RESEARCH

Open Access

CSI-based human behavior segmentation and recognition using commodity Wi-Fi



Xiaolong Yang^{1,2}, Jinglong Cheng^{1,2*}, Xinxing Tang^{1,2} and Liangbo Xie^{1,2}

*Correspondence: 2635169071@qq.com

 ¹ School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China
 ² Chongqing Key Laboratory of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing, China

Abstract

In recent years, the behavior recognition technology based on Wi-Fi devices has been favored by many researchers. Existing Wi-Fi-based human behavior recognition technology mainly uses classification algorithms to construct classification models, which has problems such as inaccurate behavior segmentation, failure to extract deep-level features from the original data and design classification models matching the proposed features in the process of behavior recognition. In order to solve above problems, this paper proposes a window variance comparison method by combining the adaptive thresholds calculated to achieve effective segmentation of multiple discontinuous human behaviors, then uses the short-time Fourier algorithm to extract timefrequency features of individual behavior, and extracts the graph structure data from the autocorrelation matrix of time-frequency features and the features themselves. A graph neural network is built for behavior recognition. The experimental results show that the segmentation accuracy of the behavior segmentation method in the two scenes is 0.964 and 0.993, which is better than the existing threshold-based behavior segmentation methods. In addition, this paper extracts graph structure data by spectral energy change method and builds behavior recognition model by using graph neural network, and the recognition accuracy is significantly improved compared with the traditional classification algorithm.

Keywords: Behavior recognition, Behavior segmentation, Channel state information, Graph neural network

1 Introduction

1.1 Motivation

The insensitivity of wireless signals to the propagation medium and the widespread deployment of wireless fidelity (Wi-Fi) devices have led to an increasing interest in behavior recognition techniques based on Wi-Fi devices. Behavior recognition techniques are based on analyzing the perturbation changes of human behavior on the wireless channel, using these changes to extract behavioral features and then constructing a behavioral classifier thus realizing human behavior recognition [1–4].

The realization of human behavior recognition based on commercial Wi-Fi devices mainly relies on received signal strength indication (RSSI) or channel state information (CSI). RSSI describes average power of received signal, which can be easily measured



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by/4.0/.

from most wireless devices. The low sensing sensitivity of RSSI and its susceptibility to noise do not allow high-accuracy sensing and identification of fine-grained behaviors. Thanks to the widespread use of orthogonal frequency division multiplexing (OFDM) technology, Halperin et al. [5] developed CSI-tools, a channel information measurement tool based on Intel 5300 NICs. The measured CSI can portray the variation of wireless signals in space. CSI is highly sensitive to changes in the signal propagation channel and has fine-grained characteristics. Therefore, a large number of researchers have started to focus on using CSI for contactless human behavior recognition, such as Wi-Hear [6], E-eyes [7], and RT-Fall [8].

For CSI-based human behavior recognition system, excluding CSI data acquisition and data pre-processing, the rest of process can be generally divided into three stages: segmentation of behavior data segments, behavior-related feature extraction and behavior category determination. The behavior segmentation stage mainly detects the start and end time points of behaviors from the continuous CSI data stream. The accuracy of segmentation of CSI data segments containing individual behavior as the input to the behavior classification model will directly affect the accuracy of subsequent behavior recognition. A large number of researchers have proposed threshold-based segmentation algorithms to implement behavior segmentation. In the feature extraction stage, features are generally extracted from the CSI signal so that the data can represent the behavior to the maximum extent and the behavior type can be discriminated by features. The features are generally time domain features or frequency domain features of the signal. In the behavior recognition stage, a behavior classification model is constructed using machine learning or neural network algorithms, and then, the trained model is used to determine the behavior.

We build a human behavior recognition system using low-cost and widely deployed Wi-Fi devices and realize behavior recognition by collecting CSI. Based on the built human behavior recognition system, we design a novel behavior segmentation method, which solves the problem that existing behavior recognition methods cannot achieve accurate segmentation of multiple behaviors and lays the foundation for subsequent behavior recognition. In addition, unlike traditional behavior recognition methods, we extract graph structure characteristics from individual behavior, and that can characterize the behavior to the greatest extent, and then builds a graph neural network to realize behavior classification.

The main work of this paper is as follows.

- 1. In this paper, a behavior segmentation method is proposed. The method can achieve high-accuracy segmentation of human behavior and solve the problems such as inaccurate behavior segmentation of existing human behavior segmentation algorithms.
- 2. We propose a new behavior recognition method. Firstly, the short-time Fourier transform is used to extract time-frequency domain features from individual behavior, and then the data containing graph structure information are innovatively extracted from human behavior features, and the graph neural network is used to realize behavior recognition.
- 3. In this paper, the behavioral segmentation method and the behavioral recognition method are validated and analyzed in the exhibition room and the laboratory,

respectively. The experimental results show that the behavior segmentation accuracy of this segmentation method is higher than the existing human behavior segmentation algorithm, and the recognition accuracy of this behavior recognition method is higher than the traditional behavior recognition method.

The rest of this paper is organized as follows. In Sect. 1.2, we briefly introduce the related work. In Sect. 2.1, we shortly introduce our activity recognition system. Section 2.2 describes the implementation process of the behavior segmentation method in detail. Section 2.3 describes the graph data acquisition method and the construction of graph neural network. Our experimental scenarios as well as experimental results are illustrated in Sect. 3. Section 4 provides the conclusion.

1.2 Related work

1.2.1 Behavior segmentation

In CSI-based human behavior recognition, it is extremely important to extract the data segments from the original data stream, which can improve the behavior recognition accuracy. Therefore, some researchers have proposed behavior segmentation algorithms.

CSI is sensitive to changes in the surrounding environmental information, and the variance of CSI in the presence of activity will be much larger than that in the absence of activity. Wu et al. [9] proposed a normalized variance sliding window algorithm to determine the start and end time points of individual behaviors. Wang et al. [10] used the variance of CSI during the occurrence of activity. Bu et al. [11] divided CSI into segments and calculated the mean and variance of each segment separately, and when a variance was larger than the threshold, the behavior was determined to exist in the data segment.

In the above method, the effectiveness of threshold is reduced when the environmental information changes. To reduce the impact of environment on the performance of behavior segmentation, some researchers have used adaptive thresholding to implement behavior segmentation. For example, Feng et al. [12] divided CSI into sub-data segments and calculated the variance in each segment, arranged them in ascending order according to the magnitude of the variance, and calculated the difference between the larger half and the smaller half of the variance. If this difference is larger than the adaptive threshold, it indicates the presence of behavior in that sub-data segment.

In addition to the threshold-based behavior segmentation algorithm, some researchers have proposed more 'novel' behavior segmentation methods. For example, in the RT-fall system [8], researchers first automatically determine the fall behavior by processing the variance of the CSI phase-difference signal. Lin et al. [13] proposed the automatic segment algorithm by combining the ideas of aggregation and mapping. Xiao et al. [14] divided the waveform of a complete behavior into four states of the same length, and then generated a behavior state model by convolutional neural network to achieve human behavior segmentation.

The above studies utilize threshold-based or other methods to achieve behavior segmentation and achieve the expected performance, but they all inevitably have certain problems. For the optimal threshold-based behavior segmentation methods, it is difficult to find the optimal threshold for mixed behaviors (multiple types of behaviors contained in a single data) because the CSI magnitudes between different behaviors can be very different. At the same time, there is a certain subjectivity in the optimal threshold set manually. To a certain extent, the behavior segmentation method based on adaptive thresholding overcomes the drawbacks of the behavior segmentation method based on optimal thresholding. Among some other behavior segmentation methods, the behavior segmentation method based on deep learning needs to collect behavior state data, build behavior state models, which can greatly increase the preliminary workload. In contrast, the behavior segmentation algorithm proposed in this paper uses the kernel density method to make the threshold adaptive, which is able to accurately segment multiple mixed behaviors. At the same time, the overall time complexity of the method is low, and it does not require the preliminary work as in the DeepSeg system, which lays the foundation for the subsequent implementation of real-time behavior segmentation.

1.2.2 Behavior recognition

Existing human behavior recognition techniques can be divided into two types: model-based behavior recognition techniques and pattern-based behavior recognition techniques. The former intends to design a physical model that is capable of linking environmental changes to signal changes [15]. Applying model-based approaches in some cases is challenging and it is difficult to develop a general mathematical model describing CSI in different scenarios. Wang et al. [16] proposed a CSI speed model and a CSI activity model for human behavior recognition.

Compared to model-based behavior recognition systems, pattern-based behavior recognition systems require a large amount of data as well as several iterations to achieve the best classification by training the network parameters. Chowdhury et al. [17] extracted correlation features from CSI subcarriers and input these features into a behavior classification model to achieve behavior recognition. Wu et al. [9] proposed a device-free behavior recognition system based on back propagation neural network to recognize seven different behaviors with an average recognition accuracy of 94.46%.Wang et al. [10] constructed a CSI behavior model for different motion states using hidden Markov model (HMM), which can recognize eight behaviors. li et al. [18] used the arrival angle difference as a reference signal, and then used a bidirectional long and short term memory network to classify behaviors. Chen et al. [19] proposed a new behavior classification algorithm by combining the attention mechanism, which uses the attention mechanism to assign different weights to all features, achieved a high accuracy of behavior recognition.

In the above study, existing human behavior recognition systems generally perform simple processing of raw data as features and then input them into the classification model for behavior recognition. Although these methods can achieve behavior recognition, they have problems such as low robustness and easy confusion of similar behaviors. The reasons for this are the failure to extract deep-level features from the raw data and the failure to design a classification model that matches the proposed features. In order to solve the above problems, we extracts the data with graph structure information from the original data and then combines graph neural network to realize behavior classification.

2 Methods

2.1 System overview

Although the existing behavior recognition system can complete the recognition of human behavior with high recognition accuracy under some conditions, there are still problems such as low segmentation accuracy, failure to extract deep-level features of human behavior and utilize classification algorithms that are compatible with the features. We propose an indoor human behavior recognition system based on Wi-Fi devices, and the overall framework is shown in Fig. 1.

2.2 Behavior segmentation method

2.2.1 Data collection

In the field of wireless communication, CSI is channel property of the communication link. Therefore, we can use CSI to evaluate the state of communication link and achieve recognition of human behavior. We extract the CSI information from the physical layer by modifying the firmware of the Intel 5300 NIC. In frequency domain, the wireless channel can be modeled as

$$Y = H \times X + \zeta \tag{1}$$

where *Y* and *X* are received signal matrix and transmitted signal matrix, respectively, ζ is additive Gaussian white noise, and *H* is a matrix containing CSI information. When the number of sampling points is 1, the matrix *H* represents the CSI value of a single packet. The CSI of the *k*th subcarrier can be expressed as

$$H(k) = |H(k)|e^{j\angle H(k)}$$
(2)

where |H(k)| and $\angle H(k)$ represent the amplitude and phase information of the CSI of the k^{th} subcarrier, respectively.

Under the role of orthogonal frequency division multiplexing system, the bandwidth of common wireless devices is divided into *K* subcarrier groups. Each CSI measurement is a matrix containing $K' = K \times Ntx \times Nrx$ dimensions, where *Ntx* is the number of transmit antennas and *Nrx* is the number of receive antennas.



Fig. 1 Framework of behavior recognition system

2.2.2 Wavelet threshold method-based noise removal

For common indoor human behavior, different types of behavior can lead to different effects on each frequency component of original signal, such that the noise of original signal cannot be removed by a band-pass filter with a fixed frequency range. In addition, using STFT to analyze CSI amplitude will lose time domain information. Therefore, this paper uses wavelet thresholding method to achieve noise removal from CSI signal amplitude. The wavelet thresholding method achieves noise removal mainly by using wavelet decomposition to remove the correlation of data. After decomposing CSI amplitude using wavelet transform, the energy of useful signal will be distributed in the larger coefficients in wavelet domain, while the energy of noise will be distributed in whole wavelet domain. Therefore, in this paper, noise cancellation is achieved by setting a threshold to retain useful signal coefficients and abate noise coefficients. The specific implementation steps of this method are as follows: (1) wavelet decomposition. Set the number of wavelet decomposition levels T_c and the basis function, and use Mallat algorithm to calculate the high-frequency coefficient CH_t and low-frequency coefficient CL_t of the *t*-level wavelet; (2) high-frequency coefficient processing. Calculate the global unified threshold, and readjust the size of the high-frequency coefficient by combining the soft-value threshold method and the relationship between the threshold and the highfrequency coefficient. First, calculate the standard deviation of noise δ :

$$\delta = \frac{1}{0.6745} \frac{1}{T_c} \sum_{t=1}^{T_c} |CH_t|$$
(3)

where 0.6745 is the adjustment coefficient. According to the VisuShrink criterion, calculate the global unified threshold *TH*:

$$TH = \delta \sqrt{2 \log N} \tag{4}$$

where N is the length of the signal. Recalculate the high-frequency coefficient according to Eq. (4), and get:

$$CH_{t}' = \begin{cases} [sign(CH_{t})] (CH_{t} - TH) & CH_{t} > TH \\ 0 & CH_{t} < TH \end{cases}$$
(5)

where $sign(\cdot)$ is a symbolic function (5). Signal reconstruction. According to CH_t and CL_t , the denoised signal is reconstructed.

The CSI amplitudes before and after wavelet denoising are shown in Fig. 2, respectively. In Fig. 2a, the high-frequency noise makes useful information in the signal drowned, while in Fig. 2b, the high-frequency noise is effectively suppressed, and the signal becomes smoother compared with that before denoising, which can reflect the fluctuation change of signal amplitude more clearly when a human behavior occurs.

2.2.3 PCA-based dimensionality reduction

The CSI amplitude is effectively suppressed by high-frequency noise and anomalies after noise cancellation, but the dimensionality of signals does not change. The CSI amplitude fluctuations caused by behavior can be observed in some subcarriers, but



the fluctuations are not clearly observed in other subcarriers. Therefore, it is unreliable to use only the amplitude information in one or some subcarriers for human behavior recognition. In addition, if the amplitude information of all subcarriers is used to realize human behavior recognition, too large amount of data will limit the running speed of the program. In order to ensure the running speed of the program while using the amplitude information of all subcarriers, we use principal component analysis to compress the dimensions of CSI matrix and extract the subspace information of this matrix, and implements human behavior recognition based on subspace information. Using PCA algorithm to reduce the dimension of the matrix, the noise can be suppressed while the fluctuation of CSI data segment reflecting human behavior information can be made more obvious. The steps of PCA algorithm implementation are as follows: (1) decentering. Subtracting the mean value of signals. (2) Calculate the covariance matrix. Suppose H' represents the denoised CSI amplitude,

$$\boldsymbol{H}' = [\boldsymbol{h}'_{k,n}]_{K'\times N} = [\boldsymbol{h}'_1\boldsymbol{h}'_2\cdots\boldsymbol{h}'_k\cdots\boldsymbol{h}'_{K'}]^{\mathrm{T}}$$

$$\tag{6}$$

where N is the number of data packets received by the receiver on a single subcarrier, $\mathbf{h}'_{k} = [h'_{k,1}h'_{k,2}\cdots h'_{k,n}\cdots h'_{k,N}]$, $h'_{k,n}$ are the elements of the k^{th} row and n^{th} column of denoised CSI matrix. Then the covariance matrix C_{H} is

$$\boldsymbol{C}_{H} = COV(\boldsymbol{H}', \boldsymbol{H}') \tag{7}$$

(3) Perform a matrix decomposition on C_H . (4)Using feature vectors obtained from the third step, calculate new projection matrix.

$$\boldsymbol{z}_i = (\boldsymbol{H}')^T \boldsymbol{q}_i \tag{8}$$

where z_i is the subspace signal corresponding to the *i*th eigenvector and q_i is the *i*th eigenvector.

From Fig. 3, it can be seen that the CSI amplitude fluctuations caused by human behavior are easier to analyze from the first five subspace signals. The subspace signals calculated according to the PCA algorithm contain both noisy and useful signals, so it becomes a problem to select appropriate subspace signals for subsequent behavior segmentation and recognition.



Fig. 3 Subspace signal amplitude

2.2.4 Relief-based subspace fusion signal acquisition

In order to improve the accuracy of behavior segmentation, it is necessary to select the optimal one from these five subspace signals that can reflect the human behavior information to the maximum extent.

In existing studies, both the CARM system and the WiAG system use PCA to achieve dimensionality reduction and denoising of CSI data, but the difference is that the former uses the second subspace signal and the third subspace signal, and the latter uses only the third subspace signal. After extensive experiments, it is inappropriate to use a subspace signal alone as the reference signal for human behavior segmentation and recognition, which will lose the information of other subspaces. In order to solve this problem, this paper uses ReliefF algorithm to calculate the weight of each subspace signal and then reconstructs the behavior segmentation reference signal by combining each subspace signal and corresponding weight.

Suppose the set of subspace signals Z is

$$\boldsymbol{Z} = [\boldsymbol{z}_1 \boldsymbol{z}_2 \boldsymbol{z}_3 \boldsymbol{z}_4 \boldsymbol{z}_5] = [\boldsymbol{g}_1 \boldsymbol{g}_2 \cdots \boldsymbol{g}_n \cdots \boldsymbol{g}_N]^T$$
(9)

where $\mathbf{z}_i = [z_{1,i}z_{2,i}\cdots z_{n,i}\cdots z_{N,i}]^T$, $i \in [1, 5]$, $n \in [1, N]$, $\mathbf{g}_n = [z_{n,1}z_{n,2}z_{n,3}z_{n,4}z_{n,5}]$, $z_{n,i}$ is the *n* row and *i* column element of \mathbf{Z} . \mathbf{g}_n is treated as a sample and the class label of \mathbf{g}_n is $c_n \in C$, $C = \{1, 2, 3, \cdots, N\}$. In other words, each row vector is treated as a sample and the row number of that row vector is treated as the corresponding class label. The difference between two samples \mathbf{g}_m and \mathbf{g}_n on subspace signal z_i can be defined as

$$df(\boldsymbol{z}_i, \boldsymbol{g}_m, \boldsymbol{g}_n) = \begin{cases} \frac{|\boldsymbol{z}_{mi} - \boldsymbol{z}_{ni}|}{\max(\boldsymbol{z}_i) - \min(\boldsymbol{z}_i)} & \boldsymbol{z}_i \quad \text{continuous} \\ 0 & \boldsymbol{z}_i \quad \text{separated}, \boldsymbol{z}_{mi} = \boldsymbol{z}_{ni} \\ 1 & \boldsymbol{z}_i \quad \text{separated}, \boldsymbol{z}_{mi} \neq \boldsymbol{z}_{ni} \end{cases}$$
(10)

The ReliefF algorithm selects d samples from all samples that are closest to g_n based on the difference in class labels. All of these selected samples that are of the same type as g_n form the set In, and all samples that are different from g_n and of type c form the set M(c). The weights w_{z_i} of z_i are updated according to Eq. (8) and In and M(c). The equation is as follows.



Fig. 4 Weights of subspace signals



Fig. 5 Subspace fusion signal

$$w_{z_i} = w_{z_i} - \frac{1}{r \times d} \sum_{x \in In} df(z_i, g_n, x) + \frac{1}{r \times d} \sum_{x \notin class(g_n)} \left[\frac{p(c)}{1 - p(class(g_n))} \sum_{x \notin M(c)} df(z_i, g_n, x) \right]$$
(11)

We use this algorithm to calculate weights for five subspace signals calculated by PCA, and the size of each weight is shown in Fig. 4.

Based on each subspace signal and its corresponding weight size, the fused signal *e*:

$$\boldsymbol{e} = \sum_{i=1}^{5} \frac{w_{z_i}}{\sum_{j=1}^{5} w_{z_j}} \times z_i \tag{12}$$

The amplitudes of subspace fusion signals are shown in Fig. 5. The fluctuations in the data segment corresponding to human behavior are more pronounced compared to the non-human behavior data segment.

2.2.5 Data detrending based on segment fitting

The CSI matrix, after being reconstructed by the PCA and ReliefF algorithms, can calculate a signal that maximally reflects the effect of human activity on CSI amplitude. From local variation of the signal, the data segment with no behavior is more moderate than the segment with behavior, but there is still a trend. The presence of this "trend" will keep variance of the local data segments at a large value, which will affect the extraction of behavior-related data segments based on sliding variance, and therefore, the subspace fusion signal needs to be detrended.

The common method is to achieve data detrending by subtracting an optimal fitted straight line, by which the global detrending of the fused signal can be achieved, but the trend of local data segments still exists. We propose the segmented polynomial detrending method to remove the overall as well as local trends of the data. First, the whole fused signal is divided into a number of data segments of length N_0 .

$$\boldsymbol{e} = [\boldsymbol{e}_{N_0}(1)\boldsymbol{e}_{N_0}(2)\cdots\boldsymbol{e}_{N_0}(i)\cdots\boldsymbol{e}_{N_0}(N/N_0)]$$
(13)

where $1 \le i \le N/N_0$. Then, a polynomial of degree *T* is chosen to fit each data segment.

$$e'_{N_0}(i) = fit(e_{N_0}(i), T)$$
(14)

where fit() denotes the polynomial fitting operator. Finally, the fitting error is obtained by subtracting the fitting results of each data segment from original data segment, and the fitting errors of each data segment are recombined to obtain detrended fusion signal ν .

$$\boldsymbol{\nu} = \boldsymbol{e} - \boldsymbol{e}' \tag{15}$$

When fitting the fused signal, the smaller the N_0 , the larger the *T*, the better the fit, but the time complexity of fit will be higher. When the time complexity of whole segment fitting operation is the lowest and overall variance of fitting error (i.e., the fused signal after detrending) is the smallest, then N_0 and *T* are the best choices.

Since time complexity of this algorithm is described in an abstract way, it cannot be calculated accurately using mathematical methods. Therefore, the computer calculates the fitting elapsed time and variance of fitting error under different parameters separately to determine the best and the length of segmented data. In this paper, the degrees of polynomials are set to 1–9, and the segment lengths of data segments are set to 50, 100, 150, and 200, respectively, and corresponding elapsed time and variance are shown in Fig. 6. In Fig. 6a, when the degree of polynomial is less than 4 and the segment length is greater than 50, the time consumption is kept low. In Fig. 6b, the variance decreases with the increase of the number of fits and increases with the increase of segment length. In summary, in the segmented polynomial fit, setting the degree of polynomial to 3 and segment length to 50 is the best choice.

Figure 7 shows the comparison of the effect of two detrending methods, where Fig. 7a and c shows fused signals and their corresponding overall and local trends, respectively, and Fig. 7b and d shows the effect of using general detrending method and segment fitting detrending method, respectively. Compared with the general



Fig. 6 The relationship between degree of polynomial and fitting time and fitting variance



trends.

(c) Fusion signals and their local trends.

Times (s)

Fig. 7 Comparison of the effect of detrending methods

detrending method, the segmented fitting detrending removes local trend in the signal while removing overall trend.

-4

0 1 2

3 4 5 6 7 8 9 10

Times (s)

(d) Local trend removal.

2.2.6 Threshold adaptive method

-10

0 1 2 3 4 5 6 7 8 9 10

Among some existing human behavior segmentation methods, achieving segmentation by setting a threshold is the most common and effective, but these methods still have some drawbacks. For example, for the optimal threshold method, the set threshold is only valid for a specific data segment, and the effectiveness of the set threshold is greatly reduced when some external environment such as the data acquisition scene changes. Some researchers have solved this problem by adapting the thresholds, but too many parameters are introduced in the calculation of adaptive thresholds, and the introduction of these parameters affects the stability of the threshold adaption.

Algorithm 1: Threshold Adaptive		
_	Data: Sliding Variance v	
	Result: Threshold δ	
1	Calculate sliding variance V of v ;	
2	Find the most stable data segment $oldsymbol{V}'$ in $oldsymbol{V}$;	
3	Calculate the data distribution of $oldsymbol{V}'$;	
4	Calculate threshold: $\delta \leftarrow \int_{-\infty}^{\delta} f(V') dV' = 0.975$;	

In this paper, a new threshold adaptive method is designed to calculate the adaptive threshold by the statistical features in the values. The implementation process of the threshold adaptive method is shown in algorithm 1. The method first calculates the sliding variance V of v, then finds the data segment with the longest stationary duration from V and performs kernel density estimation for this segment; and finally calculates the magnitude of the value corresponding to when the cumulative distribution probability reaches 0.975, which is the threshold value of this data segment.

2.2.7 Behavior segmentation

After the preliminary data processing, V is obtained. In order to effectively determine the start and end time of the behavior from V, this paper proposes behavior segmentation algorithm.

The main ideas of the algorithm include two aspects: (1) to make full use of the ideas of "window" and "comparison" to determine the start and end time of a behavior according to the relationship between the sliding variance and the threshold; (2) to set the sliding variance of the behavior that has been determined to occur (2) set the sliding variance of the identified behavior to zero, and then determine the start and end times of other behaviors according to (1). Figure 8 shows the results of human behavior segmentation.

The algorithm flow is shown in Algorithm 2. The inputs to the algorithm are V, δ , W_N , where W_N is the length of the comparison window and plays an important role in



Fig. 8 Behavior segmentation results

determining the behavior start time and end time. The output parameters are multiple t_{start} and t_{end} time pairs, the number of which is related to the number of behaviors. The detailed steps of the partitioning method are as follows.

```
Algorithm 2: Activity Segmentation Algorithm
    Data: V_1 \delta_1 W_N
    Result: A series of t_{start} and t_{end}
    [IndexMax, Max] \leftarrow max(\mathbf{V});
 2 if Max < \delta then
 з
         t_{start} \leftarrow \emptyset, t_{end} \leftarrow \emptyset;
 4 else
 5
          IndexRt, IndexLt \leftarrow IndexMax;
          FlagRt, FlagLt \leftarrow 1;
 6
         while FlagRt = 1 do
 7
               [rRt, cRt, vRt] = find(\boldsymbol{V}(IndexRt + 1 : IndexRt + W_N) > \delta);
 8
 9
               if cRt = \emptyset then
                    t_{end} \leftarrow IndexRt, \ FlagRt \leftarrow 0;
10
11
               else
                    IndexRt \leftarrow IndexRt + max(cRt);
\mathbf{12}
               end
13
14
         end
          while FlagLt = 1 do
15
               [rLt, cLt, vLt] = find(\boldsymbol{V}(IndexLt - W_N : IndexLt - 1) > \delta);
16
               if cLt = \emptyset then
17
                   t_{start} \leftarrow IndexLt, \ FlagLt \leftarrow 0;
18
               else
19
                    IndexLt \leftarrow IndexLt - (W_N - min(cLt) + 1);
20
               end
\mathbf{21}
22
         end
         if t_{end} - t_{start} < 200 then
23
               t_{end} \leftarrow 0; t_{start} \leftarrow 0;
\mathbf{24}
25
         end
26 end
27 V(IndexLt: IndexRt) \leftarrow 0;
28 Go to Line 1;
```

Step 1 Find the maximum value *Max* in *V*, and compare *Max* with δ . When *Max* is smaller than δ . when *Max* is larger than δ , the subsequent steps are continued.

Step 2 Record the index *IndexMax* of *Max*, and also define *IndexRt* and *IndexLt* corresponding to the time point when the behavior occurs and the time point when it ends, and *IndexRt* = *IndexLt* = *IndexMax*. Using *IndexMax* as starting point, find the index values in interval [*IndexMax* + 1, *IndexMax* + W_N] corresponding to greater than δ and compose the set L_1 of these index values. When L_1 is an empty set, the end time point is *IndexMax*; when L_1 is not an empty set, then find the largest element *iMax* in L_1 , and then assign the value of *iMax* to *IndexMax*.

Step 3 Using *IndexMax* as starting point, find the index values greater than δ in the interval [*IndexLt* – *W_N*, *IndexLt* – 1] and assign these index values to the set *L*₂. When *L*₂ is empty, it means that the behavior has started at *IndexMax* and the starting time point is *IndexLt*, if *L*₂ is not empty, find the smallest element *iMin* in the set, assign the value of *iMin* to *IndexLt*.

Step 4 Set the value corresponding to the interval [*IndexLt*, *IndexRt*] to 0 and proceed from step 1 to step 3.

With the above steps, this algorithm can effectively extract the data segments from the original data when the behaviors occur, and at the same time, the segmentation of the original data containing multiple discontinuous behaviors is still effective using this algorithm.

2.3 Behavior recognition methods

2.3.1 STFT-based feature extraction

STFT is a common analysis method used in signal time-frequency analysis. The method first uses a windowing function to segment the signal and then calculates a series of spectral features by sliding window function. In the process of short-time Fourier transform, the segmentation length is extremely important, which affects the time resolution and frequency resolution of the time-frequency diagram calculated by STFT. In this paper, when using STFT algorithm to calculate the time-frequency features. Considering the computer processing time and the effect of the time-frequency features, the sampling frequency is 40 Hz. In order to inhibit the spectral leakage of the short-time Fourier transform, the window function is a Hanning window with a larger side lobes attenuation. The window length is 500 packets, which keeps the same packet sending rate as data acquisition, and the overlap window size is set to 50% of the window length. Figure 9 shows the time-frequency characteristics of three different human behaviors of falling, running, and stationary, respectively. From the time dimension, the duration of running is longer than that of falling, which is consistent with the duration of the real behavior. On the other hand, the falling behavior has a higher-frequency component in the middle of the behavior, while the running has a higher frequency component in the whole period of the behavior; from the frequency dimension, the frequency component of the stationary behavior has a higher energy near the zero frequency, while the other behaviors have a higher energy in the higher frequency bands all have higher energy.



Fig. 9 Time-frequency features

2.3.2 Graph data acquisition

Before introducing graph neural networks to model human behavior features for behavior classification, it is necessary to map the time-frequency features of human behavior into graph structure data. The clearer the relationship between the nodes in the graph structure data, the better the training of the classification model. Therefore, extracting graph structure data from time-frequency features becomes a research focus of this paper.

1. Autocorrelation matrix method Considering that the adjacency matrix of graph is a real symmetric matrix, therefore, the autocorrelation is found for this matrix, and the resulting autocorrelation matrix is also a real symmetric matrix as well. Based on this, the first graph structure data method is to calculate the autocorrelation matrix for the time-frequency features of the individual behaviors, and then extract the graph structure data based on the autocorrelation matrix, as follows.

Assume that the time-frequency features extracted from individual behaviors are **Dp**,

$$\boldsymbol{D}\boldsymbol{p} = \left[\boldsymbol{D}\boldsymbol{p}_1 \boldsymbol{D}\boldsymbol{p}_2 \boldsymbol{D}\boldsymbol{p}_3 \cdots \boldsymbol{D}\boldsymbol{p}_{F_0}\right]^T \tag{16}$$

where F_0 denotes the highest frequency and is related to the sampling frequency in STFT algorithm. The ith row, $Dp_i = \{Dp_i(j)\}$, N_{SB} is related to related to the duration of the behavior and the parameter settings in STFT algorithm. Then, the auto-correlation matrix of time-frequency features is

$$Cor = E(Dp * Dp^{T})$$
⁽¹⁷⁾

where the parameter of i^{th} row and j^{th} column is

$$Cor(i,j) = \frac{\mathrm{E}(\boldsymbol{D}\boldsymbol{p}_{i}\boldsymbol{D}\boldsymbol{p}_{j}) - E(\boldsymbol{D}\boldsymbol{p}_{i})E(\boldsymbol{D}\boldsymbol{p}_{j})}{\sqrt{\mathrm{E}(\boldsymbol{D}\boldsymbol{p}_{i}^{2}) - E^{2}(\boldsymbol{D}\boldsymbol{p}_{i})}\sqrt{\mathrm{E}(\boldsymbol{D}\boldsymbol{p}_{j}^{2}) - E^{2}(\boldsymbol{D}\boldsymbol{p}_{j})}}$$
(18)

Since the correlation coefficients between different vectors exist less than zero, the derived autocorrelation matrix needs to be divided and set to zero, which is handled as follows.

$$Cor'(i,j) = \begin{cases} Cor(i,j) & Cor(i,j) \ge 0\\ 0 & Cor(i,j) < 0 \end{cases}$$
(19)

Then, the graph structure data is constructed based on matrix Cor'. Suppose the graph constructed from this matrix is G, its adjacency matrix is A and the node feature set is X. Then, $A_{i,j}$ can be expressed as

$$A_{ij} = \begin{cases} 1 \ Cor'(i,j) > 0 \\ 0 \ Cor'(i,j) = 0 \end{cases}$$
(20)

The characteristics of the i^{th} node are

$$X_{k} = \sum_{i=1}^{F_{0}} Cor'(i,k)$$
(21)

where $1 \le k \le F_0$. From this, the graph data construction can be completed according to this method.

2. Spectral energy change method

In the previous subsection, the graph structure data construction was achieved by the autocorrelation matrix method, but this method does not fully utilize the timefrequency characteristics, therefore, this paper proposes the spectral energy change method. This method constructs the graph structure data by analyzing the energy change of each frequency and the relative magnitude of energy between different frequencies. At the same time, this method can make maximum use of the original time-frequency features as the features of each node.

First, the whole time-frequency characteristics are divided into five parts: low frequency (L), low and medium frequency (ML), medium frequency (M), medium and high frequency (MH) and high frequency (H), as shown in Fig. 10. The original frequency spectrum can be expressed as follows.

$$\boldsymbol{D}\boldsymbol{p} = [\boldsymbol{D}\boldsymbol{P}_{L}\boldsymbol{D}\boldsymbol{P}_{ML}\boldsymbol{D}\boldsymbol{P}_{MH}\boldsymbol{D}\boldsymbol{P}_{H}]^{T}$$
(22)

where

$$\begin{cases} DP_{L} = [Dp_{1}Dp_{2}\cdots Dp_{F_{0}/5}]^{T} \\ DP_{ML} = [Dp_{F_{0}/5+1}Dp_{F_{0}/5+2}\cdots Dp_{2F_{0}/5}]^{T} \\ DP_{M} = [Dp_{2F_{0}/5+1}Dp_{2F_{0}/5+2}\cdots Dp_{3F_{0}/5}]^{T} \\ DP_{MH} = [Dp_{3F_{0}/5+1}Dp_{3F_{0}/5+2}\cdots Dp_{4F_{0}/5}]^{T} \\ DP_{H} = [Dp_{4F_{0}/5+1}Dp_{4F_{0}/5+2}\cdots Dp_{F_{0}}]^{T} \end{cases}$$
(23)

Different frequency bands contain the energy values of different frequency components at different time points. Depending on the moment, the average energy of different frequency bands is obtained separately. Taking the low-frequency band as an example,



Fig. 10 Five frequency bands

$$DP'_{L}(i) = \frac{5}{F_0} \sum_{j=1}^{F_0/5} Dp_j(i)$$
(24)

where DP'_L denotes the average energy of low-frequency band at i moment. The average energy values of all frequency bands at different moments form the matrix Γ .

Next, construct graph structure data based on Γ . Define a mapping relation ς ,

$$\varsigma(\Gamma) \to \Gamma'$$
(25)

where Γ' is still a matrix and $\Gamma' \in \mathbb{R}^{5*N_{SB}}$, has the same dimension as Γ . In Γ' , the elements of each column are rearranged according to the order of the size of elements in each column of Γ . That is, the average energy of different frequency bands at different moments with respect to the corresponding frequency bands is removed and replaced by the relationship between the magnitudes of the average energy of different frequency bands, while the relationship between the energy changes of the same frequency bands at different moments is retained. In short, constructing graph structure data with matrix Γ' means rearranging the average energy values of different frequency bands at the same moment according to the magnitude of energy values, and defining a node of the graph with that value, and then connecting the nodes belonging to the same frequency band at different moments.

In practice, since the average energy value of the lower frequency band is always higher than the average energy of other frequency bands, this prevents the set of nodes generated by the lower frequency band from intertwining with the set of nodes generated by the other frequency bands. Therefore, a simple processing of the original data is required before constructing the graph structure data, defining

$$\varsigma(f(\Gamma)) \to \Gamma'$$
(26)

where $f(\cdot)$ is the operator, based on the matrix row vectors, subtracting the average value of each row vector, and finally interleaving the nodes constructed with different frequency bands. Figure 11 shows the schematic diagram of the graph structure data. Each circle in the figure represents a node, and the nodes with the same color indicate that they come from the same frequency band.

After the above steps, the graph structure data construction is finally achieved. In order to retain the original time-frequency features to the maximum extent, the



Fig. 11 Graph structure data

time-frequency features are divided into features corresponding to nodes according to frequency bands and moments. Taking the moment *i* of the low-frequency band as an example, the corresponding node features are

$$\boldsymbol{x}_{L}(i) = [\boldsymbol{D}\boldsymbol{p}_{1}(i)\boldsymbol{D}\boldsymbol{p}_{2}(i)\cdots\boldsymbol{D}\boldsymbol{p}_{E_{0}/5}(i)]^{T}$$
(27)

2.3.3 Behavior identification

In previous human behavior recognition systems, behavior classification models are often constructed with algorithms such as support vector machines, decision trees, and LSTM, which generally use time-domain features such as mathematical statistical features as model input data.

In this paper, the autocorrelation matrix method and the frequency energy change method are used to successfully extract data containing graph structure information from the time-frequency features of individual behaviors, respectively, so that a graph neural network can be built to realize graph classification to achieve behavior classification. Figure 12 illustrates the structure of the graph neural network.

As shown in Fig. 12, the forward equation of this classification model is

$$Z = \operatorname{softmax}(\tilde{L}\operatorname{ReLu}(\tilde{L}SW^{(0)})W^{(1)})$$
(28)

where ReLu() is activation function, Z is the prediction class of the model, $W^{(0)} \in \mathbb{R}^{C \times H}$ denotes the weight parameter of the first convolution, where L represents the adjacency matrix of each GCN graph, which is a symmetric matrix with a principal diagonal of 0, adding an identity matrix to matrix L is equal to \tilde{L} , where C denotes the dimensionality of the features in the input layer, and H is the feature dimensionality after the first convolution. $W^{(1)} \in \mathbb{R}^{H \times F}$ is the weight parameter of the second convolution, and F denotes



Fig. 12 Graph neural network model

the feature dimensionality after the second convolution. softmax() is the activation function, which can be expressed as

$$\operatorname{softmax}(z_i) = \frac{1}{\sum_i \exp(z_i)} \exp(z_i)$$
(29)

In addition, the model training uses the gradient descent method to update the parameter matrices $W^{(0)}$ and $W^{(1)}$ in neural network for calculation, which improves the model resistance to overfitting.

3 Results and discussion

3.1 Experimental equipment parameter setting

In order to evaluate the effectiveness of segmentation method and behavior recognition method proposed in this paper, a CSI data acquisition platform based on Wi-Fi 802.11n protocol was built. The platform consists of two mini-PCs equipped with Intel 5300 NICs, with Ubuntu system and CSI-Tools toolkit installed, respectively. In the experiment, the transceiver mode is one transmitter and three receivers. The transmitter operates in 5.7 GHz band with a channel bandwidth of 40 MHz, and the CSI sends packets at a rate of 500 packets per second.

3.2 Experimental scenarios and data acquisition

In this experiment, the experimental data collection was mainly conducted in two environments: YF501 exhibition room and YF316 laboratory. The floor plans of the exhibition room and the laboratory are shown in Fig. 13. The size of the exhibition room is 7.5 m*10.5 m. There are several exhibition cabinets and a lot of miscellaneous items displayed in the room, and its environment is relatively complicated. The experimental room is 11 m*7.5 m. Compared with the exhibition room, the environment is only a large number of desks and other facilities in the room.

In both environments, the area where data are collected is the blank area between the transmitter and the receiver. In the figure, black squares and black dots are schematic diagrams of the locations where in-situ activities (e.g., falling, sitting) occur, and red dashed lines indicate schematic diagrams of activity trajectories for non-in-situ activities



Fig. 13 Environments: exhibition room and laboratory

(e.g., running, walking), and red squares and red dots represent the locations of transmitters and receivers, respectively.

In this experiment, five volunteers of different heights and body sizes were invited to perform four different behaviors in order to increase the richness of data. These behaviors were all most likely to occur in the home environment, which included walking, running, falling, and sitting. For the large range of motion behaviors of walking and running, the volunteers were asked to move back and forth in the open area between the receiver and transmitter. When capturing smaller range of motion behaviors such as falling and sitting, the volunteer was asked not to move his or her active position and to change position only after completing a set of data collection. In addition, after completing a single behavior, volunteers were required to remain stationary for a short period of time. To verify the effectiveness of this segmentation method, the number of behaviors contained in each set of data was not required to be constant when collecting data, but only to ensure that each behavior occurred 30 times for each volunteer. There are 1500 sets of sample data in the two scenarios.

3.3 Behavior segmentation accuracy

During data acquisition, it is difficult to obtain the exact time when behavior occurred due to many factors such as packet loss and low accuracy of timekeeping. Xiao et al. [14] found that human behavior can still be correctly identified by a behavior classification model even if the start and end points of the activity are not accurately detected. Based on this, this paper defines an activity segmentation result determination expression, acc = $|C_0 \cap C_1|/|C_0|$, where C_0 is the set of real packet index values corresponding to this behavior, which can be recorded at the time of occurrence and end of the activity during data collection, and C_1 is the set of packet index values obtained by the behavior segmentation algorithm. When acc is greater than 0.95, the behavior can be considered to be accurately segmented.

For the effectiveness of this segmentation method, the data collected in the exhibition room and the laboratory are segmented using our segmentation method, respectively. The segmentation results of each behavior in both scenes are shown in Fig. 14. The



Fig. 14 Behavior segmentation accuracy

average segmentation accuracy of this algorithm in the two scenes is 0.964 and 0.993, respectively, with high segmentation accuracy for multiple behaviors.

3.4 Comparative analysis of behavior segmentation methods

After the segmentation accuracy of present segmentation algorithm is calculated, this paper uses three existing threshold-based segmentation methods to segment and compare with ours in two scenarios, as shown in Fig. 15.

The experimental scenario of paper proposed the TW-see system is different from the experimental scenarios mentioned in this paper. Since the TW-see system uses the optimal thresholding method, this paper increases the threshold from 0.05 to 0.5 in the form of non-uniform variation to select the appropriate threshold. And its segmentation accuracy reaches the highest value of 0.731 when the threshold is 0.09, which is only slightly higher than Wi-Multi. WiAG and Wi-Multi have segmentation accuracies of 0.704 and 0.822, respectively. Through the above comparison, it can be concluded that the proposed human behavior segmentation method in this paper is higher in segmentation accuracy than the existing threshold-based segmentation algorithms.

3.5 Behavior recognition accuracy

In order to compare and analyze the difference of the effect of these two graph data acquisition methods in the actual test, they are used to extract graph structure data from the data collected in two scenes and then build a graph neural network to recognize human behavior. The experimental results are shown in Fig. 16.

In both scenes, the experimental results show that the behavior recognition accuracy based on the graph structure data extracted by spectral energy change method is better than that of autocorrelation matrix method. For the autocorrelation matrix method, the features of each node in the constructed graph structure data are derived by calculating the weights among all the nodes connected to the node, and a large number of original features are lost. For the spectral energy change method, the features of each node are part of original features, and the change relationships of nodes are obtained from them while retaining the original features. In addition, since the autocorrelation matrix is generally not a sparse matrix, this makes a large number of edge-weight relationships



Fig. 15 Accuracy comparison of multiple segmentation algorithms



Fig. 16 Performance of graph data acquisition methods



Fig. 17 Performance of different classification algorithms

in the construction of the graph structure using this matrix, increasing the data processing volume of subsequent graph convolutional neural network. Therefore, only the graph structure data based on the spectral energy change method were considered in the subsequent tests.

3.6 Comparative analysis of behavior recognition methods

In this section, firstly, the behavior recognition model used in this paper is compared with the traditional behavior model, and the results are shown in Fig. 17. Compared with the traditional behavior recognition model, the model used in this paper has a higher performance in both the exhibition room and the laboratory.

In order to further analyze the performance of the behavior classification model, the behavior recognition method proposed in this paper is compared with the existing classical algorithms Wi-Chase and TW-see. The experimental results are shown in Fig. 18. The recognition accuracy of the behavior data collected in the exhibition room is higher compared to the laboratory. In addition, the recognition accuracy of the stationary behaviors of the human body in both scenes is higher than that of other behaviors.



Fig. 18 Performance of different classification algorithms in different scenarios



Fig. 19 Influence of some factors on the classification model

Overall, the behavior recognition system proposed in this paper is better than Wi-Chase and TW-see in both scenarios except for running in the exhibition room.

3.7 Others

In order to verify the effect of the size of training set on the behavior recognition model proposed in this paper, the model is trained using different proportions of training data, respectively. The behavior recognition accuracies corresponding to different proportions of training data are shown in Fig. 19a. It is obvious that the behavior recognition accuracy increases as the training dataset increases.

To further illustrate the generalization ability and robustness of this behavior recognition model, this paper evaluates it by environment migration. A behavior recognition model is constructed using data collected in one scene, and then, the model is used to recognize data collected in another scene, and the test results are shown in Fig. 19b. The average recognition accuracy was 0.6351 when using the classification model in the exhibition room to validate the behavioral data collected in the laboratory, and 0.66 when using the classification model in the laboratory to validate the behavioral data collected in the exhibition room. Compared to the behavior recognition of non-scene migration, the recognition accuracy was low when using the trained classification model to validate the behavior in other scenes. This shows that the change of environment has a greater impact on the algorithm.

4 Conclusions

In this paper, a behavior segmentation method is proposed to solve the problems of low accuracy and too many intermediate parameters of existing human behavior segmentation methods; in addition, this paper extracts graph structure data from the time-frequency features of individual actions and uses graph neural network to realize human behavior recognition. The experimental results show that the segmentation accuracy of the behavior segmentation method proposed in this paper is 0.964 and 0.993 in the two scenes, which is better than the existing threshold-based behavior segmentation methods. In addition, by extracting the graph structure data through the spectral energy change method and building the human behavior recognition model using graph neural network for the feature-enhanced dataset, the behavior recognition accuracy is significantly improved compared with the traditional classification algorithm.

Abbreviations

CSI	Channel state information
HMM	Hidden Markov model
LSTM	Long short-term memory
OFDM	Orthogonal frequency division multiplexing
PCA	Principal component analysis
RSSI	Received signal strength indication
STFT	Short-time Fourier transform
Wi-Fi	Wireless fidelity

Acknowledgements

Authors acknowledge colleagues of the team in Chongqing University Of Posts And Telecommunications for their support.

Author contributions

All the authors read and approved the final manuscript.

Funding

This research was supported in part by the National Natural Science Foundation of China (61771083, 61704015, 62101085), Science and Technology Research Project of Chongqing Education Commission (KJQN201800625) and Chongqing Natural Science Foundation Project (cstc2019jcyj-msxmX0635).

Availability of data and materials

Please see Sections 6.1 and 6.2 for details on data acquisition. The datasets used during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 19 September 2022 Accepted: 17 May 2023 Published online: 09 June 2023

References

- X. Sun, H. Xu, Z. Dong, L. Shi, Q. Liu, J. Li, T. Li, S. Fan, Y. Wang, Capsganet: deep neural network based on capsule and gru for human activity recognition. IEEE Syst. J. (2022). https://doi.org/10.1109/JSYST.2022.3153503
- D. Wang, J. Yang, W. Cui, L. Xie, S. Sun, Multimodal CSI-based human activity recognition using GANs. IEEE Internet Things J. 8, 17345–17355 (2021)
- H. Fei, F. Xiao, J. Han, H. Huang, L. Sun, Multi-variations activity based gaits recognition using commodity WiFi. IEEE Trans. Veh. Technol. 69, 2263–2273 (2020)
- X. Cheng, B. Huang, J. Zong, Device-free human activity recognition based on GMM-HMM using channel state information. IEEE Access 9, 76592–76601 (2021)

- D. Halperin, W. Hu, A. Sheth, D. Wetherall, Predictable 802.11 packet delivery from wireless channel measurements. SIGCOMM Comput. Commun. Rev. 40, 159–170 (2010)
- 6. G. Wang, Y. Zou, Z. Zhou, K. Wu, L.M. Ni, We can hear you with Wi-Fi! IEEE Trans. Mob. Comput. 15, 2907–2920 (2016)
- Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, H. Liu, E-eyes: device-free location-oriented activity identification using fine-grained WiFi signatures. *Proceedings of the 20th Annual International Conference on Mobile Computing and Net*working, pp. 617–628 (2014)
- H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, S. Li, RT-Fall: a real-time and contactless fall detection system with commodity WiFi devices. IEEE Trans. Mob. Comput. 16, 511–526 (2017)
- X. Wu, Z. Chu, P. Yang, C. Xiang, X. Zheng, W. Huang, TW-See: human activity recognition through the wall with commodity Wi-Fi devices. IEEE Trans. Veh. Technol. 68, 306–319 (2019)
- F. Wang, W. Gong, J. Liu, On spatial diversity in WiFi-based human activity recognition: a deep learning-based approach. IEEE Internet Things J. 6, 2035–2047 (2019)
- F. Wang, W. Gong, J. Liu, Deep transfer learning for gesture recognition with WiFi signals. Pers. Ubiquitous Comput. 26, 543–554 (2022)
- C. Feng, S. Arshad, S. Zhou, D. Cao, Y. Liu, Wi-multi: a three-phase system for multiple human activity recognition with commercial WiFi devices. IEEE Internet Things J. 6, 7293–7304 (2019)
- 13. C. Lin, J. Hu, Y. Sun, F. Ma, L. Wang, G. Wu, WiAU: an accurate device-free authentication system with resnet. 2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), pp. 1–9 (2018)
- C. Xiao, Y. Lei, Y. Ma, F. Zhou, Z. Qin, Deepseg: deep-learning-based activity segmentation framework for activity recognition using WiFi. IEEE Internet Things J. 8, 5669–5681 (2021)
- D. Wu, D. Zhang, C. Xu, H. Wang, X. Li, Device-free WiFi human sensing: from pattern-based to model-based approaches. IEEE Commun. Mag. 55, 91–97 (2017)
- W. Wang, A.X. Liu, M. Shahzad, K. Ling, S. Lu, Understanding and modeling of WiFi signal based human activity recognition. *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pp. 65–76 (2015)
- 17. T.Z. Chowdhury, C. Leung, C.Y. Miao, WiHACS: leveraging WiFi for human activity classification using OFDM subcarriers' correlation. 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp. 338–342 (2017)
- Y. Li, T. Jiang, X. Ding, Y. Wang, Location-free CSI based activity recognition with angle difference of arrival. 2020 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6 (2020)
- Z. Chen, L. Zhang, C. Jiang, Z. Cao, W. Cui, Wifi CSI based passive human activity recognition using attention based BLS. IEEE Trans. Mob. Comput. 18, 2714–2724 (2019)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com