future internet

MDPI

Article

Multi-WiIR: Multi-User Identity Legitimacy Authentication Based on WiFi Device

Zhongcheng Wei [1,2,*] and Yanhu Dong [1,2]

[1] School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, China; q1078971707@gmail.com
[2] Hebei Key Laboratory of Security and Protection Information Sensing and Processing, Handan 056038, China
* Correspondence: weizhongcheng@hebeu.edu.cn

Abstract: With the proliferation of WiFi devices, WiFi-based identification technology has garnered attention in the security domain and has demonstrated initial success. Nonetheless, when untrained illegitimate users appear, the classifier tends to categorize them as if they were trained users. In response to this issue, researchers have proposed identity legitimacy authentication systems to identify illicit users, albeit only applicable to individual users. In this article, we propose a multi-user legitimacy authentication system based on WiFi, termed Multi-WiIR. Leveraging WiFi signals, the system captures users' walking patterns to ascertain their legitimacy. The core concept entails training a multi-branch deep neural network, designated WiIR-Net, for feature extraction of individual users. Binary classifiers are then applied to each user, and legitimacy is established by comparing the model's output to predefined thresholds, thus facilitating multi-user legitimacy authentication. Moreover, the study experimentally investigated the impact of the number of legitimate individuals on accuracy rates. The results demonstrated that The Multi-WiIR system showed commendable performance with low latency, being capable of conducting legitimacy recognition in scenarios involving up to four users, with an accuracy rate reaching 85.11%.

Keywords: WiFi sensing; channel state information (CSI); identity legitimacy authentication; multi-user recognition; multi-branch deep neural network

Citation: Wei, Z.; Dong, Y. Multi-WiIR: Multi-User Identity Legitimacy Authentication Based on WiFi Device. *Future Internet* **2024**, *16*, 127. https://doi.org/10.3390/fi16040127

Academic Editor: Gianluigi Ferrari

Received: 7 March 2024
Revised: 28 March 2024
Accepted: 29 March 2024
Published: 8 April 2024

1. Introduction

The widespread adoption of Internet of Things (IoT) technologies has spurred companies to offer more convenient and personalized services, which in turn requires accurate user identification. Traditional identification methods primarily rely on account passwords and identity markers [1]. Identity markers typically consist of account names and login codes, comprising a combination of numbers, letters, special symbols, and control symbols. For instance, login codes are commonly used to verify identity legitimacy during electronic login processes. Conversely, identity tokens are personal possessions utilized to activate electronic devices and store individual proprietary information for device identification. Examples include smart electronic cards and keys, commonly employed to access control facilities. However, traditional identification methods suffer from notable limitations: electronic cards are susceptible to loss or counterfeiting, while passwords are prone to being forgotten or stolen. Furthermore, these systems cannot differentiate between the rightful owner and an intruder who obtains the identifiers. Once another individual gains access to these identity markers, they can acquire the same privileges. Therefore, traditional identification technologies fall short of meeting demands in numerous scenarios.

Current identification methods mainly rely on biometric identification technology [2], which utilizes specialized sensing equipment to gather physiological and behavioral information. Subsequently, information processing technology is employed to match these features for identification purposes. Physiological information is intrinsic and developed in infancy, while behavioral information is habitual and cultivated in daily life; both possess

unique characteristics. Hence, biometric-based identification technology is not easily forgotten or lost, nor is it easily replicated or stolen. Moreover, it offers the advantage of user portability, enabling its utilization anytime and anywhere. Commonly employed biometric traits include iris patterns, fingerprints, facial features, and body shape. However, visual devices encounter limitations due to their inability to penetrate through walls, resulting in blind spots during facial and body image recognition. Moreover, they pose risks to personal privacy and may lead to severe privacy breaches [3]. Conversely, the utilization of biometric traits such as iris patterns and fingerprints for identification necessitates direct contact between the individual and specialized equipment. Additionally, the deployment of this equipment in advance may be required for certain temporary uses, with high associated costs [4].

In recent years, researchers have discovered that WiFi technology serves not only communication purposes but also as a means of sensing specific information. Through the utilization of received signal strength indicator (RSSI) and channel state information (CSI), various applications including indoor localization [5,6], fall detection [7,8], and monitoring physical activity [9,10] have been successfully implemented. CSI, being a form of fine-grained physical information, holds a notable advantage in its sensitivity to environmental changes, rendering it more effective in perception. Moreover, WiFi devices are cost-effective and do not require users to interact with or wear additional sensing apparatus, thus circumventing the reliance on visual and wearable devices. Additionally, WiFi sensing technology poses fewer intrusions into user privacy and enhances user comfort and security. Finally, researchers have modified the network card to facilitate stable and easy acquisition of CSI data from WiFi. Furthermore, the data we receive comprise continuous signals. Even in instances where CSI is not available or unavailable, the received data will be deemed erroneous and subsequently discarded during the decoding process. Following this, the next CSI segment will be decoded.

With the advancement of CSI-based perception technology, some scholars have integrated CSI into identity recognition systems [11–13] with notable success. However, certain shortcomings exist within current identification systems, notably the incapacity of the classification rules within classifiers to differentiate unknown samples. This limitation results in the system's inability to recognize intruders, thereby diminishing its practicality. Subsequent research efforts have addressed the issue of identity legitimacy authentication [14–16]. However, there exists an issue with the variability of the threshold for determination, which changes with variations in experimental groups, lacking adaptability. Furthermore, existing research only focuses on single-user scenarios. Therefore, their practical application poses greater challenges.

In prior studies [17], we effectively addressed the influence of changes in experimental group on judgment thresholds by introducing a particle swarm optimization (PSO) algorithm [18], yet this remained limited to single-user scenarios. To tackle the issue of multi-user legitimacy, this paper introduces a multi-user legitimacy authentication system named Multi-WiIR. The central idea leverages a multi-branch deep neural network, WiIR-Net, to extract features pertinent to individual users, and employs binary classifiers for each user. Ultimately, the legitimacy of each user is ascertained based on predefined thresholds derived from the model's output, thereby converting the multi-user legitimacy authentication issue into a single-user legitimacy authentication issue.

This paper contributes to the field in the following ways:

- We propose and implement Multi-WiIR, a multi-user legitimacy authentication system that detects the presence of trespassers in a scenario through the use of commercial WiFi devices. We evaluated the system in a real-world environment and showed that the system can concurrently authenticate the legitimacy of up to four users with an accuracy of 85.11%.
- We propose a multi-branch deep learning model termed WiIR-Net. It employs convolutional neural networks (CNN) as the backbone, with multiple BiLSTM [19] branches for feature extraction. Each branch is equipped with a binary classifier, and the legit-

imacy of each user is determined by comparing the output of each classifier with a predefined threshold. Consequently, the multi-user legitimacy authentication problem is transformed into a single-user legitimacy authentication problem.

- We conducted comparative experiments of operational efficiency and evaluation metrics between Multi-WiIR and various other models. The experimental outcomes demonstrated that the multi-branch architecture of Multi-WiIR enhanced the operational efficiency by 35.8% over multi-model multi-user legitimacy systems, and it surpassed the other classic models by nearly 5 percentage points in terms of performance.

The remaining sections of this paper are organized as follows: Section 2 provides an overview of related work. Section 3 introduces the fundamental principles of WiFi-based technology and the knowledge required for WiIR-Net. Our Multi-WiIR framework and its associated processes are presented in Section 4. Section 5 presents the experimental results. In Section 6, we discuss the limitations of our work and potential solutions. Finally, we conclude our work in Section 7.

2. Related Work

The utilization of WiFi devices for sensing applications has a significant historical background. In 2000, Bahl et al. [20] introduced the Radar system, pioneering the use of received signal strength (RSSI) for indoor localization, marking the inception of WiFi for sensing purposes. In 2012, Chetty et al. [21] achieved motion sensing through the analysis of Doppler frequency shifts. Similarly, in 2012, Halperin et al. [22] utilized a commercial network card CSI from commercial NICs for motion sensing, providing a more refined and stable foundation for WiFi-based identity sensing. In 2016, Zhang et al. [11] proposed WiFi-ID, pioneering the integration of WiFi sensing into identity recognition for the first time and leveraging CSI data. Since then, WiFi-based identification systems have undergone extensive development, including non-line-of-sight (NLOS) [12,13], cross-domain [23–25], and multi-user [26–28] WiFi identification approaches. However, in the aforementioned identification systems, the classifier's classification categories are predetermined. Consequently, when an unauthorized individual is present, their category does not match any known categories in the classifier. Despite this, the classifier erroneously classifies them as a known legitimate user. This limitation renders the system incapable of accurately determining the legitimacy of an individual's identity, leaving it vulnerable to intrusion by unauthorized individuals.

Despite the high accuracies achieved by the current identification techniques, the inability to recognize unauthorized individuals significantly restricts the applicability of these systems. In response to the challenge of authenticating personnel identity legitimacy, several researchers have incorporated this functionality into identification systems. Wang et al. [14] developed the WiFiU system, which categorizes training data into baseline personnel and target personnel. Utilizing a support vector machine (SVM) classifier, the system calculates the probability of an unknown gait instance belonging to the target personnel. Instances with probability values exceeding a threshold are considered legitimate personnel. The study in [15] presents the Wii system, which segregates a segment of the training set into legitimate and illegitimate personnel categories. While accomplishing identity recognition, the system constructs Gaussian models for both legitimate and illegitimate personnel to distinguish between them. Lin et al. [16] identified legitimate user identities and authenticated illegitimate users using specialized loss functions. However, these systems require the introduction of illegitimacy personnel data into the training set, posing challenges in realistic scenarios. Shi et al. [29] established a support vector model for each legitimate user and assessed the legitimacy of unknown personnel identities by comparing the distance between unknown personnel samples and legitimate user support vectors. Nevertheless, this approach is only suitable for a small number of users. Kong et al. [27] analyzed multipath WiFi signals to characterize users with separate CSI profiles. By training threat models with individual CSIs, they achieved an FAR of 8.8% and an FRR of 5.2%. However, this method failed in scenarios where users were positioned along

an ellipse with identical arrival times. GaitSense [30] clustered legitimate persons using k-nearest neighbors (KNN) and determined the legitimacy of a target user by assessing their proximity to legitimate persons. Wi-Sniffer [31] employed a trained deep learning model in combination with a decision tree for intruder detection, achieving a correct classification rate of 87% and an intruder detection rate of 95%.

However, these legitimacy authentication methods are only applicable to single users and not to multiple users. In conclusion, the existing legitimacy authentication systems do not meet the needs of real life.

3. Perception Principles and Key Technologies

To provide a comprehensive understanding of Multi-WiIR, this section introduces the principles of CSI and delves into the principles of WiFi sensing, CNNs, and BiLSTM technology.

3.1. Channel State Information (CSI)

CSI pertains to the physical layer data in wireless communication protocols and is not directly accessible. However, in recent years, researchers [22] have devised methods to acquire CSI from the physical layer in high throughput mode by leveraging OFDM (orthogonal frequency division multiplexing) technology as per the IEEE 802.11n standard. In this standard, CSI information is extracted from 30 specific subcarriers out of the total 64 subcarriers defined, with the aim of minimizing interference. By extracting CSI information, we can obtain the CSI matrix depicted in Equation (1),

$$H = [H(1)H(2)\dots H(30)]_{N_T \times N_R} \tag{1}$$

where N_T and N_R represent the number of antennas at the transmitter and receiver, respectively. Each antenna pair corresponds to a total of 30 carriers. Figure 1 illustrates the amplitude characteristics of 30 subcarriers generated for users by specific antennas. These magnitude features serve as our raw data and will be input into the subsequent preprocessing stages.

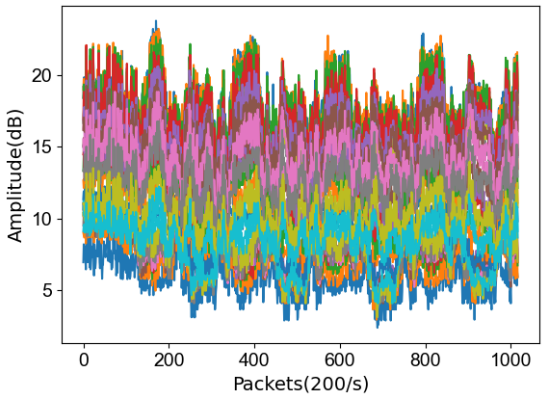


Figure 1. CSI amplitude features of user walking.

3.2. Principle of WiFi Sensing

The foundational premise of WiFi sensing is predicated on exploiting the effects that targets have on transmitted signals for the purpose of identification. As signals are emitted from a transmitter and navigate through the environment, they encounter a range of static and dynamic obstacles and entities, experiencing various phenomena such as reflection, refraction, and diffraction. Gait information quintessentially mirrors the behavioral traits correlated with an individual's identity; divergent walking patterns among users induce specific alterations in signal propagation. By meticulously examining the characteristics of these altered signals, we are able to infer the legitimacy of the user. By analyzing these signal characteristics, one can infer human activity states. This analytical process is

delineated through the intricate analysis of CSI, which represents the link state from the transmitter to the receiver.

Assuming X and Y denote the transmitted and received signals, this can be modeled as follows:

$$Y = HX + N \tag{2}$$

Here, H denotes the channel matrix and N denotes the environmental noise. WiFi signals undergo reflection from diverse objects within the environment, generating a multipath effect, as illustrated in Figure 2. The signal exhibits delay, fading, and frequency spreading on different paths, thus showing distortion at the receiver.

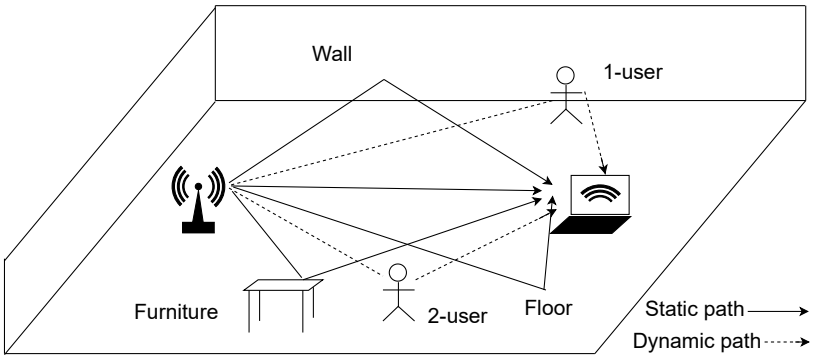


Figure 2. Illustration of the multipath effect.

From Figure 2, we can see that the signal at the receiver is reflected from multiple paths, which can be categorized into two types: static paths and dynamic paths. Therefore, the channel frequency response (CFR) can be derived as follows:

$$H(f,t) = \sum_{s=1}^{S} a_s^k(f,t) + \sum_{d=1}^{D} a_d^k(f,t) \tag{3}$$

where $H(f,t)$ is the CFR at moment t, carrier frequency f, and $\sum_{s=1}^{S} a_s^k(f,t)$ and $\sum_{d=1}^{D} a_d^k(f,t)$ represent the CFR of the static and dynamic environments, respectively. $a_s^k(f,t)$ and $a_d^k(f,t)$ denote the initial signal attenuation and phase shift of the k-th path on the static and dynamic path at moment t, the carrier frequency f, and S and D are the number of static and dynamic paths.

3.3. Convolutional Neural Network (CNN)

CNN [32], a fundamental neural network, find widespread application in diverse fields, such as image classification [33], object detection [34], and speech recognition [35]. The primary concept behind CNN is to extract features from data, such as images, through convolutional and pooling layers, followed by performing classification or regression tasks using fully connected layers. Compared to traditional fully connected neural networks, CNN offer several advantages in processing image data, including parameter sharing, local perceptibility, and hierarchical structure. With the ability to automatically learn features, CNN excel in handling large-scale datasets.

3.4. Bi-Directional Long Short-Term Memory (BiLSTM)

Wi-Fi signals, typically being time-series data, are best classified using recurrent neural networks (RNN). However, traditional RNNs, along with their variants such as long short-term memory (LSTM) networks [36] and gated recurrent units (GRU) [37], can only capture past information. For tasks like activity recognition that require consideration of sequential actions, both past and future information are equally important. Hence, the BiLSTM was introduced as a foundational neuronal structure. BiLSTM comprises forward

and backward layers capable of extracting temporal features from both past and future data, thereby enhancing recognition accuracy, as depicted in Figure 3. In this architecture, the hidden state of BiLSTM at time point is represented by

$$h_t = \overrightarrow{h_t} \oplus \overleftarrow{h_t} \tag{4}$$

where h_t represents the hidden state of the unit, $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ denote the forward and backward states, respectively, while $\oplus$ signifies the concatenation operation.

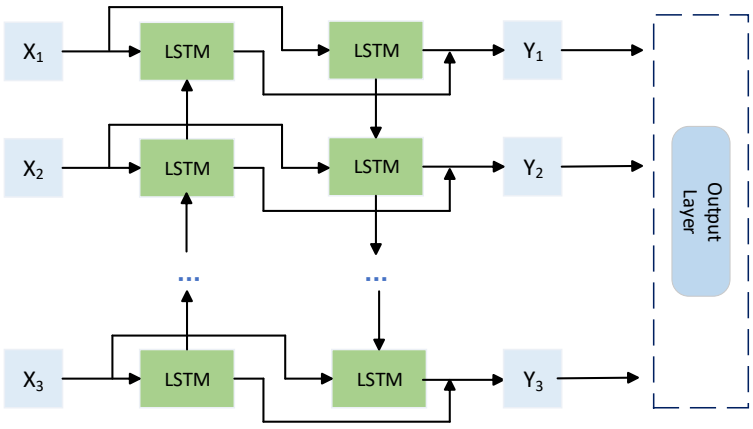


Figure 3. BiLSTM network framework.

4. Multi-WiIR

4.1. System Overview

In this section, we elaborate on the Multi-WiIR system, as depicted in Figure 4, Multi-WiIR comprises four main modules:

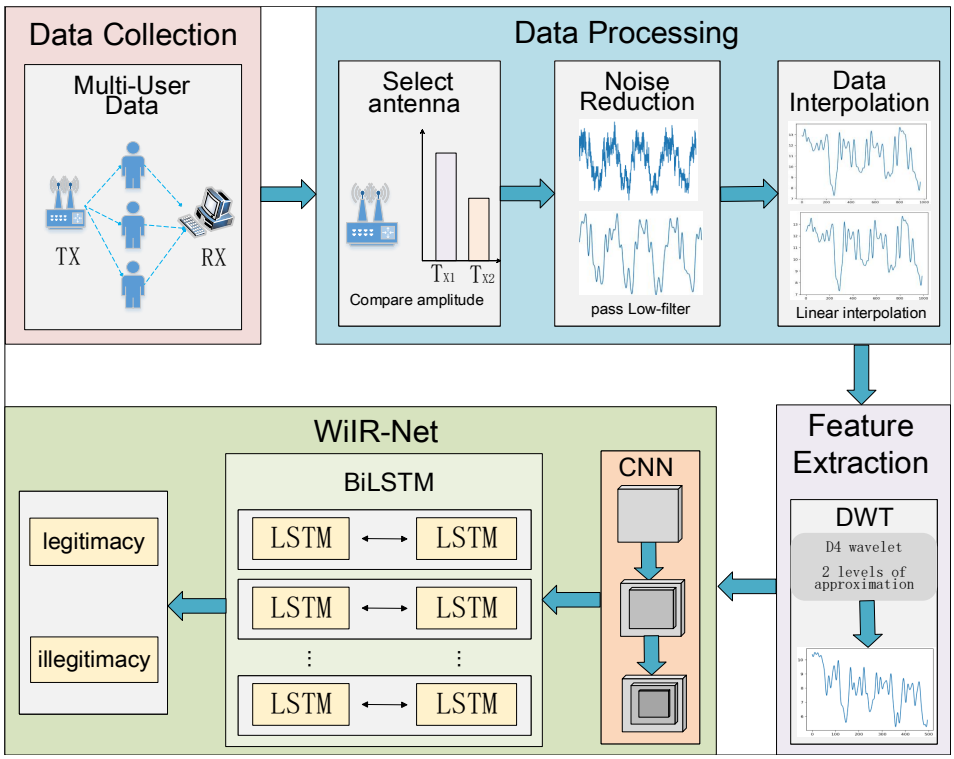


Figure 4. Multi-WiIR system framework.

(1) Data Collection: Responsible for gathering CSI data from both single-user and multi-user scenarios, this module labels the data to differentiate between legitimate and illegitimate individuals, storing them in the data collection module.
(2) Data Processing: During the data processing stage, collected data undergo a series of processes, including antenna selection, outlier detection, filtering, and linear interpolation. The processed data are then fed into the feature extraction module.
(3) Feature Extraction: The feature extraction module utilizes the D4 wavelet in the discrete wavelet transform (DWT) algorithm to process the processed CSI amplitude, calculating approximation coefficients, which are then input into the deep learning model for training.
(4) Multi-user Legitimacy Authentication: Built upon a multi-branch deep neural network comprising a multi-layer convolutional neural network and bidirectional LSTM layers, this module extracts features for each user using each branch of the WiIR-Net. Multiple binary classifiers are employed to determine the legitimacy of each user, achieving multi-user legitimacy recognition.

4.2. Data Collection

Data collection was conducted using an Ubuntu operating system equipped with an Intel 5300 Network Interface Card (NIC). The CSI tool was installed on this system to capture CSI packets. Data collection was carried out separately for single-player and multi-player scenarios. Each collected dataset was labeled to differentiate between legitimate and illegitimate individuals. Subsequently, all collected data were input into the data processing module for further processing.

4.3. Data Processing

In this section on data processing, we focus on four key aspects: antenna selection, outlier detection and filtering, linear interpolation, and feature extraction for CSI signals.

4.3.1. Antenna Selection

The sensitivity of a transmitting antenna to the environment can vary due to the existence of the Fresnel zone. To address this, we selected the transmitting antenna based on the variance in the amplitude. As illustrated in Figure 5, among the three receiving antennas, the amplitudes of the signals received by T_{X1} in R_{X1} were slightly larger than those received by T_{X2}. However, in R_{X2} and R_{X1}, the amplitude of the signals received by T_{X2} was significantly larger than that received by T_{X1}. This suggests that T_{X2} exhibited a significantly higher sensitivity than T_{X1}. Therefore, we opted to use T_{X2} as the transmitting antenna for our data collection.

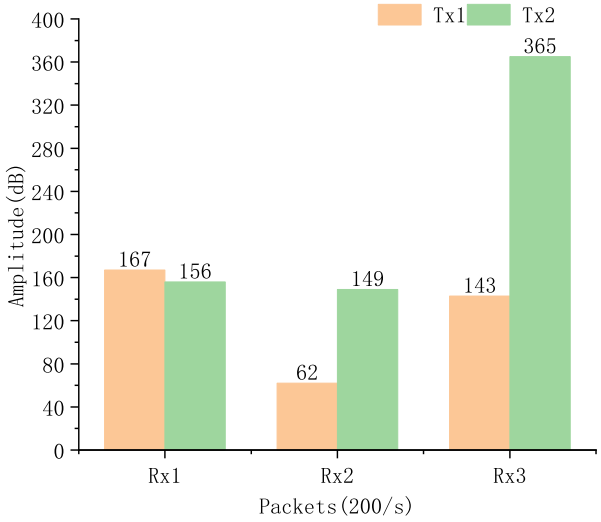
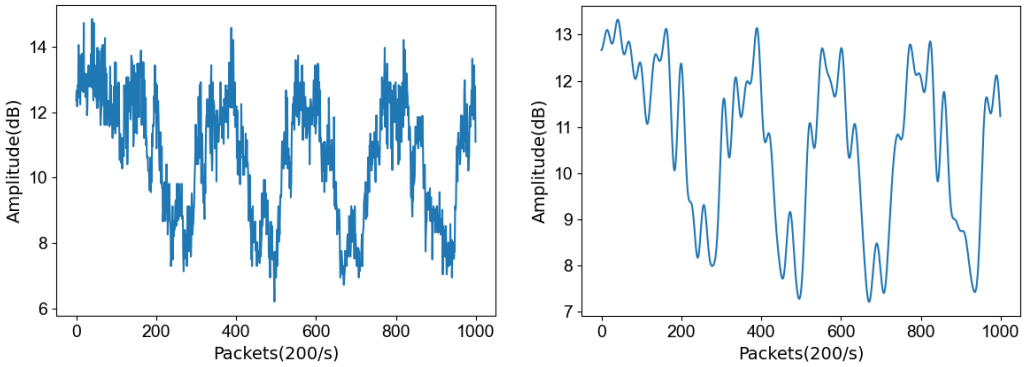


Figure 5. Comparison of transmit antenna sensitivity.

4.3.2. Outlier Detection and Filtering

Hardware defects in a device can lead to the presence of outliers in the signal, which are points deviating from the general level. A Hampel filter effectively removes outliers by setting predetermined data interval ranges. Additionally, environmental noise can introduce high-frequency pulses and bursts of noise into a signal. Since human activity typically manifests as frequency distributions in the low-frequency band around 10 Hz, a low-pass filter [38] can effectively eliminate environmental high-frequency noise, ensuring the accuracy and reliability of the subsequent signal analysis. The signal changes before and after filtering are illustrated in Figure 6.



(**a**) (**b**)

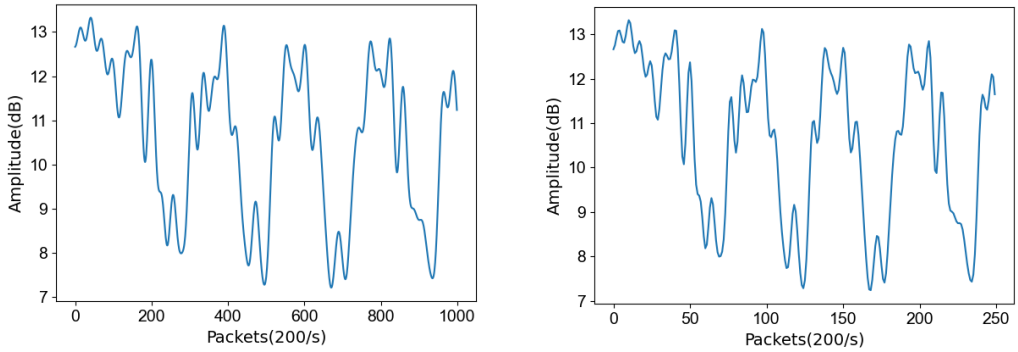
Figure 6. Comparison of amplitude low-pass filtering. (**a**) CSI before low-pass filtering; (**b**) CSI after low-pass filtering.

4.3.3. Data Interpolation

During the data acquisition process, data loss or time delay frequently arises, necessitating data interpolation to obtain CSI data uniformly distributed in time. The linear interpolation algorithm analyzed CSI data within a finite interval of values and calculates adjacent K values to estimate the approximation of discontinuous points for interpolation, thereby achieving temporal dimension uniformity.

4.4. Feature Extraction

CSI data contain abundant human motion information, but direct recognition entails substantial computation and often falls short of ideal accuracy expectations. Therefore, it is essential to perform feature extraction operations on effective activity segments within CSI data. Wavelet approximation coefficients possess the capability to extract key features from signals and mitigate data redundancy compared to time-domain and frequency-domain features. The wavelet transform effectively preserves highly variable features such as pulses and peaks of the original waveform, rendering them highly representative. Hence, we employed the discrete wavelet transform (DWT) algorithm to conduct wavelet decomposition of the CSI data and compute its approximation coefficients. As illustrated in Figure 7, This method significantly reduced the data volume without altering waveforms, thereby enhancing the efficiency of feature extraction. Compared to traditional feature extraction methods, wavelet-based feature extraction can more effectively capture dynamic information and patterns of change in CSI data, thus providing a more reliable foundation for subsequent classification and recognition tasks.



(**a**) (**b**)

Figure 7. Comparison before and after wavelet transform. (**a**) CSI before DWT; (**b**) CSI after DWT.

4.5. WiIR-Net

The WiIR-Net network model comprises multiple CNN convolutional layers and five BiLSTM branches, each dedicated to a single legitimate user. WiIR-Net employs a multi-layer convolutional network with an activation function to extract feature information from all users. Subsequently, it extracts unique features for each user through five independent BiLSTM layers. To simultaneously determine the legitimacy of multiple users, we employ five binary classifiers. Initially, features of each user are extracted using the network, and these features are then input into the five binary classifiers. By comparing the outputs of these classifiers with a threshold, we predict the target user as legitimate if the output exceeds the threshold. Conversely, if it falls below the threshold, we predict the user as illegitimate. This approach optimally utilizes the individual user identity information extracted by the deep learning model and the discrimination capability of binary classifiers to achieve multi-user legitimacy authentication. Using five binary classifiers for multi-user identification offers the advantage of enhancing system performance. In the subsequent experiments, we conducted comparative studies to validate this concept. By aggregating these judgments, we obtained the final multi-user identification result. This approach effectively addresses multi-user scenarios and provides more accurate identification capabilities. The model is depicted in Figure 8 below.

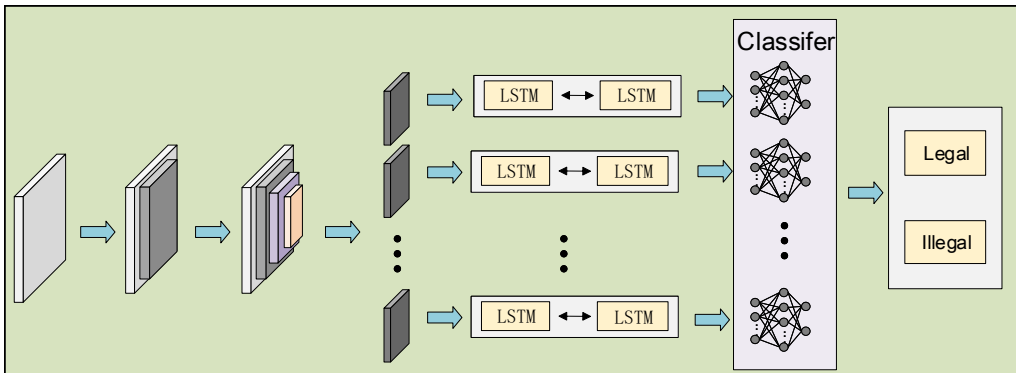


Figure 8. WiIR-Net network model.

5. Experiment Results

In this section, we outline the experimental setup and evaluate the performance of Multi-WiIR. To assess the effectiveness of WiSen-Net, we conducted comparisons with various models, utilizing standard evaluation metrics such as accuracy, precision, recall, and F1-score.

Accuracy: Accuracy is defined as the ratio of correctly predicted positive and negative cases to the total number of cases, and is calculated using Equation (5), below. The overall accuracy rate represents the average accuracy across all users.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{5}$$

Precision: Precision measures the proportion of predicted positive samples that are actually positive. It considers the accuracy of positive predictions and is computed using Equation (6) below. The overall precision rate represents the average precision across all users.

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

Recall: Recall indicates the proportion of actual positive samples that are correctly predicted as positive. It evaluates the completeness of positive predictions and is computed using Equation (7) below. The overall recall rate represents the average recall across all users.

$$\text{Re}call = \frac{TP}{TP + FN} \tag{7}$$

F1-score: The F1-score employs a harmonic mean instead of an arithmetic mean to balance precision and recall. Unlike an arithmetic mean, where each side equally contributes, to value growth and decline, a harmonic mean favors smaller values during growth and penalizes extreme cases where the precision and recall greatly differ. It achieves a balanced trade-off between precision and recall. The F1-score is calculated using Equation (8) below. The overall F1-score represents the average F1-score across all users.

$$F1\text{-}score = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

Accuracy and F1 score were the primary metrics for assessing the overall performance of the system. Accuracy measured the proportion of samples correctly predicted by the system, while the F1 score combined precision and recall, balancing the classification performance for both positive and negative classes. They provided a comprehensive evaluation of both the prediction accuracy and balance across the entire dataset.

5.1. Experimental Setup

As there was no publicly available dataset for multi-user based legitimacy authentication, we curated our own dataset for validation purposes. The experimental setup is depicted in Figure 9.

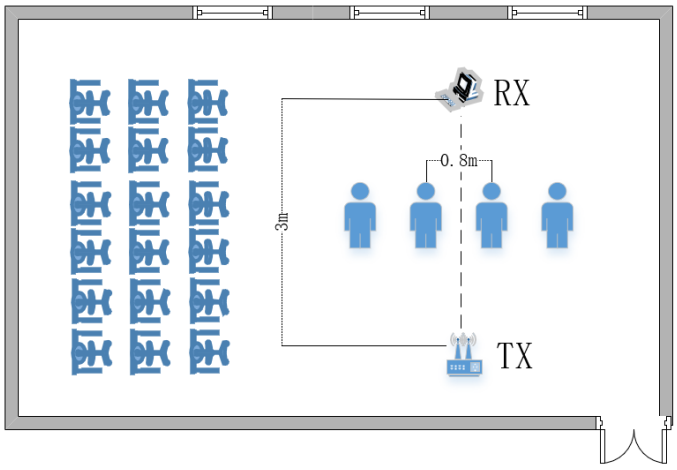


Figure 9. Experimental environment.

In our study, we employed a TP-Link router as the transmitter and a computer equipped with an Intel 5300 NIC as the receiver. The implementation was carried out using the Python 3.8 programming language in conjunction with the PyTorch framework. The training process took place on a server equipped with an NVIDIA 3060 GPU. Each batch during training comprised 64 samples, and the model underwent training for a total of 100 epochs. For detailed information regarding the experimental equipment and parameters, please refer to Table 1.

The experiment involved data collection from 12 volunteers, comprising 7 males and 5 females aged between 20 and 30 years. We collected data from 5 legitimate individuals, each contributing 200 entries as single-user data. Additionally, data collection was also conducted on legitimate and illegitimate combinations of different user counts, with each group comprising 180 records, resulting in a total of 3580 records. For example, for a triple-user combination, we collected data for three scenarios: all three being legitimate users, two legitimate users with one illegitimate user, one legitimate user with two illegitimate users, and three illegitimate users. Each scenario involved the collection of 180 records. Regarding the determination of legitimacy or illegitimacy, we considered that as long as there was at least 1 single-user in the crowd who was legitimate, then it could be considered that these people, with the permission of the legitimate individual, were also legitimate. The basic information of the volunteers is provided in Table 2. During the experiment, the volunteers maintained a distance of 0.8 m and performed stepping movements with a natural posture for approximately 5 s each time.

Table 1. Equipment parameters.

Item	Value
Transmitter (Tx)	TP-Link Router
Receiver (Rx)	Intel 5300 NIC
Transmitting antennas	2
Receiving antennas	3
Number of subcarriers	30
Working frequency antennas	5 GHz
Packet frequency	200 Hz

Table 2. Specific personnel information.

Volunteer Number	Sex	Height (cm)	Weight (kg)	Age	Legitimacy/Illegitimacy
1	female	165	56	23	legitimacy
2	male	180	81	28	legitimacy
3	female	163	59	22	legitimacy
4	male	181	82	24	legitimacy
5	male	182	82	25	legitimacy
6	female	164	55	23	illegitimacy
7	female	158	51	25	illegitimacy
8	male	175	70	21	illegitimacy
9	male	178	76	24	illegitimacy
10	female	166	53	23	illegitimacy
11	male	178	79	26	illegitimacy
12	male	175	75	24	illegitimacy

5.2. Performance Evaluation

5.2.1. Single-User Legitimacy Authentication

We initially partitioned 70% of the legitimate single-user dataset for training, allocating the remaining 30% of the legitimate and illegitimate single-user data for testing to evaluate the system's performance in single-user legitimacy authentication, particularly its effectiveness in user legitimacy authentication. The evaluation results are elucidated using a confusion matrix in Figure 10a.

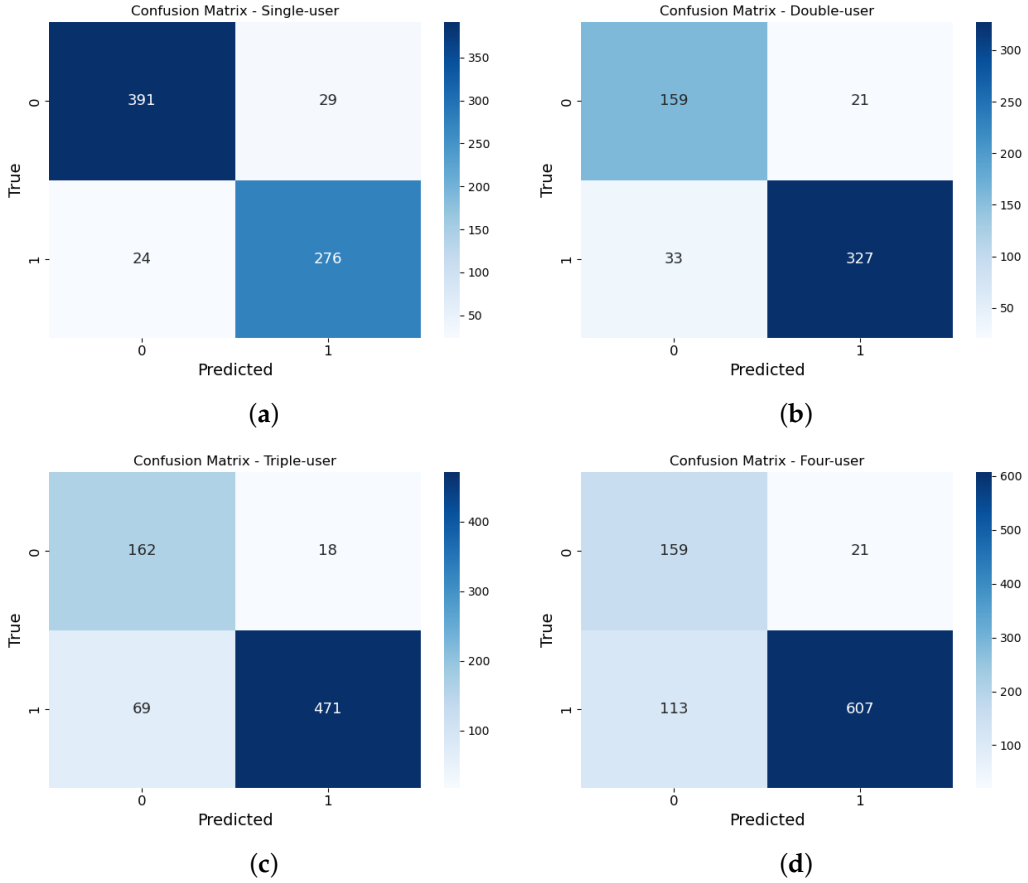


Figure 10. Confusion matrix for legitimacy authentication across different numbers of users. (**a**) Single-user legitimacy authentication; (**b**) double-user legitimacy authentication; (**c**) triple-user legitimacy authentication; (**d**) four-user legitimacy authentication.

5.2.2. Multi-User Legitimacy Authentication

In the multi-user legitimacy authentication, we conducted separate tests in scenarios involving double-user, triple-user, and four-user scenarios. The specific procedures were as follows: we initially partitioned 70% of the legitimate single-user dataset for training. The data on legitimate and illegitimate combinations of different user counts were used for testing to evaluate the system's performance in the following aspects in multi-user legitimacy authentication. The evaluation results are elucidated using a confusion matrix in Figure 10b–d.

The overall performance, as depicted in Figure 11, showed that the Multi-WiIR achieved accuracy rates of 92.64%, 90%, 87.91%, and 85.11% in the single-user, double-user, triple-user, and four-user scenarios, respectively. This result indicated that our accuracy remained, even in private settings. This achievement can be attributed to the utilization of a multi-branch approach within our system for extracting the features of a single-user, thereby mitigating potential interference from other users. Notably, within the context of triple-user and four-user scenarios, the precision significantly surpassed the other scenarios, owing to the lower presence of illicit users juxtaposed with a higher prevalence of legitimate users, consequently resulting in the highest precision upon reevaluation. Importantly, within this model, the overall false positive rate stood at a mere 3.2%, thus decisively validating the system's efficacy in preempting unauthorized intrusions.

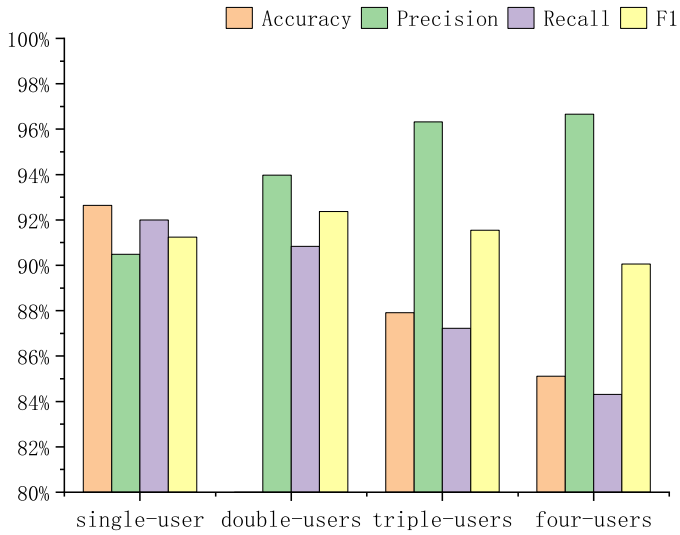


Figure 11. Overall performance of Multi-WiIR with different numbers of users.

5.2.3. Comparison of Different Models

(1) Branch model comparison

Our Multi-WiIR model utilizes a convolutional neural network (CNN) as a backbone for extracting initial features, followed by the extraction of individual user features through a multi-branch BiLSTM mechanism, considering both past and future information. To demonstrate the superiority of our approach using the BiLSTM model, we compared our model with multi-branch LSTM, GRU, and traditional CNN architectures.

As depicted in Figure 12, the average accuracy achieved by our BiLSTM model in multi-user scenarios was 88.54%, representing an improvement of nearly 5 percentage points over LSTM's 83.45% and GRU's 82.56%. This notable difference in performance substantiated the superiority of our selection of a BiLSTM architecture.

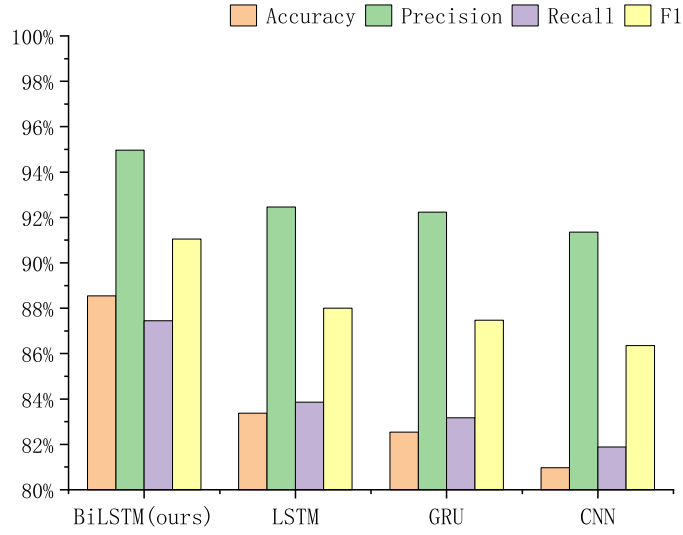


Figure 12. Comparison of the performance of the different models.

(2) Comparison of Running Efficiency

To assess the enhanced running efficiency of Multi-WiIR in contrast to multiple binary classification models, we conducted a comparative analysis of their training times and test set outcomes on the same computing platform. Figure 13 illustrates the comparison of the time consumption by both systems across various training epochs, ranging from 10 to 40 epochs. The average training time for Multi-WiIR was calculated to be 0.5088 s per

epoch, whereas the average training time for the multiple models amounted to 0.7924 s per epoch. This indicates that Multi-WiIR reduced the training time by 35.8% compared to the multi-model training duration. Moreover, it is noteworthy that the initialization time of the multi-model system significantly exceeded that of Multi-WiIR. In summary, our findings suggest that Multi-WiIR effectively enhanced the operational efficiency of the system when compared to the multi-model multi-user legitimacy authentication system.

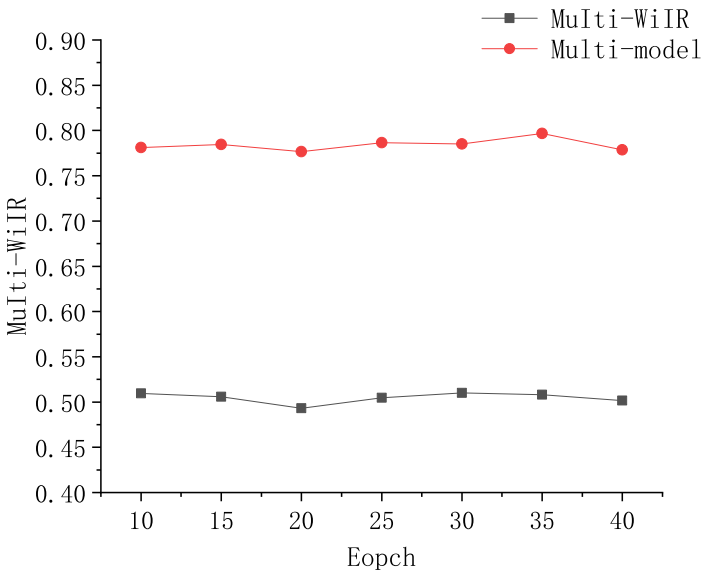


Figure 13. Comparison of model elapsed time.

(3) Comparison with Previous Legitimacy Authentication Systems

Table 3 provides a concise overview of other mature WiFi-based IoT applications. WiAU [16] necessitates the acquisition of information from untrained and unauthorized personnel, a practical infeasibility in real-world scenarios. Furthermore, the thresholds for identifying legitimate individuals necessitate adjustments with fluctuations in their numbers. Multiauth [27] requires the information of unauthorized individuals and imposes location constraints. GaitSense [30], similarly, requires adjustments to the threshold for identifying legitimate individuals with changes in the number of experimental subjects. Wi-Sniffer [31] extracts phase features, where the presence of Fresnel zones significantly influences phase variations, thus imposing location constraints. Moreover, all the aforementioned systems are tailored for single-user scenarios. In contrast, our Multi-WiIR system leverages the extraction of individual characteristics to ascertain legitimacy, unaffected by fluctuations in the number of authorized persons and capable of concurrently handling multiple individuals, irrespective of their positions.

Table 3. Comparison among works on legitimacy authentication.

System	Illegitimate Users' Information	Location Requirement	Adjust Threshold	Applicable Scenarios
WiAU	yes	no	yes	single-user
MultiAuth	yes	yes	no	single-user
GaitSense	no	yes	yes	single-user
Wi-Sniffer	no	yes	yes	single-user
Multi-WiIR	no	no	no	multi-user

6. Discussion

This approach only functions for a small number of users. As the number of users increases, the accuracy rate drops drastically, even when adding model branches. This is mainly due to the insufficient spatial resolution of WiFi signals, which leads to the mixing of signals from multiple users. Currently, there are two potential solutions: one is to improve the spatial resolution of the system by using more advanced antenna technology [39]; the other is to limit the direction of signal transmission in order to better spatially discriminate the signals of different users.

Environmental factors can affect system performance, and currently, we can only conduct experiments in trained environments, not across domains. When the environment changes, dynamic CSI containing human activity information may be obscured by environmental reflections of static CSI. Our team successfully achieved cross-domain activity recognition through the utilization of adversarial networks and domain classifiers for environment-independent activity identification [40]. Next, we will attempt to apply this approach to legitimacy recognition.

This approach will not work when CSI is unavailable or if there are errors in detecting that information.When channel estimation errors or factors such as pilot contamination render the received CSI unusable, the system may fail to operate smoothly. Noncoherent (NC) detection technology [41,42] may offer an effective solution to this problem. This technique does not require the receiver to accurately understand the channel's state information, but rather relies on the statistical properties of the signal for demodulation. In the face of complex channel conditions or communication environments, noncoherent detection technology can simplify system design and enhance system robustness.

7. Conclusions

This paper presented the Multi-WiIR framework designed for multi-user identity legitimacy authentication using commercial WiFi devices. The framework encompasses modules for data acquisition, preprocessing, feature extraction, and multi-user authentication. Utilizing a multi-branch deep neural network WiIR-Net, user features are extracted via deep learning techniques and legitimacy is determined through a binary classifier. Evaluation conducted in a real-world environment demonstrated that WiIR-Net could effectively authenticate the identity of up to four users, outperforming the alternative models. The findings of this study offer a robust framework for multi-user identification leveraging commercial WiFi devices, providing valuable insights for future research and practical applications.

Author Contributions: Conceptualization, Z.W. and Y.D.; writing—original draft, Y.D.; writing—review and editing, Z.W.; methodology, Z.W.; formal analysis, Z.W. and Y.D.; validation, Z.W. and Y.D.; software, Y.D.; funding acquisition, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Handan Science and Technology Research and Development Program under Grant (No. 21422031288), and the Provincial Innovation Funding Project for Graduate Students of Hebei Province, China under Grant (CXZZSS2024098).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on reques.

Acknowledgments: During the preparation of this work, the authors utilized ChatGPT 3.5 for grammar checking and English language enhancement. After using this tool, the author reviewed and edited the content as needed and take full responsibility for the publication's content.

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

References

1. Ibrahim, D.R.; Abdullah, R.; Teh, J.S.; Alsalibi, B. Authentication for ID cards based on colour visual cryptography and facial recognition. In Proceedings of the 3rd International Conference on Cryptography, Security and Privacy, Kuala Lumpur, Malaysia, 19–21 January 2019; pp. 164–167. [CrossRef]
2. Meltzer, D.; Luengo, D. Efficient Clustering-Based electrocardiographic biometric identification. *Expert Syst. Appl.* **2023**, *219*, 119609. [CrossRef]
3. Benedikt, L. Using 3D Facial Motion for Biometric Identification. Ph.D. Thesis, Cardiff University, Cardiff, UK, 2009.
4. Herath, S.; Harandi, M.T.; Porikli, F. Going deeper into action recognition: A survey. *Image Vis. Comput.* **2017**, *60*, 4–21. [CrossRef]
5. Li, X.; Zhang, D.; Lv, Q.; Xiong, J.; Li, S.; Zhang, Y.; Mei, H. IndoTrack: Device-Free Indoor Human Tracking with Commodity Wi-Fi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 1–22. [CrossRef]
6. Yan, B.; Cheng, W.; Li, Y.; Gao, X.; Liu, H. Joint activity recognition and indoor localization with WiFi sensing based on multi-view fusion strategy. *Digital Signal Process.* **2022**, *129*, 103680. [CrossRef]
7. Hu, Y.; Zhang, F.; Wu, C.; Wang, B.; Liu, K.J.R. DeFall: Environment-Independent Passive Fall Detection Using WiFi. *IEEE Internet Things J.* **2022**, *9*, 8515–8530. [CrossRef]
8. Xia, Z.; Chong, S. WiFi-based indoor passive fall detection for medical Internet of Things. *Comput. Electr. Eng.* **2023**, *109*, 108763. [CrossRef]
9. Tian, C.; Tian, Y.; Wang, X.; Kho, Y.H.; Zhong, Z.; Li, W.; Xiao, B. Human Activity Recognition with Commercial WiFi Signals. *IEEE Access* **2022**, *10*, 121580–121589. [CrossRef]
10. Han, B.; Wang, L.; Lu, X.; Meng, J.; Zhou, Z. Cross-modal meta-learning for WiFi-based human activity recognition. In Proceedings of the 29th Annual International Conference on Mobile Computing and Networking, ACM MobiCom 2023, Madrid, Spain, 2–6 October 2023; Costa-Pérez, X., Widmer, J., Perino, D., Giustiniano, D., Al-Hassanieh, H., Asadi, A., Cox, L.P., Eds.; ACM: New York, NY, USA, 2023; pp. 147:1–147:3. [CrossRef]
11. Zhang, J.; Wei, B.; Hu, W.; Kanhere, S.S. WiFi-ID: Human Identification Using WiFi Signal. In Proceedings of the International Conference on Distributed Computing in Sensor Systems, DCOSS 2016, Washington, DC, USA, 26–28 May 2016; IEEE Computer Society: Washington, DC, USA, 2016; pp. 75–82. [CrossRef]
12. Wu, Z.; Xiao, X.; Lin, C.; Gong, S.; Fang, L. WiDFF-ID: Device-Free Fast Person Identification Using Commodity WiFi. *IEEE Trans. Cogn. Commun. Netw.* **2023**, *9*, 198–210. [CrossRef]
13. Korany, B.; Cai, H.; Mostofi, Y. Multiple People Identification Through Walls Using Off-the-Shelf WiFi. *IEEE Internet Things J.* **2021**, *8*, 6963–6974. [CrossRef]
14. Wang, W.; Liu, A.X.; Shahzad, M. Gait recognition using wifi signals. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp 2016, Heidelberg, Germany, 12–16 September 2016; Lukowicz, P., Krüger, A., Bulling, A., Lim, Y., Patel, S.N., Eds.; ACM: New York, NY, USA, 2016; pp. 363–373. [CrossRef]
15. Lv, J.; Yang, W.; Man, D.; Du, X.; Yu, M.; Guizani, M. Wii: Device-Free Passive Identity Identification via WiFi Signals. In Proceedings of the GLOBECOM 2017—2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017; pp. 1–6. [CrossRef]
16. Lin, C.; Hu, J.; Sun, Y.; Ma, F.; Wang, L.; Wu, G. WiAU: An Accurate Device-Free Authentication System with ResNet. In Proceedings of the 15th Annual IEEE International Conference on Sensing, Communication, and Networking, SECON 2018, Hong Kong, China, 11–13 June 2018; pp. 109–117. [CrossRef]
17. Wei, Z.; Zhang, X.; Feng, H.; Lian, B.; Wang, W. Wi-Fi-based human legality verification system. *Comput. Eng. Des.* **2022**, *43*, 2423–2430. [CrossRef]
18. Lapizco-Encinas, G. Cooperative Particle Swarm Optimization for Combinatorial Problems. Ph.D. Thesis, University of Maryland, College Park, MD, USA, 2009.
19. Chen, Z.; Zhang, L.; Jiang, C.; Cao, Z.; Cui, W. WiFi CSI Based Passive Human Activity Recognition Using Attention Based BLSTM. *IEEE Trans. Mob. Comput.* **2019**, *18*, 2714–2724. [CrossRef]
20. Bahl, P.; Padmanabhan, V.N. RADAR: An In-Building RF-Based User Location and Tracking System. In Proceedings of the IEEE INFOCOM 2000, The Conference on Computer Communications, Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Reaching the Promised Land of Communications, Tel Aviv, Israel, 26–30 March 2000; IEEE Computer Society: Washington, DC, USA, 2000; pp. 775–784. [CrossRef]
21. Chetty, K.; Smith, G.E.; Woodbridge, K. Through-the-Wall Sensing of Personnel Using Passive Bistatic WiFi Radar at Standoff Distances. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1218–1226. [CrossRef]
22. Halperin, D.; Hu, W.; Sheth, A.; Wetherall, D. Tool release: Gathering 802.11n traces with channel state information. *Comput. Commun. Rev.* **2011**, *41*, 53. [CrossRef]
23. Zhang, J.; Chen, Z.; Luo, C.; Wei, B.; Kanhere, S.S.; Li, J. MetaGanFi: Cross-Domain Unseen Individual Identification Using WiFi Signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*, 152:1–152:21. [CrossRef]
24. Dai, M.; Cao, C.; Liu, T.; Su, M.; Li, Y.; Li, J. WiDual: User Identified Gesture Recognition Using Commercial WiFi. In Proceedings of the 23rd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing, CCGrid 2023, Bangalore, India, 1–4 May 2023; Simmhan, Y., Altintas, I., Varbanescu, A.L., Balaji, P., Prasad, A.S., Carnevale, L., Eds.; IEEE: Piscataway, NJ, USA, 2023; pp. 673–683. [CrossRef]

25. Li, C.; Liu, M.; Cao, Z. WiHF: Enable User Identified Gesture Recognition with WiFi. In Proceedings of the 39th IEEE Conference on Computer Communications, INFOCOM 2020, Toronto, ON, Canada, 6–9 July 2020; pp. 586–595. [CrossRef]
26. Ou, R.; Chen, Y.; Deng, Y. WiWalk: Gait-Based Dual-User Identification Using WiFi Device. *IEEE Internet Things J.* **2023**, *10*, 5321–5334. [CrossRef]
27. Kong, H.; Lu, L.; Yu, J.; Chen, Y.; Xu, X.; Tang, F.; Chen, Y. MultiAuth: Enable Multi-User Authentication with Single Commodity WiFi Device. In Proceedings of the MobiHoc '21: The Twenty-second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing, Shanghai, China, 26–29 July 2021; ACM: New York, NY, USA, 2021; pp. 31–40. [CrossRef]
28. Wei, Z.; Chen, W.; Dong, Y.; Lian, B.; Wang, W.; Zhao, J. Research on Multi-User Identity Recognition based on WiFi Sensing. *Chin. J. Internet Things* **2021**, *5*, 107–119.
29. Shi, C.; Liu, J.; Liu, H.; Chen, Y. Smart User Authentication through Actuation of Daily Activities Leveraging WiFi-enabled IoT. In Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing, Chennai, India, 10–14 July 2017; Moharir, S., Gopalan, A., Eds.; ACM: New York, NY, USA, 2017; pp. 5:1–5:10. [CrossRef]
30. Zhang, Y.; Zheng, Y.; Zhang, G.; Qian, K.; Qian, C.; Yang, Z. GaitSense: Towards Ubiquitous Gait-Based Human Identification with Wi-Fi. *ACM Trans. Sens. Netw.* **2022**, *18*, 1:1–1:24. [CrossRef]
31. Eom, J.Y.; Jang, S.U.; Jeon, W.S. Wi-Sniffer: Wifi-based intruder detection system using deep learning and decision tree. In Proceedings of the 97th IEEE Vehicular Technology Conference, VTC Spring 2023, Florence, Italy, 20–23 June 2023; pp. 1–7. [CrossRef]
32. LeCun, Y.; Bengio, Y.; Hinton, G.E. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
33. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
34. Zhao, Z.; Zheng, P.; Xu, S.; Wu, X. Object Detection with Deep Learning: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [CrossRef] [PubMed]
35. Aguirre-Peralta, J.; Rivas-Zavala, M.; Ugarte, W. Speech to Text Recognition for Videogame Controlling with Convolutional Neural Networks. In Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods, ICPRAM 2023, Lisbon, Portugal, 22–24 February 2023; Marsico, M.D., di Baja, G.S., Fred, A.L.N., Eds.; SCITEPRESS: Setúbal, Portugal, 2023; pp. 948–955. [CrossRef]
36. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
37. Chung, J.; Gülçehre, Ç.; Cho, K.; Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv* **2014**, arXiv:1412.3555. [CrossRef]
38. Xin, T.; Guo, B.; Wang, Z.; Li, M.; Yu, Z.; Zhou, X. FreeSense: Indoor Human Identification with Wi-Fi Signals. In Proceedings of the 2016 IEEE Global Communications Conference, GLOBECOM 2016, Washington, DC, USA, 4–8 December 2016; pp. 1–7. [CrossRef]
39. Xie, Y.; Zhang, Y.; Liando, J.C.; Li, M. SWAN: Stitched Wi-Fi ANtennas. In Proceedings of the 24th Annual International Conference on Mobile Computing and Networking, MobiCom 2018, New Delhi, India, 29 October–2 November 2018; Shorey, R., Murty, R., Chen, Y.J., Jamieson, K., Eds.; ACM: New York, NY, USA, 2018; pp. 51–66. [CrossRef]
40. Sheng, L.; Chen, Y.; Ning, S.; Wang, S.; Lian, B.; Wei, Z. DA-HAR: Dual adversarial network for environment-independent WiFi human activity recognition. *Pervasive Mob. Comput.* **2023**, *96*, 101850. [CrossRef]
41. Baeza, V.M.; Armada, A.G. Performance and Complexity Tradeoffs of Several Constellations for Non Coherent Massive MIMO. In Proceedings of the 22nd International Symposium on Wireless Personal Multimedia Communications, WPMC 2019, Lisbon, Portugal, 24–27 November 2019; pp. 1–6. [CrossRef]
42. Baeza, V.M.; Armada, A.G. Noncoherent massive MIMO. In *Wiley 5G Ref: The Essential 5G Reference Online*; John Wiley & Sons: Hoboken, NJ, USA, 2019; pp. 1–28.