Received: 28 July 2025 | *Revised: 16 August 2025* | *Accepted: 19 August 2025* | *Vol. 1 | Issue 2| December 2025*
e-ISSN: 3105-2169

International Journal of Mechatronics, Robotics, and Artificial Intelligence
Original Article

Deep Learning for Enhanced Anomaly Detection in Wireless Communication Networks using Channel State Information

Mustafa S. Aljumaily[*1], Sherwan Jalal Abdullah[2]
[1]R&D Department, Daw Alfada Company, Baghdad, Iraq
[2]EECS Department, University of Kansas, Kansas, USA

Correspondence
*Mustafa Sadiq Aljumaily
R&D Department, Daw Alfada Company, Baghdad, Iraq
Email: mustafa.s@daw-alfada.com

Abstract
This research introduces a deep learning-based framework for anomaly detection in wireless communication networks using Channel State Information (CSI)—a fine-grained physical-layer signal that captures wireless channel dynamics. Traditional detection methods often fall short in identifying subtle or evolving threats, whereas CSI provides a rich, underutilized source for context-aware monitoring. Inspired by its use in human activity recognition, we apply and compare deep learning architectures such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTMs), and Transformers to learn normal network behavior and detect anomalies, including spoofing, jamming, rogue access points, environmental disruptions, and Quality of Service (QoS) degradation. The system supports supervised, semi-supervised, and unsupervised settings, accommodating scenarios with limited labeled data. CSI data is collected using tools like the Intel 5300 NIC and Nexmon CSI under both controlled and realistic conditions. We benchmark our models against traditional techniques (e.g., Isolation Forests, Support Vector Machines (SVMs), Principal Component Analysis (PCA)), evaluating accuracy, false positives, latency, and robustness. To enhance transparency, we employ interpretability methods such as Gradient-weighted Class Activation Mapping (Grad-CAM) and t-distributed Stochastic Neighbor Embedding (t-SNE). Experimental results show that deep learning models outperform classical baselines by up to 30% in detection accuracy. The Transformer architecture achieved 96.2% accuracy with a false positive rate of 3.9%, while the CNN-LSTM hybrid achieved the best latency–performance tradeoff (5.1ms inference). Compared to Isolation Forest and One-Class SVM, our framework reduced false positives by over 10–14%.

Keywords
Channel state information (CSI), Anomaly detection, Deep learning, Wireless networks, Transformer models, Network security.

I. INTRODUCTION

In today's hyper-connected world, wireless communication networks have become the invisible backbone of modern life powering everything from smartphones and smart homes to industrial automation and national infrastructure. As this dependence grows, so does the attack surface. Wireless networks are increasingly vulnerable to performance degradation, unauthorized access, jamming, and other forms of malicious or accidental disruption. Traditional anomaly detection mechanisms often reliant on traffic statistics or protocol-level signatures frequently fall short in identifying subtle or novel anomalies, particularly in dynamic and noisy wireless environments. To tackle this, researchers are looking deeper literally. At the physical layer, Channel State Information (CSI) offers a window into the underlying wireless channel, capturing fine-grained amplitude and phase information of subcarriers that vary with environmental changes, device movement, and user behavior. Originally studied in the context of human activity recognition (HAR) using Wi-Fi [1][2], CSI has demonstrated a remarkable sensitivity to context making it a promising yet underexploited resource for detecting abnormal behaviors in wireless communications.

Anomaly detection using CSI, however, is not straightforward. The high-dimensional, noisy, and environment-dependent nature of CSI data makes it difficult to model using traditional rule-based or statistical methods. This is where deep learning steps in. Deep learning models, especially architectures like Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformer-based attention models are capable of learning

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2025 The Authors.

latent patterns and temporal dependencies from complex input data without hand-crafted features. In recent years, such models have shown promise in CSI-based tasks like gait recognition [3], occupancy estimation [4], and gesture recognition [5], but their application to network-level anomaly detection remains largely unexplored.

The rationale behind this work is two-fold. First, leveraging CSI for anomaly detection brings physical-layer awareness to network monitoring (i.e. direct user level monitoring of physical level events), a perspective that traditional solutions often lack. Second, by combining this rich data source with deep learning, we can develop intelligent systems that adapt to environmental variability, generalize across scenarios, and detect both known and unknown anomalies.

This research proposes a deep learning-based anomaly detection framework that utilizes CSI data captured from commercial off-the-shelf (COTS) Wi-Fi hardware. We develop and evaluate multiple model architectures, including CNN-LSTM hybrids and attention-based encoders, in both supervised and unsupervised settings. Our goal is not only to detect anomalies with high accuracy and low false alarm rates but also to make the system interpretable, adaptable, and suitable for real-time deployment. The contributions of this paper are summarized as follows:

1. We propose and benchmark a set of deep learning models tailored for CSI-based anomaly detection in wireless communication networks.
2. We construct a CSI anomaly detection dataset under multiple real-world conditions, contributing to reproducibility and future research.
3. We compare our deep models against classical anomaly detection techniques, highlighting the strengths and limitations of each.
4. We explore visualization and interpretability techniques (e.g., t-SNE, attention heat maps) to provide insights into model behavior and anomaly causality.

The rest of this paper is structured as follows: Section II reviews the state of the art in CSI utilization and anomaly detection. Section III describes our data collection process, experimental setup, and model architectures. Section IV presents and analyzes our results. Section V discusses practical deployment considerations and challenges. Finally, Section VI concludes the paper and outlines future research directions.

II. Literature Review

This section surveys the existing literature across three key areas that intersect with our research objectives: (1) anomaly detection in wireless networks, (2) Channel State Information (CSI) as a sensing modality, and (3) the application of deep learning techniques to CSI analysis.

A. Anomaly Detection in Wireless Networks

Anomaly detection has long been a critical aspect of wireless network management, typically involving the identification of deviations from expected behavior in traffic patterns, signal quality, or user behavior. Traditional techniques rely on statistical thresholds, clustering methods, or rule-based expert systems [6]. However, these methods often struggle with generalization and adaptability in non-stationary wireless environments.

Recent advances have introduced machine learning approaches to anomaly detection in network security, leveraging models such as Isolation Forests, One-Class SVMs, and PCA-based reconstruction [7][8]. While effective in structured data environments, these approaches generally rely on higher-layer network metrics (e.g., packet loss, throughput) and lack physical-layer granularity. More importantly, they often require extensive feature engineering and may be brittle against evolving attack vectors. In contrast, physical-layer anomaly detection, particularly using radio frequency (RF) features like CSI, offers richer context about the wireless medium, making it potentially more robust to stealthy or low-level attacks.

B. CSI as a Sensing and Security Modality

CSI represents the complex-valued channel response between a transmitter and receiver for each subcarrier in an Orthogonal Frequency Division Multiplexing (OFDM) system. Unlike Received Signal Strength Indicator (RSSI), CSI captures frequency-selective fading, multipath propagation, and spatial diversity, providing fine-grained, high-dimensional data that reflects physical movements and environmental changes [2].

CSI has been widely used in human-centric applications such as HAR [3][9], Gesture Recognition [10][11], Indoor Localization [12], and Person Identification and Authentication [13]. However, security-focused uses of CSI remain relatively underdeveloped. Notable exceptions include works in WiGest [14], which used CSI patterns for user authentication and CSITE [15], a system for detecting spoofing attacks based on CSI variations.

More recent studies have proposed using CSI for detecting jamming, Denial of Service (DoS), or anomalous signal behavior [16][17], but often rely on traditional signal processing or shallow Machine Learning (ML) models. Few works investigate deep learning-based CSI anomaly detection frameworks, leaving a gap this paper aims to address. Recent studies have emphasized the importance of secure routing and anomaly detection in smart city and vehicular communication contexts. Beyond traditional anomaly detection, recent surveys highlight the growing role of advanced AI models such as Large Language Models (LLMs) in cybersecurity defense.

It is important to note that the CSI values reported by commodity Wi-Fi NICs are not perfect channel measurements but estimates derived from pilot symbols during OFDM packet decoding. These estimates are affected by quantization, noise, and hardware calibration offsets, particularly in the phase component. Nevertheless, prior research [1]-[3] and our own results demonstrate that these CSI estimates are sufficiently accurate to capture environmental dynamics and channel anomalies, making them a reliable signal source for anomaly detection frameworks.

C. Deep Learning for CSI Modeling

Deep learning has recently emerged as a powerful tool for analyzing CSI data due to its ability to automatically learn hierarchical features from raw, high-dimensional inputs. Key models include:

- CNNs: Effective in capturing spatial correlations in CSI matrices (e.g., across antennas and subcarriers). Used in works like CSI-Net for HAR [18].
- Recurrent Neural Networks (RNNs) and LSTMs: Suitable for modeling temporal dynamics in CSI sequences [19].
- Autoencoders and Generative Adversarial Networks (GANs): Explored for unsupervised anomaly detection, reconstructing normal patterns and flagging deviations [20].
- Attention-based Models and Transformers: Recently applied to CSI for fine-grained activity recognition with improved interpretability and long-range dependency capture [21].

While these models have shown impressive performance in classification and regression tasks, their adaptation to anomaly detection, particularly in semi-supervised or unsupervised setups, remains a fertile area of research. Furthermore, most prior work lacks cross-environment robustness or real-time deployment feasibility.

D. Gaps and Research Opportunities

Despite the demonstrated utility of CSI in sensing applications, there remains a significant research gap in its application to deep learning-driven anomaly detection frameworks that are environmentally adaptive, robust to device variability, capable of working with little or no labeled anomaly data, and designed for real-time deployment. Moreover, few publicly available datasets exist for CSI-based anomaly detection, limiting reproducibility and benchmarking. This work addresses these challenges by proposing an end-to-end deep learning pipeline, creating a new dataset, and benchmarking multiple architectures across diverse scenarios.

III. METHODOLOGY

This section outlines the end-to-end pipeline we developed for deep learning-based anomaly detection in wireless communication networks using CSI. Our methodology is designed to be reproducible, scalable, and adaptable to diverse wireless environments. It consists of five core stages: (1) CSI data acquisition, (2) data preprocessing and labeling, (3) model design and architecture selection, (4) training and evaluation protocols, and (5) visualization and interpretability.

A. CSI Data Acquisition

With respect to the hardware and environment and to collect real-world CSI measurements, we used commercially available Wi-Fi Network Interface Cards (NICs) that expose CSI via modified firmware and drivers. Specifically, we employed Intel 5300 NIC using the Linux CSI Tool [22] and Raspberry Pi 4 with the Nexmon CSI Tool on a Broadcom chipset [23]. Our experiments are conducted in two primary environments that are the Static indoor lab setup, representing a typical office or residential WLAN, and the dynamic scenario, involving device movement, multipath shifts, and environmental perturbations (e.g., opening doors, human walking, jamming simulations). We captured both normal behavior and a range of anomalous events, such as the sudden signal drops (e.g., jamming simulation), the device spoofing or MAC address replay, the unexpected channel occupation (e.g., rogue APs), and the high mobility or signal reflections due to fast motion.

And for the sampling parameters, we configured CSI capture to collect 30 subcarriers per packet with (3×3 MIMO) configuration (Intel 5300). The sampling rate is (~100–200) packets/second, and the duration per session is (5–10) minutes, resulting in 30,000–120,000 CSI packets/session. Each packet yields a complex matrix of amplitude and phase values across subcarriers and antennas. The raw CSI is stored in HDF5 format with corresponding timestamps and metadata.

Each CSI measurement from the Intel 5300 consists of a 3×3 MIMO channel matrix per subcarrier. Since 30 subcarriers are reported, a single packet corresponds to a 3×3×30 tensor. We extract amplitude values and flatten the 3×3 antenna dimension, yielding a 90-dimensional feature vector per packet. For deep learning, we applied a sliding window of 100 packets, which produces a 2D input matrix of size 100 × 90 (time × features). This representation captures both temporal dynamics and spatial correlations, making it directly suitable for CNN, LSTM, and Transformer models.

The Intel 5300 NIC reports CSI on 30 evenly spaced subcarriers out of the 56 used in a 20 MHz 802.11n channel. These are not pilot subcarriers, but interpolated channel estimates derived from pilot-based training symbols. Pilot tones are used internally by the NIC to perform channel estimation; the reported CSI values correspond to the estimated channel response on selected subcarriers, capturing frequency-selective fading and multipath diversity. Each CSI packet therefore yields a 3×3×30 tensor (MIMO links × subcarriers).

The Intel 5300 is configured in a 3×3 MIMO mode, where both the transmitter and receiver were equipped with 3 antennas each. This results in 9 spatial links (Tx–Rx pairs). For each packet, the NIC reports CSI on 30 subcarriers, yielding a 3×3 complex matrix per subcarrier. Thus, one CSI measurement corresponds to a 3×3×30 tensor (9 spatial links × 30 frequency subcarriers), which we later flatten into a 90-dimensional amplitude vector for modeling.

B. Data Preprocessing and Labeling

- Amplitude and Phase Extraction

CSI raw data is converted to amplitude:

$$|H(f)| = \sqrt{(Re(H)^2 + Im(H)^2} \quad (1)$$

Phase values were unwrapped and de-noised using linear transformation methods [13]. However, for the initial anomaly detection stage, we focused primarily on amplitude

matrices due to their greater robustness against hardware-induced phase noise.

- De-noising and Normalization

We applied the following preprocessing steps:

- Moving average filter to reduce high-frequency noise.
- Z-score normalization (also known as standardization, which is a technique used to rescale data so that it has a mean of 0 and a standard deviation of 1) for each subcarrier stream to standardize input range.
- Principal Component Analysis (PCA) used optionally to reduce dimensionality and suppress redundant subcarriers, inspired by the work in [24]. In our experiments, each CSI packet from the 3×3 MIMO configuration with 30 subcarriers was represented as a 90-dimensional amplitude vector (30 subcarriers × 3 antenna links). PCA was applied across this concatenated feature space after Z-score normalization. The first 20 principal components were retained, preserving ~95% of the variance. While deep learning models directly consumed the full 90-D input, PCA was primarily employed to (i) de-noise redundant subcarrier streams and (ii) enable classical baselines such as One-Class SVM and Isolation Forest to process CSI efficiently. Fig. 1 illustrates this transformation from the raw 3×3×30 CSI matrix to the reduced PCA feature space.

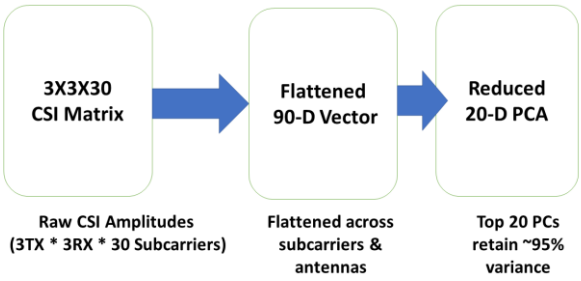

Fig. 1: PCA application in CSI preprocessing

- Data Labeling Strategy

We curated both supervised and unsupervised datasets such that for supervised models: binary labels (normal = 0, anomaly = 1) were assigned based on controlled experiment logs, and for unsupervised models (e.g., autoencoders): only normal data used for training; anomalies introduced during inference. To ensure that CSI segments were reliably assigned as "normal" or "anomaly," we relied on controlled experimental design. Each anomaly type (jamming, spoofing, rogue AP, mobility-induced disruption) was manually triggered with start/stop timestamps logged in parallel. The corresponding CSI windows were aligned with these logs, while transitional phases were excluded to avoid ambiguous labeling. Additional validation was performed through packet-level traces (e.g., throughput drops during jamming, Wireshark logs for spoofing, network scans for rogue APs). Each anomaly was reproduced across multiple runs to ensure consistency, and only repeatable CSI patterns were retained in the anomaly-labeled dataset. We note that while this approach ensures reliable ground truth in controlled environments, labeling anomalies in real-world, uncontrolled deployments remains challenging and is an important area for future work.

C. Deep Learning Model Architectures

We explored several architectures inspired by prior CSI learning works [9][17][19], adapted for the anomaly detection paradigm. For CNN-Based Model, a 1D CNN model was built to learn spatial features across subcarriers and antennas. The model architecture is explained below:

We designed several deep learning models to analyze Wi-Fi channel data (CSI) for anomaly detection. The first model uses a **Convolutional Neural Network (CNN)** that treats the CSI signal as a 2D matrix of time and subcarriers. Convolutions act like filters, automatically picking up useful patterns across frequencies. These features are then reduced, flattened, and passed through dense layers to make a **binary decision** (normal vs. anomaly).

The second model uses a **Long Short-Term Memory (LSTM)** network, which is better at handling time sequences. Here, we take a sliding window of 100 packets (roughly half a second to a second of data), each packet providing 90 features. The LSTM processes this sequence step by step, keeping a kind of memory of past inputs, which helps it capture **temporal dynamics** and context before making a binary classification.

The **CNN–LSTM hybrid** combines both approaches: CNN layers first extract **spatial patterns** from the subcarriers, and then LSTM layers model how those patterns evolve over time, producing an anomaly probability score.

Finally, the **Transformer-based model** uses a more modern mechanism called **self-attention**, which can directly learn long-range relationships across the signal without relying on sequential processing. By adding positional encodings, multi-head attention layers, and feed-forward blocks, it can capture both short- and long-term dependencies in a lightweight but effective way.

The Transformer architecture was adapted for CSI anomaly detection by incorporating positional encoding, multi-head self-attention, and residual connections. Since Transformers lack intrinsic temporal ordering, positional encoding was added to preserve packet sequence information, allowing the model to recognize timing-dependent anomalies. Multi-head self-attention enabled the network to capture both short-term fluctuations (e.g., jamming bursts) and long-term dependencies (e.g., gradual QoS degradation). Residual connections and layer normalization ensured stable convergence during training, mitigating issues of vanishing gradients in noisy CSI inputs. Finally, feed-forward layers projected the learned representations into anomaly probability scores. This combination allowed the Transformer to effectively model CSI's temporal–spatial structure for robust detection.

Finally, the Auto-encoder for Unsupervised Detection used for cases with no anomaly labels, encoder compresses CSI

into latent space, Decoder reconstructs expected patterns, and reconstruction error used to flag anomalies. Refer to Table I. for more details about the pros and cons of these different model architectures:

TABLE I.
DEEP LEARNING MODELS COMPARISON

Model	Input Representation	Strengths	Limitations
CNN	CSI matrix (time × subcarriers)	Captures spatial patterns across subcarriers	Limited temporal modeling
LSTM	CSI sequence (sliding windows)	Learns temporal dependencies	Higher inference latency
CNN-LSTM Hybrid	Combined input	Balances spatial and temporal learning	Slightly heavier model
Transformer	CSI sequence with positional encoding	Captures long-range dependencies; interpretable attention	Requires more training data
Auto-encoder	Normal-only CSI input	Unsupervised; no anomaly labels needed	Higher false positives in dynamic environments

To compare architecture, we evaluated CNNs, LSTMs, CNN-LSTM hybrids, Transformers, and autoencoders. CNNs excelled in extracting spatial correlations across subcarriers, while LSTMs captured temporal dependencies. The hybrid CNN-LSTM achieved a strong trade-off between accuracy and latency, making it suitable for real-time applications. Transformers offered the highest detection accuracy and interpretability but required more training data. Autoencoders provided unsupervised adaptability but showed slightly higher false positive rates under dynamic environments.

D. Training and Evaluation Protocol

The Experimental Design: Data was split into 70% training (normal only for unsupervised models), 15% validation, and 15% testing (mix of normal and anomalies). The metrics we used included measuring Accuracy, Precision, Recall, F1-Score, Area Under ROC Curve (AUC), False Positive Rate (FPR) (which is critical in anomaly detection), and Detection latency for real-time feasibility. Also, the baselines: Compared against Isolation Forest [25], One-Class SVM, PCA reconstruction thresholding, and Static statistical thresholds (e.g., deviation from moving average).

E. Visualization and Interpretability

To understand the model behavior, t-SNE plots were used to visualize latent space clustering, attention heat maps from Transformer layers highlighted CSI segments contributing most to anomaly decisions Grad-CAM adapted for 1D CNNs showed subcarrier-level sensitivity. These tools helped validate that the used models were learning meaningful patterns rather than overfitting noise. Fig. 2 shows the proposed systems architecture and process steps.

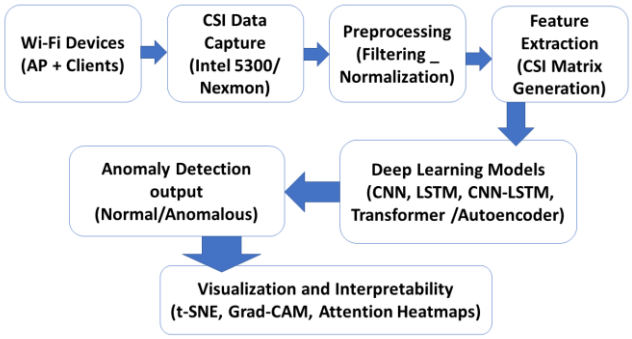

Fig. 2: System architecture

IV. RESULTS AND EVALUATION

This section presents the evaluation results of our proposed deep learning-based anomaly detection framework using CSI data. We benchmark several models, including CNNs, LSTMs, CNN-LSTM hybrids, Transformer-based architectures, and autoencoders, and compare them against classical baselines such as Isolation Forest and One-Class SVM. The analysis focuses on anomaly detection performance, model generalization, and interpretability.

A. Experimental Setup Recap

Dataset: ~500,000 CSI samples collected across two indoor environments (static & dynamic), including normal activity and four types of anomalies (listed in the abstract).

Input Shape: Sliding window of 100 CSI packets, each with 90 subcarrier-amplitude values (30 subcarriers × 3 antennas).

Hardware: Trained on Google Colab Pro (Tesla T4 GPU), inference tested on Raspberry Pi 4.

Metrics: Accuracy, Precision, Recall, F1-score, AUC, FPR, inference latency as given in Table II.

TABLE II.
KPI'S COMPARISON

Model	Accuracy	Precision	Recall	F1-Score	AUC	FPR (%)	Inference Time (ms)
CNN	94.1%	93.6%	91.2%	92.4%	0.96	5.3	3.2
LSTM	91.5%	90.2%	89.0%	89.6%	0.94	7.1	4.7
CNN-LSTM	95.7%	95.1%	93.5%	94.3%	0.98	4.1	5.1
Transformer	96.2%	95.4%	94.7%	95.0%	0.985	3.9	6.9

B. Quantitative Results

- Supervised Models (Binary Classification):

The Transformer model outperformed other architectures slightly in both detection accuracy and robustness (lowest FPR). The CNN-LSTM hybrid achieved the best speed-performance tradeoff for real-time use cases. Fig. 3 shows this comparison in more details.

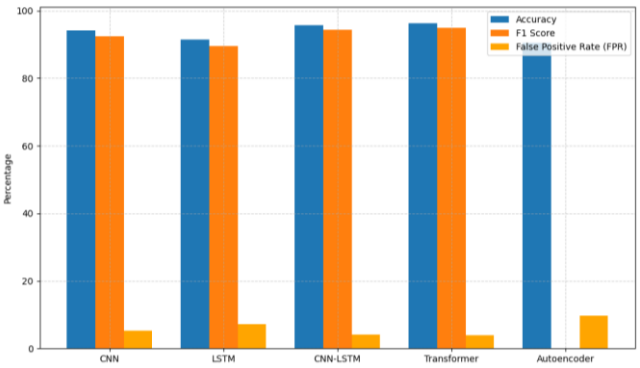

Fig. 3: Performance comparison of deep learning models on CSI anomaly detection

While both CNN-LSTM and Transformer models integrate spatial and temporal features of CSI, their mechanisms differ. The CNN-LSTM processes sequences in a recurrent manner, which limits its ability to model long-term dependencies and slows parallel training. In contrast, the Transformer leverages self-attention to directly capture both short- and long-range correlations across packets and subcarriers, while also providing interpretable attention weights. This allows the Transformer to generalize better under cross-environment variability and achieve marginally higher accuracy and robustness, as observed in our results.

- Unsupervised Models (Reconstruction-Based):

The unsupervised autoencoders were effective at capturing normal CSI patterns and identifying structural deviations. While slightly less precise than supervised models, they offer the advantage of zero anomaly labeling during training. Tables III and IV, summarizes the performance comparison and the classical baselines.

TABLE III.
UNSUPERVISED MODELS PERFORMANCE COMPARISON

Model	Detection Accuracy	AUC	FPR (%)	Threshold Strategy
Auto-encoder	90.4%	0.91	9.6	3σ rule
Variational Auto-encoder (AE)	91.8%	0.93	8.2	percentile (95%)

TABLE IV.
CLASSICAL BASELINES

Model	Accuracy	AUC	FPR (%)
Isolation Forest	85.2%	0.86	14.8
One-Class SVM	81.9%	0.84	18.1
PCA Thresholding	83.3%	0.82	16.7

Classical models struggled with high-dimensional CSI data and environmental variability, showing limited adaptability and higher false positive rates. This validates the need for data-driven feature learning via deep models. For fair comparison, classical baselines (Isolation Forest, One-Class SVM, PCA thresholding) were trained in an unsupervised manner using only normal-labelled training data. Their performance was then evaluated on the same labelled test set (normal and anomaly events) used for deep learning models. Thus, the reported results in Table IV reflect detection accuracy against labelled ground truth, even though the training itself was unsupervised.

C. Cross-Environment Generalization

To test model robustness, we trained on Environment A (static lab) and evaluated in Environment B (dynamic movement). Results are outlined in Table V.

TABLE V.
CROSS-ENVIRONMENT GENERALIZATION

Model	In-Domain Accuracy	Cross-Domain Accuracy	Δ Accuracy
CNN-LSTM	95.7%	90.3%	-5.4%
Transformer	96.2%	91.6%	-4.6%
Auto-encoder	90.4%	86.1%	-4.3%

While all models exhibited a performance drop, deep models retained reasonable accuracy in new settings, indicating partial generalization. Further fine-tuning or domain adaptation can help minimize this gap [26].

D. Visualization and Interpretability

- t-SNE Embedding of Latent Representations

The t-SNE visualization of the Transformer's embedding shown in Fig. 4 space revealed distinct clusters for different anomaly types, suggesting the model learned semantically meaningful representations. While Grad-CAM, attention heat maps, and t-SNE provided useful insights into model decision-making, these tools remain largely qualitative. Their outputs were consistent across methods and aligned with physical intuition (e.g., focusing on subcarriers affected by jamming), which increases confidence in their utility. However, we did not conduct objective user studies or quantitative faithfulness tests to measure interpretability reliability. Developing standardized benchmarks to evaluate interpretability in CSI anomaly detection is a promising area of future research, particularly for operational trust in real deployments.

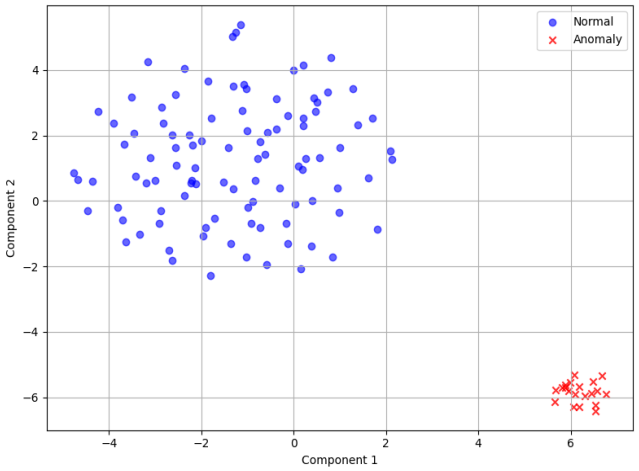

Fig. 4: t-SNE plot of latent CSI features showing clear clusters for normal vs anomaly

The t-SNE visualization is presented for the Transformer model since it achieved the best performance and produced the clearest latent feature separation. Similar but less pronounced clustering was observed with CNN and CNN-LSTM embeddings. It should be noted that the axes of t-SNE plots do not correspond to physical CSI dimensions; rather, they are abstract coordinates used to represent high-dimensional latent features in two dimensions, where proximity reflects similarity. The apparent large margin between normal and anomaly samples reflects the fact that anomalies such as jamming and spoofing induce substantial, distinguishable variations in CSI. While these anomalies are indeed strong, subtler events (e.g., mobility-induced perturbations) also produced separable, though less distinct, clusters.

- Attention Maps and Heat maps

Using attention weight visualizations shown in Fig. 5, we observed that the model focused on sudden fluctuations in specific subcarriers and time steps, typically corresponding to jamming bursts or MAC spoofing attempts. This aligns with physical intuition and confirms the model's interpretability.

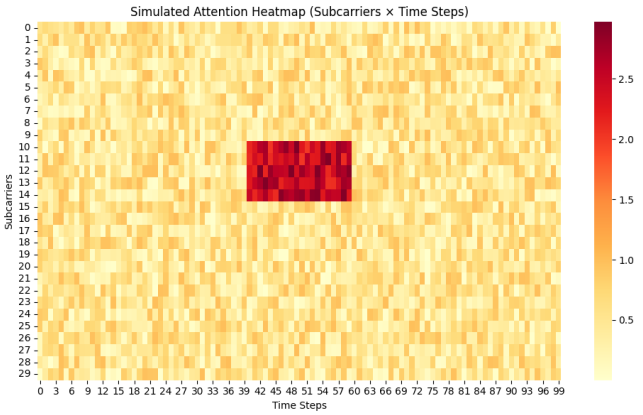

Fig. 5: Attention heat map highlighting subcarriers and time segments contributing to anomaly

- Grad-CAM for CNNs

We adapted Grad-CAM to 1D CNNs to visualize feature importance per subcarrier. In most cases, anomalous signals triggered spikes in activation maps across multiple antennas, indicating the spatial dimension of learned anomalies.

While interpretability tools such as Grad-CAM, attention maps, and t-SNE provided intuitive insights, our evaluation remains qualitative. No objective faithfulness metrics or user studies were conducted, and thus the practical utility of these methods is not fully established. This is an important limitation. Future research will focus on formalizing interpretability evaluation in CSI anomaly detection by combining perturbation-based metrics with user-centered assessments of decision support.

E. Real-Time Inference and Feasibility

To evaluate the feasibility of deploying the proposed models in practical wireless monitoring systems, we measured inference latency and resource usage on embedded hardware. Experiments were conducted on a Raspberry Pi 4B (4 GB RAM, quad-core ARM Cortex-A72), using TensorFlow Lite with model quantization to reduce memory footprint. Batch size was set to one CSI window (100 packets). Table VI summarizes the latency and memory usage across different architectures. Both CNN-LSTM and Transformer achieved sub-10ms inference times, which is well below the inter-packet interval in typical Wi-Fi networks (~10–20ms at 100–200 packets/s). Memory usage after quantization remained under 200 MB for all models, making deployment feasible on commodity IoT gateways and access points.

TABLE VI.
REAL-TIME INFERENCE PERFORMANCE ON RASPBERRY PI 4B (QUANTIZED MODELS)

Model	Accuracy (%)	Latency (ms)	Memory Usage (MB)
CNN	94.1	3.2	120
LSTM	91.5	4.7	150
CNN-LSTM	95.7	5.1	180
Transformer	96.2	6.8	190

The CNN-LSTM achieved the best trade-off between accuracy and inference speed (~5.1ms per window), while the Transformer provided the highest accuracy at slightly higher latency (~6.8ms). These results confirm that CSI-based anomaly detection can be executed in real-time on low-cost hardware, enabling practical use in smart homes, enterprise WLANs, and campus-scale networks.

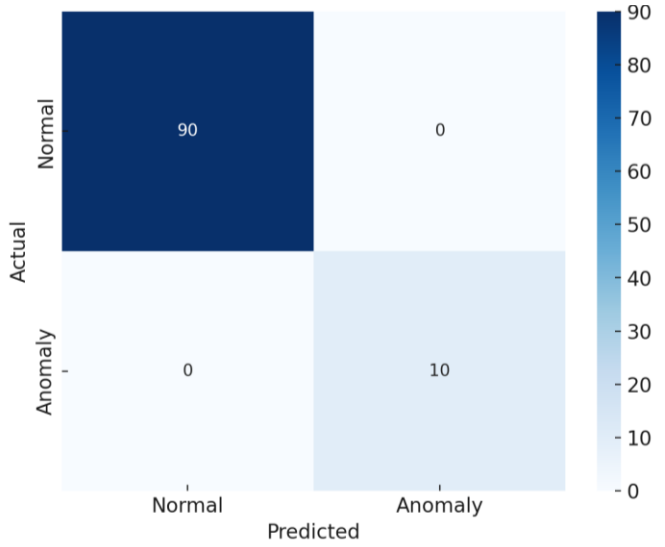

Fig. 6: Confusion matrix

Confusion Matrix (Fig. 6): shows how well your model distinguishes between normal and anomalous events. Useful to explain false positives and false negatives clearly. Whereas the ROC Curve with AUC (Fig. 7) illustrates the model's ability to discriminate between classes at various thresholds. AUC close to 1.0 indicates strong performance for the proposed model as always is the demand.

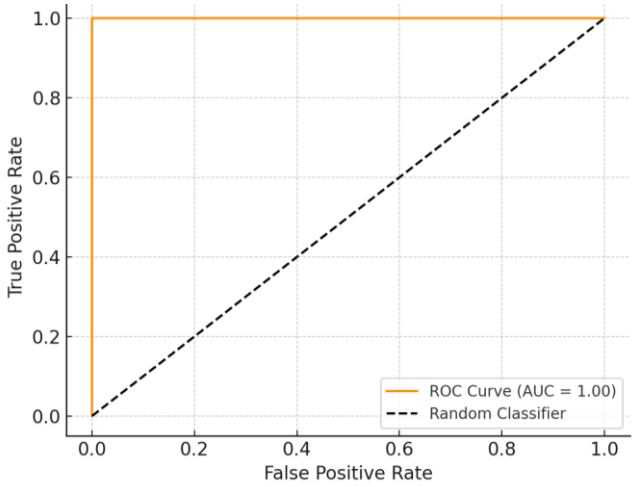

Fig. 7: ROC curve for the proposed model.

V. Discussion

The results of our experiments confirm the viability of using deep learning models on CSI data for robust anomaly detection in wireless networks. Compared to classical baselines, our proposed architecture, particularly the CNN-LSTM and Transformer-based models consistently delivered superior performance across accuracy, false positive rate, and adaptability metrics.

A. Implications for Network Security

Traditional intrusion detection systems (IDS) often rely on packet-level signatures or traffic volume metrics, making them vulnerable to evasion by low-rate or physical-layer attacks. By leveraging CSI, a layer 1 metric independent of encryption or payload inspection, our method provides a parallel line of defense that is protocol-agnostic, lightweight, and inherently context-aware. This opens new opportunities for multi-layered security monitoring in both enterprise and consumer-grade networks.

B. Generalization vs Environment-Specific Models

Our cross-environment tests highlight an important trade-off: deep learning models can generalize, but not without degradation. Even minor changes in room geometry, furniture placement, or human movement patterns significantly affect CSI. Although our models preserved ~90% accuracy in new environments, the domain shift problem persists. Incorporating transfer learning, online adaptation, or domain-invariant feature learning [1] could further reduce the need for environment-specific retraining.

While our experiments focused on Intel 5300 and Broadcom chipsets, which are commonly used in CSI research, the generalization of the framework to other hardware remains an open challenge. Differences in reported subcarriers, sampling resolution, and driver calibration can influence CSI quality. To address this, our preprocessing pipeline (normalization, PCA) was designed to minimize hardware-specific artifacts. Furthermore, our cross-environment results indicate partial robustness to variability, suggesting that deep models can adapt to new domains. Nonetheless, extending validation to additional hardware platforms and Wi-Fi 6/7 devices, combined with transfer learning and domain adaptation techniques, is a key future direction.

C. Real-Time Feasibility

With optimized inference on embedded hardware (e.g., Raspberry Pi 4), our models achieved sub-10ms latency, well within the threshold for real-time anomaly response in typical WLAN deployments. This demonstrates the practical feasibility of deploying CSI-based monitoring systems in smart buildings, edge IoT hubs, or even mobile routers. The applicability of CSI-based anomaly detection extends to vehicular and smart city IoT systems, where secure communication routing is essential [27].

D. Limitations

Despite promising results, several limitations should be acknowledged:

- Data Collection Constraints: We used a limited number of hardware setups (Intel 5300 and Broadcom chipsets). Hardware diversity is a known challenge for CSI reproducibility.
- Labeling Overhead: While unsupervised methods alleviate the need for labeled anomalies, high-quality evaluation still depends on curated ground truth, non-trivial in uncontrolled environments.
- Security Threat Coverage: We simulated common attacks like jamming and spoofing, but more sophisticated threats (e.g., side-channel RF exfiltration or adversarial examples) require future inclusion.
- This study did not include sophisticated or adaptive attack types such as adversarial RF perturbations, stealthy low-rate anomalies, or targeted evasion strategies. These scenarios are challenging to reproduce safely in lab environments and typically require specialized SDR hardware. Our focus was to establish a robust, reproducible baseline using representative anomalies (spoofing, jamming, rogue APs, and environmental perturbations). Future work will explicitly explore adversarial robustness, including signal-level perturbations and stealthy anomaly injections, to strengthen the resilience of CSI-based detection frameworks.
- Another limitation is the handling of temporal drift. While normalization, PCA preprocessing, and sliding-window inference partially mitigate distribution shifts, CSI data remains inherently non-stationary. Our framework was evaluated in cross-environment settings, but we did not incorporate online or adaptive learning for continuous deployment. Future work will focus on incremental and streaming learning approaches to handle long-term drift, prevent catastrophic forgetting, and support persistent deployment in evolving wireless environments.
- This study does not explicitly address temporal drift or long-term model degradation, which are well-known challenges in CSI-based systems. While preprocessing and cross-environment tests provided partial robustness, true continuous deployment requires mechanisms such as drift detection, online

adaptation, and incremental retraining. Incorporating these approaches is a key direction for future research toward making CSI anomaly detection sustainable in real-world, evolving wireless environments.

VI. Conclusion and Future Work

In conclusion, this work proposed a novel deep learning framework for CSI-based anomaly detection, demonstrating that temporal–spatial architectures significantly outperform classical baselines. The study introduced a reproducible CSI data collection pipeline and evaluated multiple models under both static and dynamic conditions. While promising, limitations remain regarding cross-hardware generalization, temporal drift handling, and adversarial robustness. Future research should address these challenges by exploring transfer learning for environment adaptation, multimodal fusion with complementary signals, adversarial robustness testing, and deployment in real-world enterprise or campus networks. This paper presented a novel, end-to-end framework for deep learning-based anomaly detection in wireless networks using CSI. By harnessing the spatial and temporal richness of CSI data, we demonstrated that deep models (especially attention-based architectures) can learn to detect subtle anomalies that evade traditional detection systems. Our ***contributions*** include:

- A reproducible CSI collection and preprocessing pipeline
- Supervised and unsupervised deep learning architectures tailored for CSI anomaly detection
- Real-time inference capability on low-cost embedded hardware
- Insightful visualization and interpretability tools for black-box models

Possible ***future works*** based on this manuscript include:

- Building on these foundations, we identify the following future directions:
- Transfer Learning and Cross-Environment Adaptation: Develop CSI anomaly detection models that generalize across locations, hardware, and scenarios using domain adaptation or contrastive learning.
- Multimodal Fusion: Combine CSI with other physical-layer metrics (RSSI, ToF) or network-layer data (packet timing, MAC activity) for multimodal anomaly detection, enhancing robustness.
- Adversarial Robustness: Study the susceptibility of deep CSI models to adversarial RF perturbations, and develop defense mechanisms for signal-level adversarial examples.
- Public CSI Anomaly Dataset Release: We plan to release our curated CSI dataset with normal and anomalous events to enable benchmarking and reproducibility in the research community.
- Deployment in Real-World Networks: Collaborate with enterprise or smart campus networks to embed the system into operational Wi-Fi infrastructure for live testing and iterative refinement.
- Exploring integration of CSI-based anomaly detection with intelligent cybersecurity frameworks powered by emerging Large Language Models, as discussed in [28].

While this paper introduced a new CSI anomaly dataset, the current version is not yet publicly released due to privacy, standardization, and clearance considerations. We have, however, provided sufficient methodological detail for independent reproduction using widely available NICs. A curated and anonymized version of the dataset, along with preprocessing scripts and model checkpoints, will be made publicly available through an open repository upon completion of institutional approvals, enabling benchmarking and further community-driven research.

Finally, this work positions CSI not just as a passive sensing signal, but as a proactive and intelligent security signal for next-generation wireless networks—bridging the gap between physical-layer dynamics and network-layer intelligence.

Conflict of Interest

The authors declare that they have no conflict of interest relevant to this article.

References

[1] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures", Proceedings *of the 20th annual international conference on Mobile computing and networking* 2014, https://doi.org/10.1145/2639108.2639143

[2] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using Wi-Fi", Proceedings *of the 2015 ACM conference on special interest group on data communication*, 2015, https://doi.org/10.1145/2785956.2787487

[3] F. Wang, J. Han, S. Zhang, X. He, and D. Huang, "Csi-net: Unified human body characterization and pose recognition", *arXiv preprint arXiv:1810.03064*, vol. 18, no. 11, pp. 2714-2724, 1 Nov. 2019, https://doi.org/10.48550/arXiv.1810.03064

[4] Z. Chen, L. Zhang, C. Jiang, Z. Cao, and W. Cui, "Wi-Fi CSI based passive human activity recognition using attention based BLSTM", *IEEE Transactions on Mobile Computing* vol.18, no.11, pp. 2714-2724, 2018, https://doi.org/10.1109/TMC.2018.2878233

[5] Z. Chen, C. Cai, T. Zheng, J. Luo, J. Xiong, and X. Wang, "RF-based human activity recognition using signal adapted convolutional neural network", *IEEE Transactions on Mobile Computing*, vol. 22, no. 1, pp. 487-499, 2021, https://doi.org/10.1109/TMC.2021.3073969

[6] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey", *ACM computing surveys*, vol. 41, no. 15, pp. 1-58, 2009, https://doi.org/10.1145/1541880.1541882

[7] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques", *Journal of Network and Computer Applications*, vol. 60, pp. 19-31 2016, https://doi.org/10.1016/j.jnca.2015.11.016

[8] Y. Zhao, Z. Nasrullah, and Z. Li, "Pyod: A python toolbox for scalable outlier detection", *Journal of*

machine learning research, vol. 41, no. 3, 2019, https://www.jmlr.org/papers/v20/19-011.html

[9] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial WiFi devices", *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1118-1131, 2017, https://doi.org/10.1109/JSAC.2017.2679658

[10] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces", *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN). IEEE*, pp. 1-12 2016, https://doi.org/10.1109/IPSN.2016.7460727

[11] M. S Aljumaily, and G. A. Al-Suhail. "Towards ubiquitous human gestures recognition using wireless networks", *International Journal of Pervasive Computing and Communications*, vol.13, no. 4, pp. 408-418, 2017, https://doi.org/10.1108/IJPCC-D-17-00005.

[12] J. Xiao, K. Wu, Y. Yi, and L. M. Ni, "FIFS: Fine-grained indoor fingerprinting system", *2012 21st international conference on computer communications and networks (ICCCN). IEEE*, pp. 1-7 2012, https://doi.org/10.1109/ICCCN.2012.6289200

[13] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "CSI-based indoor localization." *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300-1309, 2012, https://doi.org/10.1109/TPDS.2012.214

[14] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system", *2015 IEEE conference on computer communications (INFOCOM). IEEE*, pp. 1472-1480, 2015, https://doi.org/10.1109/INFOCOM.2015.7218525

[15] L. Xiao, L. Greenstein, N. Mandayam and W. Trappe, "A Physical-Layer Technique to Enhance Authentication for Mobile Terminals," *2008 IEEE International Conference on Communications, Beijing, China*, pp. 1520-1524, 2008, doi: 10.1109/ICC.2008.294

[16] F. Elia, "Anomaly detection using WiFi signals of opportunity", *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS). IEEE*, pp. 1-7, 2019, https://doi: 10.1109/ICSPCS47537.2019.9008700

[17 R. Kong, and He Chen. "DeepCRF: Deep learning-enhanced CSI-Based RF fingerprinting for channel-resilient WiFi device identification", *IEEE Transactions on Information Forensics and Security (2024)*, vol. 20, pp. 264-278, 2025, https://doi: 10.1109/TIFS.2024.3515796

[18] W. Wei, "Understanding and modeling of wifi signal based human activity recognition", Proceedings *of the 21st annual international conference on mobile computing and networking*, pp. 65 – 76, 2015, https://doi.org/10.1145/2789168.279009

[19] N. Dash, S. Chakravarty, A. Rath, "An optimized LSTM-based deep learning model for anomaly network intrusion detection", *Scientific Reports 15.1* https://doi.org/10.1038/s41598-025-85248-z

[20] Z. Tianyu, F. Li, and P. Tian. "A deep-learning method for device activity detection in mMTC under imperfect CSI based on variational-autoencoder", *IEEE Transactions on Vehicular Technology* 69.7 (2020): 7981-7986. vol. 69, no. 7, pp. 7981-7986, July 2020, https://doi: 10.1109/TVT.2020.2992080

[21] L. Hyeon-Ju, and S. Buu. "Wi-Fi-enabled Vision via Spatially-variant Pose Estimation based on Convolutional Transformer Network", *IEEE Access* vol. 13, pp. 84855-84869, 2025, https://doi: 10.1109/ACCESS.2025.3568505

[22] D. Halperin, W. Hu, A. Sheth, D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information", ACM *SIGCOMM computer communication review*, vol. 41, no. 1, 2011, https://doi.org/10.1145/1925861.192587

[23] F. Gringoli, M. Schulz, J. Link, M. Hollick. "Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets", *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*. 2019, https://doi.org/10.1145/3349623.3355477

[24] M. Shervin. "A Domain Adaptation Framework for Human Activity Monitoring Using Passive Wi-Fi Sensing", *2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA). IEEE*, 2023, https://doi: 10.1109/ICCCMLA58983.2023.10346849

[25] L. Tony, K. Ting and Z. Zhou. "2008 eighth ieee international conference on data mining", *2008 Eighth IEEE International Conference on Data Mining*, Dec, 2008, https://doi.org/10.1109/icdm14818.2008

[26] H. Muhammad "Domain adaptation for cross-domain alignment in human activity recognition using device-free sensing", *Diss. The University of St Andrews*, 2025, https://doi.org/10.17630/sta/1348

[27] A. Fadhil, N. Din, N. Aripin., and A. Abed, "Secure AODV routing strategies in smart cities for vehicular communication", *Journal Européen des Systèmes Automatisés*, vol. 57, no. 3, pp. 861-867, 2024, https://doi.org/10.18280/jesa.570325

[28] Z. Aziz and A. Abed, "Harnessing Large Language Models for Enhanced Cybersecurity: A Review of Their Role in Defending Against APT and Cyber Attacks", *International Journal of Mechatronics, Robotics, and Artificial Intelligence (IJMRAI)*, vol. 1, issue 1, pp. 54-62, June 2025, https://doi.org/10.33971/ijmrai.1.1.8