Received 29 January 2024; revised 9 April 2024; accepted 22 April 2024. Date of publication 24 April 2024; date of current version 27 May 2024.

*Digital Object Identifier 10.1109/OJAP.2024.3393117*

# Wi-Fi-Based Human Activity Recognition for Continuous, Whole-Room Monitoring of Motor Functions in Parkinson's Disease

**SHIH-YUAN CHEN[1] AND CHI-LUN LIN[1,2] (Member, IEEE)**

[1]Department of Mechanical Engineering, National Cheng Kung University, Tainan City 70101, Taiwan

[2]Medical Device Innovation Center, National Cheng Kung University, Tainan City 70101, Taiwan

CORRESPONDING AUTHOR: C.-L. Lin (e-mail: linc@ncku.edu.tw)

This work was supported in part by the National Technology and Science Council, Taiwan, under Grant NSTC 111-2221-E-006-038, and in part by the Higher Education Sprout Project, Ministry of Education to the Headquarters of University Advancement at National Cheng Kung University.

**ABSTRACT** Parkinson's disease is a progressive neurodegenerative disorder with significant fluctuations throughout the day, making accurate drug treatment difficult. A home-based long-term monitoring system is essential to address this challenge. Contemporary approaches to activity monitoring have focused on wearable devices and computer vision systems. Wearable devices are often uncomfortable and not ideal for long-term monitoring, while computer vision is plagued with significant privacy concerns. In this context, Wi-Fi sensing presents itself as an advantageous alternative due to its non-invasive and privacy-preserving properties. However, current human activity recognition methodologies lack the specificity to identify disease-related symptoms within everyday activities. Furthermore, the efficiency of human activity recognition methods in processing continuous data streams in real time is a crucial aspect that needs thorough assessment. This study proposes a novel approach for human activity recognition using Wi-Fi signals. Traditional methods for signal processing are avoided by converting the ratio of channel state information from antenna pairs into images. These images are then processed using a convolutional neural network to detect movements related to diseases in a large dataset. The experiments utilize a laptop PC with Intel Wi-Fi Link 5300 and a receiver equipped with three external 12 dB omnidirectional antennas in the 2.4 GHz band and cover various daily activities. The proposed method has demonstrated remarkable accuracy, with an average recognition rate of 93.8% in validation. It also showcased a consistent accuracy range of 91.9% to 95.2% in generalization tests, proving its effectiveness in different environments, with various individuals, and under assorted Wi-Fi configurations. A performance test of our method revealed that it processes raw CSI to recognition results in just 0.65 seconds per second of data, highlighting its potential for real-time applications.

**INDEX TERMS** Wireless sensing network, channel state information, human activity recognition, Parkinson's disease, health monitoring.

## I. INTRODUCTION

PARKINSON'S disease (PD) affects approximately 1% of people aged 60 years or older and 1%–3% of people aged over 80 year [1]. Currently, more than seven million people worldwide suffer from PD, and its prevalence has doubled in the past 25 years [2]. The disease is typically characterized by a combination of motor features, such as resting tremor, bradykinesia, abnormal gait, and non-motor symptoms, which significantly affect the lives of patients.

There is currently no cure for Parkinson's disease (PD). The disease can only be managed with drugs such as levodopa, monoamine oxidase inhibitors, and dopamine receptor agonists to control changes in the condition. However, the side effects of these drugs, such as dyskinesia, confusion, and hallucinations [3], can cause more suffering

than the disease itself. Therefore, personalized medication tailored to the patient's condition and home monitoring systems that help physicians monitor patients outside clinics are particularly important.

Wearable devices and computer vision are two widely accepted health monitoring methods used to assess a patient's health status. Wearable devices such as gyroscopes and accelerometers have become a mainstream research direction owing to their low cost and ease of use [4]. Nevertheless, wearing these sensors can cause inconvenience and discomfort, limiting the patient's mobility [5]. Computer-vision approaches analyze recorded videos of patients and quantify their motor functions to evaluate disease progression. However, filming the patient may invade the privacy of the patient, and the method is limited by environmental conditions, such as lighting and space.

Wireless sensing technology enables the contactless detection of human activities in a person's natural state without any equipment. This method captures signal data containing no personally identifiable information, thereby preserving privacy. Recent studies have developed sensing devices based on radio frequencies (RF) to locate and recognize indoor human movements [6]. In another study, an RF device was built to track the gait of patients with PD at home [7].

Wi-Fi-based sensing has become a popular research topic. The channel state information (CSI) of Wi-Fi contains fine-grained information that reflects the occurrence of activities and/or environmental changes. This technology can be implemented using commodity Wi-Fi devices, making it a low-cost, easy-to-deploy, and widely applicable solution for health monitoring.

Wi-Fi CSI detects movement in space by analyzing the electromagnetic waves emitted by Wi-Fi devices. In recent years, the use of CSI for human activity recognition (HAR) has gained increasing attention because of its ability to provide information on multiple channels and record the amplitude and phase information for each orthogonal frequency-division multiplexing (OFDM) subcarrier, thereby providing a wealth of features for HAR.

Recent studies developed HAR methods based on Wi-Fi CSI to identify violent activities [8], daily activities [9], and gestures [10]. These methods rely on processing the CSI amplitude or phase through denoising, filtering, and/or time-frequency transformation and using deep learning techniques to extract features from the CSI for activity recognition. However, some of these methods face the impact of changing spatial environments and device placements, which in turn affect the accuracy of the recognition model, whereas others lack a complete evaluation of the versatility of the model. Furthermore, data processing procedures, such as filtering and time-frequency transformation, often require significant computational time. Therefore, developing a deep learning model with generalizability for real-time HAR remains challenging.

Inspired by these studies, we propose a HAR system based on Wi-Fi CSI to enable the detection of disease-related movements of patients with PD for long-term health monitoring. The proposed HAR system combines the CSI ratio model [11] with a convolutional neural network (CNN) to distinguish between daily activities and disease-related movements. We use the CSI ratio for its advantage of eliminating the influence of static paths and requiring less computation than traditional filters. Moreover, the CSI ratio sequence is converted to images, which are two-dimensional arrays of data with time on one dimension and subcarriers on the other, making them suitable for a 2D-CNN to learn human activity features.

The main contributions of this paper can be summarized as follows.

1) We proposed a novel HAR system with a data processing mechanism and a highly precise deep learning model that can accurately identify human activities, particularly PD-related movements, from continuous CSI streams.
2) We took advantage of the CSI ratio model that automatically cancels the effect of environmental noise, which enhanced the adaptability of the proposed HAR system to different usage conditions. The need for a filtering process to denoise the data was also eliminated, leading to excellent computing efficiency for real-time health monitoring applications.
3) We conducted comprehensive performance evaluations of the proposed HAR system, including analysis of the recognition accuracy and performance when used in long-term scenarios and various generalization tests in terms of the environment, person, and Wi-Fi arrangement.

## II. RELATED WORKS

The development and popularity of Wi-Fi devices have led to increased attention being paid to HAR systems based on Wi-Fi signals. In early studies, most researchers used the Received Signal Strength Indicator (RSSI) of Wi-Fi signals to identify human activity. In 2006, Woyach et al. [12] conducted experiments using Crossbow's MICA2 and MICAz (equipped with an IEEE 802.15.4-compliant 2.4GHz Chipcon CC2420 radio) that sent a packet to each other every 100 ms at a transmit power of $-10$ dBm and found that the human motion caused a series of attenuation and multipath effects in the RSSI. In 2007, Youssef et al. [13] further confirmed that RSSI data can detect the presence of human bodies using Wi-Fi 2.4 GHz signals with simple feature calculation methods, such as moving averages and variances. Moreover, they proved that the position of a human body can be tracked using the RSSI value. However, owing to the attenuation and multipath effects of signals, the measurement of RSSI is unstable, creating difficulties in finer-grained activity recognition.

With the development of CSI extraction tools for commercial network interface controllers (e.g., Intel 5300 [7] and Atheros 9580 [14]), the performance of HAR technologies has significantly improved by grafting from RSSI- to CSI-based methods [15].

Existing CSI-based HAR methods can be divided into model- and pattern-based methods. Model-based approaches interpret the received signals using mathematical models to show the relationship between the signals and real movements. This type of approach is based on mathematical theories and, therefore, has better adaptability to different data collection environments than pattern-based approaches. The models proposed in current studies include the Fresnel zone, CSI speed, and models performing calculations based on the multiple signal classification (MUSIC) algorithm, such as the angle of arrival (AoA) and time of flight (ToF) [16], [17].

The focus of previous studies on model-based approaches for HAR was scarce. Few applications have focused on human localization or the detection of individual movements [11], [18], [19], such as breathing detection and finger tracking. Li et al. [19] developed an indoor localization method based on the MUSIC algorithm, which was tested in the 5 GHz frequency band employing an unutilized 40 MHz channel. Zeng et al. [11] proposed a method based on the CSI ratio model to enable wall-penetrating detection of breathing and significantly expanded the detection range of CSI. The research utilized an Intel 5300 Wi-Fi card, with the central frequency set at 5.24 GHz and a 20 MHz bandwidth, to gather a total of 56 sub-carriers of CSI data. This data was collected from the communication between a transmitter with one antenna and a receiver with two antennas. Using similar Wi-Fi specifications (5.24 GHz frequency band, 20 MHz bandwidth, 15 dBm transmitting power, and monitor mode), Wu et al. [18] used the CSI ratio to develop a sub-wavelength level tracking system for finger drawing.

However, some limitations exist in the model-based human sensing methods using Wi-Fi CSI. First, although the model-based methods are robust to environmental changes, they cannot always achieve significant performance. For example, the model-based methods may not perform as well as pattern-based methods for some coarse-grained applications. Second, it is challenging to establish mathematical relationships between CSI changes and some complex human activities. Third, the model-based methods may not work well in non-line-of-sight scenarios due to the effect of the wall or other occlusion on the signal propagation.

Pattern-based methods often employ machine and deep learning to extract features from CSI. Although this type of approach may not be as adaptable to various situations as model-based methods, it often achieves excellent results for specific HAR tasks and recognizes a wide range of human activity types. Wang et al. [20] used the Intel 5300 tool and commercial Wi-Fi devices (IEEE 802.11n AP mode at 2.4GHz) and machine learning to recognize different mouth motions and achieved a detection accuracy of 91% within six words. Bu et al. [10] used transfer learning techniques for gesture recognition and converted CSI into image matrices to transform gesture recognition problems into image classification problems. Liu et al. [8] successfully detected violent actions by converting filtered CSI into grayscale images.

Most research on human activity recognition uses pattern-based methods to automatically learn features from a denoised CSI signal, thereby simplifying the feature exploration process. However, the influence of environmental noise can be too large to be eliminated, leading to a significant decrease in HAR accuracy, particularly when an HAR method is used in a new environment.

To date, most studies have not considered HAR systems in terms of real-time or continuous detection. First, the current methods typically process CSI data with short-time Fourier transform (STFT) or discrete wavelet transform (DWT) and train the model with a time-frequency map [21], [22]. Although such data processing methods can achieve a high degree of accuracy in recognizing motion, they require a large computation time, making it difficult to achieve real-time recognition. Second, most current motion detection technologies can only detect motion for a fixed period of time rather than continuously detecting the motion and time of occurrence [23]. This makes it difficult to apply these technologies in health monitoring applications.

This study proposes a new HAR system that combines both model- and pattern-based approaches. The CSI ratio model [11] was used to cancel the static path of the signal propagation from the raw CSI data, thereby avoiding environmental noise and increasing the adaptability of the method to environmental changes. By converting the CSI ratio into images (CSI images), the pattern-based approach was used to implement a CNN-based model. The proposed HAR system was designed to identify PD-related movements from daily activities in a long-term recorded CSI sequence. Segments of the identified movements could be further quantified to understand the variation of the disease conditions on a daily basis, thereby enabling continuous health monitoring.

## III. PROPOSED METHODS

### A. SYSTEM ARCHITECTURE

The architecture of the proposed method is shown in Figure 1. The system consists of five processes: (1) CSI ratio measurement, which collects the CSI phase and amplitude values to calculate the CSI ratios; (2) CSI image conversion, which converts a sequence of CSI ratios into a grayscale image and applies contrast enhancement; (3) activity recognition, which uses the sliding window technique to produce fixed-size images from the input CSI image, classify them via a CNN model, and produces data segments of each identify activities based on the predicted probability distributions (note that the CNN model is based on the VGG19 network and is trained using CSI images). These data segments of the target categories, such as walking and tremor, are ready for motion analysis. Detailed information regarding each process is provided in the following subsections.

### B. CSI RATIO MEASUREMENT

The CSI is calculated based on the electromagnetic wave characteristics, which consider a series of influences, such as reflection, scattering, and attenuation of the wireless signal

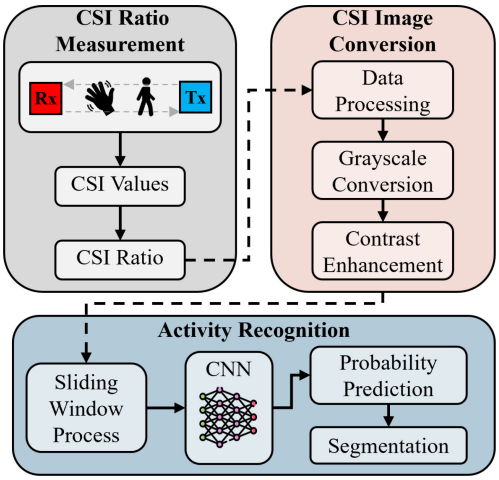


**FIGURE 1.** The system architecture of the proposed HAR system.

from the transmitter to the receiver. The CSI matrix can be expressed as:

$$\mathbf{H}^t = \begin{bmatrix} h^t_{11} & \cdots & h^t_{1k} \\ \vdots & \ddots & \vdots \\ h^t_{i1} & \cdots & h^t_{ik} \end{bmatrix} \tag{1}$$

The matrix $\mathbf{H}^t$ represents the CSI at time t, where i and k correspond to the numbers of receiving antennas and subcarriers, respectively. The element $h^t_{ik}$ in the matrix can be expressed as follows:

$$\mathbf{H}^t = a^t_{ik} e^{jsin(\theta^t_{ik})} \tag{2}$$

The variables $a^t_{ik}$ and $\theta^t_{ik}$ represent the calculated amplitude and phase, respectively.

Under ideal conditions, the static path components of CSI are composed of a combination of line-of-sight (LoS) and other static object reflections (denoted as $h_s$), whereas the dynamic components are generated by reflections caused by human body movement (denoted as $h_d$) [11], as expressed in Eq. (3), with the matrix form given in Eq. (4).

$$h^t_{ik} = h^t_{s_{ik}} + h^t_{d_{ik}} \tag{3}$$

$$\mathbf{H}^t = \mathbf{H}^t_s + \mathbf{H}^t_d \tag{4}$$

The dynamic path $h_d$ can be further represented by the change in amplitude $a^t_{ik}$ and the phase shift $e^{-j2\pi \frac{d(t)}{\lambda}}$ caused by the target displacement $d^t$. The equation for this is shown in (5), whereas (6) presents the same formula in matrix form:

$$h^t_{ik} = h^t_{s_{ik}} + a^t_{ik} e{-}J2\pi \frac{d^t}{\lambda} \tag{5}$$

$$\mathbf{H}^t = \mathbf{H}^t_s + \mathbf{A}^t e{-}J2\pi \frac{d^t}{\lambda} \tag{6}$$

Under realistic conditions, the phase $\theta^t_{ik}$ in the channel state information is affected by the carrier frequency offset (CFO) and sampling frequency offset (SFO), owing to the inconsistent carrier frequency and sampling rate between the transmitter and receiver. The attenuation and phase shift caused by these two offsets can be represented by $\mathbf{A}^t_{noise} \times \exp^{-j\theta^t_{offset}}$; thus, the channel state information affected by the offset can be written as

$$\mathbf{H}^t = \mathbf{A}^t_{noise} \exp^{-j\theta^t_{offset}} \times \left(\mathbf{H}^t_s + \mathbf{H}^t_d\right) \tag{7}$$

By calculating the CSI ratio between the two antennas (represented as $\mathbf{H}^t_R$ [11], we can eliminate the effects of the two offsets, as follows:

$$\begin{aligned} \mathbf{H}^t_R &= \frac{A^t_{noise} e^{-j\theta^t_{offset}} \times \left(\mathbf{H}^t_{S_1} + \mathbf{H}^t_{d_1}\right)}{A^t_{noise} e^{-j\theta^t_{offset}} \times \left(\mathbf{H}^t_{S_2} + \mathbf{H}^t_{d_2}\right)} \\ &= \frac{\mathbf{A}^t_1 e^{-j2\pi \frac{d^t_1}{\lambda}} + \mathbf{H}^t_{S_1}}{\mathbf{A}^t_2 e^{-j2\pi \frac{(d^t_2 - d^t_1)}{\lambda}} e^{-j2\pi \frac{d^t_1}{\lambda}} + \mathbf{H}^t_{S_2}} \end{aligned}$$

### C. CSI IMAGE CONVERSION

We used the phase of the CSI ratio to create CSI images [24], which ensured that the data fell between $\pm\pi$, so that the small motion feature would not be lost in the process of conversion to CSI images. The CSI ratio phase was calculated for two arbitrary antennas from the three antennas on the receiver, resulting in three data sequences: $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_3$. Each $\mathbf{R}_i$ contains CSI phase measurements over time from 30 subcarriers. The root mean square (RMS) values of the three $\mathbf{R}_i$ were calculated (Eq. (9)) and transformed to a range between 0 and 255 (Eq. (10)), which can later be related to the pixel values of the greyscale image. Finally, we obtained the matrix $\mathbf{M}(s, t)$, where $s$ and $t$ represent the subcarrier number and time, respectively. The formulae are as follows:

$$\mathbf{RMS}(s,t) = \sqrt{\frac{\Sigma^3_{i=1} \mathbf{R}^2_i(s,t)}{3}} \quad \{t|\ t \in \mathbb{Z} \cap 1 \le t \le L\};\ \{s|\ s \in \mathbb{Z} \cap 1 \le s \le 30\} \tag{8}$$

$$\mathbf{M}(s,t) = \mathbf{RMS}(s,t) \times \frac{255}{max(\mathbf{RMS})} \quad \{t|\ t \in \mathbb{Z} \cap 1 \le t \le L\};\ \{s|\ s \in \mathbb{Z} \cap 1 \le s \le 30\} \tag{9}$$

The data sequence $\mathbf{M}(s, t)$ was used to generate a CSI image. However, random packet loss is inevitable during network communication and should be considered. We developed an algorithm to detect occurrences of packet loss and label them properly (Fig. 2). To begin, the algorithm reads a 1 s segment of $\mathbf{M}$ (which is approximately 250 data points when the network condition is normal) and checks its data length. If the data length is less than 56, the data segment of $\mathbf{M}$ is considered packet loss, and an image matrix with size of 224 × 112 pixels with values of 255 is generated (i.e., an all-black image). Otherwise, the values in the segment were transformed into pixels of a grayscale image, where the Y-dimension was enlarged from 30 (subcarriers) to 224 pixels, and the X-dimension was enlarged or reduced to 112

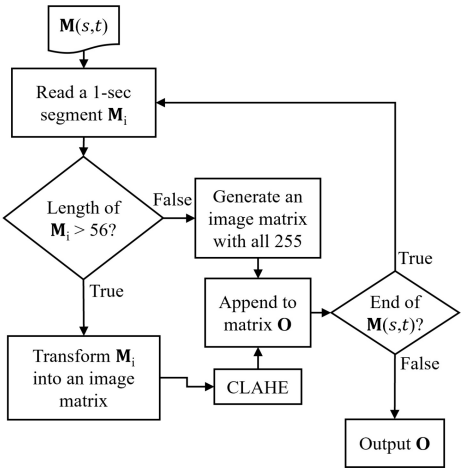


**FIGURE 2.** The process to convert a full-length M(*s*, *t*) to a CSI Image.

pixels through bilinear interpolation. The resulting image matrix was enhanced using the CLAHE algorithm [25]. Therefore, the final image matrix resulted from the 1 s segment of **M** is with size of 224 × 112 pixels that contain all-black values or meaningful content and appended to the output matrix **O** (empty at the beginning). The threshold of data length 56 for checking the packet loss ensures the converted image of 1 s has at least a half second of data. This process is repeated throughout the sequence of **M**.

### D. CNN MODEL TRAINING

While training deep learning models in the medical field, researchers frequently face the challenge of insufficient data. Therefore, we applied transfer learning techniques in the training process and selected VGG [26] as the base model to extract features from the collected CSI images. VGG is configured with multiple convolution layers of uniform size 3 × 3, a max-pooling layer of size 2 × 2, a flattened layer, and three fully connected layers, which not only improve the network depth but also make model training a time-consuming task. To address this issue, we replaced the flattened layer with global average pooling to flatten the output from the convolution layer while reducing the number of trainable parameters. Furthermore, we reduced the total number of parameters of the fully connected layers and introduced dropout layers to prevent overfitting. The model constructed using VGG19 is shown in detail in Fig. 3.

### E. ACTIVITY RECOGNITION FROM A LONG-TERM RECORDED CSI IMAGE

The trained CNN model requires an input CSI image with a specified size of 224 × 224 pixels. We implemented an algorithm to produce a sequence of images of the required size by shifting a sliding window (224 × 224 pixels) over the time length (*l*) of a long-term recorded CSI image, which had a size of 224 × 112*l*. At each step, the sliding window was shifted forward by 56 pixels (equal to 0.5 *s*) to capture an image, and the CNN model read the image and output a probability distribution over the predefined categories of

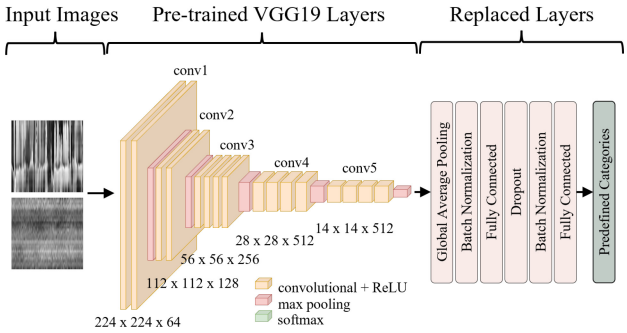


**FIGURE 3.** The VGG19-based CNN model.

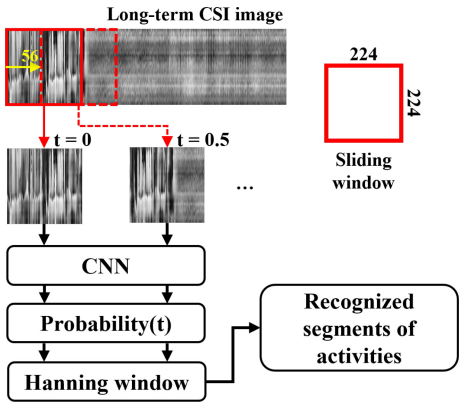


**FIGURE 4.** The process of recognizing human activities in a long-term recorded CSI image using the trained CNN model.

activities. Finally, the probability distribution was generated over all the time steps and filtered using a Hanning window (window length = 9), as shown in Fig. 4. The most likely activity at each timestep was selected.

## IV. EXPERIMENTS

### A. DATA COLLECTION

The Linux 802.11n CSI Tool [27] was used to collect CSI data for all experiments in this study. The CSI Tool was set to station (STA) mode and run on a DELL P40G i8 laptop with an Intel 5300 wireless network interface controller connected to three external 12 dB omnidirectional antennas as the receiver (Rx). The Rx was connected to a D-Link DIR-853 router as the transmitter (Tx). For data communication, a 2.4 GHz band was used, and the CSI values were measured on ping packets sent from the Tx to the Rx with a 200 Hz ping rate and 250 Hz packet sampling rate.

We aimed to identify PD-related motor features during daily activities. The following four main categories of activities were included: walking (W), tremor (T), body movements (BM), and static postures (SP). The subjects were asked to walk (1) naturally at a slow or fast pace (1–2 steps/min). For tremor (2), the subject mimicked two key features of the hand rest tremor: a frequency of 3–7 Hz and an amplitude of 5–10 cm [28]. Body movements included head nodding (3), hand shaking (4), scratching (5), and leg kicking (6), which involved movements of a single body part and were designed to produce signal responses similar

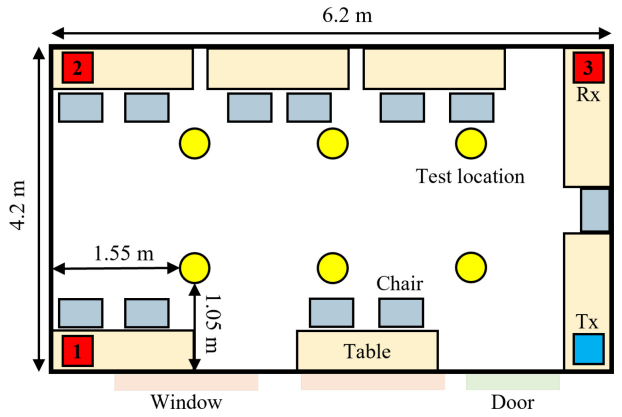


**FIGURE 5.** The interior layout diagram of room T for training data collection.

**TABLE 1.** Physical information of the subjects in the training dataset.

| Subject No. | Age | Height (cm) | Weight (kg) |
|---|---|---|---|
| 1 | 23 | 178 | 80 |
| 2 | 24 | 169 | 52 |
| 3 | 25 | 181 | 90 |

to tremor and walking, respectively. For static postures, sitting (7) and standing (8) postures were chosen. Among the four categories, walking and tremor were the targets to be recognized, whereas the others were either common daily activities or specifically selected to challenge model training and to make the model more applicable to real-use conditions.

We conducted the experiments in a university discussion room (referred to as room T), which was 6.2 m long and 4.2 m wide, as shown in Fig. 5. Room T had solid cement walls on all sides and one side had four windows and a door. During the experiments, tables and chairs were arranged against the walls, and the door and windows were closed.

Three healthy subjects aged 22–25 years were recruited to implement the experiments for collecting training data. Their physical data are presented in Table 1.

While the walking activity allowed a subject to move arbitrarily in the room, the activities in the other three categories were performed at one of six locations, as indicated by the yellow circles. At each location, the subjects were asked to perform each activity twice for 30 $s$. Therefore, the length of the data collected for training was:

$$3(sub) \times 6(loc) \times 8(act) \times 2(rep) \times 30(s) = 8640(s)$$

### B. CNN MODEL TRAINING

The collected CSI data sequences with a total length of 8,640 s were converted into CSI images, producing 3,054, 3,074, 3,048, and 2,997 labeled images (with a size of 224 × 224 pixels, representing an activity of 2 $s$) of body movements, static postures, tremor, and walking, respectively. The data were divided into training, testing, and validation sets in a ratio of 6:2:2 to train the CNN model. Table 2 lists the hyperparameters used during the training process.

**TABLE 2.** Hyperparameters for training the CNN model.

| Hyperparameters | Setting |
|---|---|
| **Optimization:** | |
| Optimizer | SGD |
| Learning rate | $5e-4$ |
| Decay | $1e-6$ |
| Momentum | 0.9 |
| **Training:** | |
| Batch size | 18 |
| Epoch | 1000 |
| Early stopping patience | 20 |
| Early stopping monitor | Validation accuracy |
| Loss function | Cross-Entropy |

### C. HUMAN ACTIVITY SEGMENTATION

The recognition performance of our HAR system was tested using 50 $s$ CSI sequences, which were collected by additional experiments in Room T with the same Wi-Fi arrangement (Fig. 5). In each experiment, subject 1 (Table 1) performed five 10 $s$ activities in the following order: walking arbitrarily in the room (walking), sitting in the nearest chair (static posture), performing hand rest tremor (tremor), sitting quietly in the same chair (static posture), and performing a random body movement from the four predefined activities (body movement).

All experiments were filmed simultaneously, and the recorded footage was manually examined to label the type of activity at intervals of 0.5 $s$. For a CSI sequence, we compared the activity type of each 0.5 $s$ interval classified by our HAR system with the labeled data and added 0.5 $s$ to the error if the classified type was incorrect. Therefore, each CSI sequence obtained an accumulated error ($E_i$), and the accuracy ($ACC$) was calculated using Eq. (11) and Eqn. 12, where $E_total$ is the sum of $E_i$ from all 10 CSI sequences, and $T_total$ is the total time of all 10 CSI sequences. Note that the time length of a CSI sequence recorded for an experiment would slightly exceed 50 $s$ owing to some buffer time at the beginning and end of the experiment. The exact duration of each experiment is presented in the results section.

$$E_{total} = \Sigma_{i=1}^{10} E_i \tag{10}$$

$$ACC = 1 - \frac{E_{total}}{T_{total}} \tag{11}$$

### D. GENERALIZATION ASSESSMENT

This section investigates the generalization of the proposed HAR system when the usage conditions differ from those in the training environment. The capability of the system to accurately classify activities for new environmental layouts, human subjects, and Wi-Fi arrangements was assessed. For each assessment, ten 50 $s$ CSI sequences were collected using the same scenario described in Section IV-C. A description of each assessment is provided in the following subsections.

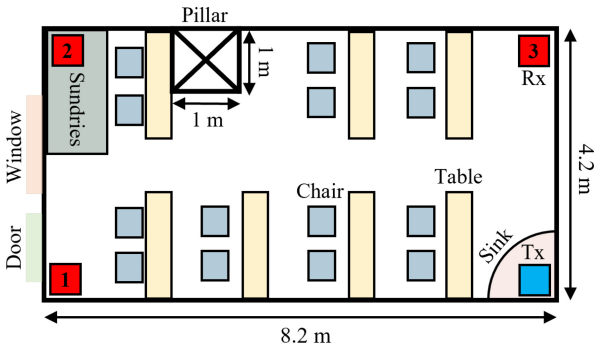


**FIGURE 6.** The interior layout diagram of room V for validation experiments.

**TABLE 3.** Physical information of the subjects in the generalization test.

| Subject | Age | Height (cm) | Weight (kg) |
|---|---|---|---|
| 4 | 23 | 178 | 73 |
| 5 | 23 | 165 | 52 |
| 6 | 24 | 180 | 70 |
| 7 | 22 | 170 | 93 |
| 8 | 23 | 169 | 60 |

#### 1) IMPACT OF ENVIRONMENTAL LAYOUT

Another discussion room (room V) with regularly arranged furniture items was selected to increase the complexity of the environment, as shown in Fig. 6. Room V is a longer space than room T and has a larger number of static objects inside it to create more complex signal propagation paths. The test subject was subject 1 (Table 1), and the transmitter and receivers were placed at the four corners, which were the settings used for the training data.

#### 2) PERSON-GENERALIZATION

For this evaluation, we included five new subjects who participated in the data collection. Their detailed physical information is presented in Table 3. Using the scenario described in Section IV-C, two 50 *s* CSI sequences were collected for each participant in room T, resulting in a dataset with ten 50 *s* CSI sequences.

#### 3) WI-FI ARRANGEMENT

In this section, we test the adaptability of the HAR system to two additional Wi-Fi arrangements (Fig. 7). In Room T, we moved the positions of the Tx and Rx from the four corners to (a) all three Rx antennas placed at a single point and (b) three Rx antennas placed along one side of the room, whereas the Tx was placed on the opposite side for both setups. Ten 50 *s* CSI sequences were collected for each setup on subject 1.

### *E. LONG-TERM TESTING*

To test the stability and efficiency of our HAR system for continuous monitoring applications, we designed a 30-minute action script for subject 1 to perform in Room T, where all tables and chairs were arranged like a normal discussion room (Fig. 8). The script contained a sequence of activities

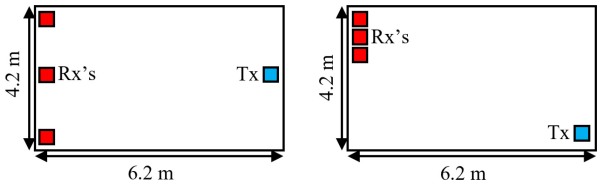


**FIGURE 7.** Two new Wi-Fi arrangements used in the generalization test: linear (left) and single-point arrangements (right).

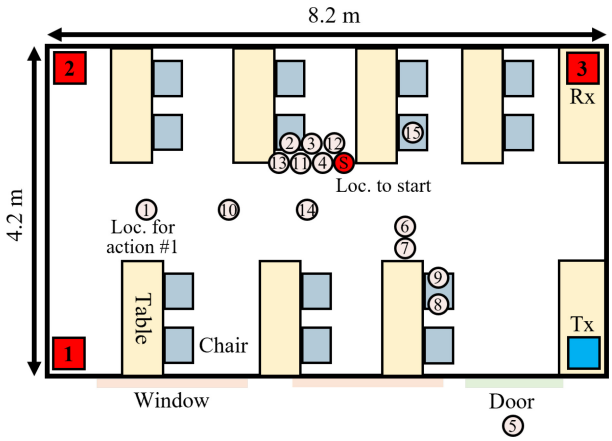


**FIGURE 8.** The experimental setup for the long-term test in room T. The circle S indicates the start location of the test, while circles 1–15 represent the sequences of activities to perform, as listed in the action script (Table 4).

and instructions on where and what activities should be performed, as shown in Figure 8. There were 15 sessions of 2 *min* each in the 30-minute experiment. The subject was reminded at the beginning of each session by a timer to perform the next activity in sequence for 5–10 s and return to a stationary state (standing, sitting, or leaving the room). Fifteen sessions of activities were conducted, including three of tremors, five of walking, six of body movements, and one in which the subject left the room, leaving no person. The entire experiment was filmed and examined manually to apply accurate timestamps for the activities as ground truth, which were used to evaluate the accuracy of the HAR system. The computational performance of the system was also evaluated.

## V. PERFORMANCE EVALUATION

### *A. HUMAN ACTIVITY SEGMENTATION*

Compared with the labeled data from the recorded footage, the errors in the 10 experiments ranged from 1 to 6 *s* over a duration of 53–55 *s*, resulting in an *ACC* of 93.82%. The largest proportion of errors was attributed to walking activities, followed by static postures, tremors, and body movements. The comprehensive results are presented in Table 5.

### *B. GENERALIZATION*

#### 1) IMPACT OF ENVIRONMENTAL LAYOUT

The proposed HAR system achieved an average *ACC* of 95.19% when segmenting the activities performed in 10 experiments in Room V, which slightly outperformed the

**TABLE 4.** The 30-minute action script for the long-term test. Each entry specifies a dynamic activity to perform for 5–10 *s* and a static posture to finish with. Activities 1, 10, and 14 require the subject to walk back and forth in the room, while activities 5 and 6 ask the subject to walk out of the room and return back in. For each walking activity, a location number is given to instruct where to stop.

| Seq | Activity | Finished with |
|---|---|---|
| 1 | walking | sitting at 2 |
| 2 | nodding | sitting |
| 3 | tremor | sitting |
| 4 | scratching | sitting |
| 5 | walking out | standing at 5 |
| 6 | walking in | standing at 6 |
| 7 | kicking | sitting at 8 |
| 8 | hand shaking | sitting |
| 9 | tremor | sitting |
| 10 | walking | sitting at 11 |
| 11 | tremor | sitting |
| 12 | nodding | sitting |
| 13 | scratching | sitting |
| 14 | walking | sitting at 15 |
| 15 | sitting | end |

**TABLE 5.** The errors (*s*) of the proposed HAR system in segmenting the CSI sequences into the four categories of human activities and overall accuracy. SP: static posture, BM: body movement, T: tremor, W: walking.

| No. | SP | BM | T | W | $E_{total}$ | Duration | $ACC$ |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 0 | 0 | 2 | 4 | 55 | 92.73% |
| 2 | 1.5 | 0 | 0 | 0 | 1.5 | 55 | 97.27% |
| 3 | 0 | 0 | 0 | 7.5 | 7.5 | 53 | 85.85% |
| 4 | 1 | 0 | 0 | 0 | 1 | 55 | 98.18% |
| 5 | 1.5 | 0 | 0 | 2.5 | 4 | 55 | 92.73% |
| 6 | 0 | 0 | 0 | 1.5 | 1.5 | 55 | 97.27% |
| 7 | 1 | 0 | 0 | 0.5 | 1.5 | 55 | 97.27% |
| 8 | 1.5 | 0 | 0 | 0 | 1.5 | 53 | 97.17% |
| 9 | 2 | 0 | 1 | 2 | 5 | 53 | 90.57% |
| 10 | 2 | 0 | 3 | 1 | 6 | 53 | 88.68% |

results obtained in Room T (93.82%). We conducted a t-test on two data series, $E_i$ of this experiment (at location V) and the $E_i$ of the human-activity segmentation experiment (at location T). The null hypothesis was that the difference between these two environments would not affect the outcomes. The t-test obtained a p-value of 0.457 ($< 0.05$), indicating that the performance of our HAR system was not affected by changes in the environment.

#### 2) PERSON-GENERALIZATION

The 10 experiments of 50 s each conducted with the five new human subjects had an average ACC of 91.92%, which was 1.9% lower than the validation result (93.82%) tested with the original three subjects (Section IV-A). The errors in the 10 experiments ranged from 2.5 to 7.5 *s* over durations ranging from 53 to 55 *s*, as shown in Table 7. However, the largest error was attributed to the segmentation of walking activities.

**TABLE 6.** The errors (*s*) of the proposed HAR system tested in Room V by segmenting the CSI sequences into the four categories of human activities and overall accuracy. SP: static posture, BM: body movement, T: tremor, W: walking.

| No. | SP | BM | T | W | $E_{total}$ | Duration | $ACC$ |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 0 | 0 | 2 | 4 | 55 | 92.73% |
| 2 | 1.5 | 0 | 0 | 0 | 1.5 | 55 | 97.27% |
| 3 | 0 | 0 | 0 | 7.5 | 7.5 | 53 | 85.85% |
| 4 | 1 | 0 | 0 | 0 | 1 | 55 | 98.18% |
| 5 | 1.5 | 0 | 0 | 2.5 | 4 | 55 | 92.73% |
| 6 | 0 | 0 | 0 | 1.5 | 1.5 | 55 | 97.27% |
| 7 | 1 | 0 | 0 | 0.5 | 1.5 | 55 | 97.27% |
| 8 | 1.5 | 0 | 0 | 0 | 1.5 | 53 | 97.17% |
| 9 | 2 | 0 | 1 | 2 | 5 | 53 | 90.57% |
| 10 | 2 | 0 | 3 | 1 | 6 | 53 | 88.68% |
| Average | 0.6 | 0.3 | 0.6 | 1.05 | 2.55 | 53 | 95.19% |

**TABLE 7.** The overall accuracy and errors (*s*) of the proposed HAR system tested with five new subjects by segmenting the CSI sequences into the four categories of human activities. SP: static posture, BM: body movement, T: tremor, W: walking.

| Sub | Exp | SP | BM | T | W | $E_{total}$ | Dur | $ACC$ |
|---|---|---|---|---|---|---|---|---|
| 4 | 1 | 0 | 0 | 0 | 0 | 0 | 55 | 100% |
| | 2 | 1.5 | 0 | 0 | 6 | 7.5 | 53 | 85.85% |
| 5 | 1 | 4 | 0 | 0 | 0 | 4 | 53 | 92.45% |
| | 2 | 0 | 0 | 0 | 2.5 | 2.5 | 53 | 95.28% |
| 6 | 1 | 2 | 0 | 0 | 2 | 4 | 53 | 92.45% |
| | 2 | 4 | 0 | 0 | 2.5 | 6.5 | 53 | 87.74% |
| 7 | 1 | 0 | 0 | 0 | 4.5 | 4.5 | 53 | 91.51% |
| | 2 | 0 | 0 | 0 | 5.5 | 5.5 | 53 | 89.62% |
| 8 | 1 | 3.5 | 0 | 0 | 1.5 | 5 | 53 | 90.57% |
| | 2 | 0 | 0 | 0 | 3.5 | 3.5 | 53 | 93.40% |
| Average | - | 1.5 | 0 | 0 | 2.8 | 4.3 | 53.2 | 91.92% |

#### 3) IMPACT OF WI-FI ARRANGEMENT

The experiments conducted using two unseen Wi-Fi arrangements, single-point and linear placement, produced average segmentation accuracies of 94.25% and 94.12%, respectively, as shown in Fig. 9. The differences between the three Wi-Fi arrangements tested in this study were within 0.5%. We also performed t-tests hypothesizing that there was no difference between the datasets obtained from the new and original Wi-Fi arrangements. The results showed that the p-value for the difference between the original (four-corner) and linear placements was 0.871, and the difference between the original and single-point placements was 0.757. This indicates that the Wi-Fi arrangement had an insignificant influence on segmentation accuracy when using the proposed HAR system.

### C. LONG-TERM TESTING

The accuracy of segmenting the 30 *min* CSI data in the long-term test achieved an average ACC of 94.74%, demonstrating an outcome similar to that obtained in the 50 *s* experiments conducted earlier in this study. Detailed information is presented in Table 8.

In terms of computing efficiency, the processing time on a computer with a CPU configuration of Intel i7-4790 and

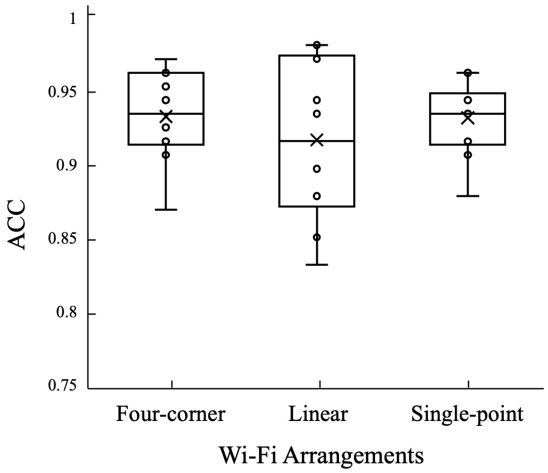


**FIGURE 9.** The activity segmentation accuracies of the proposed HAR system tested with different Wi-Fi arrangements.

**TABLE 8.** The activity segmentation accuracies of the proposed HAR system in the 30 *min* long-term test.

| Category | Time length | $E_i$ $(s)$ | Accuracy |
|---|---|---|---|
| W | 52.5 | 9 | 82.86% |
| BM | 137 | 0 | 100% |
| T | 46 | 1.5 | 96.74% |
| SP | 1608.5 | 81.5 | 94.93% |
| Total | 1844 | 92 | 94.74% |

a GPU configuration of NVIDIA Quadro RTX4000 was 155.43 *s* for converting CSI data of 1,844 *s* to CSI images and was 1205.64 *s* for recognizing the types of produced images and dividing them into segments of individual activities. In other words, the total time to recognize an activity per second was 0.65 *s*. Therefore, this system could continuously perform real-time activity recognition.

## VI. DISCUSSION

The proposed HAR system demonstrates an accuracy of 93.82% (Table 5) in recognizing predefined activities and the capability of segmenting the recognized activities from a long-term recorded CSI sequence. For health monitoring of PD, the detected target motions, such as walking and tremors, can be logged for their occurrence and quantified using existing methods [6], [29] to track motor functions.

The generalizability and capability of the proposed HAR system to process long-term CSI data were verified. When used in a new environment or Wi-Fi arrangement, the model could still segment activities with similar accuracies. The segmentation accuracy decreased slightly (1.9%) when the activities were performed by an unseen human subject. The results showed that the proposed HAR system can identify target activities regardless of individual differences. However, data from real patients with PD should be included to improve the model, as the difference between healthy persons and patients with PD can be significant. The process of receiving a CSI sequence to output the data segments of activities was computationally efficient. The processing time for each second of an input CSI sequence would only be 0.65 *s*, highlighting the potential of its use in real-time scenarios.

### A. LIMITATIONS

There are certainly limitations and challenges to be overcome in the future.

Multi-person activity recognition has always been a major challenge for HAR using Wi-Fi CSI. This study considers a situation in which the target person is the only person performing activities in the room. When more than one subject is active, interference inevitably increases and significantly affects the performance of the current system. Therefore, it is crucial to develop more dedicated signal-processing methods or deep neural networks to distinguish or separate individual signal responses from complex data.

The proposed HAR system was evaluated in two regular discussion rooms with good Wi-Fi signal coverage. In reality, obstacles, such as walls in the environment, can block or weaken the signal, leading to poor detection performance. This can be addressed in two ways. One is by enhancing the sensing capability of individual Wi-Fi links to enable through-wall detection [11], and the other is by utilizing and combining CSI from multiple Wi-Fi links that exist in the environment, such as Wi-Fi links formed by a wireless router and smart IOT devices.

In this study, the predefined activities were performed by young healthy human subjects. Therefore, the proposed HAR system might have learned to recognize the essential pattern of each activity but never saw the symptomatic features of PD. For example, abnormal walking, such as shuffling or freezing of gait and different kinds of tremors, although within the range of frequency and amplitude, might not be recognized owing to unseen features shown in the CSI image. A plan to collect data from healthy individuals of different ages and real patients with PD in different stages is required to improve the performance of the proposed HAR system.

### B. MAJOR REASONS CAUSING SEGMENTATION ERRORS

Although the overall accuracy of the proposed HAR system in segmenting the classified activities was satisfactory (93.82%), the poorest performance (85.85%) was observed in detecting walking activities. By further looking into the misjudged sessions in the CSI sequences, they were mainly transitions from a stationary state (static posture) to dynamic state (tremor, body movement, and walking) or in an opposite way. Figure 10 shows a sample of 50 s CSI image that contains four transitions between five activities, which are (1) walking to static posture, (2) static posture to tremor, (3) tremor to static posture, and (4) static posture to body movement. Transition 1 was found to begin approximately at the 8th second and last for about 4 *s*, in which the subject was stopping the momentum of walking and changing to a stationary state, leading to ambiguous recognition results.

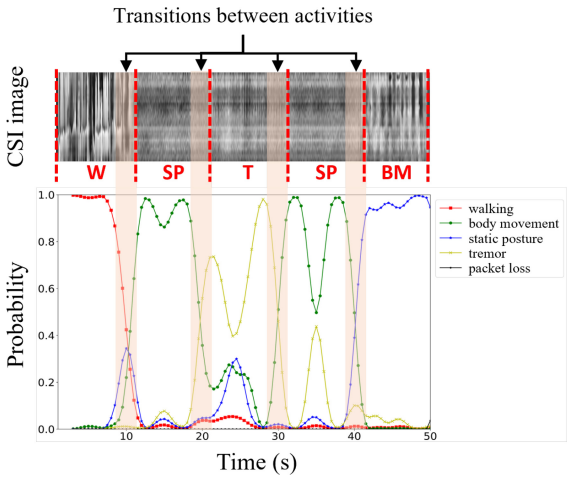


**FIGURE 10.** Recognition errors caused by transitions from one type of movement to another.

The first half of the transition was correctly recognized as walking, but the probability of walking dropped while that of static posture rose. In the second half of the transition, the probability of static posture exceeded that of walking as the motion of walking was largely reduced, leading to incorrect recognition results. Transitions 2–4 also contributed some portions of the recognition errors due to the change of state. However, the recognition error occurred more frequently in the transitions between walking and static postures, and these transitions required the longest time as the walking produced the largest motion compared to the other activities tested in this study.

To confirm this situation, we manually identified all transitions between walking and static postures for their start and end times and recalculated the ACC to be 96.03% by excluding these transition sessions, as listed in Table 9. The results suggested that the proposed HAR system can accurately segment activities during steady sessions. For health monitoring in Parkinson's disease, gait information during steady walking sessions in a day would be sufficient to provide significant insight for understanding a patient's motor fluctuations and drug efficacy. However, it is also important to observe the initial and final walking sessions. An advanced segmentation method for identifying these transition sessions is required to provide another valuable dataset for the disease.

### C. COMPARISON WITH STATE-OF-THE-ART APPROACHES

Three CSI-based HAR systems were selected from the state-of-the-art baselines for performance comparison: CSI Violent [8], WiRIM [30], and DA-HAR [9], as listed in Table 10. All three systems recognized human activities based on images generated from CSI sequences but applied different preprocessing techniques to image generation. Therefore, the three different image generation methods were compared with the proposed system, which generated the images using the CSI ratio model. The key observations are as follows:

1) CSI-based HAR methods have been developed for many applications; however, to the best of our knowledge, no method has been developed for the health monitoring of PD. Together with the three baselines, all four methods achieved over 90% accuracy in recognizing the predefined categories. However, this comparison is not intended to provide a ranking for each method, as the performance evaluation of an HAR method should depend on its usage.
2) Signal attenuation due to an increase in the distance between the subject and transmitter can significantly affect accuracy. The impact of distance was observed in the three baseline methods because their detectable areas were limited to a relatively small range. Because detection based on the CSI ratio has been shown to have an enhanced detection range [11], our method maintains an accuracy of 93.82% over a much larger space.
3) Compared with the three baselines, the proposed HAR system was comprehensively evaluated for its adaptability to changes in environmental settings, human subjects, and Wi-Fi arrangements.
4) Signal processing procedures, such as principal component analysis, filtering, and time-frequency transformations, may require more computing time. The proposed HAR system eliminated these time-consuming procedures while still obtaining high detection accuracy, making real-time HAR for health monitoring possible.

**TABLE 9.** The segmentation accuracies recalculated by excluding the transition sessions. SP: static posture, BM: body movement, T: tremor, W: walking.

| No. | SP | BM | T | W | $E_{total}$ | Duration | $ACC$ |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 0 | 0 | 0 | 2 | 55 | 96.36% |
| 2 | 1.5 | 0 | 0 | 0 | 1.5 | 55 | 97.27% |
| 3 | 0 | 0 | 0 | 4.5 | 4.5 | 53 | 91.51% |
| 4 | 1 | 0 | 0 | 0 | 1 | 55 | 98.18% |
| 5 | 1.5 | 0 | 0 | 0.5 | 2 | 55 | 96.36% |
| 6 | 0 | 0 | 0 | 0 | 0 | 55 | 100% |
| 7 | 1 | 0 | 0 | 0 | 1 | 55 | 98.18% |
| 8 | 1.5 | 0 | 0 | 0 | 1.5 | 53 | 97.17% |
| 9 | 2 | 0 | 1 | 0 | 3 | 53 | 94.34% |
| 10 | 2 | 0 | 3 | 0 | 5 | 53 | 90.57% |
| Average | 1.25 | 0 | 0.4 | 0.5 | 2.15 | 54.2 | 96.03% |

### D. PRACTICAL IMPLEMENTATION OF THE PROPOSED HAR SYSTEM

The proposed HAR system's practical implementation necessitates carefully considering its adaptability to new environments, diverse individuals, and various device configurations. Our generalization tests proved this adaptability, which showed the system's capacity to adjust to unfamiliar

**TABLE 10.** Comparisons between proposed HAR system and state-of-the-art CSI-based HAR methods.

| System | CSI tool | Antenna pairs | Frequency/ Bandwidth | Detectable area ($m^2$) | Process | Activity | Evaluation |
|---|---|---|---|---|---|---|---|
| CSI Violent [3] | Intel 5300 | Tx1, Rx3 | N.A. | 2×2 | wavelet, denoising, smoothing, filtering | walking, jumping, kicking, etc. 10 activities | 97.3% in the training environment. 92.7% in the validation environment. 5–10% reduction in an unseen environment. |
| WiRIM [19] | Atheros | Tx1, Rx2 | 5 GHz 40 MHz | 2×2 | PCA, STFT, AoA | walking, running, standing, etc. 5 activities | 97.8%–96.2% when the device and subject are within 0.2–2 m. Generalization tests were not mentioned. |
| DA-HAR [20] | Intel 5300 | Tx3, Rx1 | 5 GHz 20 MHz | 3.5×3.5 | PCA, Bandpass, STFT | walking, running, jumping, etc. 10 activities | 90% in the training environment. Generalization tests were not mentioned. |
| Proposed method | Intel 5300 | Tx2, Rx3 | 2.4 GHz 20 MHz | 8.2×4.2 | CSI ratio | walking, tremor, sitting, nodding, etc. 8 activities | 91.9%–95.2% in different use conditions. Tested for different environments, Wi-Fi arrangements, and persons. |

scenarios, including different environmental setups, individuals with diverse body sizes, and common Wi-Fi device arrangements in residential settings.

Highlighting the study's Limitations section, it's pivotal for future research to focus on collecting data from a wide array of healthy individuals of differing ages and real patients with PD in varying stages. Such comprehensive data collection is crucial to advancing the HAR system's accuracy. The system's initial deployment could be in relatively simple environments, like a room in a nursing home, to accrue more genuine and less noisy data, thereby refining the CNN model. As the system's deployment expands, advanced techniques such as few-shot learning can be incorporated to enhance the precision in recognizing individual-specific movements.

The performance test has proved that the proposed HAR system can efficiently process raw CSI into recognized motion segments within 0.65 seconds for each second of data. An initial practical application could be set up using a mini-PC (Rx) connected to a home Wi-Fi router (Tx). In this configuration, the Rx would perform HAR tasks in real-time, and the results would be sent via Tx to a remote server for comprehensive motion analysis, such as assessing gait parameters or monitoring the frequency of tremors, showcasing the system's potential for effective and immediate real-world application.

## VII. CONCLUSION

This study proposes a novel approach for implementing a CNN-based HAR system using CSI images converted from the phase of CSI ratios. The system achieved an accuracy of 93.82% in identifying PD-related activities in a CSI sequence and in dividing the data sequence into segments for motion analysis. Furthermore, the impact of the environmental layout, person, and Wi-Fi arrangement on the system performance was found to be insignificant (the variation in the detection accuracy was within 2%), demonstrating the capability of the model to adapt to unseen conditions. The system could process an input CSI sequence with high efficiency (0.65 s processing time per second of CSI data), highlighting its potential for applications in real-time health monitoring.

## REFERENCES

[1] R. L. Nussbaum and C. E. Ellis, "Alzheimer's disease and Parkinson's disease," *New England J. Med.*, vol. 348, no. 14, pp. 1356–1364, 2003.
[2] *Global Health Estimates 2020: Disease Burden By Cause, Age, Sex, by Country and by Region*, World Health Organization, Geneva, Switzerland, 2020, pp. 2000–2019.
[3] M. Stacy, "Medical treatment of Parkinson disease," *Neurol. Clin.*, vol. 27, no. 3, pp. 605–631, 2009.
[4] S. M. Iqbal, I. Mahgoub, E. Du, M. A. Leavitt, and W. Asghar, "Advances in healthcare wearable devices," *NPJ Flex. Electron.*, vol. 5, no. 1, p. 9, 2021.
[5] Y.-M. Fang and C.-C. Chang, "Users' psychological perception and perceived readability of wearable devices for elderly people," *Behav. Inf. Technol.*, vol. 35, no. 3, pp. 225–232, 2016.
[6] A. Borhani and M. Pätzold, "The impact of human walking on the time-frequency distribution of in-home radio channels," in *Proc. Asia-Pac. Microw. Conf. (APMC)*, 2018, pp. 920–922.
[7] Y. Liu et al., "Monitoring gait at home with radio waves in Parkinson's disease: A marker of severity, progression, and medication response," *Sci. Transl. Med.*, vol. 14, no. 663, 2022, Art. no. eadc9669.
[8] H. Liu, J. Chang, L. Zhang, and B. Huang, "CSI-based violent behavior detection method," in *Proc. 7th Int. Conf. Comput. Commun. (ICCC)*, 2021, pp. 1727–1732.

[9] J. Zhang et al., "Data augmentation and dense-LSTM for human activity recognition using WiFi signal," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4628–4641, Mar. 2021.
[10] Q. Bu, G. Yang, X. Ming, T. Zhang, J. Feng, and J. Zhang, "Deep transfer learning for gesture recognition with WiFi signals," *Pers. Ubiquitous Comput.*, vol. 26, pp. 543–554, Jun. 2022.
[11] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–26, 2019.
[12] K. Woyach, D. Puccinelli, and M. Haenggi, "Sensorless sensing in wireless networks: Implementation and measurements," in *Proc. 4th Int. Symp. Model. Optim. Mobile, Ad Hoc Wireless Netw.*, 2006, pp. 1–8.
[13] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proc. 13th Annu. ACM Int. Conf. Mobile Comput. Netw.*, 2007, pp. 222–229.
[14] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity WiFi," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 53–64.
[15] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 581–594, 2016.
[16] E. Soltanaghaei et al., "Robust and practical WiFi human sensing using on-device learning with a domain adaptive model," in *Proc. 7th ACM Int. Conf. Syst. Energy-Efficient Build., Cities, Transp.*, 2020, pp. 150–159.
[17] X. Ma, W. Xi, X. Zhao, Z. Chen, H. Zhang, and J. Zhao, "Wisual: Indoor crowd density estimation and distribution visualization using Wi-Fi," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10077–10092, Jun. 2022.
[18] D. Wu et al., "FingerDraw: Sub-wavelength level finger motion tracking with WiFi signals," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 4, no. 1, pp. 1–27, 2020.
[19] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-MUSIC: Accurate device-free indoor localization," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 196–207.
[20] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with wi-Fi!" in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 593–604.
[21] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using WiFi channel state information," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 98–104, Oct. 2017.
[22] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial WiFi devices," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1118–1131, May 2017.
[23] P. Fard Moshiri, M. Nabati, R. Shahbazian, and S. Ghorashi, "CSI-based human activity recognition using convolutional neural networks," in *Proc. 11th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, 2022, pp. 7–12.
[24] S.-Y. Chen and C.-L. Lin, "Subtle motion detection using Wi-Fi for hand rest tremor in Parkinson's disease," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2022, pp. 1774–1777.
[25] S. M. Pizer et al., "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.
[26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[27] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 1, pp. 53–53, 2011.
[28] L. L. Rubchinsky, A. S. Kuznetsov, V. L. Wheelock, and K. A. Sigvardt, "Tremor," *Scholarpedia*, vol. 2, no. 10, p. 1379, 2007.
[29] H.-H. Chen, C.-L. Lin, and C.-H. Chang, "WiFi-based detection of human subtle motion for health applications," *Bioengineering*, vol. 10, no. 2, p. 228, 2023.
[30] X. Shen, L. Guo, Z. Lu, X. Wen, and Z. He, "WiRIM: Resolution improving mechanism for human sensing with commodity Wi-Fi," *IEEE Access*, vol. 7, pp. 168357–168370, 2019.