Research Article

HuAc: Human Activity Recognition Using Crowdsourced WiFi Signals and Skeleton Data

Linlin Guo, Lei Wang, Jialin Liu, Wei Zhou, and Bingxian Lu

Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province School of Software, Dalian University of Technology, Dalian, China

Correspondence should be addressed to Lei Wang; lei.wang@ieee.org

The joint of WiFi-based and vision-based human activity recognition has attracted increasing attention in the human-computer interaction, smart home, and security monitoring fields. We propose HuAc, the combination of WiFi-based and Kinect-based activity recognition system, to sense human activity in an indoor environment with occlusion, weak light, and different perspectives. We first construct a WiFi-based activity recognition dataset named WiAR to provide a benchmark for WiFi-based activity recognition. Then, we design a mechanism of subcarrier selection according to the sensitivity of subcarriers to human activities. Moreover, we optimize the spatial relationship of adjacent skeleton joints and draw out a corresponding relationship between CSI and skeleton-based activity recognition. Finally, we explore the fusion information of CSI and crowdsourced skeleton joints to achieve the robustness of human activity recognition. We implemented HuAc using commercial WiFi devices and evaluated it in three kinds of scenarios. Our results show that HuAc achieves an average accuracy of greater than 93% using WiAR dataset.

1. Introduction

Human activity recognition is an important research problem in the social life, pervasive computing, and security monitoring fields [1–3]. Daily activities [4] were seen as an important means of communicating in our daily life, and we can communicate through body language like hands and head rather than speaking. Therefore, human activity recognition systems have been proposed in terms of application demand, technical support, and auxiliary devices.

Previous works related to activity recognition are roughly divided into three categories including wearable-based, vision-based, and WiFi-based. Wearable-based sensing behavior has been popular and widely used in elder healthcare, smart sensing, sports application, and tracking [1, 5, 6]. Researchers leverage the collecting information via sensors to recognize human behavior and analyze human health condition. However, it has several limitations such as *increasing the burden of users*, *the inconvenience of routine life*, and *sensors with limited power*. Vision-based activity recognition has been popular and achieves high accuracy. *The light, shadowing, privacy protection, and angle factors* increase the difficulty of activity recognition and constrain the application fields. Microsoft released Kinect technology and Kinect can provide skeleton information using built-in sensors [7, 8]. Although Kinect-based activity recognition solves the light-environment problem and can track the skeleton joints of an activity with high accuracy, it cannot recognize the imperfect activity due to *the crowded room*, *the presence of obstacles*, and *out of the monitoring range*.

With the coverage of WiFi signals and the improvement of wireless infrastructures in public places, WiFi-based activity recognition systems [4, 9–11] leverage the change pattern of WiFi signals reflected by a human body to recognize the activity. WiFi-based activity recognition systems [12–14] not only ease the burden of wearable-based users, but also can sense the presence of obstacles in comparison with Kinect-based works. For example, WiVi [14] can sense the user's behavior through the wall, and RF-Capture [11] tracks the 3D positions of a human body when the person is occluded completely and captures the human figure without wearable devices.

We are interested in BodyScan system [15], and it is estimated on the idea of the combination of the wearable sensors and WiFi signals. Moreover, it overcomes key limitations of existing wearable devices by providing a contactless and privacy-preserving approach to capture a rich variety of human activities. Based on this work, we explore the combination of CSI and skeleton data to sense human behavior. According to the works mentioned above, we explore three issues of activity recognition in this paper. First, we construct a WiFi-based activity recognition dataset named WiAR to provide a benchmark for previous works. Second, we design the mechanism of subcarrier selection to improve the robustness of activity recognition in the WiAR dataset. Third, we combine WiFi signals with crowdsourced skeleton data to improve the accuracy and robustness of activity recognition breaking the limitations of Kinect technology. The contributions of our work are summarized as follows:

(i) We propose a HuAc system to recognize human activity and also construct a WiFi-based activity recognition dataset named WiAR as a benchmark to evaluate the performance of existing activity recognition systems. We use the kNN, Random Forest, and Decision Tree algorithms to verify the effectiveness of the WiAR dataset.

(ii) We detect the start and end of the activity using the moving variance of CSI. Moreover, we leverage K-means algorithm to cluster effective subcarriers according to subcarrier's sensitivity and improve the robustness of activity recognition.

(iii) We develop a selection method of skeleton joints based on KARD's work named SSJ, and it considers the spatial relationship and the angle of adjacent joints as auxiliary information of human activity recognition to improve the accuracy of tracking.

(iv) We implement the fusion framework of CSI and skeleton data to sense the activity and solve the limitations of CSI-based and skeleton-based activity recognition, respectively. Experimental results show that HuAc achieves the accuracy of greater than 93%.

The rest of this paper is organized as follows. We introduce the related work in Section 2. Section 3 introduces preliminaries of WiFi-based activity recognition, and we describe the overview of HuAc in Section 4. Section 5 describes Kinect module, and WiFi module is shown in Section 6. Section 7 describes the process of human activity recognition. Section 8 evaluates the performance of HuAc system, and we give a case study about a motion-sensing game using WiFi signals in Section 9. Section 10 lists several discussions, and we give the conclusion of this paper in Section 11.

2. Related Work

In this section, related works on human activity recognition can be divided into two categories: Kinect-based, WiFi-based.
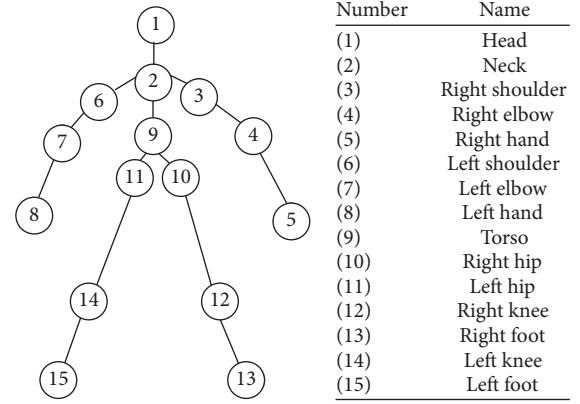

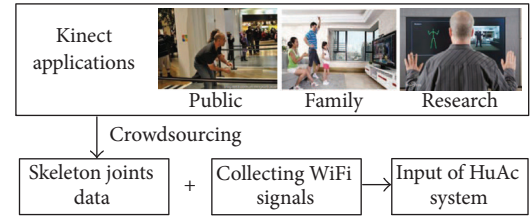| Number	Name
(1)	Head
(2)	Neck
(3)	Right shoulder
(4)	Right elbow
(5)	Right hand
(6)	Left shoulder
(7)	Left elbow
(8)	Left hand
(9)	Torso
(10)	Right hip
(11)	Left hip
(12)	Right knee
(13)	Right foot
(14)	Left knee
(15)	Left foot

FIGURE 1: Skeleton joints.

FIGURE 2: The framework of crowdsourced dataset.

2.1. Kinect-Based Activity Recognition.
Vision-based activity recognition has been proposed and developed in the computer vision field. With the release of Kinect, researchers explore the human activity recognition using depth information and skeleton joints data provided by Kinect [7, 8, 16]. Biswas and Basu [8] leverage the histogram of depth information to recognize eight gestures. Moreover, the differences between continuous frames can obtain the motion profile to describe various gestures. Other works [7, 16] leverage depth information in combination with color image to improve the accuracy of gestures recognition. The limitations of Kinect-based activity recognition contain the restriction of sensing field, skeleton joints overlapping, and position-dependence factors. HuAc system explores the spatial relationship of skeleton joints to describe the trajectory of an activity and combines with CSI to improve the robustness of human activity recognition in a dynamic environment.

2.2. WiFi-Based Activity Recognition.
Early works [17–19] explore the attenuation characteristics of WiFi signals to locate the position of someone and count the number of people in the indoor environment. Researchers study the signal pattern reflected by a human body to sense human behavior [11, 20–22]. These works describe human behavior recognition using coarse-grained RSSI information. For example, WiGest [18] studies the relationship between RSSI fluctuation and gestures to control media player actions without training. Therefore, we explore the relationship between RSSI fluctuation and human movement to detect the presence of an activity.

With the requirement of the practical application and the limitations of RSSI, an increasing number of researchers begin to explore fine-grained channel state information (CSI) to sense human behavior. Compared with RSSI, CSI can capture the tiny behavior [2, 9, 23–28] in terms of location, speed, and direction. WiFall system [2] detects a fall behavior by learning the specific CSI pattern. E-eyes [9] recognizes walking activity and in-place activity by adopting moving variance of CSI and fingerprint technique. Walking activity causes significant pattern changes of the CSI amplitude over time, since it involves significant body movements and location changes. In-place activity (watching TV) only involves relative smaller body movements and will not cause significant amplitude changes with repetitive patterns. The relationship between an activity and the place where an activity occurs motivates the novel idea on human activity recognition. CARM [10] shows the correlation between CSI value and human activity by constructing CSI-speed and CSI-activity model. WiDance [28] explores the Doppler shifts reflected by human behavior to predict the motion direction for the Exergames. We design the combination system of Kinect-based and WiFi-based methods to recognize an activity in different environments such as gaming system, supermarket, and elder health applications.

3. Preliminaries

3.1. RSSI and CSI. Received Signal Strength Indicator (RSSI) [29] in the level of packet represents signal-to-interference-plus-noise ratio (SINR) over the channel bandwidth as follows:

$$\text{RSSI} = 10 \lg \left(\|V\|^2 \right), \tag{1}$$

where V is signal voltage. RSSI is the received signal strength in decibels (dB) and mapped into the distance according to Log-distance path loss model to roughly locate users or devices.

Channel State Information (CSI) depicts multipath propagation at the granularity of OFDM subcarrier in the frequency domain. It contains amplitude and phase measurements as follows:

$$h = |h| \, e^{j \sin \theta}, \tag{2}$$

where $|h|$ and θ are the amplitude and phase, respectively. The variable h shows CSI value of each subcarrier. We study the characteristics of each subcarrier to sense activity in the following work.

3.2. Kinect Technology. Kinect (RGB-D camera) refers to the advanced RGB/depth sensing, hardware, and the software-based technology that interprets the GRB/depth information. The hardware contains a normal RGB camera, a depth sensor (infrared projector and infrared camera), and a four-microphone array, which is able to provide depth signals, RGB images, and audio signals simultaneously. Kinect-based activity recognition algorithm frequently fails due to occlusions, overlapping joints (limbs close to the body), or clutter

(other objects in the scene) [7]. A skeleton reported by Kinect contains 15 joints in Figure 1. We explore the corresponding relationship between skeleton joints and CSI to analyze the characteristics of an activity. Moreover, we explore the fusion information to improve the accuracy of human activity recognition. The details of Kinect-based activity recognition are listed in Section 5.

3.3. WiAR: Constructing WiFi-Based Activity Dataset. At present, there is no WiFi-based public activity dataset as well as vision-based public activity dataset. Due to the sensitivity of WiFi signals, it is hard for peer researchers to reproduce and evaluate previous works. Therefore, we construct the WiAR dataset which collects WiFi signals reflected by sixteen activities in three indoor environments such as empty room, meeting room, and office listed in Table 1. Each activity is performed 50 times by 10 volunteers which consist of five females and five males, and the height of human body ranges from 150 cm to 185 cm.

The environmental complexity according to the room layout divides into three levels including empty environment, normal environment, and complex environment. First, empty environment describes no people and furniture around it. We obtain the high-quality WiFi signals from the empty room due to less noise and treat it as a baseline of WiAR dataset. Then, the normal environment contains furniture and working people. Compared with the empty environment, the multipath effect reflected by the furniture enriches collecting WiFi signals. Finally, a complex environment with furniture and moving people increases the difficulty of human activity recognition. The performance of WiAR dataset is given in Section 8.

3.4. Crowdsourced WiFi Signals and Skeleton Joints. Crowdsourced-based applications [30–37] have been increasingly developed by collecting data and reducing the cost in the Internet field. For the macrolevel network, the work [30] proposed a crowdsensing-oriented mobile cyber-physical system to provide the practical usage of the vita. For the microlevel wireless network, related works [38–41] leverage crowdsensing WiFi signals to detect the user's location.

In our work, we attempt to collect WiFi signals and crowdsourced skeleton joints to reduce the training burden for collecting activity dataset. We obtain the activity label by leveraging the help from Kinect's user. The framework of crowdsourced WiFi signals and skeleton joints are shown in Figure 2.

4. Overview of HuAc

4.1. Observations. The following observations come from the combination of our results and previous works [20, 42–44].

The Impact of Indoor Environment on WiFi Signals Has a Difference with Time. RSSI and CSI keep stability in the static indoor environment, and RSSI fluctuation ranges from

TABLE 1: WiFi-based activity recognition dataset (WiAR).

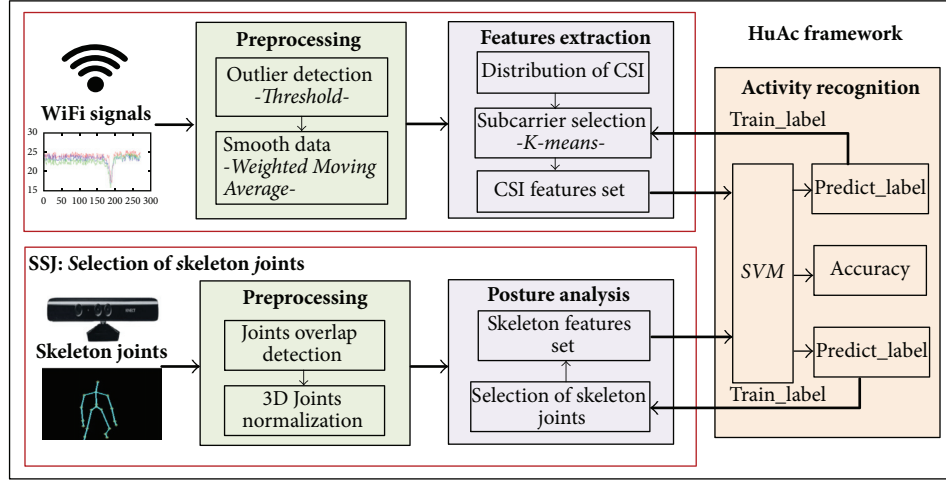| Granularity	Activities	Environments	Devices
Activity	Forward kick, side kick, bend, walk, phone, sit down, squat, drink water	Empty room, meeting room, office	Router, laptop with 5300 card
Gestures	Horizontal arm wave, two-hand wave, high throw, toss paper, draw tick, draw x, hand clap, high arm wave	Empty room, meeting room, office	Router, laptop with 5300 card

FIGURE 3: The framework of HuAc system.

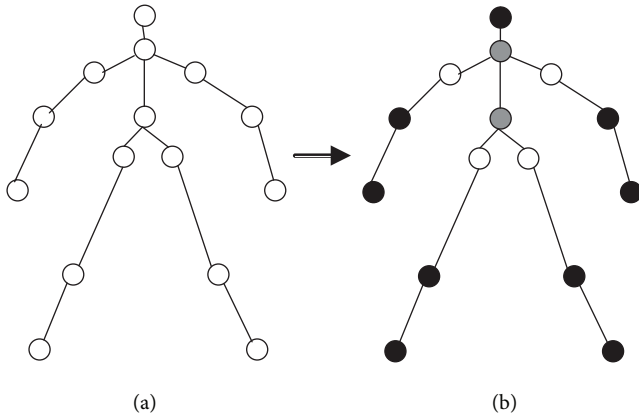(a)   (b)

FIGURE 4: Skeleton structure [7]. (a) A skeleton structure contains 15 skeleton joints. (b) The white circle represents skeleton joints without direction such as shoulder and hip. The gray circle represents the neck and the torso which has a weak effect on the upper-body activity and the lower-body activity except the squat. The black circle represents normal skeleton joints.
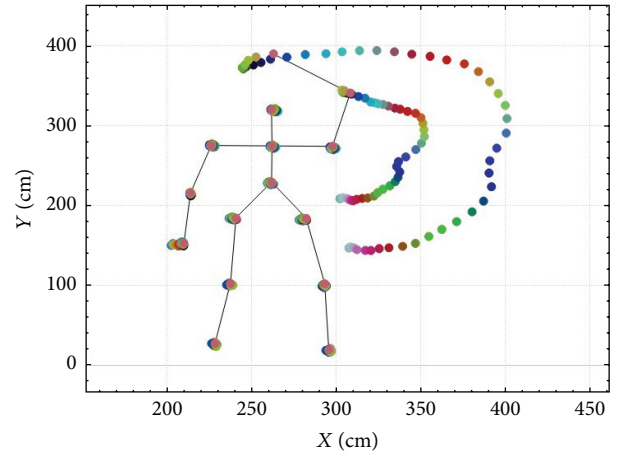

FIGURE 5: High arm wave tracking using skeleton data. The activity has two active joints (right hand, right elbow), and the direction changes with every clockwise movement. However, other joints keep varies steadily in a certain range.

0 dB to 5 dB (empty environment: 0–3 dB; home environment: 0–7 dB; office: 0–5 dB; dynamic environment: 5–10 dB). Although RSSI sharply changes with environmental change, it cannot describe the fine-grained change of indoor environment due to the multipath effect. However, CSI is able to sense the change of fine-grained environment and detects what happened in an indoor environment. Specifically, RSSI only can find the environmental change and cannot sense how the environment changes. CSI can find what causes environmental change and also can recognize how the environment changes such as tracking, sensing environment, and activity recognition.

It Is Hard to Distinguish Similar Activities. Existing works [2, 15, 45] explore the similar activity recognition. For example, WiFall [2] extracts seven features to describe fall behavior because similar activity causes the similar patterns of CSI,

and it is difficult to distinguish them only using anomaly detection. The following RT-Fall system adopts the CSI phase difference to segment fall and fall-like activities because the phase difference of CSI is a more sensitive signature than CSI amplitude for activity recognition. The phase of CSI depends on the variation of LOS (Line-of-Sight) length. Therefore, the breakthrough point of the similar activity recognition rests on the physical difference between similar activities.

The Same Activity Operated by Different People Has Various Signal Patterns. According to our observations, the amplitude of CSI reflected by the same activity changes continuously in the different time and environments. Therefore, we cannot recognize activity with high accuracy according to the amplitude of CSI. The changing pattern of signals reflected by an activity can describe the characteristic of activity as verified by Smokey [25]. Therefore, we explore the changing pattern of signals to recognize an activity.

The Impact of Activity with Different Directions on Activity Recognition. In order to explore the impact of direction on activity recognition, we design a simple and clear experiment on the playground because the playground does not have rich multipath effect and other wireless devices. We explore the impact of four directions including east, west, north, and south on the change pattern of signals, and the difference between face and back to the AP is biggest. Moreover, CSI data we collect in the playground contains less noise than that in an indoor environment.

4.2. Framework of HuAc. The HuAc framework consists of the Kinect-based module and WiFi-based module in Figure 3. We describe details of each module, respectively.

Kinect module consists of the preprocessing and posture analysis. We detect the overlap of skeleton joints using the statistical method and complete the normalization of skeleton joints. In order to obtain effective features of skeleton joints, we analyze postures of an activity according to the sequence of skeleton joints. Moreover, we design a selection method of skeleton joints named SSJ according to the result of posture analysis. Finally, we extract features of skeleton joints according to effective skeleton joints and also consider the spatial relationship of adjacent joints as auxiliary information to sense human activity.

WiFi module consists of the preprocessing and features extraction. In the preprocessing stage, we detect and remove the outlier data of an activity sequence according to the variance of RSSI reflected by an activity. After removing outlier data, we leverage the weighted moving average to smooth the activity data. For features extraction, we first analyze the amplitude distribution of CSI reflected by an activity to evaluate the sensitivity of the subcarrier on an activity. Then, we use K-means algorithm to cluster effective subcarriers. Finally, we extract important features from effective subcarriers to improve the stability of human activity recognition.

We use the combination information of CSI features set and the skeleton features set as an input of SVM to recognize human activity. Compared with the result of *predict_label*, we

give a feedback to the previous process of HuAc framework by using a *train_label*, respectively.

5. Kinect Module

We mainly describe the details of Kinect module on the human activity recognition. Kinect module contains the preprocessing and posture analysis.

5.1. Preprocessing. The collected skeleton data contain empty values due to the overlap of skeleton joints or the occlusion in the motion-sensing game. Therefore, we need to detect the overlapping joints and replace the invalid values by recovering the true value of the overlapping joints. We leverage the relationship between the coordinates of adjacent joints to detect the overlapping joints. Certainly, we discard the sample of an activity when the percent of invalid joints exceeds the threshold.

After recovering the invalid data, we normalize the coordinates of skeleton joints due to the differences of people's height and the distance between the user and the sensor. The work [7] extracts 11 joints (except right shoulder, left shoulder, right hip, and left hip) from 15 joints in Figure 4, and we explore 30 subcarriers with the similar pattern reflected by a human body. Therefore, we select 15 joints to match the 15 subcarriers. Let J_i be one of the 15 joints detected by the Kinect, and the coordinates vector f is given by

$$f = \{j_1, j_2, \ldots, j_i, \ldots, j_{14}, j_{15}\}, \qquad (3)$$

where j_i is the vector containing the 3D normalized coordinates of the ith joint J_i detected by Kinect. Thus,

$$j_i = \frac{J_i}{s} + T_i, \quad 1 \le i \le 15, \qquad (4)$$

where s is the scale factor which normalizes the skeleton according to the distance h, between the neck and the torso joints of a reference skeleton, and

$$s = \frac{\|J_9 - J_2\|}{h}. \qquad (5)$$

The translation matrix, T, needs to set the origin of the coordinate system to the torso. After preprocessing phase, we obtain high-quality skeleton data.

5.2. Postures Analysis. An activity consists of subactivity sequence over time. According to the skeleton structure, a human body is divided into two parts including upper body and lower body. Upper body contains five joints (right elbow, left elbow, right hand, left hand, and head) and two baseline joints (neck, torso) as in Figure 4. Lower body contains four joints (right foot, left foot, right knee, and left knee). We reproduce the tracking of skeleton joints using QT tool and plot the trajectory chart of each activity. We observe that the adjacent joints keep the similar track in Figure 5, and some joints have slight movement influenced by human activity.
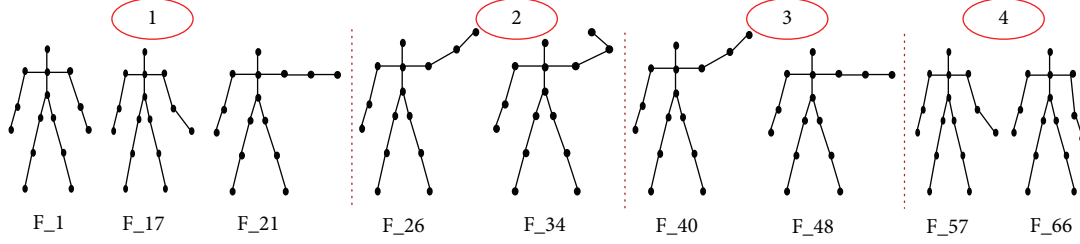
FIGURE 6: Postures of horizontal arm wave.

For example, when the right elbow and right hand move in the clockwise direction to complete the horizontal arm wave, we observe that right hip and left hip have slight movement.

According to the change of joints sequence, we can segment an activity into several subactivities in terms of direction and pause factor. Horizontal arm wave behavior consists of four postures (subactivities) as in Figure 6. Each subactivity roughly contains 14 frames and F_i represents the ith frame (packet) of the activity reported by Kinect. We can evaluate the rough activity according to the sequence of subactivity. Except for related joints of each subactivity, torso and hip joints have a weak swing. We neglect the impact of weak swing on the activity recognition. We pay more attention to the selection of skeleton joints in the following section.

5.3. SSJ: Selecting Skeleton Joints. We design a selection method of skeleton joints named SSJ to describe a fine-grained subactivity. After postures analysis, we know the relationship between a subactivity and key skeleton joints. We expend the coordinated system of human skeleton to miniature coordinated system of subactivity skeleton by the above-mentioned relationship. The miniature coordinated system needs to determine a fixed skeleton joint and different subactivities have different fixed skeleton joints. For example, we observe that shoulder joint is a fixed joint from the process of high arm wave behavior. Therefore, we determine the starting point coordinate of the miniature coordinated system corresponding to the subactivity.

6. WiFi Module

We introduce the design details of WiFi module on the human activity recognition. WiFi module consists of the preprocessing and features extraction.

6.1. Preprocessing. The collected data with noises increases the difficulty of activity recognition due to the tiny differences between noises and WiFi signals reflected by a fine-grained activity. Outlier data also weaken the quality of collecting data. Therefore, we detect outlier using the variance-based method and remove high-frequency signals using the low-pass filter. Moreover, we reduce the sawtooth wave of the filtered signal by using the weighted moving average.

6.1.1. Outlier Detection and Removing High Frequency. Outlier has an important impact on the quality of collecting data because outlier increases or decreases the fluctuation strength of WiFi signals. We analyze the RSSI distribution of an activity to evaluate the possible experience-threshold. Then, we combine with the variance of RSSI values to detect outlier. After removing outlier data, the activity corresponds to the low-frequency change of CSI according to the waveform of CSI reflected by an activity. Therefore, we adopt the low-pass filter to remove the high-frequency data in Figure 7.

6.1.2. Weighted Moving Average. For filtered signal, signal data still contain sawtooth wave. Because CSI is sensitive to indoor layout or human movement, and the received CSI fluctuation caused by the environment is hard to distinguish from the fluctuation caused by a fine-grained activity. Therefore, we smooth the CSI data using the weighted moving average as proposed in WiFall [2]. We randomly select 15 subcarriers from 30 subcarriers which correspond to 15 skeleton joints of Kinect technology. Each CSI stream contains 15 subcarriers as $\{CSI_1, CSI_2, \ldots, CSI_{15}\}$. $CSI_{t,1}$ is the first subcarrier of CSI at time t. $\{CSI_{1,1}, \ldots, CSI_{t,1}\}$ indicates the CSI sequence of first subcarrier in the time period t. The latest CSI has weight m, the second latest $m-1$, and so on. The expression of CSI series is shown as follows:

$$
\begin{aligned}
CSI_{t,1} = \frac{1}{m + (m-1) + \cdots + 1} \times \big(m \times CSI_{t,1} \\
+ (m-1) \times CSI_{t-1,1} + \cdots + 1 \times CSI_{t-m-1,1}\big),
\end{aligned}
\tag{6}
$$

where $CSI_{t,1}$ is the averaged new CSI. The value of m decides in what degree the current value is related to historical records. In our study, we select m according to the experience and trial method. We first set m as 5 which means the length of 5 packets. A weighted moving average algorithm and median filter have the similar effect on the original signals recorded by the receiver in Figure 7. They can remove the galling of signals and alleviate the sharp change of signals. With the m increasing, the weighted moving average algorithm becomes more smooth than the low-pass filter and the median filter. Finally, we set m to 10 because each activity produces a sharp change in 10 packet periods.

6.2. Feature Extraction. Plenty of related works summarize the importance of features extraction for human activity recognition in a dynamic indoor environment. We segment
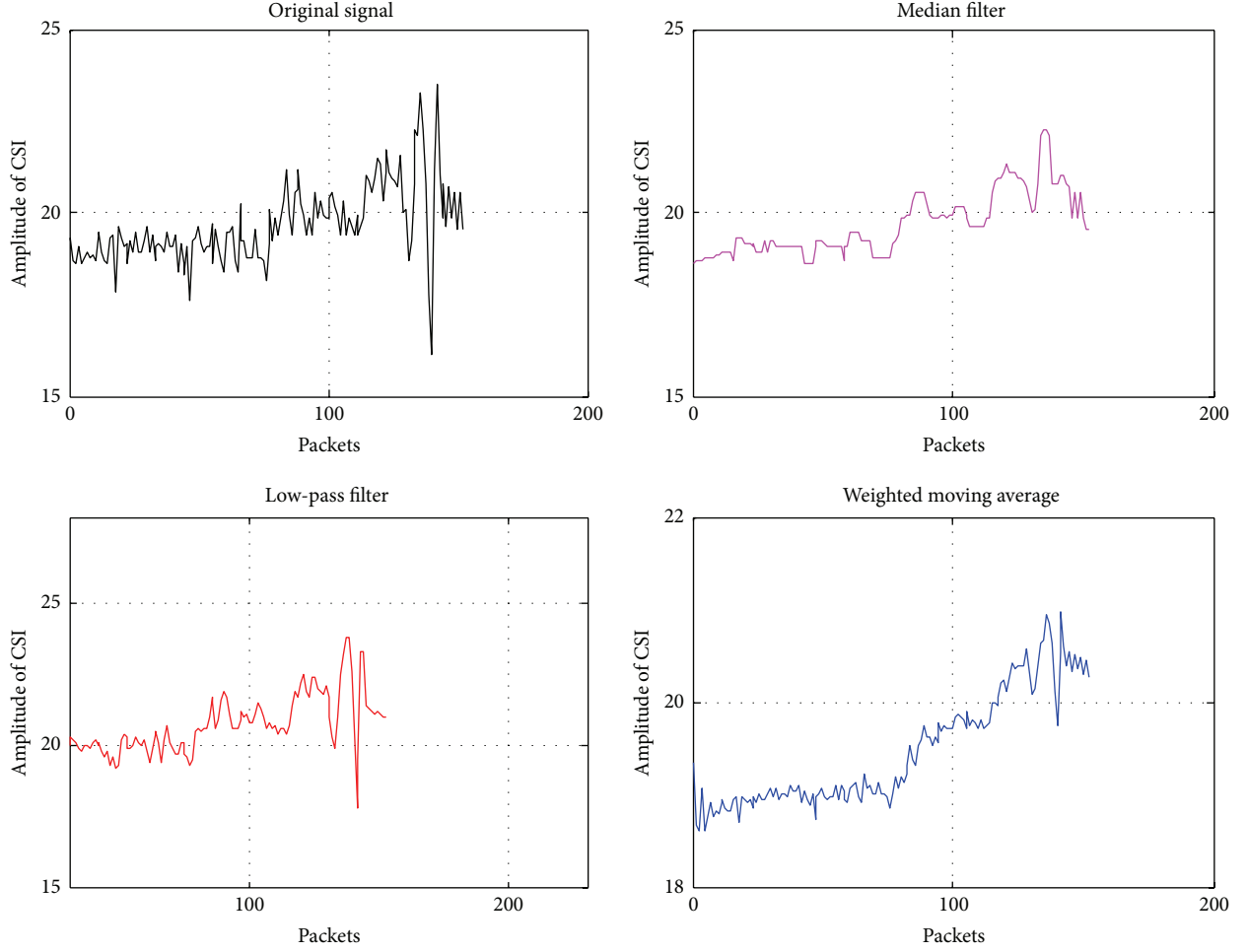
FIGURE 7: Methods of signal filtering.

activity after smoothing CSI and extract features of each activity according to activity characteristics. Kinect-based features extraction quotes the work [3].

6.2.1. Activity Segmentation. Activity segmentation mainly detects the start and end of an activity and removes the nonactivity packets from a sample which corresponds to the whole activity. We propose two methods to detect the start and end of an activity and improve the robustness of segmentation algorithm. First, we remove the first second and the last-second data sequence of an activity to reduce the error of true activity sequence in our experimental environment. But this method is invalid in the practical environment due to the unknown time which each activity starts. Therefore, we leverage moving variance of CSI to detect the start and end of each activity. Moving variance of CSI describes the difference of the local packets reflected by the activity. Packet sequences on the corresponding activity are defined as $X = \{x_1, x_2, \ldots, x_n\}$. X represents data sequence (a sample) of an activity, and x_i represents the ith packet in the

data sequence. We often use the standard deviation instead of the variance of CSI as follows:

$$\sigma_i = \sqrt{\frac{\sum_1^m \left(x_{i+j-1} - \overline{x}\right)^2}{m}}, \quad (i = 1, 2, \ldots, n - m), \quad (7)$$

where m represents step-size and $\overline{x}$ is the mean value of samples.

We construct a window per 10 packets from the packet sequence of each sample and compute the variance of the window. Then, we construct the moving variance histogram and compare with other strength windows. Finally, we can detect the sharp points of each activity and roughly recognize the start and end of each activity from the data sequence. The start and end of the activity period are shown in Figure 8. The red circle describes a sharp change of CSI at the start point of collecting data, but it is not the true start of an activity. The red rectangle represents the duration of activity. Moreover, the black dotted line roughly represents the true start and end of the activity. According to our experimental results, detecting the start and end of the activity still causes a small error due to the sensitivity of signals.
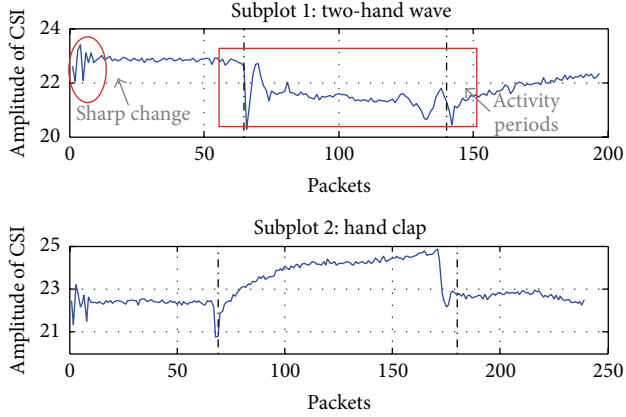
9

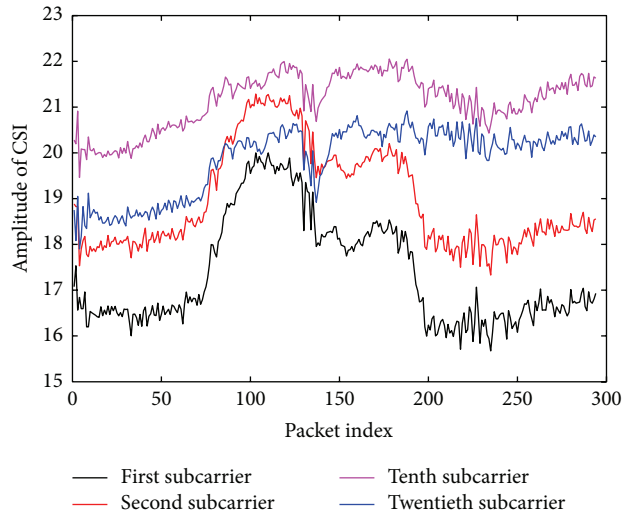FIGURE 8: Segmentation point of similar activity.

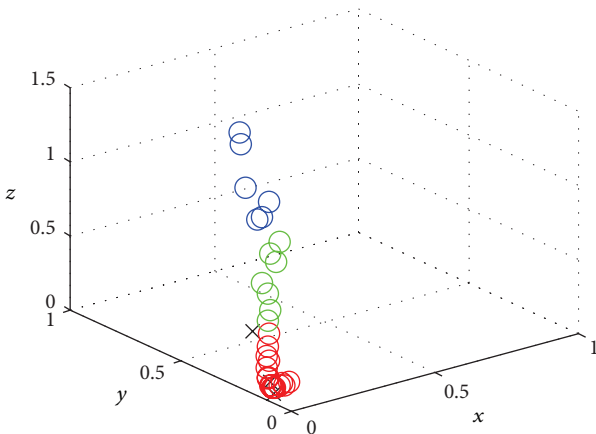FIGURE 9: The fluctuation of different subcarriers reflected by the horizontal arm wave behavior.

FIGURE 10: Clustering subcarriers.

6.2.2. Subcarrier Selection and Feature Detection. According to our observation, subcarriers have the similar tendency for the same activity in Figure 9, but they have different sensitivity. Therefore, we select the obvious subcarriers reflected by an activity using K-means to achieve the robustness of human activity recognition. Thirty subcarriers are divided into 3 clusters using K-means algorithm in Figure 10. According to the output of K-means algorithm on subcarriers, CSI features we extract include variance, the envelope of CSI, signal entropy, the velocity of signal change, median absolute deviation, the period of motion, and normalized standard deviation. Finally, we construct the features set of CSI.

7. HuAc: Activity Recognition

We explore the relationship between CSI-based and skeleton-based methods on human activity recognition in Figure 11. The CSI-based method leverages the signal pattern to recognize an activity. The skeleton-based method uses the coordinate change of skeleton joints to recognize the same activity. From the opinion of experiment results, an activity with back to the AP has more complex CSI pattern and has the smaller amplitude than that with face to AP.

We mainly introduce several classification algorithms used by the human activity recognition field including kNN, Random Forest, Decision Tree, and SVM. In the following sections, we verify that the performance of SVM outperforms others. We select SVM classification algorithm to recognize sixteen activities in the WiAR dataset. CSI features set and skeleton features set as the inputs of SVM train the optimal model to achieve the stable accuracy of activity recognition. The outputs of SVM contain the *accuracy*, *predict_label*, and *prob_estimates*. We evaluate the performance of classification algorithm according to the accuracy and achieve the accuracy of activity recognition using the *predict_label*. According to the match level between *train_label* and *predict_label*, we obtain the false positive rate and the false negative rate. We analyze the result and give a feedback on the previous step. According to the feedback, we pay more attention to the activity with low accuracy.

8. Implementation and Evaluation

8.1. Implementation

8.1.1. Experimental Setup. We use a commercial TP-Link wireless router as the transmitter operating in IEEE 802.11n AP mode at 2.4 GHz. A Thinkpad 400 laptop running Ubuntu 10.04 is used as a receiver, which is equipped with off-the-shelf Intel 5300 card and a modified firmware. During the process of receiving WiFi signals, the receiver pings 30 pkts/s from the router and records the RSSI and CSI from each packet. Three experimental environments including empty room, meeting room, and office are shown in Figure 12.

8.1.2. Experimental Data. We deal with data from three cases: For *WiFi-based activity data*, we collect activity data in different indoor environment. For *skeleton data*, we directly
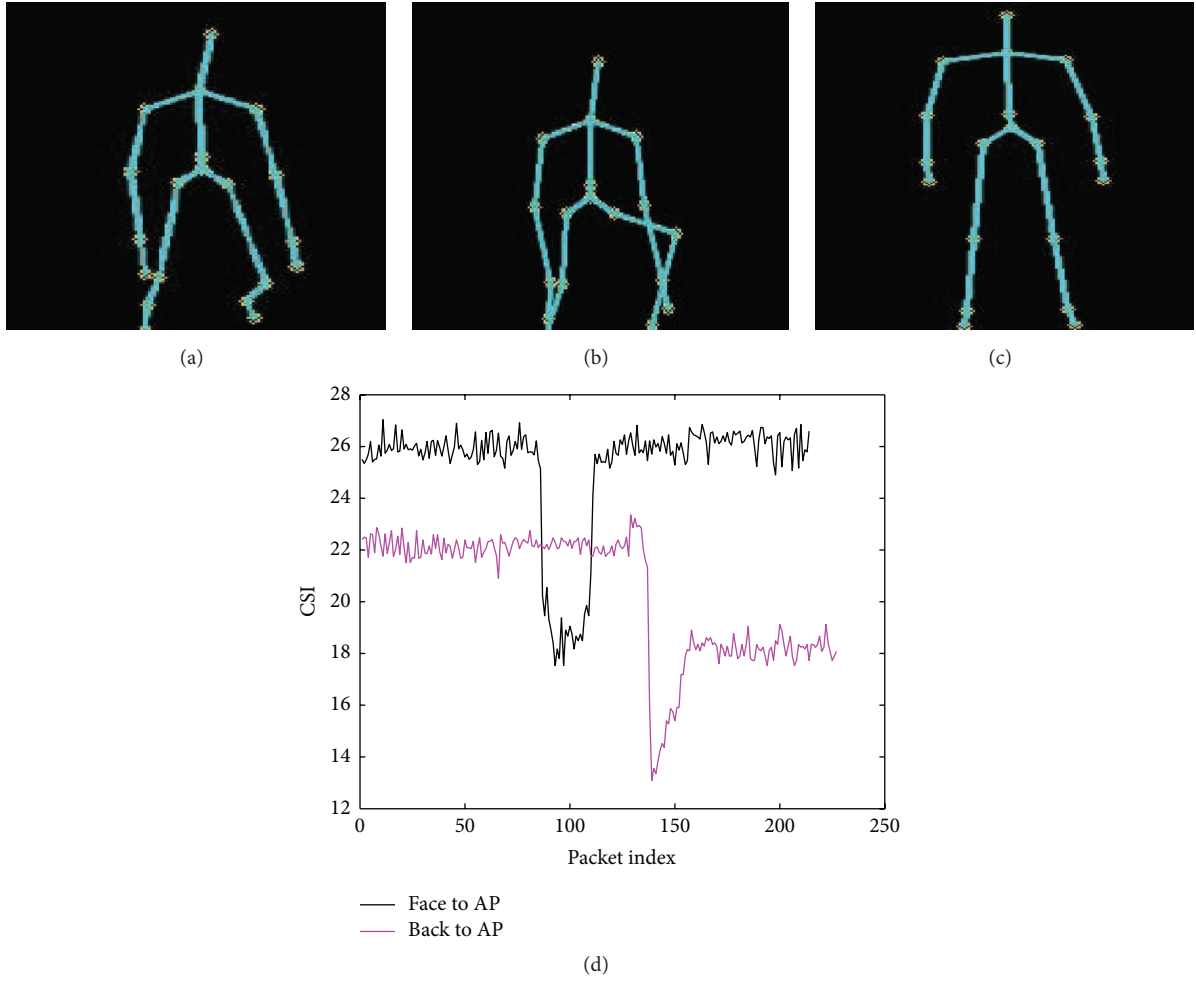
(a)

(b)

(c)

(d)

FIGURE 11: Skeleton joints sequence and CSI change of squat behavior. (a)–(c) represent the skeleton sequence of squat behavior. (d) is the CSI change reflected by squat behavior in terms of face to AP and back to AP.
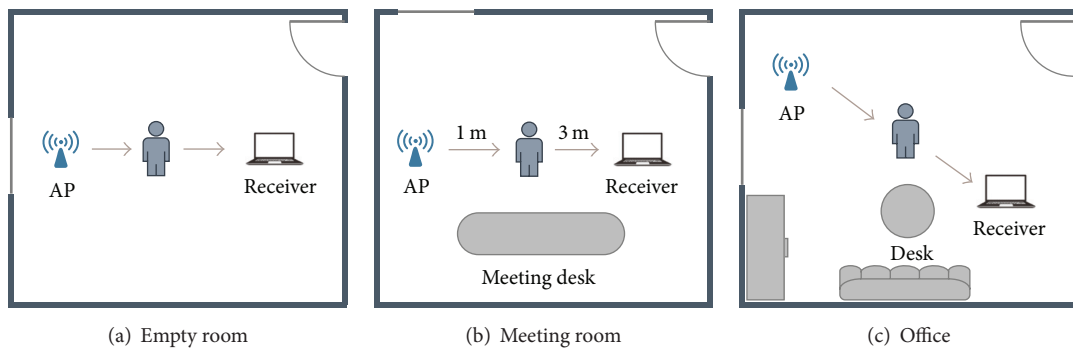

(a) Empty room

(b) Meeting room

(c) Office

FIGURE 12: Experimental scenarios.

leverage the KARD dataset [3] to get the skeleton data. For *environmental data*, we mainly collect data from the empty room, meeting room, and office with the human. Our goal is to explore the impact of the environmental factor on the WiFi signals and analyze the differences between an activity and environmental change on WiFi signals according to the above-mentioned three kinds of data.

We collect WiFi signals to construct a new dataset named WiAR which contains 16 activities with 50 times performed by ten volunteers. The details of WiAR have been introduced in Section 3. The KARD contains RGB video (.avi), depth video (.avi), and 15 skeleton points (.txt). Each volunteer performs 18 activities 3 times each with ages ranging from
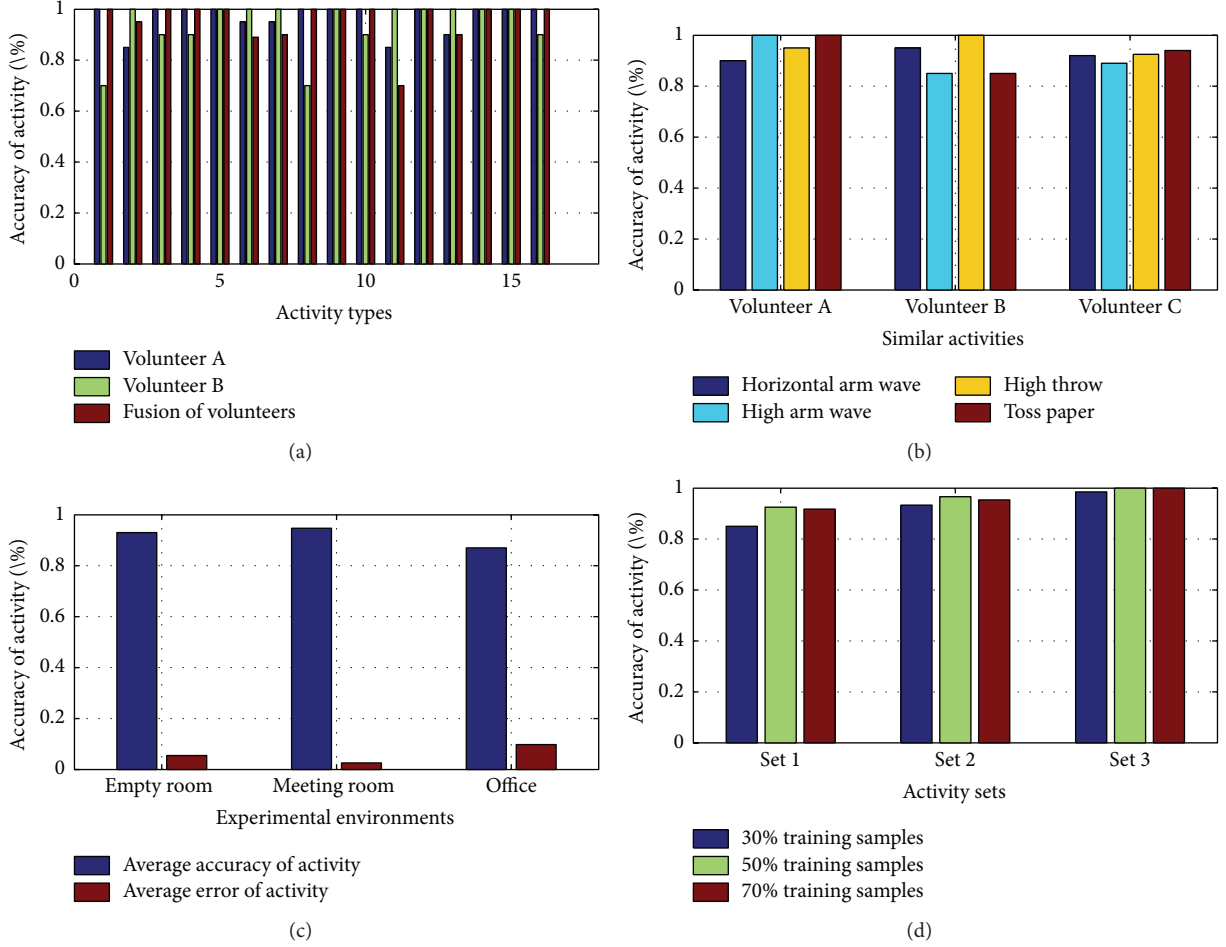
(a)

(b)

(c)

(d)

FIGURE 13: Performance analysis of activities using CSI. (a) Sixteen activities include horizontal arm wave, high arm wave, two-hand wave, high throw, draw x, draw tick, toss paper, forward kick, side kick, bend, handclap, walk, phone, drink water, sit down, and squat. (b) Four activities contain horizontal arm wave, high arm wave, high throw, and toss paper. (c) The impact of experimental environments on accuracy. (d) The impact of training samples on accuracy of three activity sets.

20–30 years and height from 150–180 cm. In this paper, we only select 16 activities as target activity listed in Table 1.

We design three experimental schemes to analyze the accuracy of activity recognition. First, we collect RSSI and CSI to recognize an activity as the reference point. Second, we leverage the skeleton data of KARD to recognize an activity by using our method and previous method [3] in the similar indoor environment. Third, we propose a fusion scheme which CSI combines with skeleton data to recognize an activity. Moreover, we design another experimental scheme in which volunteer performs an activity with repeating 10 times. The goal of the experimental scheme is to investigate the periodic regularity of CSI change influenced by the same activity.

8.2. Evaluation of WiAR Dataset.
We analyze activity data of all volunteers to evaluate the performance of WiAR dataset using kNN with voting, Random Forest, and Decision Tree algorithms.

We study the impact of subcarriers and antennae on the performance of activity recognition by using four classification algorithms shown in Table 2. It shows that the accuracy using SVM outperforms other classification algorithms and 10 subcarriers obtained by subcarrier selection mechanism increase 4.26% when compared with activity recognition using 30 subcarriers. Three antennae such as A, B, and C increase the diversity of CSI data and keep more than 80% of activity recognition accuracy. The four algorithms verify the effectiveness of WiAR dataset.

8.3. Evaluation of Activity Recognition

8.3.1. Performance of Activity Recognition Using RSSI. The section evaluates the performance of RSSI on the human activity recognition. The difficulty we encounter in the process of activity recognition using RSSI is how to deal with the multipath effect caused by indoor environment and reflection effect caused by human behavior. We select an indoor environment as a reference environment which keeps

TABLE 2: Performance comparison by four classification algorithms.

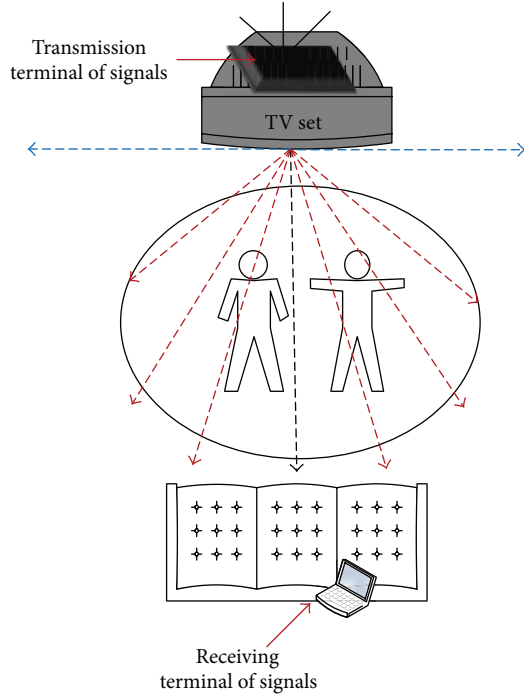| Method	10 subcarriers			30 subcarriers		
	A	B	C	A	B	C
kNN	0.875	0.916	0.947	0.916	0.895	0.947
Random Forest	0.885	0.906	0.958	0.906	0.895	0.948
Decision Tree	0.8542	0.822	0.916	0.865	0.834	0.917
SVM	**0.9625**	**0.9688**	**0.975**	**0.94375**	**0.90625**	**0.9375**

FIGURE 14: Motion-sensing game using WiFi signals.

static and only contains a volunteer and an operator. We leverage RSSI variance as an input of SVM to obtain the 89% of average recognition accuracy in the static environment. When other people move and are close to the control area of WiFi signals, the accuracy of activity recognition decreases to 77% with the high stability. Several activities face the low accuracy such as two-hand wave, forward kick, side kick, and high throw. The average false positive rate is 8.9% and increases to 15.3% in a dynamic environment. Therefore, human activity recognition using RSSI needs the help of CSI-based method to improve the accuracy and the robustness of human activity recognition.

8.3.2. Performance of Activity Recognition Using CSI. This section elaborates the impact of interference factors on human activity recognition using CSI in the following four aspects: human diversity, similar activities, different indoor environments, and the size of a training set. Moreover, we keep the fixed position of volunteers and the distance between receiver device and transmitter device in the whole experiment.

The Impact of Human Diversity on the Accuracy. Human diversity not only increases the diversity information of CSI but also raises the difficulty of activity recognition because different people have different motion styles such as speed, height, and strength. We achieve 93.42% of average recognition accuracy for all volunteers in Figure 13(a). We select two volunteers including volunteer A and volunteer B to verify the impact of human diversity on the accuracy. Volunteer A which often regularly exercises obtains 97.1% of average recognition accuracy. Volunteer B which rarely exercises in the routine lives achieves 92.3% of average recognition accuracy. Therefore, the exercise experience increases the differences between activities due to standard activity and improves the recognition accuracy.

The Impact of Similar Activity on the Accuracy. We explore two group similar activities including high arm wave, horizontal arm wave, high throw, and toss paper in Figure 13(b). The first group activity achieves 92.5% of average recognition accuracy and 94.6% for the second group. The false positive for similar activity is higher than independent activity. For example, forward kick and side kick also belong to the similar activity, and the difference between them is the moving direction. In order to obtain the better accuracy, we will consider the impact of moving direction on the signal change in the future work.

The Impact of Indoor Environment on the Accuracy. As shown in Figure 12, there are three experimental environments including empty room, meeting room, and office in terms of the complexity. The accuracy about three environments is shown in Figure 13(c). The accuracy of the meeting room with 94.7% outperforms the other two environments, and then accuracy was 93% for empty room and 87% for office due to multipath effect. The meeting room generates 2.6% of average error, and 9.8% of average error in the office due to paths excessively reflected by the body. We will deeply explore the multipath effect using the amplitude and phase of CSI in the future work.

The Impact of Training Size on the Accuracy. We design three proof schemes to analyze the accuracy of human activity recognition by using different training sizes in Figure 13(d). We first introduce three activity sets and three training sets. Activity set 1 consists of horizontal arm wave, high arm wave, high throw, and toss paper. Activity set 2 contains

Table 3: Mapping relation between WiFi and Kinect.

	WiFi	Kinect
Techniques	CSI	Skeleton joints
Granularity	Subcarriers (15)	Joints (15)
Parameters	Similarity coefficient, median absolute deviation, variance, environment factor	Distance between joints, angle between adjacent joints, variance, sequence of key joints

two-hand wave and handclap activity. Activity set 3 consists of phone, draw tick, draw x, and drink water. Moreover, these activity sets come from the same people. With the training size increasing, the accuracy of activity recognition is improved by about 10% for the activity set 1. Activity set 1 has a low accuracy because activity set 1 contains more similar activities. Although activity set 3 also contains similar activities, the accuracy is better than activity set 1 due to the strength of activity.

8.3.3. Performance between Kinect-Based and WiFi-Based Activity Recognition. It is hard for the waveform of RSSI with noise to keep the stability when controlling area changes during collecting data. Therefore, we use waveform shape of RSSI to recognize an activity that is not a better choice for the current level of technology. Waveform pattern of CSI can describe an activity with credibility and fine-grained way. The mapping relationship between CSI-based and Kinect-based activity recognition for various activities is represented by using several parameters shown in Table 3. The environmental factor is evaluated by using the number of multipaths and the complexity of the indoor environment. In order to extend the application field of activity sensing, we construct the mapping relationship between CSI-based and Kinect-based activity recognition. The mapping relationship can avoid information loss. For example, once one of the two datasets is lost, activity recognition system still works by using another dataset information.

We evaluate the performance of human activity recognition from KARD dataset [3]. The highest recognition rate is 100% (side kick, handclap), while the worst is 80% (high throw). We propose a selection method of skeleton joints named SSJ to improve the accuracy of activity recognition and reduce the computing cost. SSJ achieves 93.15% of the average recognition accuracy. Existing three activities, such as high arm wave, draw kick, and sit down, achieve the low accuracy of 80%, 75%, and 70%, respectively. Table 4 shows the performance of four methods including CSI-based, KARD-based (skeleton joints), SSJ-based, and HuAc. Table row of the bold font shows that skeleton-based method outperforms CSI-based method on the accuracy of activity recognition. Table row of the italic font shows that several activities are sensitive to CSI. HuAc improves the accuracy of activity recognition and increases the stability of activity recognition in a dynamic indoor environment. We focus attention on the stability of activity recognition algorithm or system in the future work.

9. Case Study: Motion-Sensing Game Using WiFi Signals

We introduce the application based on our work in the motion-sensing game. At present, Kinect provides the angle with limitations in which the horizontal viewing angle is 57.5° and 43.5° for vertical viewing angle, and distance with limitation ranges from 0.5 m to 4.5 m. Moreover, Kinect loses the sensing ability when barrier occurs and occludes game user in the control area. An interesting point of our work is that we pay more attention to the activity itself, and we do not care about the user location. However, Kinect needs to adjust the location of a user before activity recognition to achieve well sensing. Therefore, we will propose a framework instead of Kinect in the future when the accuracy of human activity recognition using WiFi can satisfy the requirement in an indoor environment.

We list a motion-sensing game using WiFi signals in Figure 14. One or two people are located in the middle of the transmission and receiving terminal and prolong the distance between the TV and user. The area below the blue dashed line represents the control area, and our work can sense human behavior within 10 m and achieve a better performance in the range of black circle. The user operates the same activity as well as the TV set, and receiving terminal collects corresponding data. By the phase of signals processing, we achieve an activity with the probability and match it with the game of TV set. Once the matching result satisfies the threshold value, activity recognition matches success in the motion-sensing game using WiFi signals.

10. Discussion and Future Work

10.1. Extending to Shadow Recognition. In our research, we consider the relationship between the WiFi signals and skeleton data on the human activity recognition. Moreover, we describe the interesting topic of the shadow activity recognition. Shadow is an important issue to vision-based activity recognition or monitoring; however, WiFi-based activity recognition can sense human behavior through wall or shadow. First, we explore the characteristics of CSI to enhance the sensing ability by using the high-precision device. Second, WiFi signals can help vision-based activity recognition to improve the ability of sensing environment. In this study, we also need to consider the material attenuation. According to our observations, there is a little difference between the impact of wall reflection and body reflection on the WiFi signals. WiVi [14] leverages the nulling technique to explore the through-wall sensing behavior by using CSI and

TABLE 4: Accuracy of activity for CSI-based and Kinect-based.

Activities	WiFi	KARD [3]	SSJ	HuAc
Horizontal arm wave	**90**%	**92**%	**100**%	**100**%
High arm wave	100%	96%	80%	95%
Two-hand wave	**93.1**%	**96**%	**100**%	**100**%
High throw	*90*%	*80*%	*100*%	*100*%
Draw x	100%	96%	100%	93%
Draw tick	*100*%	*90*%	*75*%	*93*%
Toss paper	*100*%	*90*%	*100*%	*100*%
Forward kick	**87**%	**96**%	**100**%	**100**%
Side kick	100%	100%	90%	100%
Bend	95.7%	96%	100%	100%
Hand clap	**92**%	**100**%	**100**%	**100**%
Walk	100%	100%	100%	100%
Phone	100%	96%	100%	100%
Drink water	*100*%	*86*%	*100*%	*100*%
Sit down	**90**%	**100**%	**70**%	**91**%
Squat	96.7%	100%	90%	90%

analyzing the offset of signals from reflection and attenuation of the wall. We recommend researchers to read this paper and their following work [11].

10.2. Extending to Multiple People Activity Recognition. Multiple people activity recognition needs multiple APs to obtain more signals information reflected by a human body. At present, existing works can locate target location [46] and detect the number [19] of multiple people using CSI in the indoor environment. Kinect-based activity recognition system recognizes two skeletons (six skeletons for Kinect 2.0) and locates skeletons of six people. Therefore, the combination of WiFi signals and Kinect facilitates the development of multiple people activity recognition. In the future, our team wants to deeply research the character of WiFi signals and propose a novel framework to facilitate the practical application of human activity recognition in the social lives.

10.3. Data Fusion. Skeleton data detect the position of each joint for each activity and track the trajectory of human behavior. CSI can sense a fine-grained activity without attaching device in the complex indoor environment. The balance point between CSI and skeleton joints and the selection method of effective features are important factors for improving the quality of fusion information. Moreover, time synchronization of fusion information is also an important challenge in the human activity recognition field.

11. Conclusion

In our work, we construct a WiFi-based public activity dataset named WiAR and design HuAc, a novel framework of human activity recognition using CSI and crowdsourced skeleton joints, to improve the robustness and accuracy of activity recognition. First, we leverage the moving variance of CSI to detect the rough start and end of an activity and adopt

the distribution of CSI to describe the detail of each activity. Moreover, we also select several effective subcarriers by using *K*-means algorithm to improve the stability of activity recognition. Then, we design SSJ method on the basis of KARD to recognize similar activities by leveraging spatial relationship and the angle of adjacent joints. Finally, we solve the limitations of CSI-based and skeleton-based activity recognition using fusion information. Our results show that HuAc achieves 93% of average recognition accuracy in the WiAR dataset.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensor-based activity recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 42, no. 6, pp. 790–808, 2012.

[2] C. Han, K. Wu, Y. Wang, and L. M. Ni, "WiFall: device-free fall detection by wireless networks," in *Proceedings of the 33rd IEEE Conference on Computer Communications (IEEE INFOCOM '14)*, pp. 271–279, Toronto, Canada, May 2014.

[3] S. Gaglio, G. Lo Re, and M. Morana, "Human activity recognition process using 3-D posture data," *IEEE Transactions on*

Human-Machine Systems, vol. 45, no. 5, pp. 586–597, 2015.

[4] H. Abdelnasser, K. A. Harras, and M. Youssef, "WiGest demo: a ubiquitous WiFi-based gesture recognition system," in *Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS '15)*, pp. 17-18, IEEE, Hong Kong, May 2015.

[5] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys*, vol. 46, no. 3, article 33, 2014.

[6] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, "Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey," in *Proceedings of the ARCS 2010*.

[7] J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced computer vision with microsoft kinect sensor: A review," *IEEE Transactions on cybernetics*, vol. 43, no. 5, pp. 1318–1334, 2013.

[8] K. Biswas and S. Basu, "Gesture recognition using Microsoft Kinect," in *Proceedings of the 5th International Conference on Automation, Robotics and Applications (ICARA '11)*, pp. 100–103, December 2011.

[9] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proceedings of the 20th ACM Annual International Conference on Mobile Computing and Networking, MobiCom 2014*, pp. 617–628, USA, September 2014.

[10] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, MobiCom 2015*, pp. 65–76, Paris, France, September 2015.

[11] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall ," *ACM Transactions on Graphics*, vol. 34, no. 6, article no. 219, 2015.

[12] J. Wilson and N. Patwari, "See-through walls: motion tracking using variance-based radio tomography networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 5, pp. 612–621, 2011.

[13] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via RF body reflections," in *Proceedings of the 12th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2015*, pp. 279–292, usa, May 2015.

[14] F. Adib and D. Katabi, "See through walls with WiFi!," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM '13)*, pp. 75–86, August 2013.

[15] B. Fang, N. D. Lane, M. Zhang, A. Boran, and F. Kawsar, "BodyScan: Enabling radio-based sensing on wearable devices for contactless activity and vital sign monitoring," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys 2016*, pp. 97–110, Singapore, June 2016.

[16] Z. Cheng, L. Qin, Y. Ye, Q. Huang, and Q. Tian, "Human daily action analysis with multi-view and color-depth data," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 7584, no. 2, pp. 52–61, 2012.

[17] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM '00)*, vol. 2, pp. 775–784, Tel Aviv, Israel, March 2000.

[18] J. Gjengset, J. Xiong, G. McPhillips, and K. Jamieson, "Phaser: Enabling phased array signal processing on commodity WiFi access points," in *Proceedings of the 20th ACM Annual International Conference on Mobile Computing and Networking, MobiCom 2014*, pp. 153–163, USA, September 2014.

[19] C. Xu, B. Firner, R. S. Moore et al., "Scpl: indoor device-free multi-subject counting and localization using radio signal strength," in *Proceedings of the 12th International Conference on Information Processing in Sensor Networks (IPSN '13)*, pp. 79–90, Philadelphia, Pa, USA, April 2013.

[20] K. Kleisouris, B. Firner, R. Howard, Y. Zhang, and R. P. Martin, "Detecting intra-room mobility with signal strength descriptors," in *Proceedings of the 11th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2010*, pp. 71–80, USA, September 2010.

[21] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: wireless indoor localization with little human intervention," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (Mobicom '12)*, pp. 269–280, August 2012.

[22] L. Sun, S. Sen, and D. Koutsonikolas, "Bringing mobility-awareness to WLANs using PHY layer information," in *Proceedings of the 10th ACM International Conference on Emerging Networking Experiments and Technologies, CoNEXT 2014*, pp. 53–65, Australia, December 2014.

[23] Y. Zeng, P. H. Pathak, and P. Mohapatra, "Analyzing shopper's behavior through WiFi signals," in *Proceedings of the 2nd Workshop on Physical Analytics, WPA 2015*, pp. 13–18, Italy.

[24] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using WiFi signals," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, MobiCom 2015*, pp. 90–102, France, September 2015.

[25] X. Zheng, J. Wang, L. Shangguan, Z. Zhou, and Y. Liu, "Smokey: Ubiquitous smoking detection with commercial WiFi infrastructures," in *Proceedings of the 35th Annual IEEE International Conference on Computer Communications, IEEE INFOCOM 2016*, USA, April 2016.

[26] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, "WiDraw: Enabling hands-free drawing in the air on commodity WiFi devices," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, MobiCom 2015*, pp. 77–89, France, September 2015.

[27] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with Wi-Fi!," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom '14)*, pp. 593–604, Maui, Hawaii, USA, September 2014.

[28] K. Qian, C. Wu, Z. Zhou, Y. Zheng, Z. Yang, and Y. Liu, "Inferring Motion Direction using Commodity Wi-Fi for Interactive Exergames," in *Proceedings of the the 2017 CHI Conference*, pp. 1961–1972, Denver, Colorado, USA, May 2017.

[29] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: indoor localization via channel response," *ACM Computing Surveys*, vol. 46, no. 2, article 25, 2013.

[30] X. Hu, T. H. S. Chu, H. C. B. Chan, and V. C. M. Leung, "Vita: a crowdsensing-oriented mobile cyber-physical system," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 1, pp. 148–165, 2013.

[31] V. Radu and M. K. Marina, "HiMLoc: indoor smartphone localization via activity aware Pedestrian Dead Reckoning with selective crowdsourced WiFi fingerprinting," in *Proceedings of the International Conference on Indoor Positioning and Indoor*

Navigation (IPIN '13), pp. 1–10, Montbeliard, France, October 2013.

[32] X. Hu, X. Li, E. C.-H. Ngai, V. C. M. Leung, and P. Kruchten, "Multidimensional context-aware social network architecture for mobile crowdsensing," *IEEE Communications Magazine*, vol. 52, no. 6, pp. 78–87, 2014.

[33] X. Hu and V. C. M. Leung, "Torwards context-aware mobile crowdsensing in vehicular social networks," *of IEEE ISCCGC*, 2015.

[34] B. Lu, Z. Zeng, L. Wang, B. Peck, D. Qiao, and M. Segal, "Confining Wi-Fi coverage: A crowdsourced method using physical layer information," in *Proceedings of the 13th Annual IEEE International Conference on Sensing, Communication, and Networking, SECON 2016*, UK, June 2016.

[35] Z. Ning, X. Hu, Z. Chen et al., "A Cooperative Quality-aware Service Access System for Social Internet of Vehicles," *IEEE Internet of Things Journal*, pp. 1-1.

[36] Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular Social Networks: Enabling Smart Mobility," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 16–55, 2017.

[37] Z. Ning, X. Wang, X. Kong, and W. Hou, "A Social-aware Group Formation Framework for Information Diffusion in Narrowband Internet of Things," *IEEE Internet of Things Journal*, pp. 1-1.

[38] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Proceedings of the 18th annual international conference on Mobile computing and networking (Mobicom '12)*, pp. 293–304, August 2012.

[39] Y. Sungwon, P. Dessai, M. Verma, and M. Gerla, "FreeLoc: calibration-free crowdsourced indoor localization," in *Proceedings of the 32nd IEEE Conference on Computer Communications (INFOCOM '13)*, pp. 2481–2489, Turin, Italy, April 2013.

[40] Z.-P. Jiang, W. Xi, X. Li et al., "Communicating is crowdsourcing: Wi-Fi indoor localization with CSI-based speed estimation," *Journal of Computer Science and Technology*, vol. 29, no. 4, pp. 589–604, 2014.

[41] Y. Chen, L. Shu, A. M. Ortiz, N. Crespi, and L. Lv, "Locating in crowdsourcing-based dataspace: Wireless indoor localization without special devices," *Mobile Networks and Applications*, vol. 19, no. 4, pp. 534–542, 2014.

[42] L. Cheng and J. Wang, "How can I guard my AP? Non-intrusive user identification for mobile devices using WiFi signals," in *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2016*, pp. 91–100, Germany, July 2016.

[43] Z. Zhou, Z. Yang, C. Wu, W. Sun, and Y. Liu, "LiFi: Line-Of-Sight identification with WiFi," in *Proceedings of the 33rd IEEE Conference on Computer Communications (INFOCOM '14)*, pp. 2688–2696, May 2014.

[44] Y. Wang, J. Yang, Y. Chen, H. Liu, M. Gruteser, and R. P. Martin, "Tracking human queues using single-point signal monitoring," in *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys 2014*, pp. 42–54, USA, June 2014.

[45] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A Real-Time and Contactless Fall Detection System with Commodity WiFi Devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, 2017.

[46] X. Guo, D. Zhang, K. Wu, and L. M. Ni, "MODLoc: Localizing multiple objects in dynamic indoor environment," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 11, pp. 2969–2980, 2014.

Composition Comments

1. Please check and confirm the author(s) first and last names and their order which exist in the last page.

2. We made the highlighted change according to the list of references. Please check.

3. We made the highlighted change(s) for the sake of clarity and correctness. Please check.

4. We rephrased the highlighted part(s) for the sake of clarity and correctness. Please check.

5. The highlighted part in the caption of Figure 5 is grammatically unclear/incorrect. Please rephrase this part for the sake of clarity and correctness.

6. We made the highlighted change(s) throughout the paper for the sake of clarity and correctness. Please check.

7. We assume that there is a missing part after the highlighted one. Please check.

8. We redrew some parts in Figures 3, 5, and 12. Please check.

9. We made the highlighted change(s) in the caption of Figure 11 for the sake of clarity and correctness. Please check.

10. We show hidden parts in Figures 9 and 13. Please check.

11. We rephrased the highlighted change(s) for the sake of clarity and correctness. Please check.

12. We removed the parts before the highlighted ones in caption of Figures 11 and 13 as per journal style. Please check.

13. We changed the gray background and the purple background in Table 4 to bold font and italic font, respectively, and their mention in the text as per journal style. Please check.

14. Comment on ref. [11]: We deleted reference [23] in the original manuscript, which was a repetition of [11]. Consequently we will replace all the citations of [23] within text with those of [11]. Please check all highlighted cases throughout.

Author(s) Name(s)

It is very important to confirm the author(s) last and first names in order to be displayed correctly on our website as well as in the indexing databases:

Author 1
Given Names: Linlin
Last Name: Guo

Author 2
Given Names: Lei
Last Name: Wang

Author 3
Given Names: Jialin
Last Name: Liu

Author 4
Given Names: Wei
Last Name: Zhou

Author 5
Given Names: Bingxian
Last Name: Lu

It is also very important for each author to provide an ORCID (Open Researcher and Contributor ID). ORCID aims to solve the name ambiguity problem in scholarly communications by creating a registry of persistent unique identifiers for individual researchers.

To register an ORCID, please go to the Account Update page (http://mts.hindawi.com/update/) in our Manuscript Tracking System and after you have logged in click on the ORCID link at the top of the page. This link will take you to the ORCID website where you will be able to create an account for yourself. Once you have done so, your new ORCID will be saved in our Manuscript Tracking System automatically.