

Article

WiPg: Contactless Action Recognition Using Ambient Wi-Fi Signals

Zhanjun Hao ^{1,2,*}, Juan Niu ¹, Xiaochao Dang ^{1,2} and Zhiqiang Qiao ¹

¹ College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China; 201671020215@nwnu.edu.cn (J.N.); dangxc@nwnu.edu.cn (X.D.); 2019221817@nwnu.edu.cn (Z.Q.)

² Gansu Province Internet of Things Engineering Research Center, Lanzhou 730070, China

* Correspondence: haozhj@nwnu.edu.cn

Abstract: Motion recognition has a wide range of applications at present. Recently, motion recognition by analyzing the channel state information (CSI) in Wi-Fi packets has been favored by more and more scholars. Because CSI collected in the wireless signal environment of human activity usually carries a large amount of human-related information, the motion-recognition model trained for a specific person usually does not work well in predicting another person's motion. To deal with the difference, we propose a personnel-independent action-recognition model called WiPg, which is built by convolutional neural network (CNN) and generative adversarial network (GAN). According to CSI data of 14 yoga movements of 10 experimenters with different body types, model training and testing were carried out, and the recognition results, independent of bod type, were obtained. The experimental results show that the average correct rate of WiPg can reach 92.7% for recognition of the 14 yoga poses, and WiPg realizes "cross-personnel" movement recognition with excellent recognition performance.

Keywords: device-free sensing; channel state information; human action standard recognition; personnel independence; generative adversarial network; principal component analysis



Citation: Hao, Z.; Niu, J.; Dang, X.; Qiao, Z. WiPg: Contactless Action Recognition Using Ambient Wi-Fi Signals. *Sensors* **2022**, *22*, 402. <https://doi.org/10.3390/s22010402>

Academic Editors: Haipeng Dai, Haisheng Tan, Ruiting Zhou, En Wang, Chi Lin, Ruipeng Gao and Yuben Qu

Received: 24 November 2021

Accepted: 31 December 2021

Published: 5 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the widespread use of wireless networks and Wi-Fi devices, Wi-Fi has become a great green sensing platform, playing a decisive role in the field of perception. Previous studies on motion recognition can be roughly divided into three categories: wearable device-based, vision-based, and Wi-Fi-based [1]. Behavior-recognition technology based on wearable devices has many limitations, such as the additional increase in user burden, inconvenience in daily life, limited power of sensors, etc. Some articles require users to carry special devices [2–4], such as radio frequency identification (RFID) tags [5,6], gloves, or mobile sensors, which can be inconvenient and costly in many real-life scenarios. For example, MRLIGHT [5] installs RFID readers on moving objects (people) and fixedly deploys RFID tags in the monitoring area, effectively and economically realizing indoor real-time position tracking. However, a large number of fixed and passive RFID tags need to be deployed, and they are prone to environmental impact in MRLIGHT, which is accompanied by a lot of inconvenience. Although vision-based behavior-recognition technology can also achieve high recognition accuracy, it is affected by problems such as light, angle, privacy, and security. The authors used cameras to detect human falling behavior, designed an inference method based on two-dimensional CNN to obtain the motion information between two consecutive images, and tested the public data set, achieving high recognition accuracy [7].

Recently, an increasing number of people have been using inexpensive and simple-to-use commercial Wi-Fi devices for behavior recognition [8–10], including human-activity recognition, gait recognition, gesture recognition, and other related applications. Because the channel state information (CSI) of the physical layer can provide more fine-grained and accurate behavior data, it describes both the amplitude attenuation and the phase shift of wireless signals, according to which various behaviors, from vital signals and basic actions

to complex activities, can be more effectively identified. Therefore, there are numerous research projects on behavior recognition based on CSI data.

For gait recognition, MAR [11] performs canonical polyadic (CP) decomposition on the signals reflected from different parts of the human body and recognizes gait according to the frequency and range of arms and legs (sometimes even the head). The uniqueness of CP decomposition is theoretically analyzed, and better accuracy of gait recognition is achieved. However, the limitation is that the anti-interference ability is weak. Multiple moving objects cause a multipath effect and other interference, which reduces the accuracy of MAR. Gate-ID [12] has proposed a method to infer people's walking direction through gait characteristics and the direction of the antenna and has designed a deep learning model based on an attention and feature-weighting algorithm to identify individuals through CSI data containing gait characteristics. However, this article stipulated the walking path and does not attempt other random walk paths, which means the method in this paper is not universal.

There are also many studies on gesture recognition based on Wi-Fi. For example, SignFi [13] is able to identify 276 sign gestures with a high recognition rate by inputting CSI data containing gesture features in Wi-Fi packets into a nine-layer convolutional neural network for recognition. However, it can only recognize a single word but not a sentence in sign language. This paper collects and evaluates CSI data of 50 gestures from 5 users, extracts different levels of gesture features through a one-dimensional parallel long- and short-memory full convolutional network model, and is robust across different gestures and different people [14]. This point is similar to the point that this paper wants to study, which is of reference significance to us. However, this paper does not consider the robustness of different environments.

In a typical study of indoor behavior recognition based on Wi-Fi CSI, Hao et al. proposed a complex human-motion recognition scheme based on CSI [15], which is combined with a restricted Boltzmann machine (RBM) and SoftMax regression correction RBM classification. This method can classify the motion data sets effectively. However, this method is only aimed at static movements, while XingYiQuan motions are dynamic movements, and there is no research on dynamic movements. Another typical human-activity recognition method based on deep learning uses attention-based bidirectional long- and short-term memory to learn representative features of activities and actions from CSI sequences, which can obtain the best recognition performance of all activities [16]. However, this method does not take into account the effects of environmental changes. Although the above two papers on human-activity recognition achieved their goals and a high recognition rate, they did not pay attention to the influence of different people and did not conduct relevant studies on different people, so they are not universally applicable to people.

Based on the above research, we consider that whether it is gesture recognition, gait recognition, or indoor activity recognition, in addition to trying various devices to improve the accuracy of identification, it is also necessary to consider the universal availability of identification methods. For example, it is worth considering whether the model trained according to the CSI data of specific experimenters can recognize the actions of other experimenters. Since the focus of our study is on human-independent motion recognition methods, we also studied the latest papers in this respect. For the sake of solving the problem of activity identification independent of people and environment, there are different methods for research. Jiang et al. [17] raised a series of questions. For example, the reflection of the wireless signal in space is affected by various environmental factors (such as glass, air, walls, etc.), resulting in a change in the reflected signal. There are also significant differences in the reflected signals received when detecting the activity of different user objects, such as adults and children with wide differences in height and weight. These problems result in activity-recognition devices that do not work well in the new environment. An activity-recognition method independent of humans and the environment is realized through Doppler translation and neural network model construction in this study. However, the recognition rate of other data sets is not very high. Ma et al. [18] built an activity-recognition model by combining the theories of the state

machine and reinforcement learning. They trained people-independent and location-independent recognition models by learning location and people's independent features from different angles of CSI and evaluated the models through multiple experiments and public data sets, with quite good results can be obtained by their work. However, it does not necessarily work in multiplayer scenarios. Widar3.0 [19] can extract domain-independent data features that only reflect the gesture itself from the original domain-dependent CSI signal by establishing a quantitative relationship model between complex gestures and CSI dynamics. Cross-domain gesture recognition is realized by the generalization technique, and the accuracy of gesture recognition based on CSI was improved to 92.4% in the cross-environment situation. The above three articles have reference significance for this paper.

As we all know, with improvement in living standards, people's work pressure is heavier and heavier. Because of work, many people suffer from a variety of "occupational diseases" at present, such as cervical spondylosis, periartthritis of the shoulder, etc. However, studies have shown that many of these "occupational diseases" can be alleviated or even cured by practicing related yoga poses, such as the half moon, which can relieve sciatica caused by sitting too long. At the same time, yoga can play a role in weight loss, shaping, regulating emotions, relieving pressure, and preventing cardiovascular and cerebrovascular diseases, as well as other aspects. However, to achieve the purpose of the exercise is the premise of yoga movements. In real life, in addition to yoga classes under the guidance of teachers, students learn by watching videos on their own, and in many cases, training results are counterproductive. Based on this background, we take CSI data of yoga actions as the data sample of action recognition in this study, which has highly practical significance. The yoga poses we adopt in this study are shown in Figure 1. The difference between the body types of different people will affect the performance of the recognition method, so the recognition method in this paper should be suitable for different people.



Figure 1. Yoga movements in WiPg.

This study solves the above problems by building a deep learning model that can effectively identify yoga movements. We obtained an action-recognition method independent of human, and a feedback update module is introduced into the model. The output results that are not within the threshold, which is set in advance and obtained experimentally, are fed back to the model for retraining. If the final output is obtained after repeated training, the action is judged to be the standard action. WiPg will provide an effective scientific guide for the user's family exercise. Similarly, Qi et al. [20] designed a multilayer recursive neural network to identify teleoperated robot gesture data obtained by multiple sensors, thus guiding human-machine collaboration tasks. The proposed method has a high recognition rate, but the data-acquisition system has limitations. Su et al. [21] modeled the dynamic motion primitive (DMP) during the operation of experienced surgeons by teaching by demonstration (TbD) technology to help surgical robots learn surgical skills and complete surgical tasks. The operational autonomy of the surgical robot is increased, and the work efficiency of the surgeon is improved, but this article only considers surgical skills, not operational strength. Qi et al. [22] combined deep learning and threshold methods to

monitor breathing and motion recognition, and compared with previous machine learning algorithms, they achieved a higher recognition rate and more robustness, although the ability to analyze breathing patterns is limited. These articles are based on the learning of real-life activities, and the learning results have a guiding role in related applications, the same intention as our article. In the last paper, the combination of deep learning and threshold value was used to solve the identification problem, which indicates that the idea proposed in WiPg is feasible and has certain reference value for us to use the deep learning model for identification.

Furthermore, “Wi” refers to Wi-Fi based, “P” refers to personnel-independent, and “g” refers to yoga in the word “WiPg”. Three contributions of WiPg are as follows:

1. Aiming at the problem that the diversity between the body types of various people affects the performance of the recognition method, the recognition method in WiPg is suitable for diverse people. Through the combination of GAN and CNN, high-level features independent of body size are extracted, and action recognition is carried out.
2. WiPg has a very good recognition accuracy. In order to achieve this goal, first of all, we need to collect a large number of real experimental data and create similar graphs of problems by using different numbers of samples. More samples provide more opportunities for learning algorithms to understand the underlying mapping of input to output, leading to better performance models. Secondly, in the process of model training, parameters are adjusted repeatedly until the best recognition result is obtained.
3. On account of the longer training time, people who practice yoga exercises cannot maintain the standard movements of the initial state from the beginning to the end of the whole training process. This can be judged by setting the threshold value in WiPg. If multiple feedback updates fail to output the result, it indicates that the action is not standard; and if it can output a classification result, it indicates that the action is standard.

2. Related Theory

2.1. Theory of CSI

CSI signals are acquired from ubiquitous Wi-Fi devices and reflect the attenuation of Wi-Fi signals in transit. The CSI signal of yoga action received by the receiver can be expressed as:

$$H(f_k) = \|H(f_k)\|e^{j\angle H(f_k)}, k \in [1, K] \quad (1)$$

where $H(f_k)$ is CSI with f_k as the center frequency representing the k th subcarrier of the CSI signal of yoga action; $\|H(f_k)\|$ is amplitude; and $\angle H(f_k)$ is phase. If $H_k = \|H(f_k)\|$, the yoga CSI sequence of N pairs of antennas is expressed as Equation (2):

$$H^n = [H_1^n, H_2^n, \dots, H_k^n], n \in [1, N] \quad (2)$$

where H_k^n stands for the k th subcarrier of CSI sequence between the n th pair of antennas. For the sake of identifying yoga movements in the detection area, we need to regularly collect continuous CSI data and take M sets of CSI sequences collected within a period of time as a group. The CSI sequence can be expressed as Equation (3):

$$\begin{bmatrix} H'_{11} & H'_{12} & \cdots & H'_{1k} \\ H'_{21} & H'_{22} & \cdots & H'_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ H'_{m1} & H'_{m2} & \cdots & H'_{mk} \end{bmatrix}, m(1, M), k(1, K) \quad (3)$$

where H'_{mk} stands for the set of the k th subcarrier at the m th time. CSI signals are affected by human activities in Wi-Fi-covered areas, as shown in Figure 2, because the human body may block the transmission path of Wi-Fi wireless signals and thus cause signal power attenuation. Otherwise, the human body can introduce more signal reflections and change the number of propagation paths. Therefore, the variance of CSI can reflect

human movement in the wireless environment. Therefore, CSI signals are used to reflect the differences of different yoga movements and predict movements in WiPg.

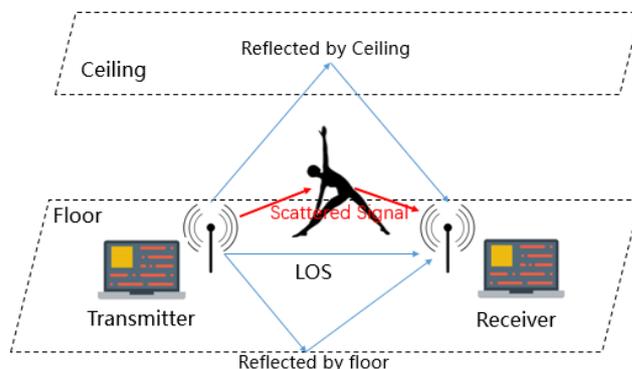


Figure 2. Wi-Fi signal propagation during human movement.

2.2. Methods of Data Preprocessing

2.2.1. An Overview of the Butterworth Filter

We selected the Butterworth filter to preprocess the collected CSI data related to yoga in this study. The filter is introduced in detail below. The amplitude square function of the Butterworth filter amplitude is shown in Equation (4).

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \left(\frac{\Omega}{\Omega_c}\right)^{2N}} \tag{4}$$

where N is the order of the filter, Ω is the cut-off frequency, and Ω_c is the cut-off frequency at -3dB [23]. It can be seen from the above formula that the n -order filter has $2N$ poles, and these poles are distributed in the Butterworth circle. When its poles are located in the left half plane of the S plane—that is, N poles in the left half plane are taken as the poles of the filter—the filter is stable. Before calculating the stable pole, we need to calculate the stopband attenuation through Equation (5) and calculate the order through Equation (6) at first. Then, we should obtain the stable pole through Equation (7).

$$A_s = -20 \log_{10} |H_a(j\Omega_s)| \tag{5}$$

$$N = \frac{1}{2} \frac{\log_{10} \left(\frac{10^{0.1A_s} - 1}{10^{0.1A_s} - 1} \right)}{\log_{10} \left(\frac{\Omega_s}{\Omega_p} \right)} \tag{6}$$

$$P_k = \begin{cases} \Omega_c e^{j\frac{2k+1}{2N}\pi}, N \text{ is even}, k = 0, 1, 2, \dots, 2N - 1 \\ \Omega_c e^{j\frac{2k}{2N}\pi}, N \text{ is odd}, k = 0, 1, 2, \dots, 2N - 1 \end{cases} \tag{7}$$

where A_s is the stopband attenuation and P_k is the stable pole. Then, the transfer function is calculated, as shown in Equation (8).

$$H_a(s) = \prod_{\text{Re}[P_k] < 0} \frac{\Omega_c}{s - P_k} \tag{8}$$

where $H_a(s)$ is the coefficient of the analog filter. Then, the variable, s , of $H_a(s)$ in Equation (8) is replaced by Equation (9) through a bilinear transformation from the analog domain to the digital domain, and the coefficient, $H(z)$, is obtained, as shown in Equation (9). Finally, the filtering result can be calculated through the differential equation.

$$s = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \tag{9}$$

2.2.2. An Overview of the PCA

As a basic dimension-reduction technique, principal component analysis (PCA) removes the redundancy of original data and maintains the inherent structure of original data after dimensionality reduction [24]. In order to remove redundancy, the original CSI data should be projected to the direction with the maximum variance because the greater the variance, the greater the difference between the data, so that the redundancy between the data will be smaller. The algorithm principle is described, as follows, with mathematical knowledge:

Assuming that there are n m -dimensional sample data sets, $D = \{x_1, x_2, \dots, x_n\}$, $x_i (1 \leq i \leq n) \in R^m$, and assuming that the characteristics of each dimension have been normalized, if the dimension of x_i is reduced from m to l and the data after dimension reduction are represented by Y , a mapping matrix, W , is required so that W satisfies the following formula.

$$Y = XW \quad (10)$$

In this formula, the reduced dimension is denoted by Y . Since X is an $n \times m$ matrix and y is an $n \times l$ matrix, W should be an $m \times l$ matrix. The variance of Y is expressed as follows:

$$\text{var}(Y) = \frac{1}{n-1} \text{trace}(Y^T Y) = \text{trace}\left(W^T \frac{1}{n-1} X^T X W\right) \quad (11)$$

$$\Sigma = \frac{1}{n-1} X^T X \quad (12)$$

where Σ is the covariance matrix of the n m -dimensional sample, X , before dimensionality reduction. In order to maximize the variance of Y , $\text{trace}(W^T \Sigma W)$ should be maximized. In order to reduce the correlation—that is, to remove redundant data— W needs to satisfy $W_i^T W_i = 1, i \in (1, 2, \dots, l)$, and then it can be converted into Equation (13) by the Lagrange multiplier method.

$$L(W, \lambda) = \text{trace}(W^T \Sigma W) - \sum_{i=1}^l \lambda_i (W_i^T W_i - 1) \quad (13)$$

where $\lambda_i (1 \leq i \leq l)$ is the Lagrange multiplier and W_i is column i of the matrix, W . Take the partial derivative of W_i and set the derivative to 0, and you get the following formula.

$$\Sigma W_i = \lambda_i W_i \quad (14)$$

$$\text{trace}(W^T \Sigma W) = \sum_{i=1}^l \lambda_i \quad (15)$$

As can be seen from the above equation, the optimized variance is equal to the sum of the characteristic roots of the covariance matrix of the original sample data. Therefore, the mapping matrix, W , can be composed of the eigenvectors corresponding to the first l largest characteristic root.

According to the principles of the two data pretreatment methods described in Section 2.2, the more important action feature data in this study and CSI data can be obtained, and the integrity of the CSI signal can be retained.

2.3. Overview of the GAN

A generative adversarial network (GAN) is composed of two important parts: discriminator network and generator network [25]. The goal of the discriminator network is to distinguish whether X comes from a real distribution or a generated model, which is actually a binary classifier. Conversely, the generator network needs to let the discrimination network distinguish the samples generated by itself into real samples. The basic structure of GAN is shown in Figure 3.

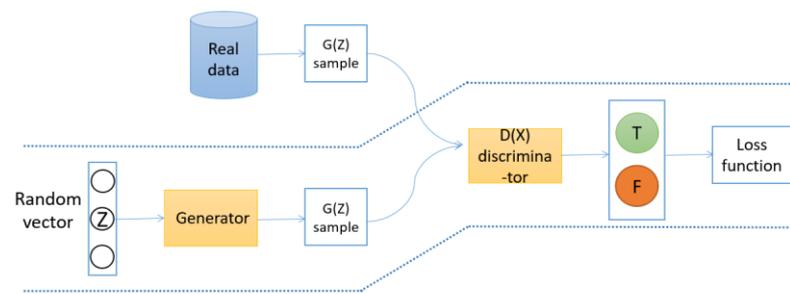


Figure 3. Overview of the GAN.

X represents the real data, Z represents the noise of the generator network, $G(z)$ means unreal data generated by the generator network, and $D(x)$ represents the probability that x belongs to the real sample distribution, where $D \in [0, 1]$. The optimization principle of GAN is simply that the generator network, G , generates $G(z)$ through continuous training and learning and makes the discriminator network, D , unable to distinguish the difference between $G(z)$ and x . D is to improve their discriminant ability through continuous training and learning, that is, to recognize that x and $G(z)$ are different.

The optimization principle of GAN can be described by a mathematical formula, as shown in Equation (16):

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (16)$$

Firstly, D is optimized, and then G is optimized. Optimization is about fixing one side and training the other. The purpose of D is to correctly distinguish x and $G(z)$. When optimizing the discriminator network, D , G should be given in advance, and efforts should be made to increase $D(x)$ and decrease $D(G(z))$. Therefore, the discriminator network is optimized to get $\max_D V(D, G)$.

Similarly, when optimizing the generator network G , D needs to be given in advance, and we just need to increase $D(G(z))$. The generation network is optimized to get $\min_G V(D, G)$.

According to the above antagonism principle of GAN, in order to remove the influence of the body size of personnel, the collected data can be trained through the mutual antagonism mode of the two networks in WiPg. The two networks are updated and iterated many times to maximize the interference ability of G to D and to minimize the discrimination ability of D to personnel tags. In the end, G makes it impossible for D to tell who the action is generated by, while the discriminator network, D , tries to tell n the action is generated by, so as to solve the personnel independence of action recognition.

3. Overview of WiPg

The overview of WiPg is shown in Figure 4. We selected recently popular yoga movements as the research object in this study, and the steps of the identification method are as follows:

1. We collected the CSI sign of 14 standard yoga poses from 10 experimenters in three real experimental environments. We eliminated the noise of the collected data by Butterworth filter and PCA, and the main characteristics of the yoga CSI data were retained. See Section 3.1 for details;
2. According to fast Fourier transform, energy changes to determine which period of data is yoga action data. See Section 3.2 for details;
3. CNN is integrated into GAN to build an action-recognition model through which to learn a piece of yoga CSI data that is personnel-independent, followed by action recognition. See Section 3.3 for details;
4. The last step is to estimate whether the yoga action is standard or not. The criterion is the threshold value obtained through a large number of training tests. If the output is within the threshold range, the action is standard, and the recognition result will be

output. On the contrary, if it is not in the range, feedback is updated, and the action is not standard if it is in a loop in which the result cannot be output all the time. See Section 3.3 for details.

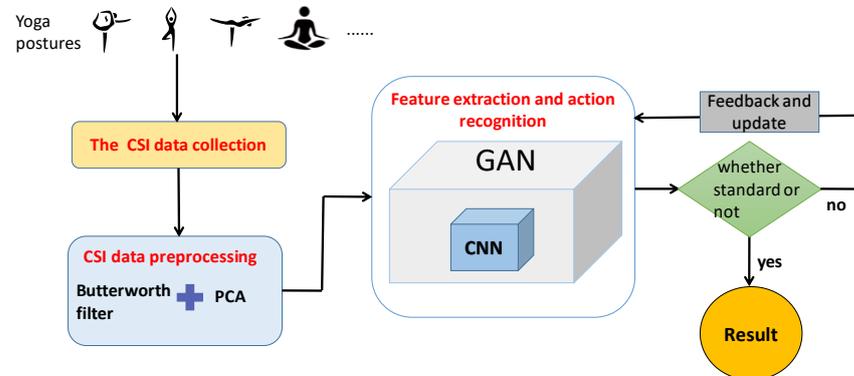


Figure 4. Overview of WiPg.

3.1. Data Preprocessing

On account of the impact of Wi-Fi equipment and the interference of environmental factors, there is a lot of interference and noise in the collected CSI data. To obtain high-quality data, we need to remove noise from the original CSI data. For this reason, to obtain more accurate motion features in CSI data, the Butterworth filter is firstly carried out on the original data to remove high-frequency noise. Secondly, because of the large amount of CSI data that belongs to the high-dimensional data, the PCA algorithm is applied to reduce the quantity of the original CSI data and retain the most crucial feature data. According to the principles of the two pretreatment techniques described in Section 2.2, the original and the preprocessed waveform are shown in Figure 5. In Figure 5a, there are three antennae from top to bottom, and each antenna contains 1900 CSI data. n curves with different colors represent the subcarrier amplitude variation curve on the n th packet. Figure 5a contains three antennas, and the total number of n is 3×1900 .

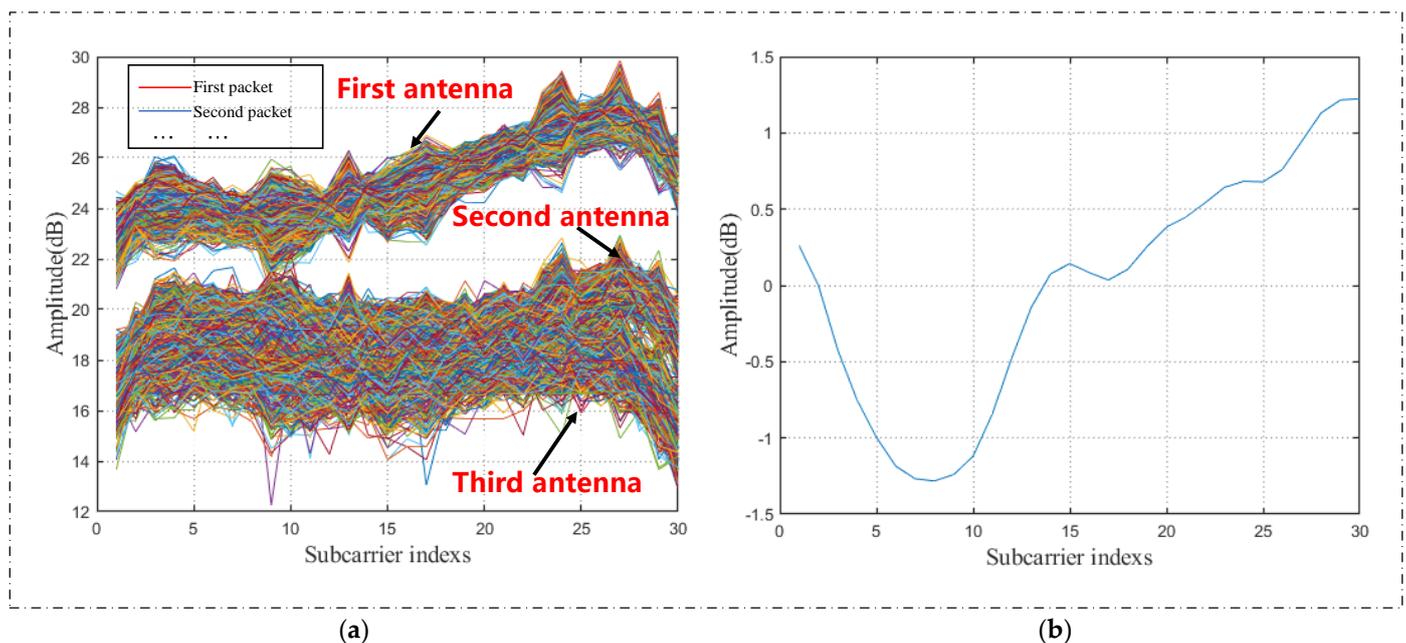


Figure 5. CSI waveform of the yoga pose. (a) Original CSI waveform; (b) CSI waveform after preprocessing.

It can be seen in Figure 5a that a large amount of noise is embraced in the original signal. According to the principles of the two methods described in Section 2.2, a Butterworth filter is first applied to remove high-frequency noise. We adopted PCA to lessen the dimensionality of superfluous data characteristics. At the same time, it retains the most substantial bits of the data characteristics and maintains the data integrity to the greatest degree. Figure 5b shows that the dimension of the CSI data, treatment by PCA, is reduced, and the prime features of yoga-movement CSI data are retained, which is fundamental to the subsequent feature extraction.

3.2. Motion Detection

Wi-Fi devices always exist in the real scene, so once people enter this environment, CSI data related to the human body are generated. Therefore, CSI data collected in the real scene include many other human-behavior data unrelated to yoga actions. To solve this problem, we need to judge these CSI data and find out which are the CSI data related to yoga movements required in WiPg. Different human behaviors cause different energy changes in the CSI frequency domain. Fast Fourier transform (FFT) can transform the CSI signal after eliminating the noise from the time domain to the frequency domain [26]. Then, by analyzing the CSI signal of yoga movements in the frequency domain, an energy threshold can be obtained to judge the occurrence and end of movements. Figure 6 is the energy change in the frequency domain caused by POSE TWO in Figure 1.

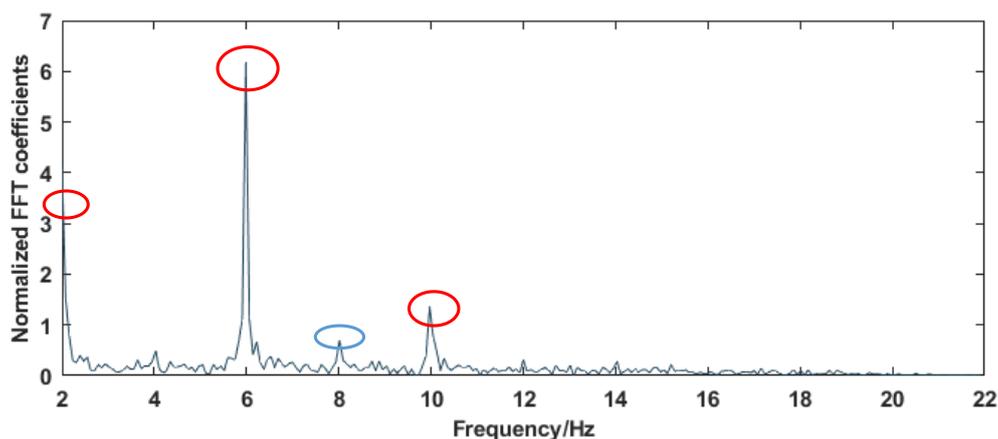


Figure 6. POSE TWO's FFT curve.

The energy distribution in the frequency domain after FFT transformation is shown in Figure 6. It can be seen that after normalized FFT conversion, there are two large peak signals, namely the 6 Hz signal and the 10 Hz signal. This segment of signal is the dynamic frequency variation range caused by POSE TWO. The frequency variation caused by DC signal is not recorded in the frequency variation range caused by the action. The FFT peak value of the blue circle in Figure 6 is lower than 1, which is considered to be the controllable signal caused by POSE TWO and is not recorded as the frequency range caused by the action. Therefore, we can use the signals between the two larger peaks to ensure that CSI can be collected effectively when the action occurs. This will be of great help to the subsequent feature extraction.

3.3. Action Recognition of Wi-Piga

A network GAN, including CNN, which is regarded as a feature-extraction algorithm, takes the central stage in WiPg. The personnel feature-discrimination module is defined to obtain a personnel-independent identification result, which is equal to the discriminator network, D . Figure 7 shows a concrete structure.

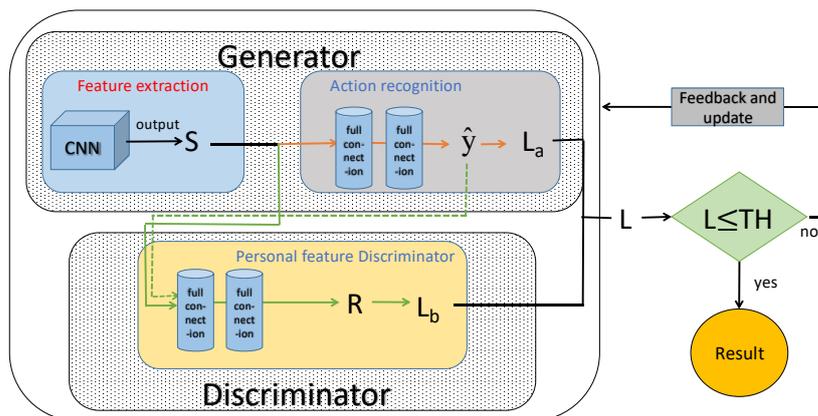


Figure 7. Action-recognition method in WiPg.

The training process of GAN in WiPg is as follows:

1. To generate parameters of network G and discriminator network, D , by initialization;
2. To extract n real data samples with personnel labels from the training set, G generates the personnel labels of the n data sample without personnel labels and forms new data;
3. To fix G and then to train D , we make D distinguish the true and false data samples as much as possible; in other words, to mark the person who generated each CSI data and to obtain a personnel-label distribution, R , through full connection;
4. To update D for k times, G is updated once. D maximizes the prediction performance of personnel labels, and G minimizes the prediction accuracy of discriminator model, D . After multiple update iterations, the people-related features are ideally eliminated, and the performance of person tag prediction is reduced.

The concrete implementation of the generation network and the discriminator network will be introduced in the following sections.

3.3.1. Generation Network, G

The Generation network, G , is divided into two parts. The first is the feature-extraction module, where we mainly adopt a convolutional neural network for yoga-movement data of CSI signal feature extraction. Due to CNN having been widely used for human-activity recognition feature extraction [27], this article also directly uses CNN to extract the features. Three layers of CNN were used in this article. Each layer of CNN contains a 2D kernel as a filter and then normalizes the mean value and variance of the data in each layer.

Finally, we add a ReLU and a maximum pooling layer to make the features between different dimensions have some numerical comparison. The result, S , after CNN convolution is shown in Equation (17).

$$S = \text{CNN}(X; \theta) \quad (17)$$

where S is the feature set, X is the input, and θ is the parameter set of CNN. The input data of the feature-extraction module is X , and all the data of X are marked. Each piece of data, X_i , has a corresponding feature-label vector, Id_i , related to personnel, where $Id_i \in D$, D represents all the Id_i sets, and each piece of data, X_i , corresponds to a real vector, y_i .

The second part is called the action-recognition module. Firstly, as shown in Equation (18), the full connection layer and an activation function are used to obtain a feature representation after learning. The softplus function is an activation function that introduces nonlinearity. F_i is then mapped into a new space, H_i , where $H_i \in R^C$, C is the amount of action data. The softmax layer is used to obtain the probability vector of the activity, as shown in Equation (19).

$$F_i = \text{Softplus}(W_S S_i + b_S) \quad (18)$$

$$\hat{y}_i = \text{Softmax}(H_i) \text{ and } H_i = W_f F_i + b_f \quad (19)$$

where W_s is the weight matrix of the feature set, S ; b_s is the deviation of the feature set, S ; W_f is the learned feature representing the weight matrix of F ; and b_f is the learned feature representing the deviation of F . The cross-entropy function is used to calculate the loss, L_a , between the predicted result, \hat{y} , and the actual result, y , as shown in Equation (20).

$$L_a = -\frac{1}{|X|} \sum_{i=1}^{|X|} \sum_{c=1}^C y_{iC} \log(\hat{y}_{iC}) \quad (20)$$

3.3.2. Discriminator Network, D

The discriminator network, D , is constituted by the personnel-feature-discrimination module. The goal of D is to identify and record the label of the person who produces the yoga action, so as to force the feature-extraction module to produce yoga action features unrelated to the personnel. We first connect the output, S , of feature-extraction module with prediction \hat{y} to get U ; that is, $U = S \oplus \hat{y}$, where \oplus is the joint operation, and S contains both human-related and human-independent features. To identify the common properties between different people, we connect the above two parts as the input of this module and then obtain a label distribution related to people by using two fully connected layers with an activation function, as shown in Equations (21) and (22).

$$T_i = \text{Softplus}(W_u U_i + b_u) \quad (21)$$

$$R_i = \text{Softmax}(W_t T_i + b_t) \quad (22)$$

where W_u is the weight matrix of U ; b_u is the deviation of U ; T_i represents the potential space; W_t is the weight matrix of T ; and b_t is the deviation of T , in order to enable the module to identify the personnel-related characteristics of the input data. In WiPg, the cross-entropy function is used to calculate the loss, L_b , between the predicted personnel-related label distribution, R , and the actual personnel-related label vector, Id , as shown in Equation (23).

$$L_b = -\frac{1}{|X|} \sum_{i=1}^{|X|} \sum_{j=1}^{|D|} Id_{ij} \log(R_{ij}) \quad (23)$$

where $|D|$ is the number of features associated with a human; Id_i is the actual personnel-tag vector. The goal of D is to minimize the loss function, L_b , so as to maximize the performance of human-label prediction, which is a contradiction for the final goal of WiPg, which is solved by Equation (24).

$$L = L_a - \beta L_b \quad (24)$$

where β is the weight value. Finally, through Equation (24), we can observe that G tries its best to deceive D by maximizing L_b while improving the performance of the action-recognition module by minimizing L_a so that we can learn all of the characteristics that are common to people, and recognition results can be obtained. By inputting the standard action data into the above two networks for training, the threshold, TH , of the loss value, L , is obtained. Then, we input CSI data of an action into the WiPg network for testing. First, a prediction result, \hat{y} , is obtained. Then, the loss value, L , is obtained. By comparing L with TH , if L is in TH interval, it means that it is standard, and the prediction result, \hat{y} , is output. On the contrary, if L is in TH interval, the network will be updated by feedback. If the predicted action type cannot be output all the time, it means that the action is not standard.

4. Experiment and Analysis

4.1. Experimental Scene

In the experiment of WiPg, we use two laptops equipped with an Intel Wi-Fi link (IWL) 5300 network card as the receiver and transmitter of CSI packets, separately. We use the Linux CSI tool to extract CSI data from the network card to obtain the original CSI data containing yoga

actions. The height of the equipment and the range of different yoga movements considered comfortable by different experimenters was counted and calculated. The average height of the equipment is 1 m, and the average distance of the equipment is 1.7 m.

However, in order to verify the performance of WiPg in different environments, CSI data in three experimental environments were collected for comparison during the experiment. Figure 8a is the plane structure of the yoga classroom. In this scene, the site is relatively empty, with fewer indoor items and less human interference, and the size of the yoga classroom is 10 m × 7 m. Figure 8b shows the plane of the laboratory. The size of the lab, which has more materials than the yoga classroom, is 7 m × 6 m. Figure 8c shows the plane of the student dormitory with the largest interference from the environmental multipath effect compared with the first two, with a size of 6 m × 4 m.

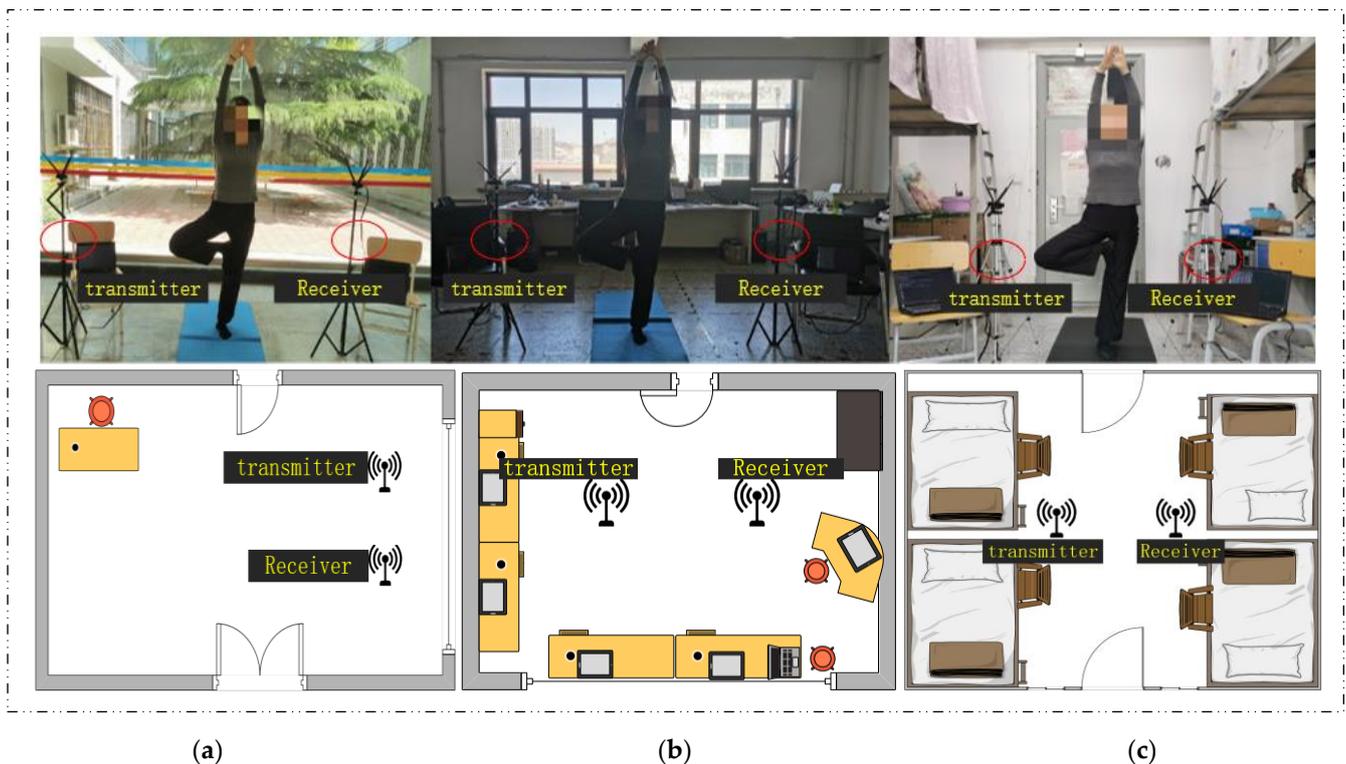


Figure 8. Experimental environment. (a) Yoga room; (b) laboratory; (c) dormitory.

During CSI data collection, subjects posed in standard yoga poses between two Wi-Fi devices under the supervision of a yoga teacher. In WiPg, five males and five females (aged from 12 to 27 years old) were invited to collect the CSI data of yoga movements. As shown in Figure 1, each experimenter was asked to do 14 sets of yoga poses. We set different packet-sending rates and packet numbers in advance. The experimenter kept still for a period of time after doing yoga movements, and each movement needed to be repeated 20 times. Table 1 shows the body characteristics of the experimental participants, including gender, height, weight, and BMI (body mass index).

It is worth noting that the static environment may change over several days, and there may be more or fewer differences between CSI data collected by different movements of the same experimenter. Therefore, we collected CSI samples of yoga movements within nearly a month, which can commendably refrain from overfitting and enhance recognition ability. In total, more than 2000 CSI data streams related to yoga poses were collected.

Table 1. Body characteristics of the experimenters.

Experimenter	Gender	Height (m)	Weight (kg)	BMI
A	Female	1.57	57	23.13
B	Female	1.62	54	20.57
C	Female	1.66	51	18.50
D	Female	1.75	52	16.97
E	Female	1.80	63	19.44
F	Male	1.69	56	20.84
G	Male	1.73	70	23.39
H	Male	1.77	80	25.53
M	Male	1.85	75	21.91
N	Male	1.89	70	19.59

4.2. Performance Evaluation

As is known to all, the common evaluation indexes of deep learning algorithms include confusion matrix, accuracy, ROC (receiver operating characteristic), AUC (area under curve), precision and recall, etc. The confusion matrix, which is also known as the error matrix, has each column representing the predicted value and each row representing the actual category [28]. F1 score is the harmonic mean of accuracy rate and recall rate. They are defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (25)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (26)$$

$$\text{F1 Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (27)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (28)$$

where *FN* is considered a negative sample, although it is, in fact, positive. *FP* is judged to be a positive sample, although it is, in fact, negative. In the same way, *TN* is judged to be a negative sample, which is correct. *TP* is judged to be a positive sample, which is correct.

4.3. Experimental Verification

4.3.1. Influence of Different People's Bodies on the Recognition Rate

In this study, we mainly discuss the influence of body size on CSI data and how to deal with this influence. Therefore, CSI data collected for people with different body types are tested in this section. Figure 9 shows the original CSI waveform was obtained by three experimenters with differing body types. There are three antennae from top to bottom, and each antenna contains 1900 CSI data. *n* curves with different colors represent the subcarrier amplitude variation curve on the *n*th packet. Figure 9 contains three antennae, and the total number of *n* is 3*1900. Figure 10 is the CSI waveform after pretreatment of the original waveform in Figure 9. CSI data collected by six experimenters with different body types were specially selected, and the experimental results obtained by training were input into the model in the three experimental environments, as respectively shown in Figure 11.

It can be seen from the original CSI waveform in Figure 9 and the preprocessed CSI waveform in Figure 10 that the signal changes are different when people of different body types do the same movement in the yoga classroom, which indicates that the different body type of experimenters has an impact on CSI data and therefore on the identification results of subsequent experiments. This is consistent with the actual situation.

However, as shown in Figure 11, when in the same environment, the experimental results obtained by different experimenters doing the same action are similar. At the same time, in combination with the contents of the above three figures, it can be shown that the

identification method of WiPg can ensure that the body shape of personnel does not affect the experiment; that is, it proves that WiPg is personnel-independent.

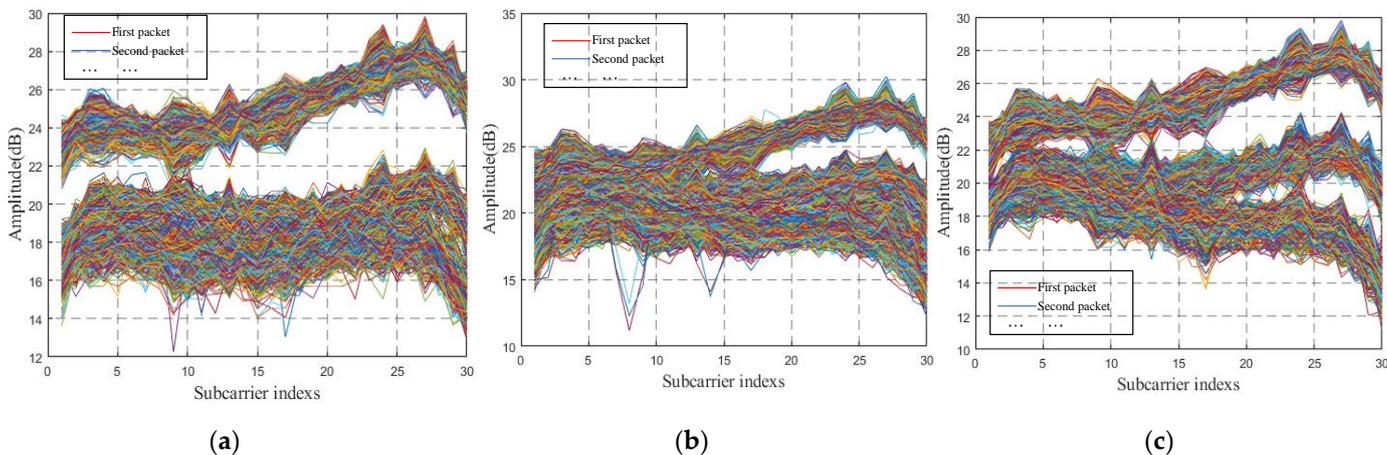


Figure 9. Original CSI signals of different people doing POSE THREE in Figure 1. (a) Experimenter A-POSE THREE; (b) experimenter B-POSE THREE; (c) experimenter C-POSE THREE.

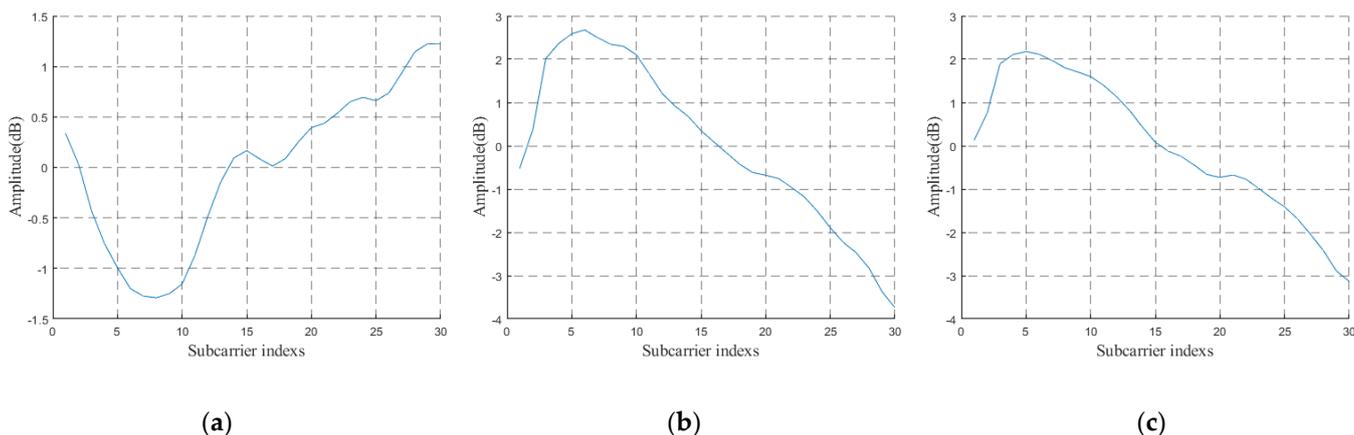


Figure 10. Preprocessed CSI signals of different people doing POSE THREE in Figure 1. (a) Experimenter A-POSE THREE; (b) experimenter B-POSE THREE; (c) experimenter C-POSE THREE.

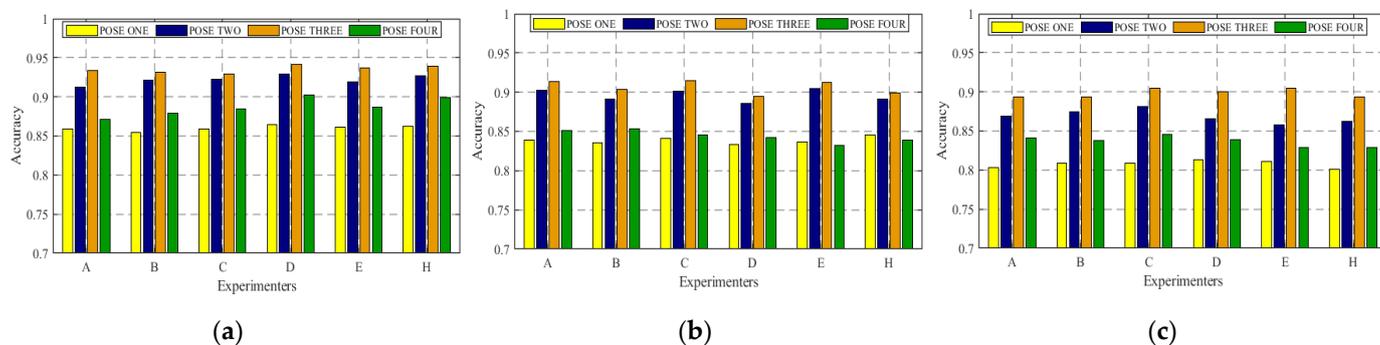


Figure 11. Recognition rate of different people practicing different actions in different experimental scenes. (a) Yoga classroom; (b) laboratory; (c) dormitory.

In the experiment in Figure 11, the first four yoga poses in Figure 1 were selected. It can be seen that the experimental results of POSE THREE are generally good, which is reasonable. Due to the large amplitude of POSE THREE, the characteristics of POSE THREE in the obtained CSI signal are more obvious, so the recognition results are relatively high.

4.3.2. Impact of Different Number of Packets, Number of Subcarriers, and Experiment Scenes on the Recognition Rate

Different numbers of CSI packets transmitted and selected subcarriers will also have a certain influence on the recognition results because the small number of subcarriers is usually not a good representation of all the information contained in the CSI data. Therefore, we set up the above two comparative experiments in three different experimental scenarios in particular, and the recognition results are shown in Figures 12 and 13.

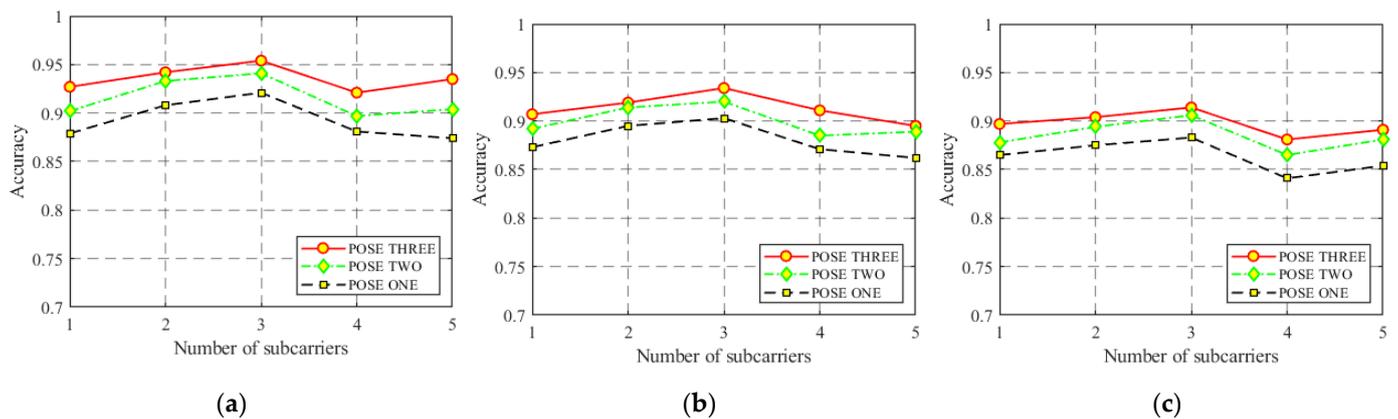


Figure 12. Accuracy of different numbers of subcarriers. (a) Yoga classroom; (b) laboratory; (c) dormitory.

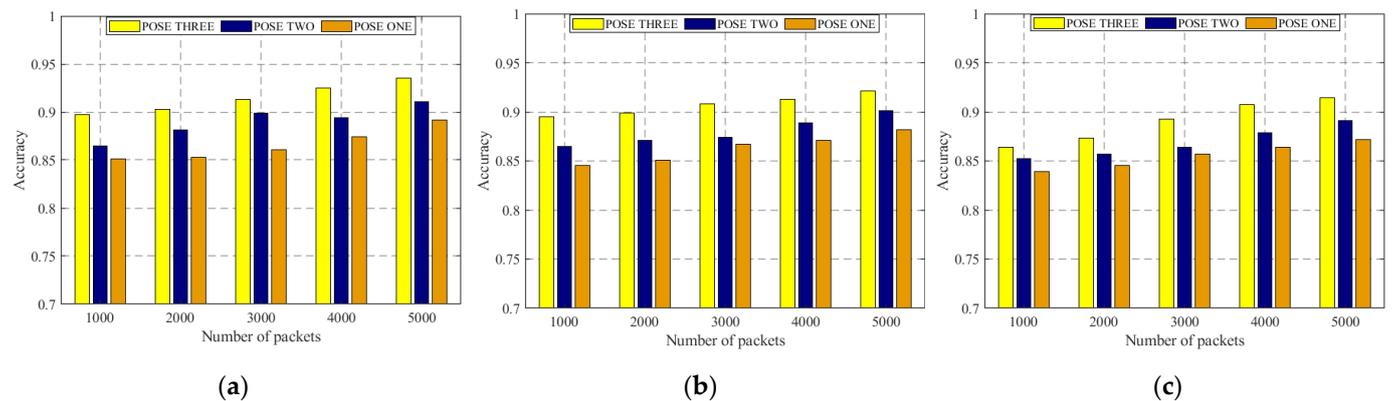


Figure 13. Accuracy of different CSI packet numbers. (a) Yoga classroom; (b) laboratory; (c) dormitory.

It can be clearly seen from Figures 12 and 13 that in three experiments, when the number of subcarriers is selected as three and the total number of transmitted packets is 5000, the model shows the best recognition effect. Therefore, the selection of appropriate experimental setting parameters (such as the above two) is very helpful to the model recognition rate. At the same time, it can also be seen that the recognition results in the yoga classroom are better than those in the other two environments, which is reasonable because the empty yoga classroom has the weakest multipath effect and the least interference of various environmental factors. Similarly, CSI signals collected in student dormitories are most affected by the multipath effect, so the recognition rate is the lowest compared with the other two.

4.3.3. Impact of the Number of Epochs and Batch Size

When a complete data set passes through the neural network once and returns once, this process is called an epoch. In other words, all training samples have a forward propagation and a backpropagation in the neural network. In the process of neural network training, it is not enough to iteratively train all the data once [29]. When all the data of an epoch is trained once, each data generate loss, which updates [the weight of the network.

The updated network has a better fit for the data, so the result curve gradually changes from under-fitting to fitting. However, when the number of epochs is too large, the resulting curve is too fitting. In layman's terms, the network now fully remembers the training data and is no longer portable.

Therefore, in this study, different numbers of epochs were set for model-training experiments so as to select the best number of epochs for neural network training in WiPg, as shown in Figure 14.

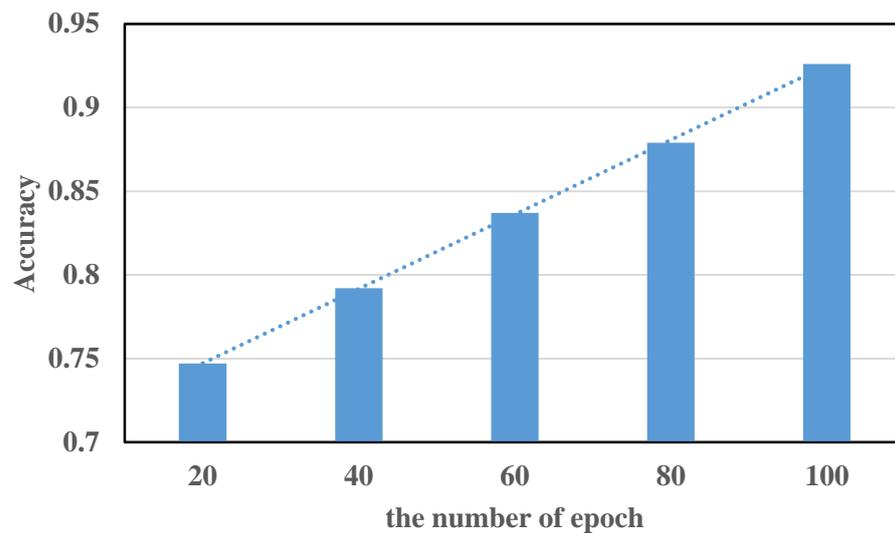


Figure 14. Different numbers of epochs.

However, when an epoch (that is, all training samples) is massive for the computer, it needs to be divided into multiple small batches for training [30]. Therefore, batch size is also an important parameter in machine learning. The appropriate batch-size range is mainly related to convergence speed and stochastic gradient noise. If the batch size is too small and there are many categories, the loss function may oscillate, which is not conducive to convergence, especially when your network is complex. If the batch size is too large, the gradient direction of different batches does not change, and it is easy to fall into a local minimum. The optimization and generalization of deep learning has problems. Therefore, an appropriate batch size is very important for the performance of the model. In order to discuss the impact of batch size on the accuracy of the method, we set different batch sizes and, in turn, set batch size to improve the model performance. The result is shown in Figure 15.

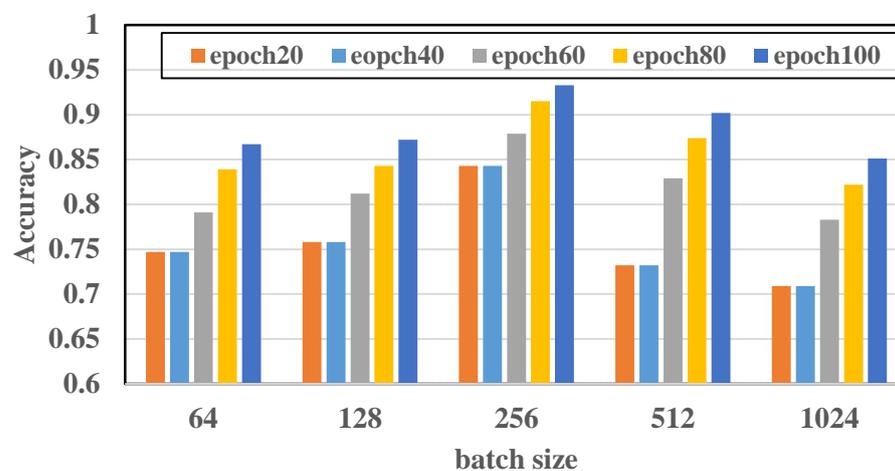


Figure 15. Different batch sizes.

It can be seen from Figures 14 and 15 that the appropriate number of epochs and correct batch sizes are of great help to the performance of the model during model training. Because the diversity of data sets determines the number of epochs, the stronger the diversity, the larger the epoch should be. Therefore, for the 14 categories of collected data sets in WiPg, the larger the value of epoch, to a certain extent, the better the performance of the model. However, an appropriate batch size needs to be selected so that the model is not under-fitted or over-fitted. In general, in this study, when the number of epochs is 100 and the batch size is 256, the performance of the model reaches the optimum.

4.3.4. Comprehensive Evaluation

The experimental parameters with the highest recognition rate can be obtained by comparing the above experiments. The specific identification method is described in Section 3.3. The comprehensive identification results of WiPg are shown in Figure 16.

ONE	0.92	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01
TWO	0.01	0.93	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
THREE	0.01	0.03	0.90	0.00	0.01	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.00
FOUR	0.00	0.00	0.00	0.93	0.01	0.00	0.02	0.01	0.01	0.00	0.00	0.01	0.01	0.00
FIVE	0.00	0.00	0.00	0.00	0.93	0.00	0.01	0.02	0.01	0.01	0.01	0.01	0.00	0.00
SIX	0.00	0.00	0.01	0.02	0.00	0.95	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
SEVEN	0.00	0.00	0.01	0.01	0.01	0.00	0.94	0.00	0.00	0.01	0.00	0.01	0.01	0.00
EIGHT	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.94	0.00	0.01	0.00	0.01	0.01	0.01
NINE	0.00	0.01	0.02	0.01	0.01	0.00	0.00	0.00	0.93	0.01	0.01	0.00	0.00	0.00
TEN	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.94	0.01	0.01	0.00	0.00
ELEVEN	0.02	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.93	0.00	0.00	0.00
TWELVE	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.01	0.02	0.00	0.92	0.00	0.01
THIRTEEN	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.92	0.00
FOURTEEN	0.02	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.02	0.01	0.91
	ONE	TWO	THREE	FOUR	FIVE	SIX	SEVEN	EIGHT	NINE	TEN	ELEVEN	TWELVE	THIRTEEN	FOURTEEN

Figure 16. Confusion matrix for synthesizing experimental results.

Each column of the confusion matrix represents the prediction category, and the total number of each column represents the number of data predicted to be in this category. As shown in the above confusion matrix, it can be seen that the comprehensive recognition result reaches more than 90%, which means that WiPg can effectively recognize and distinguish the 14 yoga poses.

(1) Comparative experiments of different classification methods

Moreover, in order to verify the overall performance of the WiPg method, we conducted comparative experiments with support vector machine (SVM) and random forest (RF), while the evaluation index was accuracy, F1 score, and receiver operating characteristic (ROC) curve. The experimental results of the three methods are shown in Table 2. In order to more intuitively show the superiority of the method of WiPg, an ROC curve was set, as shown in Figure 17.

As can be seen from above, the accuracy of the three methods can reach more than 83%, but compared with the other two methods, the accuracy of the WiPg is higher and relatively stable, and the overall recognition performance is better.

For the CSI data collected in the yoga classroom, the accuracy of all three methods reaches more than 89%. However, the WiPg method has the highest rate of accuracy, and through the value of the F1 index, AUC area, and ROC curve, we can see that the performance of the algorithm in this paper is higher and the robustness is better. Therefore, the method presented in WiPg is more suitable for the recognition of yoga movements than the other two methods.

Table 2. Comparison of different recognition methods.

Different Experimental Scenarios and Methods		Three Evaluation Indexes		
		Accuracy	F1 Score	AUC Area
Yoga Classroom	RF	0.899	0.849	0.897 4
	SVM	0.906	0.857	0.923 4
	the method of WiPg	0.933	0.907	0.942 1
Lab	RF	0.858	0.845	0.884 3
	SVM	0.891	0.853	0.904 5
	the method of WiPg	0.914	0.876	0.912 5
Dormitory	RF	0.834	0.841	0.857 6
	SVM	0.872	0.851	0.894 7
	the method of WiPg	0.899	0.867	0.909 5

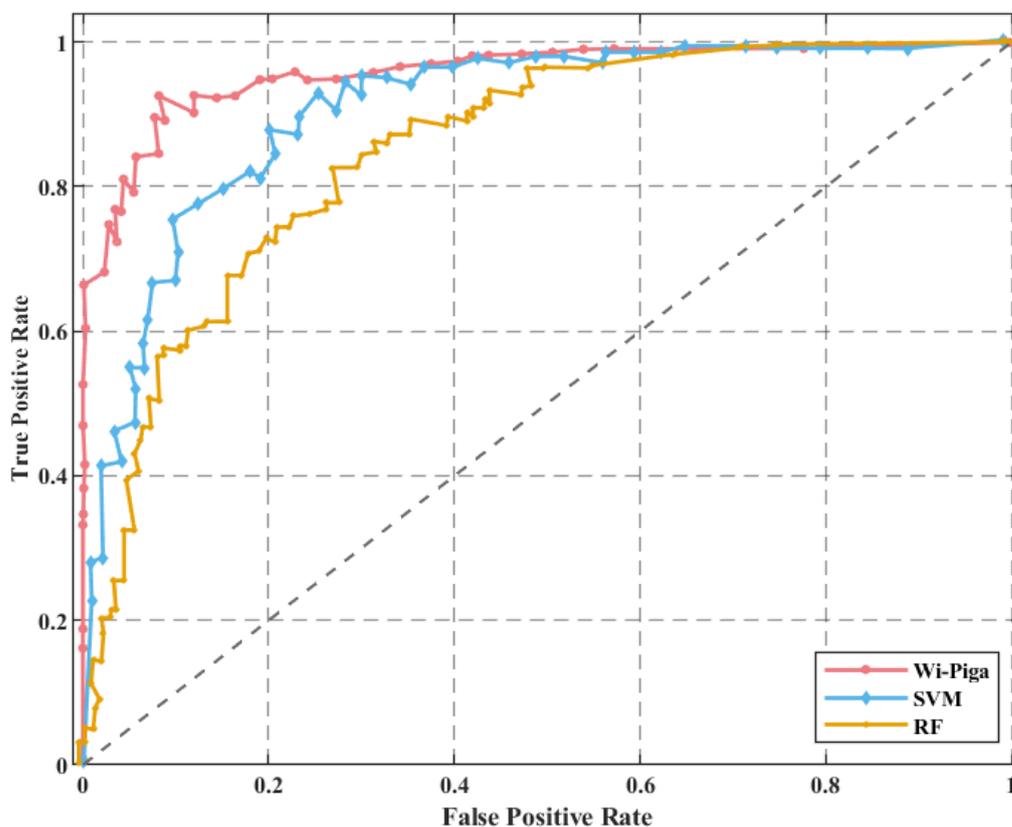


Figure 17. ROC curve comparison in the yoga classroom.

(2) Comparative experiments of different motion-recognition methods

At the same time, in order to further analyze the robustness of WiPg, we also compared WiPg with two other action-recognition methods (CSI-HC [15] and ABLSTM [16]). Similarly, we selected the three experimental environments mentioned above and 14 yoga action data sets collected by us for this comparative experiment. The evaluation results are shown in Figures 18 and 19.

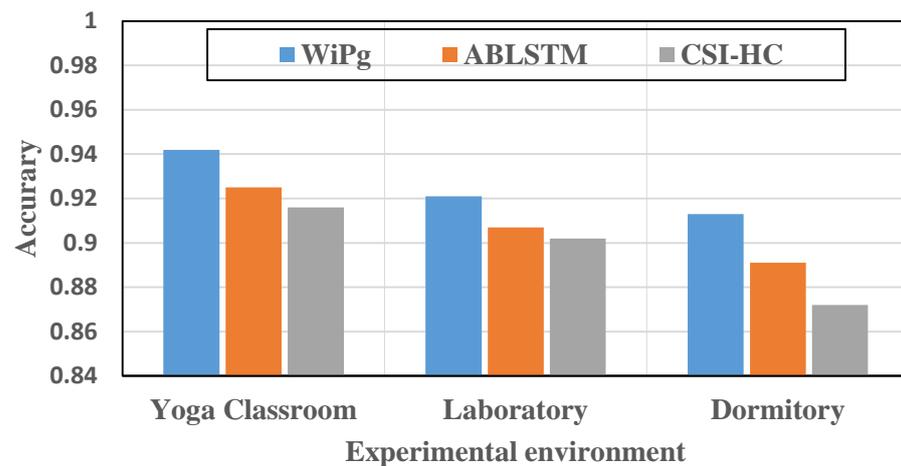


Figure 18. Robustness comparison.

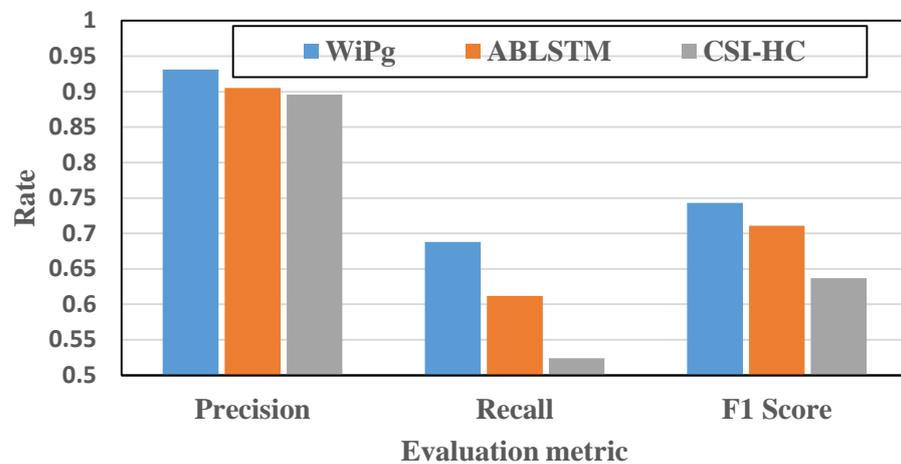


Figure 19. Comprehensive evaluation of classification performance.

Figure 18 verifies the applicability of the three methods in different environments. It can be seen that the identification accuracy of WiPg is the highest. Figure 19 shows the evaluation of the above three methods in precision, recall, and F1 score, and the results show that the effectiveness of the proposed method is superior to that of the other two methods. Through a series of comparative studies above, we effectively proved the stability and higher classification accuracy of the WiPg deep classification model.

5. Conclusions

In allusion to the matter of action recognition based on Wi-Fi, we have proposed a personnel-independent deep learning method. A mass of CSI data related to yoga poses was collected in three real environments, and the original CSI data were preprocessed by Butterworth filter and PCA. Next, CNN and GAN were combined to obtain an action-recognition model independent of human body size. The experiment proved that the recognition method proposed in WiPg is personnel-independent, and WiPg has a good performance for yoga action recognition. At the same time, WiPg uses thresholds from a large number of experiments to determine whether the yoga poses are standard or not.

This study will be beneficial to family fitness exercise guidance, and the following point can be further studied, as follows: on the basis of judging whether a yoga movement is standard, we found a method to identify the standard degree of yoga movement more accurately, which can be used in home yoga monitoring systems in the real life.

However, our paper has the following limitations. At present, the model in WiPg only studies yoga movements and does not consider the migration effect of the model on other

CSI data sets (such as gestures and falls), as well as the generalization ability of other data sets. Therefore, we believe that further research can be carried out on the following two points in the future: (1) On the basis of judging whether yoga poses are standardized, we can identify the degree of standardization of yoga poses in a more fine-grained way from the perspective of the differences between non-standard data and standard data; (2) The generalization ability of the model can be improved by training the network through data enhancement or using open-source data sets.

Author Contributions: Z.H., J.N. and X.D. conceived and designed the experiments; J.N. and Z.Q. performed the experiments; Z.H. and J.N. analyzed the data and wrote the original draft; Z.H. reviewed and edited the paper; X.D. revised the paper; Z.Q. helped in writing the related works. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Gansu Province Science and Technology Support Key R&D Program Project (20YF8GA048). 2019 Chinese Academy of Sciences “Light of the West” Talent Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: All participants in the experiment were in-formed and consented to the publication of the paper.

Data Availability Statement: Data is contained within the article.

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

References

1. Ma, Y.; Zhou, G.; Wang, S. Wi-Fi Sensing with Channel State Information: A Survey. *ACM Comput. Surv.* **2019**, *52*, 1–36. [[CrossRef](#)]
2. Yang, G.; Tan, W.; Jin, H.; Zhao, T.; Tu, L. Review wearable sensing system for gait recognition. *Clust. Comput.-J. Netw. Softw. Tools Appl.* **2019**, *22*, S3021–S3029. [[CrossRef](#)]
3. Yang, J.; Li, Q.; Wang, X.; Di, P.; Ding, H.; Bai, Y.; Dong, W.; Zhu, S. Smart wearable monitoring system based on multi-type sensors for motion recognition. *Smart Mater. Struct.* **2021**, *30*, 035017. [[CrossRef](#)]
4. Zhuang, W.; Chen, Y.; Su, J.; Wang, B.; Gao, C. Design of human activity recognition algorithms based on a single wearable IMU sensor. *Int. J. Sens. Netw.* **2019**, *33*, 193–206. [[CrossRef](#)]
5. Ma, Q.; Li, X.; Li, G.; Ning, B.; Bai, M.; Wang, X. MRLIHT: Mobile RFID-Based Localization for Indoor Human Tracking. *Sensors* **2020**, *20*, 1711. [[CrossRef](#)] [[PubMed](#)]
6. Zhao, J.; Zhou, J.; Yao, Y.; Li, D.; Gao, L. RF-Motion: A Device-Free RF-Based Human Motion Recognition System. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 1–9. [[CrossRef](#)]
7. Espinosa, R.; Ponce, H.; Gutierrez, S.; Martinez-Villasenor, L.; Brieva, J.; Moya-Albor, E. A vision-based approach for fall detection using multiple cameras and convolutional neural networks: A case study using the UP-Fall detection dataset. *Comput. Biol. Med.* **2019**, *115*, 103520. [[CrossRef](#)] [[PubMed](#)]
8. Wang, Z.; Jiang, K.; Hou, Y.; Dou, W.; Zhang, C.; Huang, Z.; Guo, Y. A Survey on Human Behavior Recognition Using Channel State Information. *IEEE Access* **2019**, *7*, 155986–156024. [[CrossRef](#)]
9. Zhao, J.; Liu, L.; Wei, Z.; Zhang, C.; Wang, W.; Fan, Y. R-DEHM: CSI-Based Robust Duration Estimation of Human Motion with Wi-Fi. *Sensors* **2019**, *19*, 1421. [[CrossRef](#)]
10. Wang, F.; Feng, J.; Zhao, Y.; Zhang, X.; Zhang, S.; Han, J. Activity Recognition and Indoor Localization with Wi-Fi Fingerprints. *IEEE Access* **2019**, *7*, 80058–80068. [[CrossRef](#)]
11. Fei, H.; Xiao, F.; Han, J.S.; Huang, H.; Sun, L. Multi-Variations Activity Based Gaits Recognition Using Commodity Wi-Fi. *IEEE Trans. Veh. Technol.* **2020**, *69*, 2263–2273. [[CrossRef](#)]
12. Zhang, J.; Wei, B.; Wu, F.; Dong, L.; Hu, W.; Kanhere, S.; Luo, C.; Yu, S.; Cheng, J. Gate-ID: Wi-Fi-Based Human Identification Irrespective of Walking Directions in Smart Home. *IEEE Internet Things J.* **2021**, *8*, 7610–7624. [[CrossRef](#)]
13. Ma, Y.; Zhou, G.; Wang, S.; Zhao, H.; Jung, W. SignFi: Sign Language Recognition Using Wi-Fi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–21. [[CrossRef](#)]
14. Tang, Z.; Liu, Q.; Wu, M.; Chen, W.; Huang, J. Wi-Fi CSI Gesture Recognition Based on Parallel LSTM-FCN Deep Space-Time Neural Network. *China Commun.* **2021**, *18*, 205–215. [[CrossRef](#)]
15. Hao, Z.; Duan, Y.; Dang, X.; Zhang, T. CSI-HC: A WiFi-Based Indoor Complex Human Motion Recognition Method. *Mob. Inf. Syst.* **2020**, *2020*, 3185416. [[CrossRef](#)]
16. Chen, Z.; Zhang, L.; Jiang, C.; Cao, Z.; Cui, W. Wi-Fi CSI Based Passive Human Activity Recognition Using Attention Based BLSTM. *IEEE Trans. Mob. Comput.* **2019**, *18*, 2714–2724. [[CrossRef](#)]
17. Jiang, W.; Miao, C.; Ma, F.; Yao, S.; Wang, Y.; Yuan, Y. Towards Environment Independent Device Free Human Activity Recognition. In Proceedings of the MobiCom '18: Proceedings of the 24th Annual International Conference on Mobile Computing and Networking, New Delhi, India, 29 October–2 November 2018; pp. 289–304.

18. Ma, Y.; Arshad, S.; Muniraju, S.; Torkildson, E.; Rantala, E.; Doppler, K.; Zhou, G. Location- and Person-Independent Activity Recognition with Wi-Fi. *Deep. Neural Netw. Reinf. Learn.* **2021**, *2*, 1–25. [[CrossRef](#)]
19. Zheng, Y.; Zhang, Y.; Qian, K.; Zhang, G.; Yang, Z. Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi. In Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '19), Seoul, Korea, 17–21 June 2019; pp. 313–325.
20. Qi, W.; Ovrur, S.; Li, Z.; Marzullo, A.; Song, R. Multi-Sensor Guided Hand Gesture Recognition for a Teleoperated Robot Using a Recurrent Neural Network. *IEEE Robot. Autom. Lett.* **2021**, *6*, 6039–6045. [[CrossRef](#)]
21. Su, H.; Mariani, A.; Ovrur, S.; Menciassi, A.; Ferrigno, G.; De Momi, E. Toward Teaching by Demonstration for Robot-Assisted Minimally Invasive Surgery. *IEEE Trans. Autom. Sci. Eng.* **2021**, *18*, 484–494. [[CrossRef](#)]
22. Qi, W.; Aliverti, A. A Multimodal Wearable System for Continuous and Real-Time Breathing Pattern Monitoring During Daily Activity. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2199–2207. [[CrossRef](#)] [[PubMed](#)]
23. Wang, F.; Li, H.; Yu, N. Investigation on impact of substrate on low-pass filter based on coaxial TSV. *IEICE Electron. Express* **2019**, *16*, 1–6. [[CrossRef](#)]
24. Wang, F.; Gong, W.; Liu, J. On spatial diversity in wifi-based human activity recognition: A deep learning-based approach. *IEEE Internet Things J.* **2019**, *6*, 2035–2047. [[CrossRef](#)]
25. Xiong, W.; Feng, C.; Xiong, Z.; Wang, J.; Liu, M.; Zeng, C. Improved pedestrian reidentification based on CNN. *Comput. Eng. Sci.* **2019**, *41*, 457–463.
26. Qiao, Y.; Li, J.; He, B.; Li, W.; Xin, T. A Novel Signal Detection Scheme Based on Adaptive Ensemble Deep Learning Algorithm in SC-FDE Systems. *IEEE Access* **2020**, *8*, 123514–123523. [[CrossRef](#)]
27. Jia, L.; Gu, Y.; Cheng, K.; Yan, H.; Ren, F. BeAware: Convolutional neural network (CNN) based user behavior understanding through Wi-Fi channel state information. *Neurocomputing* **2020**, *397*, 457–463. [[CrossRef](#)]
28. He, D.; Xie, C. Semantic image segmentation algorithm in a deep learning computer network. *Multimed. Syst.* **2020**. [[CrossRef](#)]
29. Lee, K.; Choi, C.; Shin, D.; Kim, H. Prediction of Heavy Rain Damage Using Deep Learning. *Water* **2020**, *12*, 1942. [[CrossRef](#)]
30. Koliouisis, A.; Watcharapichat, P.; Weidlich, M.; Mai, L.; Costa, P. CROSSBOW: Scaling Deep Learning with Small Batch Sizes on Multi-GPU Servers. *arXiv* **2019**, arXiv:1901.02244. [[CrossRef](#)]