Article

Wi-FiAG: Fine-Grained Abnormal Gait Recognition via CNN-BiGRU with Attention Mechanism from Wi-Fi CSI

Anming Dong [1,2], Jiahao Zhang [3], Wendong Xu [1,2], Jia Jia [4,*], Shanshan Yun [4] and Jiguo Yu [5,*]

[1] Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China; anmingdong@qlu.edu.cn (A.D.); xwd@qlu.edu.cn (W.X.)

[2] Shandong Provincial Key Laboratory of Industrial Network and Information System Security, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China

[3] School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China; 10431231249@stu.qlu.edu.cn

[4] Shandong Zhengyun Information Technology Co., Ltd., Jinan 250000, China; yunss@sdas.org

[5] School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

[*] Correspondence: jiajia@sdas.org (J.J.); jiguoyu17@uestc.edu.cn (J.Y.)

Abstract: Abnormal gait recognition, which aims to detect and identify deviations from normal walking patterns indicative of various health conditions or impairments, holds promising applications in healthcare and many other related fields. Currently, Wi-Fi-based abnormal gait recognition methods in the literature mainly distinguish the normal and abnormal gaits, which belongs to coarse-grained classification. In this work, we explore fine-grained gait rectification methods for distinguishing multiple classes of abnormal gaits. Specifically, we propose a deep learning-based framework for multi-class abnormal gait recognition, comprising three key modules: data collection, data preprocessing, and gait classification. For the gait classification module, we design a hybrid deep learning architecture that integrates convolutional neural networks (CNNs), bidirectional gated recurrent units (BiGRUs), and an attention mechanism to enhance performance. Compared to traditional CNNs, which rely solely on spatial features, or recurrent neural networks like long short-term memory (LSTM) and gated recurrent units (GRUs), which primarily capture temporal dependencies, the proposed CNN-BiGRU network integrates both spatial and temporal features concurrently. This dual-feature extraction capability positions the proposed CNN-BiGRU architecture as a promising approach for enhancing classification accuracy in scenarios involving multiple gaits with subtle differences in their characteristics. Moreover, the attention mechanism is employed to selectively focus on critical spatiotemporal features for fine-grained abnormal gait detection, enhancing the model's sensitivity to subtle anomalies. We construct an abnormal gait dataset comprising seven distinct gait classes to train and evaluate the proposed network. Experimental results demonstrate that the proposed method achieves an average recognition accuracy of 95%, surpassing classical baseline models by at least 2%.

Keywords: deep learning; gait recognition; CNN; BiGRU; time and frequency features

MSC: 68T07

1. Introduction

Gait, as a significant biological characteristic for humans, can serve as an early indicator of underlying health issues. Deviations from normal gait patterns may signal conditions stemming from brain function deterioration, neurological disorders, or musculoskeletal problems, such as Parkinson's disease [1]. Abnormal gait recognition involves the detailed analysis of biomechanical features and movement patterns during walking to identify irregularities [2]. Accurate identification of abnormal gait patterns not only facilitates prompt and targeted medical interventions but also significantly enhances treatment efficacy, potentially leading to better patient outcomes and quality of life [3].

Recently, human activities analysis based on Channel State Information (CSI) has gained much attention due to advantages of its noninvasive nature, ubiquity, better coverage, cost efficiency, and privacy protection [4–9]. The studies by [4,5] focused on identity recognition using gait features, specifically aiming to identify or verify an individual's identity based on their unique walking patterns. In our previous work [6], a deep learning architecture based on CNN and Bidirectional Long Short-Term Memory (BiLSTM) was proposed for recognizing complex continuous human activities. The results demonstrate that the proposed network achieves high accuracy in recognizing human movements involving rapid and drastic actions. In Ref. [7], a passenger-counting system based on Wi-Fi sensing is proposed and validated through practical deployment on buses, demonstrating its potential for real-time application scenarios. The study presented in [8] proposes a fine-grained finger gesture recognition system using commercial Wi-Fi. This system leverages the principal components of CSI and selects critical subcarriers for accurate gesture recognition. The extraction of principal components enables the system to adapt to individual diversity and gesture inconsistency.

These pioneering studies on Wi-Fi sensing highlight the effectiveness of wireless signals in recognizing human activities and identifying individuals. They have further inspired the work of [10], which focused on distinguishing between normal and abnormal gait patterns. Such a work belongs to a coarse-grained binary classification problem. However, in practical applications, it is often necessary to identify specific types of abnormal gaits, requiring fine-grained classification of various abnormal gait types [11]. To the best of our knowledge, research on multi-class fine-grained abnormal gait recognition remains scarce. This gap motivates our study, which aims to develop a fine-grained gait recognition system capable of distinguishing multiple types of abnormal gaits.

In this work, we propose a deep learning architecture fine-grained abnormal gait recognition from Wi-Fi CSI. The proposed framework comprises three key modules, i.e., data collection, data preprocessing, and deep learning-based classification. In the data collection phase, we construct a Wi-Fi sensing platform using two commercial Intel 5300 network interface cards (NICs), one for transmitting and the other for receiving, allowing CSI data collection and dataset building. In terms of data preprocessing, we apply wavelet filtering and linear calibration to reduce the noise and nonlinear distortion in the amplitude and phase of CSI, respectively. We construct a deep learning classification module based on CNN-BiGRU with attention mechanism for gait recognition from the processed CSI data. Here, CNN is used to extract spatial features of the motion, while BiGRU is employed to learn bidirectional temporal features of the motion's past and future. Compared to traditional recurrent network structures such as LSTM and GRU, BiGRU utilizes a bidirectional feature extraction structure (i.e., considering both past and future information) to process temporal information and capture the correlation and dependence of sequential data before and after. To verify the impact of different environments on the recognition performance of the proposed method, experiments were conducted under various conditions in different locations, demonstrating that this method can achieve high-

precision recognition of abnormal gaits with an average recognition accuracy exceeding 95%. Compared with baseline methods, the proposed approach achieves at least a 2% improvement in recognition accuracy.

The contributions of this paper are summarized as follows:

(1) We investigate a fine-grained abnormal gait recognition method using Wi-Fi CSI. Our goal is to identify seven distinct gait classes, including six abnormal and one normal gait. This work focuses on multi-class classification, an area that has not been extensively explored in the context of Wi-Fi sensing.

(2) We propose a novel deep learning architecture for fine-grained abnormal gait recognition, combining CNN, BiGRU, and an attention mechanism. This architecture captures both spatial and temporal features of CSI data through CNN and BiGRU, respectively, addressing the limitations of relying on a single feature extraction method. The attention mechanism is incorporated to enhance feature focus, further improving overall performance.

(3) Unlike traditional designs that only consider amplitude, this paper comprehensively takes into account both amplitude and phase information. Experiments demonstrate that phase information improves the recognition performance for gait.

2. System Model

Figure 1 illustrates the overall structure of the Wi-Fi perception system, which comprises three core modules: data acquisition, data preprocessing, and activity classification. Specifically, the data acquisition module captures raw CSI data using a network interface card. Subsequently, the data preprocessing module processes these raw data, including noise reduction and calibration of amplitude attenuation and phase shifts, and distinguishes between active and inactive regions of the CSI data based on amplitude variance. Finally, the activity classification module employs a classifier built using a neural network, which receives the denoised and calibrated CSI data from the active regions, automatically extracts spatial and temporal features of the CSI, and performs classification.
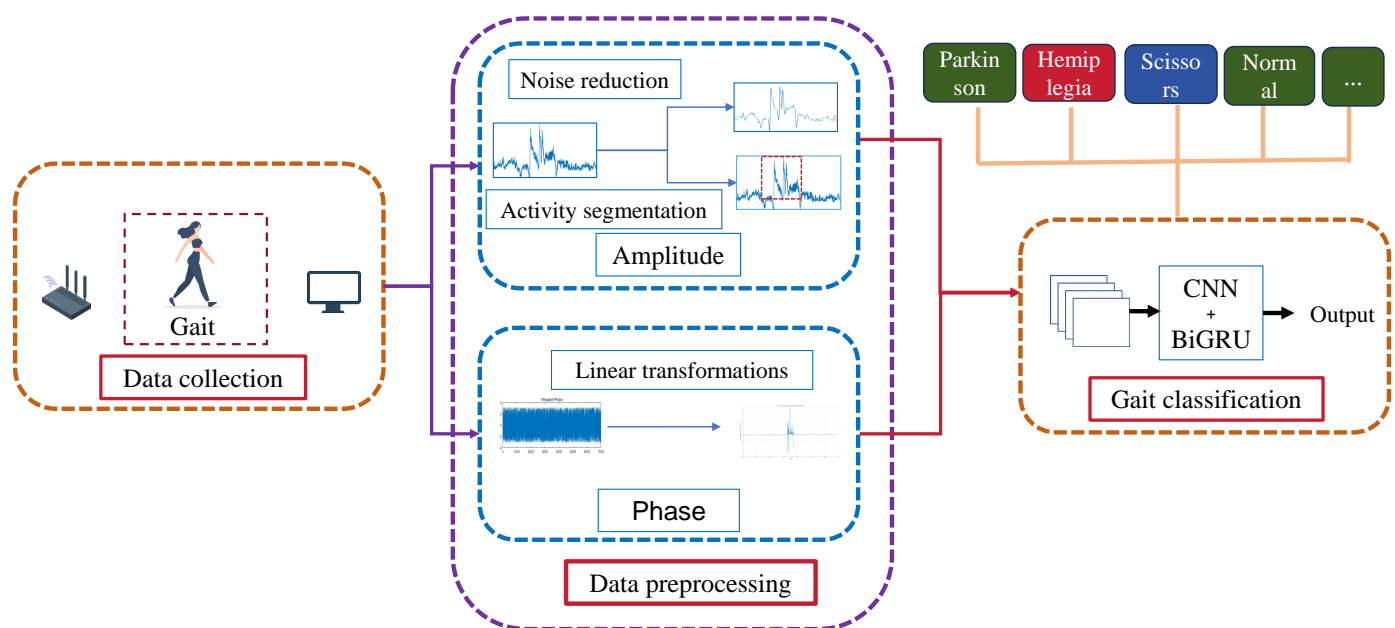

Figure 1. Overall architecture of the Wi-Fi sensing system.

2.1. Data Collection Module

We construct a data collection platform based on two personal computers (PCs), each equipped with an Intel 5300 NIC, as shown by Figure 2. It is noted that we did not using the commercial Wi-Fi router as the transmitter, since we found that it sometimes lost data packages. In order to make our data collecting high quality, we designed our specialized transceiver using the two NICs. At the transmit side, we used only one out of the three antennas of the NIC, and left the other two unused. At the receive side, we used all three antennas of the receiving NIC. In this way, a 1×3 single-input multiple-output (SIMO) Wi-Fi wireless transceiver was constructed, and we utilized the famous CSI Tool, which was proposed by [12], to parse and obtain three channels of CSI. Most existing commercial Wi-Fi technologies employ the IEEE 802.11 a/g/n wireless communication protocols, with their core using orthogonal frequency division multiplexing (OFDM) to modulate signals onto multiple subcarriers for parallel transmission.

Figure 2. Wi-Fi sensing platform.

The essence of OFDM is to convert a broadband channel into multiple parallel narrowband channels, where the channel on each subcarrier can be regarded as a flat-fading channel, thereby significantly reducing the complexity of the receiver equalizer [13]. The baseband signal after down-conversion in OFDM at the m-th receive antenna can be expressed as

$$\mathbf{Y}_m = \mathbf{H}_m \mathbf{X} + \mathbf{N}_m \tag{1}$$

where $m \in \{1, 2, 3\}$ denotes the index of the receive antennas, $\mathbf{H}_m \in \mathbf{C}^{K \times T}$ is the CSI matrix at the m-th antenna, $\mathbf{Y}_m \in \mathbf{C}^{K \times T}$ is the received baseband signal matrices, $\mathbf{N}_m$ represents the noise matrix during the transmission process, $\mathbf{X} \in \mathbf{C}^{K \times T}$ denotes the transmitted data at the transmit side. K and T denote the number of subcarriers and time slots, respectively.

The entries of the OFDM CSI matrix are dependent on the wireless signal propagation environment. Factors like path loss, reflection, scattering, and refraction affect the CSI of OFDM subcarriers. Moving objects in the physical space dynamically impact the time–frequency characteristics of CSI. To show this intuitively, examples of three-dimensional (3D) plots of the collected CSI amplitude are shown in Figure 3, with the packet index representing the time domain and the subcarrier index representing the frequency domain. The objective of Wi-Fi gate recognition is to analyze the time and frequency characteristics of the CSI data to identify types of human activity.
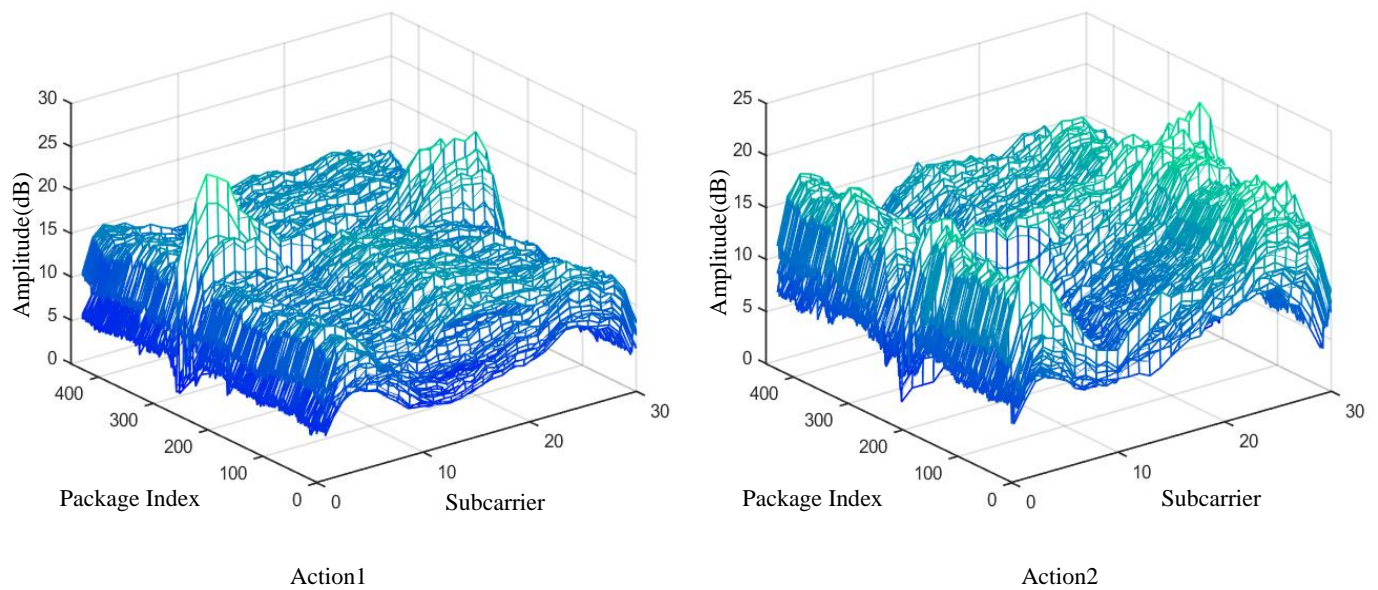
Action1   Action2

Figure 3. Examples of raw CSI data.

2.2. Data Preprocessing Module

2.2.1. Amplitude Processing

To handle outliers and noise in the CSI data, we first apply a Hampel filter with a sliding window to detect and replace outliers caused by environmental interference or equipment anomalies. Data points outside the $[\mu - \gamma\sigma, \mu + \gamma\sigma]$ range are identified as outliers and replaced with the median μ of the window, where γ is typically set to 3 [14]. Next, we use wavelet transform to reduce noise [15]. This involves decomposing the signal, applying a threshold to quantize coefficients, and reconstructing the signal to obtain a denoised version while preserving important features of the useful signal. Figure 4 shows the effect of outlier removal using the Hampel filter and noise reduction using wavelet transform (illustrated with an example of Scissors Gait data).
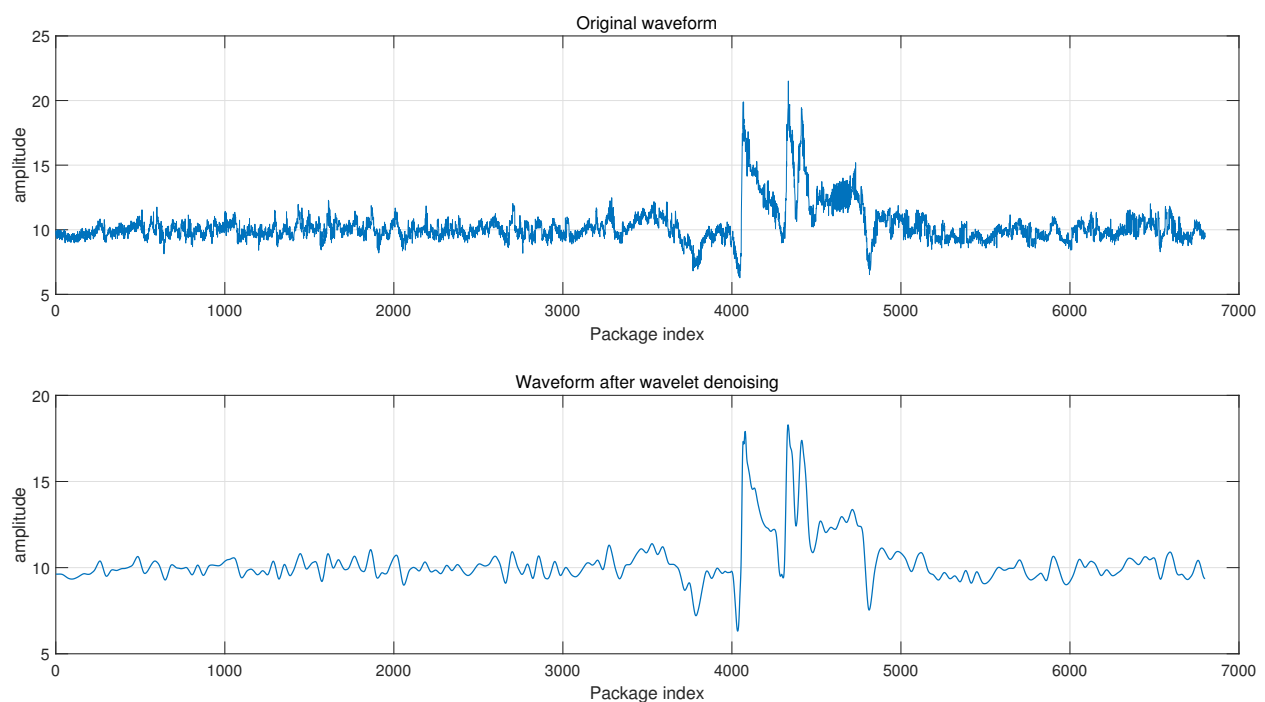

Figure 4. Comparison of amplitudes before and after data preprocessing.

2.2.2. Phase Calibration

The phase information extracted from the original CSI data contains carrier frequency offset (CFO) and sample frequency offset (SFO), which makes it unusable directly. Therefore, a linear transformation method is utilized to calibrate the phase information [16]. The original phase on the *i*-th subcarrier obtained after calibration is denoted as

$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N} \Delta t + \beta + z \tag{2}$$

In this context, $\hat{\phi}_i$ represents the original phase, ϕ_i represents the true phase, Δt is the time offset caused by the SFO (sample frequency offset), β is the unknown phase offset caused by CFO (carrier frequency offset), z is measurement noise, k denotes the index of the *i*-th subcarrier, and N is the length of the Fast Fourier Transform (in IEEE 802.11n, $N = 64$). Next, by subtracting the linear term $ak_i + b$ from the original phase, Δt and β can be eliminated, resulting in the calibrated phase. Here, the linear term is defined as

$$a = \frac{\hat{\phi}_n - \hat{\phi}_1}{k_n - k_1}, b = \frac{1}{n} \sum_{j=1}^{n} \hat{\phi}_j \tag{3}$$

After calibration using the linear transformation, the phase can be expressed as:

$$\tilde{\phi}_i = \hat{\phi}_i - ak_i - b = \phi_i - \frac{\phi_n - \phi_1}{k_n - k_1} k_i - \frac{1}{n} \sum_{j=1}^{n} \phi_j \tag{4}$$

The phase comparison before and after data preprocessing, as shown in Figure 5 (using Scissors Gait as an example), renders the phase a detectable signal.
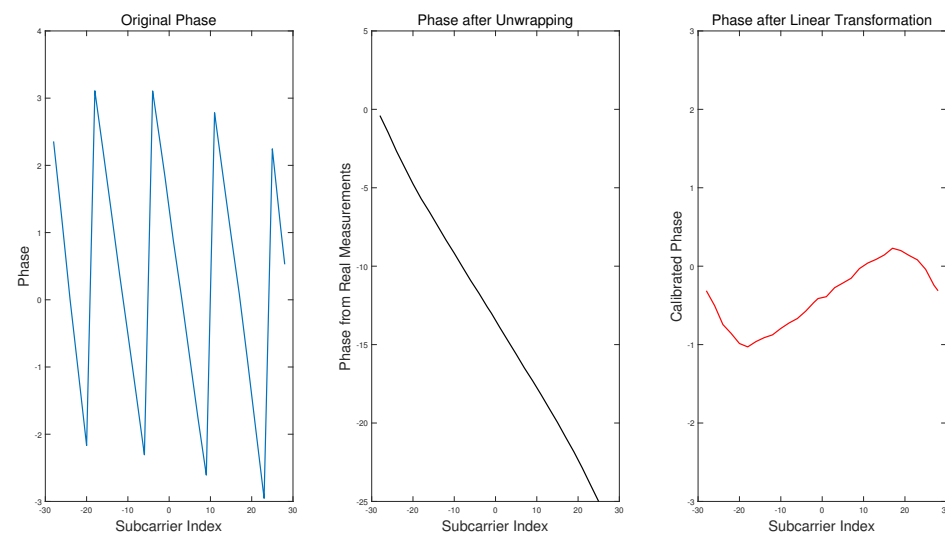

Figure 5. Comparison of phase before and after data preprocessing.

2.2.3. Activity Segmentation

Due to the presence of static information, namely the inactive parts, in CSI data, feeding this portion of the data into a neural network would increase the complexity of the algorithm. Therefore, it is essential to effectively distinguish between the active and inactive parts of the data, discard the inactive parts, and use the amplitude and phase information of the active parts as input to the neural network. Complete active data are also key to improving the classification accuracy of the neural network. For complex and vigorous continuous activities, the variance of the active part data is much greater than that

of the inactive part data. Hence, based on this phenomenon, an activity threshold ζ is preset. Additionally, due to the sensitivity of CSI, there may be brief fluctuations in the inactive parts that could be mistakenly classified as active parts. To obtain more complete and accurate active data, a window threshold η is introduced, aiming to eliminate the erroneous classifications caused by these fluctuations. The specific steps of the dual-threshold-based activity segmentation method proposed in this paper are as follows.

Step 1: Apply PCA to the matrix composed of amplitudes, automatically select the principal components that represent the most common variations in the CSI time series, and obtain the principal component matrix, which reflects the variations in subcarrier amplitudes.

Step 2: Perform activity segmentation using the first principal component. By applying a sliding window approach, calculate the variance of the data points within the window and return the data sequence composed of these variances. This results in the moving variance of the first principal component, which is used as an indicator for activity segmentation.

Step 3: Given an activity threshold ζ, activity is deemed to start when the variance of the first principal component exceeds the threshold ζ, and activity is deemed to end when the variance of the first principal component falls below the threshold ζ. Sample points with variances greater than ζ are marked as the active portions of the CSI data.

Step 4: By introducing the window threshold η, we once again label sample points with a window size (i.e., packet index) smaller than η as inactive data, thereby obtaining the final labeled data.

Figure 6 illustrates this effect, where the dashed-line boxes roughly outline the indexed segments marked as active, while solid circles represent brief fluctuations.
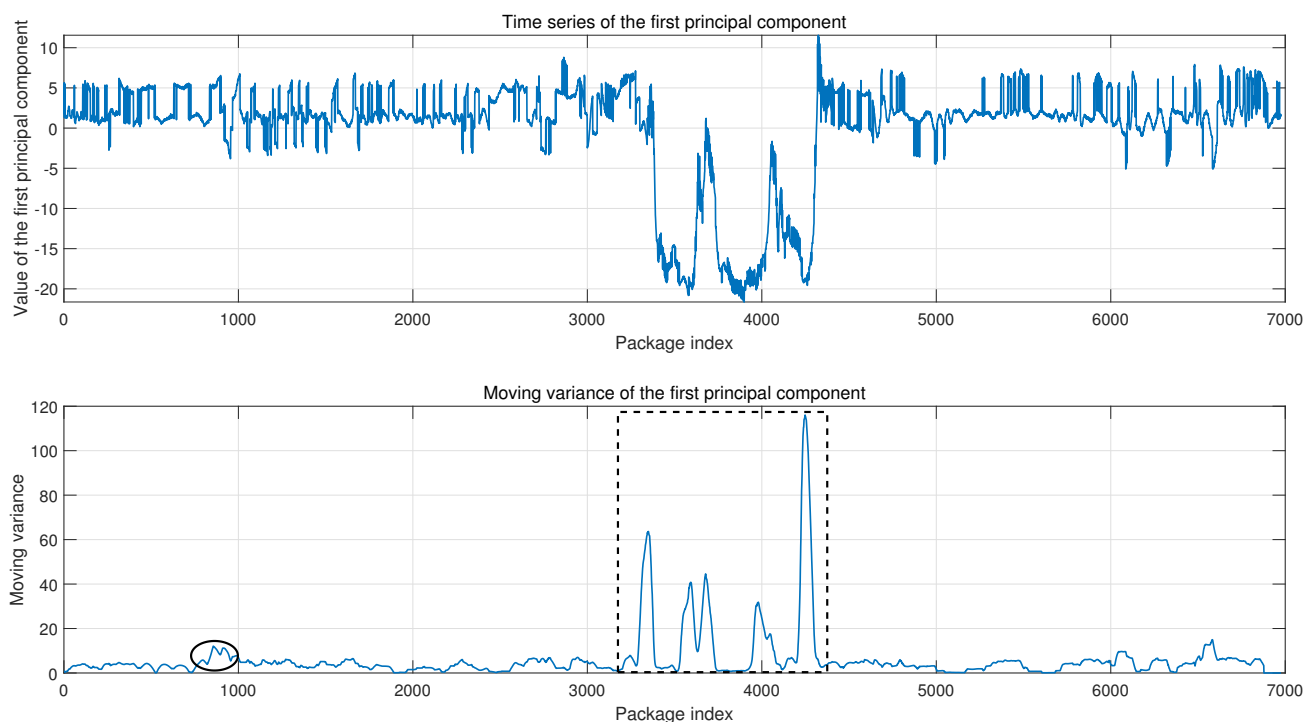

Figure 6. Activity segmentation.

2.3. Gait Recognition Module

The preprocessed CSI data are then sent to the gait classification module, where the CSI contains not only the spatial features of actions but also their temporal features. LSTM and GRU are capable of learning dependencies and correlations between long sequences of information, capturing historical information and significant events with large intervals or delays. However, LSTM and GRU networks, which possess temporal modeling capabilities,

do not consider the extraction of spatial features of actions. CNN, characterized by local connections and weight sharing, possesses powerful feature extraction capabilities but neglects the correlation between temporal information. Furthermore, due to their structural characteristic of transmitting temporal information in a single direction, LSTM and GRU can only consider past temporal information of actions, neglecting the learning of patterns from future information. BiGRU can extract temporal features from both past and future directions, but it assigns equal weights to the features of all CSI, whereas different features may contribute differently to the recognition of abnormal gait actions. Therefore, activity recognition systems in existing work have not fully exploited the spatiotemporal features of actions, leading to suboptimal recognition accuracy. To address this, this paper proposes a novel Wi-Fi-based abnormal gait perception framework that integrates CNN-BiGRU with an attention mechanism.

The model architecture is illustrated in Figure 7, where the input signal is a two-dimensional matrix obtained by stacking and expanding the amplitude and phase components of the CSI matrix $\mathbf{H}_m$, $m \in \{1, 2, 3\}$, which corresponds to the three antennas of the Wi-Fi network card.

$$\bar{\mathbf{H}} = [| \mathbf{H}_1 |, | \mathbf{H}_2 |, | \mathbf{H}_3 |, \angle \mathbf{H}_1, \angle \mathbf{H}_2, \angle \mathbf{H}_3]_{2MK \times T} \tag{5}$$

In this context, $\mathbf{H}_m$ and $\angle \mathbf{H}_m$ represent the amplitude matrix and phase matrix, respectively, obtained by extracting the amplitude and phase of each element in the matrix $\mathbf{H}_m$. Here, $M = 3$ indicates that there are data from a total of three antennas.

Neural networks require consistent input data dimensions. However, due to the varying durations of each action, the lengths of their data packets differ. Additionally, increasing the input data dimension will also increase the time complexity of the algorithm. Therefore, to ensure consistent input data dimensions and reduce the complexity of the algorithm, the designed network applies a sliding window at the input layer to segment the two-dimensional matrix along the time series direction. Segments with less than 60% of labeled active sample points are discarded to remove inactive data from the CSI, obtaining data segments of the same dimension. The retained segmented data segments serve as the final input to the network. The input data undergo feature extraction through two branches, with feature fusion serving as the basis for the final classification. The first branch is built on a one-dimensional CNN to extract features in the spatial dimension of gait movements. The second branch is built on GRU and BiGRU to extract features in the temporal dimension. The extracted features from both the spatial and temporal dimensions are integrated and used as the final basis for classification, with the softmax function employed to classify the actions.
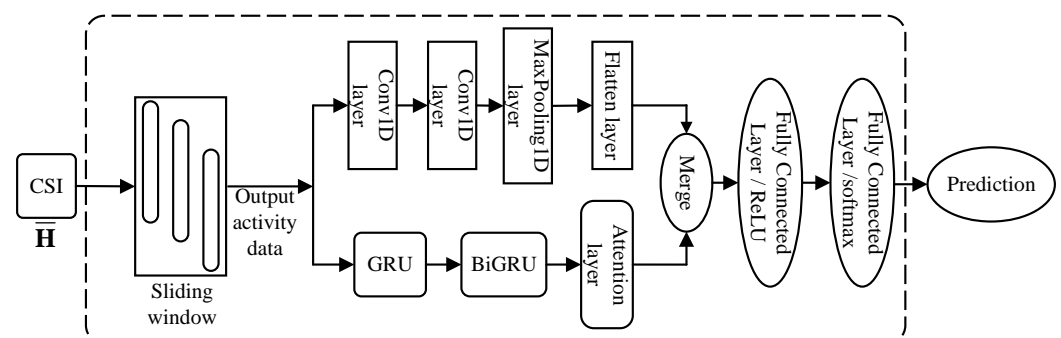

Figure 7. CNN-BiGRU architecture.

2.3.1. GRU and BiGRU

RNN (Recurrent Neural Network) has the ability of short-term memory and has significant advantages in dealing with short-term time series problems. However, when dealing with time series of high dimensionality, the issue of vanishing gradients may arise. The subsequent proposals of LSTM and GRU have improved this issue. GRU is an advanced variant of LSTM. Compared to LSTM, it simplifies the gating mechanism and does not introduce additional memory units. It controls the updating of information only through the update gate and reset gate. The GRU structure is shown in Figure 8a, which includes three parameters: the update gate z_t, the reset gate r_t, and the hidden state h_t. These parameters are updated through Equations (6)–(9).

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \tag{6}$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \tag{7}$$

$$\tilde{h}_t = \tanh(W[r_t h_{t-1}, x_t]) \tag{8}$$

$$h_t = (1 - z_t)h_{t-1} + z_t \tilde{h}_t \tag{9}$$

where W_z, W_r, and W are weight matrices, x_t represents the temporal information at time t, h_{t-1} denotes the hidden state at time $(t-1)$, and σ is the sigmoid activation function.

Since the collected gait movements are continuous actions, both past and future information are equally important for action recognition. BiGRU can extract temporal features from both past and future directions. Therefore, BiGRU is selected to learn the bidirectional patterns of motion features in order to extract more comprehensive features. The bidirectional GRU structure is shown in Figure 8b. BiGRU consists of forward and backward GRUs, and the final state h_t is jointly determined by the hidden states of both forward and backward GRUs. This state is then taken as the output of BiGRU.
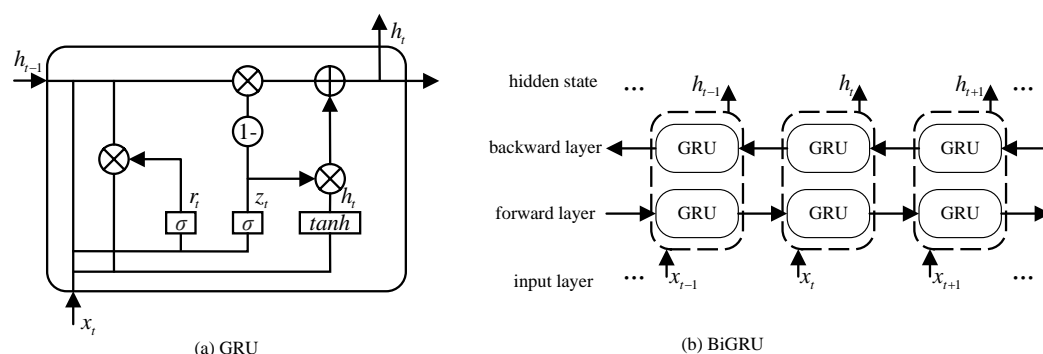

(a) GRU

(b) BiGRU

Figure 8. GRU and BiGRU architecture.

2.3.2. Attention Mechanism

The CNN-BiGRU model architecture proposed above can effectively classify actions with significantly different gait patterns. However, for actions with subtle differences in gait, such as Parkinsonian and myopathic gaits, which both exhibit a forward-leaning posture but differ in hand and foot movements, how can we focus on these fine-grained distinctions? To address this, our model architecture incorporates an attention mechanism.

The attention mechanism was initially designed for machine translation and has since been widely applied in the fields of image processing and natural language processing. This concept can be intuitively explained through the analogy of human visual perception: when a person visually perceives objects, they typically focus on specific regions of interest based on their needs. In this way, when similar scenarios reappear in the future, the individual will learn to direct their attention to those relevant areas [17]. BiGRU assigns

equal weights to all features of CSI, whereas different features may contribute differently to gait pattern recognition. For instance, both Parkinsonian and myopathic gaits exhibit forward-leaning postures, but their distinct hand and foot movements—one with knee and upper limb flexion, the other with uncoordinated limb movements—significantly impact CSI, as illustrated in Figure 9. These differences necessitate greater focus on variations in hand and foot movement-related CSI. Therefore, implementing an attention mechanism allows higher weights to be assigned to more critical features, enhancing the influence of key information, and thereby, improving the network's recognition performance.
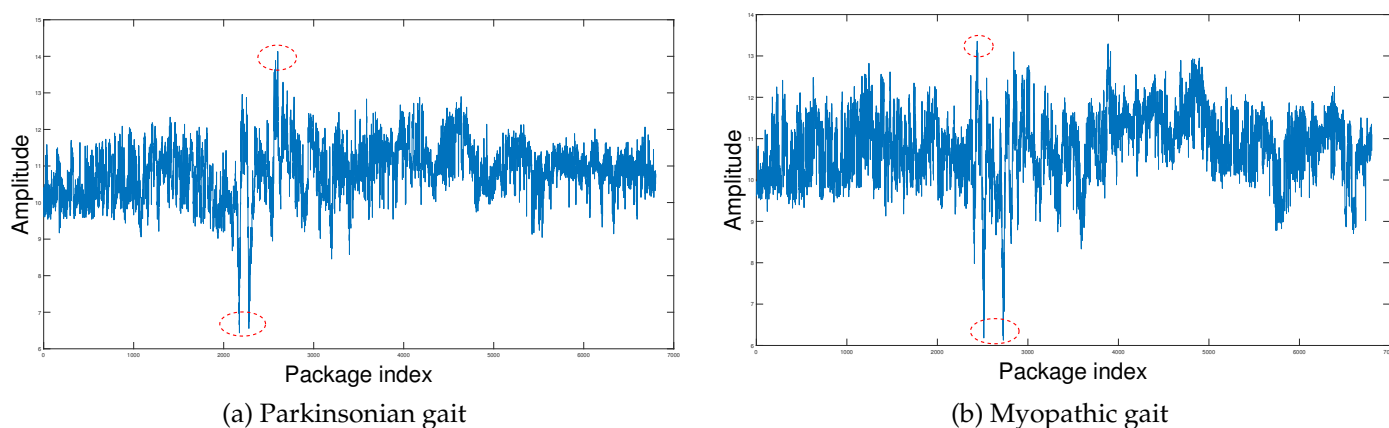

(a) Parkinsonian gait   (b) Myopathic gait

Figure 9. Parkinsonian and myopathic gait waveform comparison. The red dashed circles in the figure highlight the regions where the two gait amplitude waveforms exhibit significant differences.

The attention mechanism is illustrated in Figure 10. The input to the attention model is the sequence features learned from the BiGRU network, denoted as h_t, where $1 \leqslant t \leqslant n$. The importance score g_t for each feature vector is calculated using the tanh function, expressed as:

$$g_t = \tanh(\boldsymbol{W}^{\mathrm{T}} h_t + b) \tag{10}$$

where $\boldsymbol{W}^{\mathrm{T}}$ is the weight vector and b is the bias. Subsequently, the scores are normalized using $z_t = \mathrm{softmax}(g_t)$. Finally, the product of the feature vectors and the normalized scores is taken as the final output of the attention mechanism, expressed as:
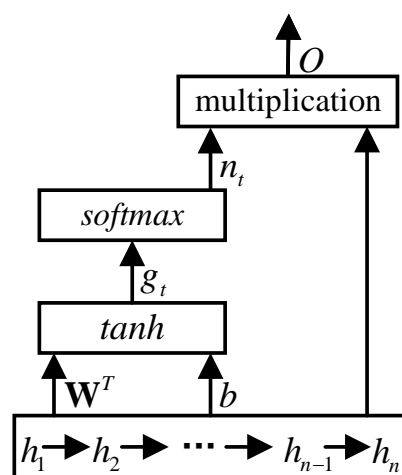
$$O = \sum_{t=1}^{n} z_t h_t \tag{11}$$

Figure 10. Attention mechanism.

3. Experiment and Result Analysis

To verify the effectiveness of the proposed method, multiple comparative experiments were conducted in this paper. The experimental scenarios and contents are as follows.

3.1. Data Acquisition

Data collection was conducted in both office and laboratory environments to evaluate the proposed method. The experimental scenarios are shown in Figure 11, where TX represents the transmitting end, RX represents the receiving end, and the test subject performs the corresponding actions between TX and RX. Seven volunteers (five males and two females) were recruited. The volunteers simulated various abnormal gait movements by watching patient videos. Each tester walked along a specified route within a given time limit (7 s), repeating each gait 30 times. The testers remained stationary at the beginning and end of each activity to reduce interference from nonwalking movements and ensure data accuracy.

The experiment designed six common abnormal gaits and one normal gait, with specific descriptions of each gait movement as follows:

- Parkinson Gait: The tester shows forward bending of the head and neck, flexed knees, and upper limbs, with extended fingers. While walking, the testers take small steps and exhibit an involuntary forward lean, leading to accelerated pacing.
- Fall Gait: The tester leans to one side due to unstable center of gravity while walking.
- Hemiplegia Gait: When walking, the tester drags one lower limb along the ground, with either the foot's heel or outer side touching the ground first.
- Scissors Gait: When walking, the tester's two thighs adduct, causing the knee joints to almost touch each other, forming a crossed-leg forward-moving posture.
- Ataxic Gait: During walking, the tester exhibits unstable side-to-side swaying, making it difficult to maintain a straight and stable walking path (similar to the gait observed after intoxication).
- Myopathic Gait: The tester exhibits uncoordinated movements of the hands and feet while walking, and demonstrates a forward-leaning posture of the body.
- Normal Gait: The tester walks along the route with normal arm swinging.

Although the dataset can be augmented by synthetic data generation methods, such as Generative Adversarial Network (GAN)-based augmentation [18], we did not adopt this approach in our current task. This decision was made due to concerns that GANs may introduce imperceptible phase shifts or amplitude distortions, which could disrupt the authentic physical relationships between signals and environmental interactions (e.g., errors in multipath reflection modeling).
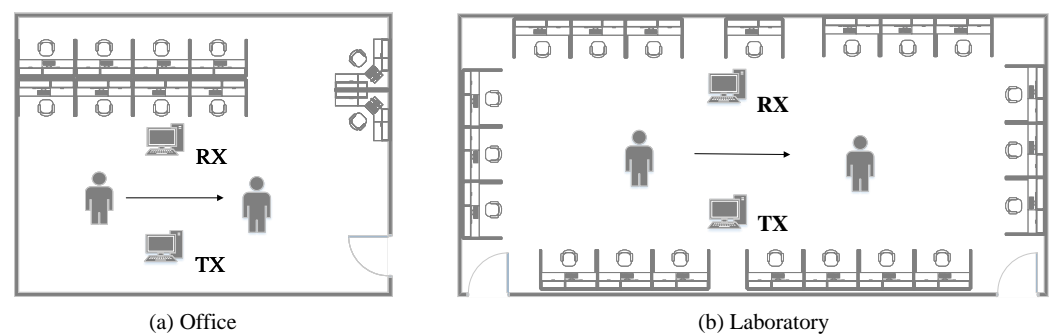

(a) Office   (b) Laboratory

Figure 11. Experimental scenes.

3.2. Model and Training Parameter Settings

This paper divides the dataset into a training set, a validation set, and a test set in a ratio of 7:2:1. The detailed parameter settings of the model are shown in Table 1, and the configuration data for the experimental environment are presented in Table 2.

Table 1. Detailed parameter settings for the model.

Parameters	Settings
Sliding window size	0.2 s
Convolution kernel size	3
Number of convolution kernels	128, 256
GRU hidden units	200
BiGRU hidden units	400
Attention mechanism hidden units	400
Fully connected layer parameters	128, 7

Table 2. Experimental environment configuration data.

Parameters	Settings
CPU	Intel Xeon Silver 4210R
Graphics Card	RTX2060-6G
Deep Learning Framework	TensorFlow
Optimizer	Adam
Loss Function	Cross-Entropy
Learning Rate	0.00003
Epochs	300
Batch Size	64

3.3. Experimental Evaluation

This paper employs various evaluation metrics to validate the reliability and effectiveness of the proposed method, including accuracy, precision, recall, and F_1 score [19].

Accuracy represents the proportion of correctly classified samples to the total number of samples, mathematically expressed as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{12}$$

Precision (also known as positive predictive value) indicates the proportion of samples predicted to be positive that are actually positive, mathematically expressed as:

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

Recall (also known as sensitivity or true positive rate) indicates the proportion of actual positive samples among the predicted positive samples relative to the total number of actual positive samples in the entire dataset, mathematically expressed as:

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

F_1 score is defined as the weighted average of precision and recall, calculated as:

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{15}$$

3.3.1. Environmental Experiments

Tests were performed in two settings: a relatively open office and a more complex laboratory, to check how well the method works. Figure 12 shows confusion matrices for both settings, with columns for predicted categories and rows for actual categories of gaits. These matrices show that the method can accurately classify seven types of gaits, six abnormal and one normal, with an average accuracy of 95.6% in the office and 95.1% in the lab. The results also show that the recognition accuracy for Hemiplegia and Ataxic Gait is relatively low. This is because the difference between these two movements lies only in the leg movements: in the former, one lower limb is dragged on the ground, while in the latter, the legs exhibit unstable swaying from side to side. The proposed CNN-BiGRU architecture is unable to achieve higher accuracy in classifying movements with very subtle differences. For example, in the laboratory environment, 7% of Ataxic Gait cases were misclassified as Hemiplegia Gait, and 4% of Hemiplegia Gait cases were misclassified as Ataxic Gait. Other misclassifications also mostly occurred among movements that involve similar motion components. Figure 13 shows the precision, recall, and F_1 scores for each activity in both the office and laboratory settings, indicating that all activities were classified within reasonable ranges.
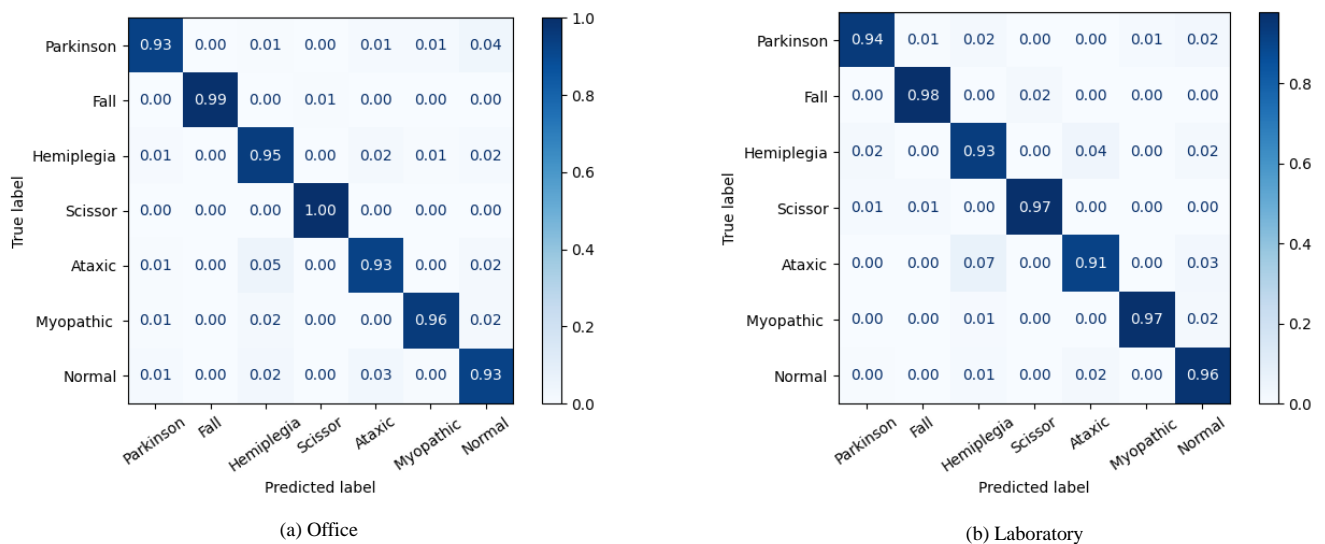

(a) Office

(b) Laboratory

Figure 12. Confusion matrices.
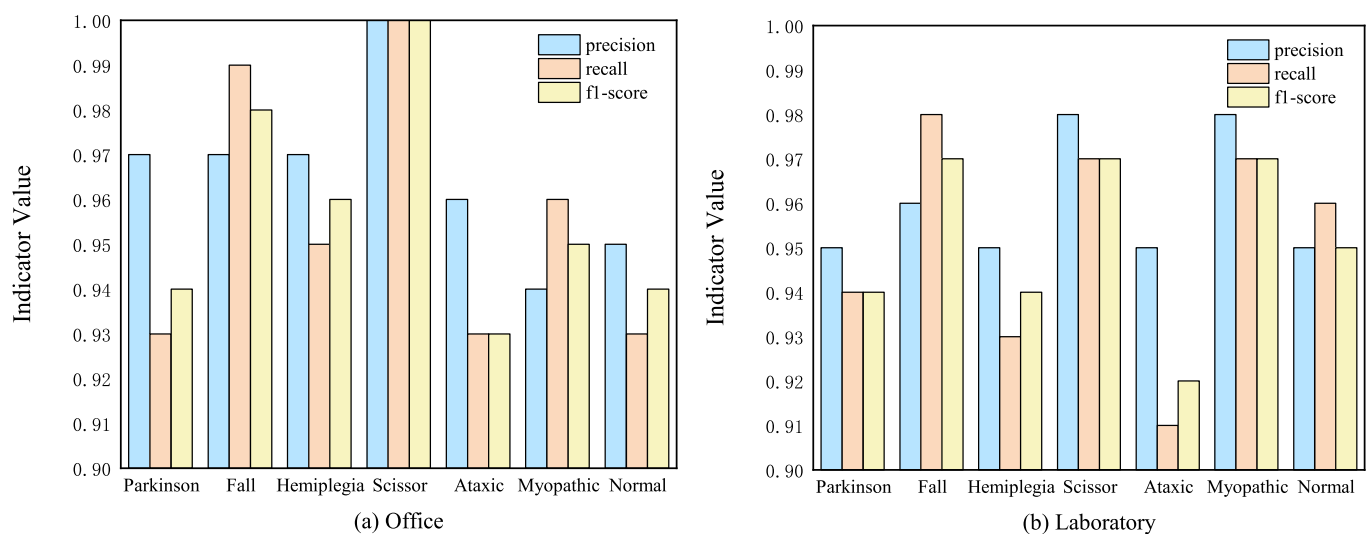

(a) Office

(b) Laboratory

Figure 13. Precision, recall, and F_1 scores.

3.3.2. Deep Learning Algorithm Experiments

This paper conducts comparative experiments using five deep-learning-based sensing algorithms, with the experimental results and time costs of different deep learning algorithms summarized in Table 3. As can be seen from the results, the accuracy of Vision Transformer (ViT) is lower than that of the proposed CNN-BiGRU architecture. This may be due to the limited generalization ability of ViT in small sample scenarios. The CNN-BiGRU-Attention architecture, on the other hand, performs more robustly in small sample scenarios due to its structural characteristics, such as the local receptive field of CNN and the recurrent memory of BiGRU. Compared to the proposed CNN-BiGRU structure, the GRU and ABLSTM models cannot achieve as high accuracy, but they do have certain advantages in terms of computation time. The proposed CNN-BiGRU architecture with an attention mechanism comprehensively considers the spatiotemporal features of gait actions, establishing a mapping relationship between CSI data and abnormal gait actions. Although this method incurs the highest training and testing time costs, it achieves recognition accuracy improvements of 2% to 20% over classical algorithms. Additionally, experimental tests show that the testing time for the proposed CNN-BiGRU structure is less than 5 s. Therefore, the proposed method not only achieves higher recognition accuracy but also maintains a certain level of real-time performance.

Table 3. Recognition accuracy and time costs of each algorithm.

Model	The Average Recognition Accuracy in Office Environment	The Average Recognition Accuracy in Laboratory Environment	Time Cost/s
MLP	78.4%	73.1%	130
CNN	90.1%	87.6%	182
LSTM	91.9%	90.3%	236
ViT	92.0%	91.2%	420
GRU	92.7%	91.5%	210
ABLSTM	93.2%	92.6%	390
CNN-BiGRU	95.6%	95.1%	519

3.3.3. Ablation Experiment

This paper evaluates the impact of each component within the proposed method framework on recognition performance, including the base signals, data preprocessing, and network architecture. The average recognition accuracy of different modules in an office environment is presented in Table 4 and the average recognition accuracy of different modules in a laboratory environment is presented in Table 5, where "✓" indicates that the experiment included that particular component.

Firstly, experiments were conducted to assess the selection of three types of base signals. Without any preprocessing, using only amplitude information resulted in higher accuracy compared to the combination of amplitude and phase. This is because unprocessed phase information suffers from severe random phase shifts, making it unsuitable for activity discrimination and thereby affecting overall recognition accuracy. The comparison of results between Experiment 4 and Experiment 6 in Tables 4 and 5 shows that the recognition accuracy is higher when using both preprocessed amplitude and phase as basic signals compared to using only preprocessed amplitude as basic signals. This also confirms the previously mentioned point that phase data can enhance gait recognition.

Secondly, when the combination of amplitude and phase was chosen as the base signal and subjected to noise reduction and calibration processing separately, the highest recognition accuracy was achieved.

Finally, this study examined the influence of individual modules within the attention-mechanism-based CNN-BiGRU model on recognition performance. Training models using CNN and BiGRU alone still yielded high recognition accuracy, which validates the effectiveness of the chosen network architecture. Furthermore, incorporating the attention mechanism into the model led to further performance improvements, achieving the optimal recognition accuracy.

Table 4. Average recognition accuracy of different modules in an office environment.

| Experiment | Whether or Not This Module Is Included | | | | | | Average Recognition Accuracy |
| | Base Signal | | Preprocessing | Network Architecture | | | |
	Amplitude	Phase		CNN	BiGRU	Attention	
1	✓			✓	✓	✓	93.2%
2		✓		✓	✓	✓	60.5%
3	✓	✓		✓	✓	✓	92.3%
4	✓		✓	✓	✓	✓	94.1%
5		✓	✓	✓	✓	✓	80.3%
6	✓	✓	✓	✓	✓	✓	95.6%
7	✓	✓	✓		✓	✓	93.7%
8	✓	✓	✓	✓			90.1%
9	✓	✓	✓	✓	✓		94.3%

Table 5. Average recognition accuracy of different modules in a laboratory environment.

| Experiment | Whether or Not This Module Is Included | | | | | | Average Recognition Accuracy |
| | Base Signal | | Preprocessing | Network Architecture | | | |
	Amplitude	Phase		CNN	BiGRU	Attention	
1	✓			✓	✓	✓	92.4%
2		✓		✓	✓	✓	59.3%
3	✓	✓		✓	✓	✓	91.3%
4	✓		✓	✓	✓	✓	93.2%
5		✓	✓	✓	✓	✓	79.4%
6	✓	✓	✓	✓	✓	✓	95.1%
7	✓	✓	✓		✓	✓	92.6%
8	✓	✓	✓	✓			87.6%
9	✓	✓	✓	✓	✓		93.7%

3.3.4. Real Scene Experiment

This paper validates the effectiveness of the proposed method in real-world scenarios. The experiments were conducted in a corridor, with the actual experimental setup and results shown in Figure 14. As illustrated in Figure 14a, test personnel simulated various abnormal gait movements between the transceiver. The results in Figure 14b present the confusion matrix of the proposed method in real-world conditions, demonstrating recognition accuracies above 85% for each action. The average recognition accuracy across seven actions reached 92.5%. Compared to office and laboratory settings, the recognition accuracy decreased slightly. This is primarily attributed to the complexity of the environment—a semi-enclosed space with significant surrounding noise (proximity to a road and human presence), which introduced greater signal interference and impacted accuracy. Overall, the proposed method achieved a recognition accuracy of 92.5% in real-world scenarios, validating its effectiveness in practical applications.
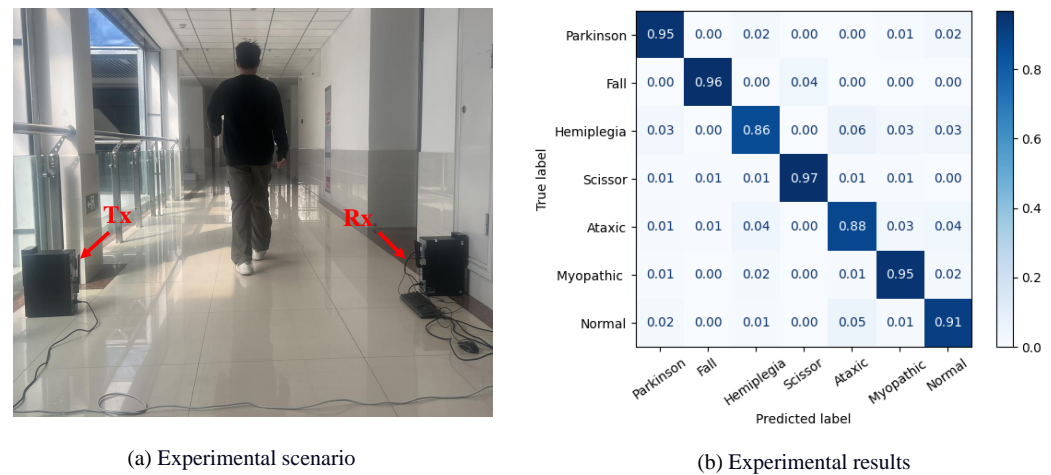
(a) Experimental scenario   (b) Experimental results

Figure 14. Real-world experimental scenarios and results.

4. Conclusions

This paper proposes a fine-grained abnormal gait recognition method that integrates an attention mechanism with CNN-BiGRU. The proposed approach effectively extracts rich spatiotemporal features from abnormal gait movements, enabling high-precision recognition of complex and continuous abnormal gaits. Experimental results demonstrate that the method achieves an average recognition accuracy exceeding 95% in ideal office and laboratory scenarios. However, its accuracy slightly decreases to 92.5% in real-world corridor environments, with most misclassifications occurring between gait patterns containing similar motion components.

Future work will focus on the following directions. The current study has only validated seven abnormal gait types in controlled environments, and its applicability to real clinical scenarios remains uncertain. Subsequent research should employ transfer learning techniques to address cross-domain generalization challenges. Additionally, model training and inference speeds will be optimized while maintaining recognition accuracy. Under the premise of preserving data fidelity, we will explore the use of GANs to enhance model generalization capability and improve classification accuracy. Furthermore, we aim to extend this research to multi-person abnormal gait recognition and advance the practical implementation of Wi-Fi sensing technologies.

Author Contributions: Conceptualization, J.Y. and J.J.; methodology, A.D. and J.Z.; validation, W.X. and S.Y.; formal analysis, A.D.; investigation, J.Z.; resources, J.J.; data curation, J.Z.; writing—original draft preparation, J.Z.; writing—review and editing, A.D.; visualization, S.Y.; supervision, J.Y.; project administration, J.J.; funding acquisition, J.J. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: Author Jia Jia and Shanshan Yun were employed by the company Shandong Zhengyun Information Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Gu, X.; Guo, Y.; Deligianni, F.; Lo, B.; Yang, G.Z. Cross-subject and cross-modal transfer for generalized abnormal gait pattern recognition. *IEEE Trans. Neural Networks Learn. Syst.* **2020**, *32*, 546–560. [CrossRef] [PubMed]
2. Paramanandam, V.; Lizarraga, K.J.; Soh, D.; Algarni, M.; Rohani, M.; Fasano, A. Unusual gait disorders: A phenomenological approach and classification. *Expert Rev. Neurother.* **2019**, *19*, 119–132. [CrossRef] [PubMed]
3. Thottempudi, P.; Acharya, B.; Moreira, F. High-performance real-time human activity recognition using machine learning. *Mathematics* **2024**, *12*, 3622. [CrossRef]
4. Zhang, L.; Wang, C.; Zhang, D. Wi-PIGR: Path independent gait recognition with commodity Wi-Fi. *IEEE Trans. Mob. Comput.* **2021**, *21*, 3414–3427. [CrossRef]
5. Zhang, L.; Wang, C.; Ma, M.; Zhang, D. WiDIGR: Direction-independent gait recognition system using commercial Wi-Fi devices. *IEEE Internet Things J.* **2019**, *7*, 1178–1191. [CrossRef]
6. Liu, Y.; Li, S.; Yu, J.; Dong, A.; Zhang, L.; Zhang, C.; Cao, Y. Wifi Sensing for Drastic Activity Recognition with CNN-BiLSTM Architecture. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Prague, Czech Republic, 9–12 October 2022; pp. 706–711.
7. Chang, W.; Huang, B.; Jia, B.; Li, W.; Xu, G. Online public transit ridership monitoring through passive WiFi sensing. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 7025–7034. [CrossRef]
8. Tan, S.; Yang, J.; Chen, Y. Enabling fine-grained finger gesture recognition on commodity WiFi devices. *IEEE Trans. Mob. Comput.* **2020**, *21*, 2789–2802. [CrossRef]
9. Wang, H.; Chen, Q.; Zhang, C.; Xu, J.; Su, C. P-CA: Privacy-Preserving Convolutional Autoencoder-Based Edge–Cloud Collaborative Computing for Human Behavior Recognition. *Mathematics* **2024**, *12*, 2587. [CrossRef]
10. Zhang, L.; Ma, Y.; Fan, X.; Fan, X.; Zhang, Y.; Chen, Z.; Chen, X.; Zhang, D. Wi-Diag: Robust Multi-subject Abnormal Gait Diagnosis with Commodity Wi-Fi. *IEEE Internet Things J.* **2023**, *11*, 4362–4376. [CrossRef]
11. Nwogo, R.O.; Kammermeier, S.; Singh, A. Abnormal neural oscillations during gait and dual-task in Parkinson's disease. *Front. Syst. Neurosci.* **2022**, *16*, 995375. [CrossRef] [PubMed]
12. Halperin, D.; Hu, W.; Sheth, A.; Wetherall, D. Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM Comput. Commun. Rev.* **2011**, *41*, 53. [CrossRef]
13. Zhang, J.; Lu, W.; Xing, C.; Zhao, N.; Al-Dhahir, N.; Karagiannidis, G.K.; Yang, X. Intelligent integrated sensing and communication: A survey. *Sci. China Inf. Sci.* **2025**, *68*, 1–42. [CrossRef]
14. Liu, W.; Chang, S.; Liu, Y.; Zhang, H. Wi-PSG: Detecting rhythmic movement disorder using COTS WiFi. *IEEE Internet Things J.* **2020**, *8*, 4681–4696. [CrossRef]
15. Sardy, S.; Tseng, P.; Bruce, A. Robust wavelet denoising. *IEEE Trans. Signal Process.* **2001**, *49*, 1146–1152. [CrossRef] [PubMed]
16. Qian, K.; Wu, C.; Yang, Z.; Liu, Y.; Zhou, Z. PADS: Passive detection of moving targets with dynamic speed using PHY layer information. In Proceedings of the 2014 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS), Hsinchu, Taiwan, 16–19 December 2014; pp. 1–8.
17. Denil, M.; Bazzani, L.; Larochelle, H.; de Freitas, N. Learning where to attend with deep architectures for image tracking. *Neural Comput.* **2012**, *24*, 2151–2184. [CrossRef] [PubMed]
18. Rajasekar, E.; Chandra, H.; Pears, N.; Vairavasundaram, S.; Kotecha, K. Lung image quality assessment and diagnosis using generative autoencoders in unsupervised ensemble learning. *Biomed. Signal Process. Control* **2025**, *102*, 107268. [CrossRef]
19. Akhtar, Z.U.A.; Wang, H. WiFi-based driver's activity recognition using multi-layer classification. *Neurocomputing* **2020**, *405*, 12–25. [CrossRef]