Wihi: WiFi Based Human Identity Identification Using Deep Learning

**JIANYANG DING[ID], (Graduate Student Member, IEEE), YONG WANG[ID], (Member, IEEE),
AND XIANGCONG FU[ID]**
School of Telecommunications Engineering, Xidian University, Xi'an 710071, China

Corresponding author: Yong Wang (johnnyly@xidian.edu.cn)

ABSTRACT Human identity identification based on channel state information (CSI) using commercial WiFi devices has drawn increasingly attention, and it can be used in many applications such as smart home, intrusion detection, building monitoring, activity recognition, etc. However, most of the existing identity identification approaches are sensitive to the influence of random noise derived from indoor environments, and thus their identification accuracies are far from satisfactory. In the present paper, a device-free CSI based human identity identification approach using deep learning (Wihi) is proposed. Wihi mainly utilizes three key techniques to identify different people. Firstly, to eliminate the influence of the random noise, discrete wavelet transform (DWT) strategy is introduced to denoise raw CSI data by leveraging signal decomposition. Secondly, in order to characterize human's gaits profoundly, several representative features are exploited from different statistical profiles, including channel power distribution in time domain (CPD), time-frequency analysis (TFA), and energy distribution in different frequency bands (ED). Thirdly, a recurrent neural network (RNN) model with long short-term memory (LSTM) blocks is employed to learn the representative gait features extracted above and encode temporal information for realizing human identity identification. The proof-of-concept prototype of the proposed Wihi approach is implemented on a set of commercial WiFi devices, and multiple comprehensive experiments have been carried out to evaluate the performance of identity identification. The experimental results confirm that the proposed Wihi can achieve a satisfactory performance compared with some state-of-the-art approaches.

INDEX TERMS Human identity identification, commercial wifi devices, channel state information (CSI), recurrent neural network (RNN).

I. INTRODUCTION

Human identity identification has been researched for many years and is of great importance for many applications, such as smart home, indoor intrusion detection, building monitoring, etc. In order to identify different people, many identity identification approaches have been proposed with different techniques such as gait-based [1]–[3], fingerprint-based [4]–[6], face recognition-based [7], [8] and iris-based [9]–[11] approaches. Generally, these biological characteristics are representative and unique for everyone and can provide a high accuracy of identity identification, and therefore

they can be widely applied to security systems. For instance, the characteristics are able to be used in the security systems to conduct identity identification when someone has access to a certain office or laboratory. Although these identification approaches have shown great promise in their applications, they suffer from a number of limitations such as needing light, personal privacy problem, high energy consumption, high installation overhead, requiring the dedicated sensor or device, etc. Consequently, these disadvantages somehow restrict their large-scale deployment in door environments (e.g., smart home and office). With the rapid development and ubiquity of commercial WiFi devices in typical indoor environments, there are increasingly applications utilizing channel state information (CSI) [12]–[27]. Since an entity

walking between a pair of transmitter and receiver could generate significant impacts on the characteristics of WiFi signal, identity identification can be feasible utilizing WiFi CSI. In addition, owing to low power consumption, easy installation, no invasion, and large-scale deployment of commercial WiFi devices in indoor environments, identity identification is able to become reality. Furthermore, the passive device-free human identity identification approach using WiFi signal does not require users to take any sensor or device. Naturally, it is an ideal one compared with the traditional approaches.

It is well known that everyone's natural walking patterns (i.e. gaits) are particular, which can be characterized by the differences in human's height, body mass, and moving speed. When an entity walks in a target area, his/her gaits could affect the indoor electromagnetic environments in a unique manner which would be changed into the impacts on the characteristics of WiFi signal, and then the significant impacts are in turn manifested as distinct perturbation in WiFi CSI. Since human's gaits are highly distinguishing for different people, it is possible to identify one person from a group of people examining his/her representative statistics features exhibited in WiFi CSI data. However, there are still many challenges we face in WiFi CSI based human identity identification. The first one is how to obtain effective CSI data. The CSI data collected is obtained from commercial WiFi Network Interface Cards (NICs), and thus it contains the random noise from various sources, such as nearby electronic devices. This random noise is able to add false edges, and then further influence the accuracy of identity identification and robustness. Therefore, it is a challenging problem that how to preserve signal details while filtering out the noise components in the WiFi signal efficiently. The second one is how to obtain multiple representative features for characterizing human's gaits profoundly. The previous works [28], [29] have shown that human's walking patterns can be described by a number of statistics features in time and frequency domain. However, the common features such as median value, mean value, maximum value, minimum value, variance, and entropy are not effective because they can be influenced easily by the random noise. In view of this, the identification accuracy is not satisfactory. Now, a natural question to ask: which gait features are better for identity identification? The third one is how to utilize the representative gait features extracted to identify different people effectively. It should be noted that feature extraction and identity identification are not jointly optimized. That is to say, it is not enough to rely on the representative gait features alone, and we should seek for suitable method to conduct identity identification.

To deal with these challenges, in this paper, a passive device-free WiFi CSI based human identity identification approach using recurrent neural network (Wihi) is proposed. To address the first challenge, discrete wavelet transform (DWT) [30] strategy is employed to eliminate the influence of the random noise through signal decomposition. It helps to reduce the interference of both complicated background environment and the random noise effectively. After that,

the aim of improving identification quality can be achieved. To address the second one, we propose to extract three representative gait features in time and frequency domain for characterizing human's walking patterns profoundly, including channel power distribution in time domain (CPD), time-frequency analysis (TFA), and energy distribution in different frequency bands (ED). To address the third one, a recurrent neural network (RNN) [31]–[33] model with long short-term memory (LSTM) blocks is introduced to learn the representative gait features above and encode temporal information for identifying different people. As the term recurrent implies, the proposed RNN model takes not only the current input data but also several previous input data. In other words, it has a memory that obtains the variation in input data. Therefore, the proposed RNN model can capture the complicated non-linear relationship between input and output data in the training phase efficiently. As a result, an entity can be identified from a group of people accurately.

Real experiments have been conducted to verify the high performance of the proposed Wihi approach. Besides, the identification performance is also compared with the existing approaches. In summary, the main contributions of the paper are as follows:

1) We propose Wihi, a novel passive device-free CSI based human identity identification using deep learning, which is capable of identifying different people and achieves excellent identification performance compared with the existing state-of-the-art approaches.
2) DWT is introduced to eliminate the influence of the random noise presented in the raw WiFi CSI data while preserving data details through signal decomposition.
3) Unlike the existing CSI based approaches, we extract several representative gait features from both time and frequency domain, including CPD, TFA, and ED, which can better characterize human's walking patterns. Thus, this helps improving identity identification accuracy.
4) To identify different people accurately, the RNN model with LSTM blocks is used to identify different people by learning the gait features extracted, instead of raw CSI data. This can reduce the influence of the random noise derived from indoor environments significantly.

The remaining paper is organized as follows. We first review the related work in Section II, and Section III shows the basic background knowledge of WiFi CSI. Section IV illustrates the architecture and design of the proposed Wihi approach. Section V describes the raw CSI data collection for experiments and presents the experimental setups. Then, the experimental results are presented in this section. We discuss the limitations related to our approach in Section VI followed by a conclusion in Section VII.

II. RELATED WORK
Owing to the importance of human identity identification, a broad range of identity identification approaches have been

proposed these years that can applied to different indoor environments. Most identity identification approaches utilize biometric characteristics such as face, iris, fingerprint, and gaits. Especially, these biometric characteristics are widely used in user authentication because they are distinguishing among different people and very stable across different time. The researchers in [1] proposed a novel human identification approach from long range gaits profiles in surveillance videos. Concretely, they investigated the role of multi view gaits images acquired from multiple cameras, importance of infrared and visible range images in ascertaining identity, and role of soft/secondary biometric in enhancing the accuracy and robustness of the identification systems. Hossain and Chetty [2] proposed a novel multi-view feature fusion of gait biometric information in surveillance videos for large-scale human identification. A fingerprint classification algorithm can be found in [4], and fingerprints were classified into five categories: arch, tented arch, left loop, right loop and whorl. Then the algorithm extracted singular points in a fingerprint image, and further conducted classification based on the number and locations of the detected singular points. A new structural approach to the fingerprint classification problem was presented in [5]. The fingerprint directional image was segmented into multiple regions by minimizing the variance of the element directions within the regions. Jain *et al.* [6] presented a fingerprint classification algorithm which can achieve a better performance than previously reported. The proposed algorithm used a novel representation and was based on a two-stage classifier to conduct a classification. Yang *et al.* [8] presented a new robust face-matching method with multi-feature fusion, combining the rotation-invariant texture feature vector, the scale-invariant feature transform vector, and the convolution neural network. Park and Park [9] proposed a novel iris recognition method based on score level fusion which used two Gabor wavelet filters and SVM. Galdi *et al.* [10] got around the sensor interoperability problem utilizing on the picture differences due to acquisition by different sensors, and then presented a novel system that combined the recognition of user's iris and user's device. Lee *et al.* [11] proposed a novel recognition approach for noisy iris and ocular images by leveraging one iris and two periocular regions, based on three convolutional neural networks. Although the identification accuracies of these approaches above were relatively high, they all required dedicated device or sensor, which could lead to high cost and limit their wide deployment.

WiFi CSI has been verified to be a reliable indicator for passive device-free identity identification and it has the special advantages of low cost, no invasion, and wide deployment in indoor environments, and thus WiFi CSI based human identity identification approaches have been widely studied [15]–[27]. The authors in [15] proposed a novel approach for device-free passive detection of moving humans with dynamic Speed. Concretely, both amplitude and phase information of CSI were extracted and shaped into sensitive metrics for target detection, and then CSI across multi-antennas

in multiple input multiple output (MIMO) systems were further exploited to improve the detection accuracy and robustness. Xi *et al.* [16] presented a device-free based on CSI crowd counting approach, and this design was motivated by the observation that CSI was highly sensitive to the indoor environment variation. Wu *et al.* [17] proposed a unified approach for non-invasive detection of stationary and moving human using commercial WiFi devices. This approach took full use of both amplitude and phase information of CSI to detect stationary or moving targets. Lv *et al.* [18] proposed an accurate approach for speed independent device-free entity detection which was suitable for intrusion detection even when the entity's moving speed was relatively slow. Domenico *et al.* [19] presented a WiFi CSI based device-free crowd counting and occupancy estimation approach that can be leveraged in several typical indoor environments different from the ones in which the training process has been performed. Xin *et al.* [20] proposed a novel approach for human identification, which took advantage of WiFi signals to perform non-intrusive human identification in domestic environments. It is based on the observation that each person has distinguishing influence patterns to the surrounding WiFi signal while moving indoors, regarding their body shape characteristics and motion patterns. The researchers in [21] presented a passive WiFi CSI based identity identification approach utilizing human's gaits based on CSI of WiFi signals. Zou *et al.* [22] presented a human identification system that leveraged the measurements from existing WiFi-enabled Internet of Things (IoT) devices and produced the identity estimation via a novel sparse representation learning technique, and utilized the unique fine-grained gait patterns of each person revealed from the WiFi CSI measurements as the "fingerprint" for human identification. Wang *et al.* [23] designed a deep learning method to analyze the gait features using CSI of COTS WiFi devices. Specially, the convolution layers were combined with LSTM layers to extract gait features automatically from CSI data and to identify persons, which effectively reduced the need for a large amount of data preprocessing by manual feature extraction. Motivated by the observation that PHY layer CSI is capable of capturing the frequency diversity of wideband channel, Hong *et al.* [25] proposed a novel feature of subcarrier-amplitude frequency (SAF). Based on this feature, the proposed approach realized human identification through a linear-kernel SVM. Liu *et al.* [26] presented a fine-grained device-free framework that can distinguish different actions and identify persons within a short duration using WiFi signal. To extract intrinsic features from the noisy CSI so as to realize high-performance device-free identification (DFI), Wang *et al.* [27] proposed a novel empirical-mode-decomposition-based identity identification framework, which decomposed raw noisy CSI measurements into intrinsic mode functions (IMF) and extracted intrinsic features from the IMF components accordingly. Zhang *et al.* [28] proposed a novel approach that analyzed the CSI data to extract unique features that were

representative of the walking patterns of that individual, and thus allowed the system to uniquely identify that person uniquely. Zeng *et al.* [29] presented a framework that can identify a person from a group of people in a device-free manner using WiFi, and showed that CSI used in recent WiFi identified a person's steps and gaits. Although these human identity identification approaches can guarantee certain accuracies as presented, they were interfered severely by the influence of the random noise derived from indoor environments, which could lead to a bad identification performance.

Different from these approaches above, we leverage DWT strategy to suppress the random noise presented in the raw CSI data. Based on this, several representative gait features are extracted from time and frequency domain to characterize human's gaits. Furthermore, the proposed RNN model with LSTM blocks is used to learn the representative gait features extracted for identifying different persons from a group of people effectively. Thus, compared with most of the existing identification approaches, our proposed approach is able to access superior performance with regard to the robustness to the random noise and the accuracy of identity identification.

III. PRELIMINARY

In this section, a short overview of CSI is presented. The most of commercial WiFi devices operate on both the 2.4 GHz and 5 GHz frequency bands and also support MIMO techniques. In addition, the modern off-the-shelf devices also leverage orthogonal frequency division modulation (OFDM) to obtain fine-grained channel measurements at the physical layer. Specially, the OFDM channel is divided into multiple subcarriers where each subcarrier has a different signal amplitude and phase with regard to each transmitted signal. Generally, the mainstream WiFi systems are based on OFDM such as 802.11 a/g/n where a relatively wideband 20 MHz channel is partitioned into 52 subcarriers. Owing to the frequency diversity of these subcarriers, both the shadow fading and multipath effect caused by minute movements at different narrowband subcarriers could lead to different amplitude and phase totally. Any time we move, we create waves in this sea of WiFi signal, so it can be known that a small body movement in indoor environments could result in the drastic change of CSI at all the subcarriers. Our proposed approach thus takes advantage of the fine-grained CSI to capture the minute movement for identity identification.

Considering that surrounding objects (e.g., furniture and wall) in indoor environments can reflect WiFi signal with different intensities, the transmitted signal arrives at the receiver through multiple different paths where each of them can introduce a different time delay, amplitude attenuation, and phase shift. Thus, the channel impulse response (CIR) can be described as follows:

$$h(\tau) = \sum_{i=1}^{N} a_i e^{-j\theta_i} \delta(\tau - \tau_i), \tag{1}$$

where N denotes the total number of paths, a_i, θ_i, and τ_i are the amplitude attenuation, phase shift, the propagation

time delay of the i-th multipath component, and $\delta(\tau)$ is the Dirac delta function, respectively. Alternatively, in frequency domain, the transmitting channel can be modeled by channel frequency response (CFR), which consists of two parts with regard to amplitude-frequency response and phase-frequency response. For that, CFR can be derived by using the Fast Fourier Transform (FFT) of CIR:

$$\mathbf{H} = FFT(h(\tau)). \tag{2}$$

With commercial WiFi Network Interface Cards such as Intel 5300 and slight firmware modification, a group of subcarriers channel measurements can be obtained in the format of CSI:

$$\mathbf{H} = \left[H_1 e^{j\angle H_1}, H_2 e^{j\angle H_2}, \ldots H_l e^{j\angle H_l}, \ldots H_{30} e^{j\angle H_{30}} \right]^T \\ l \in [1, 30], \tag{3}$$

where $[\cdot]^T$ represents the transpose operation, H_l and $\angle H_l$ are the amplitude and phase of the l-th subcarrier, respectively. Generally, the continuous raw CSI data of the l-th subcarrier is collected, and the length of sliding time window is set as m. It can be given by

$$\mathbf{H}_l = \left[H_{1,l} e^{j\angle H_{1,l}}, \ldots H_{\kappa,l} e^{j\angle H_{\kappa,l}}, \ldots H_{m,l} e^{j\angle H_{m,l}} \right] \\ \kappa \in [1, m], \tag{4}$$

where $\mathbf{H}_l$ has a dimension of $1 \times m$, $H_{\kappa,l}$ and $\angle H_{\kappa,l}$ are the amplitude and phase of the κ-th data point of the l-th subcarrier.

IV. SCHEME DESIGN
A. OVERVIEW
Our proposed Wihi approach only uses a pair of transmitter and receiver devices to collect the raw CSI data for human identification. Fig. 1 illustrates the overall architecture and block diagram of the proposed Wihi approach. It is assumed that a person without taking specified sensor or device walks in the target area. At the same time, the collected raw CSI data at the receiver would be constantly analyzed to conduct identity identification. Therefore, the proposed Wihi can be divided into three main blocks, including data preprocessing, feature extraction, and human identification.

B. DATA PREPROCESSING
As discussed above, the raw CSI data is collected from indoor environments, so it inevitably contains the random noise such as nearby electronic devices interfering. This random noise can influence the accuracy of human identity identification and robustness, and thus it is necessary to carry out data preprocessing for improving system quality.

Due to many advantages such as being computationally efficient to fit on electronic devices and preserving the details of signal, the DWT based denoising strategy is used to eliminate the random noise presented in the raw CSI data. Specially, the denoising process is divided into three
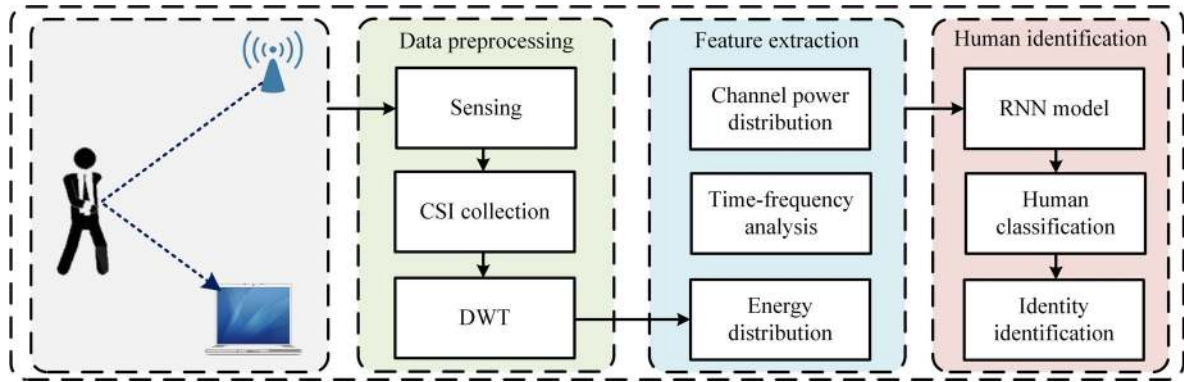
FIGURE 1. The architecture of the proposed Wihi approach.

parts with regard to signal decomposition, detail coefficients processing, and signal reconstruction. Generally, DWT can decompose the raw CSI data into two terms, including approximation coefficients and detail coefficients. Among them, the former can describe the CSI data shape, and the latter captures both the random noise and fine data details. Furthermore, in order to obtain finer details from the raw CSI data, this splitting is applied recursively a number of steps (i.e. levels), J, to the approximation coefficients. Then, a coarse approximation coefficient and a sequence of finer detail coefficients are produced by using DWT. Hence, the aim of data preprocessing can be achieved by processing detail coefficients of each level. Finally, the CSI data after denoising can be reconstructed by inverse transformation of the processed wavelet coefficients. The concrete denoising process is using following steps.

Suppose that the raw CSI data with a predefined sliding time window interfered by the random noise is given by

$$\mathbf{H}_l = \left[H_{1,l}e^{j\angle H_{1,l}}, H_{2,l}e^{j\angle H_{2,l}}, \ldots H_{\kappa,l}e^{j\angle H_{\kappa,l}}, \ldots \right.$$
$$\left. H_{m,l}e^{j\angle H_{m,l}}\right]. \quad (5)$$

Next, the wavelet coefficients in each level are able to be calculated as follows:

$$\alpha_k^{(J)} = \left\langle H_{\kappa,l}, g_{n-2^J k}^{(J)}\right\rangle_n = \sum_{n\in Z} H_{\kappa,l}g_{n-2^J k}^{(J)}, \quad k, J \in Z,$$
$$\beta_k^{(J)} = \left\langle H_{\kappa,l}, h_{n-2^j k}^{(j)}\right\rangle_n = \sum_{n\in Z} H_{\kappa,l}h_{n-2^j k}^{(j)}, \quad j\in\{1,2,\ldots J\},$$
$$(6)$$

where $\langle\cdot\rangle$ denotes the dot product operation, $\alpha_k^{(J)}$ and $\beta_k^{(j)}$ represent the approximation coefficients and the detail coefficients, respectively. $g_{n-2^J k}^{(J)}$ and $h_{n-2^j k}^{(j)}$ denote two sets of discrete orthogonal functions and also known as the wavelet basis. Then, the inverse DWT is able to be given mathematically by

$$H_{\kappa,l} = \sum_{k\in Z} \alpha_k^{(J)}g_{n-2^J k}^{(J)} + \sum_{j=1}^{J}\sum_{k\in Z}\beta_k^{(j)}h_{n-2^j k}^{(j)}. \quad (7)$$

Then, a soft threshold method is applied to the detail coefficients because it is able to remove the random noise component while retaining sufficient details for human identity identification. We have

$$S = \begin{cases} \beta_k^{(j)} - T, & \beta_k^{(j)} > T \\ \beta_k^{(j)} + T, & \beta_k^{(j)} < -T \\ 0, & \left|\beta_k^{(j)}\right| \leq T, \end{cases} \quad (8)$$

where T is the threshold. In order to obtain the optimal detail coefficients, the coefficients with a small absolute value are set to zero, and on the other side, the coefficients with a large absolute value are decreased. As a result, the CSI data can be reconstruct by leveraging the processed detail coefficients, which is denoted mathematically by

$$\tilde{H}_{\kappa,l} = \sum_{k\in Z} \alpha_k^{(J)}g_{n-2^J k}^{(J)} + \sum_{j=1}^{J}\sum_{k\in Z}\tilde{\beta}_k^{(j)}h_{n-2^j k}^{(j)}, \quad (9)$$

where $\tilde{\beta}_k^{(j)}$ is the processed detail coefficients. Then, the CSI data after denoising is able to be obtained, which is given mathematically by

$$\tilde{\mathbf{H}}_l = \left[\tilde{H}_{1,l}e^{j\angle H_{1,l}}, \tilde{H}_{2,l}e^{j\angle H_{2,i}}, \ldots \tilde{H}_{\kappa,l}e^{j\angle H_{\kappa,l}}, \ldots \tilde{H}_{m,l}e^{j\angle H_{m,l}}\right]. \quad (10)$$

The data amplitude of two people's walking activity is shown in Fig. 2(a) and 2(c), and Fig. 2(b) and 2(d) show the results after denoising. It is obvious that the CSI data is smoother after denoising which will be utilized for feature extraction next.

C. FEATURE EXTRACTION

Based on the previous work [34], it is well known that there are significant differences in human's height, body mass, and moving speed for different people, which would lead to different impacts on the characteristic of WiFi signal. Thus, the received WiFi signal can generate different signatures for different people. Concretely, when a person walks in the
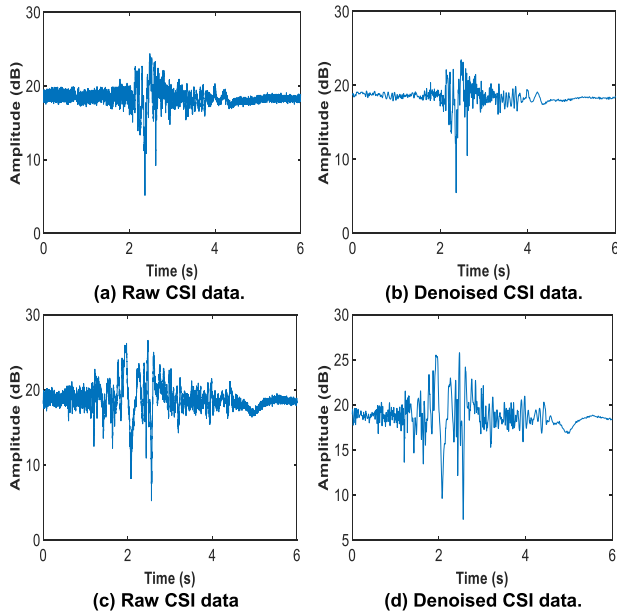
FIGURE 2. The CSI amplitude variation results for different people. (a) and (b) The first person's gaits. (c) and (d) The second person's gaits.

target area, the received signal is affected by mixed effects, including line-of-sight (LoS), static and dynamic reflection paths. Since the static reflection paths and LoS path do not contain the person's gaits information, only the dynamic reflection paths can be used to analyze how the person's gaits impact on the received WiFi signal. More importantly, this could offer us particular insights on how to find the representative features that allow us to uniquely identify the testing target. Based on this, three representative features related to the dynamic reflection paths are analyzed and extracted, including CPD, TFA, and ED.

1) CHANNEL POWER DISTRIBUTION

The human body can be modeled as a conducting cylinder when study its impact on WiFi signal [34]. Empirically, it is known that most people walk in a certain direction and have relatively steady pace in indoor environments (e.g., office and home). In view of this, it is assumed that people walk as they normally do between a pair of transmitter and receiver in the present study. Furthermore, considering that indoor objects can reflect WiFi signal to human body, and then reflect from that to the receiver. That is to say, the dynamic reflection paths are composed of $1, 2, \ldots L$ times reflection of WiFi signal. Based on this, the impact of an entity's gaits on CPD can be accessed.

Firstly, we have

$$p(t_0) = \frac{1}{\tau_{\max}} \int_0^{\tau_{\max}} \tilde{h}(t) \tilde{h}^*(t) \, dt, \qquad (11)$$

where $p(t_0)$ is the channel power at time t_0, $\tilde{h}$ is the result of inverse fast Fourier transform of denoised CSI data, and $\tau_{\max}$

is the maximum delay. Besides, it is more accurate as

$$
\begin{aligned}
p(t_0) &= \frac{1}{\tau_{\max}} \int_0^{\tau_{\max}} \mathrm{Re} \left\{ \begin{array}{l} \sum_{\iota=1}^{N} \sum_{i=1}^{N} \tilde{a}_\iota \tilde{a}_i \delta(t - \tau_\iota) \delta(t - \tau_i) \\ \times \exp\left(j\left(\theta_\iota - \theta_i\right)\right) \end{array} \right\} dt \\
&= \frac{1}{\tau_{\max}} \int_0^{\tau_{\max}} \left(\sum_{\rho=1}^{N} \tilde{a}_\rho^2 \delta^2 \left(t - \tau_\rho\right) \right) dt = \sum_{\rho=1}^{N} \tilde{a}_\rho^2.
\end{aligned}
$$
(12)

Then, the value of CPD can be accessed and denoted as

$$CPD = \frac{1}{t - t_1} \sum_{\vartheta=t_1}^{t} p(\vartheta) \quad t \in [t_1, t_2], \qquad (13)$$

where t_1 and t_2 represent the beginning time and end time of a sliding time window, respectively. Obviously, it can be found that the CPD has a significant relationship with the amplitude attenuation. The reason is that the human's gaits generate influence on the characteristic of WiFi signal, and different people's gaits lead to different multipath for the propagation of WiFi signal. Fig. 3 shows the comparison result between two people's CPD. It can be confirmed that human's CPD is different from each other when walking in the same target area.
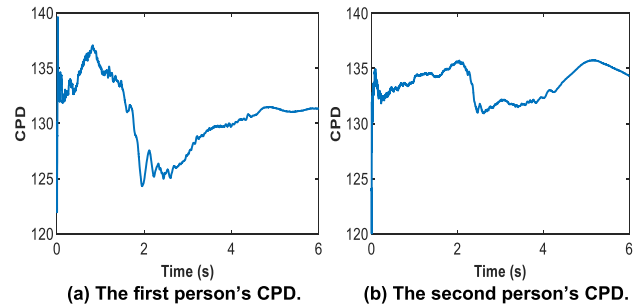

FIGURE 3. The CPD results for different people's gaits.

2) TIME-FREQUENCY ANALYSIS

The work in [35] shows that there is a function relationship between the moving speed of walking activity and Doppler frequency. In addition, the moving speed differs from each other when human walks in a target area. Thus, in order to extract representative gait features, it is very necessary to conduct a transformation to the time-frequency domain. Specifically, for time-frequency analysis, there are varieties of time-frequency feature extraction methods where Short-Time Fourier Transform (STFT) has a better spectrogram characteristics compared with the others. Furthermore, TFA is conducted to separate the Doppler component caused by walking activity. Depending on this, human's gaits can be modeled by profiling the each frequency component by leveraging STFT method.

Fig. 4 illustrates the comparison between two people's TFA. It is evident from the spectrograms that the first one has a higher Doppler frequency compared with the second
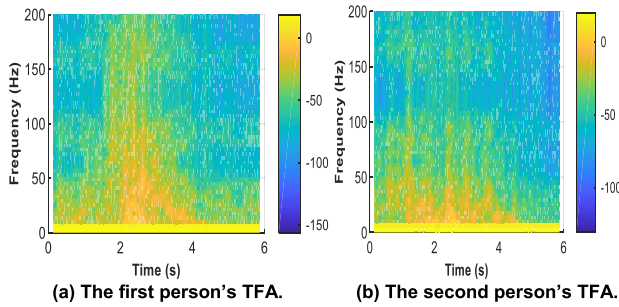
FIGURE 4. The TFA results for different people's gaits.

FIGURE 5. The ED results for different people's gait.

one. In other words, the first person has faster speed when walking in the target area.

3) ENERGY DISTRIBUTION

Prior work [36] has shown that different parts of human's body present different speed when walking. This means that the impact of one person's gaits is likely to be the most pronounced in several specific frequency bands of denoised CSI data. In order to observe the energy distribution in different frequency bands, and get more subtle gait features, the wavelet packet decomposition (WPD) is employed to plot the spectrograms for characterizing human's gaits. Specially, WPD can decompose the approximate and detail information of CSI data many times equally, and then the CSI data can be separated into Ω subspaces. In addition, in order to determine which parts of the CSI data are most likely to exhibit distinguishing features for a person's gaits, three levels ($J = 3$) decomposition is conducted for further investigating in the present study. The process of WPD can be represented as orthogonal sum of different subspaces, which is able to be given by

$$\left\| \tilde{\mathbf{H}}_l \right\| = U_{J-1}^1 \oplus U_{J-2}^1 \oplus U_{J-3}^1 \oplus U_{J-3}^0,$$
$$U_{j+1}^\varpi = U_j^{2\varpi} \oplus U_j^{2\varpi+1} \quad j \in Z, \; \varpi \in Z_+, \quad (14)$$

where U_j^ϖ is the subspaces of the CSI data, and $\oplus$ is the orthogonal sum operation.

By WPD, eight specific frequency bands can be obtained, and the bandwidth of each frequency band is 20 Hz. Fig. 5 illustrates the comparison between two people's ED. It can be found that the frequency band of "0-20 Hz" contains the least energy for different people from the spectrograms. For the first one, the frequency bands of "80-100 Hz", "100-120 Hz", and "140-160 Hz" contain the most energy. In contrast, it is manifested by the less energy observed in the other frequency bands relatively. Besides, for the second one, the frequency bands of "80-100 Hz" and "140-160 Hz" have the most energy while other frequency bands have less energy relatively. Specially, the second one's ED appears to have more energy compared with the first one across all frequency bands. It is observed that the same frequency bands contain different energy for two people's walking activity at the same time. By investigating ED in different frequency
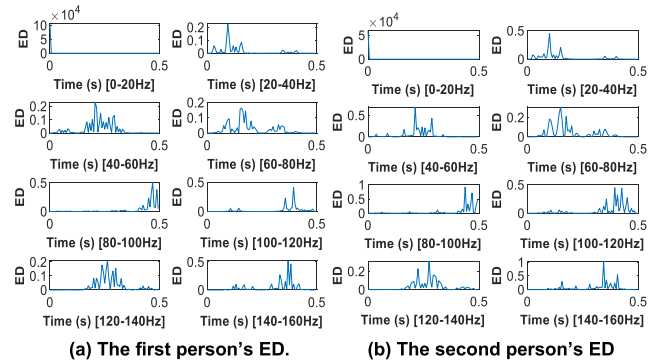
bands, the separated data and their corresponding energy can give us more information about gait features.

D. RNN MODEL

In addition to the gait features above, we also make full use of eight statistics features which have been used in human identity identification, including maximum, minimum, mean, variance, standard deviation, median, energy, and entropy. The common statistics features in time and frequency domain are listed in Table 1. To consider a series of gait features, the RNN model with LSTM blocks is introduced in this section. Then an appropriate input vector from these human's gait features is created, which can be denoted by

$$\mathbf{x} = [CPD, TFA, ED, Common\ statistics\ features]^T, \quad (15)$$

where the input vector has a dimension of D_x. In order to achieve the aim of better identification performance, the selected RNN model leveraging P sets of representative features is considered to make a decision by extending gait features. Fig. 6 illustrates the detailed structure of the RNN model, where $\mathbf{x}^{(p)}$ is the input vector and $\mathbf{h}^{(p)}$ denotes the hidden layer at the p-th time step, respectively. Besides, the dimension of the hidden layer is set as D_h. In addition to the hidden layer, the LSTM blocks has also one unit, called cell state $\mathbf{c}^{(p)}$, which has the capacity for controlling the information flow by taking use of three significant gates with regard to forget, input, and output gates. Each gate vector can be obtained with regard to the input vector, which is expressed

TABLE 1. Common statistics features.

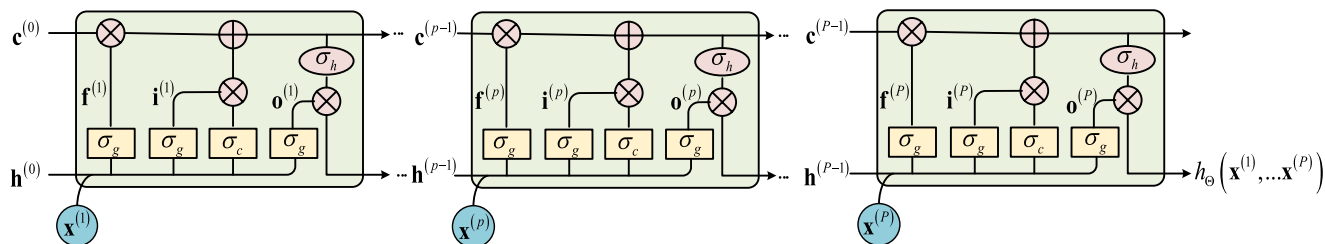| ID	Feature name	Domain
1	Maximum	Time
2	Minimum	Time
3	Mean	Time
4	Variance	Time
5	Standard deviation	Time
6	Median	Time
7	Energy	Frequency
8	Entropy	Frequency

FIGURE 6. The RNN model with LSTM blocks.

mathematically as

$$
\mathbf{f}^{(p)} = \sigma_g \left(\mathbf{W}_f \mathbf{x}^{(p)} + \mathbf{U}_f \mathbf{h}^{(p-1)} + \mathbf{b}_f \right),
$$
$$
\mathbf{i}^{(p)} = \sigma_g \left(\mathbf{W}_i \mathbf{x}^{(p)} + \mathbf{U}_i \mathbf{h}^{(p-1)} + \mathbf{b}_i \right),
$$
$$
\mathbf{o}^{(p)} = \sigma_g \left(\mathbf{W}_o \mathbf{x}^{(p)} + \mathbf{U}_o \mathbf{h}^{(p-1)} + \mathbf{b}_o \right), \quad (16)
$$

where $\mathbf{W}$, $\mathbf{U}$, and $\mathbf{b}$ denote input weight, cyclic weight, and bias, respectively. Besides, $\mathbf{W}$ has a dimension of $D_h \times D_x$, $\mathbf{U}$ have a dimension of $D_h \times D_h$, and $\mathbf{b}$ have a dimension of $D_h \times 1$. In addition, $\sigma_g(\cdot)$ represents an element-wise activation function with regard to all gates, and it is given by a sigmoid function $\sigma_g(z) = 1/(1 + e^{-z})$. Additionally, it should be noted that the previous cell state and hidden at the first time step are initialized as zero vectors. Due to using these gate vectors, both the cell state and hidden layer are able to be updated timely. We have

$$
\mathbf{c}^{(p)} = \mathbf{f}^{(p)} \circ \mathbf{c}^{(p-1)} + \mathbf{i}^{(p)} \circ \sigma_c \left(\mathbf{W}_c \mathbf{x}^{(p)} + \mathbf{U}_c \mathbf{h}^{(p-1)} + \mathbf{b}_c \right),
$$
$$
\mathbf{h}^{(p)} = \mathbf{o}^{(p)} \circ \sigma_h \left(\mathbf{c}^{(p)} \right), \quad (17)
$$

where $\circ$ denotes the Hadamard product operation, $\sigma_c(\cdot)$ and $\sigma_h(\cdot)$ represent element-wise activation functions with regard to the cell state and hidden layer. For activation functions above, we take use of the hyperbolic tangent in the present study. After that, the RNN output is able to be obtained from the hidden layer at the last time step, and it is denoted by

$$
h_\Theta \left(\mathbf{x}^{(1)}, \dots \mathbf{x}^{(P)} \right) = \sigma \left(\mathbf{V} \mathbf{h}^{(P)} + b \right), \quad (18)
$$

where $\mathbf{V}$ denotes the row vector and its dimension is D_h, and b represents a bias constant, respectively. Besides, every parameter in the proposed RNN model is contained in the set Θ.

Every parameter in the RNN model is able to be adjusted appropriately by training gait features dataset. We denote the sequence of P input data vectors as $\mathbf{X} = (\mathbf{x}^{(1)}, \dots \mathbf{x}^{(P)})$, and the corresponding identity label as y, where y is viewed as person A, person B, etc. Then, a set of input and output pairs $(\mathbf{X}, y)$ is produced to train and verify the RNN model from the gait features and corresponding identity labels. Specially, ς is introduced as such a gait features dataset. Besides, all parameters of the RNN model are adjusted in the direction of

minimizing the cost function, and we can have

$$
J(\Theta) = -\frac{1}{|\varsigma|} \sum_{g \in \varsigma} C(g), \quad (19)
$$

where $|\cdot|$ represents the number of elements in a set, $C(g)$ denotes the cost of the g-th input and output pair and measures the accuracy of the output of the RNN model compared with the ground truth data.

In order to evaluate the cost performance, cross-entropy is selected from many choices, we have

$$
C(g) = y^{(g)} \log h_\Theta \left(\mathbf{X}^{(g)} \right) + \left(1 - y^{(g)} \right)
$$
$$
\times \log \left(1 - h_\Theta \left(\mathbf{X}^{(g)} \right) \right), \quad (20)
$$

where the superscript represents the index of the input and output pairs. Then, all the parameters of the proposed RNN are able to be updated in an iterative method by leveraging the gradient descent method. Besides, every parameter is updated in the direction of the steepest descent at each iteration, which can be denoted as

$$
\Theta_{n+1} = \Theta_n - \eta \nabla_\Theta J(\Theta), \quad (21)
$$

where ∇_Θ represents the gradient operation for Θ, and η denotes the learning rate and it determines the step size. In addition, there are varieties of variations of the gradient descent methods, such as Adam optimizer [37]. Concretely, these optimizers are able to adaptively change the learning rate in the RNN model to achieve the aim of minimum cost precisely and efficiently. Regardless of which optimizer is utilized, the gradient of the cost function needs to be taken at each iteration.

Due to many significant parameters, the non-linearity relationship between the input and output layers is able to be captured by the RNN model efficiently, and meanwhile, there is also a high risk of over-fitting. In order to avoid this key issue, early stopping strategy is taken into consideration, and here the Adam optimizer is leveraged. In addition, we separate the dataset of input and output pairs, ς, into three non-overlapping datasets, including training dataset ς_{tr}, validation dataset ς_v, and testing dataset ς_{te}. Then, the proposed RNN is trained to adjust every parameter by using the training dataset with true label. Furthermore, the RNN model evaluates and

tracks these values of the cost function (19) for the validation dataset at each iteration, and the optimal parameters are selected in the case when the cost function is minimized.

The proposed RNN model is trained utilizing the Matlab application running on a workstation with 40 CPUs in this study. Specially, the dimension of the input data is 7606×1, the number of hidden layers is set as $D_h = 10$, the batch size is set as 128, and the Adam optimizer with the learning rate is set as $\eta = 0.1$, respectively. As mentioned above, the gait features dataset is separated into three parts with regard to training, validation, and testing sets. Besides, considering identification performance of the proposed RNN model is mainly depending on the number of gait features, certain amounts of gait features are selected and the corresponding portion of the selected datasets becomes 67.7%, 6.6%, and 26.7%, respectively. In addition, the optimization algorithm runs up to 1000 iterations where one iteration is called epoch, which means that every parameter in the proposed RNN model is updated using the entire training dataset. As a result, when all parameters of the RNN model are optimal, the proposed approach is able to identify the corresponding person from a group of people. If adding another person, the proposed model needs to be retrain. By doing so, an entity can be identified from a group of people effectively.

V. PERFORMANCE AND EVALUATION

To validate the architecture of the proposed Wihi approach, a prototype is designed and tested. Specifically, some real-world experiments are conducted to evaluate the performance of identity identification in two typical indoor environments, i.e. an office and a laboratory. Besides, in order to verify the high performance, the proposed Wihi is compared with five state-of-the-art identification approaches, including WiFi-ID, WiWho, Wii, AutoID, and CSIID. Firstly, WiFi-ID extracts and utilizes the common statistics features for characterizing the walking patterns, and identifies people from a group of 2 to 6 people using sparse approximation based classification (SAC). Secondly, WiWho uses decision tree-based machine learning classifier to identify people from a group of 2 to 6 based on the common statistics features. Thirdly, Wii extracts the common statistics features to characterize walking gaits, and then selects the most effective gait features according to information gain. Depending on these features, Wii realizes stranger recognition through Gaussian mixture model (GMM) and identities identification through a support vector machine (SVM) with radial basis function (RBF) kernel. Fourthly, AutoID uses the amplitude information of WiFi CSI as gait feature, and identifies people from a group of 2 to 20 through the optimization-based shapelet learning framework. Finally, CSIID extracts the amplitude of WiFi CSI as gait feature, and leverages a deep learning method to identify people from a group of 2 to 6. Specially, in CSIID, the convolution layers are combined with LSTM layers. Besides, these approaches are implemented on commercial off-the-shelf WiFi devices, and the design scenarios of these approaches are almost the same as Wihi that all of them

work in the office or laboratory environment. In the paper, these approaches are implemented utilizing the same data collected from the testing area as Wihi. From the comparison, the identification performance in terms of accuracy and estimation error is discussed. Moreover, the impact of different experimental settings on the performance of the proposed Wihi approach is also investigated.

A. DATA COLLECTION

To collect the raw CSI data, a commercial WiFi device with one antenna is as the transmitter and a Lenovo laptop with three antennas is as the receiver, which is able to form 1×3 data transmission links. Thus, in this experiment, we have three links where each of them has 30 subcarriers. The WiFi device runs on 5 GHZ with bandwidth channel of 20 MHz, and the data sample rate is set as 1 KHz. With WiFi NICs such as Intel 5300 and slight firmware medication, the raw CSI data can be collected at the receiver by leveraging CSI tool [38]. Besides, a sliding window with a window size of 6 seconds is utilized for data segmentation. In the present study, human's gaits features dataset is collected in typical indoor environments in terms of an office and a laboratory. Besides, the layouts of the two experimental environments are shown in Fig. 7. Of the two rooms, the office containing three tables has a size of $5m \times 6m$ where the transmitter and receiver are placed on the top of the tables, and the laboratory containing eight tables and a variety of electronic devices has a size of $5m \times 8m$ where the transmitter and receiver devices are also placed on the top of the tables.

Eight healthy students are recruited in our experiments and the basic information of these participants is shown in Table 2. Concretely, each of all the testing objects is asked to perform walking activity for a period of 8 seconds during CSI data collection, so that we are able to get enough gait information. In addition, the one needs to remain stationary at the beginning and the end of walking activity to reduce the noise. It should be noted that these testing objects are asked to naturally walk on the path that crosses the LoS of the transceivers without any constraint of walking speed of style. Totally, eight people are involved for the raw WiFi CSI data collection, and each one needs to perform walking activity 375 times in the same target area. After data preprocessing and feature extraction, we are able to get about 3000 sets of

TABLE 2. Basic information of participants.

ID	Gender	Age	Height(cm)	Weight(kg)	Speed(m/s)
1	Male	25	172	62	1.23
2	Female	24	161	50	1.11
3	Male	26	176	65	1.18
4	Female	25	155	49	1.14
5	Male	25	178	61	1.21
6	Female	24	165	52	1.05
7	Male	25	170	60	1.22
8	Male	25	180	68	1.24
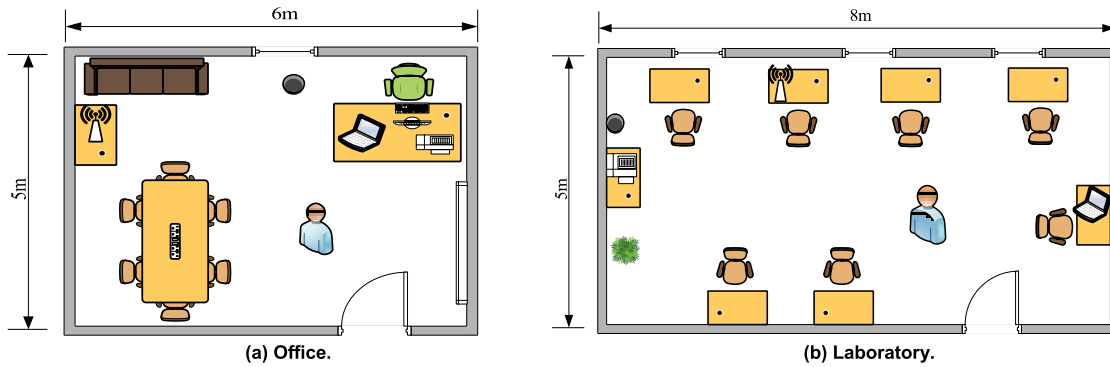
FIGURE 7. The experimental environments for the proposed Wihi approach.

(a) Office. (b) Laboratory.

gait features labeled with corresponding identity information. As mentioned above, the training, validation, and testing sets account for 66.7%, 6.6% and 26.7%, respectively. Therefore, the size of the corresponding sets is about 2000, 200, and 800, respectively. In addition to evaluating the performance in different experimental environments, the gait features dataset collected from the laboratory is also utilized for investigating the impact of different numbers of features and hidden nodes, the impact of window size, the identification performance of different approaches, the identification accuracy of different people, the impact of the random noise, and so forth.

Moreover, gait features dataset is collected with multiple different experimental settings for comprehensive evaluation in the present study. In addition to the above walking path, other walking paths in the laboratory scenario is considered as well. Fig. 8 shows three typical human walking paths in the target area, including walking across LoS (p1), circling (p2), and walking on the direct LoS (p3). In the above experiments, the testing objects walking across LoS is only considered. In order to investigate the impact of different walking paths, we need to collect data of the two walking paths with regard to p2 and p3, and the total experimental setup is the same as the above. In addition, it is highly crucial to analyze the impact of some unrelated people moving into or out of this laboratory. In view of this, conducting data collection is very necessary to verify the high identification performance, and the walking path is set as p1 during data collection. Besides, in order to explore the significant impact of different walking speed, multiple walking speed is taken into consideration, including slow, normal, and fast, when one person walks in the target area. Note that, all the data collected need to go through data preprocessing and feature extraction steps for obtaining gait features dataset.

B. IDENTIFICATION RESULTS

In this section, the performance of identity identification with different group sizes is evaluated. The experiments are conducted in typical indoor environments, i.e. an office and a laboratory. Besides, we only consider group of people in the range from 2 to 8. For each of the group sizes, Wihi
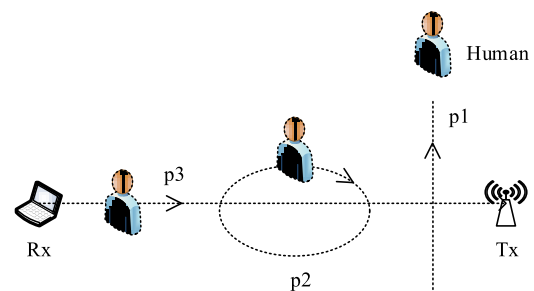

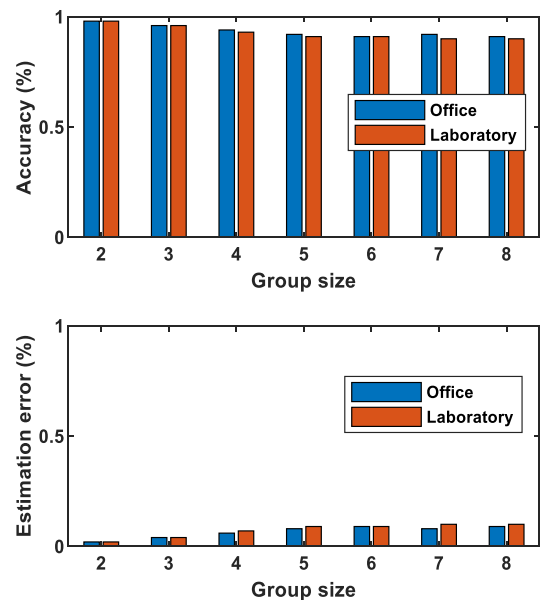**FIGURE 8.** Human's walking paths p1, p2, p3.

FIGURE 9. The accuracies for identity identification with different group sizes in typical indoor environments.

uses gait features to perform human identity identification. Fig. 9 shows the average accuracies and estimation errors of identity identification with different group sizes in two typical indoor environments. As it is indicated, with the group size increasing, the accuracy of identity identification decreases. The reason is that the chances of people sharing similar

walking patterns could increase with the increasing of the number of people. Specially, Wihi is able to achieve average accuracies of identity identification of 98% and 96% with the group size of 2, and the accuracies decrease to 92% and 91% with the group size of 8 respectively in two typical indoor environments. Besides, the identification accuracies in the laboratory are the worst compared with that in the office. The possible reason is that the laboratory area is much larger than that of the office, which means the laboratory has a more complicated multipath environment. Additionally, it is observed that the estimation errors are the highest when the group size of 8 is considered. In contrast, the errors are the lowest with the group size of 2.

C. RESULTS WITH DIFFERENT APPROACHES

To validate its high effectiveness, we compare the proposed Wihi approach with WiFi-ID, WiWho, Wii, AutoID, and CSIID fairly using the data collected from the laboratory. Specially, Wihi takes use of the gait features extracted and common statistics features after feature extraction, WiFi-ID and WiWho use the common statistics features, and AutoID and CSIID utilize the amplitude information as the input of RNN. The identification accuracies for the existing state-of-the-art approaches and the proposed Wihi are shown in Fig. 10. It can be observed that WiFi-ID performs the worst and AutoID outperforms WiFi-ID slightly. Besides, Wii yields a better performance compared with WiFi-ID and AutoID. The deep learning based approach of CSIID considers the temporal dependencies in sequential data, so it achieves a better performance than the Wii approach. WiWho yields a superior performance compared with the CSIID approach. The reason is that WiWho utilizes the deep learning to learn the common statistics features, instead of the amplitude of WiFi CSI. The proposed Wihi approach is able to achieve the best performance in human identity identification for all groups of people. The possible reason is that the extracted gait features theoretically contain more information which can characterize human's walking patterns profoundly, and they are more suitable for identity identification.

D. ACCURACY OF DIFFERENT PEOPLE

Depending on the data collected from the laboratory, the identification accuracies of the eight students in the above experiments are shown in the confusion matrix in Fig. 11. It shows the details of the identification of different people in the laboratory scenario. As it is indicated, it is observed that different people can yield different identification accuracies which vary from 88% to 95%. According to Table 2, it can be found that there are only three female here, the others are male in the testing group, which indicates that the walking patterns of female share less similarity with that of male. Besides, if two people share similar height and weight, they are more likely to have much difficulty in being identified correctly. Student 8 has the highest identification accuracy as he is higher and heavier than the other students in the test group, which means that the walking patterns of Student 8 have larger influence on
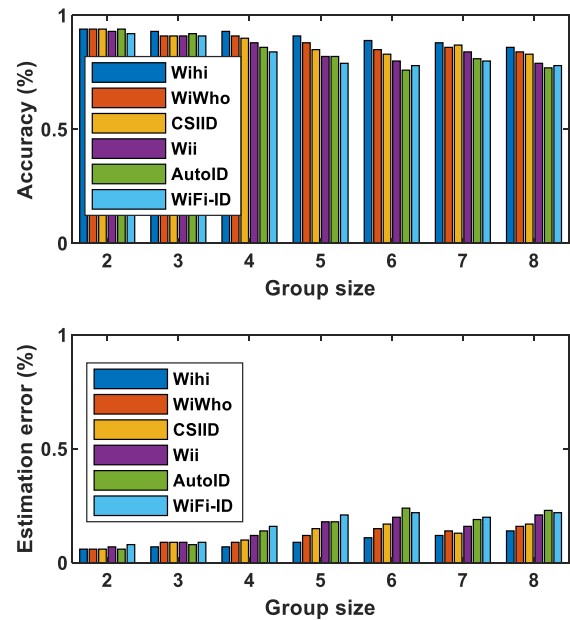

FIGURE 10. The identification accuracies for different approaches in the laboratory.

FIGURE 11. The confusion matrix of human identification with eight people in the laboratory.

the characteristics of WiFi signals. In contrast, Student 1 and 7 yield the lowest identification accuracies due to sharing the similar height and weight. In other words, height and weight have a significant impact on the propagation of WiFi signal.

E. IMPACT OF DIFFERENT NUMBERS OF FEATURES

The number of features has important impact on both the computation complexity and the identification accuracy. So, the data collected from the laboratory is used to investigate the significant impact of different numbers of features on the identification performance. We select first eight features from the extracted gait and common statistics features, and the group size of 8 students is involved for this evolution.
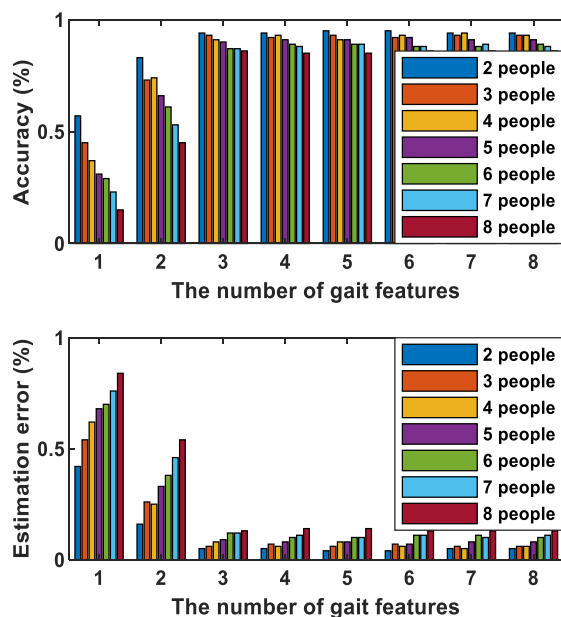
FIGURE 12. The accuracies for identity identification with different number of gait features in the laboratory.

FIGURE 13. The accuracies for identity identification with different walking paths in the laboratory.

The experimental result is depicted in Fig. 12, and it can be observed that the identification accuracy first rises with the increasing of the number of gait features, and then it starts to remain relatively stable when the number of gait features becomes larger than 3. The reason is that the gait features extracted contain the highest information gain. Depending on this, a good trade-off between computation complexity and identity identification accuracy can be given.

F. IMPACT OF DIFFERENT WALKING PATHS
Another important factor in evaluation of Wihi is walking paths. Thus, three walking paths is considered in this study, including p1, p2, and p3. All the data are collected from the laboratory and the experimental setup is the same. Besides, the identification accuracies of different walking paths are shown in Fig. 13. As seen, the identification performance of the path p3 is the worst. The reason is that a moving object blocks the propagation of signal and incurs the shadowing of different extent, which means that the walking patterns have small influence on the WiFi signal. The path p2 yields a better performance compared with the path p3. From Fig. 13, it is observed that the path p1 has the best identification performance due to inducing significant changes on WiFi signals continuously. The experimental result demonstrates that the different walking paths can yield different influence on the characteristic of WiFi signal.

G. IMPACT OF THE NUMBER OF HIDDEN NODES
Considering the number of hidden nodes is an important parameter for the RNN model, an additional experiment is conducted to investigate the impact of this parameter on the system performance. The data collected from the laboratory
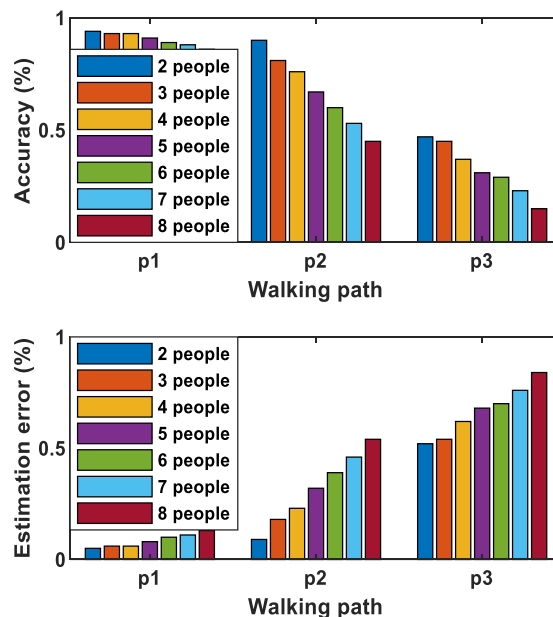
is used and the experimental result is shown in Fig. 14. As seen, in the case when the number of hidden nodes is only 1, the identification accuracies are very low for all groups of people. In addition, when the number of hidden nodes increases from 1 to 7, the identification accuracies also increase with that. Furthermore, with the increasing of the number of hidden nodes from 7 to 10, it can be found that the identification accuracies remain relatively stable. The possible reason is that when the number of hidden nodes is small, some parameters of the RNN model are not optimal, which could lead to a bad identification performance. Table 3 shows the training time of different numbers of hidden nodes in the case when the group size is 8. It should be noted that more hidden nodes would lead to longer training time. Therefore, 10 hidden nodes are chosen in this study.

TABLE 3. The training time of different hidden nodes.

Number	1	3	5	7	9	10
Training (sec)	408	652.8	1821	2325	3182	5362

H. IMPACT OF PRESENCE OF OTHER HUMANS
In real environments, some unrelated people presented in the testing area is inevitable, which means that we need to conduct an additional experiment to investigate the impact of this parameter on the identification performance. This experiment is also conducted in the laboratory scenario, and there are four unrelated people performing random actions at random locations in the same room. In particular, the random location needs to keep a certain distance from the transceiver and the
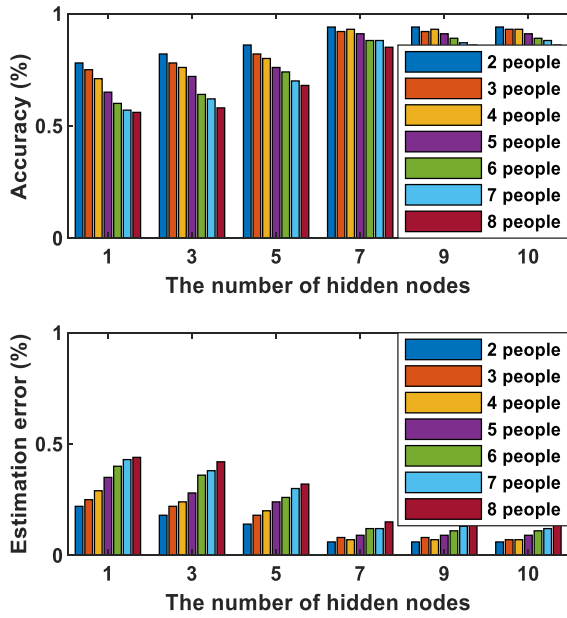
FIGURE 14. The accuracies for identity identification with different number of hidden nodes in the laboratory.
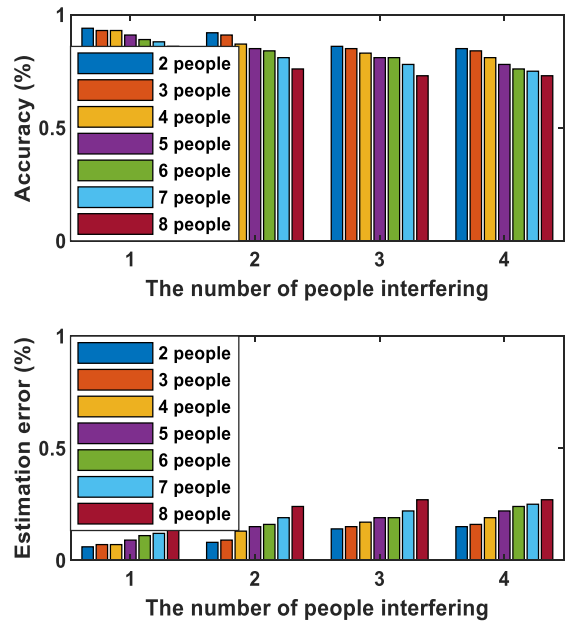

FIGURE 15. The accuracies for identity identification with different number of people interfering in the laboratory.

FIGURE 16. The accuracies for identity identification with different window sizes in the laboratory.

Specifically, in the case when only a receiver is deployed, the average accuracies are as high as 85% and 73% with the group sizes of 2 and 8 respectively in the presence of four people interfering. The reason is that there is a closer distance between the Wihi actual user and the transceiver devices compared with the unrelated people. Thus, the proposed Wihi can yield a high performance even though there are some people interfering.

I. IMPACT OF WINDOW SIZE
In the experiments above, the window size of predefined sliding window is set as 6 seconds. In order to investigate whether there is significant impact of this parameter on the system performance, the window size of sliding window is flexibly changed. In addition, the data collected from the laboratory is leveraged, and the window size is set from 4 seconds to 8 seconds to segment the data. The experimental result with different window sizes is shown in Fig. 16, and it can be found that the identification accuracies increase slightly in the case when the window size increases, and then become stable gradually. The possible reason is that when the window size is large enough, the human' walking patterns would generate less influence on the characteristics of WiFi signal. Considering that longer time would increase the computation complexity, the window size of 6 seconds is appropriate for human identity identification.

J. IMPACT OF DIFFERENT WALKING SPEED
In this section, an additional experiment is carried out to analyze the significant impact of different walking speed on the system performance. In particular, three typical walking

experimental objects in the same room, and more importantly, these unrelated people cannot walk across LoS and interfere with the participants' walking. In addition, the random action denotes some activities in daily life, such as walking, jumping, sitting, crouching, and so forth. Furthermore, compared with the actual participants' movement, the movement of these random actions of these unrelated people is small. Fig. 15 depicts the overall human identification accuracies with the increasing of the number of people interfering.
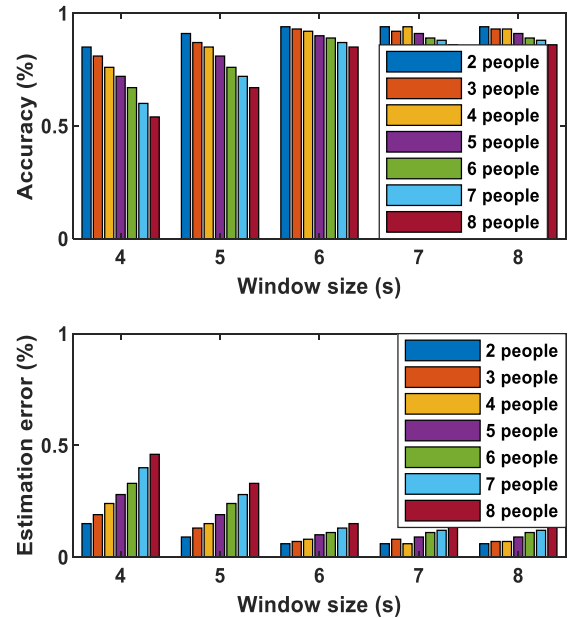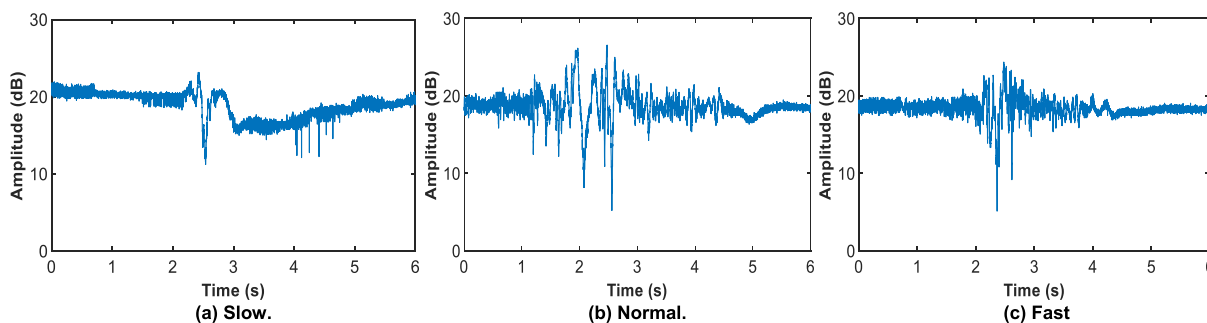
FIGURE 17. The CSI amplitude variation results when one person walks with different speed in the target area.

speed is taken into consideration, including slow, normal, and fast. Additionally, this experiment is conducted in the laboratory and the walking path is p1. Fig. 17 shows the experimental result in the case when one person walks with different speed in the same room. From these plots, it can be observed that different walking speed yields different influence on the characteristics of WiFi CSI. There is a large difference in CSI amplitude variation, especially for the slow and normal walking. The possible reason is that different walking speed introduces different Doppler in CSI for the same person. Moreover, in order to investigate the identification performance, the data of the normal walking is utilized for training the model and the data of the slow and fast walking is leveraged for testing. The identification accuracy is shown Fig. 18 and it can be found that the slow walking performs the worst. On the other hand, the fast walking yields a better identity identification performance compared with the slow one. The possible reason is that the fast walking shares similar patterns with the normal one on WiFi CSI amplitude variation.
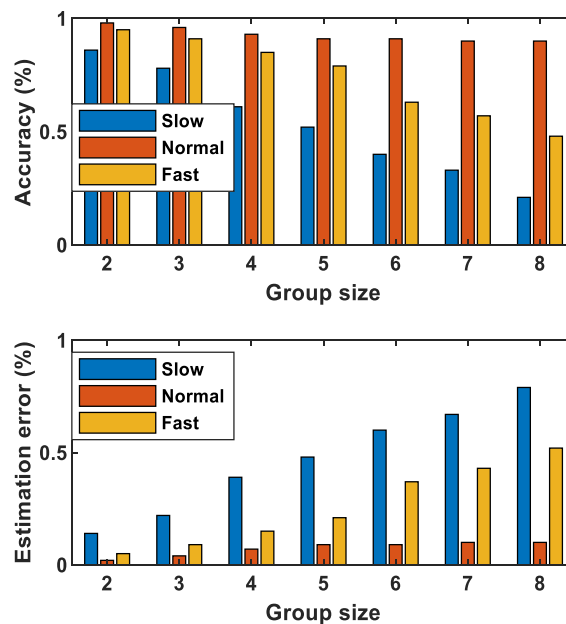
K. IDENTIFICATION ACCURACY OF THE DATA MIXED FROM TWO INDOOR ENVIRONMENTS

As mentioned above, a mount of CSI data is collected from two indoor environments. Now, a natural question to ask: can the data from two indoor environments be mixed for identity identification? Therefore, an additional experiment needs to be performed to investigate the impact of this parameter on the identification accuracy. Specifically, the data from the office is employed for training and the data from the laboratory is used for testing. Besides, the walking path is p1. The experimental result is depicted in Fig. 19. As it is indicated, compared with the identification accuracy in the office, the accuracy is unsatisfactory utilizing the data from the laboratory. The identification performance in the office outperforms that in the laboratory slightly. It can also be found that the accuracies of identity identification in the office scenario are the highest for all the group sizes. The possible reason is that the laboratory is able to generate a more complex multipath environment compared with the office, which means that they yield different influence on the characteristics of WiFi CSI.

FIGURE 18. The accuracies for identity identification with different walking speed in the laboratory.

L. IDENTIFICATION ACCURACY OF THE DATA MIXED FROM DIFFERENT WALKING PATHS

In this section, an additional experiment is conducted to analyze whether the data collected can be mutually trained and tested when the same person walks with different paths in the same room. Specially, the p1 data is used for training and the p2 and p3 data are utilized for testing. Besides, this experiment is conducted in the laboratory scenario as well. The experimental result is shown in Fig. 20. Concretely, the identification accuracy using the p3 data performs the worst, and the identification performance utilizing the p2 data is better compared with that using the p3 data. The reason is that the walking path p2 yields a significant influence on the propagation of WiFi signal. Moreover, the identification accuracies using the p1 data are the highest for all the group sizes.

M. IMPACT OF THE RANDOM NOISE

In the above experiments, all the data collected need to be processed by the DWT method, which could suppress the serious
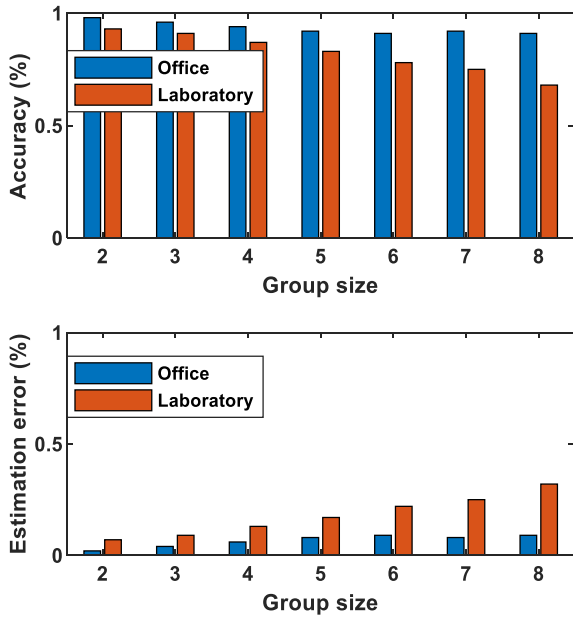
FIGURE 19. The identification accuracies of the data mixed from two indoor environments.
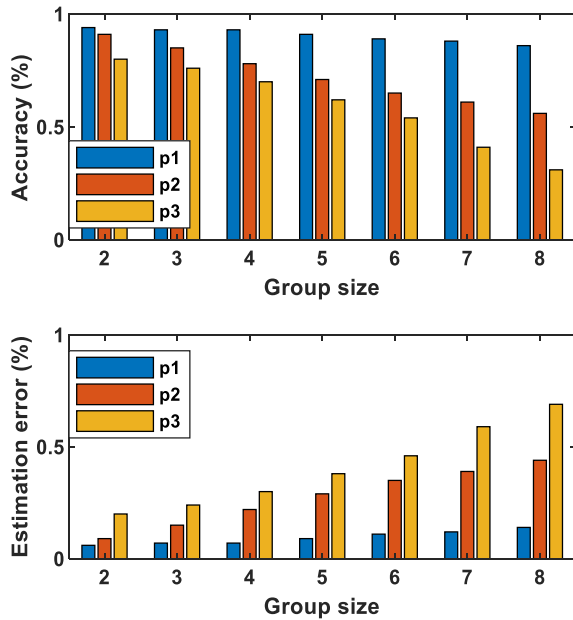

FIGURE 20. The identification accuracies of the data mixed from different walking paths in the laboratory.

influence of the random noise derived from typical indoor environments. In order to verify its effectiveness, an additional experiment is conducted to investigate the impact of this parameter on the identification performance. Besides, the data from the laboratory is used and the walking path is set as p1. Specially, one set of data is processed by the DWT while the other is not. Fig. 21 shows the identity identification accuracy. From this plot, it can be found that the identification performance using the data after the DWT denoising is better than that using the unprocessed data for all the group sizes.
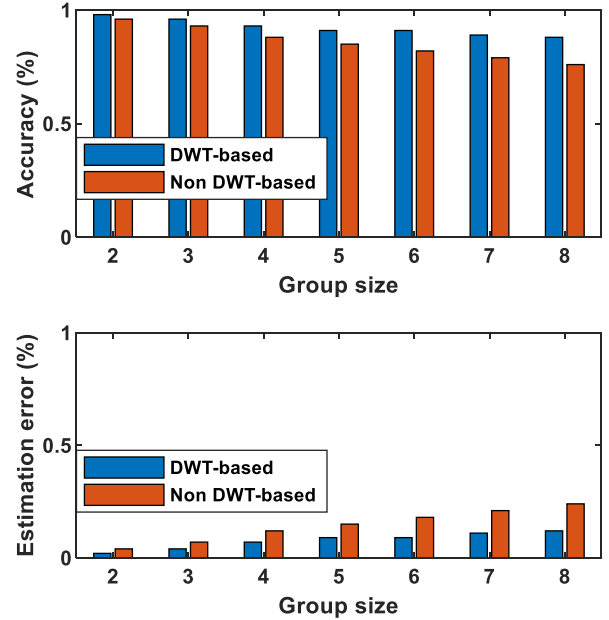

FIGURE 21. The accuracies for identity identification with different data processing methods in the laboratory.

The possible reason is that the random noise can add false edges and affect identification accuracy and robustness. As a consequence, data denoising based on DWT is essential for improving system quality.

VI. LIMITATIONS
A. MULTI-TARGET IDENTIFICATION
Multi-target identification is a well-known challenging problem. Although most of the existing approaches are able to achieve a good performance of identity identification, they yield a low accuracy in the multi-target identity identification. To handle this problem, the mobile paths need to be differentiated from different moving targets. Besides, this problem becomes even more difficult when several targets are near to each other. In the future work, we plan to explore the possibility of utilizing multiple WiFi devices to separate gait signal from multiple targets. With a higher density deployment, we believe the proposed approach can achieve the aim of multiple targets identification.

B. WALKING PATH
The targets must walk on a predefined path in a predefined walking direction, e.g., walking across LoS as in our experiments. Besides, the proposed RNN model trained for a given walking path and direction cannot be leveraged for testing set obtained on different walking paths and directions. The reason is that different walking paths and directions are able to cause somewhat influence on the propagation of WiFi signal.

C. THE TESTING RANGE
While Wihi can support up to 5m×8m interaction range, achieving whole-home coverage is not an easy task. As the

operation distance increases, the reflected power attenuates exponentially while the interference from static signals and noise remains unchanged. Meanwhile, the field of directions with Doppler shifts over the sensitivity narrows continuously. In addition, given the lowest signal to interference and noise ratio (SINR) and sensitivity of an NIC, these two factors determine the maximum coverage of the proposed approach. Nevertheless, whole-home can still be achieved by deploying more WiFi devices in the area of interest, or by sensing human activities with larger Doppler frequency shifts.

VII. CONCLUSION

In this paper, we propose a WiFi CSI based passive identity identification approach utilizing recurrent neural network. To verify the effectiveness of the proposed Wihi approach in typical indoor environments, some real experiments are conducted. Based on those real-world dataset collected, the performance in the two scenarios is first presented. Next, we compare the proposed Wihi approach with the existing identification approaches. According to the experimental results, Wihi, WiWho, CSIID, Wii, AutoID, and WiFi-ID can achieve the identification accuracies of 96%, 96%, 95%, 95%, 96%, and 94% with the group size of 2, respectively. In addition, the identification accuracies are 91%, 90%, 89%, 84%, 82%, and 83% in the case when the group size is 8. Thus, the experimental results demonstrate that the proposed Wihi can achieve much better performance than the existing approaches in the laboratory. Furthermore, we evaluate its high identification performance with different settings, such as the impact of different walking paths, the impact of different walking speed, the impact of the random noise, in two indoor environments. The experimental results show that the proposed Wihi approach is able to achieve a satisfactory performance, which makes it an ideal candidate for large-scale deployment in indoor environments. Finally, the limitations of the proposed Wihi are also analyzed and discussed, including multiple target identification, walking paths, and the testing range.

In our future work, we will try our best to collect enough dataset used for training and testing to further improve the performance of human identity identification. Furthermore, these limitations and challenging issues in this study, such as more people, through the wall, more realistic scenarios, will continue to be studied combining with new theoretical perspective and multi-disciplinary integration.

REFERENCES

[1] E. Hossain and G. Chetty, "A multi-modal gait based human identity recognition system based on surveillance videos," in *Proc. 6th Int. Conf. Signal Process. Commun. Syst.*, Gold Coast, QLD, Australia, Dec. 2012, pp. 1–4.

[2] E. Hossain and G. Chetty, "Gait based human identity recognition from multi-view surveillance videos," in *Proc. 12th Int. Conf. Algorithms Archit. Parallel Process.*, Fukuoka, Japan, 2012, pp. 319–328.

[3] E. Hossain, G. Chetty, and R. Goecke, "Multi-view multi-modal gait based human identity recognition from surveillance videos," in *Proc. IAPR Workshop Multimodal Pattern Recognit. Social Signals Human-Computer Interact. (MPRSS)*, Tsukuba, Japan, 2012, pp. 1–12.

[4] K. Karu and A. K. Jain, "Fingerprint classification," *Pattern Recognit.*, vol. 29, no. 3, pp. 389–404, Mar. 1996.

[5] D. Maio and D. Maltoni, "A structural approach to fingerprint classification," in *Proc. 13th Int. Conf. Pattern Recognit.*, Vienna, Austria, 1996, pp. 348–359.

[6] A. K. Jain, S. Prabhakar, and L. Hong, "A multichannel approach to fingerprint classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 4, pp. 348–359, Apr. 1999.

[7] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, no. 5, pp. 705–741, May 1995.

[8] Y.-X. Yang, C. Wen, K. Xie, F.-Q. Wen, G.-Q. Sheng, and X.-G. Tang, "Face recognition using the SR-CNN model," *Sensors*, vol. 18, no. 12, pp. 1–24, Nov. 2018.

[9] H.-A. Park and K. R. Park, "Iris recognition based on score level fusion by using SVM," *Pattern Recognit. Lett.*, vol. 28, no. 15, pp. 2019–2028, Nov. 2007.

[10] C. Galdi, M. Nappi, and J.-L. Dugelay, "Multimodal authentication on smartphones: Combining iris and sensor recognition for a double check of user identity," *Pattern Recognit. Lett.*, vol. 82, no. 2, pp. 144–153, Oct. 2016.

[11] M. B. Lee, H. G. Hong, and K. R. Park, "Noisy ocular recognition based on three convolutional neural networks," *Sensors*, vol. 17, no. 12, pp. 1–26, Dec. 2017.

[12] D. Halperin, W. J. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," in *Proc. ACM Special Interest Group Data Commun.*, New Delhi, India, 2010, pp. 159–170.

[13] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 1–32, Dec. 2013.

[14] Y. Ma, G. Zhou, and S. Wang, "WiFi sensing with channel state information: A survey," *ACM Comput. Surv.*, vol. 52, no. 3, pp. 1–36, Jun. 2019.

[15] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, "PADS: Passive detection of moving targets with dynamic speed using PHY layer information," in *Proc. 20th IEEE Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Hsinchu, Taiwan, Dec. 2014, pp. 1–8.

[16] W. Xi, J. Zhao, X.-Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang, "Electronic frog eye: Counting crowd using WiFi," in *Proc. IEEE Conf. Comput. Commun.*, Toronto, ON, Canada, Apr. 2014, pp. 361–369.

[17] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-invasive detection of moving and stationary human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2329–2342, Nov. 2015.

[18] J. Lv, W. Yang, L. Gong, D. Man, and X. Du, "Robust WLAN-based indoor fine-grained intrusion detection," in *Proc. IEEE Global Commun. Conf.*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[19] S. Di Domenico, G. Pecoraro, E. Cianca, and M. De Sanctis, "Trained-once device-free crowd counting and occupancy estimation using WiFi: A Doppler spectrum based approach," in *Proc. IEEE 12th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, New York, NY, USA, Oct. 2016, pp. 1–8.

[20] T. Xin, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "FreeSense: Indoor human identification with Wi-Fi signals," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–7.

[21] J. Lv, W. Yang, and D. Man, "Device-free passive identity identification via WiFi signals," *Sensors*, vol. 17, no. 11, pp. 1–17, Nov. 2017.

[22] H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. J. Spanos, "WiFi-based human identification via convex tensor shapelet learning," in *Proc. 32nd AAAI Conf. Artif. Intell.*, New Orleans, OR, USA, 2018, pp. 1711–1719.

[23] D. Wang, Z. Zhou, X. Yu, and Y. Cao, "CSIID: WiFi-based human identification via deep learning," in *Proc. 14th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Toronto, ON, Canada, Aug. 2019, pp. 326–330.

[24] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using WiFi signals," in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput.*, Heidelberg, Germany, Sep. 2016, pp. 363–373.

[25] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "WFID: Passive device-free human identification using WiFi signal," in *Proc. 13th Int. Conf. Mobile Ubiquitous Syst.*, New York, NY, USA, 2016, pp. 47–56.

[26] S. Liu, Y. Zhao, and B. Chen, "WiSmart: Robust human access and identification for smart homes using WiFi signals," in *Proc. 3th Int. Conf. Cloud Comput. Secur.*, Nanjing, China, 2017, pp. 855–861.

[27] J. Wang, Y. Zhao, X. Fan, Q. Gao, X. Ma, and H. Wang, "Device-free identification using intrinsic CSI features," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8571–8581, Sep. 2018.

[28] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "WiFi-ID: Human identification using WiFi signal," in *Proc. Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, Washington, DC, USA, May 2016, pp. 75–82.

[29] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, Vienna, Austria, Apr. 2016, pp. 1–12.

[30] S. Sardy, P. Tseng, and A. Bruce, "Robust wavelet denoising," *IEEE Trans. Signal Process.*, vol. 49, no. 6, pp. 1146–1152, Jun. 2001.

[31] A. Murad and J. Y. Pyun, "Deep recurrent neural networks for human activity recognition," *Sensors*, vol. 17, no. 11, pp. 1–17, Nov. 2017.

[32] M. Inoue, S. Inoue, and T. Nishida, "Deep recurrent neural network for mobile human activity recognition with high throughput," *Artif. Life Robot.*, vol. 23, no. 2, pp. 173–185, Jun. 2018.

[33] D. Singh, E. Merdivan, I. Psychoula, J. Kropf, S. Hanke, M. Geist, and A. Holzinger, "Human activity recognition using recurrent neural networks," in *Proc. Int. Cross-Domain Conf. Mach. Learn. Knowl. Extraction*, Reggio, Italy, 2017, pp. 267–274.

[34] A. Dehghaninejad and S. Saviz, "Plasma effects on the growth rate of a helix magnetron-type conducting cylinder," *Eur. Phys. J. Plus*, vol. 131, no. 10, p. 353, Oct. 2016.

[35] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1328–1337, May 2009.

[36] P. van Dorp and F. C. A. Groen, "Feature-based human motion parameter estimation with radar," *IET Radar, Sonar Navigat.*, vol. 2, no. 2, pp. 135–145, May 2008.

[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015, pp. 1–15.

[38] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 1, p. 53, Jan. 2011.

JIANYANG DING (Graduate Student Member, IEEE) was born in Henan, China. He received the B.E. degree from Xidian University, Xi'an, China, in 2018, where he is currently pursuing the Ph.D. degree in indoor positioning. His current research interests include wireless communication and signal processing.
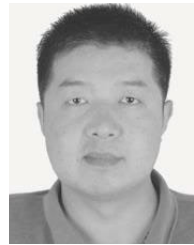
YONG WANG (Member, IEEE) was born in Shaanxi, China. He received the B.S., M.E., and Ph.D. degrees from Xidian University, Xi'an, China, in 1997, 2002, and 2005, respectively. He is currently a Professor with the State Key Laboratory of Integrated Services Networks, Xidian University. His research interests include full-duplex wireless communications and nonlinear predistorter technique.

XIANGCONG FU was born in Hebei, China. He received the B.E. degree from China Three Gorges University, Yichang, China, in 2017. He is currently pursuing the M.E. degree in signal detection with Xidian University. His current research interests include wireless communication and signal processing.

• • •