attacks. Overall, these findings underscore the efficacy of ensemble methods as a promising avenue for fortifying defenses against such security vulnerabilities.

E. SIGNIFICANCE OF PROPOSED APPROACH

This study also performs cross-validation to analyze the significance of the proposed model SVM-CNN. Table 7 illustrates the cross-validation results for the SVM-CNN model for classifying GPS spoofing attacks. A model's performance can be reliably analyzed using cross-validation where k folds are considered to investigate the model's ability and robustness to provide accurate results with lower standard deviation. The "average" row at the bottom of the table presents the mean values of these performance metrics across all folds. The Proposed Ensemble (SVM-CNN) model exhibits consistently high performance across the folds, with an average accuracy, precision, recall, and F-Score of 99.72%, 99.75%, 99.79%, and 99.77% respectively. This suggests that the ensemble model is effective and reliable in accurately classifying GPS spoofing attacks across various subsets of the dataset, indicating its robustness and generalization capability.

F. COMPARISON WITH STATE-OF-THE-ART EXISTING MODELS

To evaluate the significance of the proposed SVM-CNN model, this study undertakes a performance analysis by comparing it with other advanced approaches commonly used for detecting GPS spoofing attacks.

TABLE 7. Results for k-fold cross-validation of the proposed ensemble (SVM-CNN).

Fold Number	Accuracy	Precision	Recall	F-Score
Fold-1	99.45	99.76	99.79	99.77
Fold-2	99.78	99.78	99.83	99.81
Fold-3	99.67	99.82	99.88	99.85
Fold-4	99.55	99.70	99.76	99.73
Fold-5	99.79	99.86	99.76	99.81
Average	**99.72**	**99.75**	**99.79**	**99.77**

TABLE 8. Experimental results of ensemble models.

Reference	Year	Model	Accuracy
[31]	2021	SVM	92.31%
[55]	2022	Nu SVM	92.78%
[37]	2023	PCA-CNN-LSTM	99.49%
Proposed Ensemble	2023	(SVM-CNN)	99.72%

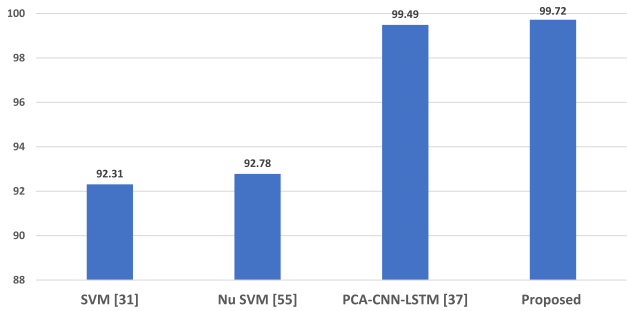

FIGURE 3. Accuracy comparison with state-of-the-art models.

Table 8 facilitates a comparative analysis between the proposed ensemble (SVM-CNN) model and several state-of-the-art existing models in the field of GPS spoofing attack detection. Notably, the study [36] introduces a PCA-CNN-LSTM model achieving an impressive accuracy of 99.49% in detecting GPS spoofing attacks. In contrast, the proposed ensemble (SVM-CNN) model, featured in this study, surpasses the other models with a notably higher accuracy of 99.72%. This comparative assessment

emphasizes the effectiveness of the proposed model in GPS spoofing attack detection and its competitive standing among contemporary state-of-the-art methodologies as shown in Figure 3.

VI. CONCLUSION

This work presented a complete way for detecting GPS signal spoofing using an ensemble model. The dataset was rigorously collected using a composite-wing UAV and a deceptive misleading spoofing blocker. The following phases involved data analysis and preparation, most notably z-score normalization. Following that, an ensemble of machine learning and deep learning models is devised to detect instances of GPS signal spoofing. The use of confusion matrices as a rigorous assessment tool was critical in determining the model's computational efficiency. Experimental results demonstrate the ensemble model's efficiency, with an outstanding accuracy rate of 99.74%. Notably, this work used actual GPS spoofing signal data, assuring the preservation of critical data features that support GPS signal manipulation while providing greater dependability than simulation-derived datasets. Finally, a comprehensive model capable of detecting GPS signal spoofing was developed that is an ensemble of both machine learning methods and deep neural networks. The proposed ensemble SVM-CNN has shown significant performance in identifying spoofing actions in UAVs. Furthermore, this study explains the move from complex machine learning models to simpler deep learning models in UAV GPS signal spoofing detection. Future initiatives will incorporate cutting-edge deep learning methodologies and models to improve spoofing attack detection.

REFERENCES

[1] Y. Liu, S. Li, Q. Fu, and Z. Liu, "Impact assessment of GNSS spoofing attacks on INS/GNSS integrated navigation system," *Sensors*, vol. 18, no. 5, p. 1433, May 2018.

[2] C. Liang, M. Miao, J. Ma, H. Yan, Q. Zhang, X. Li, and T. Li, "Detection of GPS spoofing attack on unmanned aerial vehicle system," in *Proc. Int. Conf. Mach. Learn. Cyber Secur.*, Xi'an, China. Cham, Switzerland: Springer, Sep. 2019, pp. 123–139.

[3] T. Benarbia and K. Kyamakya, "A literature review of drone-based package delivery logistics systems and their implementation feasibility," *Sustainability*, vol. 14, no. 1, p. 360, Dec. 2021.

[4] M. Moshref-Javadi and M. Winkenbach, "Applications and research avenues for drone-based models in logistics: A classification and review," *Expert Syst. Appl.*, vol. 177, Sep. 2021, Art. no. 114854.

[5] S. Mohsan, N. Othman, Y. Li, M. Alsharif, and M. Khan, "Unmanned aerial vehicles (UAVs): Practical aspects, applications, open challenges, security issues, and future trends," *Intell. Service Robot.*, vol. 16, no. 1, pp. 109–137, 2023.

[6] N. H. Motlagh, T. Taleb, and O. Arouk, "Low-altitude unmanned aerial vehicles-based Internet of Things services: Comprehensive survey and future perspectives," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 899–922, Dec. 2016.

[7] P. Jiang, H. Wu, and C. Xin, "DeepPOSE: Detecting GPS spoofing attack via deep recurrent neural network," *Digit. Commun. Netw.*, vol. 8, no. 5, pp. 791–803, Oct. 2022.

[8] S. Khandker, H. Turtiainen, A. Costin, and T. Hämäläinen, "Cybersecurity attacks on software logic and error handling within ADS-B implementations: Systematic testing of resilience and countermeasures," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 4, pp. 2702–2719, Aug. 2022.

[9] N. Souli, P. Kolios, and G. Ellinas, "Online relative positioning of autonomous vehicles using signals of opportunity," *IEEE Trans. Intell. Vehicles*, vol. 7, no. 4, pp. 873–885, Dec. 2022.

[10] J. Kim, S. Lee, and M. Jung, "Case study on the user interface of GPS plotters to enhance their usability," *J. Mar. Sci. Eng.*, vol. 9, no. 1, p. 57, Jan. 2021.

[11] M. A. Mehdi, S. Z. N. Zukhraf, and H. Maryam, "Analysis of vulnerabilities in cybersecurity in unmanned air vehicles," in *Studies in Computational Intelligence*, Cham, Switzerland: Springer, 2022, pp. 131–143.

[12] S. Mohanti, N. Soltani, K. Sankhe, D. Jaisinghani, M. Di Felice, and K. Chowdhury, "AirID: Injecting a custom RF fingerprint for enhanced UAV identification using deep learning," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[13] G. Muralikrishna, G. Mallesham, and M. Kannan, "Autonomous integrity monitoring of INS/GPS integrated navigation system under multipath environment," in *Proc. 6th Int. Conf. Electron., Commun. Aerosp. Technol.*, Dec. 2022, pp. 55–62.

[14] Z. Feng, N. Guan, M. Lv, W. Liu, Q. Deng, X. Liu, and W. Yi, "Efficient drone hijacking detection using two-step GA-XGBoost," *J. Syst. Archit.*, vol. 103, Feb. 2020, Art. no. 101694.

[15] T. T. Khoei, S. Ismail, and N. Kaabouch, "Boosting-based models with tree-structured Parzen estimator optimization to detect intrusion attacks on smart grid," in *Proc. IEEE 12th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Dec. 2021, pp. 165–170.

[16] J. Noh, Y. Kwon, Y. Son, H. Shin, D. Kim, J. Choi, and Y. Kim, "Tractor beam: Safe-hijacking of consumer drones with adaptive GPS spoofing," *ACM Trans. Privacy Secur.*, vol. 22, no. 2, pp. 1–26, Apr. 2019.

[17] M. L. Psiaki and T. E. Humphreys, "GNSS spoofing and detection," *Proc. IEEE*, vol. 104, no. 6, pp. 1258–1270, Jun. 2016.

[18] K.-C. Kwon and D.-S. Shim, "Performance analysis of direct GPS spoofing detection method with AHRS/accelerometer," *Sensors*, vol. 20, no. 4, p. 954, Feb. 2020.

[19] M. R. Manesh, J. Kenney, W. C. Hu, V. K. Devabhaktuni, and N. Kaabouch, "Detection of GPS spoofing attacks on unmanned aerial systems," in *Proc. 16th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2019, pp. 1–6.

[20] Y. Qiao, Y. Zhang, and X. Du, "A vision-based GPS-spoofing detection method for small UAVs," in *Proc. 13th Int. Conf. Comput. Intell. Secur. (CIS)*, Dec. 2017, pp. 312–316.

[21] M. Varshosaz, A. Afary, B. Mojaradi, M. Saadatseresht, and E. G. Parmehr, "Spoofing detection of civilian UAVs using visual odometry," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 1, p. 6, Dec. 2019.

[22] M. Y. Arafat, M. M. Alam, and S. Moh, "Vision-based navigation techniques for unmanned aerial vehicles: Review and challenges," *Drones*, vol. 7, no. 2, p. 89, Jan. 2023.

[23] P. Srinivasan and S. Sathyadevan, "GPS spoofing detection in UAV using motion processing unit," in *Proc. 11th Int. Symp. Digit. Forensics Secur. (ISDFS)*, May 2023, pp. 1–4.

[24] H. Engwerda, M. Snijders, and J. C. Sadlier, "Experimental results of integrity monitoring for UAV flights in urban environment leveraging image based masking," in *Proc. Int. Tech. Meeting Inst. Navigat.*, Feb. 2023, pp. 413–427.

[25] M. Majidi, A. Erfanian, and H. Khaloozadeh, "Prediction-discrepancy based on innovative particle filter for estimating UAV true position in the presence of the GPS spoofing attacks," *IET Radar, Sonar Navigat.*, vol. 14, no. 6, pp. 887–897, Apr. 2020.

[26] L. Meng, L. Yang, S. Ren, G. Tang, L. Zhang, F. Yang, and W. Yang, "An approach of linear regression-based UAV GPS spoofing detection," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–16, May 2021.

[27] E. Schmidt, N. Gatsis, and D. Akopian, "A GPS spoofing detection and classification correlator-based technique using the LASSO," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4224–4237, Dec. 2020.

[28] A. Shafique, A. Mehmood, and M. Elhadef, "Detecting signal spoofing attack in UAVs using machine learning models," *IEEE Access*, vol. 9, pp. 93803–93815, 2021.

[29] H.-J. Yoon, W. Wan, H. Kim, N. Hovakimyan, L. Sha, and P. G. Voulgaris, "Towards resilient UAV: Escape time in gps denied environment with sensor drift," *IFAC-PapersOnLine*, vol. 52, no. 12, pp. 423–428, 2019.

[30] X. Zhu, T. Hua, F. Yang, G. Tu, and X. Chen, "Global positioning system spoofing detection based on support vector machines," *IET Radar, Sonar Navigat.*, vol. 16, no. 2, pp. 224–237, Oct. 2021.

[31] Y. Dang, C. Benzaïd, Y. Shen, and T. Taleb, "GPS spoofing detector with adaptive trustable residence area for cellular based-UAVs," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.
[32] T. Talaei Khoei, S. Ismail, and N. Kaabouch, "Dynamic selection techniques for detecting GPS spoofing attacks on UAVs," *Sensors*, vol. 22, no. 2, p. 662, Jan. 2022.
[33] R. A. Agyapong, M. Nabil, A.-R. Nuhu, M. I. Rasul, and A. Homaifar, "Efficient detection of GPS spoofing attacks on unmanned aerial vehicles using deep learning," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2021, pp. 01–08.
[34] Y. Dang, C. Benzaïd, B. Yang, and T. Taleb, "Deep learning for GPS spoofing detection in cellular-enabled UAV systems," in *Proc. Int. Conf. Netw. Netw. Appl. (NaNA)*, Oct. 2021, pp. 501–506.
[35] P. Mykytyn, M. Brzozowski, Z. Dyka, and P. Langendoerfer, "GPS-spoofing attack detection mechanism for UAV swarms," in *Proc. 12th Medit. Conf. Embedded Comput. (MECO)*, Jun. 2023, pp. 1–8.
[36] Y. Sun, M. Yu, L. Wang, T. Li, and M. Dong, "A deep-learning-based GPS signal spoofing detection method for small UAVs," *Drones*, vol. 7, no. 6, p. 370, Jun. 2023.
[37] S. K. Goudos and G. Athanasiadou, "Application of an ensemble method to UAV power modeling for cellular communications," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, pp. 2340–2344, 2019.
[38] Y.-H. Sung, S.-J. Park, D.-Y. Kim, and S. Kim, "GPS spoofing detection method for small UAVs using 1D convolution neural network," *Sensors*, vol. 22, no. 23, p. 9412, Dec. 2022.
[39] S. Wu, Y. Li, Z. Wang, Z. Tan, and Q. Pan, "A highly interpretable framework for generic low-cost UAV attack detection," *IEEE Sensors J.*, vol. 23, no. 7, pp. 7288–7300, Apr. 2023.
[40] Y. Dang, C. Benzaid, B. Yang, T. Taleb, and Y. Shen, "Deep-ensemble-learning-based GPS spoofing detection for cellular-connected UAVs," *IEEE Internet Things J.*, vol. 9, no. 24, pp. 25068–25085, Dec. 2022.
[41] O. Jullian, B. Otero, M. Stojilović, J. J. Costa, J. Verdú, and M. A. Pajuelo, "Deep learning detection of GPS spoofing," in *Machine Learning, Optimization, and Data Science*. Cham, Switzerland: Springer, 2022, pp. 527–540.
[42] F. Almutairy, L. Scekic, M. Matar, R. Elmoudi, and S. Wshah, "Detection and mitigation of GPS spoofing attacks on phasor measurement units using deep learning," *Int. J. Electr. Power Energy Syst.*, vol. 151, Sep. 2023, Art. no. 109160.
[43] E. Shafiee, M. R. Mosavi, and M. Moazedi, "Detection of spoofing attack using machine learning based on multi-layer neural network in single-frequency GPS receivers," *J. Navigat.*, vol. 71, no. 1, pp. 169–188, Aug. 2017.
[44] H. Rajadurai and U. D. Gandhi, "A stacked ensemble learning model for intrusion detection in wireless network," *Neural Comput. Appl.*, vol. 34, no. 18, pp. 15387–15395, May 2020.
[45] G. A. Aissou, S. B. Benouadah, H. E. A. EL Alami, and N. K. Kaabouch. (2022). *A Dataset for Gps Spoofing Detection on Autonomous Vehicles*. [Online]. Available: https://ieee-dataport.org/documents/dataset-gps-spoofing-detection-autonomous-vehicles
[46] S. J. Rigatti, "Random forest," *J. Insurance Med.*, vol. 47, no. 1, pp. 31–39, 2017.
[47] D. A. Pisner and D. M. Schnyer, "Support vector machine," in *Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 101–121.
[48] R. E. Wright, "Logistic regression," in *Reading and Understanding Multivariate Statistics*, L. G. Grimm and P. R. Yarnold, Eds. American Psychological Association, 1995, pp. 217–244.
[49] L. Breiman, *Classification and Regression Trees*. Oxfordshire, U.K.: Routledge, 2017.
[50] O. Kramer and O. Kramer, "K-nearest neighbors," in *Dimensionality Reduction With Unsupervised Nearest Neighbors*. Berlin, Germany: Springer-Verlag, 2013, pp. 13–23.
[51] A. Sharaff and H. Gupta, "Extra-tree classifier with metaheuristics approach for email classification," in *Advances in Intelligent Systems and Computing*. Cham, Switzerland: Springer, 2019, pp. 189–197.
[52] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D, Nonlinear Phenomena*, vol. 404, Mar. 2020, Art. no. 132306.
[53] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: An overview and application in radiology," *Insights Imag.*, vol. 9, no. 4, pp. 611–629, Aug. 2018.
[54] G. Aissou, S. Benouadah, H. El Alami, and N. Kaabouch, "Instance-based supervised machine learning models for detecting GPS spoofing attacks on UAS," in *Proc. IEEE 12th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2022, pp. 208–214.

ALA' ABDULMAJID ESHMAWI is currently an Assistant Professor of cybersecurity with the University of Jeddah, Saudi Arabia. His research interests include cybersecurity, machine learning, data mining, and the IoT.

MUHAMMAD UMER received B.S. and M.S. degrees from the Department of Computer Science, Khwaja Fareed University of Engineering and IT (KFUEIT), Pakistan, in October 2014 and October 2020, respectively, and the Ph.D. degree in computer science from KFUEIT, in February 2024. He served as a Research Assistant at Fareed Computing and Research Center, KFUEIT. Currently, he is serving as Head of the Software Engineering Department at The Islamia University of Bahawalpur, Pakistan. His recent research interests are related to data mining, mainly working with machine learning and deep learning-based IoT, text mining, and computer vision tasks.

IMRAN ASHRAF received the M.S. degree in computer science from the Blekinge Institute of Technology, Karlskrona, Sweden, in 2010, and the Ph.D. degree in information and communication engineering from Yeungnam University, Gyeongsan, South Korea, in 2018. He was a Postdoctoral Fellow with Yeungnam University, where he is currently an Assistant Professor with the Information and Communication Engineering Department. His research areas include indoor positioning and localization, advanced location-based services in wireless communication, smart sensors (LIDAR) for smart cars, and data mining.

YONGWAN PARK received the B.E. and M.E. degrees in electrical engineering from Kyungpook University, Daegu, South Korea, in 1982 and 1984, respectively, and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York at Buffalo, USA, in 1989 and 1992, respectively. He was with the California Institute of Technology, as a Research Fellow, from 1992 to 1993. From 1994 to 1996, he was a Chief Researcher of developing IMT-2000 system with SK Telecom, South Korea. Since 1996, he has been a Professor of information and communication engineering with Yeungnam University, South Korea. From January 2000 to February 2000, he was an Invited Professor with the NTT DoCoMo Wireless Laboratory, Japan. He was also a Visiting Professor with UC Irvine, USA, in 2003. From 2008 to 2009, he was the Director of the Technology Innovation Center for Wireless Multimedia by Korean Government. From 2009 to March 2017, he was the President of the Gyeongbuk Institute of IT Convergence Industry Technology (GITC), South Korea. He is also the Chairman of the 5G Forum Convergence Service Committee, South Korea. His current research interests include 5G systems in communication, OFDM, PAPR reduction, indoor location-based services in wireless communication, and smart sensors (LIDAR) for smart cars.