



HAL
open science

Wi-PIGR: path independent gait recognition with commodity Wi-Fi

Lei Zhang, Cong Wang, Daqing Zhang

► **To cite this version:**

Lei Zhang, Cong Wang, Daqing Zhang. Wi-PIGR: path independent gait recognition with commodity Wi-Fi. *IEEE Transactions on Mobile Computing*, 2022, 21 (9), pp.3414 - 3427. 10.1109/TMC.2021.3052314 . hal-03363404

HAL Id: hal-03363404

<https://hal.science/hal-03363404>

Submitted on 3 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wi-PIGR: Path Independent Gait Recognition with Commodity Wi-Fi

Lei Zhang, *Member, IEEE*, Cong Wang, and Daqing Zhang, *Fellow, IEEE*

Abstract—Wi-Fi based gait recognition has many potential applications. However, the gait information derived from Wi-Fi changes with the walking path. This makes the human identification through gait really challenging, the existing Wi-Fi based gait recognition systems require the subject walking along a predetermined path. This path dependence restriction impedes Wi-Fi based gait recognition from being widely used. In this paper, a path independent gait recognition system for a single subject, Wi-PIGR, is proposed. In Wi-PIGR, the subject is identified through the gait regardless of the walking path. Specifically, an extra receiver is introduced to get CSI data in orthogonal directions. A series of signal processing techniques are proposed to eliminate the differences among signals introduced by walking along the arbitrary paths and generate a high quality path independent signal spectrogram. Furthermore, a deep learning approach is integrated into the feature extraction. The experiment results in typical indoor environment demonstrate the superior performance of Wi-PIGR, with the average recognition accuracy of 77.15%, when the number of subjects is 50.

Index Terms—Channel State Information (CSI), Gait Recognition, Fresnel model, Device-free Sensing.



1 INTRODUCTION

Compared with computer vision [1], [2], [3], floor sensor [4], [5] and wearable sensor [6], [7], [8], Wi-Fi as a new sensing technology has its predominant advantages of ubiquity, better coverage, non-invasive, less or no special modification needed and no privacy violation. As a milestone application of Internet of Things (IoT), Wi-Fi based activity identification attracts extensive attention, more and more efforts are made to explore the frontier of this area.

Recently, there are significant advances in taking advantage of physical layer Channel State Information (CSI) in Wi-Fi signals to identify the human gait [9], [10], [11], [12], [13], [14]. Due to multipath effect of wireless signals, the Wi-Fi signals reflected by human body generates unique CSI dynamics. Intuitively, the extracted gait features from CSI can be applied to identify human through a series of signal processing techniques. However, the path dependence leads the fact that if the subjects walk along an arbitrary path, the gait information derived from the receiving signals changes arbitrarily with the random change of walking path. This leads to the fact that the features derived from the gait information change dramatically and the gait recognition based on the derived features becomes very challenging.

In order to tackle this problem, the existing Wi-Fi based gait recognition systems set a strict restriction that the subject must walk along a predetermined path. This strong constraint impedes Wi-Fi based gait recognition from being widely used. It is necessary to propose a gait recognition system without path dependence.

Before relaxing the restriction of current gait recognition systems, one thing needs to be ensured that the gaits walking along arbitrary paths are consistent so that the stable gait features can be derived and the gait recognition can be achieved. As found by a study, when a subject walks along a straight line path or a curve path, his legs and feet change most, about a few centimeters [15]. The gaits for the same subject have a little bit variation due to body change. Comparing the gait change of one subject walking along two different arbitrary paths with the gait changes of two subjects walking along the same path, the latter has larger gait variation [15], [16]. Therefore, for the same subject, the gaits walking along different paths are consistent and can be employed for the gait recognition. This fact provides a strong support for the gait recognition when the subject walks along an arbitrary path. Note that though, the gaits walking along arbitrary paths are actually consistent, the gait features derived from Wi-Fi signals changes randomly with the random changes of walking path due to the well-known multipath effect of wireless signals. This is the fundamental issue to be tackled in this paper.

Based on the solid foundation, we propose the first Wi-Fi based Path Independent Gait Recognition system, called Wi-PIGR. Wi-PIGR is designed for single subject gait recognition, under the assumption that the subject walks at the speed of $1 \sim 1.5\text{m/s}$ without a sharp turn. The maximum walking area for Wi-PIGR is fixed, which is the limitation of the specific measurement setup. The Wi-Fi transmission power can be increased by using higher gain antennas. The corresponding maximum walking area will be enlarged accordingly. In Wi-PIGR, the path dependence is removed

This work is supported by the NSFC under Grant No.61772364, National Instrument Program under Grant 2013YQ030915, and the EU CHISTCERA RADIOSENSE project. (Correspondence to: C Wang, D Zhang.)

- L. Zhang is with College of Intelligence and Computing, Tianjin University, Tianjin 300050, China, and also with Tianjin Key Laboratory of Advanced Network Technology and Application, Tianjin 300050, China, and also with Key Laboratory of Grain Information Processing and Control, Henan University of Technology, Ministry of Education, Henan 450001, China. (e-mail: lzhang@tju.edu.cn)
- C. Wang is with the 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, Hebei Province 050081, China, and with College of Intelligence and Computing, Tianjin University, Tianjin 300050, China, and also with Tianjin Key Laboratory of Advanced Network Technology and Application, Tianjin 300050, China. (e-mail: congwang@tju.edu.cn)
- D. Zhang is with Institut Mines-Télécom/Télécom SudParis, IP Paris, France. (e-mail: daqing.zhang@telecom-sudparis.eu)

and the subject is identified through the gait regardless of the walking path.

The Wi-PIGR enables many potential applications that require user identification. For instance, for a smart home equipped with Wi-PIGR, an unauthorized subject can be identified after he enters the home. Another application of Wi-PIGR can be this: a smart building equipped with Wi-PIGR in each room can locate a subject in an emergency. Besides these, Wi-PIGR can also be customized to adjust temperature, light and background music according to the taste of each individual.

In summary, the main contributions are as follows:

The first contribution is the strong dependence between the random walking path and gait information is removed by introducing an extra Wi-Fi receiver and sensing the comprehensive gait information. Specifically, when a subject walks along different arbitrary paths, with the dynamic change of walking path, the measured gait information changes dynamically. And the consistent gait features can not be extracted from the measured gait information. The essential reason is the strong dependence between the gait information and walking path. In order to tackle this issue, based on the fact that a pair of transceivers can only obtain the movement information perpendicular to Line of Sight (LoS) but not parallel to LoS [17], [18], [19], [20], [21], an extra transceiver is introduced. This is to form a 2D plane and obtain the both orthogonal direction movement information. The sliding time window is first employed to the denoised orthogonal direction CSI data and cut them into slices. Every small walking path in each slice can be treated as a straight line. Then, each sub-spectrogram is derived from CSI waveform in one window slice and is mapped from the orthogonal directions to the movement direction to remove the direction dependence. The resulting direction independent sub-spectrogram can truly reflect the gait information along the small path in window slice. Finally, all the resulting sub-spectrograms in all the window slices are integrated to form a complete path independent spectrogram and represent the true gait information along the whole path. In this way, regardless of the walking path, the spectrograms can always be mapped to the movement path and the gait information can be derived. In summary, the path dependence is removed by adding an Wi-Fi receiver to obtain the comprehensive gait information.

After the removal of path dependence, the consistent gait features can be extracted from the comprehensive gait information. The second contribution is to leverage a deep learning technology to derive gait features. In [9], [11], the gait features, such as gait speed and stride length are derived manually. However, a small number of features can not guarantee the accurate gait recognition when the number of subjects to be identified increases. There will be more likely several subjects having similar gait features. Besides, the gait recognition with the limited number of features are easier to be influenced by environmental changes. It is challenging to portray the gait comprehensively and accurately with the limited number of features derived from a few absolute gait parameters. This issue is tackled by tailoring the deep learning technology to the common feature extraction. In particular, both Convolutional Neural Network (CNN) [22] and Long Short Term Memory (LSTM) [23] network are

employed to automatically extract the spatial and temporal features. In summary, as the second contribution, the deep learning technology is integrated into the common feature extraction to derive the fine grained features.

Finally, a prototype of the Wi-PIGR with commodity Wi-Fi devices is built and its superiority in various indoor environment is demonstrated with extensive experiments. With 50 subjects, the average recognition accuracy of 77.15% is achieved in Wi-PIGR.

The rest of the paper is organized as follows. We first review the related work in Section 2 and introduce some preliminaries in Section 3. The detailed design and implementation of Wi-PIGR are provided in Section 4. Then, the performance is evaluated in Section 5. Some limitations and discussions are presented in Section 6. Finally, the work is concluded in Section 7.

2 RELATED WORK

Gait Recognition. As a milestone application of Internet of Things (IoT), human gait related recognition attracts extensive attention. There are various sensing techniques, such as vision-based [1], [2], [3], ambient environment sensing-based [4], [5], [24], specialized hardware based [25] and wearable sensor based [6], [7], [8]. However, these techniques have their own limitations. In the vision based system [1], [2], [3], a sequence of human silhouettes are first generated from the video frames by separating the moving objects from the background, and then the sequence of human silhouettes is used to do recognition. However, the privacy violation, the necessity of line of sight and intensive workload for real-time processing make the vision based system less desirable. The floor sensor based and the wall sensor based recognition systems are two typical ambient environment sensing-based recognition systems. In floor sensor based systems [4], [5], force plates installed under the floor, are used for capturing the gait features. WiGait [24] equipped with a sensor hangs on the wall can obtain the absolute gait parameters for the clinical uses, however it does not have the capability of human identification through the unique gait. For the ambient environment sensing-based technology, the special devices are required to install beforehand, the activity induced data are mixed with the environment dependent data. It is hard to extract the activity induced data in some cases. Equipped with specialized hardware (FMCW), RF-Pose [25] does an excellent job on accurately tracking the 2D human pose through walls and obstructions. However, it does not specialize in human recognition through the unique gait. In addition, though more accurate signals can be derived, it is difficult to deploy and cost inefficient. In wearable sensor based systems [6], [7], [8], accelerometers and gyroscopes equipped in smart phones and wearable devices, are used to collect gait information and extract features. For the wearable sensor based sensing technology, people always need to wear and frequently charge wearable sensors, which is obviously inconvenient.

Passive Human Sensing using Commodity Wi-Fi. With the extensive deployment of Wi-Fi devices, significant progress has been made in Wi-Fi based human sensing and recognition [17], [18], [26], [27], [28], [29], [30], [31], [32],

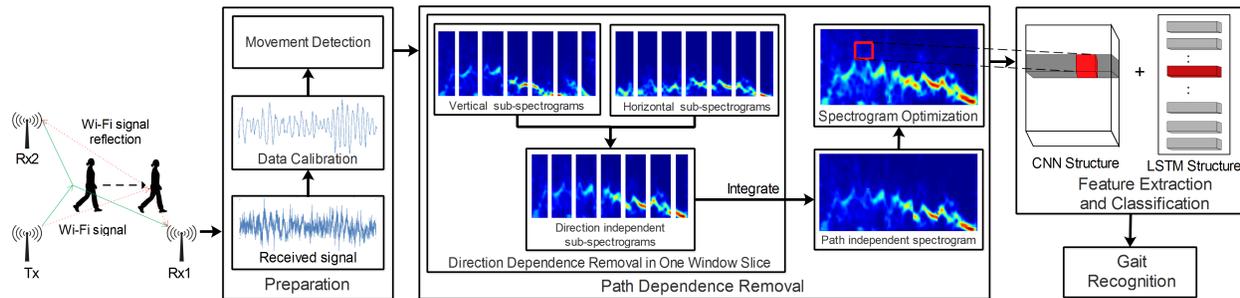


Fig. 1: Overview of Wi-PIGR.

[33], [34], [35], [36], [37], [38], [39]. Various approaches have been proposed for macro-movement and micro-movement recognitions identification and single and multi-person activity recognition. In principle, these are based on the fact that human motions distort multipath profiles during signal propagation. These Wi-Fi based approaches are more appealing since they do not require carrying devices, support omnidirectional coverage even in Non-Line-Of-Sight (NLOS) scenarios, widely deployed and cost-effective, and preserve user privacy gracefully. However, compared with the activity recognition, so far there is very few research for gait recognition with commodity Wi-Fi devices. So far, there are only five typical gait recognition systems: WiDIGR [9], WiWho [10], WifiU [11], WiFi-ID [12] and WiAU [13]. In WiWho [10], the subject is asked to walk along a path 1 meter parallel to LoS and the gait patterns are profiled. In WifiU [11], the subject is required to walk along the straight line path perpendicular to LoS. The time-frequency analysis is utilized to distinguish the radio signal reflections from different body parts. In WiFi-ID [12], the subject is required to walk along the path across LoS. The time and frequency domain features are derived to portray the walking pattern in a simple scenario. In WiAU [13], the subject is required to walk along the predetermined paths. ResNet algorithm with two dedicated loss functions is leveraged to extract gait features. In WiDIGR [9], the gait information is mapped from the orthogonal directions to the movement direction to remove the direction dependence so that the subject can walk along multiple straight line paths. In a word, all current existing gait recognition systems have the constraint of the path dependence and require the subject walking along the predetermined paths. We believe that Wi-PIGR opens up a new direction for the development of commodity Wi-Fi based gait recognition.

3 PRELIMINARY

3.1 Channel State Information

Channel State Information (CSI) depicts how a signal propagates from the transmitter to the receiver and represents the combined effect of scattering, fading and power decay with distance. Suppose $X(f, t)$ is the frequency domain component of the transmitted signals with carrier frequency f at time t , and $Y(f, t)$ is the corresponding frequency domain component received by the receiver. The relationship between $X(f, t)$ and $Y(f, t)$ is $Y(f, t) = H(f, t) \times X(f, t)$, where $H(f, t)$ is the complex

valued Channel Frequency Response (CFR). In practice, there are many propagation paths between transceivers.

$$H(f, t) = \sum_{k=1}^N \alpha_k(f, t) e^{-j2\pi f \tau_k(t)} \quad (1)$$

where N is the total number of multipaths, $\alpha_k(f, t)$ and $\tau_k(t)$ represent the attenuation factor and propagation delay for the k^{th} path respectively.

3.2 Retrieve Path Independence Movement

When a subject walks along the arbitrary paths, the Wi-Fi signals reflected by the human body will be different at the receiver due to the well-known multipath effect of wireless signals. As proposed in [17], [18], [19], [20], [21], a pair of transceivers can only obtain the movement information perpendicular to LoS but not the parallel to LoS. For instance, when a subject walks along the direction perpendicular to LoS, the speed can be obtained by Wi-Fi devices. In contrast, when the walking direction is parallel to LoS, the speed measured by Wi-Fi devices is close to 0. This implies that the obtained gait information changes with the walking direction. In addition, a complete walking path (e.g., a curve path) may be consist of various walking directions. With the dynamic change of walking path, the obtained gait information changes dynamically and the extracted gait features are not consistent for various walking paths. The fundamental issue lies in the strong dependence between the dynamic change of walking path and dynamic change of derived gait information. However, the existing gait recognition systems always require the subject to walk along a predetermined path to avoid the impact of different walking paths.

In order to capture the walking patterns along the arbitrary paths and identify the gait. The sliding time window is adopted. In this paper, another transceiver is introduced to form a 2D observation plane and obtain the orthogonal direction movement information. Since the direction of the walking path can be regarded as constant in a short timing window, the walking path is a straight line. Thus, the whole walking path can be cut into slices. In order to remove the direction dependence and reflect the accurate walking patterns along the small path, the sub-spectrograms derived from CSI waveform in each window slice are mapped on the orthogonal directions to the movement direction. Then, all the direction independent sub-spectrograms are integrated to form a complete path independent spectrogram. Finally, some effective actions are taken to denoise the integrated

spectrogram. The detailed discussion of path independent spectrogram generation is provided in the Path Dependence Removal Section. In this way, regardless of the walking path, the consistent gait information can be derived.

4 THE RECOGNITION SYSTEM

4.1 System Architecture

The designed Wi-PIGR system can effectively achieve the gait recognition by using commodity Wi-Fi devices. As shown in Fig. 1, Data Collection module works as the input to Wi-PIGR, while the results of gait recognition is the output. Wi-PIGR mainly consists of four core modules: Data Collection, Preparation, Path Dependence Removal and Feature Extraction and Classification. The Data Collection module collects CSI signals from the commodity Wi-Fi devices. The Preparation module including two sub-modules: Data Calibration and Movement Detection, works for the gait recognition preparation. The Path Dependence Removal module including two sub-modules: Direction Dependence Removal in One Window Slice and Spectrogram Optimization, produces a high quality path independent spectrogram. The Direction Dependence Removal in One Window Slice produces a series of direction independent sub-spectrograms. The Spectrogram Optimization sub-module optimizes the integrated path independent spectrogram for accurate and stable recognition. The Feature Extraction and Classification module adopting deep learning model, extracts features automatically and generates the recognition results. Thus, the gait is recognized.

4.2 Data Collection

As proposed in [17], [18], [19], [20], [21], a pair of transceivers can only obtain the movement information perpendicular to LoS but not the parallel to LoS. In order to derive the complete movement information, an extra receiver is deployed to form a 2D Fresnel zone [40], as shown in Fig. 2. Two pairs of transceivers are placed vertically, and the vertical and horizontal CSI measurements can be obtained. The Fresnel zone is a series of concentric ellipsoidal regions whose foci are the pair(s) of transceivers [40]. For each CSI stream, CSI data are obtained from 30 OFDM subcarriers by 802.11n [41]. For each pair of transceivers, one antenna works as a transmitter and three antennas work as a receiver. Hence, in each received 802.11n frame, there are $1 \times 3 \times 30 = 90$ CSI values. As the sampling rate is 1000 Hz, there are 90000 CSI values collected by each pair of transceivers.

4.3 Preparation

4.3.1 Data Calibration

CSI measurements contain noises generated by the internal state transitions in the transmitter and the receiver, including transmission power changes, transmission rate adaptation, and internal CSI reference level changes [29]. The phases of CSI are too noisy to be used for gait recognition due to carrier frequency offset [29], only the amplitude is adopted for human gait recognition. The bandpass filtering and Principal Component Analysis (PCA) are leveraged to mitigate the noises as much as possible. The detailed data calibration can be found in WiDIGR [9].

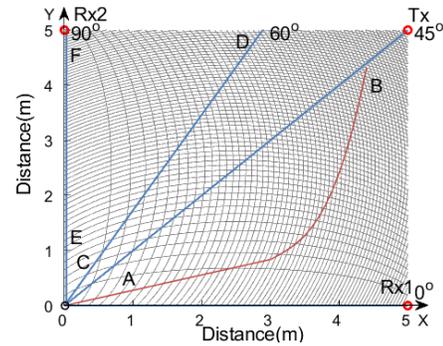


Fig. 2: Placement of Wi-Fi devices and the Fresnel zones.

4.3.2 Movement Detection

After data calibration, Wi-PIGR begins the threshold based walking start detection. In this paper, the walking detection scheme is designed based on the fact that human movement generates higher variances in CSI measurements than background noises. In the absence of movement, there is smaller variance of background noises, and the noise level changes slowly over the time due to the relatively fixing environment. In order to track the noise level, a dynamic threshold algorithm is leveraged. Specifically, the Exponential Moving Average algorithm is adopted to update the noise level L_t during the silent period.

$$L_t = (1 - \alpha)L_{t-1} + \alpha \times var_t \quad (2)$$

where the coefficient α is set to be 0.1 empirically. And the variance var_t for i^{th} sliding window with 200 samples is calculated in the first PCA component. The detection threshold is set to be three times of the noise level L_t when the variance var_t in one sliding window deviates from the average noise level. In this way, the start and end of walking can be discovered.

4.4 Path Dependence Removal

The purpose of this section is to obtain the path independent gait information and prepare for the following feature extraction. As mentioned in 3.2, in order to tackle the problem of path dependent gait recognition along arbitrary path, a sliding time window is applied to cut the whole walking path into slices. The small path in each window slice can be treated as a straight line and the window size is set to be 1s empirically. Then, sub-spectrogram is derived from CSI waveforms in each window slice. As proposed in [17], [18], [19], [20], [21], a pair of transceivers can only obtain the movement information perpendicular to LoS but not parallel to LoS. According to this, the small sub-spectrogram in each window slice is mapped from the orthogonal direction to the movement direction and generate a direction independent sub-spectrogram. This sub-spectrogram can reflect the accurate gait information of walking along the small path within the window slice. Note that the sub-spectrogram refers to a small spectrogram in one window slice. Then, all the sub-spectrograms in all the window slices are integrated to form a complete path independent spectrogram and represent the accurate gait information along the whole path. The whole processing

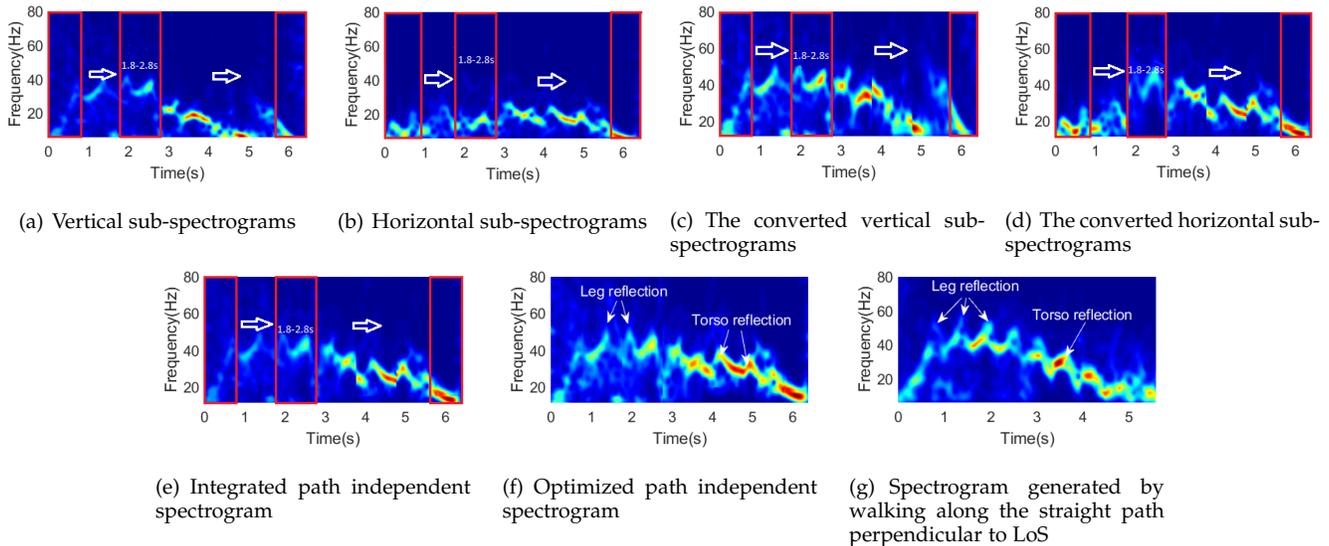


Fig. 3: Spectrograms in the Path Dependence Removal.

involves in two parts, the Direction Dependence Removal in One Window Slice and the Spectrogram Optimization, which will be discussed in the details in the following.

4.4.1 Direction Dependence Removal in One Window Slice

In this part, how to derive a direction independent sub-spectrogram $S(i)$ from the CSI in the i^{th} ($1 \leq i \leq N$) window slice is discussed, where N is the number of the window slices. Since the small path in each window slice can be treated as a straight line and contains one walking direction, the gait information needs only to be mapped from the orthogonal directions to the movement direction. The resulting direction independent gait information can truly reflect the walking patterns along the small path in window slice.

When walking, different portions of the human body move at different speeds, the radio signal reflections from different body portions have different frequency and are mixed together in the CSI waveform [11]. In order to distinguish the radio reflections of different body portions, Short-Time Fourier Transform (STFT) is leveraged to transform the vertical and horizontal CSI measurements in the window slice into sub-spectrograms so that the gait information can be analyzed in the time-frequency domain. The spectrogram is adopted to profile human gait patterns. In order to maintain the well balanced time and frequency resolution, the FFT size is set to be 1024 samples and the sliding window step size is set to be 6 samples. The sub-spectrograms between 1.8 ~ 2.8s in Fig. 3(a) and 3(b) are generated when the walking time is 1.8 ~ 2.8s. The spectrograms in Fig. 3(a) and 3(b) are generated by walking along the curve path \overline{AB} in Fig. 2. The red color indicates higher reflection energy. Then, each sub-spectrogram is mapped from the orthogonal directions to the movement direction to remove the direction dependence by a series of signal processing techniques proposed by WiDIGR [9]. The technique includes the methods of direction estimation, sub-spectrogram generation, transformation, alignment and fusion. This is out of the scope of this work, the details can

be found in WiDIGR [9]. The sub-spectrograms shown in the range of 1.8 ~ 2.8s of Fig. 3(c) and 3(d) are formed during the processing of sub-spectrogram transformation. Finally, a high quality direction independent sub-spectrogram $S(i)$ in the i^{th} window slice can be obtained after the processing, as demonstrated in the range of 1.8 ~ 2.8s in Fig. 3(e). The resulting direction independent sub-spectrogram can truly reflect the gait information along the small path in one window slice.

4.4.2 Spectrogram Optimization

So far, each sub-spectrogram has been mapped to the small path in the corresponding window to reflect the gait information along the small path in one window slice. Similar to the small paths in all the window slices can be integrated to form the whole path, all the sub-spectrograms $\{S(1), \dots, S(N)\}$ in all the window slices can be integrated to form a complete path independent spectrogram, as demonstrated in Fig. 3(e). The resulting spectrogram can reflect the gait information along the whole path. However, when the walking information is piecewise estimated, it is easy to make information in two adjacent segments discontinuous. As shown in the range of 3 ~ 5s of Fig. 3(e), the integrated spectrogram has obvious dislocation deviations at the segment points. The reasons behind this fact can be summarized as follows. First, the walking path in each small segment is regarded as a straight line, which causes the path curve unsmooth. Although the path dependence problem is tackled, it will lead to the dislocation deviation of the spectrogram. Second, since CSI is sensitive to noise, the direction estimation in each small segment will be easily affected by noise. In order to eliminate the dislocation deviation in the integrated spectrogram, some actions need to be taken to optimize the integrated information. The dislocation deviation can be eliminated by shifting the sub-spectrogram in each window slice up or down. The total shifting distance of sub-spectrograms should be as little as possible to maintain the optimization of the spectrogram. The implementation details are provided in *Algorithm 1*.

Algorithm 1 Spectrogram Optimization

Input: all the direction independent sub-spectrograms $\{S(1), \dots, S(N)\}$;
Output: An optimized path independent spectrogram S ;
1: **for** $i := 1$ to $(N - 1)$ **do**
2: Obtain the relative position between adjacent sub-spectrograms using the cross-correlation;
3: **end for**
4: Derive the absolute position from the relative position;
5: Calculate the median of the absolute position;
6: Eliminate the dislocation deviation by moving all the sub-spectrograms $\{S(1), \dots, S(N)\}$ to the same level as that in the median;
7: Obtain a optimized path independent spectrogram S by integrating all the moved sub-spectrograms and applying a Gaussian low-pass filter;
8: **return** S ;

The first step is to determine the absolute position of sub-spectrograms in each window slice. Note that the position refers to the energy frequency level. Specifically, for a aggregated spectrogram with N sub-spectrograms, cross-correlation is applied to the last column of the matrix representing i^{th} sub-spectrogram and the first column of the matrix representing $(i + 1)^{th}$ sub-spectrogram to obtain the relative position between the adjacent sub-spectrograms. Then, the absolute position of each sub-spectrogram can be derived based on the relative position.

The second step is to determine the shifting distance of each sub-spectrogram and eliminate the dislocation deviation by the shifting distance. In order to minimize the shifting distance, the positions after shift is needed. Since the absolute position of each sub-spectrogram is known, the problem becomes shifting all to the same position to find the shortest shifting distance. Specifically, all the absolute positions of sub-spectrograms are first sorted to obtain the median. And then, all the sub-spectrograms are shifted to the same level as the sub-spectrogram in the median. Finally, a 2-dimensional Gaussian low-pass filter with a size of 5 and σ of 0.8, is applied to the resulting spectrogram for optimization.

So far, the dislocation deviation in the spectrogram is eliminated, and an optimized path independent spectrogram S is derived. The resulting spectrograms give us detailed human gait information similar to those generated by Doppler radars [42], [43], and the different parts of the body’s reflection can be extracted from CSI spectrograms using insights from Doppler radars [11], [42], [43]. In particular, the moving speed changes of different body parts (e.g., torso and legs) can be clearly observed by the percentile method [9], [42], [43]. Since the torso has a larger reflection area than other portions of the body, the signal reflected from the torso is the strongest in the spectrogram. In Fig. 3(f) or Fig. 3(g), a high energy (“hot” colored) band corresponding to torso reflection can be identified easily, and the torso reflection energy fluctuates regularly about twice per second. In addition, some “blue flame” around the high energy band in the frequency region of 35 ~ 55 Hz represents the reflections of swinging legs. The magnitude

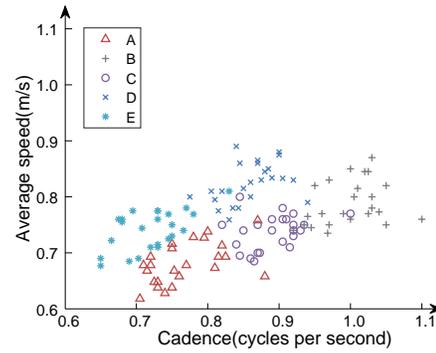


Fig. 4: Distribution of average speed and cadence.

of leg reflections is weak due to the much smaller reflection area of legs. Compared with the direction dependent spectrograms with the energy frequency below 40 Hz in Fig. 3(a) or Fig. 3(b), the path independent spectrogram in Fig. 3(f) or Fig. 3(g) has a higher energy frequency, up to 55 Hz. All in all, Fig. 3 records spectrogram changes step by step when the path dependence is removed. Initially, there are hardly distinguished gait patterns. And then, the gait patterns begin to appear. Finally, after the removal of path independence, clear gait patterns are noticeable. Even the moving speed changes of different body parts, e.g., torso and legs, can be obtained from the spectrogram. This analysis of spectrograms not only verifies the necessity of the path dependence removal, but also provides a deep understandings of gait pattern changes.

The walking path is plotted as the curve \overrightarrow{AB} in Fig. 2, and the optimized path independent spectrogram is demonstrated in Fig. 3(f). Different from the raw spectrograms in Fig. 3(a) and 3(b), the optimized spectrogram in Fig. 3(f) is similar to the spectrogram in Fig. 3(g). Fig. 3(g) is derived by walking along the straight path perpendicular to LoS, e.g., along the straight line path \overrightarrow{EF} in Fig. 2. This implies the fact the resulting spectrogram is path independent and is not affected by the change of the walking path. Therefore, the gait information derived from the resulting spectrogram is also path independent and can be directly applied to the gait recognition.

In order to verify the effectiveness of the resulting path independent spectrograms on gait recognition, we design an experiment. As concluded in [11], [16], a human subject keeps his footstep length nearly constant, but has small variations in his cadence so that his average speed is linearly related to the cadence. If we can verify the linear relationship between average speed and the cadence extracted from the path dependence removal spectrograms, the effectiveness of the spectrograms on gait recognition can be proved. Specifically, five randomly selected subjects A, B, C, D and E were asked to walk along the arbitrary paths one by one. Their gait patterns were recorded and corresponding path independent spectrograms were generated. Two typical gait features, average speed and cadence were derived from the path independent spectrograms. The reason of only these two features were extracted is their linear relationship can verify the effectiveness of the spectrograms. The distributions of average speed and cadence for five subjects were

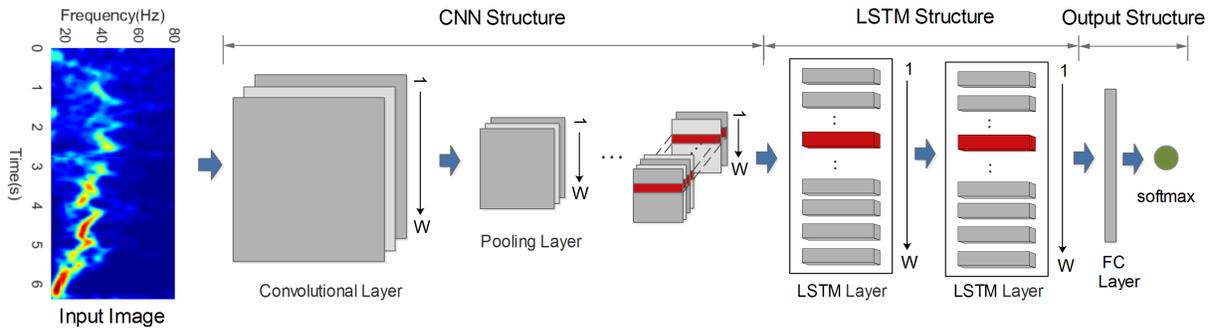


Fig. 5: The deep learning architecture of Wi-PIGR.

illustrated in Fig. 4. In Fig. 4, X-axis is the cadence, defined as the gait cycles a person produces per second. From Fig. 4, we can see that the subjects samples are scattered around the two straight lines, which indicates the linear relationship between average speed and the cadence extracted from the path dependence removal spectrogram though there are a few outliers possibly caused by measurement errors. This experiment roughly verifies the effectiveness of the spectrograms on gait recognition. However, the manually extracted features are not appropriate for the gait recognition. First, the manual feature extraction approaches are extracted by experience, do not perform well when the number of subjects increases. Second, the feature extraction is easily influenced by environmental changes. Last, the number of features extracted is very limited. So it is necessary to find an effective method compensating for these defects. Deep learning with the potential to eliminate the need for manually specifying appropriate feature representations through automatically learning data representations, integrated into an overarching classification model is very appealing. To make recognition more accurate, deep learning methods is leveraged and automatically extract the comprehensive features from the path independent spectrograms.

4.5 Feature Extraction and Classification

Wi-PIGR employs an advanced deep learning model to analyze gait patterns through supervised learning. The hidden layer in our deep learning architecture includes a CNN structure and an LSTM structure. CNNs are widely used for sequential data analysis by employing a sliding window preprocessing step that carves out consecutive sample data. However, just as with the caveat of sliding-window procedures in conventional machine learning, great care has to be taken when parameterizing the approach. Inappropriate window lengths inevitably lead to suboptimal recognition results as the filters of the CNN will capture irrelevant or only partial information that shall be analyzed. LSTM models are particularly attractive, as their specialized internal structure implements a memory that includes a forget function to very effectively and selectively focus on the data stream parts that are actually relevant for the recognition process. As such, these models not only integrate representation learning and classification, but also effective segmentation, which is of utmost importance for gait recognition. The capabilities of the various modeling variants have been effectively analyzed so that specific suggestions

can be given on when to use a particular modeling variant as well as what (hyper-) parameters should be subject to optimization efforts [44], [45]. Currently, the most effective and thus successful modeling variants are combinations of CNNs and LSTM models that integrate the great feature learning capabilities of CNNs with the sequential modeling capabilities of LSTM.

4.5.1 Deep Learning Model

As shown in Fig. 5, the network consists of a CNN structure with three convolutional layers and pooling layers, a LSTM structure with two LSTM layers, and an output structure with one fully-connected layer and a softmax function. The number of layers is determined based on the preliminary experiment.

(1) *Convolutional Layer.* The convolutional layer replaces matrix multiplications with convolution operations. In Wi-PIGR, two-dimensional convolution is adopted:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n) \quad (3)$$

where I is the input, K is the kernel, and S is the output [46]. The convolutional layers use sliding filters, and their outputs have roughly the same aspect ratio as the inputs. These outputs are known as feature maps [47]—involving not only the strength of the responses, but also the spatial positions. The convolutional layer learns the feature maps while processing the input. Since the convolution layer shares the same feature map for the whole input, it significantly reduces computation overhead for both training and testing. In this research, the input size is set to be 84×1000 , obtained by downsampling of the optimized path independent spectrogram. In Wi-PIGR, three 3×3 kernels are used in each layer and the stride is set to be 1 in both vertical and horizontal directions. The number of kernels controls the number of channels in the output of the convolutional layer. Thus, there are 27 channels in the output of the last convolutional layer.

In addition, each of convolutional layer is followed by pooling layer. These pooling layers can also be considered as “convolutional”, in the sense that they use sliding windows. After the spectrogram processed by CNN, a 3D matrix is generated. On the time axis of the 3D matrix, the 3D matrix is divided into a series of small 3D matrices with the same size. Then, all the small 3D matrices as one sample are input into LSTM in order, where each small 3D matrix is reshaped into a one-dimensional vector by a window filter [48].

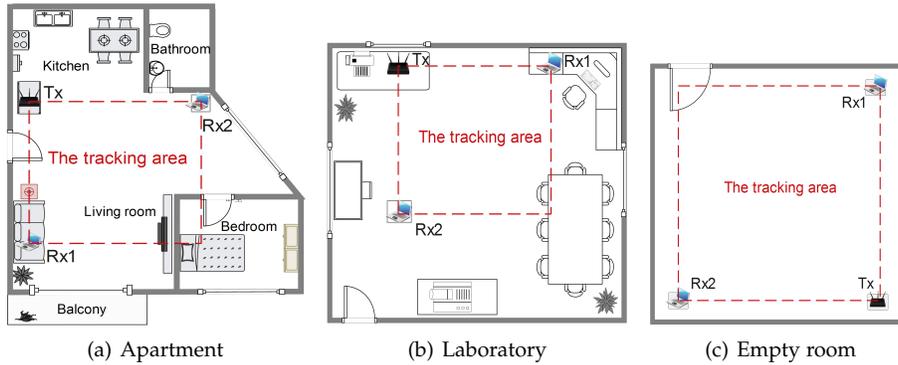


Fig. 6: Different layouts.

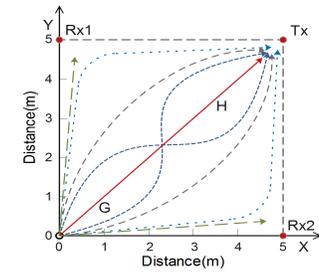


Fig. 7: Placement of Wi-Fi devices and examples of walking paths.

(2)*LSTM Layer.* The recurrent layers (LSTM layers) is employed to learn the temporal dynamic features extracted by the convolutional layers. LSTM as a RNN architecture with memory cells, permits to the learn temporal relationships at a long time scale. A standard RNN fails to learn the long-term dependencies because of the vanishing gradient problem [49]. In order to solve this problem, LSTM employs memory cells holding past information and updates the cell state [50]. Therefore, it has good performance in time-series data processing. In this paper, the two LSTM layers employ a sigmoid activation function, and the number of nodes for each layer is 128.

(3)*Fully-connected Layer.* The fully-connected layer widely used to avoid overfitting [51], connects all its neurons to the neurons in the former layer, namely, the second LSTM layer. The effect is to combine all the features learned in the former layer together to classify the input. The size of the fully-connected layer is equal to the number of all possible classes.

The last component of the deep learning approach is the output, which receives the outputs from the last fully-connected layer and normalizes them with a softmax function. This function computes the distribution probabilities of each one's gait and the one higher than a given threshold is finally recognized as the predicted subject.

5 EXPERIMENTS AND EVALUATION

5.1 Implementation

Experimental Environment: We implement a prototype of Wi-PIGR with three ThinkPad T-series laptops equipped with Intel 5300 wireless NICs. Specifically, one laptop with an antenna works as the transmitter, and each of the others with three antennas works as the receiver. Two pairs of transceivers on the same heights (80cm) are placed vertically so that a 2D Fresnel zone is formed, as illustrated in Fig. 2. The detailed impact of the transceiver placements on Wi-PIGR can be found in WiDIGR [9]. The sender and receivers are placed on tables with a height of 80cm and the distance between them are 5 meters. Each receiver provides 3 streams in total as well as 90 streams of CSI values working in IEEE 802.11n monitor mode on Channel 161 at 5.825 GHz. All laptops are installed with Linux 802.11n CSI Tool [52] on Ubuntu desktop 11.04 LTS OS and process the CSI

measurements by MATLAB. The transmission rate is set to be 1000Hz. The transmitter power is 15dBm by default.

Experimental Methodology: As shown in Fig. 6, Wi-PIGR's performance is evaluated in three typical indoor scenarios. An apartment consists of a bathroom, a bedroom, a kitchen and a living room with one sofa and one television. A laboratory with size 8m×10m has five tables, ten chairs and computers. An empty room has the size of 7m×8m. As presented in Fig. 7, Tx is located at the point (x = 5, y = 5) on the coordinate. On x axis, Rx1 has 5m distance to Tx. On y axis, Rx2 has 5m distance to Tx. The tracking area is a 5m×5m square.

To study the impact of the user diversity, 60 volunteers including 35 males and 25 females were recruited and tested over 8 months. They vary in age from 18 to 42 years old, in weight from 43kg to 95kg, and in height from 155cm to 195cm. During the experiments, each subject was asked to walk along each arbitrary path as demonstrated in Fig. 7. Each subject was asked to walk along arbitrary paths 170 ~ 200 times in one indoor scenario, and 16140 gait instances were collected in three indoor scenarios, including 8960 from the laboratory, 4450 from the empty room and 2730 from the apartment. The subject may walk along the similar paths or different paths at different time as long as he felt comfortable. Moreover, the walking duration was 5 seconds to guarantee data sufficiency. In order to obtain the ground truth, HD videos were taken to record the testing.

Evaluation Metrics: To analyze the performance of Wi-PIGR, True Positive Rate (TPR) and False Positive Rate (FPR) are adopted for single user identification and classification accuracy is adopted for multiple user recognition. TPR is defined as the rate that the true subject is correctly identified:

$$TPR = \frac{TP}{TP + FN} \quad (4)$$

and FPR is defined as the rate that a stranger is wrongly identified as a target subject:

$$FPR = \frac{FP}{TN + FP} \quad (5)$$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives and false negatives respectively. Classification accuracy is defined as the proportion of correctly classified instances to total instances.

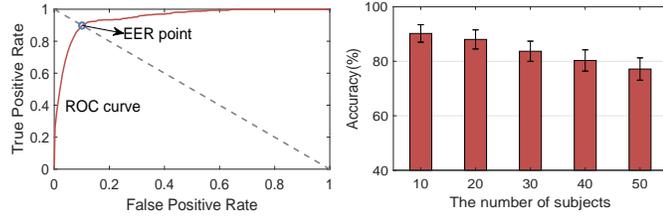


Fig. 8: ROC curve of the average identification accuracy for one subject identification.

Fig. 9: The classification accuracy of multiple subjects.

TABLE 1: The identification accuracy in three typical indoor scenarios.

	True Positive Rate	False Positive Rate
Apartment	92.84%	8.37%
Laboratory	93.49%	7.43%
Empty room	94.18%	6.65%

5.2 Overall Performance

To evaluate the identification accuracy for single subject, 10 subjects were chosen from the 60 subjects as the benchmark set. And then, for each subject not in the benchmark set, a two class classifier that treats the target subject as the positive class and treats benchmark set as the negative class was trained. The rest 49 subjects were treated as strangers in the testing phase. The identification accuracy was calculated using 10 fold cross validation, where one tenth of the walking data set were used for testing and the remaining were used for training. None of the testing data were available to the deep learning model during training. The final accuracy was averaged based on 100 randomly selected benchmark sets.

Wi-PIGR was implemented with commodity Wi-Fi devices and evaluate its performance in in three typical indoor scenarios, empty room, laboratory and apartment. The results were summarized in Table 1. It is obvious that Wi-PIGR demonstrates an excellent performance in different scenarios, with mean TPR of 93.50% and mean FPR of 7.48%, when setting the probability threshold for identification to 0.5. The encouraging results exhibit the superior of Wi-PIGR in different scenarios though, there are still slightly differences. In the apartment, Wi-PIGR performs worst, TPR is 92.84% and FPR is 8.37%. The reason is the apartment has the most complicated layout and the corresponding multipath interference is most serious. In an empty room, Wi-PIGR performs best since the simplest layout and stable multipath effect. In addition, the Receiver Operating Characteristic (ROC) curve of the average identification accuracy is plotted in Fig. 8. The ROC curve exhibits the tradeoff between TPR and FPR over various settings. In Fig. 8, Wi-PIGR achieves average TPR and FPR of 92.52% and 7.48% respectively, at the Equal Error Rate (EER) point. The encouraging results demonstrate the feasibility of Wi-PIGR for single subject identification.

To evaluate classification accuracy for multi-subject recognition, in each experiment, N subjects were randomly picked up from the walking data set, and N ranged from 10 to 50. The classification accuracy was obtained via 10

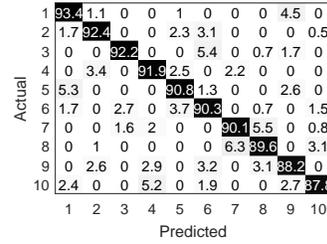


Fig. 10: Confusion matrix of recognition rate for 10 subjects.

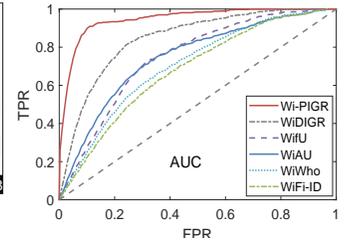


Fig. 11: Compared with the baseline gait recognition systems.

fold cross validation. This experiment was performed 100 times to calculate the average classification accuracies. Fig. 9 shows the results when the number of picked subjects ranges from 10 to 50. The classification accuracies decreases with the increase of the picked subjects: the classification accuracy decreases from 90.21% to 77.15%. The reasons behind this fact can be summarized as following. More subjects introduces more similar cadences and more chances of gait similarity. In addition, it is obvious that when the number of subjects increases, the number of classes increases, the identification becomes more challenging and accuracy drops accordingly. All the factors have a negative influence on the overall performance. However, even if the classification accuracy is at the lowest level of 77.15%, the system can still have a good achievement. Note that the task of multi-subject recognition is more complicated than that of the single subject recognition since not only determining whether the test subject is a true subject is needed, but also recognizing exactly who he or she is necessary.

In order to further examine the performance of multi-subject identification, Wi-PIGR was tested from a different perspective. The confusion matrix of recognition rate for 10 subjects is provided in Fig. 10. This indicates the satisfactory results. The numbers on the confusion matrix indexes are the subject numbers. The confusion matrix is coincident with overall performance and it can be regarded as some special cases of the overall performance. It is clear that Wi-PIGR is able to maintain a fairly well performance in most cases.

5.3 Performance Evaluation

5.3.1 Performance Comparisons

The Wi-PIGR was compared with five state-of-the-art baselines gait recognition systems, including WiDGR [9], WiWho [10], WiU [11], WiFi-ID [12] and WiAU [13] in three typical indoor scenarios. In each testing, there is only one subject walking in his natural way without any restriction. The walking paths change arbitrarily. The performance of the five systems is portrayed in Fig. 11. The larger the area under the ROC Curve (AUC) is, the better performance of the system is. There is a distinguishing performance gap between Wi-PIGR and other four systems. The reason of the obvious degradation lies in that the four gait recognition systems are direction dependent and features extracted cannot reflect the exact movement information. While Wi-PIGR is path independent recognition system and maintains excellent performance record.

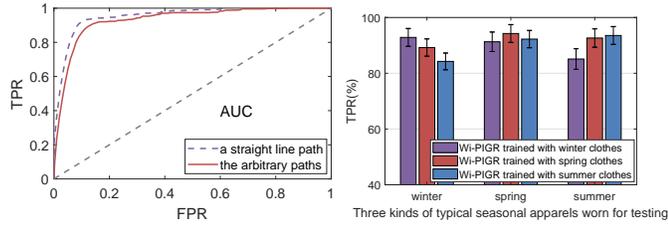


Fig. 12: Verification of path independence. Fig. 13: Impact of apparel.

TABLE 2: The computational complexity and the time consumption of processing ten seconds CSI data.

Modules	Complexity	Time consumption (s)
Data Calibration	$O(\min(L^3, N_c^3 N_s^3))$	0.8863
Movement Detection	$O(L)$	0.0502
Path Dependence Removal	$O(LN_v \log_2 L)$	0.6042
Feature Extraction and Classification	$O(\sum_{l=1}^D M_l^2 \cdot K_l^2 \cdot C_{in} \cdot C_{out} + W)$	0.1042
Total	$O(\min(L^3, N_c^3 N_s^3))$	1.6449

5.3.2 Verification of Path Independence

In order to verify the path independence, Wi-PIGR was first trained with training samples of a subject walking along a straight line path \overrightarrow{GH} in Fig. 7. And then, its accuracy was evaluated on testing samples when the same subject walked along the same straight line path \overrightarrow{GH} and walked along the arbitrary paths as in Fig. 7. None of the testing data were available to the learning model during training, and the final accuracy was averaged based on 100 randomly selected training data. The results of gait recognition for different walking paths are provided in Fig. 12. Although the recognition accuracy of Wi-PIGR for arbitrary paths decreases slightly, it still maintains a pretty good performance. This implies that in Wi-PIGR, the path dependence is removed and the subject is identified successfully through the gait regardless of the walking path.

5.3.3 System Complexity Analysis

In order to deep understand real-time performance of Wi-PIGR, its efficiency and computational complexity of each processing procedure were calculated. Wi-PIGR was evaluated on a laptop equipped with NVIDIA Geforce GTX 1060 GPU card. The computational complexity and the time consumption for processing ten seconds of CSI measurements containing one gait sampling is provided in Table 2, where L is the length of the CSI amplitude, $N_c = 3$, $N_s = 30$ and $N_v = 10$ denote the number of CSI streams, subcarriers and the first ten PCA components respectively, D is the number of convolutional layers, M is the size of output feature map, K is the size of kernel, C_{in} and C_{out} are the number of input channels and output channels respectively [46], and W is the number of weights [50]. The major computational cost is data calibration (involves the PCA processing) because this step needs to process the raw CSI measurements having $90 \times 1000 \times 10 = 900k$ samples. Since the data calibration reduces the data size to one tenth of the raw CSI measurements, the remaining steps take a small

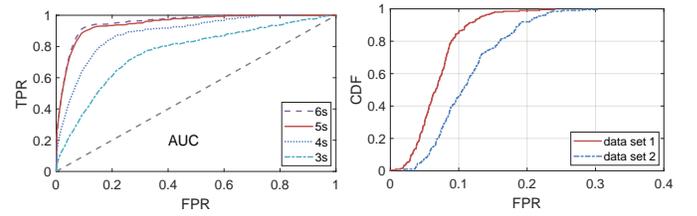


Fig. 14: Impact of walking time. Fig. 15: Impact of human gait evolution.

amount of computational time. For the model that need to recognize 50 subjects, the average training time is about 15.5 minutes. While the testing time is only 0.1 seconds for recognizing one gait instance. The highly efficient testing process enables Wi-PIGR to continuously recognize the gait. In summary, the system complexity analysis implies that the trained Wi-PIGR can successfully work in real-time.

5.3.4 Impact of Apparel

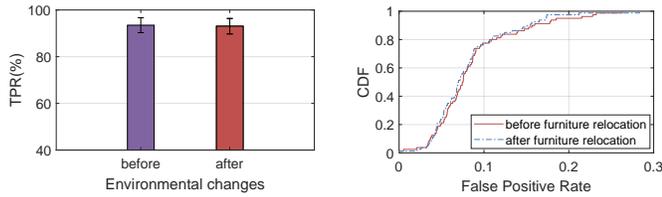
To study the impact of apparel change on the accuracy, Wi-PIGR was first trained with training samples of a subject wearing winter clothes, spring clothes and summer clothes respectively. And then, the accuracy of Wi-PIGR trained with each seasonal clothes was tested and evaluated with three different seasonal apparels. The accuracy was calculated using 10 fold cross validation. None of the testing data were available to the learning model during training. Fig. 13 plots the TPR of Wi-PIGR with different seasonal apparels when FPR is set at 9.85%. In Fig. 13, TPR decreases when the subjects wear different seasonal apparels. This verifies the fact that human gait changes when wearing different types of apparel [6]. Despite the fact that Wi-PIGR are never trained on the samples with other two seasonal apparels, it can still maintains TPR above 84.26%.

5.3.5 Impact of Walking Time

The walking time for correct gait recognition is adopted as another factor to evaluate the performance of Wi-PIGR. Intuitively, the longer the walking time, the richer the gait information can be derived. Fig. 14 shows the results when the walking time is set to be 3s, 4s, 5s, and 6s respectively. It is obvious that there is a diminishing improvement of the identification accuracy with walking time increase. Furthermore, the system with walking time of 5s gives nearly the same identification accuracy as that of 6s. In summary, the longer walking time can improve the accuracy of the system until it converges to a limit at 5 seconds.

5.3.6 Impact of Human Gait Evolution

To study the performance of Wi-PIGR against gait evolution with time, data set 1 was first collected. After eight months, data set 2 for the same subjects in the same environment was collected. Then, Wi-PIGR was trained with the data randomly selected from data set 1, and tested with the remaining data in set 1 and all data in set 2 respectively. None of the testing data was available to the learning model during training. The final accuracy was averaged based on 100 randomly selected training data. Wi-PIGR achieves mean TPR of 94.18% by testing with data set 1, and 90.04%



(a) Mean TPR before and after furniture relocation (b) CDF before and after furniture relocation

Fig. 16: Result of gait recognition before and after environmental changes.

by testing with data set 2, when the probability threshold for gait recognition is set to be 0.5. Fig. 15 presents the Cumulative Distribution Functions (CDF) of FPR with different testing data. Similar to the result of TPR, FPR with testing data set 1 is superior to that of with the testing data set 2. However, even if the identification accuracy decreases when using data set 2 for testing, Wi-PIGR can still maintain a pretty good performance. This implies that the evolution of human gait with time does not have a significant influence on the accuracy of Wi-PIGR.

5.3.7 Resilience to Environment Change

In order to test Wi-PIGR's resilience to the environment change, a two-step experiment was carried out in the laboratory. The settings of the first step are illustrated in Fig. 6(b). And then, in order to simulate the environment changes, the tea table with two seats was moved to another side of the room and the table with six chairs was moved 2 meters away. Wi-PIGR was trained in the first scenario, and tested in both scenarios. The identification accuracies of two scenarios were compared when the probability threshold for identification is set to be 0.5.

Fig. 16 presents the results of gait recognition before and after furniture relocation. The overall mean FPR before and after furniture relocation are 7.43% and 7.08%. The mean TPR are 93.49% and 93.06%, respectively. It is obvious that Wi-PIGR has a certain resilience to the environment changes. In order to further mitigate the negative impacts led by the environment changes, the transfer learning method, such as in [53] might be adopted. And this is out of the scope of this paper.

5.3.8 Impact of Slicing Window Size

The major design question is the choice of window size: if the window size is too small, it cannot reliably estimate walking direction due to noise sensitivity; if the window size is too large, people may walk in different direction during this window, and thus also lead to imprecise estimation. The size of sliding window is affected by walking speed and turning angle. During the experiments, the subjects were asked to walk at a speed of 1 ~ 1.5m/s without intentional speed up or slow down. In addition, the training speed and the testing speed were kept the same as much as possible, otherwise the identification accuracy can not be guaranteed. If the walking speed changes arbitrarily, the gait will change arbitrarily. This arbitrary change is not beneficial to gait recognition. In order to guarantee the accuracy of gait recognition, the subjects were asked to walk without

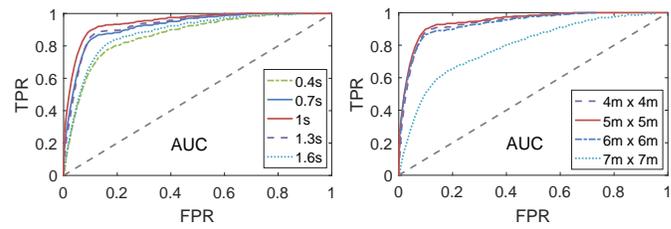


Fig. 17: Impact of Slicing Window Size.

Fig. 18: Impact of Walking Area.

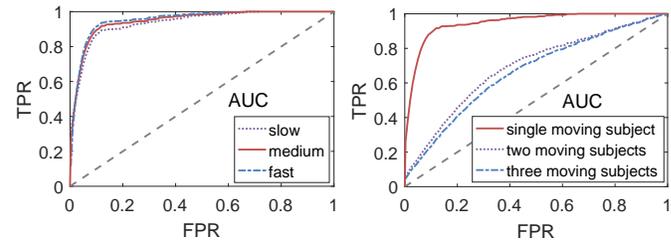


Fig. 19: Impact of Walking Speed Diversity.

Fig. 20: Impact of Multiple Moving Subjects.

arbitrary speed change as long as they felt comfortable. Normally, an adult takes 1 ~ 2 steps a second. When people keep walking at a constant speed, the turning angle within two steps should not be greater than 90 degrees, otherwise the body will easily lose balance and fall [16]. Therefore, under such constraints, different sliding window sizes from 0.4 to 1.6s were set and compared, as shown in Fig. 17. When the sliding window size is set to be 1s, Wi-PIGR performs best. When the sliding window size is set to be 0.4s, Wi-PIGR performs worst. The reason is the influence of signal noise on the walking direction estimation is significantly amplified when the walking path within each sliding window is short. When the sliding window size is set to be 1.6s, Wi-PIGR performs worse than 1s since 1.6s is too long to regard the path in the window as a straight line. Therefore, in this study, the best sliding window size is 1s.

5.3.9 Impact of Walking Area

To study the impact of walking area change on the performance, Wi-PIGR was evaluated with various walking areas from 4m×4m to 7m×7m when the transmitter power is 15dBm by default. Fig. 18 shows the results across different walking areas. When the walking area changes from 4m×4m to 6m×6m, Wi-PIGR can maintain a pretty good performance. However, when the walking area reaches 7m×7m, the identification accuracy drops significantly. When the walking area increases, transmission distance increases, the signal strength weakens and packet loss increases significantly. This makes the fine grained gait identification more difficult. The signal strength is defined by the TX/RX gain and power of the Wi-Fi system. So the maximum walking area for Wi-PIGR is 6m×6m, which is the limitation of the specific measurement setup. The Wi-Fi transmission power can be increased by using higher gain antennas. The corresponding maximum walking area will be enlarged accordingly.

5.3.10 Impact of Walking Speed Diversity

According to the literature [54], [55], the human comfortable walking speed v is $1 \sim 1.5$ m/s. In order to check the impact of walking speed, the subjects were asked to walk at three different speed levels: slow (below $(v - 0.3)$ m/s), medium ($(v - 0.3)$ m/s $\sim (v + 0.3)$ m/s) and fast (above $(v + 0.3)$ m/s). The accuracy was calculated using 10 fold cross validation. None of the testing data were available to the learning model during training. As shown in Fig. 19, even with very different walking speeds, Wi-PIGR still achieves consistent high gait recognition accuracies.

5.3.11 Impact of Multiple Moving Subjects

To test the impact of multiple moving subjects on its performance, Wi-PIGR was evaluated with various number of moving subjects. The moving subjects in each group were randomly chosen from 60 volunteers and each group had two or three moving subjects. The moving subjects walked randomly without any intention. For each group, 170~200 gait instances were collected, and a two class classifier treating the group as the positive class and treating others as the negative class was trained. The identification accuracy was calculated using 10 fold cross validation. Since it is challenging to distinguish the gait of different subjects under such complicated conditions. As shown in Fig. 20, when the number of moving subjects increases, the performance of Wi-PIGR declines significantly. It is difficult to restore blocked signals from lots of moving subjects. When multiple people co-exist in the target area, the signals reflected by multiple persons are more likely to cover each other, they will cause additive effect at the receiver side and weaken the received signal strength. In addition, the reflections off the nearest subject has much more power than distant reflections. Thus, Wi-PIGR's performance is severely influenced by the surrounding moving subjects.

6 LIMITATIONS AND DISCUSSION

The Wi-PIGR system shows the feasibility of commodity Wi-Fi based gait recognition system in which the path dependence is removed and the subject is identified through the gait regardless of the walking path. However, the current implementation of the Wi-PIGR has three limitations.

Multiple Moving Subjects. Although the Wi-PIGR system focuses on one subject gait recognition in typical indoor environment, some static subjects such as those sitting in the same room have no impact on the performance of the Wi-PIGR system. When there are multiple subjects moving at the same time, the receiving signals reflected from different subjects will be mixed together at the receiver. Thus, it is difficult to separate gait signals from the moving subjects' signals. There are some applications, such as [56], [57], the additive effects led by multiple moving subjects can be sequentially estimated based on the rationale of successive cancellation. However, the successive cancellation based techniques rely on separation of distance and azimuth of multiple subjects [56], which is not feasible for our scenario. In addition, CSI data collected from commodity Wi-Fi devices have rich hardware distortions in both power strength [29] and phase [58]. This makes the commodity Wi-Fi based activity recognition even more challenging [11], [17], [19],

[29], [32], [39]. How to separate the gait signals of each subject from the mixed signals will be a valuable future work.

Scalability Issue. Though Wi-PIGR has wide range of applications, such as smart home or smart building, it has some scalability issues, which impede its wide use. The first issue is the deployment. When the smart home is large, with current deployment, Wi-PIGR's performance gets worse. It is necessary to wisely deploy more Wi-Fi devices, balance the interference and sensing range and maintain satisfied identification accuracy. The second issue is the confliction between the intensive computation workload and limited computation resources. Substantial efforts, such as automated optimization have been devoted to make deep learning models usable for real-time inference on mobile platforms [59]. The optimization criterion is to reduce the number of layers, nodes, and weights without hurting the overall recognition performance too much. A number of effective techniques have been proposed, for example, aim at distillation (compression) or elimination (pruning) of parts of a model [59]. In order to improve its computation capability, Wi-PIGR system can be customized by the proposed techniques. Another approach to tackle this confliction issue is to upgrade Wi-PIGR system to client-server architecture. Specifically, the client (transmitter) is a router equipped with wireless NIC and the server (receiver) is both GPU and wireless NIC equipped computer. The server can receive and process the real-time data. Through this new architecture, the scalability issue will be mitigated. Therefore, the scalability issue will be a promising future work.

Application Scenarios. The core technology of Wi-PIGR is to identify the subject through the gait regardless of the walking path, which can be applied to various scenarios requiring human identification. However, since the multi-class learning methods have the disadvantage of the more classification and the lower accuracy [60], the recognition accuracy drops to 77.15% when the number of human subjects increases to 50. This restricts the application scenarios of the Wi-PIGR severely. This is also a common problem faced by other Wi-Fi based recognition systems [11], [29], [32], [39]. In order to maintain a good recognition accuracy, Wi-PIGR is only applicable for the scenarios with fewer people such as an apartment or an office room, not for scenarios with more people such as a railway station or a shopping mall. The restriction will motivate us to seek the feasible solutions in the future.

Hard Access Control. Recall from Section 1.1 that one of the application scenarios for Wi-PIGR is soft access control, such as the temperature, light and background music being adjusted based on who is at the smart home. Although Wi-PIGR can be successfully applied to such soft access control scenarios, due to the non-zero error rate, Wi-PIGR is not suitable for hard access control scenarios, such as unlocking a drawer in an office containing confidential documents, or allowing access to certain rooms in a bank.

7 CONCLUSION

In this paper, Wi-PIGR, the first path independent single subject gait recognition system using commodity Wi-

Fi devices, has been proposed. The key technical novelty of Wi-PIGR lies in the path dependence removal so that regardless of the walking path, the gait can be identified. In order to achieve this, an extra Wi-Fi receiver is introduced to form a 2D plane and obtain the both orthogonal direction movement information. And then, a series of signal processing techniques are proposed to eliminate the signal difference led by walking along the arbitrary paths and generate a high quality path independent spectrogram. This path independent spectrogram can represent the gait information along the whole path. Finally, the deep learning approach is tailored to the feature extraction in order to extract the comprehensive features from the path independent spectrogram. Extensive experiments under different types of indoor environment have been conducted to validate the superiority of Wi-PIGR from different aspects. Over the 50 subjects, the average recognition accuracy of 77.15% is achieved in Wi-PIGR. We are confident to believe that Wi-PIGR has a wide range of potential applications in human computer interactions.

REFERENCES

- [1] H. Chao, Y. He, J. Zhang, and J. Feng, "GaitSet: Regarding Gait as a Set for Cross-View Gait Recognition," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019.
- [2] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, "On Input/Output Architectures for Convolutional Neural Network-Based Cross-View Gait Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 9, pp. 2708–2719, 2017.
- [3] J. Lu, E. Zhang, and C. Jing, "Gait recognition using wavelet descriptors and independent component analysis," in *International Conference on Advances in Neural Networks*, pp. 232–237, 2006.
- [4] H. Ma and W.-H. Liao, "Human Gait Modeling and Analysis Using a Semi-Markov Process With Ground Reaction Forces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 6, pp. 597–607, 2017.
- [5] S. Pan, N. Wang, Y. Qian, I. Velibeyoglu, H. Y. Noh, and P. Zhang, "Indoor person identification through footstep induced structural vibration," in *International Workshop on Mobile Computing Systems and Applications*, pp. 81–86, 2015.
- [6] D. Gafurov, "A survey of biometric gait recognition: Approaches, security and challenges," in *Annual Norwegian Computer Science Conference*, 2007.
- [7] M. Maaaz and R. Mayrhofer, "Smartphone-based gait recognition: From authentication to imitation," *IEEE Transactions on Mobile Computing*, vol. 16, no. 11, pp. 3209–3221, 2017.
- [8] W. Xu, G. Lan, Q. Lin, S. Khalifa, M. Hassan, N. Bergmann, and W. Hu, "Keh-gait: Using kinetic energy harvesting for gait-based user authentication systems," *IEEE Transactions on Mobile Computing*, vol. 18, no. 1, pp. 139–152, 2018.
- [9] L. Zhang, C. Wang, M. Ma, and D. Zhang, "WiDIGR: Direction-Independent Gait Recognition System using Commercial Wi-Fi Devices," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1178–1191, 2020.
- [10] Y. Zeng, P. H. Pathak, and P. Mohapatra, "Wiwho: Wifi-based person identification in smart spaces," in *International Conference on Information Processing in Sensor Networks*, p. 4, 2016.
- [11] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 363–373, 2016.
- [12] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *International Conference on Distributed Computing in Sensor Systems*, pp. 75–82, 2016.
- [13] C. Lin, J. Hu, Y. Sun, F. Ma, L. Wang, and G. Wu, "Wiau: An accurate device-free authentication system with resnet," in *15th Annual IEEE International Conference on Sensing, Communication, and Networking, SECON 2018, Hong Kong, China, June 11-13, 2018*.
- [14] T. Xin, B. Guo, Z. Wang, M. Li, and Z. Yu, "Freesense: indoor human identification with wifi signals," in *Proceedings of the 2016 IEEE Global Communications*, pp. 1–7, 2016.
- [15] G. Courtine and M. Schieppati, "Human walking along a curved path. i. body trajectory, segment orientation and the effect of vision," *European Journal of Neuroscience*, vol. 18, no. 1, pp. 177–190, 2003.
- [16] M. P. Murray, "Gait as a total pattern of movement," *American Journal of Physical Medicine*, vol. 46, no. 1, pp. 290–333, 1967.
- [17] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "Widir: walking direction estimation using wireless signals," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 351–362, 2016.
- [18] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, "Widar: Decimeter-level passive tracking via velocity monitoring with commodity wi-fi," in *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2017.
- [19] K. Qian, C. Wu, Z. Zhou, Y. Zheng, Z. Yang, and Y. Liu, "Inferring motion direction using commodity wi-fi for interactive exergames," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1961–1972, 2017.
- [20] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, "Indotrack: Device-free indoor human tracking with commodity wi-fi," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2017.
- [21] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2.0: Passive human tracking with a single wi-fi link," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, 2018.
- [22] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [23] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [24] C. Y. Hsu, Y. Liu, Z. Kabelac, R. Hristov, D. Katabi, and C. Liu, "Extracting Gait Velocity and Stride Length from Surrounding Radio Signals," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017.
- [25] M. Zhao, T. Li, M. A. Alsheikh, Y. Tian, and D. Katabi, "Through-Wall Human Pose Estimation Using Radio Signals," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.
- [26] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-Invasive Detection of Moving and Stationary Human With WiFi," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2329–2342, 2015.
- [27] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2017.
- [28] L. Cheng and J. Wang, "Walls have no ears: A non-intrusive wifi-based user identification system for mobile devices," *IEEE/ACM Transactions on Networking*, vol. 27, no. 1, pp. 245–257, 2019.
- [29] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *International Conference on Mobile Computing and Networking*, (Paris, France), pp. 65–76, Association for Computing Machinery, Sep 2015.
- [30] H. Zhu, F. Xiao, L. Sun, R. Wang, and P. Yang, "R-ttwd: Robust device-free through-the-wall detection of moving human with wifi," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1090–1103, 2017.
- [31] D. Zhang, H. Wang, and D. Wu, "Toward centimeter-scale human activity sensing with wi-fi signals," *Computer*, vol. 50, no. 1, pp. 48–57, 2017.
- [32] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "Rt-fall: A real-time and contactless fall detection system with commodity wifi devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, 2017.
- [33] H. Wang, D. Zhang, J. Ma, Y. Wang, Y. Wang, D. Wu, T. Gu, and B. Xie, "Human respiration detection with commodity wifi devices: do user location and body orientation matter?," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 25–36, 2016.
- [34] X. Zheng, J. Wang, L. Shangguan, Z. Zhou, and Y. Liu, "Design and Implementation of a CSI-Based Ubiquitous Smoking Detection System," *IEEE/ACM Transactions on Networking*, vol. 25, no. 6, pp. 3781–3793, 2017.
- [35] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Zero-effort cross-domain gesture recognition with wi-fi,"

- in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, 2019.
- [36] K. Qian, C. Wu, Z. Yang, Y. Liu, F. He, and T. Xing, "Enabling contactless detection of moving humans with dynamic speeds using CSI," *ACM Transactions on Embedded Computing Systems*, vol. 17, no. 2, pp. 1–18, 2018.
- [37] X. Wang, C. Yang, and S. Mao, "PhaseBeat: Exploiting CSI Phase Data for Vital Sign Monitoring with Commodity WiFi Devices," in *IEEE International Conference on Distributed Computing Systems*, pp. 1230–1239, 2017.
- [38] X. Wang, C. Yang, and S. Mao, "TensorBeat: Tensor Decomposition for Monitoring Multi-Person Breathing Beats with Commodity WiFi," *Acm Transactions on Intelligent Systems and Technology*, vol. 9, no. 1, 2017.
- [39] F. Wang, W. Gong, and J. Liu, "On spatial diversity in wifi-based human activity recognition: A deep learning based approach," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2035–2047, 2019.
- [40] T. S. Rappaport, "Wireless communications: Principles and practice," *Prentice Hall PTR*, 1996.
- [41] "Enhancements for higher throughput," *IEEE Standard 802.11n*, 2009.
- [42] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, "Simulation and analysis of human micro-dopplers in through-wall environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 2015–2023, 2010.
- [43] P. V. Dorp and F. C. A. Groen, "Feature-based human motion parameter estimation with radar," *IET Radar, Sonar and Navigation*, vol. 2, no. 2, pp. 135–145, 2008.
- [44] F. J. Ordóñez and D. Roggen, "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, p. 115, 2016.
- [45] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," in *International Joint Conference on Artificial Intelligence*, pp. 1533–1540, 2016.
- [46] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning," *MIT Press*, 2016.
- [47] Y. Lecun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.
- [48] Z. Tian, W. Huang, H. Tong, P. He, and Y. Qiao, "Detecting text in natural image with connectionist text proposal network," in *European Conference on Computer Vision*, 2016.
- [49] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [50] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with lstm," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
- [51] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *International Conference on Neural Information Processing Systems*, vol. 1, pp. 1097–1105, 2012.
- [52] D. C. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," in *ACM SIGCOMM 2010 Conference*, vol. 40, (New Delhi, India), pp. 159–170, OCT 2010.
- [53] L. Zhang, Z. Wang, and L. Yang, "Commercial wi-fi based fall detection with environment influence mitigation," in *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking*, 2019.
- [54] R. W. Bohannon, "Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants.," *Age and Ageing*, vol. 26, no. 1, pp. 15–19, 1997.
- [55] S. L. Fritz and M. Lusardi, "Walking speed: the sixth vital sign," *Journal of Geriatric Physical Therapy*, vol. 32, no. 2, pp. 46–49, 2009.
- [56] C. Xu, B. Firner, R. S. Moore, Y. Zhang, W. Trappe, R. Howard, F. Zhang, and N. An, "Scpl: indoor device-free multi-subject counting and localization using radio signal strength.," in *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks*, 2013.
- [57] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via rf body reflections," in *Proceedings of the 12th USENIX Conference on Networked Systems Design and Implementation*, pp. 279–292, 2015.
- [58] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wi-fi," *IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1342–1355, 2019.
- [59] N. D. Lane, S. Bhattacharya, A. Mathur, P. Georgiev, C. Forlivesi, and F. Kawsar, "Squeezing deep learning into mobile and embedded devices," *IEEE Pervasive Computing*, vol. 16, no. 3, pp. 82–88, 2017.
- [60] J. Li, Y. Liu, R. Yin, H. Zhang, L. Ding, and W. Wang, "Multi-class learning: From theory to algorithm," in *Annual Conference on Neural Information Processing Systems*, pp. 1593–1602, 2018.



Lei Zhang received her Ph.D. degree in Computer Science from Auburn University (Auburn, AL, USA) in 2008. She worked as an assistant professor from 2008-2011 in the Computer Science Dept. at Frostburg State University (Frostburg, MD, USA); She is now an associate professor in College of Intelligence and Computing at Tianjin University (Tianjin, P.R. China), where she is a member of Tianjin Key Laboratory of Advanced Network Technology and Application. Her research interests include mobile computing,

computer networks, data mining. She is a member of the ACM and IEEE.



Cong Wang received the B.E. degree from Central South University, Changsha, China, in 2017, and M.Eng. degree from Tianjin University, Tianjin, China, in 2020. He is an assistant engineer in the 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, Hebei Province, China. His research interests include wireless communication and data mining.



Daqing Zhang (Fellow, IEEE) received the Ph.D. degree from the University of Rome La Sapienza, Rome, Italy, in 1996. He is with the Computer Science Department, Peking University, Beijing, China, and Telecom SudParis, IP Paris, Paris, France. He has published more than 200 technical papers in leading conferences and journals. His research interests include context-aware computing, urban computing, mobile computing, big data analytics, and pervasive elderly care. He is the winner of the Ten-years CoMoRea Impact Paper Award at IEEE PerCom 2013, the Honorable Mention Award at ACM UbiComp 2015 and 2016, the Best Paper award at IEEE UIC 2015 and 2012. He served as the general or program chair for more than 15 international conferences, giving keynote talks at more than 20 international conferences. He is an Associate Editor of the ACM Transactions on Intelligent Systems and Technology and IEEE Pervasive Computing.