

## Article

# Light Weight CSI-Based Physical Layer Authentication Model for IoT Networks

Monika Roopak <sup>1,\*</sup> , Yachao Ran <sup>2</sup>, Simon Parkinson <sup>3</sup>  and Jonathon Chambers <sup>4</sup>

<sup>1</sup> School of Computing, Engineering and Creative Industries, University of Bedfordshire, Luton LU1 3JU, UK

<sup>2</sup> School of Mechanical Engineering, Sichuan University, Chengdu 610207, China; yachaoran@qq.com

<sup>3</sup> Department of Computer Science, University of Huddersfield, Huddersfield HD1 3DH, UK; s.parkinson@hud.ac.uk

<sup>4</sup> College of Science and Engineering, University of Leicester, Leicester LE1 7RH, UK; jonathon.chambers@leicester.ac.uk

\* Correspondence: monika.roopak@beds.ac.uk

## Abstract

This paper introduces a novel physical layer authentication technique for Internet of Things (IoT) networks, leveraging channel state information (CSI) data from Wi-Fi signals to distinguish between authorised and unauthorised nodes and thereby enhancing security without compromising performance. The core novelty lies in its integrated framework, which employs non-negative matrix factorization (NMF) for efficient feature selection and a Gaussian mixture model (GMM) for identifying complex patterns within the CSI data specifically adapted to the dynamic nature of IoT networks. NMF is utilised to mitigate the high dimensionality and redundancy inherent in raw CSI metrics, reducing processing load, extracting salient features, alleviating overfitting risks, and exhibiting superior resilience to noise. Following NMF, the GMM component is used for data classification, capitalising on its probabilistic and soft clustering attributes to represent intricate distributions and handle heterogeneous CSI data characteristics. This integrated proposed methodology not only exploits the inherent nonlinear and probabilistic characteristics of CSI data but also upholds computational efficiency, making it highly suitable for implementation in resource-constrained IoT wireless networks. The model achieves exceptional classification proficiency, with an accuracy rate of 99.83 percent and a recall of 100 percent, which are crucial for cybersecurity and anomaly detection. Furthermore, the system is designed for efficiency and minimal resource consumption, exhibiting good computational efficiency, reduced training duration, and lower energy consumption compared with more complex, heavily exploited architectures for CSI data processing like CNN and CNN + LSTM, making it particularly suitable for resource-constrained IoT environments.

**Keywords:** IoT; GMM; NMF; CSI; lightweight; authentication; ESP32



Academic Editor: Christos J. Bouras

Received: 4 June 2026

Revised: 19 July 2026

Accepted: 20 July 2026

Published: 29 July 2026

**Copyright:** © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

The Internet of Things (IoT) is a network of connected objects and sensors that is vital in a variety of fields, including industries, smart cities, healthcare, and agriculture. It enables the automation, observation, and management of diverse systems, predominantly depending on a wireless connection for the interchange of data [1]. In order to address security problems in source-constrained situations, efforts have been made to build lightweight encryption algorithms for wireless sensor networks (WSNs) and the Internet

of Things [2]. The effort to adopt a lightweight paradigm in diverse technological frameworks, particularly concerning security within wireless IoT infrastructures, necessitates the navigation of a multifaceted landscape of trade-off. One significant hurdle associated with this task is achieving a balance between improving performance and the critical necessity of maintaining stringent security measures. In efforts to decrease resource demands, touching on processing potential, memory allocation, and energy levels, there appears to be a related risk of compromising security soundness. Conventional cryptography techniques are frequently computationally expensive for nodes [3], and key management presents difficulties for authentication systems that depend on secret keys [4,5]. This is pertinent when additional complexities, such as key management protocols, are integrated. Consequently, although lightweight methodologies aim to improve performance metrics, they must simultaneously address intrinsic vulnerabilities that materialise from the adoption of simplified or inadequately secure key management practices. Established cryptographic techniques, which customarily rely on set frameworks and notions about device performance, may experience obstacles in adequately securing such diverse situations. This change might uncover security flaws, highlighting the urgent demand for solutions that are equally flexible and strong across various devices and scenarios [6]. Compared with typical cryptographic solutions, Physical (PHY) layer techniques reveal a promising substitute. These methodologies capitalise on the intrinsic physical attributes of the communication medium to strengthen security measures. As a case in point, individuals could use the unique aspects of wireless signals, which encompass signal power and fluctuations over time, to set up secure communication channels. By focusing on the physical layer, these techniques can offer intrinsic security enhancements, as they do not require extensive computational resources or intricate key management systems [7]. This groundbreaking technique strengthens security while ensuring that the mobility and efficiency essential for the operational success of IoT devices are retained. In essence, although the drive for lightweight solutions within IoT networks encounters various challenges, particularly those around performance and security, investigating the PHY layer techniques uncovers a workable pathway to attain both targets. By re-evaluating the implementation of security and focusing on the distinctive attributes of the physical communication layer, it becomes feasible to devise solutions that are both effective and efficient, thus facilitating the sustained advancement and deployment of IoT technologies.

Using PHY layer authentication, lightweight authentication can be achieved, and the trade-off between IoT network security and performance may be addressed. The PHY layer authentication algorithms can identify if a node is authorised or unauthorised based on the physical characteristics of wireless communication. The number of characteristics of the physical layer that are often exploited, including carrier frequency offset (CFO), in-phase quadrature phase imbalance (IQI), channel impulse response (CIR), and the signal strength indicator (SSI) [8,9] are received by CSI [10,11].

In this article, CSI is utilised to improve the authentication process by harnessing its inherent capacity to distinctly characterise wireless channels and accurately differentiate between authorised and unauthorised IoT nodes. In contrast to other attributes of the physical layer, CSI captures the nuanced spatial-temporal dynamics of the wireless medium, facilitating more robust device identification even in challenging conditions, such as suboptimal signal-to-noise ratios [12]. By reflecting fluctuations in signal propagation, variations in power, and temporal channel behaviours, CSI furnishes a more comprehensive and discriminative fingerprint of communicating devices. This dependence on fundamental physical-layer characteristics—intrinsic attributes of wireless communication rather than environment-specific artefacts—enhances security while promoting broader applicability beyond strictly controlled laboratory environments [13]. As a result, the inherent

resilience and versatility of CSI render it a compelling basis for scalable and robust IoT authentication frameworks.

To address the challenges of security and resource limitations in IoT networks, our proposed physical layer authentication method leverages NMF and GMM. NMF plays a crucial role by mitigating the high dimensionality and redundancy in CSI metrics, effectively reducing processing load and extracting salient features essential for accurate authentication in resource-limited IoT devices [14]. This approach enhances computational efficiency, alleviates overfitting risks, and demonstrates superior resilience to noise, making it robust in dynamic environmental conditions. Following NMF, the GMM component further enhances suitability by adeptly representing intricate distributions and handling the heterogeneous characteristics of CSI data, which is critical for dynamic IoT networks. GMMs also offer resilience to noise, preserve data integrity, and are scalable for managing extensive datasets. This integrated methodology not only exploits the inherent nonlinear and probabilistic characteristics of CSI data but also upholds computational efficiency, making it highly suitable for implementation in resource-constrained IoT wireless networks. Promising applications of CSI are demonstrated in applications as a behavioural biometric [15–17].

The main contributions of this paper are as follows:

- We introduce a novel low-resource-consuming physical layer authentication method specifically designed for IoT devices with Wi-Fi capabilities. This method successfully addresses the unique difficulties that arise in IoT contexts, guaranteeing improved security and flexibility in the intricacies of networked systems.
- We achieve initial authentication by exploiting the CSI data obtained from the nodes, using the different physical placements of devices within the system to enhance the authentication accuracy.
- Using the CSI data from these nodes, an enhanced NMF + GMM-based method is proposed to categorise the nodes as permitted or unauthorised. This model offers a more efficient and reliable authentication mechanism that adapts to the dynamic and constantly changing nature of IoT networks.
- We employ the amplitude of the CSI data in the suggested model for both testing and training. The robustness of the model in various circumstances is demonstrated by its thorough validation of efficacy and reliability through intensive training and rigorous testing procedures, as well as by comparison with deep learning models and state-of-the-art methods.

## 2. Literature Review

Traditional CSI-based authentication methods, such as dynamic time warping (DTW) and the Neyman–Pearson (NP) test, require significant knowledge of channel behaviour, limiting their performance in dynamic communication situations [10,18]. These methods do not perform as well in real-world situations because they have trouble with complicated patterns and time-variant channels. The research gap for these methods lies in their inability to adapt to dynamic conditions and their reliance on extensive prior channel knowledge, making them impractical for many IoT deployments. Deep neural networks (DNNs), such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid convolutional recurrent neural networks (CRNNs), were developed in [19] to overcome these constraints. While RNNs are better at capturing the temporal correlations between frequencies, CNNs are better at extracting local characteristics from CSI. By combining the two, the CRNN model increases robustness and accuracy. The proposed method has the advantage of DNNs automatically adapting to varying conditions and also having the capability to reduce false alarms and missed detections in comparison with other

techniques, along with being capable of capturing local and temporal patterns in CSI data. However, this methodology requires a high amount of resources to execute in real time, and there is the possibility that DNNs overfit if diversity in training CSI data is not available.

CSI in OFDM systems provides rich data for Wi-Fi sensing applications such as device-free localisation, human pose recognition, and person identification. Historically, research has relied on the Intel 5300 Network Interface Card (NIC) to gather CSI data. Its mobility and versatility are limited by the requirement for a host computer, which prevents researchers from investigating scenarios such as mobile receiver (RX) devices or active repositioning. The authors of [20] presented a low-cost CSI collection method that uses the ESP32 microcontroller and the Espressif IoT Development Framework. The ESP32 offers mobility and repositioning capabilities that were previously unattainable with conventional equipment, allowing data collection with a smartphone or in a single pocket. They demonstrated that by including the mobility to the Rx nodes, they achieved a 29.4 percent increase in accuracy. ESP32 has the advantage of being compact, independent, and able to be repositioned for optimised sensing tasks. The proposed technique is not applicable to complex real-time processing, which might need more processing power. ESP32 has limited computation power, and as the accuracy is directly dependent on the location of the node, it might decrease according to the location, therefore bringing another challenge to the proposed methodology.

The authors of [21] proposed work to tackle the challenge of other physical layer authentication methods based on channel characteristics that rely on quantisation algorithms to simplify the authentication, but they perform worse due to quantisation mistakes and difficulties in determining the right thresholds. Their research introduced the Multiple CIRs PLA (MCP) scheme and the Enhanced Multiple CIRs PLA (EMCP) scheme, two novel CIR-based physical layer authentication (PLA) methods designed to overcome these problems. By eliminating the need for quantisation, the MCP approach increases the accuracy of authentication. By adding the channel correlation coefficient, the EMCP scheme improves performance even further and offers higher authentication reliability. Their proposed methodology benefits from directly exploiting CSI data by bypassing quantisation to obtain high accuracy in low-signal-to-noise (SNR) ratio conditions, but certainly, the effectiveness of this scheme can decline with reduced channel correlation and a low SNR, which can restrict its use in certain environments.

The work presented in [22] assesses the viability of implementing Wi-Fi-sensing systems on the edge by exploiting CSI data, with an emphasis on inexpensive microcontrollers such as the ESP32. The study investigated machine learning and signal processing techniques that are currently in use and can be used with CSI data on edge devices. The study addresses critical difficulties and methodologies necessary for signal processing, dimensionality reduction, and real-time machine learning inference right at the edge. The work fine-tuned machine learning models to improve accuracy across several scales of sensing tasks, from large-scale activity and location sensing to hand gesture detection using tree-structured Parzen estimator (TPE) hyperparameter optimisation. This method benefits from being feasible for real-time CSI data processing and machine learning inference on edge devices using optimised signal processing techniques, and they employed Tensorflow Lite (TFlite) to execute a machine learning interface on ESP32 with limited computation power for online predictions without the requirement of a powerful cloud-based infrastructure. However, because of the resource constraint, it may not be applicable for high-dimensional signal processing to run complex models.

In [23], a ResNet-based method for physical layer authentication in the Industrial Internet of Things (IIoT) environment is presented to address the issue of building a scheme to achieve high authentication accuracy in complex and diverse channels due to

the occlusion of the metal device that contains dynamic mobile nodes. In their proposed method, they implemented an exponentially averaging data augmentation algorithm and transfer learning parameter tuning for training the model to speed up network coverage and increase the detection accuracy. The proposed method leverages the transfer learning using the ResNet50 pretrained model to achieve an incredibly high 99.64 percent accuracy and data augmentation to resolve the issue of overfitting. ResNet offers high accuracy but requires high computational resources, which can be a challenge in the case of resource-constrained connected IoT devices. In addition, the complexity of this method may increase with the decrease in the number of nodes in the IIoT.

An improved deep learning-based physical layer security (PLS) model named the Bahdanau Attention AutoEncoder (BAAE) is presented in [24] to address the issue of traditional deep learning-based PLS models that rely on large datasets to obtain high authentication accuracy, which face challenges in dynamic environments due to Doppler shifts and multipath fading caused by the varying mobility velocities of the users. The proposed model focuses on extracting relevant channel features and mitigating interference, such as inter-symbol interference (ISI) and inter-carrier interference (ICI). This method has the advantage of not being dependent on large data availability for the training, and it achieved 99.6 percent accuracy, not being affected by the Doppler shift effect. The model was tested in a simulation based on China Telecom's 5G network parameters, and thus it needs to be tested on real-world CSI data. This model is tailored to dynamic CSI predictions only and may give different results in other types of communication scenarios where different factors influence PLS performance, and since this method is a deep learning-based method, it demands high amounts of computational power to execute.

The variability in mobile surroundings poses difficulties, including spatial position ambiguity and channel instability, which reduces PLA's dependability. Maintaining constant authentication performance in mobile contexts is challenging due to oscillations in CSI caused by device mobility and environmental dynamics. The authors of [25] suggested Time-Sensitive Secure Transmission-Physical Layer Authentication (TSST-PLA), a knowledge-enhanced PLA algorithm that takes into account past knowledge about wireless channels and mobile devices, as a solution to these problems. The program seeks to increase the reliability of authentication by extracting stable channel features and confirming the rationality of mobile trajectories through the use of a novel three-scale feature model and trajectory verification. The proposed algorithm leverages prior knowledge to mitigate the effect of channel instability and spatial variations, and it is able to extract the stable channel characteristics from unstable CSI to improve the accuracy of location predictions. However, the performance of the proposed method can be limited in environments where such knowledge is sparse or unavailable, and also, the performance is poor for distant Rx-Tx because of the low SNR. A comprehensive survey of deep learning methods used in human activity recognition (HAR) that leverages CSI data in the context of Wi-Fi-based sensing technologies is presented in [26]. The paper identifies several limitations regarding the application of CNNs in HAR. One key limitation arises from the input adaptation process: when each dimension of time series sensor readings is treated as a separate channel for 1D convolution, the approach tends to ignore the crucial dependencies that exist between these different dimensions and sensors, potentially affecting performance. Furthermore, a broader challenge for traditional AI algorithms, including convolutional models, is their inherent focus on local features. This characteristic can be restrictive, as it may not adequately capture the intricate spatial and temporal relationships present across diverse sensor modalities and varying time intervals within time series data from wearable devices. A new security approach was proposed in [27] that can address the demands of 6G networks and overcome the limitations of conventional complex-based cryptographic protocols. In this

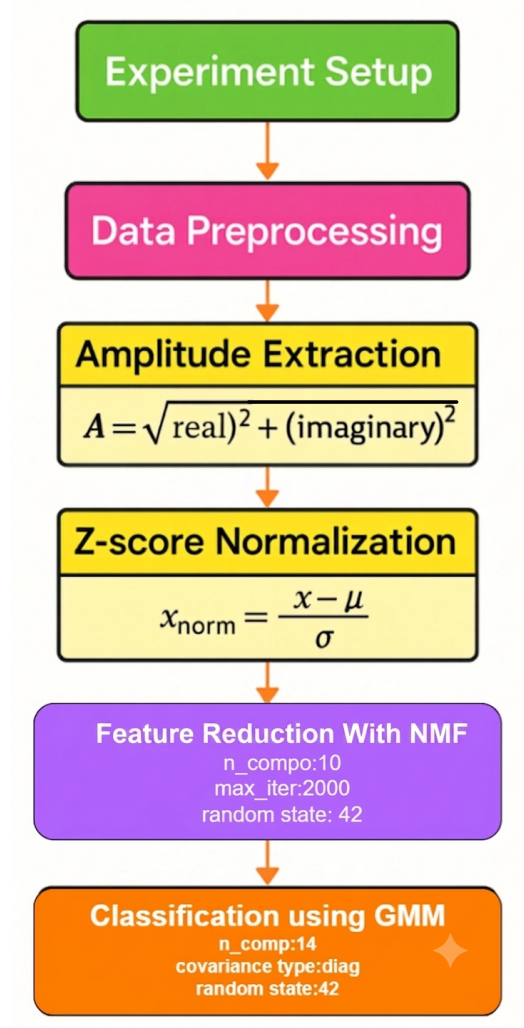
paper, PLS was identified as a key candidate technology to meet these evolving security requirements in this area, and the authors proposed a novel security system that leverages RNNs for channel prediction by utilising predicted CSI parameters as an authentication mechanism. Channel prediction is presented as an effective strategy to acquire up-to-date CSI, even in challenging conditions like channel ageing and mobility. By continuously tracking the wireless channel through predicted CSI, the system aims to prevent network attacks, such as spoofing, where illegitimate users attempt to impersonate legitimate communication. The accuracy results indicate that this method can generate alarms when spoofing is detected. Furthermore, the proposed approach is designed to complement existing upper-layer authentication protocols, thereby enhancing the overall security performance in wireless networks. This suggests a promising direction for improving wireless security in the context of 6G systems by integrating physical layer characteristics with advanced machine learning techniques. A channel-based PLA proposed in [28] leverages the randomness of wireless multipath fading channels. It predicts future channel impulse responses using RNNs, making it suitable for time series prediction and anomaly detection. The CGAN architecture features a generator for synthetic CSI metrics and a discriminator to identify legitimate versus fake metrics, using previous CSI magnitudes as conditional input. Various loss functions were explored, including binary cross-entropy (BCE) and the mean-squared error (MSE), with a hybrid approach to mitigate vanishing gradients. The results showed that the standalone RNNs achieved lower MSEs in channel prediction than the CGAN generators. However, a CGAN-trained discriminator with GRU cells and a hybrid loss function reached 98.5% accuracy in authentication, comparable to a standalone GRU network. The CGAN discriminator learned the boundaries for legitimate transmitters without a predefined error threshold. The GRU RNNs outperformed the LSTM RNNs in channel prediction, but the absence of false positives for illegitimate transmitters does not guarantee they will always be denied authentication.

The specific research gap addressed by the proposed work is the need for a physical layer authentication technique that achieves high accuracy and reliability in dynamic IoT networks while remaining computationally efficient and consuming minimal resources.

### 3. Proposed Methodology

The methodology used in this research, as shown in Figure 1, entails a systematic framework for the classification of authorised and unauthorised nodes using CSI data. The experimental configuration comprises four authorised nodes alongside one unauthorised node, from which CSI data are meticulously gathered under controlled experimental conditions utilising ESP32 microcontrollers. These unrefined data undergo an initial preprocessing phase that includes noise reduction, normalisation, and segmentation to maintain both quality and consistency. In order to mitigate the high dimensionality inherent in CSI data, NMF is employed, which facilitates the extraction of the most salient features while maintaining the non-negative characteristics of the signals. The resultant reduced feature set is subsequently integrated into a GMM, which capitalises on its probabilistic and soft clustering attributes to execute the classification of the data. The GMM's capability to allocate data points to multiple clusters with varying probabilities enables a nuanced analysis of CSI patterns, rendering it particularly adept at differentiating between legitimate and illegitimate transmissions within dynamic wireless environments. Furthermore, the proposed framework employs a GMM-based classification approach built around two distinct, class-specific models. One GMM is trained solely on legitimate (normal) traffic to capture its underlying statistical distribution, while the other is trained exclusively on malicious (attack) traffic to model adversarial behaviour. At inference time, each incoming data instance is evaluated against both trained models in parallel. The resulting likelihood

scores are then compared, and the instance is assigned to whichever class produces the higher posterior probability under its respective GMM.



**Figure 1.** Proposed methodology's flow graph.

### 3.1. Experiment Set-Up and Data Collection

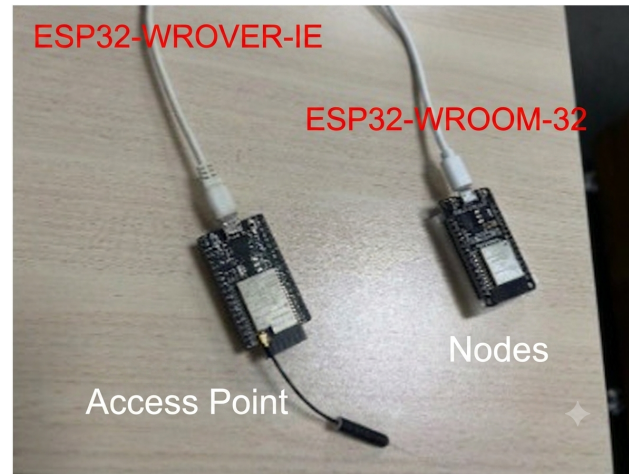
CSI embodies the detailed subcarrier-level channel response inherent in orthogonal frequency-division multiplexing (OFDM). It documents the variations in amplitude and phase as the frame transitions from the transmitter to the receiver through the frequency lattice [22,29]. To investigate this variation, an OFDM transmitter introduces known pilot tones, either interspersed within the payload or aggregated as a preamble, thereby enabling the receiver to deduce the channel vector  $H$  as CSI via the following equation:

$$y^{(i)} = H^{(i)} x^{(i)} + \eta^{(i)} \quad (1)$$

In this formulation,  $x^{(i)}$  denotes the  $i$ th pilot symbol transmitted over the channel,  $y^{(i)}$  represents the signal received at the antenna corresponding to that pilot, and  $\eta^{(i)}$  is the additive noise component affecting the transmission. The term  $H^{(i)}$  is a complex scalar coefficient characterising the channel for the  $i$ th subcarrier, with its real and imaginary parts jointly capturing the multipath propagation effects experienced by the signal.

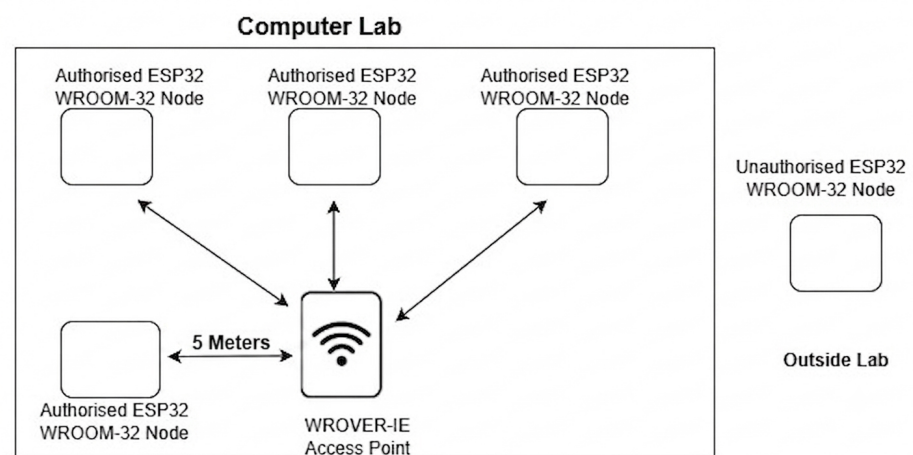
Our experimental apparatus consisted of two Espressif modules: ESP32-WROVER-IE, designated as the access point (AP), and ESP32-WROOM-32, designated as the client transmitters, as shown in Figure 2. Both modules were equipped with a dual-core Xtensa® LX6 CPU (240 MHz, approximately 600 MIPS) and provided Wi-Fi 802.11n at 2.4 GHz (with

a maximum throughput of  $150 \text{ Mb s}^{-1}$ ), supplemented by Bluetooth v4.2/BLE capabilities. The primary distinction between the two lies in their RF configurations: the WROVER-IE was equipped with an external u.FL connector, whereas the WROOM-32 was integrated with an onboard PCB antenna, and this architectural divergence facilitated modulation of the link budget without disrupting the silicon commonalities.



**Figure 2.** ESP32 modules used in the experiment.

CSI data were meticulously collected within the computer laboratory at Sichuan University using the ESP32-CSI-Tool [20]. Four sanctioned WROOM-32 nodes were strategically distributed throughout the laboratory, perpetually transmitting data to the WROVER-IE access point (AP). A fifth unauthorised node was positioned externally, simulating adversarial signals directed towards the same AP as presented in Figure 3.



**Figure 3.** Experiment set-up in lab explaining the positions of the nodes.

An individual, discrete transmission event was received by a node. Each received frame yielded exactly one raw CSI payload. Each acquired frame revealed 128 CSI coefficients, which corresponds to two complex taps per 64 subcarriers derived from the legacy long training field (LLTF), the Wi-Fi preamble segment designated for coarse synchronisation and channel estimation. To ensure statistical robustness, we recorded approximately one million CSI instances per node, culminating in an aggregate of  $\approx 5 \times 10^8$  complex samples. Each instance occupied 128 bytes, encoding magnitude-phase pairs that were essential for the subsequent feature extraction pipeline.

This extensive, high-resolution dataset constitutes the foundation for the training and validation of the proposed method in the following sections, thereby facilitating a rigorous distinction between authorised and unauthorised nodes in realistic indoor wireless environments.

### 3.2. Data Preprocessing

Raw CSI data, at their core, consist of complex-valued vectors that jointly characterise the coordinated interaction between amplitude and phase components, thereby providing a rich, multidimensional representation of the signal. Nonetheless, to simplify analysis and alleviate the burdensome dimensionality, these CSI datasets were transformed in a way that emphasised the amplitude rather than the phase. This choice was motivated by the fact that the phase component is prone to erratic fluctuations, which are commonly intensified by the chaotic effects of multipath propagation and environmental reflections. Even though phase information holds theoretical potential, its practical deployment faces numerous challenges. It is well known for its vulnerability to stochastic distortions driven by many sources; hardware imperfections, carrier frequency offsets, clock asynchrony, and the unpredictable behaviour of multipath fading all contribute to a setting in which rapidly changing environmental reflections disrupt the signal severely. These factors can produce abrupt, nonlinear phase shifts that resist stabilization, thereby often requiring extensive calibration and compensation. As a result, phase-based features are inherently noisy and unreliable, which complicates coherent modelling pursuits [30]. The amplitude component exhibits significantly greater resilience to such disturbances because it directly reflects signal power attenuation and channel gain characteristics, which remain comparatively stable across time and spatial variations. Furthermore, the decision to focus exclusively on the amplitude component of the CSI data was a deliberate optimization choice driven by two core factors: computational efficiency for edge deployment and inherent feature stability. First, implementing full-phase calibration loops introduces a heavy computational burden to the data preprocessing pipeline. Because our target framework prioritises high throughput and low latency, avoiding complex phase recovery transformations allows the overall system to remain highly lightweight. Second, the CSI amplitude component directly reflects physical signal power attenuation and spatial channel gain characteristics. Our empirical data demonstrates that these amplitude features remain comparatively stable across both temporal and spatial variations without requiring intensive mathematical compensation, thereby providing sufficient discriminative power on their own to achieve highly accurate classification. Consequently, the amplitude component exhibits significantly greater resilience to such disturbances because it directly reflects signal power attenuation and channel gain characteristics, which remain comparatively stable across time and spatial variations.

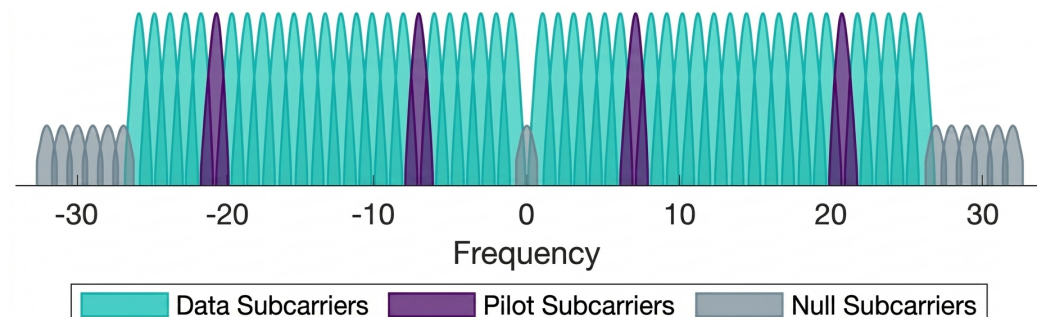
Following the computation of amplitude, the dataset undergoes dimensional refinement by eliminating null subcarriers, ensuring a more efficient feature set for analysis. This process culminates in a refined dataset that enhances the reliability of subsequent analyses and model training.

Formally, for each data point within the CSI framework, the amplitude  $A$  is calculated through the Euclidean norm of its complex components:

$$A = \sqrt{(\text{real})^2 + (\text{imaginary})^2}$$

where each CSI vector consists of 128 raw values that represent 64 subcarriers, with the corresponding real and imaginary components. The structure of the CSI data is presented in Figure 4. Following the computation of amplitude, the dataset undergoes dimensional refinement by eliminating null subcarriers and those channels that demonstrate near-zero

signal-to-noise ratios and therefore possess negligible informational value, culminating in a refined feature set comprising 53 amplitude attributes. The data subcarriers that convey authentic payloads, along with pilot subcarriers designated for channel estimation, are preserved to ensure the integrity of the CSI.



**Figure 4.** Layout of subcarrier types in the Wi-Fi frequency domain.

To alleviate the detrimental effects of ambient noise and interference, Z-score normalisation is employed on the amplitude features, thereby standardising the data distribution to a mean of zero and a variance of one, as illustrated by the equation

$$x_{\text{norm}} = \frac{x - \mu}{\sigma}$$

where  $x$  represents an amplitude feature,  $\mu$  denotes the population mean, and  $\sigma$  signifies the standard deviation. This normalisation process enhances the convergence characteristics and stability for ensuing analytical modelling.

The CSI data are not continuous time series, and thus the normalised dataset was methodically partitioned into training and testing subsets using an 80:20 random split ratio. This random splitting strategy was specifically chosen to align with the operational realities of the targeted network environment, which is characterised by highly dynamic, bursty, and fast-switching connection events. Because devices in this framework rapidly connect, authenticate via an instantaneous connections, and disconnect at sporadic intervals, consecutive CSI data instances do not exhibit a strong temporal correlation. Instead, each data instance represents a unique, statistically independent authentication attempt. Utilising a random split thus accurately simulated the system's real-world capability to verify devices based solely on the distinct spatial and multipath features of isolated connection events without relying on a prior sequence of events

This stratified partitioning promotes rigorous iterative optimisation and unbiased assessment.

### 3.3. Feature Selection via Non-Negative Matrix Factorization

To mitigate the elevated dimensionality and inherent redundancy frequently observed in raw CSI metrics, NMF was employed as the principal mechanism for feature extraction. In contrast to alternative decomposition methodologies, NMF facilitates a parts-based additive representation through the imposition of non-negativity constraints upon the factorised metrics. This inherent non-negativity proves particularly beneficial in the modelling of signal strength characteristics, which are intrinsically non-negative.

NMF represents a robust methodology that provides numerous benefits for the analysis of data. A significant characteristic of this technique is the imposition of a non-negativity constraint, which corresponds to the physical attributes of diverse datasets, thereby enhancing interpretability, particularly in contexts extending beyond physical layer authentication [31]. Moreover, NMF proficiently decreases the dimensionality of high-dimensional datasets while maintaining critical features, which not only augments computational ef-

iciency but also alleviates the risk of overfitting [32]. The methodology yields sparse representations that highlight the most pertinent characteristics, a factor that is essential for a variety of applications in the domain of machine learning [33]. In addition, NMF exhibits superior resilience to noise present in datasets when juxtaposed with conventional feature selection methodologies, rendering it suitable for practical applications frequently affected by environmental variables. The components derived from the NMF function as parts-based representations offer valuable insights into the fundamental structures of the data, a feature that is often absent in alternative feature extraction techniques [14]. Finally, NMF's flexibility enables its customisation for a variety of data types and applications, establishing it as a versatile instrument for various feature extraction contexts [34]. NMF feature extraction, compared with PCA, shows a superior feature representation capability at lower feature dimensions, particularly for attack categories with scarce or localised features, and this leads to faster convergence speeds during the training process of abnormal detection methods, outperforming both PCA feature extraction and direct training with raw data [35,36].

Let  $X \in \mathbb{R}^{m \times n}$  denote the CSI measurement matrix, where  $m$  is the number of observations and  $n$  is the number of subcarriers or spatial dimensions. NMF seeks to approximate  $X$  through the product of two lower-rank non-negative metrics:

$$X \approx WH$$

where  $W \in \mathbb{R}^{m \times k}$  and  $H \in \mathbb{R}^{k \times n}$ , with  $k \ll \min(m, n)$  representing the latent dimension. In this study,  $k$  was empirically set to 10, balancing model compactness with the preservation of salient signal features.

The factor matrix  $W$ , which encodes the low-dimensional projection of the original CSI data, serves as the discriminative input feature set for the subsequent classification task. Its structure encapsulates the dominant components of signal variation while eliminating noise and redundancy, thus facilitating efficient and interpretable modelling for physical layer authentication.

The decomposition process was executed using the scikit-learn NMF implementation, configured with a maximum of 2000 iterations, a random initialisation seed of 42, and a number of components as 10 to ensure reproducibility across experimental trials. The values of the hyperparameters were determined through iterative runs of the algorithm, as this value yielded the best results in terms of the employed metrics.

### 3.4. Classification Using Gaussian Mixture Model

Subsequent to the process of dimensionality reduction, the resultant latent features are subjected to probabilistic modelling by utilising a Gaussian mixture model. A GMM offers numerous advantages in comparison with alternative algorithms for the classification of physical layer authentication utilising CSI data. One principal advantage is their adaptability to representing intricate distributions; a GMM can proficiently amalgamate multiple Gaussian components to conform to the heterogeneous characteristics of CSI data, as elucidated in [37]. Furthermore, GMMs demonstrate resilience to noise, which is essential for practical applications where data may be susceptible to various interferences; this proficiency in preserving data integrity under adverse conditions was underscored in [38]. The soft clustering attribute of GMMs permits data points to be affiliated with multiple clusters with differing probabilities, thereby enabling more refined interpretations and enhancing classification precision in physical layer authentication endeavours [39]. Moreover, GMMs are scalable and can adeptly manage extensive datasets, a critical necessity in physical layer authentication scenarios [40], highlighting the efficacy of the expectation-maximization algorithm in high-dimensional CSI data. This model operates on the premise that the

observed data originate from a superposition of multiple multivariate Gaussian distributions. Such a modelling approach is particularly effective in discerning the multimodal characteristics and inherent stochastic variability present within wireless channel states, distinguishing between authorised and unauthorised nodes. The reduced features, denoted as  $\mathbf{W} \in \mathbb{R}^{m \times k}$ , are formally modelled by the following probability density function:

$$p(\mathbf{x}) = \sum_{i=1}^K \pi_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$$

where  $\pi_i$  are the mixing coefficients,  $\boldsymbol{\mu}_i$  are the mean vectors, and  $\boldsymbol{\Sigma}_i$  are the diagonal covariance metrics corresponding to the  $i$ th Gaussian component. In this study, the number of components  $K$  was determined to be 14 through iterative runs of the algorithm, with  $K = 14$  yielding the best results in terms of the employed metrics in this work.

The imposition of a diagonal covariance structure was implemented to streamline computational complexity and mitigate the risk of overfitting, particularly given the moderate dimensionality of the features. The classifier subsequently yielded posterior probabilities which quantified the likelihood of a given sample's affiliation with a specific class (i.e., authorised or unauthorised).

The classifier output posterior probabilities that indicate the likelihood of a sample belonging to a specific class (authorised or unauthorised), using  $M_0$  to denote the GMM trained exclusively on legitimate (authorised) traffic distributions and  $M_1$  to denote the GMM trained on malicious (unauthorised) traffic distributions. For an incoming, dimensionally reduced CSI feature vector  $\mathbf{x}$ , the system computes the log-likelihood under both models in parallel using

$$\log L(\mathbf{x}|M_k) = \sum_{i=1}^N \log P(\mathbf{x}_i|M_k), \quad k \in \{0, 1\}$$

The final binary authentication decision  $\hat{y}$  is determined strictly using a maximum log-likelihood selection rule (equivalent to a maximum a posteriori decision with equal prior probabilities):

$$\hat{y} = \arg \max_{k \in \{0, 1\}} (\log L(\mathbf{x}|M_0), \log L(\mathbf{x}|M_1))$$

where:  $\hat{y} = 0$  maps directly to the unauthorised class if  $\log L(\mathbf{x}|M_1) > \log L(\mathbf{x}|M_0)$  and  $\hat{y} = 1$  maps directly to the authorised class if  $\log L(\mathbf{x}|M_0) \geq \log L(\mathbf{x}|M_1)$ .

The integrated methodology not only exploits the inherent nonlinear and probabilistic characteristics of CSI data but also upholds computational efficiency, rendering it highly suitable for implementation in resource-constrained IoT wireless networks, ultimately contributing to the enhancement of node authentication and the robustness of security measures.

## 4. Results and Discussion

The utilization of computational hardware resources by an algorithm plays a pivotal role concerning performance in practical applications, as exemplified in [41–45]. The utilization of computational hardware resources by an algorithm plays a crucial role in determining performance in practical applications, as exemplified in [41]. This becomes especially significant for artificial intelligence (AI) algorithms as they learn and train from the data fed to them, and based on those, they perform the classification of authorised and unauthorised nodes. The efficient use of hardware resources is crucial to optimising the performance of AI algorithms, especially when dealing with large datasets and complex computations in real time, where models are trained with new data periodically [46,47]. Efficient utilisation of hardware not only improves performance but also contributes to

energy savings, which makes it essential for the deployment of AI in resource-constrained environments [48]. The system configuration utilised for this research paper was based on a high-performance device, specifically the L-W-CAT-5031BFX, equipped with a 13th Generation Intel(R) Core(TM) i9-13950HX processor operating at 2.20 GHz with 32.0 GB of installed RAM (31.5 GB usable). This configuration supports a 64-bit operating system on an x64-based processor, ensuring efficient processing capabilities for computational tasks. The research leveraged Python 3.12.7 as the primary programming language, utilising the Scikit-learn library for machine learning tasks, which provides robust tools for data analysis and model evaluation. Additionally, TensorFlow was used for deep learning applications, enabling the implementation of complex neural networks and facilitating advanced model training. Furthermore, the experiments for both our proposed model and the CNN benchmark were executed under a strictly isolated environment with all non-essential OS daemons, background processes, and network interfaces disabled.

For the evaluation of resource consumption, we measured five important hardware resources [49]:

1. CPU usage: The utilisation of the central processing unit (CPU) was quantified as the proportion of the processor's comprehensive available time that was effectively employed by the program throughout its execution period. This metric was derived by taking the sum of the total CPU time (comprising both user time and system time) and dividing it by the total elapsed wall clock time, subsequently multiplying the result by 100 to represent it in percentage form. This particular metric serves as an indicator of the degree to which the CPU was actively engaged in processing tasks:

$$\begin{aligned}\text{CPU Usage} &= \left( \frac{\text{CPU Time}}{\text{Wall-clock Time}} \right) \times 100 \\ &= \left( \frac{T_{\text{user}} + T_{\text{system}}}{T_{\text{wall}}} \right) \times 100\%\end{aligned}$$

2. Total CPU time: The CPU time represents the cumulative duration that the CPU allocates to the execution of a program's directives, encompassing both user-mode and system-mode functionalities. This metric encapsulates the genuine computational effort expended in the execution of the code, thereby omitting intervals of inactivity or waiting. The CPU time was determined by aggregating both the user and system time:

$$\text{CPU Time} = T_{\text{user}} + T_{\text{system}}$$

3. Average memory usage [43]: The mean memory usage represents the average memory consumption throughout the execution duration. This metric was derived by performing multiple assessments of memory utilisation during the execution phase and subsequently computing the arithmetic mean of these observations. This approach provides a comprehensive understanding of the use of sustained memory:

$$\text{Average Memory Usage} = \frac{1}{N} \sum_{i=1}^N \text{Memory Usage}(t_i)$$

where  $N$  is the number of samples taken during the runtime.

4. Peak memory usage: Peak memory usage denotes the largest memory allocation documented throughout the execution of the programme. Employing tracemalloc, it quantifies the peak volume of memory used at any given instant, reflecting the most extreme memory requirements imposed by the programme on the system:

$$\text{Peak Memory Usage} = \max_{t \in [0, T]} (\text{Memory Usage}(t))$$

#### 5. Estimated RAM usage [44]:

This is a calculation relating to the amount of data that are processed and the efficiency of the algorithms employed, often expressed through a general formula applicable in machine learning or data processing tasks:

$$\text{RAM Usage} = n \times d \times m$$

where  $n$  is the number of data samples (instances),  $d$  is the number of features (attributes) per sample, and  $m$  is the memory overhead (constant factor representing additional memory usage due to the algorithm, data structures, etc.)

This formula provides a rough estimate of the memory required to hold the dataset in memory, along with any additional overhead associated with processing the data [50]. It is important to distinguish this metric from the average memory usage reported during execution. While the estimated RAM isolates the intrinsic, mathematical footprint of the algorithm itself ( $n \times d \times m$ ), the average memory usage captures the dynamic, empirical system consumption during live runtime. Consequently, the average memory usage includes additional real-world system overheads—such as loading the execution framework and libraries, managing high-throughput data I/O buffering, and generating transient runtime variables—which are naturally absent from the theoretical formula.

#### 6. Estimated energy consumption:

The energy consumption associated with the proposed algorithm denotes the aggregate electrical energy necessary to execute the algorithm on a specified hardware platform. This is conventionally quantified in joules and computed as the integral of power over a designated time interval. For pragmatic considerations, it can be approximated by using

$$E = P \times t$$

where  $E$  is the energy consumed,  $P$  is the average power drawn (in watts), and  $t$  is the execution time (in seconds). Power  $P$  may be further derived from voltage and current measurements, where

$$P = V \times I$$

This metric offers critical insights into the energy efficiency of the algorithm and is essential to assess performance, particularly in energy-constrained environments such as embedded systems, mobile devices, or expansive cloud infrastructure [45].

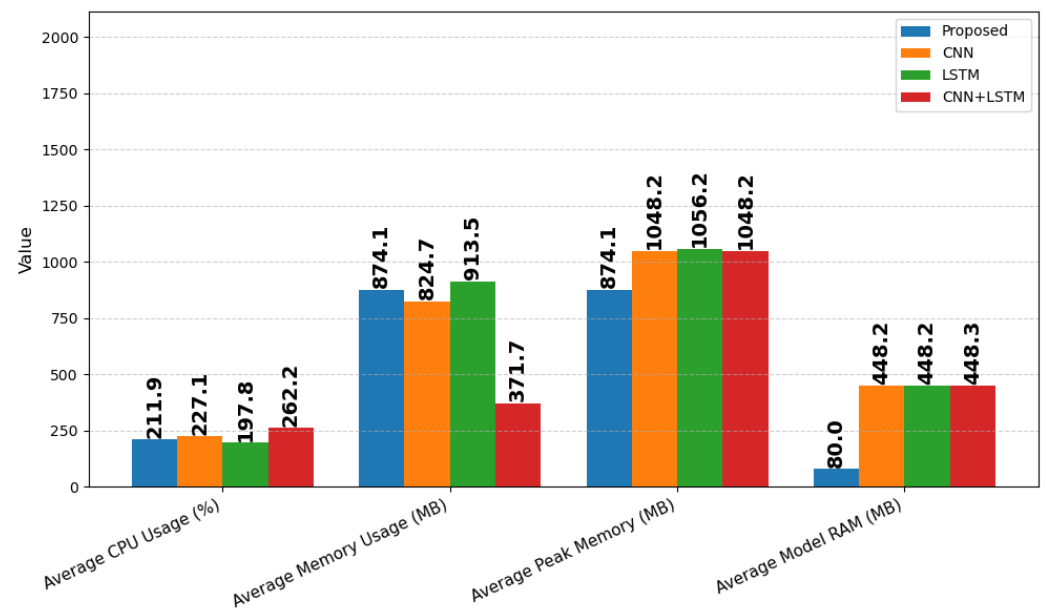
An empirical assessment of the computational resource requirements was performed during the training phase in four distinct architectures: the proposed method, a CNN [51], a long short-term memory (LSTM) model, and a hybrid CNN + LSTM model [52].

In this experimental study, we carefully trained several deep learning architectures, specifically the CNN, LSTM model, and our novel CNN-LSTM hybrid model, using a single, consistent set of environmental parameters. The CNN architecture consisted of a one-dimensional convolutional layer with 32 filters, a kernel size of three, and ReLU activation. This was followed by a max pooling layer (MaxPooling1D) with a pooling size of two, a dropout layer with a rate of 0.2, a flattening layer, and a fully connected dense layer with 64 units and ReLU activation. The network concluded with a final two-unit dense layer that applied Softmax activation designed for binary classification tasks. In contrast, the LSTM model was implemented with an LSTM layer containing 64 hidden units that processed sequences using `return_sequences = False`. After this, the model used a dropout layer (rate = 0.2), a dense hidden layer with 64 units and ReLU activation, and finally a two-unit dense layer with

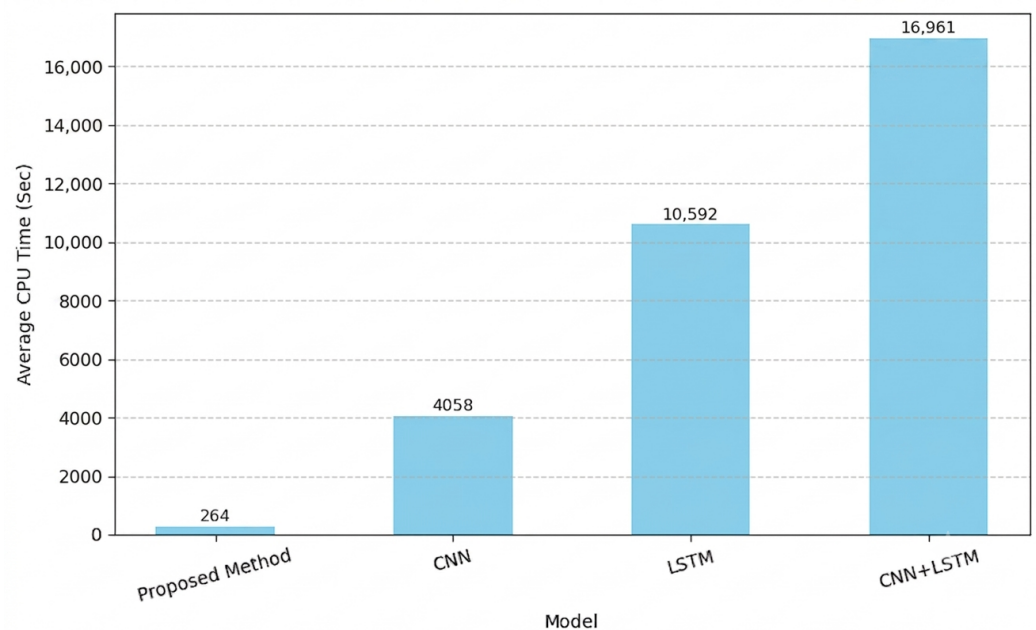
Softmax activation. The CNN + LSTM model provided an integrated approach that combines spatial feature extraction with the capacity for sequential modelling. It started with a Conv1D layer (64 filters, kernel size = 3, ReLU activation) and then applied a MaxPooling1D layer (pool size = 2) and a dropout layer (rate = 0.3). The reduced features were then passed to an LSTM layer (64 hidden units, *return\_sequences* = *False*, followed by another dropout layer (rate = 0.3), a dense hidden layer with 64 units, and a final two-unit Softmax output layer. Each model underwent careful optimization with the Adam optimiser, using a fixed batch size of 32 across 25 epochs. Importantly, the CNN + LSTM model used a dropout rate of 0.3, whereas the CNN and LSTM models used a dropout rate of 0.2; all configurations were chosen to enhance strong generalization. The input tensors were consistently normalised and reshaped to (N, features, 1), preserving data integrity throughout this comparative investigation, where N represents the batch size. The results from this thorough evaluation are shown in the accompanying figure. The principal goal was to assess the models' efficiency and scalability by examining key metrics such as the CPU utilization, training time, memory usage, and energy consumption, considerations that are especially critical in real-world deployment contexts, where resource limitations frequently present themselves.

The comparison of average CPU usage and average CPU time for each model during the training phase, as presented in Figures 5 and 6, respectively, highlights significant differences in their computational demands and training efficiency. The Hybrid CNN and LSTM models exhibited the highest average CPU usage at 262.24% and a lengthy training duration of 16,960.56 s, indicating their substantial computational requirements. The CNN model trailed closely at 227.12% with a training time of 4057.84 s, primarily due to its operations being parallelised, effectively capitalising on the multicore processing capabilities. The proposed method, designed for efficiency, showed an average CPU usage of 211.91%, likely resulting from the integration of custom regularisation techniques and dynamic training components; however, it excelled in terms of training duration, completing the process in just 264.48 s. Finally, the LSTM model, which is defined by its sequential data dependencies and the absence of inherent parallelism, had the lowest average CPU usage at 197.75% but faced an extensive training time of 10,591.95 s, reflecting the trade-off between its lower computational load and longer training duration. Overall, while the hybrid model had the highest CPU usage and longest training time, the proposed method achieved a remarkable balance between efficiency and reduced training duration.

The memory footprint plays a crucial role in assessing the scalability of a model, particularly within edge or resource-limited environments, and the obtained memory from the conducted experiment is presented in Figure 5. Although the proposed method exhibited a comparatively elevated average memory usage of 874.13 MB, its maximum memory consumption remained consistent at the same figure, suggesting stable memory allocation and effective utilisation without significant fluctuations. Conversely, both the CNN and CNN + LSTM models attained peak memory utilizations exceeding 1048 MB, which indicates the presence of extensive intermediate activation maps and hierarchical feature representations. The LSTM model, which requires the retention of temporal hidden states across successive time steps, similarly demonstrated a peak memory utilisation of 1056.15 MB. Notably, the CNN + LSTM model exhibited an unexpectedly low average memory usage of 371.68 MB, potentially attributable to proactive memory management strategies between layers or batched computations involving smaller input windows.

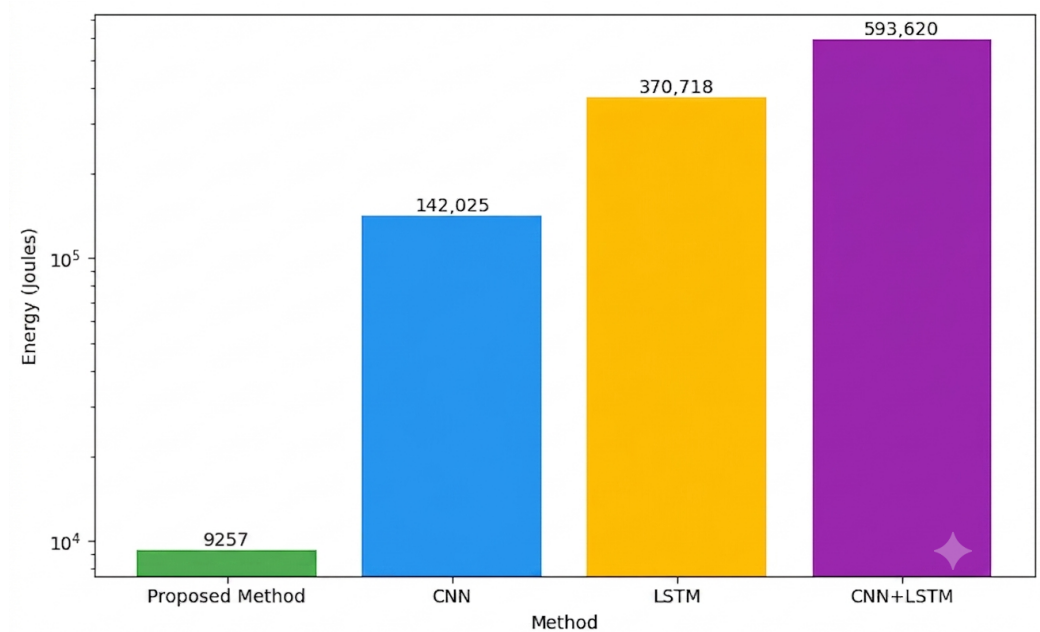


**Figure 5.** Comparison of computer resource consumption on training data for proposed method, CNN, LSTM, and CNN + LSTM models.



**Figure 6.** Comparison of CPU time consumption for proposed method, CNN, LSTM, and CNN + LSTM models.

The energy consumption was inferred from the power profiles of the CPU and the duration of training. The proposed method exhibited energy consumption of only 9256.95 J, which reinforces its character as consuming a low amount of resources and being energy-efficient, as shown in Figure 7. This contrasts sharply with CNN + LSTM, which exhibited energy expenditure of an astounding 593,619.69 J, followed closely by LSTM at 370,718.36 J. The CNN model required 142,024.53 J, thus reaffirming the notion that deeper or hybrid architectures, while potentially yielding improved accuracy, impose substantial cost. These findings are of great significance in situations where energy limitations are crucial, particularly in mobile health monitoring and IoT systems.



**Figure 7.** Comparison of energy consumption for proposed method, CNN, LSTM, and CNN + LSTM models.

The proposed method also demonstrated exceptional memory efficiency, using only 80.02 MB of RAM, in contrast to 448.2 MB for the CNN and LSTM models and 448.26 MB for the CNN + LSTM model, as presented in Figure 5. This reduction implies a lesser number of trainable parameters, optimised layer architecture, and minimal overhead in internal data structures. It also highlights the model's relevance for implementation on devices with limited RAM, particularly embedded systems.

The comparative analysis of resource utilisation distinctly demonstrates that the proposed method attained enhanced training efficiency, characterised by markedly reduced training duration, diminished energy expenditure, and negligible memory overhead, all while sustaining competitive performance indicators. Although more intricate models, such as CNN + LSTM, may yield slightly superior predictive accuracy, they incur substantially greater computational and energy costs. These conclusions significantly favour the application of the proposed method in cases that require elevated precision and computational resourcefulness.

We compared the results of the proposed method with the CNN, LSTM, and CNN + LSTM models on unseen test data. To evaluate the performance of the proposed model, the following metrics were exploited:

1. Accuracy:

The metric of accuracy is defined as the ratio of instances that have been accurately predicted (including both true positives and true negatives) to the aggregate number of instances. This metric serves to assess the comprehensive efficacy of a classification model by reflecting the ratio of total correct predictions to the entirety of the predictions executed:

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

where *TP* is true positives, *TN* is true negatives, *FP* is false positives, and *FN* is false negatives.

2. Precision:

Precision refers to the proportion of accurately predicted positive instances (true positives) relative to the aggregate of the predicted positive instances (the total of true

positives and false positives). This metric evaluates the correctness of positive predictions, addressing the inquiry, “Among all instances classified as positive, how many are genuinely positive?”:

$$\text{Precision} = \frac{TP}{TP + FP}$$

### 3. Recall:

Recall, frequently referred to as sensitivity or the true positive rate, represents the proportion of accurately identified positive instances relative to the total number of actual positive instances (which is the aggregate of true positives and false negatives). This metric evaluates the efficacy of a model in locating all pertinent positive cases, addressing the inquiry, “Of the total true positive instances, how many were recognised by the model?”:

$$\text{Recall} = \frac{TP}{TP + FN}$$

### 4. F1 score:

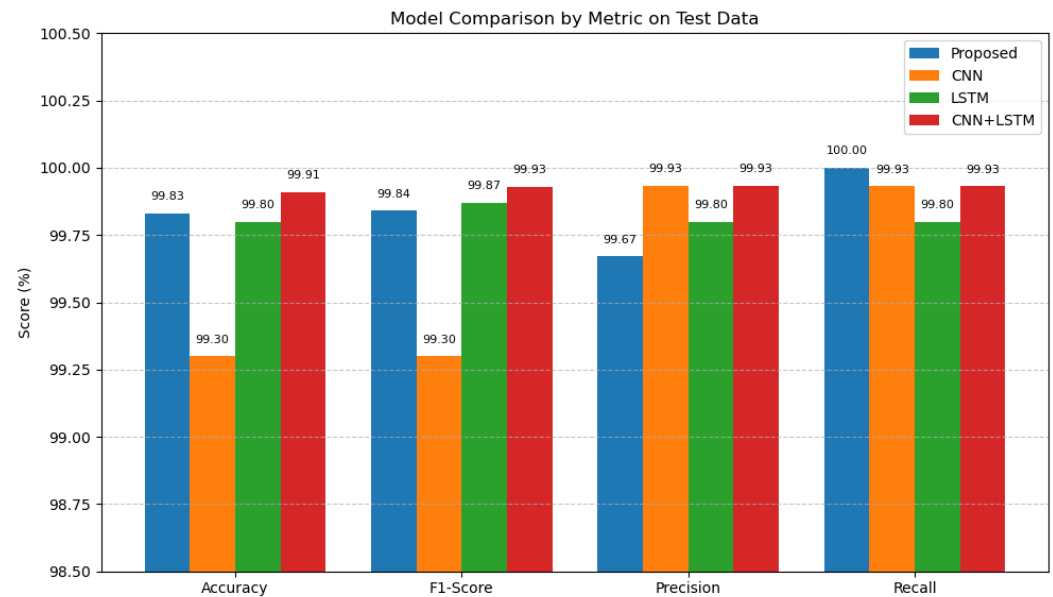
The F1 score represents the harmonic mean of precision and recall, thereby offering a singular metric that effectively harmonises both dimensions. Its application is particularly advantageous when one aims to reconcile the trade-off between precision and recall or when addressing the challenges associated with imbalanced datasets:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

An in-depth review of the four structures—the proposed method, CNN, LSTM, and CNN + LSTM—carried out on the dataset of the assessment test demonstrates slight variations in performance statistics, specifically in relation to precision, recall, and F1 score, as shown in Figure 8. The proposed method exhibited exceptional classification proficiency, attaining an accuracy of 99.83% and a flawless recall of 100%, which is of paramount importance in critical fields such as cybersecurity and anomaly detection, where failure to identify a threat (false negatives) can yield far more detrimental consequences than generating false positives.

Although it is correct that the CNN and CNN + LSTM models demonstrated a slight advantage in precision, being 99.932% compared with 99.67% for the proposed method, this apparently minor trade-off was more than offset by the proposed method’s notable lack of false negatives, as reflected in its strong recall metric. In addition, the F1 score of 99.84% indicates a solid equilibrium between precision and recall, making it a highly competitive option relative to the hybrid CNN + LSTM model (99.93%) and the LSTM model (99.87%).

Significantly, what distinguished the proposed method from its counterparts was not solely its elevated detection capabilities but also its enhanced computational efficiency throughout the training phase. The model recorded a mere 264.48 s of CPU time with an average CPU utilisation of 211.91%, markedly less than the CNN + LSTM model’s training duration, which exceeded 16,000 s. It also showcased superior memory management, using an average of 874.13 MB of memory with a peak at the same value, alongside an estimated model RAM footprint of only 80.02 MB, which was approximately 5.6 times smaller than the memory demands of the CNN + LSTM model (448.26 MB). Furthermore, the expected energy requirement for the training of the proposed method was just 9256.95 J, illustrating its exceptional efficiency compared with the 593,619.69 J utilised by the CNN + LSTM framework.



**Figure 8.** Comparison of results on test data of proposed method with CNN, LSTM, and CNN + LSTM models.

The proposed method is distinguished by its substantial benefits for the training time, being characterised by minimal CPU usage, lower memory consumption, and a reduced energy footprint. As a result, it serves not only as a high-performing classifier for previously unseen data but also as a resource-efficient option. These advantages are especially important during real-world deployment, particularly when computational resources are constrained or when rapid processing is required. In summary, although the CNN + LSTM model may deliver marginally stronger outcomes in composite test metrics, the proposed method offers a well-regarded trade-off between performance and computational cost. Supported by outstanding recall, a near-optimal accuracy and F1-score, and a less resource-intensive training process, it is well suited to environments in which reliability, speed, and efficiency in detection are critical.

The proposed method was further compared to other state-of-the-art methods, as presented in Table 1. The thorough investigation demonstrates the greater effectiveness of the proposed method when applied to wireless authentication leveraging CSI. Among the methodologies evaluated, the proposed method achieved the highest documented precision of 99.83%, surpassing a variety of deep learning and machine learning frameworks. For example, although VGG11 [10] establishes a robust image-centric CSI authentication framework and achieves over 90% accuracy through sophisticated augmentation and loss strategies, its performance remains conspicuously inferior to that of the proposed method. In addition, the CNN-RNN blended configuration [19], which combines spatial and spectral learning to refine channel representation, achieved an accuracy of 99.7%, slightly lower than our recommended strategy.

The CNN + LSTM model [53] achieved an accuracy of 99.9%, which was closest to the proposed method; however, it frequently incurred elevated computational expenses and complexity as a result of its hybrid architecture. In contrast, the proposed method maintained its competitive performance while embodied in a more efficient design. The GMM [54], which is based on statistical modelling using Euclidean and correlation characteristics, achieved a detection probability of 97%, a figure markedly lower and less adaptable compared with deep learning methodologies. Similarly, ELM [55], a rapid learning algorithm refined through AdaBoost, reported an accuracy of 98.91% with fewer than

900 training samples, but it confronted challenges in performance consistency attributable to the initialisation of random weights.

**Table 1.** Performance comparison of different methods based on accuracy and probability of detection.

| Methods         | Value    | Matrix                   |
|-----------------|----------|--------------------------|
| VGG11 [10]      | Over 90% | Accuracy                 |
| CNNRNN [19]     | 99.7     | Accuracy                 |
| CNN + LSTM [53] | 99.9     | Accuracy                 |
| GMM [54]        | 97%      | Probability of detection |
| ELM [55]        | 98.91    | Accuracy                 |
| CSI-PUF [56]    | 97.17    | Accuracy                 |

Moreover, the CSI-PUF approach [56], which circumvents traditional classification by depending on unique hardware-based identifiers, achieved a maximum accuracy of 97.17%, yet it exhibited a deficiency in adaptability to novel scenarios and failed to facilitate classification-based learning. Within this context, the proposed method differentiates itself by effectively balancing high accuracy, generalisation, and computational efficiency, thereby positioning itself as a state-of-the-art solution for CSI-based device authentication in dynamic wireless environments.

## 5. Conclusions

In conclusion, the proposed physical layer authentication method leveraging CSI marks a significant advancement in securing and streamlining authentication processes for IoT devices equipped with Wi-Fi capabilities. By harnessing the distinct characteristics of CSI data, the proposed innovative approach effectively distinguishes between authorised and unauthorised nodes, successfully addressing the challenges presented by dynamic wireless environments. The extensive experimental validation demonstrates the method's exceptional performance, achieving an impressive accuracy rate of 99.83% alongside a flawless recall of 100%. This reliability is crucial for critical applications such as cybersecurity and anomaly detection. Moreover, the method demonstrated remarkable computational efficiency, characterised by minimal CPU utilization, a reduced training duration, and significantly lower energy consumption compared with more complex architectures like the CNN and CNN + LSTM models. This advantageous balance between high performance and resource efficiency makes the proposed methodology particularly well suited for deployment in environments with limited computational resources. The ESP32 platform, known for its mobility and repositioning capabilities, enhances data collection, enabling scenarios that were previously unattainable with conventional equipment. However, it is important to note that the ESP32's limited computational power may affect the accuracy based on the node's location, presenting a challenge for the proposed methodology. Overall, this research contributes significantly to the evolving landscape of secure IoT communications, offering a robust solution that enhances both security and operational efficiency in real-world applications. Future research endeavors will focus on transitioning the proposed model to facilitate real-time online inference in resource-constrained environments, particularly targeting platforms like the ESP32 and other microcontroller units (MCUs). This deployment pipeline can leverage TensorFlow Lite Micro and post-training quantization (PTQ) to compress models into an optimised eight-bit integer format. By evolving the system from a high-performance computational baseline to a quantised MCU realization, we can achieve autonomous, low-latency physical layer security at the forefront of IoT sensing. Looking ahead, future research will also evaluate the model's robustness against a wider array of spoofing scenarios, including multiple attacker locations, varying

distances, heterogeneous hardware devices, and dynamic environments. Additionally, the model's adaptability to dynamic environmental changes, such as human movement, furniture rearrangements, and device repositioning, will be thoroughly assessed to ensure its real-world scalability and effectiveness.

**Author Contributions:** Conceptualization, M.R. and Y.R.; methodology, M.R.; validation, M.R., J.C., and S.P.; formal analysis, S.P. and J.C.; investigation, S.P. and J.C.; data curation, M.R. and Y.R.; writing—original draft preparation, M.R., S.P., and J.C.; writing—review and editing, M.R., S.P., and J.C.; supervision, J.C. and S.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** Data are available at <https://iee-dataport.org/documents/csi-datasets> (accessed on 3 June 2026).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Javanmardi, S.; Shojafar, M.; Mohammadi, R.; Alazab, M.; Caruso, A.M. An SDN perspective IoT-Fog security: A survey. *Comput. Netw.* **2023**, *229*, 109732. [\[CrossRef\]](#)
2. Rana, M.; Mamun, Q.; Islam, R. Lightweight cryptography in IoT networks: A survey. *Future Gener. Comput. Syst.* **2022**, *129*, 77–89. [\[CrossRef\]](#)
3. Garg, P.; Singh, D.K. Analysis of cryptographic encryption algorithm design to Secure IoT Devices: A review. *Mater. Today Proc.* **2022**, *51*, 810–814. [\[CrossRef\]](#)
4. Rana, S.; Parast, F.K.; Kelly, B.; Wang, Y.; Kent, K.B. A comprehensive survey of cryptography key management systems. *J. Inf. Secur. Appl.* **2023**, *78*, 103607. [\[CrossRef\]](#)
5. Rao, P.M.; Deebak, B.D. A comprehensive survey on authentication and secure key management in internet of things: Challenges, countermeasures, and future directions. *Ad Hoc Netw.* **2023**, *146*, 103159. [\[CrossRef\]](#)
6. El-Hajj, M.; Mousawi, H.; Fadlallah, A. Analysis of lightweight cryptographic algorithms on IoT hardware platform. *Future Internet* **2023**, *15*, 54. [\[CrossRef\]](#)
7. Ding, Z.; He, D.; Qiao, Q.; Li, X.; Gao, Y.; Chan, S.; Choo, K.K.R. A lightweight and secure communication protocol for the IoT environment. *IEEE Trans. Dependable Secur. Comput.* **2023**, *21*, 1050–1067. [\[CrossRef\]](#)
8. Ran, Y.; Al-Shwailly, H.; Tang, C.; Tian, G.Y.; Johnston, M. Physical layer authentication scheme with channel based tag padding sequence. *IET Commun.* **2019**, *13*, 1776–1780. [\[CrossRef\]](#)
9. Shi, C.; Liu, J.; Liu, H.; Chen, Y. WiFi-enabled user authentication through deep learning in daily activities. *ACM Trans. Internet Things* **2021**, *2*, 13. [\[CrossRef\]](#)
10. Wang, S.; Huang, K.; Xu, X.; Zhong, Z.; Zhou, Y. CSI-based physical layer authentication via deep learning. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 1748–1752. [\[CrossRef\]](#)
11. Jiang, P.; Wu, H.; Xin, C. A channel state information based virtual MAC spoofing detector. *High-Confid. Comput.* **2022**, *2*, 100067. [\[CrossRef\]](#)
12. Pecoraro, G.; Di Domenico, S.; Cianca, E.; De Sanctis, M. CSI-based fingerprinting for indoor localization using LTE signals. *EURASIP J. Adv. Signal Process.* **2018**, *2018*, 49. [\[CrossRef\]](#)
13. Zhang, J.; Ardizzon, F.; Piana, M.; Shen, G.; Tomasin, S. Physical Layer-Based Device Fingerprinting for Wireless Security: From Theory To Practice. *IEEE Trans. Inf. Forensics Secur.* **2025**, *20*, 5296–5325. [\[CrossRef\]](#)
14. Friedman, H.; Maina-Kilaas, A.R.; Schalkwyk, J.; Ahmed, H.; Haddock, J. An Interpretable Joint Nonnegative Matrix Factorization-Based Point Cloud Distance Measure. In *Proceedings of the 2023 57th Annual Conference on Information Sciences and Systems (CISS)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
15. Custance, O.; Khan, S.; Parkinson, S. Transformer Network-based Gait Identification using WiFi. *IEEE Trans. Biom. Behav. Identity Sci.* **2025**, *8*, 311–326. [\[CrossRef\]](#)
16. Custance, O.; Khan, S.; Parkinson, S.; Sheng, Q.Z. Parameter-Efficient Domain Adaption for CSI Crowd-Counting via Self-Supervised Learning with Adapter Modules. *arXiv* **2026**, arXiv:2601.02203.
17. Custance, O.; Khan, S.; Parkinson, S. Why Commodity WiFi Sensors Fail at Multi-Person Gait Identification: A Systematic Analysis Using ESP32. *arXiv* **2026**, arXiv:2601.02177.

18. Vogt, H.; Li, C.; Sezgin, A.; Zenger, C. On the precise phase recovery for physical-layer authentication in dynamic channels. In *Proceedings of the 2019 IEEE International Workshop on Information Forensics and Security (WIFS)*; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
19. Wang, Q.; Li, H.; Zhao, D.; Chen, Z.; Ye, S.; Cai, J. Deep Neural Networks for CSI-Based Authentication. *IEEE Access* **2019**, *7*, 123026–123034. [[CrossRef](#)]
20. Hernandez, S.M.; Bulut, E. Lightweight and Standalone IoT Based WiFi Sensing for Active Repositioning and Mobility. In *Proceedings of the 2020 IEEE 21st International Symposium on “A World of Wireless, Mobile and Multimedia Networks” (WoWMoM)*, Cork, Ireland, 31 August–3 September 2020; pp. 277–286. [[CrossRef](#)]
21. Xie, N.; Chen, J.; Huang, L. Physical-Layer Authentication Using Multiple Channel-Based Features. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 2356–2366. [[CrossRef](#)]
22. Hernandez, S.M.; Bulut, E. WiFi Sensing on the Edge: Signal Processing Techniques and Challenges for Real-World Systems. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 46–76. [[CrossRef](#)]
23. Jing, T.; Huang, H.; Gao, Q.; Wu, Y.; Huo, Y.; Wang, Y. Multi-User Physical Layer Authentication Based on CSI Using ResNet in Mobile IIoT. *IEEE Trans. Inf. Forensics Secur.* **2024**, *19*, 1896–1907. [[CrossRef](#)]
24. Han, J.; Li, Y.; Liu, G.; Ma, J.; Zhou, Y.; Fang, H.; Wu, X. Model-Driven Learning for Physical Layer Authentication in Dynamic Environments. *IEEE Commun. Lett.* **2024**, *28*, 572–576. [[CrossRef](#)]
25. Wang, Q.; Liang, W.; Zhang, J.; Wang, K.; Jiang, X. Knowledge-Enhanced Physical Layer Authentication for Mobile Devices. *IEEE Trans. Consum. Electron.* **2024**, *70*, 7436–7448. [[CrossRef](#)]
26. Kaseris, M.; Kostavelis, I.; Malassiotis, S. A comprehensive survey on deep learning methods in human activity recognition. *Mach. Learn. Knowl. Extr.* **2024**, *6*, 842–876. [[CrossRef](#)]
27. Martins, J.; Gomes, M.; Silva, V.; Dinis, R. Deep learning-based channel prediction for wireless physical layer security. In *Proceedings of the 2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*; IEEE: Piscataway, NJ, USA, 2024; pp. 114–118.
28. Germain, K.S.; Kragh, F. Channel prediction and transmitter authentication with adversarially-trained recurrent neural networks. *IEEE Open J. Commun. Soc.* **2021**, *2*, 964–974. [[CrossRef](#)]
29. Luo, F.; Khan, S.; Jiang, B.; Wu, K. Vision Transformers for Human Activity Recognition Using WiFi Channel State Information. *IEEE Internet Things J.* **2024**, *11*, 28111–28122. [[CrossRef](#)]
30. Du, L.; Shang, S.; Zhang, L.; Li, C.; Yang, J.; Tian, X. Multidomain Correlation-Based Multidimensional CSI Tensor Generation for Device-Free Wi-Fi Sensing. *CMES-Comput. Model. Eng. Sci.* **2024**, *140*, 1749–1767. [[CrossRef](#)]
31. Zhang, Y.; Wang, L.; Li, J. Non-negative matrix factorization for feature selection: A review. *J. Mach. Learn. Res.* **2021**, *22*, 1–30. [[CrossRef](#)]
32. Saberi-Movahed, F.; Berahman, K.; Sheikhpour, R.; Li, Y.; Pan, S. Nonnegative matrix factorization in dimensionality reduction: A survey. *arXiv* **2024**, arXiv:2405.03615.
33. Feng, X.R.; Li, H.C.; Wang, R.; Du, Q.; Jia, X.; Plaza, A. Hyperspectral unmixing based on nonnegative matrix factorization: A comprehensive review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 4414–4436. [[CrossRef](#)]
34. Del Corso, G.M.; Romani, F. Adaptive nonnegative matrix factorization and measure comparisons for recommender systems. *Appl. Math. Comput.* **2019**, *354*, 164–179. [[CrossRef](#)]
35. Yuan, Y.; Yu, N.; Zheng, Z.; Yang, Y.; Ma, K.; Liu, Z.; Chen, C.; Zhang, J. Enhancing Network Abnormal Detection with NMF-SECNN: Leveraging Deep Feature Learning for High-Precision Traffic Analysis. *IEEE Trans. Netw. Sci. Eng.* **2025**, *12*, 2069–2080. [[CrossRef](#)]
36. Li, F.; Gao, Q.; Wang, Q.; Yang, M.; Deng, C. Tensorized Soft Label Learning Based on Orthogonal NMF. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *36*, 13219–13231. [[CrossRef](#)] [[PubMed](#)]
37. Reynolds, D. Gaussian mixture models. In *Encyclopedia of Biometrics*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 827–832.
38. Passah, A.K.A.; Chorti, A.; de Lamare, R.C. Enhanced Multiuser CSI-Based Physical Layer Authentication Based on Information Reconciliation. *IEEE Wirel. Commun. Lett.* **2024**, *14*, 544–548. [[CrossRef](#)]
39. Cavicchia, C.; Vichi, M.; Zaccaria, G. Gaussian mixture model with an extended ultrametric covariance structure. *Adv. Data Anal. Classif.* **2022**, *16*, 399–427. [[CrossRef](#)]
40. Lai, Z.; Chang, Z.; Sha, M.; Zhang, Q.; Xie, N.; Chen, C.; Niyato, D. A Comprehensive Survey on Physical Layer Authentication Techniques: Categorization and Analysis of Model-Driven and Data-Driven Approaches. *ACM Comput. Surv.* **2025**, *57*, 117. [[CrossRef](#)]
41. Bermejo-Sabbagh, C.; Orozco-del Castillo, M.G.; Valdiviezo-Navarro, J.C.; Ortiz-Sánchez, P.A.; Cuevas-Cuevas, N.L. Hardware performance analysis for deep learning-based mistletoe detection in UAV imagery. *Comput. Electron. Agric.* **2025**, *237*, 110629. [[CrossRef](#)]
42. Somvanshi, S.; Islam, M.M.; Chhetri, G.; Chakraborty, R.; Mimi, M.S.; Shuvo, S.A.; Islam, K.S.; Javed, S.A.; Rafat, S.A.; Dutta, A.; et al. From Tiny Machine Learning to Tiny Deep Learning: A Survey. *arXiv* **2025**, arXiv:2506.18927.

43. Maitra, S.; Richards, D.; Abdelgawad, A.; Yelamarthi, K. Performance Evaluation of IoT Encryption Algorithms: Memory, Timing, and Energy. In *Proceedings of the 2019 IEEE Sensors Applications Symposium (SAS)*; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6. [[CrossRef](#)]
44. Dennis, D.K.; Gaurkar, Y.; Gopinath, S.; Goyal, S.; Gupta, C.; Jain, M.; Jaiswal, S.; Kumar, A.; Kusupati, A.; Lovett, C.; et al. EdgeML: Machine Learning for Resource-Constrained Edge Devices. 2020. Available online: <https://github.com/Microsoft/EdgeML> (accessed on 15 May 2026).
45. Pereira, G.C.C.F.; Alves, R.C.A.; Silva, F.L.d.; Azevedo, R.M.; Albertini, B.C.; Margi, C.B. Performance Evaluation of Cryptographic Algorithms over IoT Platforms and Operating Systems. *Secur. Commun. Netw.* **2017**, *2017*, 2046735. [[CrossRef](#)]
46. Zeng, Z.; Sapatnekar, S.S. Energy-efficient hardware acceleration of shallow machine learning applications. In *Proceedings of the 2023 Design, Automation & Test in Europe Conference & Exhibition (DATE)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
47. Kazmierczak, J.; Salama, K.; Huerta, V.; Bahadur, S.K.J. MLOps: Continuous Delivery and Automation Pipelines in Machine Learning. 2024. Available online: <https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning> (accessed on 28 August 2024).
48. Chandra, V.; Chen, Y.; Yoo, S. Introduction to the special section on energy-efficient AI chips. *ACM Trans. Des. Autom. Electron. Syst. (TODAES)* **2022**, *27*, 41. [[CrossRef](#)]
49. HPC Wiki Contributors. Performance Metrics. 2025. Available online: [https://hpc-wiki.info/hpc/Performance\\_metrics](https://hpc-wiki.info/hpc/Performance_metrics) (accessed on 9 July 2025).
50. Estrada, R.; Valeriano, I.; Aizaga, X. CPU usage prediction model: A simplified VM clustering approach. In *Proceedings of the Conference on Complex, Intelligent, and Software Intensive Systems*; Springer: Cham, Switzerland, 2023; pp. 210–221.
51. Roopak, M.; Tian, G.Y.; Chambers, J. Deep learning models for cyber security in IoT networks. In *Proceedings of the 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*; IEEE: Piscataway, NJ, USA, 2019; pp. 452–457.
52. Roopak, M.; Tian, G.Y.; Chambers, J. An intrusion detection system against DDoS attacks in IoT networks. In *Proceedings of the 2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*; IEEE: Piscataway, NJ, USA, 2020; pp. 562–567.
53. Roopak, M.; Ran, Y.; Chen, X.; Tian, G.Y.; Parkinson, S. Channel state information based physical layer authentication for Wi-Fi sensing systems using deep learning in Internet of things networks. *IET Wirel. Sens. Syst.* **2024**, *14*, 441–450. [[CrossRef](#)]
54. Qiu, X.; Jiang, T.; Wu, S.; Hayes, M. Physical layer authentication enhancement using a Gaussian mixture model. *IEEE Access* **2018**, *6*, 53583–53592. [[CrossRef](#)]
55. Yan, J.; Ma, C.; Kang, B.; Wu, X.; Liu, H. Extreme learning machine and AdaBoost-based localization using CSI and RSSI. *IEEE Commun. Lett.* **2021**, *25*, 1906–1910. [[CrossRef](#)]
56. Song, Y.; Liu, W.; Zhu, H.; Gong, Y.; Li, Y.; Huang, J.; Deng, Y. Enhancing wireless channel authentication in industrial control: Attack-resistant CSI-based PUF. *Ad Hoc Netw.* **2025**, *178*, 103919. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.