Indoor sneezing and coughing detection based on COTS wireless device

Zhanjun Hao
Northwest Normal University
School of Computer Science and Engineering
LanZhou, China
zhanjunhao@126.com

Daiyang Zhang
Northwest Normal University
School of Computer Science and Engineering
LanZhou, China
0000-0002-7131-693X

Yu Duan
Xidian University
School of Computer Science and Technology
Xi'an, China
yuduaniot@126.com

Xiaochao Dang
Northwest Normal University
School of Computer Science and Engineering
LanZhou, China
dangxc@nwnu.edu.cn

***Abstract*—Coughing and sneezing are important routes of virus transmission. Droplets carrying the virus enter the air and spread rapidly, increasing the spread of the disease. Therefore, how to accurately detect coughing and sneezing behaviors in a timely manner so as to effectively warn the spread of the virus has become an urgent problem. To solve this problem, we designs a coughing and sneezing detection scheme for indoor people on commercial wireless devices. First, the Doppler shift feature image caused by the action is segmented using a clustering algorithm, which reduces the computational overhead of the system. Then, the HOG features of the segmented image are extracted and input to the two-dimensional SOM network for action classification and recognition, which effectively improves the detection accuracy of target actions. Finally, a dataset consisting of real coughing and sneezing actions is constructed and open-sourced in this paper. The performance of this solution was tested and analyzed in several dimensions under two typical application scenarios. The results show the robustness of this scheme and the accuracy up to 93.1% in real-world scenarios. Our solution offers a new technology and method for disease prevention detection.**

Keywords—Wireless Sensing, Cough and sneeze detection, Device-free, CSI

I. INTRODUCTION (*HEADING 1*)

Sneezing and coughing is a coordinated protective respiratory reflex in humans[1]. When a person sneezes or coughs, the virus enters the air with droplets and is spread[2]. According to Venugopal Arumuru et al[3]. The main route of transmission of COVID-19 infectious agents is through droplets emitted during coughing and sneezing. And it has been observed that unobstructed droplets in air can reach 25 feet in 22 seconds in a stationary environment. Even if blocked, droplets can still travel 2.5 feet. Therefore, the detection of human sneezing and coughing can effectively lock the spread of the virus, so as to cooperate with professional personnel for effective monitoring and timely warning.

In the past few years, there have also been numerous studies on coughing and sneezing movements. Usually they use the distinctive sound of sneezing and coughing and detect the occurrence of sneezing and coughing by the corresponding audio. Sneezing and coughing can also be effectively detected by having people wear relevant motion sensors. However, sound-based detection methods are limited by the environment and do not work effectively in noisy environments. Wearable sensor-based detection, on the other hand, requires the person being detected to wear the relevant sensor, which violates the user's privacy and causes inconvenience to the user.

The advent of Wi-Fi technology overcomes the limitations of the above technologies. Wi-Fi devices are inexpensive and have three main characteristics: wireless, sensorless, and contactless[4]. In addition, the ubiquitous Wi-Fi signal gives a solid signal foundation for Wi-Fi-based wireless sensing. And the Channel State Information (CSI) extracted from the Wi-Fi signal captures the wireless characteristics of the nearby environment, allowing for a more fine-grained portrayal of human behavior[5]. These advantages make human motion sensing based on Wi-Fi signals one of the hot spots of current research. However, there are few studies on sensing human coughing and sneezing based on Wi-Fi signals. Most of the relevant studies have been conducted to determine the fatigue state of the human body by detecting movements such as yawning[6].Or sleep quality monitoring by detecting the body's breathing status[7].

In this paper, we propose a scheme to implement indoor person sneezing and coughing detection on COTS wireless devices. Detects coughing, sneezing and yawning movements of people in the room using CSI extracted from the Wi-Fi signal. Specifically, there are three challenges in this paper: firstly, coughing and sneezing are not routine actions and there is no publicly available dataset for validation. Secondly, for the same action, when the action occurs at different heights there will be different changes, so how to adapt the device to all heights of action is also one of the current challenges. Finally, since coughing and sneezing are extremely similar to actions such as yawning, it is necessary to find a way to distinguish them effectively.

We implemented this solution on a commercial Wi-Fi device, using Doppler shift to distinguish different actions, and applying k-means clustering to segment the Doppler images and extract the HOG features of the segmented images to make the feature information of different actions more obvious. Finally, a self-organizing neural network (SOM) is introduced to classify the different actions. Specifically, the contributions of this paper are as follows:

- We propose a method for detecting human coughs and sneezes on commercial wireless devices, providing a low-cost, feasible early warning solution for virus prevention and control.
- In this paper, the Doppler image of the action is segmented using a clustering algorithm, and its HOG

This work was funded by National Natural Science Foundation of China (61762079). Gansu Province Science and Technology Support Key R&D Program Project (20YF8GA048)2019 Chinese Academy of Sciences "Light of the West" Talent Program.

features are extracted and then processed using a SOM network. This method improves the recognition accuracy of actions while effectively distinguishing different state actions.

- we obtain real data of sneezing and coughing from several groups of volunteers in two scenarios close to the application and build an open-source dataset by relying on commercial wireless devices. By verifying in real scenarios, the method obtains 93.1% action recognition accuracy with high robustness.

The rest of this paper is organized as follows. The related technical studies are presented in Section 2, and the basic theoretical model of this paper is described in Section 3. Section 4 shows the detailed design of the scheme in this paper, followed by experiments and evaluation of this scheme in Section 5, and finally concludes the paper in Section 6.

II. Related Work

In this section we present the current research on human-related motion sensing from both non-Wi-Fi based and Wi-Fi based aspects.

A. Non-Wi-fi Based

For coughing, sneezing and other movements detection is currently mainly based on audio detection and sensor-based detection. The audio-based detection technique detects the action by comparing the sound signal of normal breathing with that of coughing. Madhurananda Pahar et al.[8] used the bandwidth and characteristics of cough sounds that are completely different from the complete speech signal for cough detection. Extracting the MEL frequency spectrum coefficients of the speech signal as features to analyze the audio, and enhancing the sound features to achieve better detection results. Renard Xaviero Adhi Pramono et al. [9] used a logistic regression model to classify and detect cough sounds in audio and determine whether a toddler has whooping cough by the detected cough. But the sound signal is very easy to be disturbed by noise or other sounds, resulting in the captured audio signal being unusable. So it can only be used in the more quiet scenes, can not adapt to complex scenes. Sensor-based detection technology mainly uses professional sensors or sensors that come with the phone to detect relevant movements. Helia Mohammadi et al.[10] achieved better experimental results by extracting biaxial acceleration signals from a single accelerometer deployed in the patient's neck and using support vector machine for classification to detect patient's coughing, swallowing, and tongue movements. Sidi Ahmed Mahmoudi et al.[11] could provide effective cough detection and classification by deploying multiple sensors including ECG, thermistor, chest strap, accelerometer, oximeter, audio microphone, etc. However, most of the sensor-based detection technologies require users to carry sensor devices with them, which causes inconvenience to users and violates their privacy.

B. Non-Wi-fi Based

Currently there are fewer technologies to detect coughing, sneezing and other motions based on Wi-Fi technology, and most of the research is to detect the user's breathing, yawning and other related actions. Hashim Saeed et al.[12] used CSI information extracted from in-vehicle Wi-Fi signals to detect behaviors such as driver yawning as a way to determine whether a driver is driving fatigued or not. Better detection results were achieved by conducting experiments in real vehicles. Muhammad Bilal Khan et al.[13] proposed a contactless sensing platform. The platform uses CSI information to detect abnormal breathing, coughing and other movements in humans to help doctors diagnose Covid-19 symptoms early and monitor human activity and health during an outbreak quarantine. Alexander Tataraidze et al.[14] used CSI information extracted from Wi-Fi signals connected to cell phones and routers to detect human respiration. Sleep monitoring of the human body by detecting the breathing status of the three sleeping positions. Abdelwahed Khamis et al.[15] proposed WiRelax. WiRelax can map changes in channel state information to transient breathing states and can correlate the relative phase of the received signal with the movement of the chest during human breathing. Qinghua Gao et al.[16] designed DAM, a two-dimensional Doppler AOA map constructed by a super-resolution method based on CSI signals, and by clustering and analyzing DAM can monitor the respiratory rate of three people simultaneously. All of the above studies demonstrate that Wi-Fi-CSI-based cough and sneeze detection is theoretically feasible and has considerable advantages over audio and sensor-based technologies.

III. Preliminary

This section introduces the model of human perception and decomposes the perceptual actions, followed by a detailed exposition of the theory related to channel state information and Doppler shift.

A. Human Effect Model

In this paper, six actions were selected for validation: coughing, sneezing, yawning, cover mouth to cough, cover mouth to sneeze, and cover mouth to yawn. According to the literature [2], the coughing action can be divided into four stages: (a) Irritation phase. (b) Inspiration phase. (c) Compression phase. (d) Expulsion phase. After extracting the effective actions from them, the coughing action can be decomposed into four steps as shown in Figure 1. First, the body performs an inhalation action, with the chest protruding forward, followed by a rapid compression of the chest, forcing the internal gas to flow out quickly, while the head begins to move to form a coughing action, and finally a cycle of the above actions is performed until the coughing action stops.

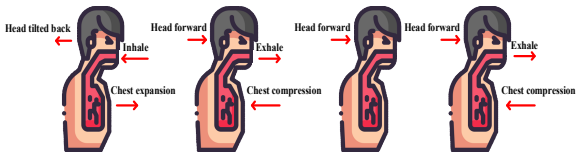


Fig. 1. Breakdown of coughing.

The action process of sneezing is described in detail in the literature [1]. The sneezing action can be broken down into two stages: inspiration and expiration. After extracting the effective action as shown in Figure 2. In the first stage, the chest protrudes to inhale a large amount of gas and the head is tilted back, and in the second stage, the chest is squeezed while the head moves forward rapidly to exhale gas to form a sneezing action.

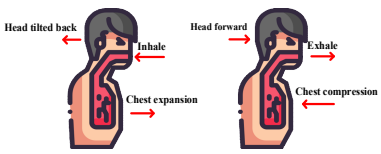


Fig. 2. Breakdown of sneezing .

Although the duration of both coughing and sneezing is brief, unlike the one-time process of sneezing, coughing may occur with multiple exhalations in one inhalation. Yawning is similar to coughing and sneezing in terms of action flow but has a longer duration, which makes the effect of the three actions on the CSI signal different.

B. Channel State Information and Doppler Shift

In wireless communication, CSI is physical layer information with fine-grained properties that describes the amplitude and phase information of each subcarrier to express the channel characteristics. As shown in Figure 3(a), each channel can be divided into 30 subcarriers and each subcarrier has a different amplitude [17]. In addition, CSI can capture the wireless characteristics of the nearby environment and is highly resistant to environmental noise. So CSI can be specifically expressed as:

$$H(f,t)=\sum_{i}^{N}a_i(t)e^{-j2\pi f\tau_i(t)} \tag{1}$$

where a_i is the amplitude decay of the i th path, N is the total number of propagation paths, $a_i(t)$ and $\tau_i(t)$ are the complex decay factor and the flight time of the i th path [18].

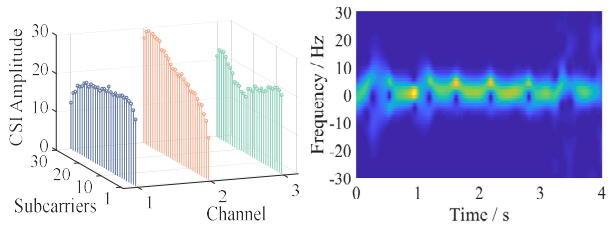


Fig. 3. CSI signal and Doppler shift.(*Figure (a) is the CSI subcarrier with 3 channels. Figure (b) is the Doppler shift*)

According to the literature [19], the Doppler shift is the change in the oscillation frequency of the reflected signal, as shown in Figure 3(b). Usually, the Doppler frequency shift reflecting the object signal can be expressed as:

$$F_D=-\frac{1}{\mu}\frac{d}{dt}d(t) \tag{2}$$

where μ denotes the wavelength of the signal and $d(t)$ is the length of the reflected path. The Doppler shift of the CSI signal can be extracted using the above equation. According to Eq. 1 and Eq. 2, the channel response can be expressed by the Doppler shift on each path as:

$$D(f,t)\approx H_s(f)+\sum_{i\in H_d}a_i(t)W(F_{D_i}(t)) \tag{3}$$

where $H_s(f)$ is the sum of static path responses, H_d is the set of dynamic paths, and $W(F_{D_i}(t))$ is the window function that cuts the signal segment of interest. The Doppler shift of the CSI signal can be extracted using the above equation.

IV. Method Design

In this section, the paper first introduces the architecture of the entire method, followed by a detailed description of each part of the architecture.

A. system overview

The system proposed in this paper is divided into four main parts: data acquisition, data processing, feature extraction, and action classification. The system flow is shown in Figure 4.

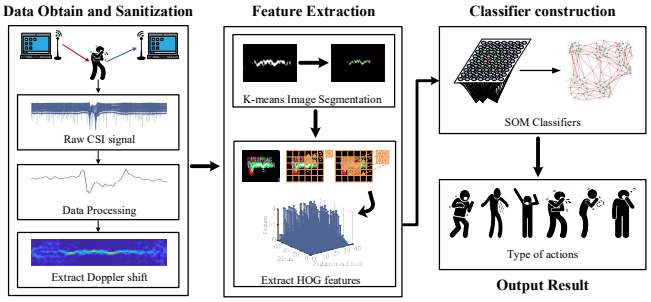


Fig. 4. System flow chart.

In the data obtaining and processing section, two commercial laptop devices equipped with Intel 5300 NICs are used for communication, one as the transmitter and the other as the receiver. The receiver device records and stores the raw CSI action signals. The raw CSI data is then sanitized and converted to Doppler shift. First, the raw CSI action data is denoised using Butterworth low-pass filtering and Gaussian filtering. Then the optimal subcarriers are selected by Principal Component Analysis (PCA) algorithm, and finally the selected subcarriers are converted to Doppler shifts.

The feature extraction stage uses the k-Means algorithm to segment the part of the Doppler frequency shift images where the action occurs, after which the HOG features of the segmented image are extracted to form the feature vector of the action.

The action recognition stage inputs all action feature arrays as feature sample sets into the SOM network for classification and recognition, and finally outputs the action recognition results.

B. Data Obtain and Sanitization

In this paper, we collect raw CSI action data from Intel 5300 NIC chip, set one transmitting antenna and three receiving antennas to receive data from three channels in total.

In order to reduce the data complexity, one channel that is most sensitive to the action needs to be selected as a representative. Figure 5(a) shows the sensitivity of the three channels to the action, and it can be seen from Figure 5(a) that the second channel is more sensitive to the action compared to the other two channels. Due to the multipath effect, the original CSI signal contains a large number of outliers and ambient noise that mask the real signal changes caused by the action.

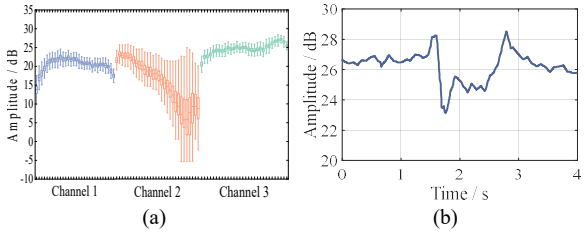


Fig. 5. Data Processing (*Figure (a) is the sensitivity of the three channels. Figure (b) is the best subcarrier extracted by the PCA algorithm*)

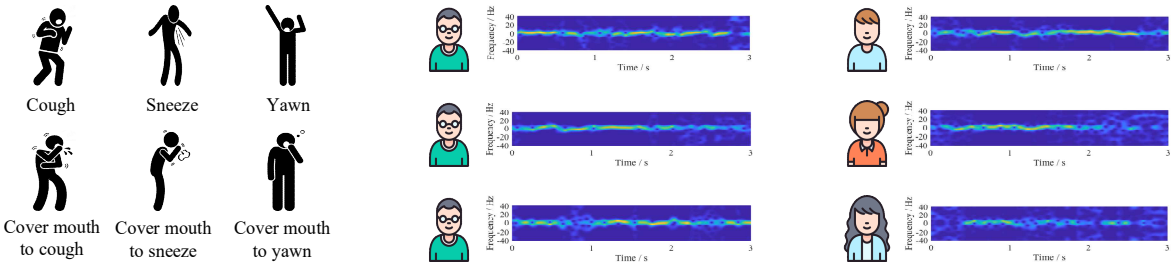


Fig. 6. Acquisition of action schematic and action-induced Doppler shift（*Figure (a) shows the diagram of the 6 movements. Figure (b) shows the Doppler frequency shift of different movements of the same person, and Figure (c) shows the Doppler frequency shift of the same movement of different people.*）

Therefore, in this paper, the Butterworth low-pass filter as well as the moving Gaussian filter are selected. In addition, to further reduce the data complexity, the optimal subcarrier of this channel is selected using the PCA algorithm in this paper, and the processed action signal is shown in Figure 5(b).

Finally, the selected subcarriers are converted into Doppler shift images. Figure 6(a) shows the schematic of the six actions, Figure 6(b) shows the Doppler shift of different actions of the same person, and Figure 6(c) shows the Doppler shift of different people doing the same action.

From the Figure 6, it can be seen that although the human body does not move much during coughing and sneezing, it can still cause significant Doppler shift changes, and the Doppler shift caused by different movements is obviously different, while the Doppler shift caused by the same movement has some similarity Feature Extraction

In order to make the correlation of actions more prominent to improve the classification results the correlation feature of Doppler shift needs to be extracted. In this paper, we first use the k-Means algorithm to segment the image and separate the part with action-induced Doppler frequency shift from the original image. Since the image with Doppler shift belongs to RGB mode, the pixels of the original image can be classified into 3 categories using the calculated Euclidean distance.

The Histogram of Oriented Gradient (HOG) feature is a feature descriptor used in image recognition for object detection[20]. Construct features by computing and counting histograms of gradient directions in local regions of the image. Compared with other feature description methods, HOG features can maintain good invariance to both geometric and optical deformations of motion Doppler shift images, so this paper extracts HOG features of motion Doppler shift images. The process of action segmentation and HOG feature extraction is shown in Figure 7.

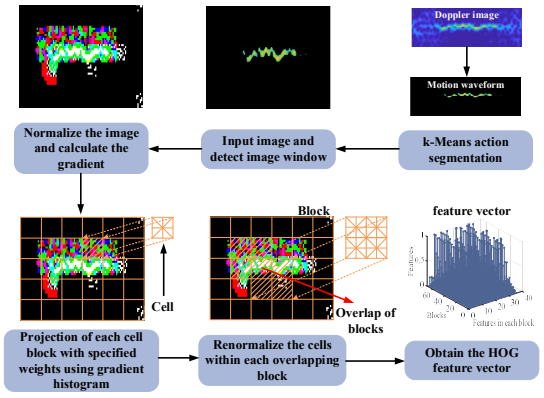


Fig. 7. HOG feature extraction

The image is first converted to a grayscale map and the input image is normalized using the Gamma correction method, and the Gamma compression formula is as follows.

$$G(x,y) = G(x,y)^{Gamma} \tag{4}$$

where $G(x,y)$ denotes the input image and $Gamma$ takes the value of 0.4. The gradient in the horizontal and vertical coordinate directions of the image is calculated after normalization, and the gradient direction value of each pixel position is calculated accordingly. So the gradient of each pixel point in the figure is:

$$G_x(x,y) = P(x+1,y) - P(x-1,y) \tag{5}$$

$$G_y(x,y) = P(x,y+1) - P(x,y-1) \tag{6}$$

where $G_x(x,y)$ is the horizontal gradient of the input image at point (x,y), $G_y(x,y)$ is the vertical gradient, and $P(x,y)$ is the pixel value at that point. So the amplitude of the gradient at point (x,y) and the direction of the gradient are:

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \tag{7}$$

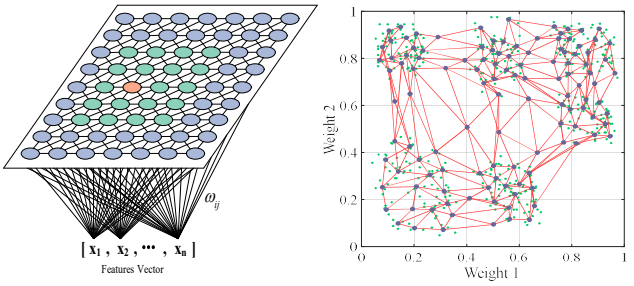
$$\varpi(x,y) = \tan^{-1}\left(\frac{G_y(x,y)}{G_x(x,y)}\right) \tag{8}$$

The HOG characteristics of each block can be derived from the above equation. Since each block overlaps each other, this causes the feature vector of each cell to appear multiple times in the final feature vector with different results. So finally the HOG features of so overlapping blocks in the whole image need to be collected and form the final feature vector for that action.

C. Motion Recognition

Self-organizing map network is an unsupervised artificial neural network that competes neurons with each other through a competitive learning strategy for progressive optimization [21]. Since the SOM network is very similar to the biological neural system, it works well for different dimensional action feature datasets. Therefore, this paper uses SOM network to classify the actions.

The SOM network used in this paper consists of an input layer and a competing layer with bidirectional connection of neurons between the two layers and no hidden layer, and its structure is shown in Figure 8(a). The competitive layer is the core of the SOM network. In this paper, we use the competitive layer of two-dimensional planar line array and use



(a) (b)

Fig. 8. SOM Network. (*Figure (a) shows the structure of SOM network, and Figure (b) shows the classification result of SOM network*)

Kohonen learning algorithm. First, all neurons in the competitive layer are normalized to obtain $\omega_i(i=1,2,.....m)$. A random value between 0 and 1 is given as the initial weight of this neuron, an initial superior neighborhood N_j is established, and an initial value is assigned to the learning rate. where, m is the number of competing layer neurons, and its minimum value is calculated according to Equation 9 and rounded upward:

$$m_{\min}=5\sqrt{L} \tag{9}$$

Where L is the number of training samples, in this paper, 120 sets of data are collected for each action, so the action feature data set is 600 sets in total, so the minimum number of competing layer neurons is 123.

After that, one input pattern is randomly selected from the training set and normalized to obtain $\hat{X}=\{x_i \ i=1,2.....s\}$. s is the number of neurons in the input layer. Each neuron in the competitive layer is traversed and the Euclidean distance of the input data from each neuron is calculated.

$$d_j=\| \hat{X}-\hat{\omega}_i \| \tag{10}$$

The neuron with the smallest Euclidean distance d_j from the input data is selected as the winning node, and the nodes that will be included in the winning neighborhood are determined according to the neighborhood radius δ and their respective updated magnitudes are calculated by the Neighborhood function, and a Gaussian function is chosen to represent the relationship between the influence strength and distance in the winning neighborhood. After determining the Neighborhood of Superiority, the weights of the nodes in the Neighborhood of Superiority are updated.

$$\omega_{ij}(t+1)=\omega_{ij}(t)+\mu(t,N)[x_i^p-\omega_{ij}(t)] \tag{11}$$

where $\omega_{ij}(t)$ is the weight of neuron i from 0 to n at moment j. $\mu(t,N)$ is a function of the training time and the topological distance N between the j th neuron and the winning neuron in the neighborhood. the SOM network is characterized by a gradual decrease in the neighborhood range and learning rate as the iteration progresses. In this paper, we set the output of the final SOM classifier to stop iterating when the learning rate is 0.

Figure 8(b) shows the SOM classifier trained using a sample set composed of action feature vectors. Since the SOM network is unsupervised learning so the samples do not need to be labeled, the initial number of neurons is set to 123. From the results in the figure, it can be seen that the trained SOM classifier can classify the samples into 6 categories, each category corresponds to an action, and even two similar actions can be distinguished more effectively.

V. Experimentation and Evaluation

In this section, we first describe the experimental environment, experimental equipment and other parameters, and in the subsequent subsections, we present our experimental and evaluation results in detail.

A. Experimental setup

In this paper, two Thinkpad x201i laptops loaded with Intel 5300 NICs are used as experimental devices, and both devices are installed with CSI-Tool[22] for transmitting and receiving CSI signals. In order to ensure the communication quality, both laptops are equipped with extended antennas, each with a signal gain of 6dBi, and one transmitting antenna and three receiving antennas are set in this paper.

In this paper, the experimental environment is set up as an empty hall and a classroom in two different scenarios, and the experimental scenario is shown in Figure 9.

(a) (b) (c)

Fig. 9. Experimental equipment and experimental environment.(Figure (a) shows the experimental equipment. Figure (b) shows the hall environment. Figure (c) shows the classroom environment)

The equipment was arranged in the two scenes shown in Figure 9 separately in order to conduct the experiments, and the resulting multipath effects were smaller and had a lower impact on the CSI signal due to fewer interferents in the hall. The conference room, on the other hand, contains more interferents and has an increased impact on the CSI signal. In addition, a total of 5 volunteers, 3 males and 2 females, were selected for this paper. Each person collected 120 sets of action data at different time periods, and a total of 600 action data constituted the sample set. The specific information of the volunteers is shown in Table 1.

In order to make the collected data more random, this paper collects data from different people at different time periods between May 29th and June 11th. To further ensure the randomness of the collected data, the same actions of some volunteers were also divided into two different time periods for collection. In addition, since the occurrence of sneezing

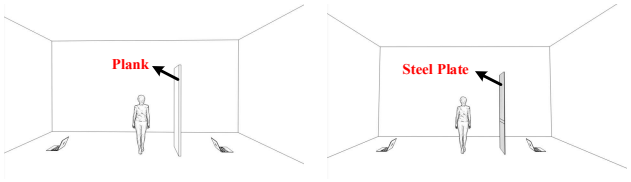
TABLE I. Data Collection Summary

Volunteer	Age	Height/Weight	Data Collection Date	Motion duration	Number of samples
Volunteer 1	24	186cm/110kg	May.29,2021;Jun.07,2021	3h	120
Volunteer 2	22	178cm/72kg	Jun.11,2021	1.5h	120
Volunteer 3	22	172cm/68kg	Jun.11,2021	1.5h	120
Volunteer 4	26	163cm/55kg	Jun.09,2021;Jun.10,2021	2h	120
Volunteer 5	26	159cm/53kg	Jun.07,2021	1.5h	120

action is random, in order to make the collected action data more realistic, this paper uses safe irritants such as pepper to trigger the real sneezing action of the experimenter. In order to make the experimental results more experimenter. In order to make the experimental results more concise, Cover mouth to cough, Cover mouth to sneeze and Cover mouth to yawn are abbreviated as: CMC, CMS, and CMY respectively in this paper.

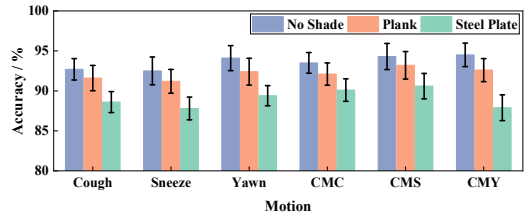
B. Impact of environments and Interfering items

In order to verify the robustness of this scheme to the environment and obstacles, the experiments are conducted in 2 scenarios, the experimental scenarios are an empty hall and a classroom. And in the actual situation, the movement of the personnel will be obscured by the indoor distractions, so as shown in Figure 10 this paper set up in both scenes, such as the thickness of 2cm of wood and steel panels of common indoor materials as a mask. The experimental results are shown in Figure 11.

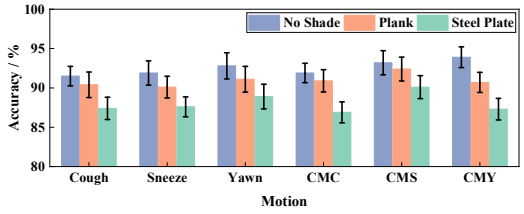


(a) (b)

Fig. 10. Different obstacle placement schematic .(*Figure (a) is the obstacle is a plank Figure (b) is the obstacle is a steel plate*)



(a)



(b)

Fig. 11. The influence of obstacles in different environments .(*Figure (a) shows the results for the lobby environment Figure (b) shows the results for the classroom environment*)

Figures 11(a) and 11(b) show the experimental results for the hall and classroom, respectively. From the environmental comparison, it can be seen that the action recognition result of the hall is higher than that of the classroom, which is due to the presence of a large number of interference items in the classroom and the enhanced multipath effect of the signal, resulting in a decrease in the action recognition rate. For different occlusions, the action recognition rate under different occlusions is above 85% from the experimental results. Since the signal is partially absorbed when it passes through the steel plate, it is reasonable that the recognition rate of the system is lower than that of the wooden plate when the obscuring object is the steel plate. From the comprehensive results, both for different environments and different occlusions, this solution shows better performance.

C. Impact of user diversity

We explores the impact of user diversity in three dimensions: different personnel, the magnitude of the personnel's movements, and the orientation of the movements. In this paper, an additional 3 male and 3 female volunteers, totaling 6 volunteers, were recruited to test this scheme, first performing 6 actions in a hall environment to validate the action recognition rate. Next, six volunteers were allowed to perform all movements according to different movement amplitudes, and finally six volunteers were allowed to perform the experiment with six movements in four directions as shown in Figure 12(a). The results of the three experiments are shown in Figure 12.

The results in Figure 12(b) show that the system has a high recognition probability for the actions performed by all six experimenters. As can be seen in Figure 12(c), the average recognition result is higher when the body is oriented perpendicular to the LOS path, which is due to the fact that when the body is perpendicular to the LOS path, the movement cuts the signal propagation path and therefore causes a more pronounced signal change. Figure 12(d), on the other hand, shows that strenuous movements do increase the accuracy of action recognition due to the greater impact of strenuous movements on the signal, but even for minor movements, the system still has an action recognition rate of over 90%. From the combined results, the system has high robustness to all personnel diversity.

D. System boundary detection

In order to further verify the performance of this scheme, this paper explores the boundaries of this scheme in three aspects: personnel interference, detection distance, and inter-travel action recognition. Firstly, one experimental person is allowed to perform the relevant actions, and then one interfering person, one, two, and three persons doing slightly interfering actions are set up respectively, and the experimental results are shown in Figure 13(a). Secondly, this paper explores the maximum detection distance of this program, and sets the initial interval of two devices in two environments to 0.5m, and the height is set to 1.7m, and then measures every 0.5m and records the results for comparison, and the comparison results are shown in Figure 13(b). Finally, this paper tested the action recognition between walking, this paper let the experimenter walk at a uniform speed first, and six actions were performed and detected during the walking process to verify whether the system could detect sneezing, coughing and other actions while keeping walking. The experimental results are shown in Figure 13(c).

Figure 13(a) shows that when the interferers are stationary, the action recognition rate of the system is basically unaffected. When the number of interferers rises the action recognition rate of the system also decreases. When there are 1-2 interferers, the average recognition rate of action remains above 75%, and when the number of interferers reaches 3, the system can still maintain an average recognition rate above 60%, and the experimental results prove that the system can resist the interference of personnel more effectively. The result of Figure 13(b) shows that when the equipment spacing is about 1.5m, the average action recognition rate reaches 93.1%, and with the increasing distance, the action recognition rate decreases continuously,

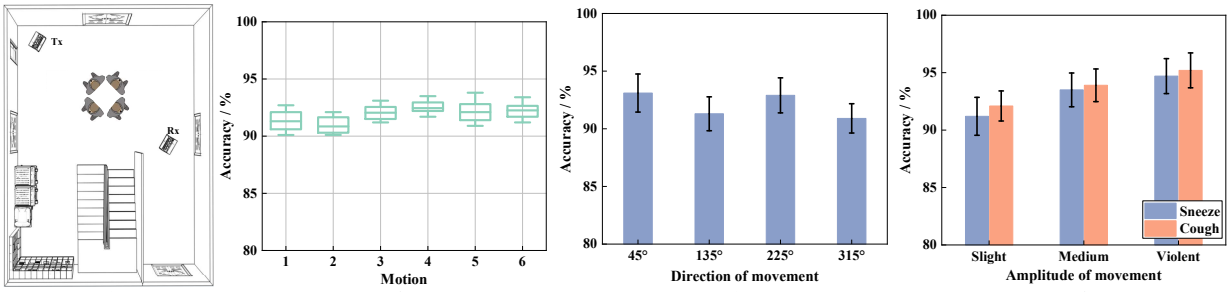


Fig. 12. The impact of personnel diversity.(*Figure (a) shows the orientation of people. Figure (b) shows the motion recognition rate of different people. Figure (c) shows the motion recognition rate of different orientations. Figure (d) shows the motion recognition rate of different magnitudes.*)

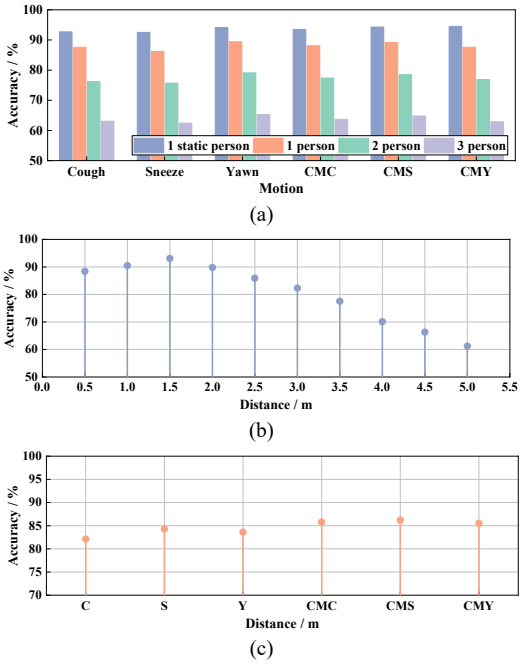


Fig. 13. System Boundary Exploration. (*Figure (a) shows the motion recognition rate in the presence of disturbed people. Figure (b) shows the test of detection distance. Figure (c) shows the motion recognition rate between people moving.*)

and when the equipment spacing is about 5m, the action recognition rate decreases to 61.2%. From the experimental results, the action recognition probability can still be maintained above 80% when the equipment spacing is around 3m, and remains above 60% when the equipment spacing reaches 5m, indicating that the system can maintain a better performance within 5m. From the results of Figure 13(c) can be seen for the action detection results between the marching can be seen, after adding the interference of the human body's walking causes the average action recognition rate to drop, because this paper let the volunteers walk at a uniform speed, so still the action recognition rate still remains above 80%. According to the analysis of the experimental results, the action recognition rate decreases significantly if the volunteer performs irregular motion or strenuous motion. However, the experimental results validate that it is feasible to recognize specific movements for a human body in motion.

E. Cross-Platform Testing

In order to verify the portability of the system, experiments were also conducted on two other experimental platforms of Atheros NICs in this paper. They are Atheros AR9580 NIC platform and Atheros AR9380 NIC platform. The action data collected from the two Atheros platforms are identified after the same processing and the experimental results are compared with those of Intel 5300 platform. The NIC parameters and the comparison results are shown in Figure 14.

From the experimental results, it can be seen that the system has good recognition results for all actions on all three NIC platforms, and the system can effectively differentiate each action on the other two Atheros NIC platforms as well. The experimental results show that this solution is less restricted by the NIC platform and has strong portability.

F. Impact of different classification algorithms

The performance of the classification algorithm is also an important indicator of the capability of this scheme. To verify the performance of the classification algorithm, the data were processed using the method proposed in this paper and then processed with random forest (RF), support vector machine (SVM), and convolutional neural network (CNN), and the classification results were compared with ROC curves. The comparison results are shown in Figure 15(a).

From the comparison results, we can see that when the True Positive Rate (TPR) reaches 0.8, the False Positive Rate (FPR) of SOM is only 0.03, which has the best performance. the FPR of CNN is 0.11, which is slightly lower than that of SOM. the FPR of SVM is 0.43, which is slightly higher than that of RF. while the FPR of RF reaches 0.61, which has the worst performance. In summary, SOM has the best performance, higher than the CNN algorithm. rf has the worst performance. Compared with these three algorithms, SOM has better performance.

G. Exploration of system performance parameters

In order to further evaluate the system performance, this paper conducts a comprehensive investigation in terms of two parameters: the data processing method and the number of neurons. In the feature extraction stage, this paper divides the data into three parts, one part is subjected to feature extraction according to the method proposed in this paper, another part extracts the HOG features directly without the k-means algorithm to segment the actions, and the last part extracts the grayscale co-occurrence matrix as features after using k-means for action segmentation. The features of the remaining two parts of the data are used to form a sample set to train a corresponding classifier again, and the three parts of the data are input into the corresponding classifier and tested with different numbers of neurons. The experimental results are shown in Figure 15(b). Since the number of samples is 600, the minimum number of neurons in the competitive layer is 120 according to Equation 9, so the experiment starts with 120 neurons for testing. From the experimental results, it can be seen that the recognition results gradually increase with the

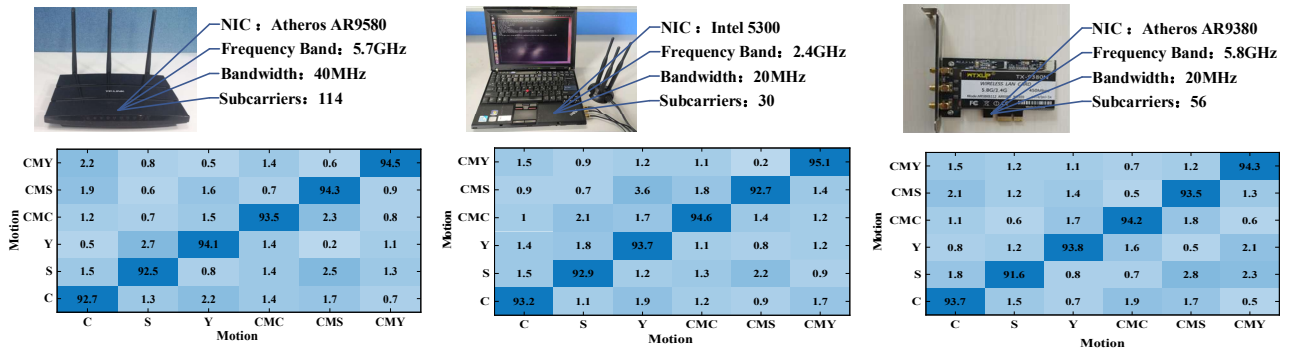


Fig. 14. Cross-platform experimental results. (*Figure (a) shows the device parameters and experimental results of the Atheros AR9580 NIC. Figure (b) shows the device parameters and experimental results of the Atheros AR9580 NIC. Figure (c) shows the device parameters and experimental results of the Atheros AR9580 NIC.*)

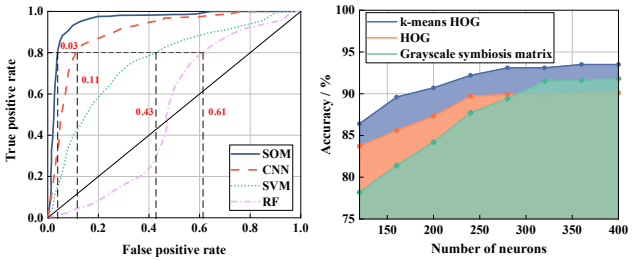


Fig. 15. Performance Evaluation Results. (*Figure (a) shows the comparison results of different classification algorithms. Figure (b) shows the effect of different feature extraction methods and the number of neurons.)*

increase of the number of neurons in the three processing cases and finally stabilize. In addition the results of both HOG and Grayscale symbiosis matrix are lower than k-means HOG, indicating that the method proposed in this paper is due to the remaining two methods. In addition, the average action recognition rate reaches an optimal 93.1% when the number of neurons reaches 280 or more, and when the number of neurons is greater than 280, although the average action recognition rate increases slightly. However, as the number of neurons increases, the processing time of the system also increases gradually, so under comprehensive consideration, it is optimal when the number of neurons is 280.

VI. Conclusion

This paper presents a scheme to implement human coughing and sneezing detection on a commercial wireless device. The CSI signal of human motion is first collected on the commercial wireless device and the original signal is sterilized and processed, after which the Doppler shift of the motion CSI signal is extracted. The action Doppler images are segmented and their HOG features are extracted as feature vectors using the k-Means algorithm. The feature vectors of all actions form a dataset to train the SOM network as an action classifier to output action recognition results. In this paper, various experiments are conducted to validate the system in a real environment, and the experimental results show that the proposed system in this paper has strong robustness and high performance. In future work, this paper will further investigate how to detect the coughing and sneezing behavior of human in action and upgrade the scheme in terms of applicability.

References

[1] Songu, Murat, and Cemal Cingi, “Sneeze Reflex: Facts and Fiction,” Therapeutic Advances in Respiratory Disease, 2009, pp. 131–141.

[2] Umayahara Y, Soh Z, Sekikawa K, Kawae T, Otsuka A, Tsuji T. “Clinical Significance of Cough Peak Flow and Its Non-Contact Measurement via Cough Sounds: A Narrative Review,” Applied Sciences. 2020, vol. 10, no. 8, pp. 2782.

[3] Venugopal Arumuru, Jangyadatta Pasa, and Sidhartha Sankar Samantaray, “Experimental visualization of sneezing and efficacy of face masks and shields,” Physics of Fluids, 2020,vol. 32, pp. 1-11.

[4] Z Yang, Z Zhou, Y Liu. “From RSSI to CSI: Indoor Localization via Channel Response,” ACM Computing Surveys, 2013, vol.46, no.2, pp. 1-32.

[5] Hao Z , Duan Y , Dang X , et al., “Wi-SL: Contactless Fine-Grained Gesture Recognition Uses Channel State Information,” Sensors, 2020, vol. 20, no.14 pp. 1-25.

[6] Zain Ul Abiden Akhtar, Hongyu Wang, “WiFi-based driver’s activity recognition using multi-layer classification,” Neurocomputing, 2020, vol. 405, pp. 12-25.

[7] F. Zhang, C Wu, B Wang, et al., "SMARS: Sleep Monitoring via Ambient Radio Signals," IEEE Transactions on Mobile Computing, 2021, vol. 20, no. 1, pp. 217-231.

[8] Tusar Kanti Dash, Soumya Mishra, Ganapati Panda, et al., “Detection of COVID-19 from speech signal using bio-inspired based cepstral features,” Pattern Recognition, 2021, Vol. 117, pp.1-13.

[9] Pramono RXA, Imtiaz SA, Rodriguez-Villegas E A, “Cough-Based Algorithm for Automatic Diagnosis of Pertussis,” PLoS ONE, 2016, vol. 11, no. 9 pp. 1-20.

[10] H Mohammadi, Ali-Akbar Samadani, C Steele, et al., “Automatic discrimination between cough and non-cough accelerometry signal artefacts,” Biomedical Signal Processing and Control, 2019, Vol. 52, pp. 394-402.

[11] Sidi Ahmed Mahmoudi, Paulo Da Cunha Possa, Thierry Ravet, et al., “Sensor-based System for Automatic Cough Detection and Classification,” 8th ICT Innovations Conference, pp. 1-10, September 2016. *(references)*

[12] H. Saeed, T. Saeed, M. Tahir and M. Uppal, “Risky Driving Behavior Detection Using In-Vehicle WiFi Signals,” 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall) Chicago, IL, USA, pp. 1-5, Aug 2018. *(references)*

[13] Khan M B , Zhang Z , Li L , et al. “A Systematic Review of Non-Contact Sensing for Developing a Platform to Contain COVID-19,” Micromachines, 2020, vol.11, no. 10, pp. 912-935.

[14] A. Tataraidze, R. Olesyuk and M. Pikhletsky, "Can We Monitor Breathing During Sleep via Wi-Fi on Smartphone?," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, pp. 6710-6713, July 2019. *(references)*

[15] Abdelwahed Khamis, Brano Kusy, Chun Tung Chou, et al., “WiRelax: Towards real-time respiratory biofeedback during meditation using WiFi,” Ad Hoc Networks, 2020, Volume 107, pp. 1-17.

[16] Q. Gao, J. Tong, J. Wang, Z. Ran and M. Pan, "Device-Free Multi-Person Respiration Monitoring Using WiFi," IEEE Transactions on Vehicular Technology, 2020, vol. 69, no. 11, pp. 14083-14087.

[17] K. Niu, Fusang Zhang, et al., "WiMorse: A Contactless Morse Code Text Input System Using Ambient WiFi Signals," IEEE Internet of Things Journal, 2019, vol. 6, no. 6, pp. 9993-10008.

[18] Y. He, Y. Chen, Y. Hu and B. Zeng, "WiFi Vision: Sensing, Recognition, and Detection With Commodity MIMO-OFDM WiFi," IEEE Internet of Things Journal, 2020, vol. 7, no. 9, pp. 8296-8317.

[19] Yue Zheng, Yi Zhang, Kun Qian,et al., “Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi,” In Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '19). Association for Computing Machinery, New York, NY, USA, pp. 313–325, June 2019.

[20] Navneet Dalal, Bill Triggs, “Histograms of Oriented Gradients for Human Detection,” International Conference on Computer Vision & Pattern Recognition (CVPR ’05), San Diego, United States. pp.886–893, Jun 2005. *(references)*

[21] Bin Dong, Guirong Weng, Ri Jin, “Active contour model driven by Self Organizing Maps for image segmentation,” Expert Systems with Applications, 2021, vol. 177, pp.1-9.

[22] Halperin Daniel, et al., “Tool release: Gathering 802.11n traces with channel state information,” ACM SIGCOMM Computer Communication Review. 2011, vol.41, no.1, pp. 53-53.