SafeRoute: Autoencoder-Driven Unified Models for GPS Spoofing Detection in Autonomous Vehicles

***Abstract*—*Index Terms*:-**

I. Introduction

Rapid advancement of autonomous vehicles (AVs) has transformed the transportation system by minimizing human intervention and enhancing safety and efficiency [1]. AVs rely on various sensors for localization, navigation, and communication. Among all sensors, the Global Positioning System (GPS) plays a crucial role in determining global positions and optimizing routes [2]. However, GPS is vulnerable to several attacks, such as jamming and spoofing [3]. In a spoof attack, an adversary transmits fake signals to manipulate the vehicle's Global Navigation Satellite System (GNSS), causing erroneous positioning and navigation decisions. Such attacks can have serious consequences, including route deviations, traffic interruptions, or even accidents.

Motivation:- We consider an AV is navigating through a smart transportation system, relying on GPS real-time positioning and route optimization. As it approaches a critical intersection, an attacker nearby transmits spoofed GPS signals, tricking the AV into believing it is on a different road. As a result, the vehicle makes an incorrect turn, leading to a traffic disruption or, in a worst-case scenario, a collision.

To solve the above-mentioned problem, existing literature has proposed several GPS spoofing attack detection mechanisms to enhance AV security. The existing work focuses only on the objective of detecting attacks with high accuracy [4], [5]. However, solutions often face challenges such as multicollinearity among features, which can negatively impact model performance. Furthermore, existing solutions lose valuable information during the feature selection process, leading to a high false alarm rate. Furthermore, previous works [5], [6] have employed unsupervised learning approaches to detect GPS spoofing attacks. Unfortunately, these unsupervised models are trained exclusively on historical GPS data, which often leads to misclassification by mistakenly identifying deviations from normal traffic behavior as anomalies. On the other hand, most existing works [7], [8], and [9] develop models using a specific dataset, limiting their generalizability across diverse data sets and their ability to adapt to the evolving patterns of GPS spoofing attacks. Moreover, some methods [][][] depend solely on sensors on board for both detection and mitigation, limiting their effectiveness against GPS spoofing attacks. In summary, GPS spoofing attack detection requires a more robust, adaptive approach that can generalize across multiple datasets while ensuring accurate attack detection on AVs.

We attempt to solve the following research gaps.

- How can robust feature selection techniques be developed to mitigate the impact of multicollinearity and information loss for detecting GPS spoofing attack?
- What novel methodologies can be developed to enhance the generalizability of GPS spoofing detection mechanisms across diverse datasets?

Contributions:- To fill the critical research gaps mentioned above, in this letter, we introduced a feature selection methodology based on autoencoders to effectively decouple correlated features and minimize information loss. In addition, a hybrid model has been proposed by integrating the recurrent neural network with the long-short-term memory (RNN-LSTM) and the random forest (RF) classifier. The hybrid model enhances generalizability across diverse data sets and adapts to evolving GPS spoofing patterns to detcet the attacks accurately.

In summary, the letter makes the following significant contributions:-

- **Autoencoder-driven feature selection:-** Unlike static feature selection, we employ an autoencoder for dynamic feature selection, that minimizes information loss by learning non-reduntant feature representation. This effectively mitigates multicollinearity and chooses a more accurate and reliable set of features for the detection of GPS spoof attacks on AVs.
- **Generalized hybrid model:-** We propose a generalized hybrid model for GPS spoofing detection, combining RNN-LSTM and RF classifier. The RNN-LSTM extracts temporal patterns from AV traffic, while the RF classifier performs accurate attack detection. This hybrid model reduces dataset-specific biases, enabling robust detection across diverse Datasets.
- **Evaluation:-**Evaluation results demonstrate a 99. 98% accuracy in detecting GPS spoofing attacks for AVs. Furthermore, The proposed model exhibits enhanced generalizability across diverse datasets and detects GPS spoofing attacks accurately. In particular, our proposed scheme also exhibits remarkable efficiency, with a throughput of Mbps in classifying AV traffic.

II. RELATED WORK

Recent researchers have been developing new techniques to enhance the AV security against GPS spoofing attacks. In this section, the literature review is classified as (1) Enhancing the Security of AV, and (2) GPS spoofing attack detection in AV. A detailed description of each classification is described below.

Enhancing the Security of AV:- The paper [10] develops a security-by-design framework for autonomous vehicles (AVs), incorporating mutual authentication, cryptographic protection, and secure communication. On the other hand, Biraja Prasad

Nayak et al. [11] analyze key security threats in AV communication, including impersonation, Sybil attacks, jamming, and GPS spoofing, while evaluating authentication techniques, cryptographic mechanisms, and intrusion detection systems for enhanced security. In addition, blockchain and AI-based automation improve access control and authentication in vehicle-to-thing (V2X) communications in [12]. Next, in [13], a security-based Vehicle Security Operations Center (VSOC) architecture with Management, Orchestration, Detection, and Response (MODR) components is proposed, integrating real-time monitoring, threat intelligence, and automated response. Furthermore, [14] systematically reviews the cybersecurity vulnerabilities of AV and evaluates security countermeasures, including cryptography, intrusion detection, authentication, and access control.

GPS spoofing attack detection in AV:- Ghilas Aissou et al. [7] explored various supervised machine learning algorithms, including KNN, Radius Neighbors, Linear SVM, C-SVM and Nu-SVM, for detection of GPS spoofing attacks on AVs. The paper collected real-time features and trained models using three types of spoofing attacks. Next, the paper [6] proposed an anomaly detection method that uses historical trajectories and a decision tree classifier for the detection of GPS spoofing. Similarly, Opt-attack, a deep learning-based detection method using long-short-term memory (LSTM), was introduced in [8] to bypass Kalman filter (KF)-based localization systems. In addition, [5] has developed GPS-IDS, an anomaly-based intrusion detection system that uses a vehicle behavior model based on physics to detect GPS spoofing attacks on AVs. Furthermore, Bhawana Poudel Devkota et al. [4] employed a Random Forest Multiclass Classifier (MRFC) to categorize GPS spoofing attacks into authentic signals, simplistic spoofing, intermediate spoofing, and sophisticated spoofing. Their approach incorporated Spearman's correlation and Gini index for feature selection and used SMOTE-ENN for data set balance. Finally, XGBoost was applied in [9] to develop a GPS spoofing attack detection model for AVs, further highlighting its effectiveness in detecting spoofing incidents.

Synthesis:- Existing literatures such as [4] and [5] use static feature selection methods, which introduce multicollinearity and degrade model performance. In contrast, we leverage an autoencoder for dynamic feature learning, effectively mitigating multicollinearity, and minimizing the information loss. Further, [6] and [5] employ unsupervised learning based solely on historical GPS data, often leading to misclassification of GPS spoofing attacks. In addition to this, [7], [8] and [9] develop models on specific Datasets, limiting their generalizability. Unlike these approaches, our work focuses on generalized model to enhance adaptibility across diverse Datasets.

III. Attack Methodology

This section introduces a methodology for generating GPS spoofing attacks on AVs. We consider an attack scenario where the adversary introduces a fixed offset to the true GPS coordinates. This causes the AV to perceive its location as consistently shifted from its actual position. The magnitude of the offset was varied to assess the detection sensitivity to different levels of positional error. Specifically, the spoofed position $(lat_{spoofed}, lon_{spoofed})$ at time t can be represented in Eq. 1 and 2.

$$lat_{spoofed(t)} = lat_{true} + \Delta lat, \tag{1}$$

$$lon_{spoofed(t)} = lon_{true} + \Delta lon, \tag{2}$$

where, lat_{true} and lon_{true} are the true latitude and longitude at time t, and Δlat and Δlon are the constant offsets. This can lead to

IV. Proposed System Framework

This section outlines the workflow of our proposed system, which develops a generalized hybrid model for GPS spoofing attack detection. The model integrates a Recurrent Neural Network-Long Short-Term Memory (RNN-LSTM) network with a Random Forest (RF) classifier, trained on the AV-GPS-Dataset and the UAV Attack Dataset. Initially, the datasets undergo preprocessing and relevant feature selection.

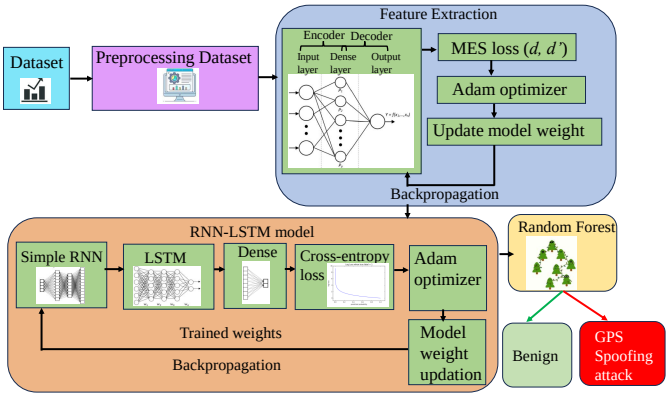

Fig. 1: Architecture

A. Feature selection

Initially, the AV-GPS-dataset containing 45 features was reduced to 34, while the UAV-Dataset's 84 features were narrowed to 14. To reduce feature dimensionality, we employed an autoencoder-based feature extraction methodology. Specifically, autoencoder comprises an encoder and a decoder. Basically, the encoder maps the high-dimensional input to a lower-dimensional latent space. In this work, the encoder consists of a fully connected layer with 64 neurons and a ReLU activation function, represented by Eq.3

$$h = f(W_e d + b_e), \tag{3}$$

Where, d is the set of input features, W_e is the encoder weight matrix, b_e is the bias vector, f is the ReLU activation function and h is the latent representation with 64 neurons. Subsequently, decoder reconstructs the original input from this latent representation, as shown in Eq. 4.

$$d' = g(W_d h + b_d), \tag{4}$$

Where, W_d is the decoder weight matrix and b_d is the bias vector. Further, the autoencoder minimizes the Mean Squared

Error (MSE) between the input *d* and the reconstructed output *d'*, as depicted in Eq. 5

$$L(d, d') = \frac{1}{n}\sum_{i=1}^{n}(d_i - d'_i), \tag{5}$$

Where, *n* is the numnber of features. Therefore, by training the autoencoder to minimize this error, the encoder effectively identifies non-redundant feature set, as depicted in Algorithm 1(lines 13-14) . After selecting relevant, non-redundant features, the generalized hybrid model is developed to identify GPS spoofing attacks on AVs. A detailed description of the model development is given in the following subsection.

Algorithm 1 Feature calculation

Input:-Dataset $d \in R^{n*m}$, *n* is the number of samples and *m* is the number of features

Output:-Reduced feature set $d' \in R^{n*k}, k < m$

1: Normalize d
2: **for** each feature $j \in d$, j=1 to m **do**
3: $d^{norm} = \frac{d-min(d)}{max(d)-min(d)}$
4: **end for**
5: define autoencoder with 64 neurons
6: apply ReLu activation function f(j)
7: **if** $j > 0$ **then** f(j)=j
8: **else**
9: f(j)=0
10: **end if**
11: define decoder
12: calculate mean square error *L (d,d')*
13: use adam optimizer to update autoencoder weight θ
14: $\theta = \theta - \eta * \nabla_\theta \zeta(d, d')$

B. Development of generalized hybrid model

The output of the autoencoder is divided into training (d'_{train}) and a testing set (d'_{test}), which serve as input to the sequential RNN-LSTM model. This model is designed for temporal pattern recognition in GPS spoofing attack detection. Specifically, the RNN comprises 32 units with a ReLU activation function. Additionally, LSTM has two layers, one with 64 unit and other with 128 units to capture long-term dependencies. Additionally, there is a dense layer that uses the softmax activation function for classification of GPS spoofing attacks, which is expressed as Eq. 6.

$$P(y = c|x) = \frac{e^{z_c}}{\sum_{i=1}^{c} e^{z_i}}, \tag{6}$$

where, where z_c is the output of the dense layer for class *c*, and *C* is the number of classes. This yields a probability distribution across the classes, enabling the identification of GPS spoofing attacks accurately. Further, the RNN-LSTM model is trained using categorical cross-entropy loss and the Adam optimizer, as shown in the equation. 7.

$$L_{Cross\_entropy} = -\sum_{i=1}^{n}\sum_{c=1}^{C} y_{ic} log(p_{ic}), \tag{7}$$

where y_{ic} is 1 if sample *i* belongs to class *c*, and 0 otherwise, and p_{ic} is the predicted probability of sample *i* belonging to class *c*. However, there is a chance of overfitting during model training. To mitigate overfitting, an early stop mechanism is implemented. Finally, the output of the last LSTM layer is fed into a Random Forest (RF) classifier for the identification of GPS spoofing attacks on AVs.

Following the RNN-LSTM model, an RF classifier is developed for the final classification of GPS spoof attacks and benign traffic. In particular, RF is an ensemble learning method that constructs an ensemble of decision trees. For each tree *t (t = 1, 2, ..., T)* within the forest, a bootstrap sample is generated by randomly drawing *N* samples with replacement from the training dataset d'_{train}. In addition, at each node of each tree, a random subset of k features is selected ($F_k \subset F$, where *F* is the total set of features). This random feature subset is then utilized to build the individual decision trees. The nodes within these trees are split recursively based on the Gini impurity index, as defined by Eq. 8.

$$Gini(p) = 1 - \sum_{i=1}^{c} p_i^2, \tag{8}$$

where p_i is the proportion of samples belonging to class *i* (either GPS spoofing attack or benign) within the node, and *C* is the number of classes. Upon reaching a leaf node in each tree, a vote is cast for the corresponding class. The final classification result for a given input is determined by the class that receives the majority of votes across all *T* trees in the forest. This ensemble voting mechanism improves the robustness and generalization capability of the classification between GPS spoofing attacks and benign traffic in AVs.

V. Implementation and Result Analysis

A. Performance of proposed solutions

This subsection presents the performance evaluation of our proposed hybrid model, which demonstrates its efficacy in detecting GPS spoofing attacks on AVs. To quantify the performance of the model, we calculated key metrics, including precision, precision, recall, and F1-Score, on two distinct datasets. The results of this evaluation are summarized in Table I. As indicated in Table I, our proposed hybrid model achieves a high level of accuracy, exceeding 99% on both datasets in detecting GPS spoofing attacks. Furthermore, the model exhibits strong performance across other evaluation metrics, resulting in robust F1 scores of 0.95 and 0.99 for the AV-GPS and UAV data sets, respectively.

TABLE I: Performance of proposed model

Dataset	Accuracy(%)	Precision	Recall	F1-Score
AV-GPS Dataset []	99.86	0.97	0.98	0.95
UAV Dataset []	99.98	0.99	0.99	0.99

B. performance of generalization of proposed solution

This subsection evaluates the generalizability of our proposed model across various datasets by comparing its performance with existing state-of-the-art mechanisms. To assess generalizability, we tested models developed in prior works

on datasets different from those they were originally trained on, and vice versa. Specifically, models from the GPS-IDS framework [5] were evaluated in the UAV data set, and models from the work [] were tested in the AV-GPS data set. Performance comparison is presented in Table II. The table demonstrates the superior generalization performance of our proposed model compared to existing approaches. Our proposed model achieves an accuracy of 99. 98% using the UAV Dataset, which significantly outperforms the models of the GPS-IDS framework [5]. Similar result has been obtained when tested with the AV-GPS dataset. These findings highlight the robustness and adaptability of our proposed generalized model across different datasets.

TABLE II: Performance of proposed model

Existing work	Dataset	Model	Accuracy(%)
GPS-IDS [5]	UAV Dataset	Random Forest	93.9
		Decision tree	93.45
		Logistic regression	82.3
		SVM	82
		CNN	84.52
		Proposed generalized model	99.98
et al. []	AV-GPS Dataset	MLP	88.69
		Gradient Boosting	89
		CNN	86.82
		ANN	86.21
		Proposed generalized model	99.86

C. Resource utilization

The measurement of CPU utilization plays a crucial role in understanding the performance of the proposed system. This experiment calculates CPU utilization for 60 seconds, as shown in Fig.2 (a). It is observed from the figure that the average CPU utilization of the hybrid model is greater than RF and RNN-LSTM. This is because the hybrid architecture involves sequential processing of data through the RNN-LSTM layers, followed by the Random Forest classifier for the final identification of GPS spoofing attacks. In addition, we evaluated the training time of RF, RNN-LSTM and the proposed hybrid model in a varying number of AV traffic, as shown in Fig. 2 (b). According to the figure, the proposed hybrid model incurs a higher training time compared to the RF and RNN-LSTM model. This is due to the multi-stage architecture, which involves sequential training of both components. Despite the additional time required, this hybrid model leads to a more accurate and generalized detection of GPS spoofing attacks on AVs.

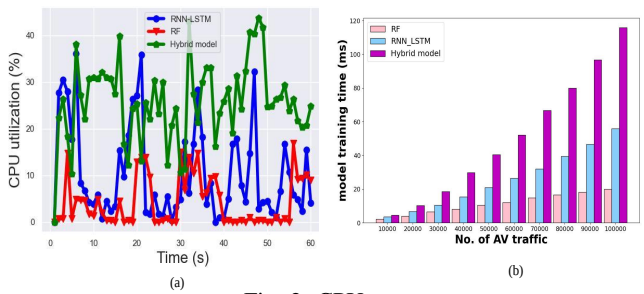

Fig. 2: CPU usage

D. Comparison with state-of-the-art solutions

In this subsection, Table III accurately compares the proposed solution with GPS-IDS [5], [6], [4] and . [9]. The proposed solution achieves an accuracy of 99.98%, which is higher than all the existing literature. Additionally, the solution detects attacks in 6.8 seconds, which is significantly faster than existing works.

TABLE III: Comparison table

Works	Accuracy(%)	Attack detection time (s)
GPS-IDS [5]	97.45	10
Yang et al. [6]	98	12.7
devkota et al. [4]	96.29	✗
Isleyen et al. [9]	97	78.2
Proposed work	99.98	6.8

VI. Conclusion and Future Works

This letter proposes a novel security mechanism to detect GPS spoofing attacks on AVs. The proposed solution leverages an autoencoder-driven feature selection method to mitigate the impact of multicollinearity and information loss, which is enhancing the robustness of feature selection. Further, a generalized hybrid model, integrating RNN-LSTM and RF, is introduced to improve the generalizability of GPS spoofing detection across diverse datasets and adapt to evolving attack patterns. The evaluation results demonstrate the effectiveness of this approach, achieving a high accuracy of 99.98% in detecting GPS spoofing attacks. The proposed model exhibits enhanced generalizability, representing a significant advancement in securing autonomous vehicles against GPS spoofing. In the future, the solution can be extended to detect wide range of GPS spoofing attacks dynamically.

References

[1] M. Kamal, A. Barua, C. Vitale, C. Laoudias, and G. Ellinas, "Gps location spoofing attack detection for enhancing the security of autonomous vehicles," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*. IEEE, 2021, pp. 1–7.

[2] S. Dasgupta, M. Rahman, M. Islam, and M. Chowdhury, "A sensor fusion-based gnss spoofing attack detection framework for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 23 559–23 572, 2022.

[3] S. Filippou, A. Achilleos, S. Zukhraf, C. Laoudias, K. Malialis, M. K. Michael, and G. Ellinas, "A machine learning approach for detecting gps location spoofing attacks in autonomous vehicles," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*. IEEE, 2023, pp. 1–7.

[4] B. P. Devkota, L. Saunders, R. Dhakal, and L. N. Kandel, "Gps spoofing detection with a random forest multiclass classifier," in *MILCOM 2024-2024 IEEE Military Communications Conference (MILCOM)*. IEEE, 2024, pp. 202–208.

[5] M. Mehrab Abrar, R. Islam, S. Satam, S. Shao, S. Hariri, and P. Satam, "Gps-ids: An anomaly-based gps spoofing attack detection framework for autonomous vehicles," *arXiv e-prints*, pp. arXiv–2405, 2024.

[6] Z. Yang, J. Ying, J. Shen, Y. Feng, Q. A. Chen, Z. M. Mao, and H. X. Liu, "Anomaly detection against gps spoofing attacks on connected and autonomous vehicles using learning from demonstration," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 9462–9475, 2023.

[7] G. Aissou, S. Benouadah, H. El Alami, and N. Kaabouch, "Instance-based supervised machine learning models for detecting gps spoofing attacks on uas," in *2022 IEEE 12th annual computing and communication workshop and conference (CCWC)*. IEEE, 2022, pp. 0208–0214.

[8] Q. Chen, G. Li, P. Liu, and Z. Wang, "Anomaly detection and secure position estimation against gps spoofing attack—a security-critical study of localization in autonomous driving," *IEEE Transactions on Vehicular Technology*, 2024.

[9] E. İşleyen and Ş. Bahtiyar, "Gps spoofing detection on autonomous vehicles with xgboost," in *2024 9th International Conference on Computer Science and Engineering (UBMK)*. IEEE, 2024, pp. 500–505.

[10] A. Chattopadhyay, K.-Y. Lam, and Y. Tavva, "Autonomous vehicle: Security by design," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7015–7029, 2020.

[11] B. P. Nayak, L. Hota, A. Kumar, A. K. Turuk, and P. H. Chong, "Autonomous vehicles: Resource allocation, security, and data privacy," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 117–131, 2021.

[12] G. Bendiab, A. Hameurlaine, G. Germanos, N. Kolokotronis, and S. Shiaeles, "Autonomous vehicles security: Challenges and solutions using blockchain and artificial intelligence," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 3614–3637, 2023.

[13] J. Han, Z. Ju, X. Chen, M. Yang, H. Zhang, and R. Huai, "Secure operations of connected and autonomous vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 11, pp. 4484–4497, 2023.

[14] B. A. Tanaji and S. Roychowdhury, "A survey of cybersecurity challenges and mitigation techniques for connected and autonomous vehicles," *IEEE Transactions on Intelligent Vehicles*, 2024.