Borhani-Darian *et al.*
EURASIP Journal on Advances in Signal Processing (2024) 2024:14
https://doi.org/10.1186/s13634-023-01103-1

EURASIP Journal on Advances in Signal Processing

RESEARCH Open Access

Detecting GNSS spoofing using deep learning

Parisa Borhani-Darian[1*], Haoqing Li[1], Peng Wu[1] and Pau Closas[1]

*Correspondence: borhanidarian.p@northeastern.edu

[1] Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Avenue, Boston, USA

Abstract

Global Navigation Satellite System (GNSS) is pervasively used in position, navigation, and timing (PNT) applications. As a consequence, important assets have become vulnerable to intentional attacks on GNSS, where of particular relevance is spoofing transmissions that aim at superseding legitimate signals with forged ones in order to control a receiver's PNT computations. Detecting such attacks is therefore crucial, and this article proposes to employ an algorithm based on deep learning to achieve the task. A data-driven classifier is considered that has two components: a deep learning model that leverages parallelization to reduce its computational complexity and a clustering algorithm that estimates the number and parameters of the spoofing signals. Based on the experimental results, it can be concluded that the proposed scheme exhibits superior performance compared to the existing solutions, especially under moderate-to-high signal-to-noise ratios.

Keywords: GNSS acquisition, Spoofing detection, Machine learning, Deep learning

1 Introduction

For the successful operation and implementation of modern applications such as Intelligent Transportation Systems and location-based services, a continuous and precise source of navigation, positioning, and timing information is essential. Global Navigation Satellite Systems (GNSS) serve as the primary source of such information, forming the backbone of Positioning, Navigation, and Timing (PNT) data[1–3], when it is available [4–6]. The susceptibility of GNSS receivers to intentional interference renders them highly sensitive and vulnerable. This vulnerability creates opportunities for malicious actors seeking to compromise GNSS-based systems or infrastructure, potentially leading to severe consequences. The absence of built-in security features in GNSS systems leaves numerous applications exposed to potential risks, as has been documented in multiple articles [7, 8]. Deliberate attacks on GNSS receivers can be classified into two categories: physical attacks on the receiver (non-signal attacks) or attacks at the GNSS signal-in-space (SIS) level (signal attacks). Physical attacks on the receiver involve physical tampering or manipulation, while signal attacks target the GNSS signals transmitted by the satellites and can cause disruption or degradation of the receiver's ability to accurately determine position, velocity, and timing [9]. The focus of this paper is on

intentional attacks aimed at GNSS signals, which can manifest in three distinct forms: jamming, meaconing, and spoofing. This work concentrates on spoofing, a technique where fabricated GNSS-like signals are transmitted to manipulate the position output of the victim's receiver without causing any disruption to GNSS operations, essentially giving the attacker control over the receiver. It is worth noting that jamming attacks have the objective of obstructing GNSS positioning services, whereas spoofing interference aims to deceive the receiver into providing incorrect position information. These goals are diametrically opposed. The goal of this article is to use the so-called Cross Ambiguity Function (CAF), computed by GNSS receivers, to detect spoofing attacks from the I &Q samples. GNSS receivers [10–13] implement a statistical hypothesis test, in which the receiver determines the presence or absence of a signal from a specific satellite in the received signal, while also providing a basic estimation of the delay and Doppler frequency when the signal is detected. To carry out this test, it is common practice to maximize the Cross Ambiguity Function (CAF) between the received signal and a local code replica [14]. The performance of the acquisition method in a satellite communication system is typically characterized by three probabilities: detection probability, false-alarm probability, and miss-detection probability. The detection probability, also known as the probability of detection, is the probability that the acquisition method will correctly detect the presence of a signal when a signal is present. It is important to have a high detection probability to ensure that the system can reliably detect and acquire the satellite signal. The false-alarm probability, also known as the probability of false detection or the probability of false alarm, is the probability that the acquisition method will incorrectly detect the presence of a signal when there is no signal present. A high false alarm probability can lead to unnecessary processing and waste of system resources. The miss-detection probability, also known as the probability of miss or probability of false negative, is the probability that the acquisition method will mistakenly decide on the null hypothesis and not detect a signal when a signal is present. A high miss-detection probability can lead to a failure to acquire the signal, which can result in a loss of data or communication. The Receiver Operating Characteristics (ROC) plot is an assessment method of the probability of detection against the probability of a false alarm [15, 16]. The methodology of maximizing the CAF is considered optimal in certain model conditions (e.g., channel Gaussianity) for signal acquisition, which is founded on reliable statistical principles. For instance, under the presence of spoofing attacks where the CAF is distorted by the appearance of additional maxima for each spoofing signal present [17].

With the increase in computing capabilities of GNSS receivers [18], the use of deep learning has grown in that field [19], [20] presented a deep learning-based beamforming technique that was introduced to counteract multipath, revealing the drawbacks of traditional beam-forming algorithms. Applying the deep neural network (DNN) model in various contexts can result in a reduction of root-mean-squared error (RMSE). [21] discussed that GNSS is a satellite-based system that allows users to determine their precise location, velocity, and time information. However, in urban areas, GNSS measurements can be affected by various factors, including multipath interference and signal attenuation due to obstructions like buildings and trees. To address this issue, the study used a DNN to extract important features from the data and then used this information to learn GNSS measurement quality. The results

showed that by learning the GNSS measurement quality using the DNN, the study was able to make more accurate predictions of satellite visibility and pseudo-range errors in urban areas. This could have important applications in fields like transportation, where accurate GNSS measurements are essential for navigation and tracking.

The work in [22, 23] proposed a new approach to replace the traditional local code correlation-based CAF calculation with a DNN technique. This new method was designed to better understand the complexities of the multipath channel, which can affect GNSS signal quality in areas with obstructions or signal reflections. By using a DNN, the researchers aimed to improve the accuracy of the CAF calculation and enhance the overall performance of the GNSS system. The study displayed encouraging outcomes when the DNN approach was implemented in traditional tracking loops. Those, and other works [24, 25] highlight the GNSS multipath mitigation issue as a significant challenge in ensuring the accuracy and reliability of GNSS measurements. Over the years, numerous methods and techniques have been proposed to mitigate the effects of multipaths, including using specialized antennas, signal processing algorithms, and machine learning approaches like DNNs. These methods have gained popularity due to their effectiveness in improving GNSS performance in areas with obstructions or signal reflections[26, 27]. On another set of GNSS applications, the effect that using deep learning techniques has on improving the accuracy and effectiveness of GNSS spoofing detection [28–32] and jamming [33, 34] attacks is presented in several works. The researchers used a DNN to analyze GNSS data and identify patterns that could indicate an earthquake [35], hurricane monitoring [36], ice detection [37], and ionospheric scintillation [38–40] and the survey article [41].

By considering the use of deep learning in the context of GNSS spoofing attacks, researchers can continue to develop and refine techniques for detecting and mitigating these attacks. In [42] to achieve an improved detection probability of GPS spoofing, a decision fusion-based identification system is employed. The singular values of the wavelet transform coefficients of both authentic and spoofed signals are utilized as feature vectors and fed into three classifiers, namely support vector machines (SVM), probabilistic neural networks (PNN), and decision tree (DT), to identify GPS spoofing. By merging the outcomes of the three classifiers using a K-out-of-N decision rule, the ultimate classification outcome yields a greater probability of detection and a lower rate of false alarms. [43] presents a technique for identifying GPS spoofing that involves using a Multi-Layer Neural Network that takes in feature indices as inputs and leverages traditional machine learning algorithms like K-Nearest Neighbor (KNN) and Naive Bayesian classifier for detecting spoofing. [30] The approach also entailed employing a Multi-Layer Perceptron (MLP) neural network classifier that was trained via particle swarm optimization (PSO), where the received signal power and correlation function distortion were used as input. [44] introduced a MLP and two classes of Convolution Neural Networks (CNNs) to detect the existence of spoofing signal, with CAF as an input feature. A turn-by-turn spoofing attack detection technique that relies on deep reinforcement learning (RL) and utilizes low-cost in-vehicle sensor data is employed. Various machine learning (ML) algorithms have been utilized for spoofing detection, such as RNN based on long short-term memory (LSTM), classification SVM (C-SVM) that incorporates principal component analysis

(PCA), SC-SVM, and a method based on general adversarial networks (GAN), as seen in [31, 32, 45, 46], respectively.

This paper is an extended version of the methodology of applying deep neural network for GNSS spoofing detection in [29] that incorporates the concepts of dataset splitting and DNN parallelization from [47]. In [29], by utilizing efficient data-driven models trained over large datasets, the authors presented initial findings indicating that multimodal distributions or moderate-to-severe nonlinearities that impact the received signal can be effectively learned. The study suggested utilizing DNN models to detect or classify the existence of a single spoofing signal. While results were promising, the naive use of a DNN applied to the full CAF map (known to be very sparse) was not deemed to provide sufficiently satisfying results and be computationally manageable. This article's contributions are therefore to *i*) Enhancement of the spoofing detection accuracy and reduction of computational complexity can be achieved by dividing the dataset and processing it in parallel using DNN; *ii*) allow for non-coherent integration times within the DNN spoofer-detection with the process of combining multiple sources of data (data fusion steps) to improve the accuracy and reliability of spoofer detection from[47]; and *iii*) expanding the range of applicability of the technique to estimate the number of spoofing signals, performed through the incorporation of a Gaussian Mixture Model (GMM)-based clustering algorithm at the output of the DNN model.

The paper is structured in the following manner: section GNSS signal and spoofing models reviews the necessary concepts on GNSS signal acquisition and how it is impacted by spoofing signals. Section Data-driven GNSS Spoofing detection details the proposed DNN approach for spoofing detection, which includes a discussion on the deep learning scheme and structure, model training setup, the extension to non-coherent integration times, and a methodology to estimate the number of spoofing signals. Results are discussed in results section and conclusions are provided in conclusions section.

2 GNSS signal and spoofing models

This section describes the signal model considered in this article, as well as the fundamental signal processing that a GNSS receiver is in charge of. Essentially, the optimization of the so-called CAF. The section also discusses how such CAF is distorted by the presence of spoofing signals, thus making the signal detection problem potentially ambiguous.

2.1 GNSS signal acquisition

The discrete-time signal, which is shown in the following obtained after downconversion and sampling (at a rate of $f_s = 1/T_s$) from M satellites along with noise, is observed by a receiver.

$$\begin{aligned} y[n] &= \sum_{i=1}^{M} x_i[n;\boldsymbol{\theta}_i] + \eta[n] \\ x_i[n;\boldsymbol{\theta}_i] &= \alpha_i b_i(nT_s - \tau_i) c_i(nT_s - \tau_i) e^{j2\pi f_d nT_s + j\phi_i} \end{aligned} \tag{1}$$

The received signal from the i-th satellite is characterized by several parameters, including the amplitude α_i, data bits $b_i(\cdot)$ of the navigation message, spreading code $c_i(\cdot)$ of the satellite, time-evolving delay τ_i, Doppler frequency $f_{d,i}$, carrier-phase term ϕ_i introduced by the channel, and random noise $\eta[n]$ at the receiver. The noise is typically complex, zero-mean, and Gaussian distributed with variance σ^2. To simplify notation, all the signal parameters for the i-th satellite are combined into a vector $\boldsymbol{\theta} i = (\alpha_i, \phi_i, \tau_i, fd, i)^\top$ for clarity.

Signal acquisition is an essential preliminary step that a receiver must carry out. By performing signal acquisition, the receiver can determine which satellites are in view and available for use in subsequent navigation or communication tasks. This process involves a search through different code delay and Doppler frequency values to find the correct ones that match the incoming signal. Once acquired, the signal can be tracked, and more accurate estimates of the code delay and Doppler frequency can be obtained [14]. Hence, in the search for the i-th satellite, the problem can be formulated as a hypothesis-testing scenario with two competing hypotheses

$$\begin{aligned} &\mathcal{H}_0 : y[n] = \eta[n] \\ &\mathcal{H}_1 : y[n] = x_i[n; \boldsymbol{\theta}_i] + \eta[n] \end{aligned} \tag{2}$$

In this case, where $n = 0, \ldots, N-1$ represents the N samples used in the acquisition, the parameters in $\boldsymbol{\theta}_i$ being unknown, the optimal detection framework in terms of maximum likelihood (ML) is the Generalized Likelihood Ratio Test (GLRT). The GLRT requires the ML estimation (MLE) of the vector $\boldsymbol{\theta}_i$. Given a set of N observations denoted as $\mathbf{y} = (y[0], y[1], \ldots, y[N-1])^\top$, the MLE of $\boldsymbol{\theta}_i$ is obtained as:

$$\hat{\boldsymbol{\theta}}_i = \arg\max_{\boldsymbol{\theta}_i} p(\mathbf{y}|\boldsymbol{\theta}_i) \,, \tag{3}$$

Assuming that the parameters in $\boldsymbol{\theta}_i$ are piece-wise constant within the N samples of $\mathbf{y}$, and that the codes have ideal cross-correlation properties, it is commonly assumed in signal acquisition that these codes can be processed independently at the receiver.

The GLRT statistic is computed as the ratio of the maximum correlation value to the average correlation value, as evident from the equation above. By maximizing the correlation between the received signal and the local code, the GLRT approach can effectively distinguish between the desired signal and any interfering signals or noise [10–12, 14]. The Cross Ambiguity Function (CAF), which represents the correlation between $y[n]$ and the spreading code of the i-th satellite at a specific delay/Doppler pair in discrete time, encodes this correlation operation, as illustrated in the equation as follows:

$$\mathcal{C}_i(\tau, f_d) = \frac{1}{N} \sum_{n=0}^{N-1} y[n] \underbrace{c_i(n\mathrm{T_s} - \tau) \exp\{-j2\pi f_{d,i} n\mathrm{T_s}\}}_{\text{Local replica}} \,, \tag{4}$$

which can be expressed more compactly in vector notation after gathering N samples from the samples and the local code as $\mathbf{y}, \mathbf{c}_i \in \mathbb{C}^{N\times 1}$ as $\mathcal{C}_i(\tau, f_d) = \frac{\mathbf{c}_i^H \mathbf{y}}{N}$.

The Maximum Likelihood Estimation (MLE) of $\boldsymbol{\theta}_i$ can be expressed in terms of the CAF as follows:

$$(\hat{\tau}_i, \hat{f}_{d,i}) = \arg\max_{\tau, f_d} \left\{ \left| \mathcal{C}_i(\tau, f_d) \right|^2 \right\} \tag{5}$$

$$\hat{\alpha}_i = \left| \mathcal{C}_i(\hat{\tau}_i, \hat{f}_{d,i}) \right| \tag{6}$$

$$\hat{\phi}_i = \angle \mathcal{C}_i(\hat{\tau}_i, \hat{f}_{d,i}), \tag{7}$$

and we make a decision on the presence of the i-th satellite by setting a detection threshold β, which is designed to achieve a desired false alarm probability, on the test statistic in the optimization problem in:(567) such as

$$\left| \mathcal{C}_i(\tau, f_d) \right|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \beta \ . \tag{8}$$

The CAF is a function that is dependent on the delay τ and the Doppler frequency f_d of the local replica. The optimization in (567) is performed over a grid of possible τ and f_d values, typically evaluating the CAF on a set of discrete values. This grid, known as the search space, consists of a set of cells that include different values of delay and Doppler, represented by vectors $\boldsymbol{\tau} \in \mathbb{R}^{n_\tau}$ and $\mathbf{f}_d \in \mathbb{R}^{n_f}$, respectively. Typically, the number of delay values n_τ is much larger than the number of Doppler values n_f.

There are several strategies for evaluating this grid, which trade-off search speed and performance [14].

2.2 GNSS signal spoofing effects on acquisition

The spoofer is an interference transmission of a forged GNSS-like signal generated with the purpose of manipulating a victim's receiver's estimated position and time. The spoofer sends a set of false signals that mimic the legitimate satellite signal, except for those parameters that would eventually cause a different position estimate at the receiver unless properly detected. The received GNSS signal, with a spoofing attack, is therefore as follows:

$$y[n] = \sum_{i=1}^{M} x_i[n; \boldsymbol{\theta}_i] + \sum_{j=1}^{M_s} x_j[n; \boldsymbol{\theta}_{s,j}] + \eta[n] \tag{9}$$

where M_s, denotes the number of spoofed signals. In order to deceive the receiver [48], each spoofed signal must have the same spreading code c_i of the satellite it is trying to supersede and broadcast a valid navigation message b_i. The spoofed amplitude, code phases, and carrier phases are gathered in (9) by $\boldsymbol{\theta}_{s,j}$ for the j-th spoofer. When building a spoofer detector, two hypotheses are tested:

1 The null hypothesis is ($\mathcal{H}_0$) that the legitimate signal and noise are present, but there is no spoofing signal,

$$\mathcal{H}_0 : y[n] = \sum_{i=1}^{M} x_i[n; \boldsymbol{\theta}_i] + \eta[n] \ .$$

2 The alternative hypothesis is ($\mathcal{H}_1$) that the legitimate signal, spoofed signal, and noise are present in the dataset;

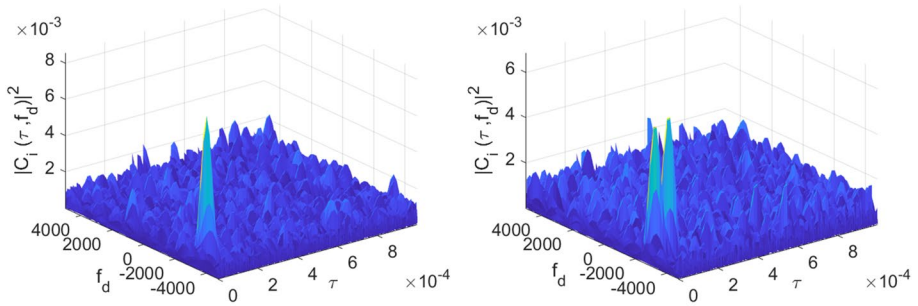
$$\mathcal{H}_1 : y[n] = \sum_{i=1}^{M} x_i[n; \boldsymbol{\theta}_i] + \sum_{i=1}^{M_s} x_{s,i}[n; \boldsymbol{\theta}_{s,i}] + \eta[n]$$

The effect of a spoofing signal on the CAF is well-known and shown in Fig. 1 for clarity. Figure 1a shows an arbitrary CAF under $\mathcal{H}_0$ (Fig. 1b) shows the situation when a spoofing signal is present as well, causing the appearance of a secondary peak on the CAF. This work proposes to train a deep neural network (DNN), data-driven model to learn to classify between spoofed or clean signal receptions.

3 Data-driven GNSS spoofing detection

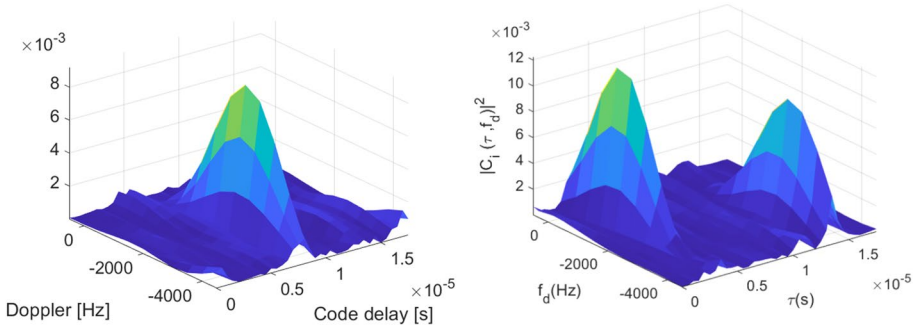
The purpose of this work is to design and use a neural network (NN) in order to recognize the spoofed signal from CAF *images*. Neural networks are computational models that consist of individual processing units called neurons. These neurons work together to perform complex data analysis tasks. A typical neural network architecture comprises an input layer, one or more hidden layers, and an output layer. Each layer is connected to adjacent layers via predefined activation functions that determine the output of the layer. During the training process, the neural network learns to perform a specific task by analyzing large amounts of labeled data. This involves adjusting the weights of the connections between the neurons using a technique called back-propagation. The weights are adjusted iteratively until the network's output matches the desired output for a given input. This process allows the neural network to learn from the data and improve its performance over time [49, 50]. In this paper, the task of the DNN model is to classify CAF maps as either spoofed or clean. This classification is performed in a probabilistic manner, yielding probabilities for both hypotheses. The input for the NNs is derived from the sub-image obtained by sliding across the delay/Doppler grid of the CAF evaluation. This sub-image can be viewed as an *image* and is further elaborated in the Deep Neural Networks training section.

The images (refer to Fig. 2 for an exemplary situation) derived from the sub-images have specific attributes that enable the identification of spoofed signals, including: *i*) the image should display a single peak corresponding to the authentic satellite signal



(a) Legitimate signal, $\mathcal{H}_0$ (b) Legitimate and spoofing signal, $\mathcal{H}_1$

Fig. 1 CAF evaluation at the delay/Doppler grid with $C/N_0 = 45$ dB-Hz



(a) Legitimate signal only, $\mathcal{H}_0$ (b) Legitimate and spoofing signal, $\mathcal{H}_1$

Fig. 2 Portions of the CAF fed for processing to the NN with $\Delta_m = 18$ and $\Delta_n = 5$ defining the size of the $\{m, n\}$-th sub-image. The resulting sub-image $\mathbf{Z}_i^{(m,n)}$ is shown on the reduced delay/Doppler grid in the case of **a** absence and **b** presence of a GNSS spoofed signal with $C/N_0 = 45$ dB-Hz

(provided it has sufficient power) in the absence of a spoofing signal. Another assumption is that the coherent integration time makes multipath effect negligible to the detector, which is commonly the case when long coherent times are employed in acquisition [10, 51]. If this assumption is not satisfied, the algorithm could confuse multipath detections with spoofing signals detections, as happens to the vast majority of spoofing detection methods in the current literature, although some solutions have been proposed in the past [52–54]; and *ii*) if a spoofing signal is present, the CAF *image* should comprise at least two peaks in each sub-image, or we have more than one sub-image with a single peak in the CAF *image*. This is used to train a NN model to classify between $\mathcal{H}_0$ and $\mathcal{H}_1$, the hypotheses described earlier. The details of trained NN are explained in Deep Neural Networks training section.

The suggested approach operates on a per-satellite basis, with the input data for the NNs consisting of the associated CAF image for each satellite, which we denote with $\mathbf{Z}_i \in \mathbb{R}^{n_\tau \times n_f}$ in the sequel.

It is possible to integrate this information into the hypothesis test, leading to the adaptation of a threshold γ. The details of calculating the test statistic can be found in [29]. The test statistic that is derived as a result of this approach is such that

$$\mathcal{T}(\mathbf{Z}_i) \triangleq \frac{p(\mathbf{Z}_i|\mathcal{H}_1)}{p(\mathbf{Z}_i|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \ , \tag{10}$$

The threshold γ serves as a tuning parameter for our spoofing detection algorithm. Since the test statistic is a ratio of probabilities, we have that $0 < \mathcal{T}(\mathbf{Z}_i) < \infty$. Similarly, as in [44], the DNNs provide estimated probabilities for each of the hypotheses in (10). Therefore, the input data would be $\mathbf{Z}_i$ and the output of the DNN would be the estimated probabilities in the dataset $\mathbf{y}$ used to build $\mathbf{Z}_i$.

3.1 Deep neural networks structure

Convolutional Neural Networks (CNNs) are a class of deep learning models that are particularly effective for image classification tasks. These models typically consist of tens or even hundreds of layers, with each layer responsible for recognizing specific features or patterns in an image [55, 56]. CNNs are designed to process images by applying a series

of convolutional filters to the input image. Each filter performs a specific task, such as detecting edges or recognizing shapes. The output of one layer serves as the input to the next layer, allowing the network to learn increasingly complex representations of the input image.

One of the key strengths of CNNs is their ability to automatically learn features from raw data, without the need for manual feature extraction in situations where data processing is otherwise challenging. This makes them particularly useful for tasks such as image recognition, where the features that distinguish one object from another may be difficult to define explicitly. This study employs a CNN structure, which is depicted in Fig. 3. More details of deep neural network structure can be found in [47].

The aim of this study is to use a neural network model to classify the presence or absence of the signal and spoofed signal in CAF maps, and accurately estimate the corresponding delay/Doppler parameters in the event of positive signal and spoofed signal detection. To achieve this objective, a CAF map is first generated in a densely sampled delay/Doppler grid and fed into the NN model to obtain posterior class probabilities. However, the size of the input matrix could be potentially large, with dimensions of $(n_\tau n_f)$, leading to computational complexity issues and increased costs associated with the need for a GPU with larger memory on the processing device. To address the computational complexity and memory constraints associated with the large input matrix in detecting spoofing attacks using CAF maps, [47] proposes a sliding scheme. The scheme involves scanning the large input CAF matrix using lower dimensional images as input to the NN classifier. This enables the classification of the presence/absence of the signal and spoofed signal in CAF maps and accurate estimation of the corresponding delay/Doppler parameters while reducing computational costs and memory requirements.

The concept of acquisition method, where the CAF map $\mathbf{Z}_i$ is split into smaller sub-images $\mathbf{Z}_i^{(m,n)}$, which are fed to a bank of parallel DNN binary classifiers to produce probability ratio maps is sketched in Fig. 4, where the $\{m, n\}$-th sub-image corresponds to the correct location of the delay/Doppler pair. The output of the DNN model is a probability ratio map, which is used in the subsequent Bayesian hypothesis test, labeled as $K = 1$ in the plot. It is worth noting that the probability ratio map may contain false peaks, as shown in Fig. 4 under $K = 1$. To mitigate those potential false detections, the section probabilistic signal detection describes a methodology to fuse non-coherent integrations of K DNN outputs. The impact of these integrations is illustrated in Fig. 4 In the probability ratio map obtained with $K = 6$ non-coherent integrations, as shown in the

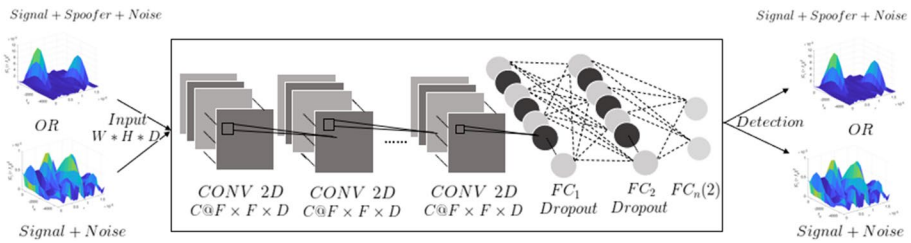


Fig. 3 Classification of signal ($\mathcal{H}_0$) or signal ($\mathcal{H}_0$) and spoofer ($\mathcal{H}_1$) in CAFs as part of the proposed GNSS signal acquisition scheme. Particularly, a set of convolutional layers followed by fully connected layers provides the capabilities of deep learning from large datasets

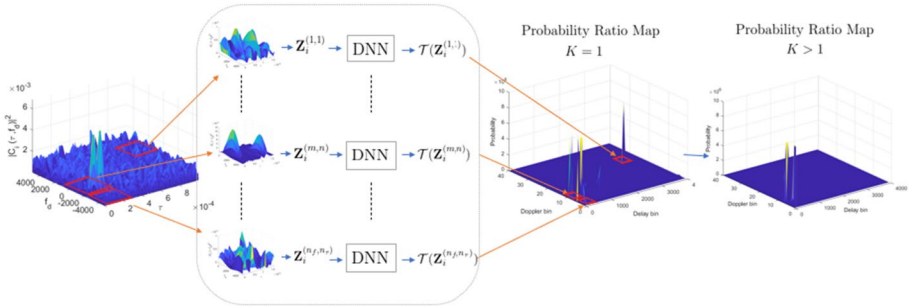


Fig. 4 Proposed acquisition method where the CAF map $\mathbf{Z}_i$ is split into smaller sub-images $\mathbf{Z}_i^{(m,n)}$, which are fed to a bank of parallel DNN binary classifiers to produce probability ratio maps. To increase accuracy, several ($K > 1$) probability ratio maps can be non-coherently fused, as shown on the rightmost plot

rightmost panel, the signal probability is emphasized in the correct delay/Doppler bin, while false peaks generated by noise are reduced.

3.2 Deep neural networks training

This section describes the methodology for training the model, which involved using a realistic GNSS signal simulator to generate I &Q samples from GPS L1 C/A satellites. The simulator was used to vary parameters according to the training plan outlined in this section. To improve the detection and localization accuracy, a higher sampling frequency may be desirable, since it increases the correlation between samples around the CAF peak. However, this also impacts the number of samples to be processed, and a trade-off must be considered. Therefore, the sampling frequency was increased to 4 MHz from the previously used 2 MHz in our initial work [29]. Raising the sampling frequency can result in a high-dimensional CAF image, which could complicate the use of a DNN model for classification and increase the computational cost. To mitigate this issue, the full CAF image can be split, and a sliding DNN scheme can be employed to reduce the complexity and cost, as described in [47] to detect the GNSS signal and applied in this work to detect the spoofed attack. This approach allows for efficient processing of the GNSS acquisition data, even at higher sampling frequencies, without requiring a device with a larger memory or GPU.

For model training purposes, thousands of snapshots of GPS L1 C/A, I &Q samples were generated to create a dataset. On the order of 10^4 images were used for training the model and 3000 images for testing. The dataset represented carrier-to-Noise-density ratios (C/N_0) ranging from 36 to 45 dB-Hz, and each snapshot had a length of 1 ms, which was equal to the duration of a code, making it the coherent integration time of the approach. In addition, to increase the diversity of the dataset and simulate real-world scenarios, random delays ranging from 0 to 1 ms and Doppler frequencies ranging from −4000 to 4000 Hz were introduced. The obtained I &Q samples were then used to compute the CAF maps over the Doppler-delay grid. These CAF maps can be viewed as images, where each pixel corresponds to a Doppler/delay cell, and the CAF value is denoted as $\mathbf{Z}_i$. However, processing such large CAF maps with a single NN model can be computationally expensive, as discussed earlier in the Section, Deep Neural Networks structure. For instance, when generating the CAF for a GPS L1 C/A signal with 50 Doppler bins (corresponding to 200 Hz bins for improved CAF peak identification),

the resulting images would be 4000×50 dimensional at the sampling frequency considered in this study. To address this challenge, a sliding DNN scheme can be employed to reduce the complexity and cost, as described in Section Deep Neural Networks structure.

To avoid this computational expense, $\mathbf{Z}_i$ is split into smaller images of size 11×36 (Doppler $\times$ delay). This results in a total of 158600 low-dimensional sub-images, which can be efficiently processed by the NN in parallel. A sub-image size of 11×36 was selected to achieve a reasonable trade-off between sub-image size and model complexity. This approach ensures that the complete CAF peak falls exactly in the middle of the sub-image. Choosing an appropriate sub-image size is crucial, as selecting smaller sizes could lead to missing the CAF peak, while larger sizes could result in multiple peaks and increased computational complexity.

The sliding approach can be seen as analogous to the convolutional layers utilized in a CNN. In this method, the CAF is scanned in smaller windows with a stride of one cell. These windows can include both the primary signal peak and any spoofed signals. Notably, these peaks exhibit correlations in both the delay and Doppler domains, which can be utilized by the NN classifier, as opposed to random noise-generated peaks.

To train the NN-based classifier, the dataset generated comprised of snapshots with either spoofed signals and signal-plus-noise ($\mathcal{H}_1$) or signal-plus-noise ($\mathcal{H}_0$), which were then divided into sub-images as illustrated in Fig. 2. These types of snapshots were used as input for the NN training. In a supervised manner, the classifier learned its parameters by observing a set of 3000 input/output pairs. A *softmax* layer with dropout was used as the NN output, producing the binary class probabilities needed to compute the test in (10) or its non-coherent version, which is described in [47] and section probabilistic signal detection.

3.3 Probabilistic signal detection

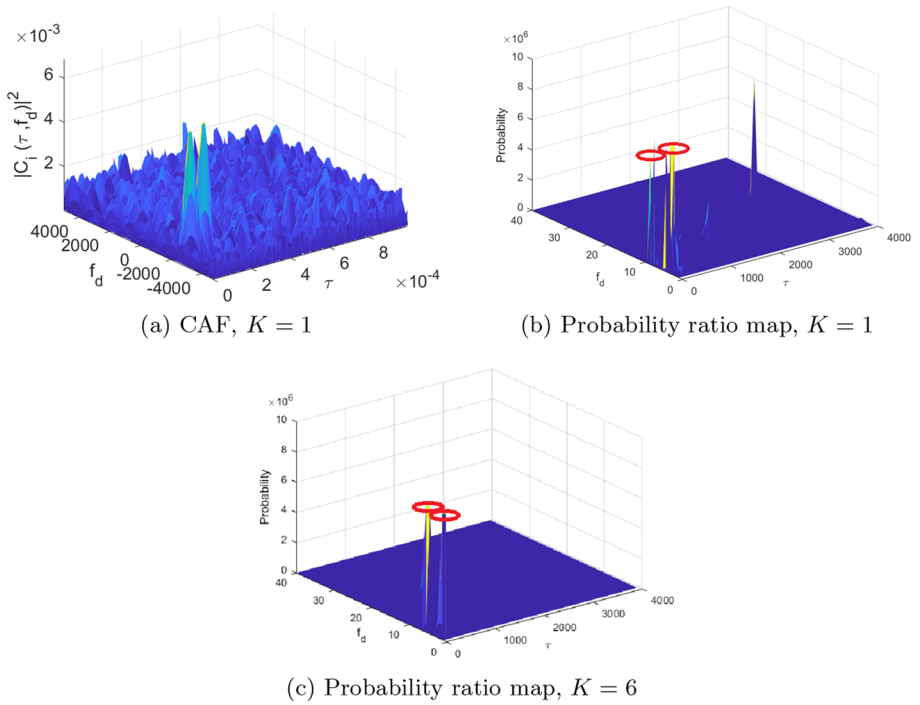
One way to implement non-coherent integration is to merge the multiple probability ratio maps obtained from processing CAF images using the NN architecture explained in Section, Deep Neural Networks structure We denote by $K \in \mathbb{Z}^+$ the total number of non-coherent integrations. This section discusses the data fusion of such multiple classifiers, which was first introduced in [47] and applied here in the presence of spoofer. Increasing the integration time, whether coherently or non-coherently, is known to enhance the overall detection performance of the acquisition process. Similarly, the proposed data-driven classifier also benefits from non-coherent integrations (i.e., fusing multiple classifier solutions) that improve the reliability of the probability maps by reducing falsely detected peaks and amplifying the locations where actual signals exist.

When dealing with non-coherent data snapshots, a group of K CAF maps are generated. Traditionally, this set corresponds to complete CAF maps $\mathbf{Z}_{i,k}$, where $k = 1, \ldots, K$. However, when using the sub-image approach, a distinct sub-image is produced for each integration period $\mathbf{Z}_{i,k}^{(m,n)}$. We apply Bayes' rule to obtain an optimal fusion rule that combines the class probabilities of the K classifiers, assuming that they are conditionally independent given their own data, which contain the binary class probabilities of the K (snapshots) classifiers explicitly: $p(\mathcal{H}_0|\mathbf{Z}_{i,k}^{(m,n)})$ and $p(\mathcal{H}_1|\mathbf{Z}_{i,k}^{(m,n)})$. The statistical test can then be formulated as

$$\mathcal{T}(\mathbf{Z}_{i,1}^{(m,n)},\ldots,\mathbf{Z}_{i,K}^{(m,n)}) = \prod_{k=1}^{K} \frac{p(\mathcal{H}_1|\mathbf{Z}_{i,k}^{(m,n)})}{p(\mathcal{H}_0|\mathbf{Z}_{i,k}^{(m,n)})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \tag{11}$$

such that the decision threshold becomes $\gamma = 1$ when $\mathbb{P}(\mathcal{H}_0) = \mathbb{P}(\mathcal{H}_1)$. The optimal fusion rule can be observed to involve the multiplication of the binary class probabilities of the K, (as previously shown in [57]). Obtained from processing the NN classifier on the K non-coherent integrations, these results represent the probability maps to be combined using the optimal fusion rule. The decision threshold plays a crucial role in determining P_d and P_{fa} of the overall classifier. A reasonable assumption is that both hypotheses are equally likely, which yields the choice of $\gamma = 1$ as the decision threshold. This intuitively implies that the detector selects the hypothesis with the largest a posteriori probability. Indeed, having access to the test statistics distribution under both hypotheses when using a NN becomes challenging, which would be required for optimal adjustment of the threshold. However, in the experiments, we considered $\gamma = 1$ as the design choice, yielding satisfactory results.

An example of how the classifier's performance is affected by the fusion rule is illustrated qualitatively in Fig. 5. The CAF delay/Doppler map utilized in standard signal acquisition in the presence of a spoofer is displayed in Fig. 5a. This map is generated without any non-coherent integration and only 1 ms coherent integration. It is well-known in the GNSS literature that outside the true peak (indicated by a red circle), the noise floor is relatively spiky and can result in significant false alarms, particularly at low



(a) CAF, $K = 1$

(b) Probability ratio map, $K = 1$

(c) Probability ratio map, $K = 6$

Fig. 5 Comparison of the delay/Doppler grid in the presence of spoofer for **a** standard CAF map with coherent integration only, **b** probability map produced by the data-driven classifier with coherent integration only, and **c** probability map after fusing $K = 6$ non-coherently classifier outputs. The GNSS signal had a C/N_0 of 42 dB-Hz and the red circle highlights the location of the peak generated by the GNSS signal

C/N_0 values. In contrast, the proposed data-driven method uses the CAF values to produce the probability ratio maps, as defined on the right-hand side of the figure.(11).

Figure 5b illustrates the probability ratio map obtained after processing the CAF shown in Fig. 5a which displays a reduced variability in the noise floor, although some residual spikes can still be detected in delay/Doppler bins where no signal was present. This effect is further smoothed with the fusion method, as illustrated in Fig. 5c where the NN model uses sub-images as inputs to produce a class probability pair, as shown in Fig. 2. The posterior probabilities produced by the NN take into account the delay/Doppler correlations of the CAF around the signal peak, which is in contrast to the standard method. The standard method only considers the maximum value of the CAF and neglects the waveform generated by the noise floor (i.e., the autocorrelation function of the corresponding spreading code).

3.4 Estimating number of signals

As opposed to the case in [47], this paper deals with the possible situation where multiple signals might be present in the data. To estimate that number of signals (that is, the legitimate signal and an arbitrary number of spoofers), the proposed DNN scheme is connected to a Gaussian Mixture Model (GMM)-based clustering method. In particular, the probability ratio map produced by the DNN is fed into the GMM, which is in charge of determining the number of peaks (i.e. signals) and their location in order to approximate that probability map as a linear combination of Gaussian functions. This scheme leverages the notable performance of the scheme in [47] to detect signals and complements it with a GMM clustering method that deals with the estimation of the number of signals. This approach aims at avoiding having to train a potentially more complex deep learning model to predict the number of signals as well as the signal presence.

GMM consists of a linear superposition of Gaussian components, providing a richer class of density models than single Gaussian models [58], particularly relevant in the case at hand where the probability map is known to be multimodal in the presence of spoofers. The Gaussian mixture distribution for the problem at hand can then be written as:

$$p(\boldsymbol{\psi}) = \sum_{\ell=1}^{L} \mathcal{N}(\boldsymbol{\psi}|\boldsymbol{\mu}_\ell, \boldsymbol{\Sigma}_\ell) \tag{12}$$

where $\boldsymbol{\psi} = (\tau, f_d)^\top$ is a two-dimensional vector with the delay/Doppler values at which the probability map $\mathcal{T}(\mathbf{Z}_i)$ is evaluated. L determines the number of components in the mixture, which in this case represents an estimate of the number of spoofing signals plus legitimate signals. The parameters $\{\boldsymbol{\mu}_\ell, \boldsymbol{\Sigma}_\ell\}_{\ell=1}^{L}$ denote, respectively, the mean and covariance of each of the L Gaussian in the mixture.

The assumption is that the satellite signal is always observed, that is, when the spoofing signal is absent, $L = 1$ and $\boldsymbol{\psi}$ is Gaussian distributed; when the spoofing signal occurs, $L > 1$ and $\boldsymbol{\psi}$ would be a Gaussian mixture. Given the observed data, an Expectation-Maximization (EM) algorithm is used in order to learn the parameters L and $\{\boldsymbol{\mu}_\ell, \boldsymbol{\Sigma}_\ell\}_{\ell=1}^{L}$. In order to compare different mixture complexities (i.e. values of L), the Bayesian Information Criterion (BIC) is employed. BIC is a popular criterion for model

selection among a finite set of models, a model with lower BIC is generally preferred. It is defined as:

$$\mathrm{BIC} = m \ln n - 2 \ln(\mathcal{L}) \tag{13}$$

where $\mathcal{L}$ is the marginal likelihood of the model; n is the sample size, which in this work is a number of the points that passed the threshold; m is the number of parameters estimated by the model, being $m = 6K$ in our case. By fitting different GMMs with varying L values using the EM algorithm, BIC can be used to measure the performance of each and assess their modeling capabilities, ultimately useful in estimating the number of signals given the observed data.

Figure 6 explains the process that is followed by the GMM clustering method in order to detect and localize the signal and spoofing signals. In particular, a running example where three signals (i.e. one legitimate and two spoofers) are present in the dataset. Initially, the proposed method uses the computed CAF delay/Doppler map, shown in Fig. 6a, to feed a DNN classifier. The DNN model is in charge of producing probability ratio maps, used to determine the presence or absence of signals in the map. This model can operate with coherent processing of codes (e.g. 1 ms long codes for GPS L1 C/A signals) or non-coherently over K coherently computed CAFs. After the DNN process, the probability ratio map is obtained, as shown in Fig. 6b. The data in Fig. 6b are thresholded and then fed to the GMM clustering algorithm, which uses BIC to decide the number of

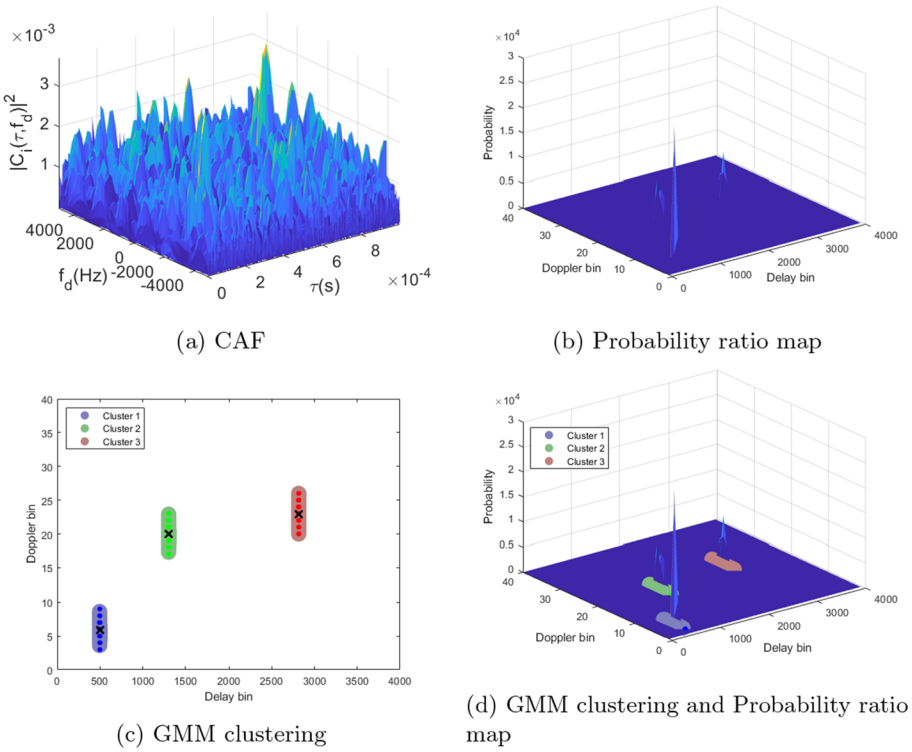


Fig. 6 Running example showing the process followed by the proposed algorithm. The experiment consists of a legitimate signal and two spoofers with $C/N_0 = 42$ dB-Hz. The various panels show the corresponding **a** CAF; **b** probability ratio map; **c** top view of the threshold probability ratio maps, after clustering is applied; and **d** three-dimensional perspective of the latter with the probability ratio maps overlaid

clusters L. As shown in Fig. 6c, d, it can be seen how the clustering algorithm is capable to associating the probability values corresponding to each of the signals together. In addition, notice that the resulting estimate of the mean of each mode in the mixture (that is, $\boldsymbol{\mu}_\ell$) is also an estimate of the delay/Doppler for that signal.

4 Results

The proposed data-driven spoofer detection scheme was validated using synthetic data. Particularly, a first set of experiments was performed to evaluate the ROC plots of the proposed classifier for the case of one spoofing signal. In these experiments, different C/N_0 values were considered and the DNN model training considered CAF images produced by 1 ms coherent correlation process without non-coherent integration times, as done in [47]. In training, the use of $K = 6$ DNN outputs was fused according to the methodology described earlier. Figure 7 shows the ROC results for the proposed method (dashed lines). These are compared to the theoretical performance of the standard method (solid lines), although in that case, the theoretical results are for the case of detecting a single signal in noise for which this result is available [14], while it is not the case for multiple signals detection. The result shows that for low C/N_0 values, the proposed scheme performs poorly, mostly caused by the CAF peaks being too weak for the DNN to discern from noise. However, when the C/N_0 values are increased, the CAF peaks become higher in the sub-images and the proposed method eventually outperforms the standard methods since it can distinguish the difference between signal/spoofer and noise. The explanation for this performance improvement is that, especially in the moderate to high signal-to-noise ratios, the NN uses more information than CAF-maximization methods. The latter is based on a single correlator sample of the CAF, while the NN-based solution leverages the CAF correlation (that is, the CAF waveform in time and frequency) [47]. The results seem to indicate that this additional information

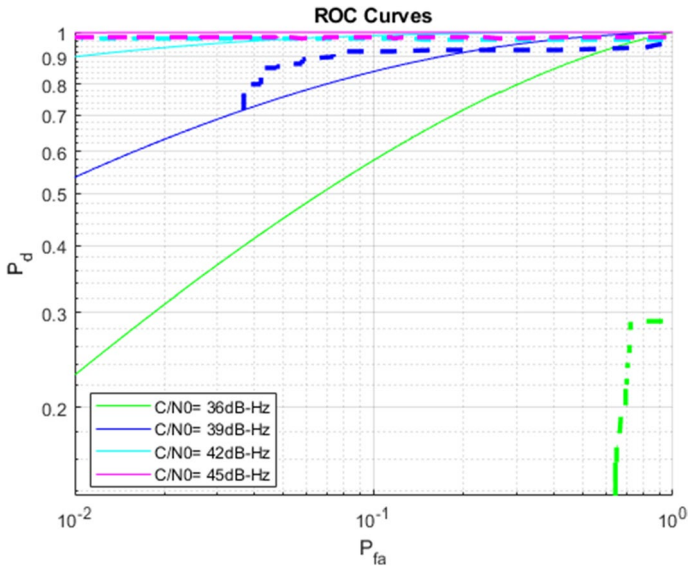


Fig. 7 ROC curves for a 1 ms coherently integrated snapshot and $K = 6$ non-coherently processed blocks for a variety of C/N_0 values for signal and spoofer. The performance of the proposed scheme (dashed lines) is compared to the theoretical performance of standard methods (solid lines)

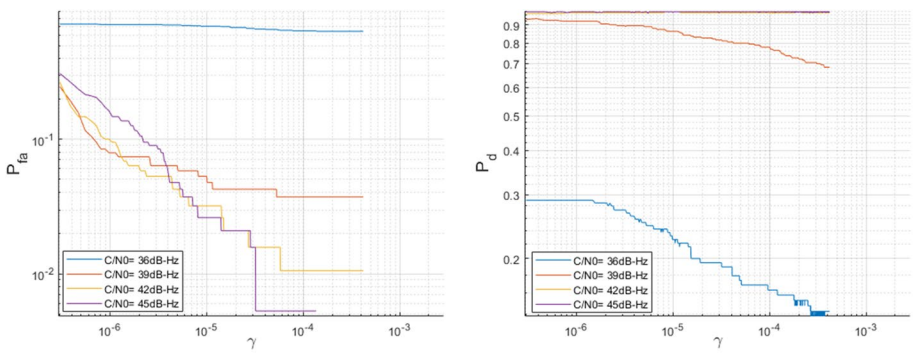
can be exploited by the NN classifier to better distinguish the difference between signal and noise.

The probabilities of false alarm and detection for all C/N_0 corresponding to ROC in Fig. 7 are shown in Fig. 8. As it is shown in Fig. 8a for the lowest C/N_0 the probability of a false alarm is high because the NN detects the noise as a signal/spoofer, while for other C/N_0 values increasing the threshold (γ) decreases the probability of false alarm. Additionally, the probability of detection is depicted in Fig. 8b where a similar behavior can be observed.

Another relevant experiment that was performed with the DNN (plus GMM) scheme proposed in this paper was to evaluate its performance for different separations of the signal and spoofing signals. The objective was to assess the impact on both the DNN probability ratio maps and the GMM-based clustering method. It is indeed relevant to understand, how close signals can be before the detector starts degrading. Figure 9 shows the probability of the detection according to the relative delay ($\Delta\tau$) for three different Doppler bins. The red solid line with ellipses represents the case where the two signals have the same Doppler frequency; the green solid line with circles reports the results when they have 1 Doppler bin separation; and the blue solid line with stars represents the case when they are two Doppler bins separated. As it is expected, by decreasing the delta delays the probability of detection decreases as well. Similarly, when they are in the same Doppler bin, they have the worst probability of detection as they cannot be distinguished as they get closer.

5 Conclusions

This paper investigated the use of deep neural networks as a method to detect GNSS spoofing attacks. This work builds on previous promising works whereby deep learning was used to detect legitimate GNSS signals from noise. In the case of spoofing, the situation is slightly more challenging, which makes the use of deep learning models more relevant. In addition, to efficiently implement the data-driven classifier through an image-splitting process (enabling parallelization), the article considers a Gaussian mixture model approach to determine the number of spoofing signals. Compared to standard GNSS signal detection schemes, the proposed method is more computationally



(a) $P_{fa}(\gamma)$ (b) $P_d(\gamma)$

Fig. 8 **a** $P_{fa}(\gamma)$ and **b** $P_d(\gamma)$ probabilities for a 1 ms coherently integrated snapshot, $K = 6$ non-coherent processing, and a variety of C/N_0 values when signal and spoofer are present

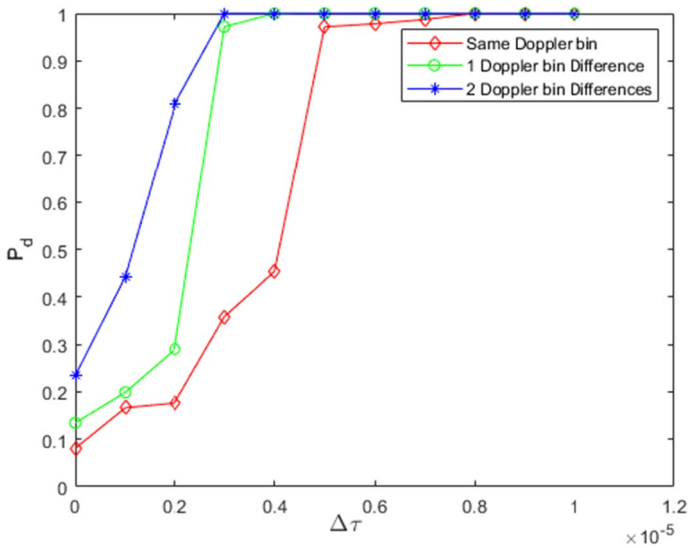


Fig. 9 Probability of correctly detecting both legitimate and spoofing signals, as well as determining their number, as a function of their relative delay separation. Also, the same Doppler, one bin, and two bin separations were considered. Bin size being 500 Hz

demanding since it requires the use of multiple NN models. Although the architecture can be parallelized, improving its computational performance is required for certain applications requiring spoofing detection at high rates, while for others where only sporadic spoofing detection is sough the method could operate. Results show that the proposed deep learning method can outperform current approaches, especially in the moderate to high signal-to-noise ratios. Similarly to state-of-the-art spoofing detection methods, the proposed solution might not be able to discriminate between multipath and spoofing signals due to their inherent similarities. Current and future work in the topic of spoofing detection is foreseen to be along the lines of discerning it from multipath occurrences, particularly in environments such as the urban canyon. Additional, with the objective of exploring the generalization capabilities of the proposed solution, future work involves the testing of the trained model with publicly available spoofing datasets such as TEXBAT [59] or OAKBAT [60].

Acknowledgements
This work was partially supported by the National Science Foundation under Awards ECCS-1845833 and CCF-2326559, as well as NU-FY21-TIER1 Award.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 3 May 2023 Accepted: 22 December 2023
Published online: 18 January 2024

References

1. D. Dardari, E. Falletti, M. Luise, *Satellite and Terrestrial Radio Positioning Techniques: A Signal Processing Perspective* (Academic Press, Boston, 2011)
2. M.G. Amin, P. Closas, A. Broumandan, J.L. Volakis, Vulnerabilities, threats, and authentication in satellite-based navigation systems [scanning the issue]. Proc. IEEE **104**(6), 1169–1173 (2016)

3. D. Dardari, P. Closas, P.M. Djurić, Indoor tracking: theory, methods, and technologies. IEEE Trans. Veh. Technol. **64**(4), 1263–1278 (2015)
4. N. Williams, P.B. Darian, G. Wu, P. Closas, M. Barth, Impact of positioning uncertainty on connected and automated vehicle applications. SAE Int. J. Connect. Autom. Veh. **6**(12-06-02-0010) (2022)
5. Z.M. Kassas, P. Closas, J. Gross, Navigation systems panel report navigation systems for autonomous and semi-autonomous vehicles: current trends and future challenges. IEEE Aerosp. Electron. Syst. Mag. **34**(5) (2019)
6. K. Yu, S.-H. Fang, A. Broumandan, P. Closas, G. Retscher, A. Dempster, IEEE access special section: positioning and navigation in challenging environments. IEEE Access **11**, 12636–12639 (2023). https://doi.org/10.1109/ACCESS.2023.3240354
7. R. Ioannides, T. Pany, G. Gibbons, Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques. Proc. IEEE **104**(6), 1174–1194 (2016)
8. H. Sathaye, G. LaMountain, P. Closas, A. Ranganathan, SemperFi: anti-spoofing GPS receiver for UAVs, in Network and Distributed Systems Security (NDSS) Symposium 2022 (2022)
9. F. Dovis, *GNSS Interference Threats and Countermeasures* (Artech House, Norwood, 2015)
10. E. Kaplan, C. Hegarty, *Understanding GPS: Principles and Applications* (Artech house, Norwood, 2005)
11. P. Misra, P. Enge, Global Positioning System: signals, measurements and performance second edition. Global Positioning System: Signals, Measurements And Performance Second Editions (2006)
12. J.B. Tsui, *Fundamentals of Global Positioning System Receivers: A Software Approach*, vol. 173 (Wiley, Hoboken, 2005)
13. D. Akos, J. Arribas, M.Z.H. Bhuiyan, P. Closas, F. Dovis, I. Fernandez-Hernandez, C. Fernández–Prades, S. Gunawardena, T. Humphreys, Z.M. Kassas et al., GNSS software defined radio: history, current developments, and standardization efforts, in Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022), pp. 3180–3209 (2022)
14. D. Borio, A statistical theory for GNSS signal acquisition. Ph.D. Dissertation Polytecnico di Torino (2008)
15. H. Mathis, P. Flammant, A. Thiel, An analytic way to optimize the detector of a post-correlation FFT acquisition algorithm. Quadrature **1000**, 1 (2003)
16. A. Whalen, *Detection of Signals in Noise* (Academic Press, Boston, 2013)
17. P. Closas, J. Arribas, C. Fernández-Prades, Spoofing detection by a reduced acquisition process, in Proceedings of the 2016 International Technical Meeting of The Institute of Navigation, pp. 726–731 (2016)
18. J. Dampf, T. Pany, W. Bär, J. Winkel, C. Stöber, K. Fürlinger, P. Closas, J. Garcia-Molina, More than we ever dreamed possible: processor technology for GNSS software receivers in the year 2015. Inside GNSS **10**(4), 62–72 (2015)
19. A. Siemuri, H. Kuusniemi, M.S. Elmusrati, P. Välisuo, A. Shamsuzzoha, Machine learning utilization in GNSS-use cases, challenges and future applications, in 2021 International Conference on Localization and GNSS (ICL-GNSS), pp. 1–6 (2021). https://doi.org/10.1109/ICL-GNSS51451.2021.9452295
20. A.A. Abdallah, Z.M. Kassas, Deep learning-aided spatial discrimination for multipath mitigation, in 2020 IEEE/ION Position, Location and Navigation Symposium (PLANS), pp. 1324–1335 (2020). IEEE
21. G. Zhang, P. Xu, H. Xu, L.-T. Hsu, Prediction on the urban GNSS measurement uncertainty based on deep learning networks with long short-term memory. IEEE Sens. J. **21**(18), 20563–20577 (2021)
22. H. Li, P. Borhani-Darian, P. Wu, P. Closas, Deep learning of GNSS signal correlation, in Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020), pp. 2836–2847 (2020)
23. H. Li, P. Borhani-Darian, P. Wu, P. Closas, Deep neural network correlators for GNSS multipath mitigation. IEEE Trans. Aerosp. Electron. Syst. (2022)
24. C. Savas, F. Dovis, Multipath detection based on K-means clustering, in Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019), pp. 3801–3811 (2019)
25. T. Suzuki, K. Kusama, Y. Amano, NLOS multipath detection using convolutional neural network, in Proceedings of the 33rd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2020), pp. 2989–3000 (2020)
26. E. Munin, A. Blais, N. Couellan, Convolutional neural network for multipath detection in GNSS receivers, in 2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT), pp. 1–10 (2020). IEEE
27. G. Caparra, P. Zoccarato, F. Melman, Machine learning correction for improved PVT accuracy, in Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021), pp. 3392–3401 (2021)
28. M.R. Manesh, J. Kenney, W.C. Hu, V.K. Devabhaktuni, N. Kaabouch, Detection of GPS spoofing attacks on unmanned aerial systems, in 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC), pp. 1–6 (2019). IEEE
29. P. Borhani-Darian, H. Li, P. Wu, P. Closas, Deep neural network approach to detect GNSS spoofing attacks, in Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020), pp. 3241–3252 (2020)
30. S. Tohidi, M.R. Mosavi, Effective detection of GNSS spoofing attack using a multi-layer perceptron neural network classifier trained by PSO, in 2020 25th International Computer Conference, Computer Society of Iran (CSICC), pp. 1–5 (2020). IEEE
31. R. Calvo-Palomino, A. Bhattacharya, G. Bovet, D. Giustiniano, Short: LSTM-based GNSS spoofing detection using low-cost spectrum sensors, in 2020 IEEE 21st International Symposium On" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM), pp. 273–276 (2020). IEEE
32. S. Semanjski, A. Muls, I. Semanjski, W. De Wilde, Use and validation of supervised machine learning approach for detection of GNSS signal spoofing, in 2019 International Conference on Localization and GNSS (ICL-GNSS), pp. 1–6 (2019). IEEE
33. R. Morales Ferre, A. de la Fuente, E.S. Lohan, Jammer classification in GNSS bands via machine learning algorithms. Sensors **19**(22), 4841 (2019)

34. A. Louis, M. Raimondi, Neural network based evil waveforms detection, in Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020), pp. 1984–1989 (2020)
35. D. Brum, M.R. Veronez, E. Menezes de Souza, I.r. Koch, L. Gonzaga, I. Klein, M.T. Matsuoka, V. Francisco Rofatto, A.M. Junior, G. Eliane dos Reis Racolte, F. Bordin, E.K. Nzinga, A proposed earthquake warning system based on ionospheric anomalies derived from GNSS measurements and artificial neural networks, in IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, pp. 9295–9298 (2019). https://doi.org/10.1109/IGARSS.2019.8900197
36. M. Alshaye, F. Alawwad, I. Elshafiey, Hurricane tracking using Multi-GNSS-R and deep learning, in: 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS), pp. 1–4 (2020). IEEE
37. Q. Yan, W. Huang, Sea ice sensing from GNSS-R data using convolutional neural networks. IEEE Geosci. Remote Sens. Lett. **15**(10), 1510–1514 (2018)
38. N. Linty, A. Farasin, A. Favenza, F. Dovis, Detection of GNSS ionospheric scintillations based on machine learning decision tree. IEEE Trans. Aerosp. Electron. Syst. **55**(1), 303–317 (2018)
39. Y.L. Liu, Y. Morton, Y.J. Jiao, Application of machine learning to characterization of GPS L1 ionospheric amplitude scintillation, in 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS), pp. 1159–1166 (2018)
40. M.O. Selbesoglu, Prediction of tropospheric wet delay by an artificial neural network model based on meteorological and GNSS data. Eng. Sci. Technol. Int. J. **23**(5), 967–972 (2020)
41. J. Vilà-Valls, N. Linty, P. Closas, F. Dovis, J.T. Curran, Survey on signal processing for GNSS under ionospheric scintillation: detection, monitoring, and mitigation, navigation. J. Inst. Navig. **67**(3), 511–536 (2020)
42. M. Sun, Y. Qin, J. Bao, X. Yu, GPS spoofing detection based on decision fusion with a k-out-of-n rule. IJ Netw. Secur. **19**(5), 670–674 (2017)
43. E. Shafiee, M. Mosavi, M. Moazedi, Detection of spoofing attack using machine learning based on multi- layer neural network in single-frequency GPS receivers. J. Navig. **71**(1), 169–188 (2018)
44. P. B.-D., P. Closas, Deep Neural Network Approach to GNSS Signal Acquisition, in 2020 IEEE/ION Position, Location and Navigation Symposium (PLANS), pp. 1214–1223 (2020)
45. S. Semanjski, I. Semanjski, W. De Wilde, S. Gautama, Use of supervised machine learning for GNSS signal spoofing detection with validation on real-world meaconing and spoofing data-part ii. Sensors **20**(7), 1806 (2020)
46. J. Li, X. Zhu, M. Ouyang, W. Li, Z. Chen, Q. Fu, GNSS spoofing jamming detection based on generative adversarial network. IEEE Sens. J. **21**(20), 22823–22832 (2021)
47. P. Borhani-Darian, H. Li, P. Wu, P. Closas, Deep learning of GNSS acquisition. Sensors **23**(3), 1566 (2023)
48. M.L. Psiaki, T.E. Humphreys, GNSS spoofing and detection. Proc. IEEE **104**(6), 1258–1270 (2016)
49. T. O'Shea, T. Roy, T. Clancy, Over-the-air deep learning based radio signal classification. IEEE J. Sel. Top. Signal Process. **12**(1), 168–179 (2018)
50. Q. Yan, W. Huang, C. Moloney, Neural networks based sea ice detection and concentration retrieval from GNSS-R delay-Doppler maps. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. **10**(8), 3789–3798 (2017)
51. Y.J. Morton, F. van Diggelen, J.J. Spilker Jr., B.W. Parkinson, S. Lo, G. Gao, *Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications* (Wiley, Hoboken, 2021)
52. A. Broumandan, A. Jafarnia-Jahromi, G. Lachapelle, R.T. Ioannides, An approach to discriminate GNSS spoofing from multipath fading, in 2016 8th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC), pp. 1–10 (2016). IEEE
53. J.N. Gross, T.E. Humphreys, GNSS spoofing, jamming, and multipath interference classification using a maximum-likelihood multi-tap multipath estimator, in Proceedings of the 2017 International Technical Meeting of The Institute of Navigation, pp. 662–670 (2017)
54. K. Ali, X. Chen, F. Dovis, On the use of multipath estimating architecture for spoofer detection, in 2012 International Conference on Localization and GNSS, pp. 1–6 (2012). IEEE
55. S. Liu, W. Deng, Very deep convolutional neural network based image classification using small training sample size, in 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 730–734 (2015). IEEE
56. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
57. F. Pastor, J. García-González, J.M. Gandarias, D. Medina, P. Closas, A.J. García-Cerezo, J.M. Gómez-de-Gabriel, Bayesian and neural inference on LSTM-based object recognition from tactile and kinesthetic information. IEEE Robot. Autom. Lett. **6**(1), 231–238 (2020)
58. C.M. Bishop, N.M. Nasrabadi, *Pattern Recognition and Machine Learning*, vol. 4 (Springer, Berlin, 2006)
59. T.E. Humphreys, J.A. Bhatti, D. Shepard, K. Wesson, The Texas spoofing test battery: toward a standard for evaluating GPS signal authentication techniques, in Radionavigation Laboratory Conference Proceedings (2012)
60. A. Albright, S. Powers, J. Bonior, F. Combs, Oak Ridge Spoofing and Interference Test Battery (OAKBAT)-Pure Tones. Technical report, Oak Ridge National Lab (ORNL), Oak Ridge Leadership Computing Facility (OLCF), Oak Ridge, TN (US) (2020)

Publisher's Note