

On Cross-Testing Datasets for RF-Fingerprinting based Deep-Learning GNSS Spoofing Detection

Leatile Marata*, Mohammad Zahidul H. Bhuiyan[†], Elena Simona Lohan*

* Tampere Wireless Research Centre, Electrical Engineering Unit, Tampere University, Finland

{leatile.marata,elena-simona.lohan}@tuni.fi

[†] Finnish Geospatial Research Institute, National Land Survey of Finland, zahidul.bhuiyan@nls.fi

Abstract—Global Navigation Satellite Systems (GNSS) are crucial for acquiring position, velocity, and time (PVT) information for various applications such as air navigation, precision agriculture, logistics, and emergency services. GNSS is a backbone technology that supports many crucial aspects of modern-day life and industry. Unfortunately, GNSS is currently facing significant security risks due to a surge in signal interference from jammers and spoofers (i.e., fake GNSS signals), which can render it unreliable. To address this ever-growing challenge, novel GNSS receiver algorithms are currently being developed using machine learning (ML). These are fully reliant on available GNSS spoofing repositories such as TEXBAT and OAKBAT. However, in the current state-of-the-art, there are still inconsistencies on how models are trained using a single dataset and cross-tested with others. To address this issue, our work offers an overview of the impact of cross-testing with different datasets and demonstrates that results vary significantly depending on the dataset used for training and that with proper training datasets, cross-testing accuracy above 90% can be achieved.

Index Terms—Convolution Neural Networks (CNN), Deep learning, Radio Frequency Fingerprinting (RFF), OAKBAT and TEXBAT datasets.

I. INTRODUCTION

Global Navigation Satellite Systems (GNSS) play a key role in enabling modern day automation technologies such as autonomous driving and automated robotics. For instance, unmanned aerial vehicles (UAVs) and self-driving cars rely on GNSS sensors for precise positioning, velocity, and timing (PVT) parameters [1]. There has been a rapid technological advancement of GNSS in the past decade, which has included the modernization of the conventional Medium Earth Orbiting (MEO)-based GNSS satellites. Similarly, there are some efforts aimed at provision of GNSS via Low Earth Orbit (LEO) satellites. However, such advancements in satellite-based positioning come with an increase in threats from interference signals. This may be due to the fact that modern technology, such as software-defined radio (SDR), has enabled users to generate interference at low costs. Among the various sources of interference, there is the spoofing category, which is intended to deceive the receiver by mimicking the exact signal meant for the receiver [2]. The spoofer interference is a major security concern since it can alter sensitive PVT parameters and can render GNSS unreliable for many applications, including the safety-critical ones. For instance, Fig. 1 shows a scenario where a spoofer intends to alter the PVT of a delivery

robot. Therefore, an increase in spoofing incidents can render the vision of autonomous systems impractical, as it can distort crucial data.

There has been a significant amount of research studies aimed at detecting spoofing signals in GNSS. Among these, there are the classical approaches, which involve the use of Signal Quality Monitoring (SQM) techniques that rely on the statistical signal processing frameworks [3]. However, one major drawback of these approaches is that they are highly reliant on mathematical models [1]. Unfortunately, spoofers can appear in different forms, from simple to very complex ones. Mathematically-derived solution are in general based on fixed assumptions, with less flexibility than Machine-Learning (ML)-based approaches and therefore they cannot fully capture different forms of spoofers. Of recent, new approaches such as [2], that leverage ML and artificial intelligence (AI) have been on the rise due to their ability to learn complex forms of signal interference by using data that has been previously collected from GNSS signals.

As mentioned earlier, AI-inspired receiver algorithms can exploit available datasets from which they learn to differentiate between a clean signal and a spoofed signal. For instance, this can be done by relying on Radio Frequency Fingerprinting (RFF), which is of particular interest in our paper. RFF is a technique that exploits the unique hardware imperfections of various electronic components of the transmitters, which differentiate them from one another [4]. Some of these imperfections include variations caused by the phase noise, inphase/quadrature (I/Q) imbalance, and power amplifier distortions. Therefore, the accuracy of AI-inspired receiver solutions is directly tied to the amount of data that can be leveraged for training the AI models. Though still deficient, there are some common repositories with different spoofer types that have been publicly available such as Texas Spoofing Test Battery (TEXBAT) and Oak Ridge Spoofing and Interference Test Battery (OAKBAT). However, there is a general lack of information on developing models that can be fully generalized for anti spoofing. In an endeavour to synthesize research, this paper studies different datasets from TEXBAT and OAKBAT, with the aim of uncovering the non-triviality of choosing the datasets to work with, particularly for the RFF techniques.

To overcome the growing security concerns in GNSS re-

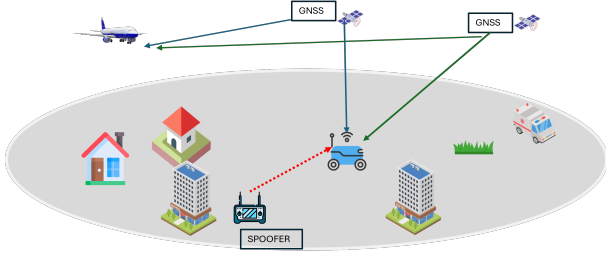


Figure 1: A spoofer attack scenario in an autonomous driving vehicle.

ceiver, there have been huge efforts aimed at developing spoofer detection solutions. For example, Mao *et al.* presented a SQM-based on long-short-term memory (LSTM) in [5] using TEXBAT datasets ds2, ds3, ds4, ds5, ds6 and reported a detection capability of more than 98.5%. Also, Iqbal *et al.* [6] proposed a two-stage artificial neural network (ANN) that relied on the multi correlator finger outputs and used the variational auto-encoders to enhance detection in unseen datasets. Results in [6] revealed a detection accuracy of about 99% in all the datasets in TEXBAT, except the ds7, which is reported to be detected on average with a rate of 50.68%. In a similar vein, Sun *et al.* [7] proposed spatial temporal variational autoencoders, with spoofer detection capability results that are around 99%. Feng *et al.* in [8] proposed a second-order differential pseudorange method for detecting spoofers for distributed internet of things that is supported by GNSS for time synchronization. Results presented using Receiver Operating Curves (ROC) from this work revealed that their work made a 10% improvement to existing methods used on TEXBAT datasets ds2 and ds7. Deng *et al.* in [9] proposed an SQM method that used the energy information and achieved a detection capability of more than 99% using ds2 and ds3. In [10], Wang *et al.* proposed a multi-domain joint spoofing detection method based on semi-supervised learning, specifically for TEXBAT ds7, which is a static spoofing complex scenario. Results of their work showed a performance that is around 99%, while Zhou *et al.* in [11] proposed a spoofing detection method using the auto-correlation function. Results of this work revealed a 90% detection capability in the TEXBAT datasets. To further substantiate this state of the art, we present Table I, that summarizes results and approaches used when using TEXBAT and OAKBAT datasets. To the best of our knowledge, the literature does not contain any works specifically focusing on RF fingerprinting using spectrogram feature extraction while also investigating the impact of cross-testing these datasets.

As paper contributions, we study the impact of cross-testing with different datasets from TEXBAT and OAKBAT in order to find out the generalization capability of RFF ML algorithms for GNSS spoofer detection. We classify the signal based on spectrograms of the I/Q data using ML. The main novelty of the paper comes from studying the cross-testing capability with raw I/Q data and ML. We show that not all spoofing types have good capability of generalization and that, the best dataset for training, among the studied ones, is TEXBAT ds8,

which is based on a sophisticated spoofing scenario.

II. METHODOLOGY

GNSS signals are generally identified by their unique parameters such as Pseudo-Random Noise (PRN) codes. Recently, RFF techniques have been employed to distinguish GNSS signal sources based on hardware-induced distortions [4]. Similar to human fingerprints, RFF arise from unique hardware impairments in GNSS transmitters and they are specific to each device. These impairments result from variations in manufacturing processes, leading to measurable differences in transmitted signals. This is due to the fact that even though GNSS transmitters are designed to identical specifications, slight variations in components such as oscillators, power amplifiers, and filters introduce unique distortions, which can be exploited via ML, and thus overcoming the limitation of mathematical-based solutions. A clean GNSS signal from a GNSS satellite l in a set of L of GNSS satellites can, in the absence of interference and on a single carrier be modeled as

$$s_l(t) = \sqrt{P_t} \sum_n b_n x_n(t), \quad (1)$$

where P_t is the transmission power of the GNSS signal, while b_n and $x(t)$ are the navigation data bits and the modulated pseudorandom waveforms for the n -th bit, respectively. The received signal would ideally, in the absence of spoofers have a form $r(t) = \sum_{l=1}^L s_l(t) + w(t)$, where L is the number of satellites, $w(t)$ denotes the Gaussian noise component, with zero mean and variance σ^2 , i.e., $w(t) \sim \mathcal{N}(0, \sigma^2)$. Unfortunately, due to the imperfections of the electronic components, the transmitter introduces the RFF, which alters the structure of the signal. Assume that the RFF of a GNSS signal is denoted by Φ_l , $l \in \mathcal{L}$, such that each non-ideal GNSS signal will be received as a function of the ideal signal and the RFF. Similarly, assume that there is a set $\mathcal{M} = \{1, \dots, M\}$ of spoofers, where each of the spoofers is affected by the RFF of its own transmitting device, that we denote by Ω_m . In this case, we can define a function h_i , $i \in \{l, M\}$ that models the RFF imperfection. Then, the non-ideal received signal in the presence of spoofers and assuming single-path channels is

$$\begin{aligned} \tilde{r}(t) = & w(t) + \sum_{l=1}^L \beta_l h_l(s_l(t - \tau_l), \Phi_l) \exp(j2\pi f_l t) \\ & + \sum_{m=1}^M \alpha_m \beta_m h_m(v_m(t - \tau_m), \Omega_m) \exp(j2\pi f_m t), \end{aligned} \quad (2)$$

where f and τ , $i \in \{l, m\}$ are the Doppler spread and the path delays, respectively, while $\alpha_m \in \{0, 1\}$ is a binary function that denotes the presence (1) or absence (0) of a spoofer m , and β_{dot} is the path attenuation. For robust GNSS, receivers should be able to identify the GNSS satellites by their PRNs, and also detect if there is a spoofing signal present. From (2), this can be done by estimating α_m . Unfortunately, such a solution is non-trivial as it relies on non-convex optimization and cannot be fully captured by mathematically-inspired solutions

Table I: Comparison of different methods used for TEXTBAT and OAKBAT datasets.

Reference	Method	Datasets	Metric	Accuracy	Cross Testing
[12]	k-nearest neighbors algorithm (k-NN) in post-correlation data	TEXTBAT (ds3 and ds8)	Detection accuracy per class	94.5%	Not reported
[13]	Spoofing correlation peak cancellation (SCPC)	TEXTBAT (ds3)	Not specified	98%	Not reported
[14]	Enhanced SQM using Kolmogorov-Smirnov test	TEXTBAT (ds2, ds3, and ds7)	Detection accuracy per class	99.76% for 10% false alarm	No cross-testing
[2]	Autoencoders and GAN using multiple stages of the receiver	TEXTBAT (ds2-ds7)	Detection accuracy per class	More than 95%	Not reported, but some generalization via autoencoders
[15]	SVM on multiple postcorelation using PCA for feature decorrelation	TEXTBAT (ds2, ds3, ds7) and OAKBAT (os2, os3)	F1 score	93.97% on ds2, ds3, ds7 and 97% on os2 and os3	Not reported
[16]	Two-stage statistical detection based on CNR	TEXTBAT (ds1-ds6)	Only visual results	No accuracy numbers given	Not reported
[17]	Statistical detection metric based on CNR	TEXTBAT (ds2 and ds3)	Detection accuracy per class	90.13%	Not reported
[18]	AdaBoost with multiple features	TEXTBAT (ds2, ds3, ds7) and OAKBAT (os2, os3)	Average detection	97.04% for ds2, ds3 and ds7, 98.08% for os2 and os3	Not reported
[3]	Statistical method based on quadrature channel energy	TEXTBAT (ds2, ds3)	Detection accuracy per satellite	93.33% for ds2 and 81.57% for ds3	Not reported
[19]	Spiking NN applied to CAF feature	TEXTBAT (ds2-ds8)	Detection accuracy per satellite	more than 98%	Not reported
Our work	DNN applied on I/Q-based spectrogram	TEXTBAT (ds2-ds4, ds8) and OAKBAT (os2)	Detection accuracy per class	between 90% and 99.9% with adequate cross-training	Cross testing included

due to the fact that spoofers can take many different forms. Moreover, the exact RF features infused by both Φ_l and Ω_m vary significantly, which makes it difficult to directly model the RFF. We focus on ML-based solutions that can learn the RFF based in raw I/Q data. To extract the features, we use the Space-Time Frequency Transform (STFT) (i.e., spectrogram), due to its ability to reveal signal variations. The datasets from OAKBAT and TEXTBAT are quantized at 16 complex bits (i.e., 8 bit per I and 8 bit per Q channel) in TEXTBAT data 32 complex bits in OAKBAT data, with sampling frequencies of 25MHz (TEXTBAT) and 5MHz (OAKBAT). When sampled at such frequencies they generate very large amounts of data can be challenging to process. We use a 4ms data window for all the datasets to process smaller chunks of the data. For both the clean signal (between 0 and 110 seconds in TEXTBAT) and the spoofed signal (above 110 seconds in TEXTBAT), we only read 60 seconds of data to have a fair comparison. In total, this yields 15000 spectrograms each for clean and spoofed data aggregating a total of 30000 samples for TEXTBAT. When developing a model with a particular dataset, 80% data samples are selected for training, while the rest are used for testing. Since the main idea of this work is to study the impact of different datasets on generalization of ML on GNSS spoofer detection, we do not develop ML algorithms from scratch but we used existing ones, namely CNN, k-NN, Random Forest and Support Vector Machines (SVM). We first compared the performance of different algorithms, configured using the parameters listed in Table II on the training data, and the results are presented in Fig. 2. It can be noted that all the algorithms perform well and have a comparative performance,

with an accuracy of more than 95%, while CNN slightly outperforms the other three considered methods. Therefore, without any loss of generality, we will use only CNN for the results present in the sequel. The CNN is configured with a 2D convolutional layer with 32 filters, followed by a 2D max pooling layer. Then, another 2D convolutional layer with 64 filters is applied. Finally, these layers are followed by another 2D max pooling layer, a flattening layer, and a dense layer, with other parameters reflected in Table II.

Algorithm	Parameter
Random Forest	No. of Estimators: 100
SVM	kernel: Radial Basis Function
k-NN	No. of neighbors: 5
CNN	Optimizer: Adam Learning rate: 0.001 Loss: Categorical Crossentropy Epochs: 10, Batch size: 32

Table II: Machine-Learning Model Parameters

III. CNN-BASED RESULTS AND DISCUSSIONS

In this section, we present a comparative study of the impact of the cross-testing effects on spoofer detection models. We have selected the datasets ds2, ds3, ds4, and ds8 from TEXTBAT, and os2 from OKABAT, primarily due to their divergent characteristics (from simplistic to sophisticated spoofers), which allow a comprehensive analysis. We present the results using the confusion matrix and average two-class detection accuracy, as we discuss next.

In Fig. 3, we present the performance of the CNN in terms of classification accuracy for the training dataset ds8. In general, the results show that the RF magnitude spectrogram consistently achieves an accuracy above 90% across all datasets

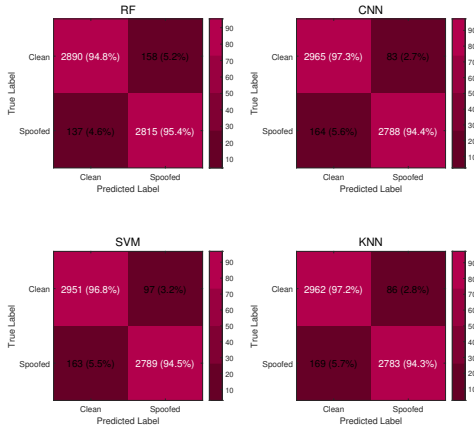


Figure 2: Confusion matrices for different data matrices using magnitude spectrogram and four ML approaches, for Texbat ds2.

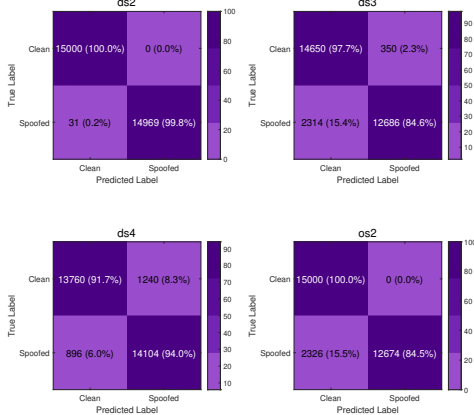


Figure 3: Confusion matrices for different datasets using magnitude spectrograms training with TEXBAT ds8.

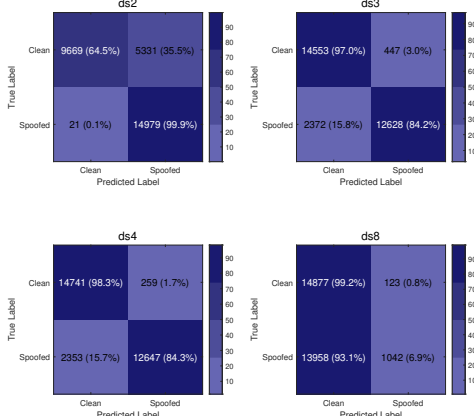


Figure 4: Confusion matrices for different datasets using magnitude spectrograms training with TEXBAT os2.

(ds2, ds3, ds4, and os2) when the CNN model is trained on ds8. This indicates that the CNN model generalizes well when trained on ds8 dataset. One of the reasons for this could be that ds8 involves zero-delay spoofing scenarios, which are complex GNSS spoofing cases that may reveal crucial features for detecting various forms of spoofing [20]. Furthermore, these results are consistent with the results reported in [6] for ds7, which is also a complex spoofing scenario, similar to ds8.

In Figs. 4, 5, and 6, it can be seen that the CNN model generalized well for cross-training with most of the spoofing

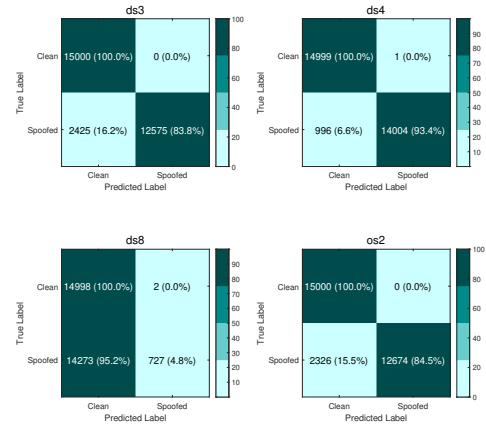


Figure 5: Confusion matrices for different datasets using magnitude spectrograms training with TEXBAT ds2.

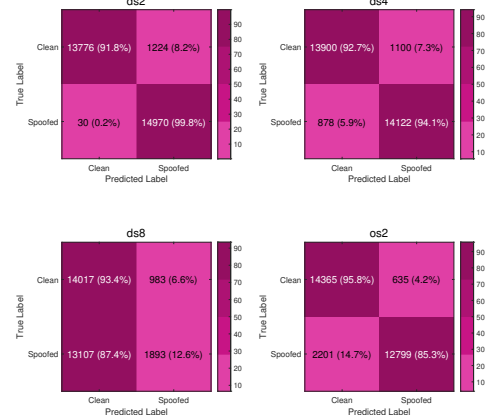


Figure 6: Confusion matrices for different datasets using magnitude spectrograms training with TEXBAT ds3.

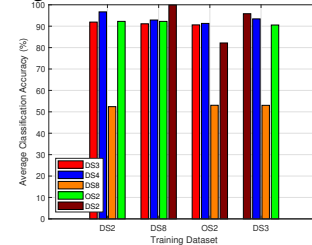


Figure 7: Accuracy as a function of the datasets during training.

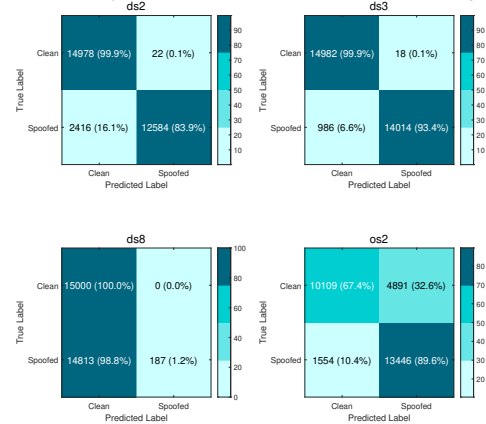


Figure 8: Confusion matrices for different datasets using magnitude and phase spectrograms training with TEXBAT ds2.

scenarios except when cross-testing with ds8 (in particular, training with ds2 or ds3 and testing with ds8 dataset gives very poor results). This is consistent with the fact that scenarios such as os2, ds4, and ds3 are much simpler than the ds8 spoofing scenario, which renders the model trained on simplistic datasets insufficiently developed for sophisticated datasets (however, training on a sophisticated dataset such as ds8 does give excellent testing results also for simplistic datasets). In Fig. 7, we provide the average accuracy in cross-testing, where it can be observed that the feasibility of cross-testing on GNSS based on ML models is highly reliant on the data used. The x-label in Fig. 7 shows what dataset was used in training and the legend shows what dataset was used in testing; clearly, when training and testing with the same dataset, the accuracy is the best; however, training with ds8 and testing with the other datasets also provides an accuracy higher than 90%.

We have also analyzed the impact of CNN models that rely on both the magnitude and phase spectrograms. The results reported in Fig. 8 reveal that the combination of both magnitude and phase spectrograms adds very little information compared to the magnitude-only information. In addition, this combination can essentially decrease generalization, as seen with the results corresponding to os2, which decreased by 10% from the results of the magnitude-only spectrogram (Fig. 5).

IV. CONCLUSIONS AND FUTURE DIRECTIONS

In this work, we analyzed the cross-testing effects of datasets from the TEXBAT and OAKBAT spoofing repositories on ML-based spoofing detection algorithms. This was done using I/Q raw pre-correlation data, specifically GNSS signal spectrograms. We observed that the performance of cross-testing heavily relies on the dataset used during training. For instance, training on complex spoofing scenarios (such as ds8) results in better generalization capability to other scenarios than training ML models on simpler scenarios (such as ds2). Furthermore, we have shown that the magnitude-plus-phase information from I/Q-based spectrograms carries almost the same information (in terms of RFF) as the magnitude-only spectrogram, and thus there is no advantage in using the phase of the spectrogram as a feature in the ML algorithm.

Although this work provides some insights into the cross-testing capabilities with raw I/Q data in spoofing detection, it does not consider an optimized approach for enhancing the ability of training methods to generalize across different datasets. Future research directions include training with additional features from both the pre-correlation and post-correlation domains, such as carrier phase, DLL discriminator, multi-correlation output, or carrier-to-noise ratios, and focusing also on multi-class detection (i.e., determining whether each PRN comes from a spoofer or a clean signal instead of two-class spoofing-vs-clean signal detection). Additionally, future work could involve training on multiple datasets, such as the FGI-Spoofing Repository (FGI-SpoofRepo) [21] and TUNI datasets [22], where possible generalization mechanisms could include training across multiple datasets, as well as leveraging transfer and meta-learning frameworks to enhance adaptability.

ACKNOWLEDGMENT

This work was supported by the Research Council of Finland number 359846 (RESILIENT project).

REFERENCES

- [1] M. L. Psiaki and T. E. Humphreys, "GNSS spoofing and detection," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1258–1270, 2016.
- [2] A. Iqbal, M. N. Aman, and B. Sikdar, "A Deep Learning Based Induced GNSS Spoof Detection Framework," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 457–478, 2024.
- [3] C. Sun *et al.*, "Robust Spoofing Detection for GNSS Instrumentation Using Q-Channel Signal Quality Monitoring Metric," *IEEE Trans. on Instrumentation and Measurement*, vol. 70, pp. 1–15, 2021.
- [4] N. Soltanieh *et al.*, "A review of radio frequency fingerprinting techniques," *IEEE Journal of RFID*, vol. 4, no. 3, pp. 222–233, 2020.
- [5] W. Mao, J. Ren, and S. Ni, "Fast GNSS spoofing detection based on LSTM-detect model," *GPS Solutions*, vol. 29, no. 1, pp. 1–17, 2025.
- [6] A. Iqbal, M. N. Aman, and B. Sikdar, "Machine and Representation Learning-Based GNSS Spoofing Detectors Utilizing Feature Set From Generic GNSS Receivers," *IEEE Trans. on Consumer Electronics*, vol. 70, no. 1, pp. 574–583, 2024.
- [7] P. Sun *et al.*, "Multi-Parameter Joint GNSS Spoofing Detection Based on TSVAE," *IEEE Trans. on Aerospace and Electr. Sys.*, pp. 1–15, 2024.
- [8] F. Feng *et al.*, "A GNSS time synchronization attack detection method for commercial off-the-shelf receivers: Cumulative second-order difference of pseudoranges," *IEEE Internet of Things Journ.*, vol. 12, no. 3, pp. 2322–2333, 2025.
- [9] M. Deng *et al.*, "GNSS Continuous Spoofing Detection Technology Based on Q-Channel Composite Information," *IEEE Sensors Journal*, vol. 24, no. 23, pp. 39 540–39 551, 2024.
- [10] S. Wang *et al.*, "Multi-Domain Joint Spoofing Detection Based on a Semi-Supervised Detection Network for GNSS-Based Train Positioning," *IEEE Trans. on Aerospace and Electronic Systems*, pp. 1–15, 2024.
- [11] W. Zhou *et al.*, "Anti-spoofing technique based on vector tracking loop," *IEEE Trans. on Instrument. and Meas.*, vol. 72, pp. 1–16, 2023.
- [12] A. Mahroof *et al.*, "Machine Learning-Based Detection of Spoofing Attacks in GNSS: A Study Using TEXBAT Dataset," in *14th Int. Conf. on Electrical Eng.*, 2024, pp. 90–95.
- [13] B. Yang *et al.*, "Research on GNSS Spoofing Mitigation Technology Based on Spoofing Correlation Peak Cancellation," *IEEE Communications Letters*, vol. 26, no. 12, pp. 3024–3028, 2022.
- [14] W. Zhou *et al.*, "Detection of Spoofing Attacks on Global Navigation Satellite Systems Using Kolmogorov–Smirnov Test-Based Signal Quality Monitoring Method," *IEEE Sensors Journal*, vol. 24, no. 7, pp. 10 474–10 490, 2024.
- [15] Z. Chen *et al.*, "GNSS Multiparameter Spoofing Detection Method Based on Support Vector Machine," *IEEE Sensors Journal*, vol. 22, no. 18, pp. 17 864–17 874, 2022.
- [16] M. Ali, F. T. Zahra, and A. A. Raja, "Robust spoofing detection in GNSS-SDR systems: A two-stage method for real-time signal integrity," in *3rd Int. Conf. on Emerging Trends in Electrical, Control, and Telecomm. Eng.*, 2024, pp. 1–6.
- [17] M. Deng *et al.*, "GNSS Spoofing Detection Based on Abnormal Receiver Noise and Carrier-to-Noise Ratio Metric," in *IEEE 22nd Int. Conf. on Comm. Technology*, 2022, pp. 1306–1311.
- [18] J. Li *et al.*, "The GNSS Spoofing Detection Method Based on AdaBoost," in *6th Int. Symposium on Autonomous Systems*, 2023, pp. 1–6.
- [19] S. Wang *et al.*, "Post-correlation Identification of GNSS Spoofing based on Spiking Neural Network," in *2024 IEEE Intelligent Vehicles Symposium (IV)*, 2024, pp. 3248–3254.
- [20] A. Lemmenes, P. Corbell, and S. Gunawardena, "Detailed analysis of the TEXBAT datasets using a high fidelity software GPS receiver," in *Proceedings of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation*, 2016, pp. 3027–3032.
- [21] S. Islam *et al.*, "An open GNSS spoofing data repository: characterization and impact analysis with FGI-GSRx open-source software-defined receiver," *GPS Solutions*, vol. 28, no. 176, 2024.
- [22] W. Wang, J. Sakari, and E. S. Lohan, "Raw I/Q measurement data and software for GNSS RFF fingerprinting," Zenodo open-access repository, <https://zenodo.org/records/13846381>, accessed 29.3.2025.