WILEY | Hindawi

Research Article

Air-CSL: Chinese Sign Language Recognition Based on the Commercial WiFi Devices

Honghong Chen,[1,2] **Danyang Feng** [iD],[1] **Zhanjun Hao** [iD],[1,2] **Xiaochao Dang,**[1,2] **Juan Niu,**[1] **and Zhiqiang Qiao**[1]

[1]*College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China*
[2]*Gansu Province Internet of Things Engineering Research Center, Lanzhou 730070, China*

Correspondence should be addressed to Zhanjun Hao; zhanjunhao@126.com

Artificial intelligence and Internet of Things (IoT) devices are experiencing explosive growth. Currently, the commonly used gesture recognition methods are difficult to deploy and expensive, so this paper uses the Channel State Information (CSI) for Chinese sign language recognition. Aiming at the problems of current gesture recognition methods, such as strong personnel dependence, high computational resource consumption, and low robustness, we proposed a Chinese sign language gesture recognition method named Air-CSL. In this method, the Local Outlier Factor (LOF) removal algorithm and the Discrete Wavelet Transform (DWT) are used to reduce the noise in the data, and the subcarriers that best represent the gesture data are selected by principal component analysis. After denoising, mathematical statistics were extracted from the gesture waveform as the eigenvalues, and the features were fused by the Deep Restricted Boltzmann Machine (DBM). Finally, the result of gesture classification and recognition is obtained by the Gated Recurrent Unit (GRU). In this way, the prediction model realizes as well as the classification of sign language gestures. The results show that the proposed method can effectively recognize Chinese sign language gestures of different people in different environments and has good robustness.

1. Introduction

Wireless communication technology has evolved to the point where WiFi and mobile devices use traffic transmission, accounting for 68% of the total global Internet information [1]. In 2011 Halperin et al. obtained the Channel State Information (CSI) of WiFi signals from Intel 5300 NICs. The research on human perception and indoor localization based on CSI has become the focus of academic attention [2]. Some scholars have focused on gesture recognition and worked on real-life applications. At present, wireless communication technology pays attention to its application in the field of special education [3]. According to statistics, there are about 1.57 billion people with hearing impairment worldwide [4]. In China, the number of deaf people exceeds 20.8 million, accounting for 1.69% of the country's total population. As a universal language for deaf people, sign language is a necessary means of communication and learning for the hearing impaired and speech impaired. With the advancement of gesture recognition technology, wireless Internet of Things environments enable people to interact freely with environmental devices.

In special education, deaf people learn Chinese characters through pinyin and eventually reach the purpose of learning sign language. Meanwhile, as a special language, if sign language can be converted into corresponding characters, it will greatly facilitate the communication between deaf people and normal hearing people. Thus, the application of sign recognition through wireless communication technology to the teaching of sign language to special people such as deaf people is a very effective way of teaching.

Sign language recognition systems based on multiple methods can break the barrier that exists between deaf people and people who do not know sign language. For
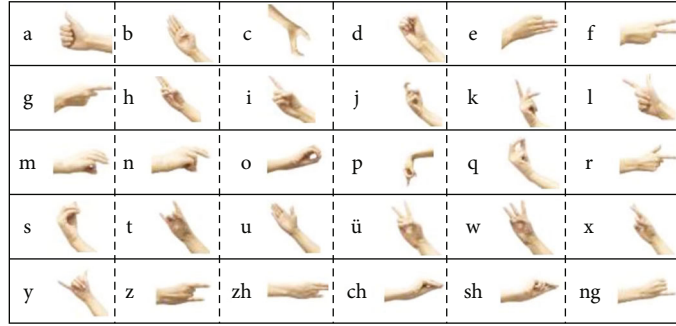
Figure 1: Hanyu Pinyin sign language sign.

example, some sign language recognition systems use cameras [5–9] or Kinect [10, 11] to capture sign language gesture data literature. However, the data acquisition equipment is affected by the lighting conditions, while the camera captures data in such a way that violates the privacy of the person. Some gesture recognition systems propose the use of somatosensory controllers for gesture data acquisition. However, this data collection method is more sensitive to the distance and displacement between the device and the person. Some gesture recognition works propose that sensors can be used for motion data acquisition [12–14]. However, for gesture recognition, the user needs to wear the sensor in a position such as a finger or an arm, which will greatly affect the user's daily life.

Since 2013, academia has begun to pay attention to CSI-based research on human perception and indoor positioning. At present, human behavior research based on CSI has achieved good results. Currently, CSI can be used in intrusion detection [15–17] with gesture recognition [18–20]. Moreover, WiFi devices open new paths for gesture recognition due to their simplicity and ease of deployment. Currently, most of the gesture recognition objects using WiFi are American sign language gestures or daily actions, numbers, etc. And our goal is to recognize Chinese sign language gestures, which are large in number and complicated in action. When different experimenters complete the gesture actions multiple times, it causes problems such as a large amount of data and high computational overhead, so we should choose recognition algorithms with high recognition accuracy and low computational overhead cost.

Although there has been some progress in gesture recognition, however, CSI is very sensitive to the indoor environment, and the whole transmission process is still affected by the multipath effect. Moreover, the feature extraction of gesture recognition is challenging, and the actual application environment greatly affects the gesture recognition results. When the amount of gesture data is large, the calculation cost is high. To solve the above problems, we propose a sign language gesture recognition method based on CSI, which can recognize Chinese deaf people's Hanyu Pinyin sign language, and the Hanyu Pinyin sign language gestures are shown in Figure 1. Hanyu Pinyin differs from other languages and has a total of thirty letters. The sign language gesture actions are described by the CSI measurements. After denoising the raw data, Air-CSL classifies and recog-

nizes the preprocessed CSI values. We validated Air-CSL in both laboratory and classroom environments and verified that the method achieves good robustness in perceiving and recognizing Pinyin sign language gestures. In summary, we have made three contributions.

(1) Due to the multipath effect, this paper combines the Local Outlier Factor (LOF) removal algorithm with the Discrete Wavelet Transform (DWT) to remove noise from gesture data

(2) Commonly used mathematical statistics such as kurtosis cannot fully describe the characteristics of sign language gestures, resulting in insufficient gesture information. There are many kinds of gestures, which can easily lead to the misclassification of classifiers. Therefore, the Deep Restricted Boltzmann Machine (DBM) is used to train mathematical statistics and extract more abstract and comprehensive sign language gesture features

(3) The sample size of sign language gesture data collected by us is large, and the calculation time is too long, which makes the feedback time of human-computer interaction extremely challenging. Therefore, the Gated Recurrent Unit (GRU) is used for Chinese sign language gesture recognition

(4) Finally, we propose the Air-CSL model based on commercial WiFi devices. Through a large number of experiments, the superiority of the model in recognizing different people's sign language gestures in different experimental environments is verified, and the recognition rate of the empty hall can reach 91.77%.

2. Related Work

Currently, researchers have proposed various perceptual techniques for human action recognition, mainly based on sensors, computer vision, and wireless devices.

The first category recognizes human actions through sensors. Yuan et al. proposed the use of sensor-equipped accessories as data acquisition devices to capture gestural actions, with recognition results of more than 90% in all cases [21]. However, wearable device-based recognition

methods require the user to wear a special device, which affects the action description and greatly reduces the comfort of the user.

The second category recognizes actions through computer vision. The gesture recognition is through the Microsoft Kinect sensor [22, 23]. For instance, Liu et al. used double and tenfold cross-validation to achieve more than 91% recognition of Arabic numerals (0-9) and English letters (A-Z). In the literature [24], the projector camera system is used to extract the spatial information of human actions, so as to realize the control of dynamic gestures on the Augmented Reality (AR) model. In this paper, the four gestures are recognized and combined with the projected AR environment construction, correction, and registration to achieve surgical guidance. In the literature [25], the surgical robot was placed on a mobile platform, and the advantages of AR visual guidance information, surgeon's experience, and robotic surgery were integrated to enhance the surgical field of vision. The eight gestures mentioned in this paper can control the robot and AR, so as to achieve high-quality interaction of surgical information. The literature [26] recognizes hand gestures based on video. In this paper, by modeling the human skeleton and retaining the key frame, the time dimension information is taken as the classification key. For static gestures, this method achieves a very satisfactory recognition rate. However, visual recognition usually requires lighting conditions and involves personal privacy issues, which has limitations in practical use.

The third type of wireless device action recognition method can be done by ultrabroadband radar [27], radio frequency identification technology [28], received signal strength indication, or CSI signal [29]. Among them, RF technology and ultrabroadband radar require dedicated equipment and high deployment complexity, and currently, received signal strength indication or CSI is mostly used for contextual awareness. Such sensing alleviates the need for people to constantly wear sensor devices. Researchers used WiFi for indoor person behavior detection [30, 31].

Kang et al. [32] used adversarial learning schemes and feature de-entanglement modules to remove the influence of irrelevant factors in gestures and used an attentional scheme based on the output of source domain discriminator to reflect the similarity differences between multiple source domains and target domains, thus reducing negative transfer. On Widar 3.0 data sets, their model evaluation improved by an average of 3% to 12.7%. In 1D, WiTrace [33] used the synthetic signal to derive the phase of the hand-reflected signal and measured the phase change to obtain the distance of movement. For 2D space, WiTrace proposed the Kalman Filter to filter out the noise of tracking. The method achieved an average accuracy of 6.23 cm for initial position estimation. And WiTrace achieved the average tracking errors of 1.46 cm and 2.09 cm for 1D tracking and 2D tracking. The DFGR used the deep network. The deep network could learn discriminative deep features. What is more, the method could use the transferable similarity to evaluate ability in test conditions [34]. The WiGAN system proposed by Jiang et al. uses generative adversarial networks to extract and generate gesture features, fuses the features, and classifies human activities through support vector machines with a final average recognition accuracy of over 95% [35].

Through experiments, we found that the gesture data reaches a certain magnitude when using support vector machines (SVM) for feature extraction, gesture recognition requires a large overhead, and the existing gesture recognition lacks sufficient research on Chinese gestures. In this paper, the LOF algorithm is used to remove the outliers in sign language gesture data. Meanwhile, we use the DWT to process the low-frequency information in gesture data. After data preprocessing, this paper selects the subcarrier that can best represent sign language gesture information through the PCA, so as to effectively remove part of the environmental interference. Through the above methods, we have effectively solved the problem of large computation overhead. To solve the problems of large computational overhead and low accuracy of gesture recognition, this paper analyzes the statistical features of CSI amplitude and fuses the features by the DBM. Finally, the recognition results are obtained when the gesture data are input into the model, and the GRU is recognized.

3. Overview of the Gesture Recognition Method

Sign language gesture recognition by CSI requires four steps: sign language gesture data sensing, noise removal, feature extraction, and sign language gesture recognition, and the workflow is shown in Figure 2. We used two laptops configured with Intel 5300 NIC for data acquisition, one working in IEEE 802.11n Monitor mode as the transmitter and the other as the receiver.

3.1. Data Acquisition and Preprocessing. There are multiple reflection paths caused by gesture actions. The traditional CSI-based action recognition generally uses the channel features of a single antenna and a single link as the sensed data. However, such an approach tends to lose more action feature data easily. Multiple antennas provide sufficient CSI, and we collected the raw information as shown in Figure 3. The channel impact response can be used to describe the merit of the propagation path, expressed as

$$H(f, t) = H_s(f) + \sum_{k \in P_d} \alpha_k(t) e^{j2\pi \int_{-\infty}^{t} f_{D_k}(u) du}, \qquad (1)$$

where $H_s(f)$ denotes the static response, P_d denotes the sum of the number of dynamic reflection paths, and $\alpha_k(t)$ denotes the complex attenuation factor. f_{D_k} denotes the Doppler shift of the reflected signal.

The amplitude can describe the static response of CSI, and since the experimenter is on the central link and there is a strong LOS signal in the indoor environment, we choose an antenna with a smaller amplitude. Meanwhile, different antennas have different sensitivities to hand gesture actions. The variance can fully describe the CSI changes caused by gesture actions, and a larger variance indicates a better reflection of the dynamic response. Therefore, we choose
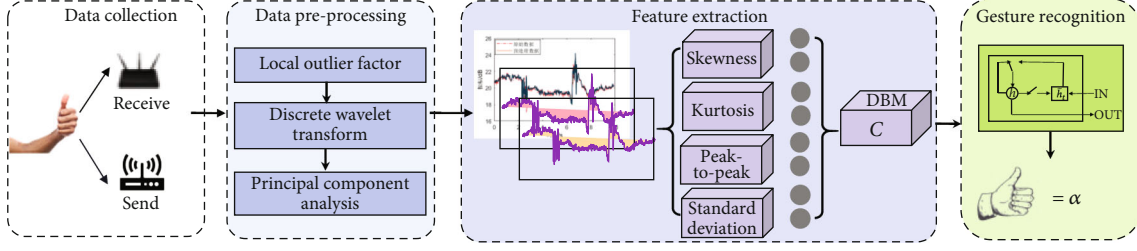
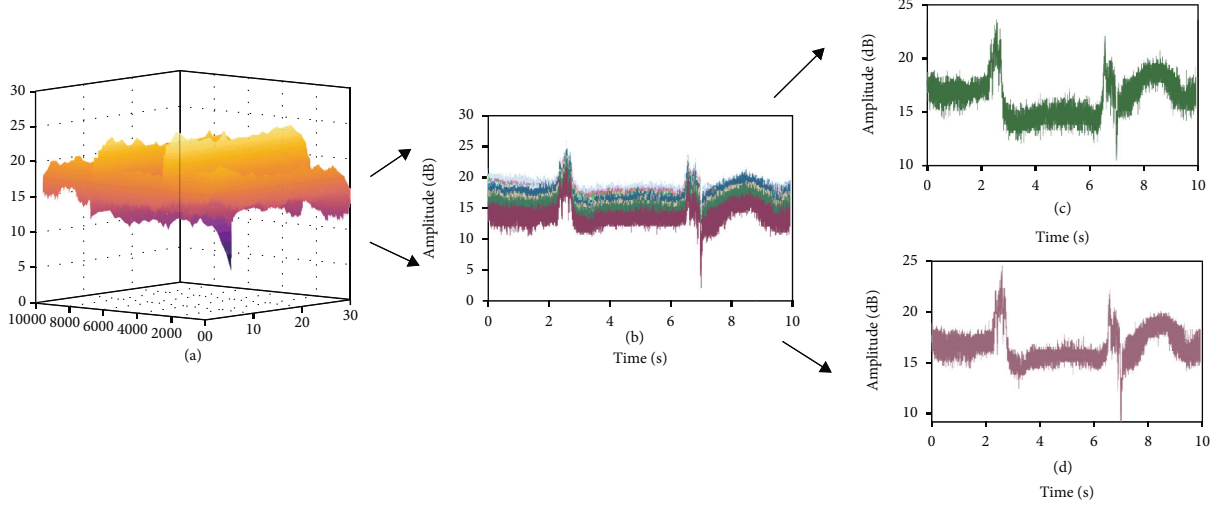FIGURE 2: Sign language gesture recognition workflow diagram.

FIGURE 3: Graph of raw data of sign language "e."

the antenna with the maximum variance of CSI and relatively small amplitude.

The filtered antennas are affected by the multipath effect and inherent noise, and there are spikes and burrs in the data waveform. To remove the outliers while preserving the gesture information as much as possible, we choose the LOF algorithm to remove the outliers from the gesture data, as shown in Figure 4(a). It can be expressed as

$$\mathrm{LOF}_k(X) = \frac{\sum_{Y \in N_k(X)} (\mathrm{Ird}_k(Y)/\mathrm{Ird}_k(X))}{|N_k(X)|}, \qquad (2)$$

where $N_k(X)$ denotes the total number of points in the k-distance neighborhood of the point X. $\mathrm{Ird}_k(X)$ is the local reachable density of the point X.

The principle of $\mathrm{LOF}_k(X)$ anomaly removal is when $\mathrm{LOF}_k(X)$ tends to 1, it indicates that the density of neighboring points of the measured point is almost equal and belongs to the same cluster; the more $\mathrm{LOF}_k(X)$ is greater than 1, there is a difference in the domain density of the measured point, then the point is considered as an anomaly; if $\mathrm{LOF}_k(X)$ is much less than 1, the measured point is considered as a dense point.

When there is more interference in the environment, DWT is selected for multipath effect removal as in Figure 4(b). DWT performs multiscale analysis of fine-grained actions and uses Symlet5 for signal decomposition

into approximate coefficients and multiple detail coefficients to remove high-frequency noise while retaining the approximate characteristics and data details of the gesture waveform, and the detail coefficients describe the random noise and CSI data details in the device. This can be expressed as

$$\begin{aligned} \alpha_k^{(J)} &= \sum_{n \in Z} X(n) g_{n-2^J k}^{(J)}, J \in Z, \\ \beta_k^{(L)} &= \sum_{n \in Z} X(n) h_{n-2^L k}^{(L)}, 1 \le L \le J, \end{aligned} \qquad (3)$$

where $\alpha_k^{(J)}$ is the approximation coefficient, $\beta_k^{(L)}$ is the detail coefficient, and $X(n)$ denotes the gesture data sample points. A soft thresholding algorithm is applied to the detail parameters, and the inverse Discrete Wavelet Transform is used to reconstruct the denoised gesture waveform $X'(n)$, expressed as

$$X'(n) = \sum_{n \in Z} \alpha_k^{(J)} g_{n-2^J k}^{(J)} + \sum_{L=1}^{J} \sum_{n \in Z} \beta'_k^{(L)} h_{n-2^L k}^{(L)}. \qquad (4)$$

The 30 subcarriers after removing the noise contain subcarriers with less correlation with the gesture actions, so the principal component analysis algorithm is used to reduce the dimensionality and select the subcarriers with high similarity to those before the reduction. Firstly, the mean value $\bar{x}$

(a) The waveform after LOF

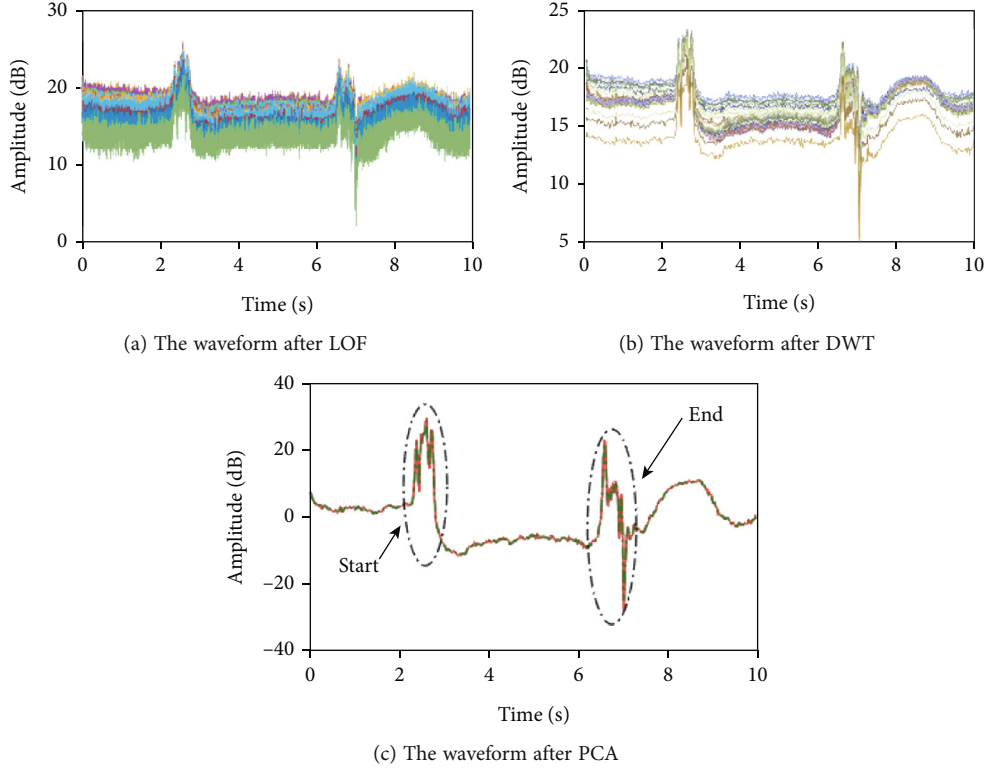(b) The waveform after DWT

(c) The waveform after PCA

FIGURE 4: Preprocessing diagram of gesture data.

of the gesture sample set $X = \{x_1, x_2, \cdots, x_m\}$ is obtained and expressed as

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix}. \tag{5}$$

$Y = \{y_1, y_2, \cdots, y_m\}$ is the sample set after the sample normalization process, where $y_i = x_i - \bar{x}$. The covariance matrix C of the reconstructed sample Y is obtained and expressed as

$$C = \begin{bmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_1) & \cdots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_1) & \cdots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_1) & \cdots & \text{cov}(X_n, X_n) \end{bmatrix},$$

$$\text{cov}(X, Y) = \frac{\left[\sum_{i=1}^{m}(X_i - \bar{x})(Y_i - \bar{y})\right]}{n-1}. \tag{6}$$

The eigenvalue matrix of this covariance matrix C is $\lambda = [\lambda_1, \lambda_2, \cdots, \lambda_n]$. The first k eigenvalues are taken in descending order, and $Q = \{q_1, q_2, \cdots q_k\}$ is the eigenvector matrix composed of the vectors corresponding to the eigen-

values. The eigenvector matrix is multiplied with the original sample set to obtain the reduced-dimensional matrix $R = Q^T X$. The first principal component is finally retained as the CSI waveform for gesture recognition, and the results of the subcarrier extraction by the principal component analysis algorithm are shown in Figure 4(c).

3.2. Feature Extraction. There are differences in the way and speed of describing different people when completing gestures, and in addition, it is difficult to guarantee that the waveforms are the same when someone performs the same gesture. Different people perform different gestures as shown in Figure 5. To achieve consistency in the waveform of the same gesture and to highlight the differences between different gestures, multiple eigenvalues can be selected, but too many eigenvalues are likely to cause fitting problems. Therefore, in this paper, we select the feature values: skewness, kurtosis, standard deviation, and peak-to-peak value, and the four are described in Table 1.

$$S = \frac{1/n\sum_{i=1}^{N}(x_i - \bar{x})^3}{\left(1/n\sum_{i=1}^{N}(x_i - \bar{x})^2\right)^{3/2}}, \tag{7}$$

where x_i denotes the data points in the sample and $\bar{x}$ denotes the mean of the data points.

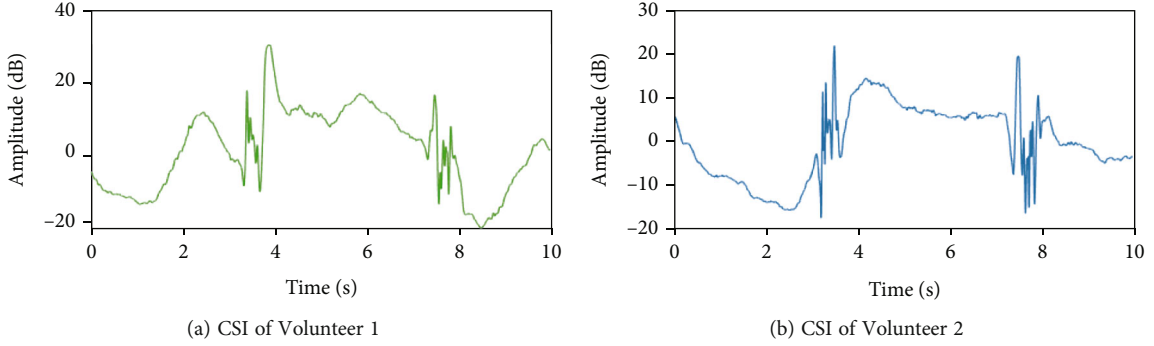$$\text{Kurtosis} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^4}{N\sigma^4}, \tag{8}$$

(a) CSI of Volunteer 1

(b) CSI of Volunteer 2

FIGURE 5: Different people perform different gestures.

TABLE 1: Description of eigenvalues.

Feature name	Feature description
Skewness	Measure the direction of skew, degree of skew, and asymmetry of the data distribution, expressed as Equation (7)
Kurtosis	The number of characteristics that indicate the peak height of the probability density distribution curve at the mean is expressed as Equation (8)
Standard deviation	The square root of the squared and averaged data deviation from the mean is expressed as Equation (9)
Peak-to-peak value	The value of the difference between the highest and lowest values of the signal in a period is expressed as Equation (10)

where $(x_i - \bar{x})^4$ represents the fourth-order central moment, x and $\bar{x}$ represent the sample points and sample means, respectively, and σ represents the standard deviation.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}}, \bar{x} = \frac{1}{N}(x_1 + x_2 + \cdots + x_n), \quad (9)$$

$$pk - pk = \max(x_i) - \min(x_j), \quad (10)$$

where $pk - pk$ refers to the value of the difference between the highest and lowest value of the signal in one cycle, which is the range between the maximum and minimum.

When there are too many gesture data and a variety of gesture types, it is easy to cause inadequate feature description. Currently, we know that existing restricted Boltzmann machines usually perform well in extracting high-dimensional features from the data. Stacking Restricted Boltzmann Machines (RBM) to form the DBM can fuse and downscale gesture features, thus compensating for the recognition errors caused by single features. The multilayer nonlinear transform structure of DBM can accomplish the simulation of complex nonlinear functions. The key component of the DBM is the RBM, which detects, identifies, and classifies the input data by combining multiple layers of RBMs with a final classifier. We input the raw mathematical statistics of the data waveform into the DBM, which is mapped to depth features through multiple feature reconstruction.

Feature fusion through DBM-3 can be roughly divided into the following three steps:

Step 1. Input the base feature values of the gesture data: skewness, kurtosis, standard deviation, and kurtosis into the DBM-1 layer for the first deep fusion to obtain the new feature eigenvalues c_1, c_2, c_3, c_4.

Step 2. Combine the features two by two and input them into DBM-2 for the second fusion, feature values c_5, c_6.

Step 3. The features are combined and input into DBM-3 to get the final fused features, which are used as the input of GRU in the gesture recognition stage. The feature fusion via DBN is shown in Figure 6.

3.3. Recognition Model. Recurrent neural network (RNN) is divided into input layer-hidden layer-output layer when processing time-series data, and the output result is related by the current input and previous hidden state, but it ignores the law between information with long interval and current input information. Based on the recurrent neural network and its existing deficiencies, the RNN is improved and derived from neural network models such as long short-term memory network (LSTM) and Gated Recurrent Unit (GRU) [36]. However, LSTM requires four linear layers for each unit, and the training efficiency is low. Compared with RNN and LSTM, sequence modeling with GRU involves fewer parameters and is easier to train.

The GRU contains an update gate and a reset gate. The update gate determines what information needs to be added to the new state and what information needs to be removed at the previous time $t - 1$. The reset gate determines the degree of forgetting of the hidden state at the previous time $t - 1$. When the time is t, the inputs of both the reset gate and update gate are the gesture waveform timing features
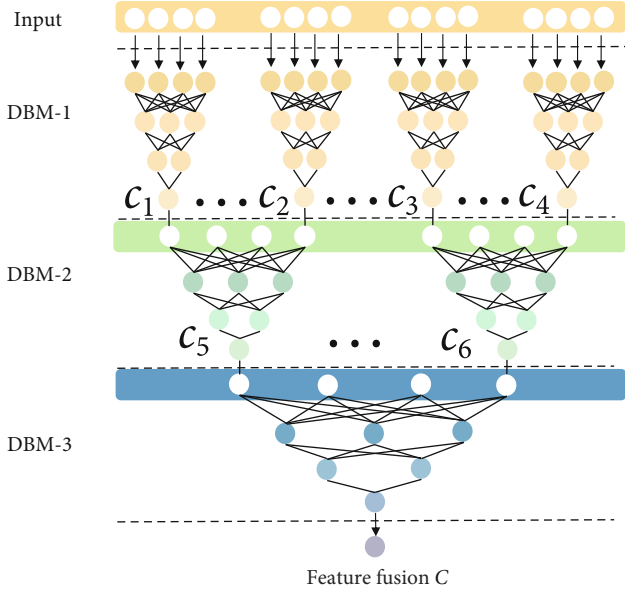
Figure 6: Model diagram of feature fusion by DBM-3.

Figure 7: GRU structure diagram.

extracted by the DBM and the hidden states of the previous time $t-1$. The network structure of GRU is shown in Figure 7, and the model is derived from the literature [37].

The hidden state $\tilde{h}_t$ processed by the reset gate r_t, update gate z_t, and tanh activation function in GRU and the final hidden state h_t parameters are updated by the following equations.

$$z_t = \text{sigmoid}(W^z \cdot [h_{t-1}, x] + b_z), \tag{11}$$

$$r_t = \text{sigmoid}(W^r \cdot [h_{t-1}, x] + b_r), \tag{12}$$

$$\tilde{h}_t = \tanh\left(W^{\tilde{h}} \cdot [r_t \cdot h_{t-1}, x] + b_{\tilde{h}}\right), \tag{13}$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t, \tag{14}$$

where x represents the feature value information input at moment t and h_{t-1} is the hidden state at moment $t-1$. And sigmoid is the activation function.

After the feature information $x = \{x_1, x_2, \cdots, x_n\}$ enters the GRU, it is multiplied with the hidden state of the previous moment by the activation function sigmoid to get the gating signals r_t and z_t. As a result, the hidden state h_{t-1} of the previous moment and the updated reset gating with x are multiplied by the elements through the activation function tanh to get the current candidate state value $\tilde{h}_t$. The "updated memory" forgets some unimportant information in h_{t-1}, and the "enhanced memory" for some important information to get a new sequence of hidden states h_t for the next time step. After several training sessions, the probability of the gesture action category is finally derived by the softmax function as

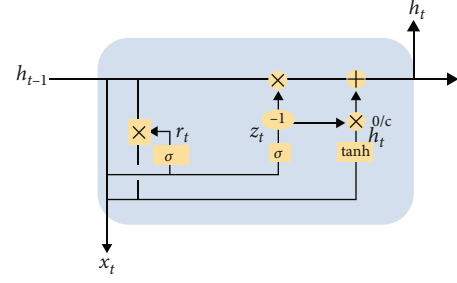$$p_m = \frac{e^{y_m}}{\sum_{j=1}^{m} e^{y_m}}, \tag{15}$$

where y_m is the probability value of the classification action. It can be calculated by Equation (16). The reason why we choose ReLU is that the calculation speed is fast, the calculation amount is small, and the training time is shorter

$$y_t = \text{ReLU}(d \cdot W_n + b_n). \tag{16}$$

The pseudocode of the recognition framework training process of Air-CSL is shown in Algorithm 1. For the test of gesture data, we use the trained model to output the probability.

4. Experiments and Evaluation

4.1. Implementation. We collected gesture data from two laptops with Intel 5300 NIC inside as a pair of transceivers with one antenna at the receiving end and three at the transmitting end. We consider the actual lecture environment, so we set the horizontal distance between the devices to 2 meters. The experimental environments are an empty hall and a classroom, and the scene schematic is shown in Figure 8. To reduce the impact of action completion time on the overall recognition rate, the data acquisition time was set to 10 s. Among them, 0-3 s were stationary, the experimenter described the gesture action in the 4th second, the action description took about 2 s, and the action was retracted in the 7th s. Each action was repeated 20 times. Seventy percent of the experimental data were used for the training of the recognition model, and 30% were used as a test set to test the model. Twelve experimenters were randomly selected. The height and weight information of volunteers is shown in Table 2. The experimenters numbered 1-6 are randomly selected interested people in our college who participated in all the comparative trials involved in this paper, where (∗) represents the tester added later, namely 7∗-12∗. The experimenters in this section (∗) only participated in the experiments in Section 4.4.

4.2. Influence of Different Heights. A pair of transceivers was chosen for the experimental equipment. However, considering the practical situation, we propose that the device will have some effect on gesture recognition when it is at different heights. We speculate that with the height where the average center of gravity of the experimenter is located as the middle value, the sensitivity of the signal to the gesture action decreases as the height of the device increases or decreases, and when the height reaches a certain limit, the

Input: (*hidden_size*)
Output: The loss value *total_loss*
Initialize the model parameters, the hidden layer dimension (*hidden_size*), the training rounds (*n_epochs*), the number *input_size*,
training loss value (*test_loss*)
1:   *total_loss* ⟵ 0
2:   for *i* to *input_size* do
3:       use *hidden_size* to update the reset door $r_t = sigmoid(W^r \cdot [h_{t-1}, x] + b_r)$, the update door $z_t = sigmoid(W^z \cdot [h_{t-1}, x] + b_z)$, and
the hidden state $\tilde{h}_t = \tanh(W^{\tilde{h}} \cdot [r_t \cdot h_{t-1}, x] + b_{\tilde{h}})$
4:       forget unimportant information, update $h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t$
5:   End for
6:   pass the data to the tensor
7:   for *y* to *n_epochs*;
8:       if the current *loss* < the predicted loss *prev_loss*
9:       save model parameters
10:     and *prev_loss = loss*
11:     else if *loss* < (*test_loss*)
12:         output the training rounds *n_epochs* and *loss*
13:     End if
14:   End for
15:   Return *total_loss*

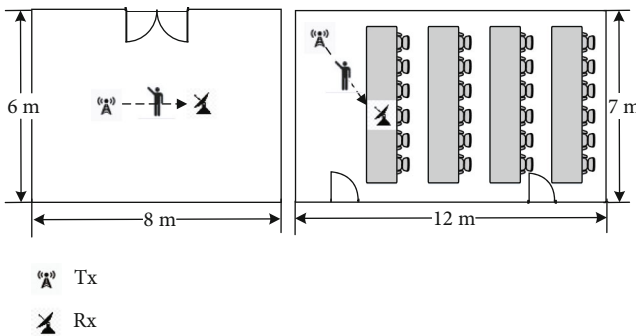ALGORITHM 1:Overall framework of Air-CSL.

Figure 8: Schematic diagram of the scenario.

effect of the sign language action on the CSI stream almost disappears. The accuracy of sign language gestures decreases with the increase or decrease of the device height. Because a stronger signal can respond to hand movements, signal sensitivity is high; conversely, sensitivity decreases. The accuracy rates at different heights are shown in Figure 9.

As seen in Figure 9, the sign language recognition performance is best when the height of the device is 1.35 m, and the performance of gesture recognition decreases to different degrees as the height increases or decreases. We calculated the average of the comfortable heights considered by the six subjects, and combined with the real-life teaching scenarios, we finally chose the experimental device spacing of 1.25 m. In the subsequent comparison experiments, we all used the height setting of 1.25 m.

4.3. Influence of Different Packet Sending Rates. Sign language gesture actions are fine-grained actions, and the signal changes caused by fine-grained actions may be weak or cause brief and fast changes in the signal. Therefore, the high sampling rate can fully describe each sign language gesture. When the CSI value collected is large enough, it is beneficial

to the feature extraction and classification of sign language gestures. To evaluate the performance at different sampling rates, we chose to vary from 400 to 1200 packets/sec. The experimental results show that the accuracy of the method reaches its maximum at a packet rate of 1000 packets/sec. Therefore, the packet rate is uniformly set to 1000 packets/sec in the subsequent experiments. At the same time, we found that the packet rate and gesture recognition accuracy are roughly positively correlated, but when the packet rate reaches 1000 packets/second or more, better hardware is needed to support the store as shown in Figure 10.

4.4. Influence of Different Experimenters

4.4.1. Influence of Different Individuals. Since different personnel differs in the method and time spent on action description when completing sign language gestures. To compare the recognition accuracy of different personnel in the same environment, we compared the recognition results of six experimental personnel in an empty hall and a conference room. The recognition rates of different experimenters are shown in Figure 11, which shows that there are some differences in the recognition rates of different experimenters, and it is concluded that the recognition results of female students are better than those of male students due to the inherent difference in body size.

The recognition rate is the average value of the recognition rate of this experimenter in both scenes. We found that the recognition rate is relatively low for slightly fat people and people with too fast gesture description and higher for people with proportional body and even gesture description processes. Since the height of the device is set at 1.25 m, when the height of the experimenter is too high or too low, it will also have some influence on the recognition results. Overall, the average gesture recognition rate of different personnel in an open environment stays above

TABLE 2: Information table of physical signs of experimental personnel.

Number	1	2	3	4	5	6	7*	8*	9*	10*	11*	12*
Gender	M	F	M	F	M	F	M	F	M	F	M	F
Height/cm	175	158	170	173	182	168	169	172	183	159	173	162
Weight/kg	71	54	67	63	93	51	63	57	71	60	67	48
BMI	23.18	21.63	23.18	21.05	28.08	18.07	22.06	19.27	21.20	23.73	22.39	18.29
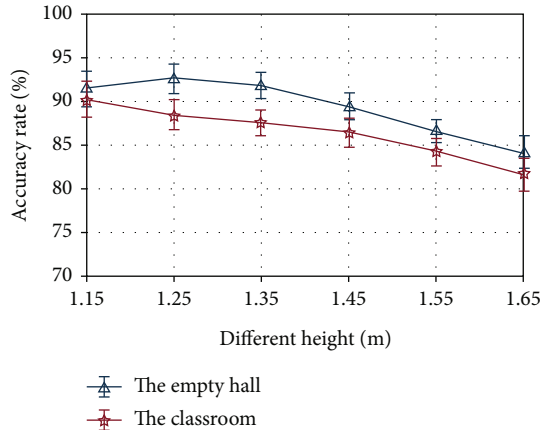
FIGURE 9: Impact of different heights.
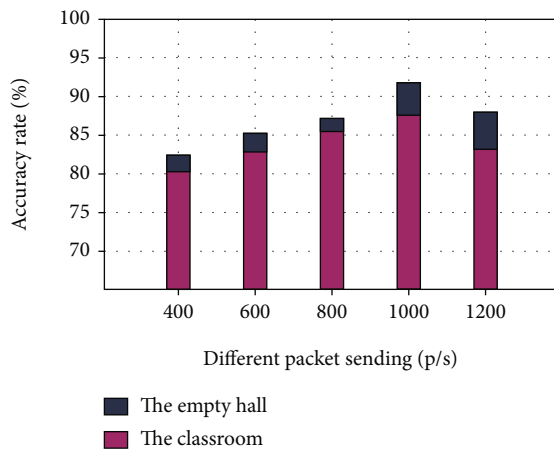
FIGURE 10: Impact of different transmission rates.

91.77%. This indicates that Air-CSL has strong adaptability to different personnel.

4.4.2. Influence of Different Genders. To fully study the influence of different genders on the recognition rate, we added four males and four females based on the existing experimenters. We conducted data collection for the additional six people under the same experimental scenario. Each gesture was repeated 10 times in each scene. This shows that Air-CSL data can present good recognition rates even for untrained personnel. At the same time, it can also reflect the influence of gender differences on gesture recognition rate.

As can be seen from Figure 12, compared with the trained sample, the gesture sample of the untrained sample

also achieved good recognition results. The average recognition rate of people without sample training was 88.27%. Figure 12(a) shows the recognition rate of different females. By comparing with Figure 12(b), we can get the following results: Among several randomly selected experimental personnel, the recognition rate of females is better than that of males. The average recognition rate was 88.89% for females and 87.66% for males. Our discussion leads to the following conclusion: women with the same body mass index (BMI) have a slower motion description process than men. In addition, among the subjects we selected, women had less fat distribution in their hands than men, so their fingers were thinner, which is conducive to sign language gesture recognition.

4.5. Influence of Different Experimental Environments. To verify the robustness of the method in this paper, we added static interference and dynamic interference to the two existing experimental environments. The static interference (I) is set as follows: a chair is placed at a horizontal distance of 0.5 m from the transmitting end and the receiving end, respectively; the dynamic interference (I) is set as follows: an experimenter is allowed to walk at a uniform speed at a place parallel to the line-of-sight path distance of 1 m; the dynamic interference (II) is set as follows: the door of the room is artificially flapped to simulate a door disturbed by external factors in the teaching environment. The recognition rates of different environments are shown in Figure 13.

The experimental results show that static interference (I) has relatively little effect on both environments. The average recognition rate in the empty hall can still reach 90.78% on the other hand because the classroom has been arranged with more furniture thus leading to a decrease in the recognition rate in the classroom. When dynamic interference (I) was added to the environment, the gesture recognition rate in both environments decreased significantly to 88.38% and 84.32%, respectively, due to the large gait motion amplitude that interfered with CSI. However, these values are within the acceptable range. Dynamic interference (II) in the two environments caused a more obvious impact, due to the human control of the door swing amplitude being larger, and each swing speed cannot be determined; therefore, gesture recognition accuracy decreased significantly were 85.7% and 81.8%. From Figures 13(b) and 13(c), it can be seen that, when the interference in the environment increases, the original gesture recognition rate of low experimental personnel is affected more seriously.

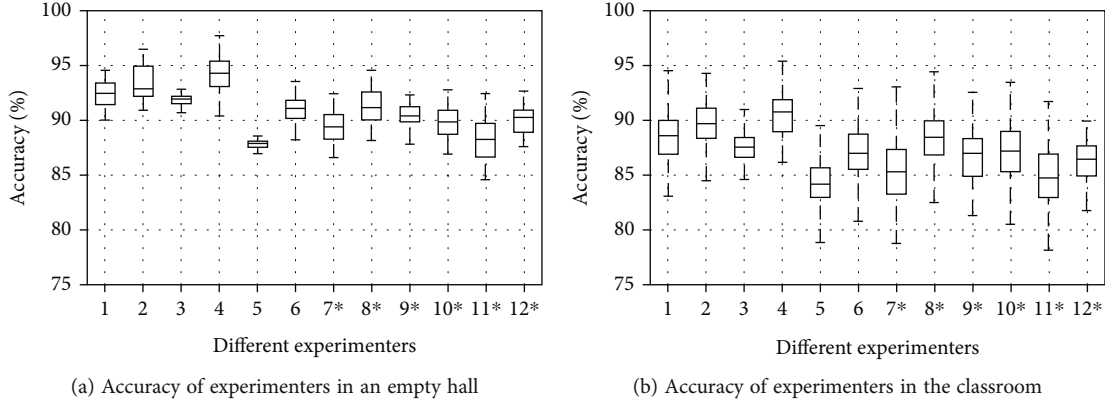*4.6. Influence of Different Feature Values on Gesture Recognition.* Feature extraction directly affects the gesture

(a) Accuracy of experimenters in an empty hall

(b) Accuracy of experimenters in the classroom

FIGURE 11: The accuracy of different experimenters in different scenarios.

(a) Male identification results
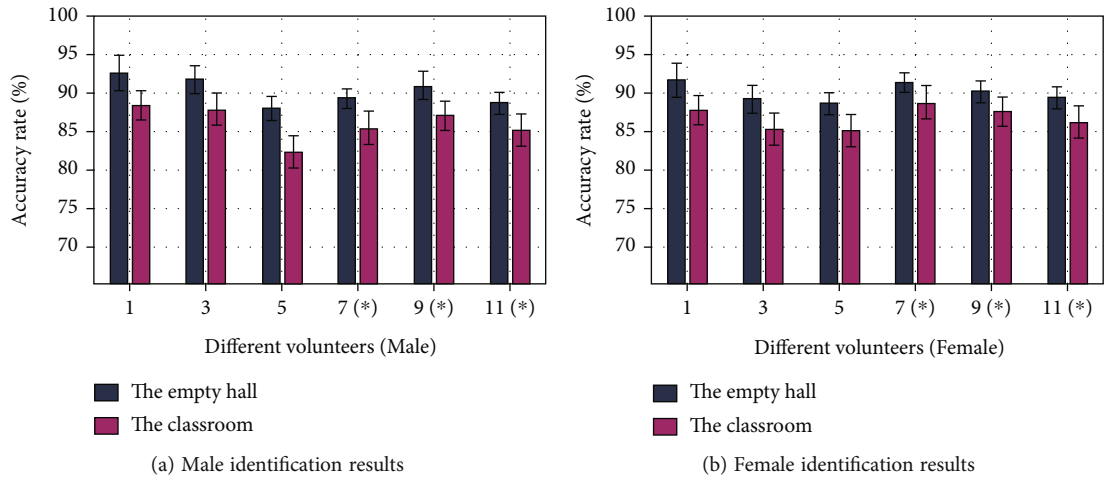

(b) Female identification results

FIGURE 12: Accuracy of different experimenters.

recognition results. To verify the effectiveness of the feature extraction method selected, we use the unfused mathematical statistics-valued features and the fused feature values to test in different environments, respectively. Figure 14 shows the average recognition rate of the original feature combination in two different experimental scenarios. It can be seen that the recognition accuracy is lower when a single mathematical-statistical value is used as a feature. And the recognition rate is distributed in [84.24,89.36], which is because the original statistics contain errors. Figure 15 shows the average recognition rate of the features fused by the method of this paper in two different experimental scenarios. We can find after the fusion of feature values that the gesture recognition rate is significantly improved in both, and the average recognition rate can reach 89.85% in the two experimental environments, which indicates the effectiveness of feature fusion.

4.7. Influence of Different Data Preprocessing Methods. Due to the large amount of noise contained in the original data and to facilitate the extraction of gesture features, we propose to combine LOF, DWT, and PCA data processing schemes. To prove the effectiveness of the method, we remove each data preprocessing scheme separately. In addition, existing gesture data processing schemes were added

for comparison after removing some data preprocessing schemes. In the following, we use the initial six experimenter gesture recognition data to do the evaluation.

(a) We compare the previous methods after removing the LOF algorithm, and the results are shown in Figure 16. From Figure 16, it can be concluded that the overall recognition rate of gestures decreases after the LOF algorithm is removed. Because the data collection is random and the external environment is variable, there are unknown outliers in the collected CSI data, and these outliers are scattered for various reasons and are far from the normal data set. The core of the LOF algorithm is to calculate and compare the density of data distribution to detect the scattered outliers. Therefore, we use the LOF algorithm to preprocess the data to facilitate the recognition of gestures

(b) The DWT algorithm is removed and compared with the previous method, and the results are shown in Figure 16. From Figure 16, it can be seen that the gesture recognition rate decreases in different environments after the DWT algorithm is removed. The decrease in recognition rate is more obvious in
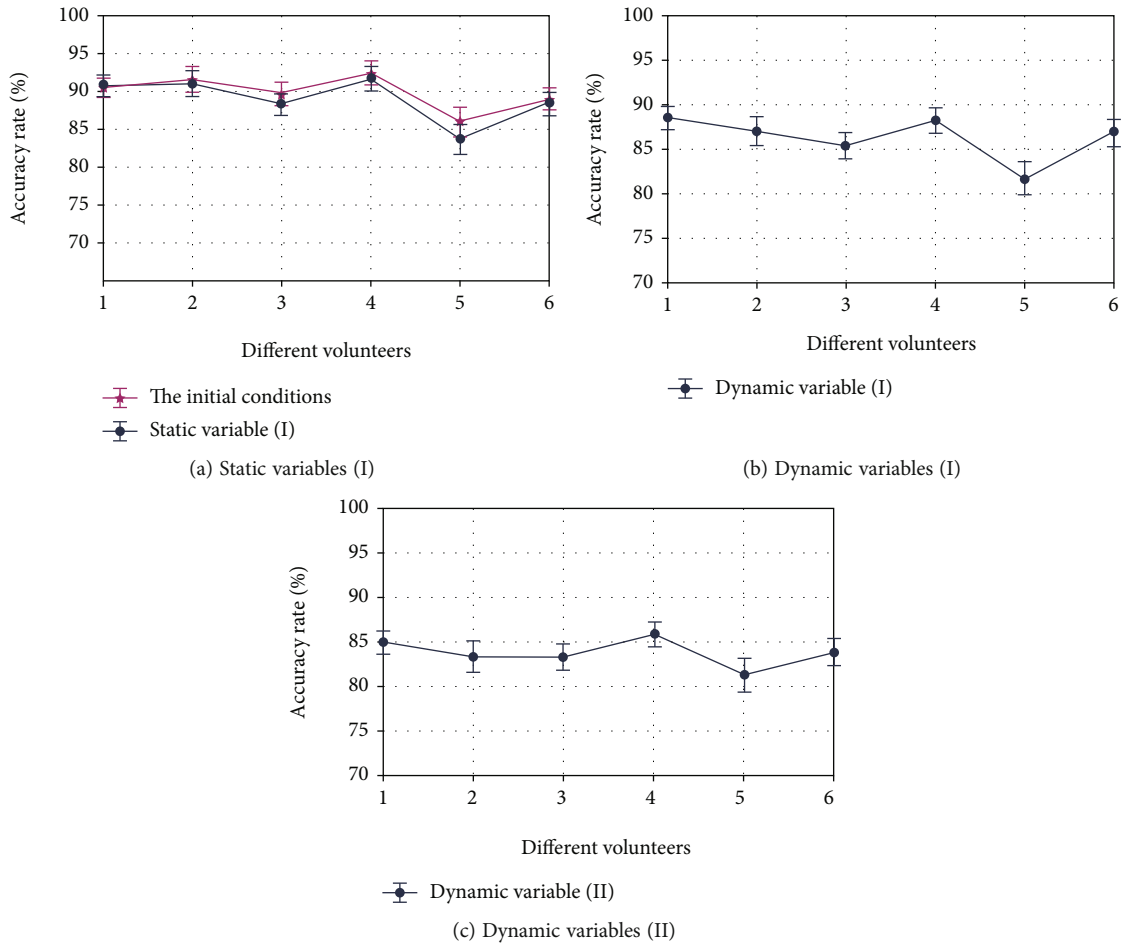
(a) Static variables (I)

(b) Dynamic variables (I)

(c) Dynamic variables (II)

Figure 13: Accuracy of different experimental environment variables.

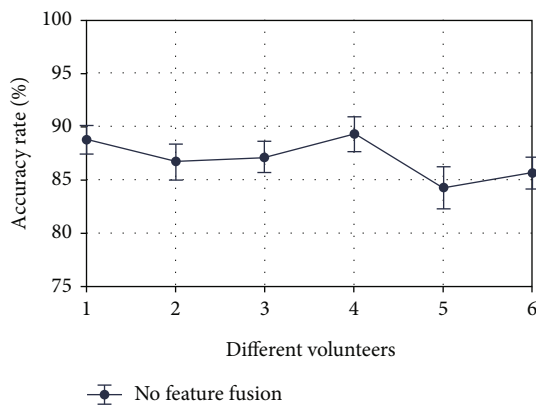Figure 14: Recognition rate of different eigenvalues.

Figure 15: Recognition rate of different eigenvalues.

the classroom compared to the empty hall. This is because the gesture data is low-frequency information, and the data without DWT processing has mixed low-frequency signals and high-frequency signals. The high-frequency information includes those caused by hand gestures and those caused by environmental factors. In order to effectively distinguish the above information, DWT is chosen to retain the peak and abrupt change parts of the useful
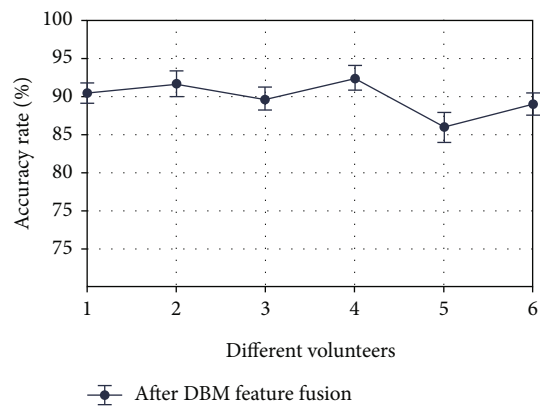
signal needed in the original signal. Detail coefficients and approximation coefficients in DWT can filter out the high-frequency information caused by environmental noise while retaining the signal abrupt change caused by gestures

(c) The PCA algorithm is removed and compared with the previous method, and the results are shown in Figure 16. From Figure 16, it can be seen that the
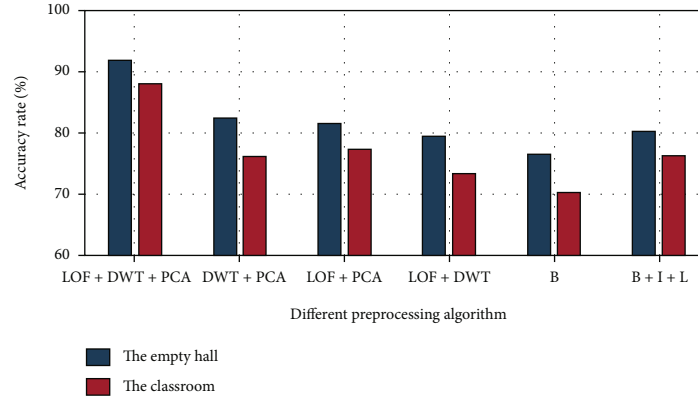
FIGURE 16: The recognition rate of different preprocessing algorithms

recognition rate of gestures has significantly decreased after removing the PCA algorithm. This is because environmental factors cannot be avoided in the gesture data collection process, and PCA can eliminate the influence between CSI data caused by different factors as much as possible by using covariance matrix. And in the actual data processing stage, we tried to select different numbers of principal components, and the results showed that the higher the number of selected principal components, the higher the cumulative contribution rate, but it also contains the interference factors in the environment, and at the same time will increase the computational cost. After several attempts, we finally selected the first ten principal components

(d) In addition to self-comparison, we added comparison with other data pretreatment methods. It uses a Butterworth low-pass filter and sets the cutoff frequency to 100 Hz (we will abbreviate it as "B") [38]. In actual comparison, we need to set the cutoff frequency of the Butterworth low-pass filter lower to conform to the frequency of the gesture (we will abbreviate it as "B + I + P") [39]. There are studies on linear interpolation of the original data and then Butterworth low-pass filter for filtering and finally PCA processing of the data. The comparison results are shown in Figure 16, indicating that LOF + DWT + PCA is more suitable for gesture data preprocessing. We think that only one filter can roughly filter the data, and linear interpolation of the original data is easy to cause problems such as low interpolation accuracy and poor smoothness of the data curve. To sum up, the combination of LOF + DWT + PCA can better process the data

4.8. Influence of Different Algorithms. In recent years, scholars have proposed a variety of recognition methods for gesture recognition, such as LSTM, SVM, and convolutional neural network (CNN). To fully demonstrate the performance of Air-CSL, Air-CSL (GRU) is compared with the existing CNN, LSTM, SVM, k-NN, and random forest (RF).
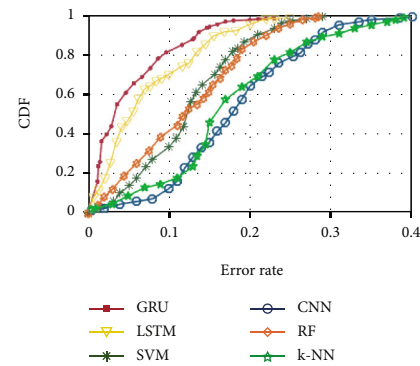

FIGURE 17: Recognition rate of different recognition algorithms.

TABLE 3: Performance comparison table of different algorithms.

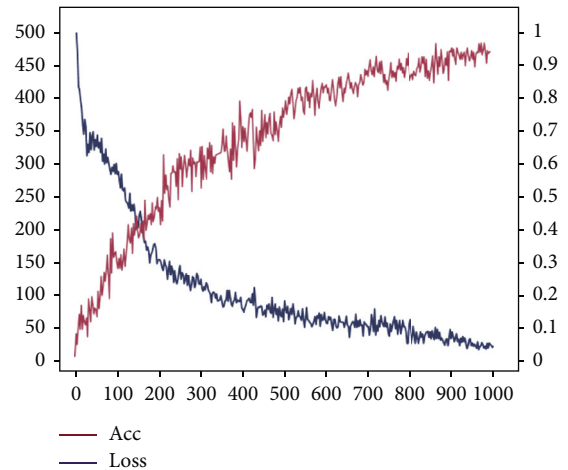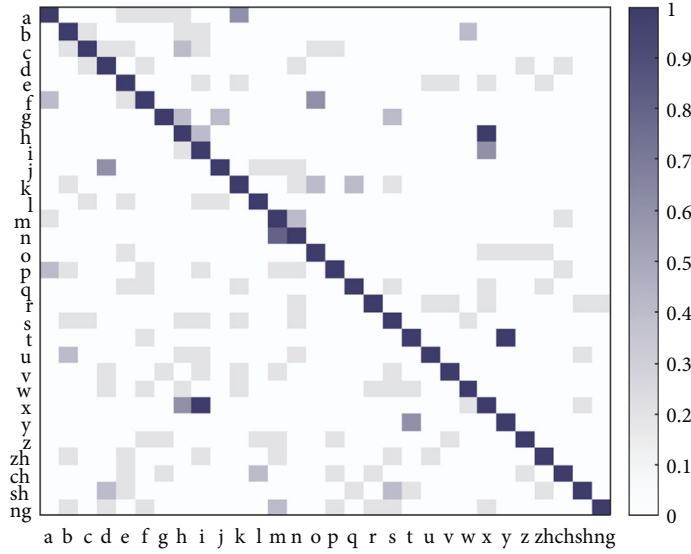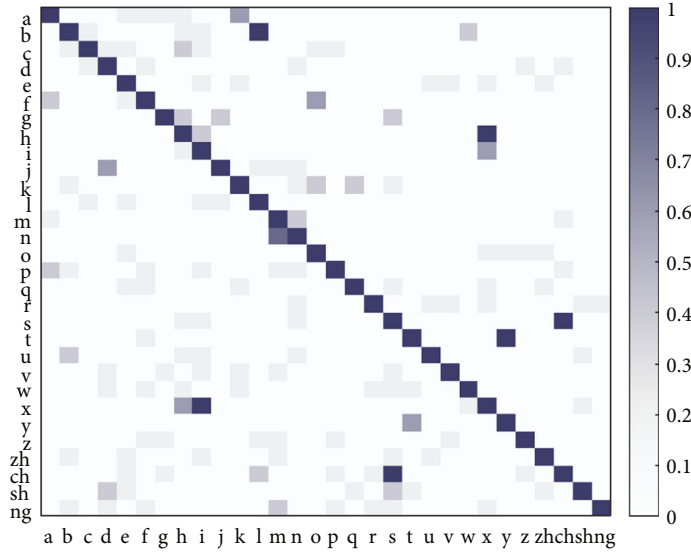| Algorithm	Accuracy rate (%)	F1-score (%)	Time (s)
Air-CSL(GRU)	91.77	95.32	342.37
LSTM	86.91	90.19	328.48
CNN	81.63	85.19	374.63
SVM	83.75	86.68	412.29
k-NN	81.59	86.34	394.82
Random forests	83.06	87.14	425.38

FIGURE 18: The Loss and Acc curves during model training.

(a) Confusion matrix in the empty hall

(b) Confusion matrix in the classroom

FIGURE 19: Confusion matrix with different sign language gestures.

The above methods are used to identify the data collected in the two scenarios, as shown in Figure 17.

We evaluate different algorithms by the cumulative distribution functions (CDF). In Figure 17, x-axis represents the recognition error rate and y-axis represents CDF. We chose the initial device height of 1.25 m. Figure 17 shows that about 85% of the Air-CSL test data have an error rate less than 10%. When CNN was selected as the recognition algorithm, the performance was poor, and the error rate of about 57% of the test data was less than 20%. This indicates that Air-CSL can maintain high accuracy in the recognition of human sign language gestures in the two scenes selected by us.

To further compare the performance of the Air-CSL method with other gesture recognition models, we perform cross-validation of the above models. In this paper, we compare the performance of the four different methods in the

empty hall by accuracy, recall, and F1-score values, as shown in the performance comparison table of different algorithms in Table 2. F1-score is the weighted average of accuracy and recall, and the larger the F1 value is, the better the model performance is. From Table 3, we can conclude that the recognition rate and F1-score of the Air-CSL method are higher than other methods. It is proved that Air-CSL can effectively recognize the sign language gestures of deaf people and has better overall performance.

In order to fully the superiority and robustness of the proposed method in this paper, we evaluate the algorithm performance in terms of the processing time dimension of the algorithm. The number of training samples is 3200, and the number of testing samples is 960. As can be seen from Section 4.7, the recognition of gestures by multiple methods shows better recognition in open environments, while at the same time, the recognition results in classrooms

are generally lower. In terms of computation time, the best algorithm is LSTM, followed by GRU, followed by CNN and SVM. Compared to the GRU model chosen in this paper, the LSTM algorithm performs better in terms of computation time, but combined with its recognition rate in different environments, we believe that the GRU model is more suitable for the sign language gestures recognized in this paper.

4.9. System Performance Evaluation. The core of Air-CSL is essentially a feature fusion recognition algorithm. It carries out feature fusion on commonly used mathematical statistics through the DBM and transforms the feature matching in a general gesture recognition scheme into the feature fusion problem. The curves of Loss and Acc of model training are shown in Figure 18. It can be seen from Figure 18 that with the increase of the training cycle, the loss value gradually decreases, which proves that the model we trained tends to converge, and the Acc value gradually increases, which proves that the accuracy of training gradually improves and reaches the expected value of model training.

We used data sets collected from 6 experimentarians in two different environments to conduct multiple sets of comparison tests for evaluation and combined with multiple sets of comparison tests. Finally, we obtained the recognition rate of the empty hall which can reach 91.77%, and the average recognition rate of the classroom is 87.97%. In order to fully describe the recognition accuracy of the sign language gesture recognition method proposed in this paper, we describe the comprehensive recognition results of 30 sign language gestures by confounding evidence, as shown in Figure 19. In general, the recognition results of all gestures were satisfactory, but due to similar gestures, such as "M" and "N" or "CH" and "C," the misjudgment rate of these gestures was relatively high.

5. Conclusion

This paper proposes a CSI-based sign language gesture recognition method, combining real-life applications and the influence of environmental factors on the gesture features of people, using the LOF algorithm for outlier removal and noise reduction and filtering of the collected gesture data by DWT and PCA, and extracting and fusing the time-domain information by DBM. Finally, the gesture data feature values are put into the GRU network for gesture classification and recognition. After various comparison tests, the results show that the average recognition rate of Air-CSL for sign language gestures is 88.93%.

Although we achieved satisfactory accuracy, there were still some limitations. The follow-up work of this paper mainly focuses on the following aspects: (1) improve the robustness of this method and apply it to continuous sign language gesture recognition in different environments; (2) add gesture features with frequency-domain information on the existing basis to describe human gesture features as comprehensively as possible; (3) based on the existing recognition models such as CNN, LSTM, K-NN [40], and SVM,

we will try to use their upgraded versions to compare with our recognition model in the future work.

Data Availability

Data are available on request to the authors. The data that support the findings of this study are available from the corresponding author upon reasonable request. However, as the data involved in this paper are still used in other studies, only part of the data are provided.

Conflicts of Interest

The authors declare no potential competing interests in the paper.

Acknowledgments

References

[1] S. K. Tyagi, S. R. Pokhrel, M. Nemati, D. K. Jain, G. Li, and J. Choi, "Redesigning compound TCP with cognitive edge intelligence for WiFi-based IoT," *Future Generation Computer Systems-The International Of Escience*, vol. 125, pp. 859–868, 2021.

[2] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: gathering 802.11n traces with channel state information," *ACM SIGCOMM Computer Communication*, vol. 41, no. 1, pp. 53–53, 2011.

[3] H. Mora, F. A. Pujol-López, J. C. Mendoza-Tello, and M. R. Morales-Morales, "An education-based approach for enabling the sustainable development gear," *Computers in Human Behavior*, vol. 107, no. 107, article 105775, 2020.

[4] L. M. Haile, K. Kamenov, P. S. Briant et al., "Hearing loss prevalence and years lived with disability, 1990–2019: findings from the Global Burden of Disease Study 2019," *Lancet*, vol. 397, no. 10278, pp. 996–1009, 2021.

[5] W. Cheng, Y. Sun, G. Li, G. Jiang, and H. Liu, "Jointly network: a network based on CNN and RBM for gesture recognition," *Neural Computing& Applications*, vol. 31, no. S1, pp. 309–323, 2019.

[6] S. Saraswat, A. Gupta, H. P. Gupta, and T. Dutta, "An incremental learning based gesture recognition system for consumer devices using edge-fog computing," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 1, pp. 51–60, 2020.

[7] H. Rahaman and V. Dyo, "Tracking human motion direction with commodity wireless networks," *IEEE Sensors Journal*, vol. 21, no. 20, pp. 23344–23351, 2021.

[8] L. Yu, H. Abuella, M. Z. Islam, J. F. O'Hara, C. Crick, and S. Ekin, "Gesture recognition using reflected visible and infrared lightwave signals," *IEEE Transactions On Human-Machine Systems*, vol. 51, no. 1, pp. 44–55, 2021.

[9] H. Brock, I. Farag, and K. Nakadai, "Recognition of non-manual content in continuous Japanese sign language," *Sensors*, vol. 20, no. 19, p. 5621, 2020.

[10] Y. Huang and J. Yang, "A multi-scale descriptor for real time RGB-D hand gesture recognition," *Pattern Recognition Letters*, vol. 144, pp. 97–104, 2021.

[11] Y. Zhou, "Role of human body posture recognition method based on wireless network Kinect in line dance aerobics and gymnastics training," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 9208891, 12 pages, 2021.

[12] Y. Ling, X. Chen, Y. Ruan, X. Zhang, and X. Chen, "Comparative study of gesture recognition based on accelerometer and photoplethysmography sensor for gesture interactions in wearable devices," *IEEE Sensors Journal*, vol. 21, no. 15, pp. 17107–17117, 2021.

[13] Y. Zeng, H. Xiang, N. Zheng, X. Cao, N. Wang, and Z. L. Wang, "Flexible triboelectric nanogenerator for human motion tracking and gesture recognition," *Nano Energy*, vol. 91, no. 91, article 106601, 2022.

[14] H. Han and S. W. Yoom, "Gyroscope-based continuous human hand gesture recognition for multi-modal wearable input device for human machine interaction," *Sensors*, vol. 19, no. 11, p. 2562, 2019.

[15] Y. N. Yang, J. N. Cao, X. L. Liu, and X. Liu, "Door-monitor: counting in-and-out visitors with COTS WiFi devices," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1704–1717, 2020.

[16] J. C. Wang, Z. S. Tian, M. Zhou, J. Wang, X. Yang, and X. Liu, "Leveraging hypothesis testing for CSI based passive human intrusion direction detection," *IEEE Transactions On Vericular Technology*, vol. 70, no. 8, pp. 7749–7763, 2021.

[17] W. Zhuang, Y. Shen, L. Li, C. Gao, and D. Dai, "Develop an adaptive real-time indoor intrusion detection system based on empirical analysis of OFDM subcarriers," *Sensors*, vol. 21, no. 7, p. 2287, 2021.

[18] J. F. Yang, H. Zou, Y. X. Zhou, and L. Xie, "Learning gestures from WiFi: a Siamese recurrent convolutional architecture," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10763–10772, 2019.

[19] K. Niu, F. S. Zhang, Y. H. Jiang et al., "WiMorse: a contactless Morse code text input system using ambient WiFi signals," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9993–10008, 2019.

[20] H. F. Ahmed, H. Ahmad, K. Narasingamurthi, H. Harkat, and S. K. Phang, "DF-WiSLR: device-free Wi-Fi-based sign language recognition," *Pervasive and Mobile Computing*, vol. 69, no. 69, article 101289, 2020.

[21] G. Yuan, X. Liu, Q. Y. Yan, S. Qiao, Z. Wang, and L. Yuan, "Hand gesture recognition using deep feature fusion network based on wearable sensors," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 1–547, 2021.

[22] C. Tan, Y. Sun, G. Li, G. Jiang, D. Chen, and H. Liu, "Research on gesture recognition of smart data fusion features in the IoT," *Neural Computing & Applications*, vol. 32, no. 22, pp. 16917–16929, 2020.

[23] F. L. Liu, W. Zeng, C. Z. Yuan, Q. Wang, and Y. Wang, "Kinect-based hand gesture recognition using trajectory information, hand motion dynamics and neural networks," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 563–583, 2019.

[24] R. Wen, B. P. Nguyen, C. B. Chng, and C. K. Chui, "In situ spatial AR surgical planning using projector-Kinect system," in *Proceedings of the Fourth Symposium on Information and Communication Technology*, pp. 164–171, Danang, Vietnam, 2013.

[25] R. Wen, W. L. Tay, B. P. Nguyen, C. B. Chng, and C. K. Chui, "Hand gesture guided robot-assisted surgery based on a direct augmented reality interface," *Computer Methods and Programs in Biomedicine*, vol. 116, no. 2, pp. 68–80, 2014.

[26] K. C. Lim, S. H. Sin, C. W. Lee et al., "Video-based skeletal feature extraction for hand gesture recognition," in *Proceedings of the 4th International Conference on Machine Learning and Soft Computing (ICMLSC2020)*, pp. 108–112, Haiphong City, Viet Nam, 2020.

[27] J. Park, J. Y. Jang, G. Lee, H. Koh, C. Kim, and T. W. Kim, "A time domain artificial intelligence radar system using 33-GHz direct sampling for hand gesture recognition," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 4, pp. 879–888, 2020.

[28] H. Ding, L. Guo, C. Zhao et al., "RFnet: automatic gesture recognition and human identification using time series RFID signals," *Mobile Networks & Application*, vol. 25, no. 6, pp. 2240–2253, 2020.

[29] 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*, vol. 7, no. 9, pp. 8296–8317, 2020.

[30] Y. Gu, Y. T. Wang, Z. Liu, J. Liu, and J. Li, "SleepGuardian: an RF-based healthcare system guarding your sleep from afar," *IEEE Network*, vol. 34, no. 2, pp. 164–171, 2020.

[31] D. H. Zhang, Y. Hu, Y. Chen, and B. Zeng, "BreathTrack: tracking indoor human breath status via commodity WiFi," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3899–3911, 2019.

[32] H. Kang, Q. Zhang, and Q. Y. Huang, "Context-aware wireless-based cross-domain gesture recognition," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13503–13515, 2021.

[33] L. Wang, K. Sun, H. P. Dai et al., "WiTrace: centimeter-level passive gesture tracking using OFDM signals," *IEEE Transactiona On Mobile Computing*, vol. 20, no. 4, pp. 1730–1745, 2021.

[34] X. R. Ma, Y. N. Zhao, L. Zhang, Q. Gao, M. Pan, and J. Wang, "Practical device-free gesture recognition using WiFi signals based on metalearning," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 228–237, 2020.

[35] D. Jiang, M. Li, and C. Xu, "Wigan: a WiFi based gesture recognition system with GANs," *Sensors*, vol. 20, no. 17, p. 4757, 2020.

[36] N. Q. Khanh Le, Q. H. Nguyen, X. Chen, S. Rahardja, and B. P. Nguyen, "Classification of adaptor proteins using recurrent neural networks and PSSM profiles," *BMC genomics*, vol. 20, no. S9, pp. 1–9, 2019.

[37] M. Z. Alom, T. M. Taha, C. Yakopcic et al., "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, p. 292, 2019.

[38] H. Yan, Y. Zhang, Y. Wang, and K. Xu, "WiAct: a passive WiFi-based human activity recognition system," *IEEE Sensors Journal*, vol. 20, no. 1, pp. 296–305, 2020.

[39] J. Yin, Q. Zhang, R. Zhao, and D. Wang, "WiLAR: a location-adapted action recognition system based on WiFi," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, Nanjing, China, 2021.

[40] B. P. Nguyen, W. L. Tay, and C. K. Chui, "Robust biometric recognition from palm depth images for gloved hands," *IEEE Transactions On Human-Machine Systems*, vol. 45, no. 6, pp. 799–804, 2015.