IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 25 April 2024, accepted 13 May 2024, date of publication 16 May 2024, date of current version 24 May 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3401685*

RESEARCH ARTICLE

# Research on Human Behavior Intention Perception Method Based on Wearable Sensors

**FANG YANG[1], QIANQIAN ZHENG[1], LIANQING CHEN[1], XIONG TAN[2], AND PENGCHENG CHE[1]**

[1]College of Nursing and Rehabilitation, North China University of Science and Technology, Tangshan 063210, China
[2]School of Mechanical Engineering, Hebei University of Technology, Tianjin 300401, China

Corresponding author: Pengcheng Che (17713126668@163.com)

This work was supported in part by the Natural Science Foundation of Hebei Province under Grant H2023209021 and in part by the Hebei Key Laboratory for Rehabilitation Engineering and Regenerative Medicine under Grant SZX202327.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of North China University of Science and Technology under Approval No. 2022256.

**ABSTRACT** With the increasing aging of the population, exoskeleton robots, gait assistance in the medical field, and condition monitoring in the rehabilitation field have been widely concerned by scholars. Accurate recognition of human gait and the perception of behavioral intention are essential for understanding the human motion state. Addressing issues of intricate sensor layout, low resolution of the plantar pressure sensor, and inaccurate gait classification, this paper presents a wearable sensor-based method for perceiving human behavior intentions. Firstly, a foot motion data acquisition system is built, and the foot motion feature data are obtained by combining the self-designed tactile pressure sensor and the Inertial Measurement Unit. Secondly, the rule of human movement is analyzed and the phase of gait in different motion modes is divided. Then, considering the temporal characteristics of human motion, the behavior pattern recognition algorithm is constructed by combining a Convolutional Neural Network and a Long Short-Term Memory network. Simultaneously, a Convolutional Neural network is established to recognize gait information and enhance the recognition accuracy of human behavior. Finally, the experiment of human motion intention recognition is carried out. The experimental results show that the proposed method has an average recognition accuracy of 96% for human motion patterns, and an average recognition accuracy of 93.38%, 94.57%, and 92.57% for the gait phase in various modes, respectively. The experimental results show that this method can meet the requirements of common human behavior recognition.

**INDEX TERMS** Intention perception, gait recognition, foot motion data, wearable sensors.

## I. INTRODUCTION

At present, society is gradually entering the aging stage, the proportion of the elderly population is increasing year by year, and the concern about the health monitoring of the elderly and the shortage of social labor is increasing. With the advancement of robotics technology, exoskeleton robots can not only provide assistance to the main labor force, but also help stroke patients to walk normally, improve movement ability, and help rehabilitation treatment. The exoskeleton robot is composed of human wearable devices and human-robot interaction systems, which can provide help for the human, enhance strength, withstand the load, and assist the human to move, overcome movement obstacles caused by disease or age, and complete various actions [1], [2], [3]. In the field of medical rehabilitation, the exoskeleton robot can assist patients with body movement incoordination and muscle atrophy, impacting their mobility. It can facilitate repetitive rehabilitation exercise training, replacing the manual rehabilitation training provided by rehabilitators [4], [5]. A complete exoskeleton robot system includes four parts: mechanical structure, sensing system, control system, and actuator [6]. When working, the sensing system captures the real-time motion state and intention of the human. It then

The associate editor coordinating the review of this manuscript and approving it for publication was Angelo Trotta.

sends control signals to the actuator through the control system to achieve the corresponding motion. For the assisted lower limb exoskeleton robot, the study of human gait characteristics and the accurate perception of human intention are crucial for achieving precise control and human-robot collaborative movement.

Researchers employ a range of sensors to collect human motion data, enabling real-time recognition of human motion gait and accurate perception of motion intention. This information is then utilized to determine walking state and motion intention, offering an effective control strategy for the system. The brain-computer interface bypasses the nervous system and muscles to process Electroencephalogram (EEG) signals, using external devices to directly acquire human EEG signals and detect movement intentions in patients with mobility disorders caused by conditions such as stroke [7], [8]. Despite its high recognition accuracy, electrodes require implantation into the brain, posing risks of surgical complications and infection. Moreover, implanted electrodes must address biocompatibility and high performance, leading to increased costs; Surface Electromyography (sEMG) uses a non-invasive sensor electrodes attached to the skin on the surface of the human to analyze the correlation between neuromuscular and gait phase or movement patterns [9], [10]. Leveraging the mechanism where EMG signals are generated before movement enables the sensing of human motion. Nevertheless, EMG signals are weak and complex. Furthermore, electrode connections are vulnerable to skin surface factors, and the presence of wires may hinder human movement. The Kinect camera captures the human gait by acquiring point cloud data on the human surface and a model of the human skeleton [11], [12]; The optical motion capture system can dynamically capture high-precision 3D position and motion posture information of the human to detect and identify gait [13]. Nonetheless, the method of collecting human motion data through visual detection has great limitations in its application scenario and may intrude upon individuals' privacy.

The aforementioned sensors may disrupt normal human activities, thereby limiting practical applications. Conversely, Wearable sensors [14], [15] like pressure sensors and Inertial Measurement Units (IMU) offer greater comfort and portability, seamlessly integrating with users' daily activities to capture more realistic and reliable data. IMU [16], [17] comprise acceleration sensors and gyroscopes, enabling the measurement of human movement acceleration and joint angle changes. Plantar pressure data are crucial biomechanical information in human motion, reflecting various motion states [18], [19], [20], [21]. Plantar data acquisition using pressure sensors is non-invasive and unaffected by the surface state of the human skin. Currently, there are two methods for collecting plantar pressure: one is to use a plantar pressure measurement platform [22], but this platform lacks mobility and is mainly used for laboratory and hospital tests. Another method is to use pressure measuring insoles for measurement: these insoles are placed inside the shoe, allowing the wearer to perform various movements without scene restrictions. Under different motion states, the force output needed for the collaborative movement of an exoskeleton robot and the human varies, and gait recognition can supply fundamental information for the perception and control system of an exoskeleton robot [23]. Lou [24] developed a flexible pressure sensor with a complex fabrication process based on the piezoresistive effect of multi-layer graphene films on polyester textiles. This sensor was utilized for real-time plantar pressure distribution and quantitative gait analysis. However, the sensor consisted of eight layers of materials, making the fabrication process complex. Zhang et al. [25] manufactured piezoresistive composite insoles by uniformly dispersing multi-walled carbon nanotubes into a polydimethylsiloxane matrix. The sensor consists of eight sensing elements with a laminated structure, the upper layer is a piezoresistive composite sensing layer, and the lower layer is a flexible printed circuit board as an electrode. However, the sensor's sensitivity and resolution are relatively low. Han et al. [26] proposed a sandwich-like capacitive pressure sensor with electrodes and dielectric layers made of carbon nanotube-polydimethylsiloxane and porous polymer, respectively. This design aimed to enhance the sensor's sensitivity, albeit with a complicated production process. Although the plantar pressure sensor utilized in the aforementioned method can recognize human motion states, the preparation process of the sensor is complex, overlooking the wearer's comfort. Furthermore, the resolution of collected plantar pressure data is insufficient, hindering a clearer depiction of the distribution and changes in plantar pressure data. Addressing issues of wearable comfort, simplicity in manufacturing process, and precision in pressure data, the flexible capacitive pressure sensor employed in this paper is easy to produce and capable of measuring pressure data from 512 points on the sole, thereby presenting abundant plantar pressure information. Moreover, when coupled with foot movement data obtained through IMU, it can attain accurate recognition of high-precision human movement patterns and gait phases, thereby effectively discerning human behavioral intentions.

The rest of this paper is as follows: Section II introduces the foot motion data acquisition system, which obtains high-resolution plantar pressure data and foot kinematics characteristics; section III analyzes the laws of human movement and categorizes the gait phase under different movement modes; in section IV, different models are set up to recognize human motion patterns and sense human motion intentions; in section V, gait recognition experiments are used to prove the effectiveness of the proposed method; Finally, it is summarized in section VI.

## II. FOOT MOTION DATA ACQUISITION SYSTEM

Gait data forms the foundation of gait recognition, and a system capable of efficiently and accurately collecting gait data is pivotal for successful gait recognition. As shown in Fig.1, this paper builds a foot motion data acquisition system,

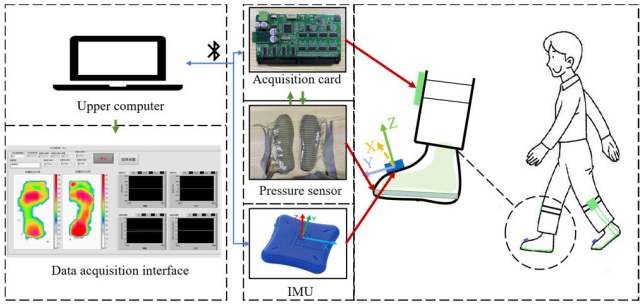


**FIGURE 1.** Foot motion data acquisition system.

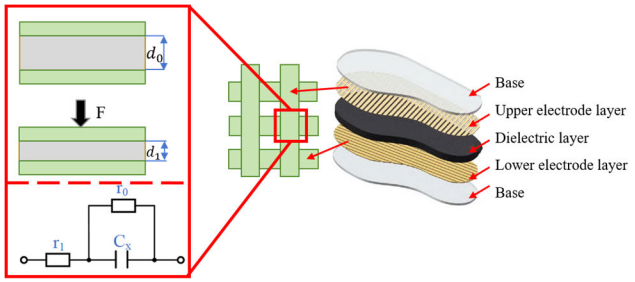


**FIGURE 2.** Working principle and structure of capacitive flexible pressure sensor.

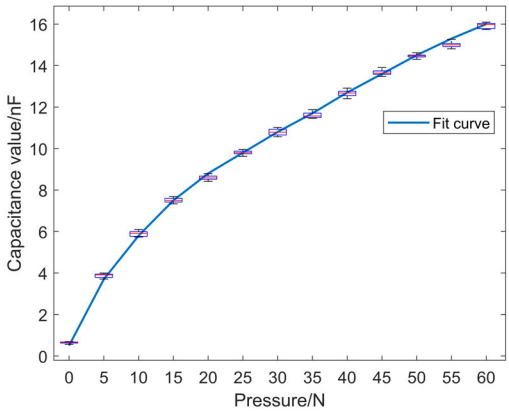


**FIGURE 3.** Capacitive pressure sensor performance.

including a plantar pressure sensor, data acquisition card, IMU, and upper computer data acquisition program. Among them, the plantar pressure sensor adopts flexible capacitance technology, and the upper and lower electrodes are distributed in array mode, which can form multiple capacitance sensing units in a limited area and be arranged on the sole to achieve high-resolution pressure data acquisition. The data acquisition card collects the capacitance signal of the flexible pressure sensor, converts the capacitance signal into a voltage signal through the signal conditioning circuit, and then transmits it to the PC through Bluetooth for subsequent data processing. The IMU is set above the toe bone to collect foot motion characteristics. It integrates 3 single-axis accelerometers, 3 single-axis gyroscopes, and 3 single-axis magnetometers, and communicates with the upper computer through Bluetooth. The acquisition system is easy to wear, highly portable, and capable of efficiently acquiring, storing, and processing foot movement data to fulfill the requirements of intention perception.

Capacitive sensors are highly favored by researchers for their high sensitivity, temperature independence, and suitability for large areas. They are extensively employed in detecting human physiological information and developing artificial skin [27], [28]. The structure of the capacitive pressure sensor designed in this paper is illustrated in Fig.2. Utilizing Thermoplastic polyurethane (TPU) film, the electrode material consists of nickel-plated conductive cloth with excellent electrical conductivity, whereas the intermediate dielectric layer is composed of a soft polyurethane sponge with good resilience. On the TPU substrate, the upper electrode layer consists of 32 rows of conductive cloth pasted parallel to each other, and the lower electrode layer consists of 16 rows of conductive cloth pasted in the same manner. One side of the conductive cloth from the two electrode layers is positioned opposite to the two sides of the dielectric layer. The upper and lower layers of conductive cloth are perpendicular to each other in space, creating a total of 512 cross-point array distributions, with each electrode cross-point serving as a capacitor unit. When the capacitor unit is subjected to stress, the capacitance value changes due to variations in the dielectric layer thickness. The capacitance value of each capacitor unit reflects the pressure changes in the corresponding region. Thus, distributed pressure detection can be achieved by collecting the capacitance value of each capacitor unit on the sensing element. Each capacitor unit resembles a parallel plate capacitor and can be expressed as:

$$C = \varepsilon s/d \tag{1}$$

where $s$ is the area of a single capacitor unit, $d$ is the distance between two electrodes, the dielectric constant $\varepsilon = \varepsilon_0 \cdot \varepsilon_r$,$\varepsilon_0 = 8.85 \times 10^{-12} F/m$ and $\varepsilon_r$ is the relative dielectric constant.

To evaluate the performance of the capacitive pressure sensor in this paper, the ZQ-990A pressure testing machine was employed. Ten capacitive units were randomly chosen from various positions on the pressure sensor, and a range of 0 to 60N pressure was continuously applied. Subsequently, the output capacitance value of each unit was determined. Fig.3 illustrates the test results of the pressure sensor, depicted in a box plot with a fitted curve. With the pressure incrementally rising from 0 to 60N, there was a corresponding gradual increase in the output capacitance, displaying a quasi-linear trend. Under identical pressure conditions, the maximum disparity in output capacitance among different capacitor units was 1.1nF, indicating favorable capacitance uniformity, thus satisfying the requirements for gait recognition and intention perception.

## III. ANALYSIS OF HUMAN MOVEMENT GAIT INFORMATION

Human walking is a periodic event, where two consecutive heel touches of the same leg make up a complete gait cycle.

The proper division of the gait period is essential for recognizing the walking state. Based on the characteristics of leg movement, the gait cycle can be divided into the supporting phase and the swinging phase. The support phase refers to the process from heel contact to toe off the ground, where the sole contacts the ground and bears the weight of the human, constituting about 60% of the entire gait cycle; The swinging phase refers to the process from the toe leaving the ground to the next heel touching the ground, constituting about 40% of the total gait cycle [29]. The division of the gait period should be adjusted based on the requirements of system recognition. Inadequate division of gait phases leads to insufficient captured motion information, making later-stage gait recognition challenging and impacting the overall system accuracy. Throughout the support phase, the pressure sensor can capture the contact details between the human foot and the ground. Conversely, during the swing phase, the pressure dissipates. Thus, integrating gait information from the swing phase, identified by the IMU sensor, enables the acquisition of motion characteristics for the complete gait cycle.

To accurately identify human gait information, three common motion modes are categorized based on the actual scenario: flat walking, up stairs and down stairs. The characteristics of human gait in different motion modes are analyzed, and the gait phases in each mode are defined, providing the foundation for human gait recognition. The gait cycle is calculated from the moment the right calcaneal bone makes contact with the ground until it touches the ground again. To enhance the accuracy of the acquisition system, phase division was conducted based on the plantar pressure distribution characteristics and IMU characteristics. The time interval between each landmark event and the subsequent landmark event was defined as a gait phase, named after the landmark event at the phase's initiation.

During walking, based on plantar pressure characteristics and IMU characteristics, gait phases were subdivided into Heel Contact (HC), Arch Contact (AC), Stance (ST), Heel Off (HO), and Forefoot Support (FS), Initial Swing (IS), Mid Swing (MS), and Terminal Swing (TS). Each phase is named after the foot state in the initial phase. The following is the definition of gait phases during walking. The division of gait phases in walking mode is depicted in Fig.4.

Heel Contact (HC): The right foot makes contact with the ground, and the right heel bone touches the ground. Simultaneously, the left heel bone starts to lift off the ground, entering the Heel Off phase.

Arch Contact (AC): The right foot continues to make contact with the ground, the ball of the foot starts to bear pressure, and the left arch lifts off the ground, entering the Forefoot Support phase.

Stance (ST): The right foot fully touches the ground, the sole is in contact with the ground, and the right foot alone supports the weight of the human, creating a single-support process during the gait cycle, during which the left foot completely lifts off the ground and enters the Swing phase.

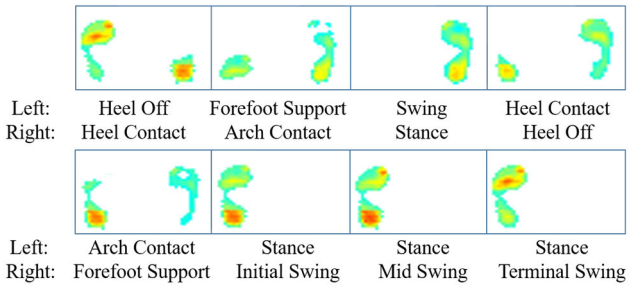


**FIGURE 4.** Gait division of flat walking mode.

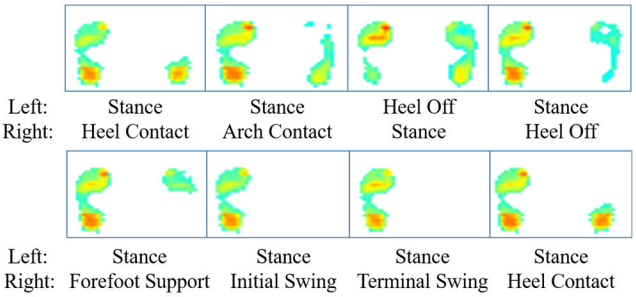


**FIGURE 5.** Gait division of flat walking up stairs mode.

Heel Off (HO): The right heel bone starts to lift off the ground, the weight of the human shifts forward, the front foot experiences increased pressure, and the left heel bone begins to make contact with the ground, entering the Heel Contact phase.

Forefoot Support (FS): The heel bone of the right foot is entirely lifted off the ground, and the forefoot exerts a force on the ground to propel the lower limb forward, while the forefoot of the left foot makes contact with the ground, entering the Arch Contact phase.

Initial Swing (IS): The right foot is entirely off the ground, the right thigh starts swinging forward, and the left foot is fully in contact with the ground, entering the Stance phase.

Mid Swing (MS): The right calf starts to swing with the knee as the center, while the left foot remains in the Stance phase.

Terminal Swing (TS): The right calf swings, causing the right foot to reach the lowest point of the swing arc until the heel bone of the right foot touches the ground, while the left foot remains in the Stance phase.

The foot support phase during walking up stairs is similar to that of flat walking. The support phase during the gait cycle can be divided into Heel Contact, Arch Contact, Stance, Heel Off, and Stance phase. The main movement of the foot in the process of walking up stairs is to step forward and upward, without the complicated swing of walking. Therefore, the swing phase is divided into the Initial Swing, where the pressure data disappears, and the Terminal Swing, where the foot begins to fall. Due to the limited size of the stairs and the short swing distance of the lower limbs during walking up stairs, the left foot remains in the Stance phase when the right foot contacts and lifts off the ground. The phase division of gait in walking up stairs mode is shown in Fig.5.

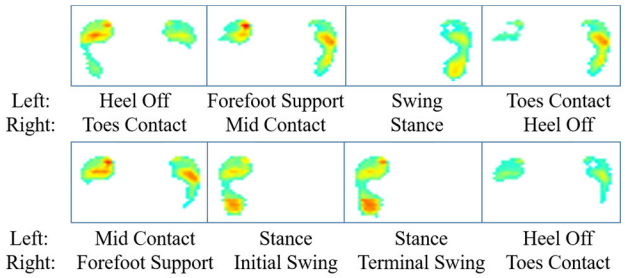


**FIGURE 6.** Gait division of flat walking down stairs mode.

When walking down stairs, the foot movement is that the toe bone area first contacts the surface of the stairs, so the gait phase is quite different from the walking and walking up stairs mode. Two gait phases are added to the gait phase: Toes Contact (TC) and Mid Contact (MC). The following is the definition of the gait phase during stair descent. The division of the gait phase in walking down stairs mode is shown in Fig.6.

Toes Contact (TC): The toe bone of the right foot begins to touch the surface of the next stair, while the heel bone of the left foot begins to move away from the surface and into the Heel Off phase.

Mid Contact (MC): The right foot continues to descend to the arch of the foot and makes contact with the surface of the staircase, while the left foot only makes contact with the surface of the staircase and enters the Forefoot Support phase.

Stance (ST): The right foot fully touches the surface of the stairs, supporting the weight of the human, while the left foot begins to Swing phase.

Heel Off (HO): The heel bone of the right foot begins to leave the surface of the staircase and the right knee begins to bend. At the same time, the toe bone of the left foot begins to contact with the surface of the next stair and enters the Toe Contact phase.

Forefoot Support (FS): The heel is completely removed from the surface of the staircase, the pressure on the ball of the foot is increased, and the left full arch begins to contact the surface of the staircase and enters the Arch Contact phase.

Initial Swing (IS): The foot completely leaves the surface of the stairs and begins to swing, while the human relies on the support of the left foot and enters the Stance phase.

Terminal Swing (TS): Foot swing begins to slow down until the toe bone area touches the surface of the staircase again, while the left foot remains Stance phase.

## IV. HUMAN MOTION INTENTION RECOGNITION METHOD

In this paper, a pattern recognition and intention perception method are designed using a deep learning network. Convolutional neural network (CNN)-Long Short-Term Memory (LSTM) model is used to identify human movement patterns, and the CNN model is used to identify the gait phase to further identify human behavior intention.

### A. CONVOLUTIONAL NEURAL NETWORKS

The convolutional neural network represents a highly parallel technique renowned for its capability to autonomously extract feature information from images [30], [31]. It operates on the principles of forward and backpropagation algorithms, rendering it the predominant model in visual-related domains encompassing tasks like image classification, object detection, and semantic segmentation. A complete CNN consists of an input layer, a convolutional layer, a downsampling layer, a fully connected layer, and an output layer. First, the image data to be learned are input in the input layer, and the input layer is connected with the convolution layer. The image features are extracted through the convolution operation to obtain the feature map.

The calculation process of the convolution operation can be expressed by the following:

$$x_j^k = f\Big(\sum_{i \in M_j} x_i^{k-1} \times \omega_{ij}^k + b_j^k\Big) \tag{2}$$

where $x_j^k$ is the $j$ neuron of the $k$ layer; $f(\cdot)$ is the activation function; $M_j$ is the set of input graphs; $w_{ij}^k$ is the convolution kernel parameter; $b_j^k$ is a biased value.

The convolutional layer operation completes the reduction and feature extraction of the input image. Considering that the dimension of the feature image is still high, the downsampling operation is needed to reduce the dimension of the feature to reduce the complexity of the model. The specific calculation process is as follows:

$$x_j^k = f(\beta_j^k down(x_j^{k-1}) + b_j^k) \tag{3}$$

where $x_j^k$ is the $j$ neuron of the $k$ layer; $\beta_j^k$ is the weighting coefficient; $down(\cdot)$ indicates the maximum downsampling operation; $b_j^k$ is a biased value.

After multiple convolution layers and downsampling layers, one or more fully connected layers are connected at last, all the features are connected, and the signals are summarized into one output signal, which is classified and processed. *Softmax* logistic regression is usually used for classification. This layer is called the output layer. The expression of the *Softmax* logistic regression function is as follows:

$$y(x_i) = \frac{\exp(x_i)}{\sum_i^M (\exp(x_i))} \tag{4}$$

where $x_i$ is the confidence degree of the Class $i$ target, and $y(x_i)$ is the probability of the class $i$ target.

The network training process is to keep looking for the minimum parameters of the loss function, the loss function of the network is:

$$E = \frac{1}{2} \sum_j^l (\hat{y}_{j\max}^k - y_{j\max}^k)^2 \tag{5}$$

where $\hat{y}_{jmax}^n$ is the identification result of batch $j$, $y_{jmax}^n$ is the expected identification result of batch $j$, and $l$ is the number of batches.

The backpropagation algorithm is adopted for optimization, and the optimal parameters are found based on the

gradient descent method to minimize the cost function:

$$\omega_{hi}^{k} = \omega_{hi}^{k-1} - \eta \frac{\partial E}{\partial \omega_{hi}^{k-1}} \tag{6}$$

$$b_{i}^{k} = b_{i}^{k-1} - \eta \frac{\partial E}{\partial b_{i}^{k-1}} \tag{7}$$

where $\eta$ is the learning rate, the most critical part above is to solve the partial derivative of the loss function, the back-propagation algorithm is a very effective calculation method, the whole calculation process is: (1) First set the weight $w$ and the initial value of the bias $b$; (2) Train the sample data $(x, y)$, input the sample data into the multi-layer perceptron, and all the weights $w$ and bias $b$ can be obtained. This process is forward propagation; (3) Calculate partial derivatives of ownership weight $w$ and bias $b$ under the loss function $\frac{\partial E}{\partial w}$, $\frac{\partial E}{\partial b}$ according to the chain rule; (4) The gradient descent method was used to update the weight $w$ and offset $b$.

### B. LONG SHORT-TERM MEMORY NETWORK

Neural networks excel in processing spatial data but struggle with time series data; Recurrent Neural Network (RNN) emerges as a prominent solution for such challenges. Nevertheless, RNN encounters issues like gradient explosion and vanishing gradients, both of which LSTM [32], [33] effectively mitigates. LSTM employs various gate units to selectively update relevant information, thereby preventing the accumulation of irrelevant data in long-term memory. LSTM consists of a cell state and a ''gate'' mechanism (input gate $f_t$, forgetting gate $i_t$, output gate $o_t$), where $c_t$ represents cell state, representing long-term memory. Adding or deleting state information at time $t$ through the gate structure allows the control to pass the modified state information to the next time step. The activation of each LSTM unit is as follows:

(1) Initially, determine if the unit state should be retained in the subsequent calculation process via the forgetting gate:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{8}$$

where $\sigma$ is the *Sigmoid* function; $W_f$ is the weight; $b_f$ is offset.

(2) The input gate determines which signals to update based on $h_{t-1}$ and $x_t$:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{9}$$

(3) Next, acquire the candidate cell state $\tilde{C}_t$ through the *tanh* layer, update the cell state based on the forgetting gate and the input gate, resulting in the new cell state $C_t$:

$$\tilde{C}_t = tanh(W_C[h_{t-1}, x_t] + b_C) \tag{10}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{11}$$

(4) Once the unit status update is complete, it's essential to identify the characteristics of the output unit status. Obtain the judgment conditions through the *Sigmoid* layer of the output gate and calculate with the *tanh* layer to determine the actual output signal:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{12}$$

**TABLE 1.** CNN-LSTM network parameters.

| Layer | Type | Convolution Kernel Size and Number | Number | Step Size | Feature Map | Neuron |
|---|---|---|---|---|---|---|
| Input | Input layer | - | - | - | 34×32 | - |
| C1 | Convolution layer | 4×4 | 6 | 1 | 30×18 | 6 |
| S2 | Pooled horizon | 2×2 | 1 | 2 | 15×9 | 6 |
| C3 | Convolution layer | 5×5 | 16 | 1 | 24×14 | 16 |
| LSTM | LSTM | - | - | - | 128 | - |
| F6 | Fully connected layer | 1×1 | - | - | - | - |
| Output | Output layer | 1×1 | - | - | - | - |

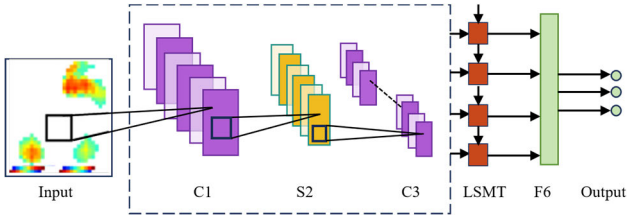


**FIGURE 7.** CNN-LSTM network structure.

$$h_t = o_t \times tanh(C_t) \tag{13}$$

### C. ESTABLISHING A CNN-LSTM NETWORK FOR MOTION PATTERN RECOGNITION

Human motion is a cohesive sequence of movements, exhibiting more pronounced motion characteristics compared to isolated frame data. While CNN excels in image recognition and classification tasks, it faces challenges in recognizing and processing time series data correlations. Conversely, LSTM networks can retain past recognition outcomes and merge them with present data, thereby providing benefits in recognizing time series data. Therefore, to better capture temporal changes, this paper proposes a method that combines CNN and LSTM to construct a human behavior recognition model. It extracts image features from foot movement data using a CNN and then feeds them into the LSTM to recognize human movement patterns. The network performance is significantly influenced by the number of layers and the combination mode. This paper establishes the network model structure through multiple tests. The parameters of each layer are presented in Table 1, and the specific network structure is illustrated in Fig.7. Additionally, a dropout layer is inserted between the LSTM layer, randomly deactivating 30% of neurons to prevent overfitting during model training.

As seen above, the phase combination of bipedal gait varies significantly under different movement modes.

**TABLE 2. Convolution neural network parameters.**

| Layer | Type | Convolution Kernel Size | Number | Step Size | Feature Map | Neuron |
|---|---|---|---|---|---|---|
| Input | Input layer | - | - | - | 34×16 | 1 |
| C1 | Convolution layer | 5×5 | 6 | 1 | 30×12 | 6 |
| S2 | Downsampling layer | 2×2 | 1 | 2 | 15×6 | 6 |
| C3 | Convolution layer | 5×5 | 16 | 1 | 11×2 | 16 |
| S4 | Downsampling layer | 2×2 | 1 | 1 | 6×2 | 16 |
| C5 | Convolution layer | 1×1 | 120 | 1 | 6×2 | 120 |
| F6 | Fully connected layer | 1×1 | - | 1 | - | - |
| Output | Output layer | 1×1 | - | 1 | - | - |

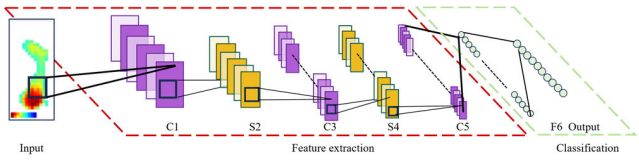


**FIGURE 8. Convolutional neural network structure.**

Consequently, the input layer data for the movement pattern recognition network consists of images representing the movement of the left and right feet. Consequently, the input layer data for the movement pattern recognition network consists of images representing the movement of the left and right feet. Max pooling is chosen for the pooling layer, and the *ReLU* function is selected as the activation function. The expression is as follows:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \tag{14}$$

The extracted features are input into the LSTM network. The activation function of the LSTM layer is the *tanh* function. The expression is as follows:

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{15}$$

The fully connected layer and *Softmax* function serve as network classifiers. The output layer has 3 nodes, with output values 1 to 3 representing different motion modes: 1 for flat walking mode, 2 for walking up stairs mode, and 3 for walking down stairs mode.

### D. ESTABLISHING THE GAIT RECOGNITION CNN NETWORK

The CNN-LSTM network is utilized to identify the human movement pattern, followed by the design of a gait phase intention perception model based on the gait information in the motion pattern. The CNN is employed to identify the human movement gait phase, enabling more accurate identification of human gait information. The parameters of each layer in the CNN are presented in Table 2, and the specific structure of the CNN is depicted in Fig.8.

The CNN structure comprises eight layers: Input layer, convolutional layer C1, downsampling layer S2, convolutional layer C3, downsampling layer S4, convolutional layer C5, fully connected layer F6, and Output layer output. The input layer encompasses the entire motion data image. Effective features of the image are extracted through the convolution layer, with the pooling layer utilizing maximum pooling. The activation function adopts the *ReLU* function, and the fully connected layer and output layer act as network classifiers, mapping the features extracted through convolution downsampling to the marker space. Subsequently, the probability value of each category is calculated for classification. The output layer is assigned numbers from 1 to 10, representing the states of the gait phase. Specifically, 1 denotes Heel Contact, 2 is Arch Contact, 3 is Stance, 4 is Heel Off, 5 is Forefoot Support, 6 is Toes Contact, 7 is Mid Contact, 8 is Initial Swing, 9 is Mid Swing, and 10 is l Terminal Swing.

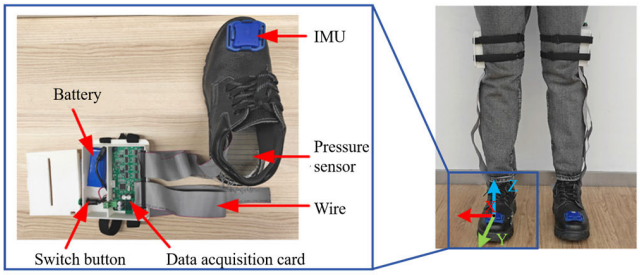


**FIGURE 9. Foot motion data collection system device wearing schematic diagram.**

## V. GAIT RECOGNITION EXPERIMENT

In this paper, three movement modes of flat walking, up stairs, and down stairs, and their gait phases are identified. The movement data for walking, up stairs, and down stairs are recorded using a foot motion data acquisition system. To ensure the reliability of the sample data, 20 volunteers, comprising 10 men and 10 women, were enlisted to partake in the experiment. Subsequently, movement data for flat walking, up stairs, and down stairs were collected. The stairs utilized in the experiment measured 20cm in height and 30cm in width each. Fig.9 illustrates the device setup for the foot motion data acquisition system, comprising a sole pressure measurement insole, a data acquisition card, and an IMU. Plantar pressure sensors are arranged in the shoe. To prevent IMU installation errors caused by vamp deformation during walking, rigid shoes are selected, and the IMU is installed in the toe position of the vamp. The X direction points to the right side of the human, the Y direction points forward, and the Z direction is determined by the right-hand coordinate

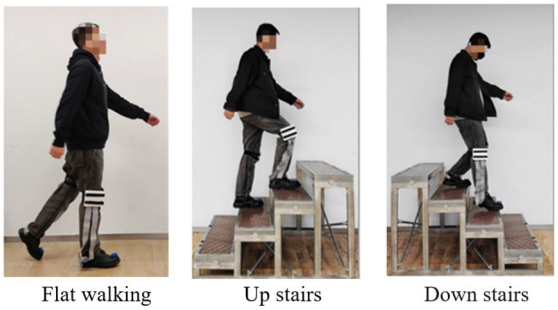


**FIGURE 10.** Actual motion pattern test diagram.

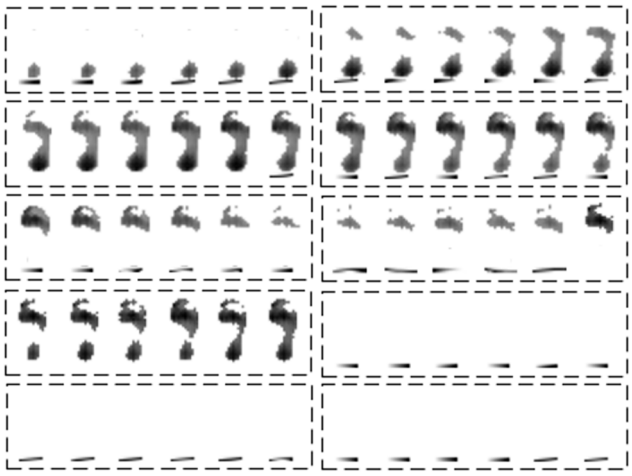

**FIGURE 11.** Foot motion data image.

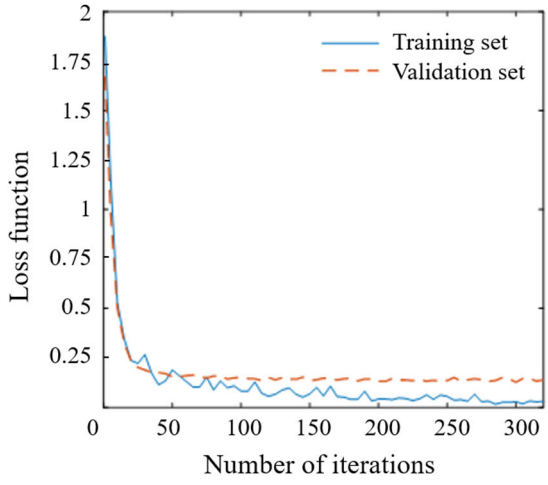


**FIGURE 12.** CNN-LSTM network training loss function.

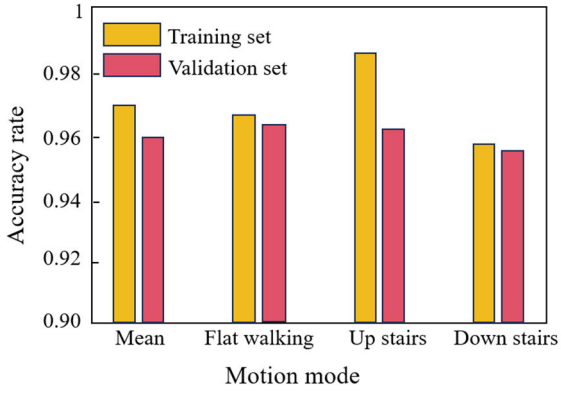


**FIGURE 13.** Motion pattern recognition accuracy.

system. The pressure sensor has a sampling interval of 150ms, and the IMU has a sampling interval of 10ms. The data acquisition card and IMU transmit data to the gait recognition and acquisition control system of the upper computer through the Bluetooth module. Upon receiving data from the plantar pressure data acquisition system, the system preprocesses and assesses the current human gait and intention.

### *A. MOTION PATTERN RECOGNITION EXPERIMENT*

The foot motion data acquisition system was utilized to collect data for flat walking, walking up stairs, and walking down stairs. 10 gait cycles were collected for each mode. For flat walking, as shown in Fig.10, 6460 frames were collected, for walking up stairs, 7060 frames were collected, and for walking down stairs, 6521 frames were collected. To more clearly observe changes in plantar pressure data, the data are displayed as a color image. However, color images have large storage requirements and relatively slow processing speeds. Therefore, this paper opts for grayscale images to create the dataset. The segmented plantar pressure data and foot motion data are combined to form a foot motion data image. The pressure sensor in this paper has a sampling interval of 150ms, and the IMU has a sampling interval of 10ms. Therefore, 1 frame of plantar pressure data corresponds to 15 frames of foot position shift data. The 15 frames of foot movement data corresponding to every 1 frame of pressure data are normalized. The data are then enlarged by 60 times to align with the pressure data range, resulting in an image size of $2 \times 15$. Among them, the first set of behaviors is Y-axis displacement data, and the second set is Z-axis displacement data. The pressure data image was combined with the foot movement data image to form the foot movement data image. The combined image size is $34 \times 16$, comprising plantar pressure images from the 1st to the 32nd, and foot movement data images on the 33rd and 34th lines. The foot motion data image data for different gait phases is displayed in Fig.11.

The flat walking data were labeled as 1, the up stairs data as 2, and the down stairs data as 3. Subsequently, the human movement data were collected and divided into the training set and validation set. Foot motion data images were input into the network for learning, and the network's hyperparameters were adjusted to minimize the loss function and achieve optimal recognition. It was found that setting the learning rate to 0.001, employing a batch size of 32, and conducting training for 11 epochs results in the network's loss function reaching its minimum value. Fig.12 and Fig.13 depict the loss function and recognition accuracy of the training set and validation set, respectively, for different activities. The outcome indicates that the average recognition accuracy of the training set is 0.97, and that of the validation set is 0.96.

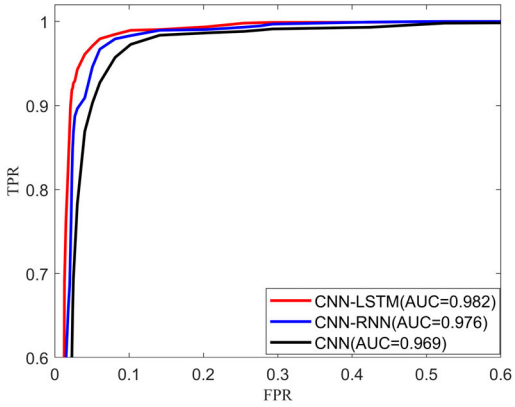


**FIGURE 14.** ROC curves of different algorithm models.

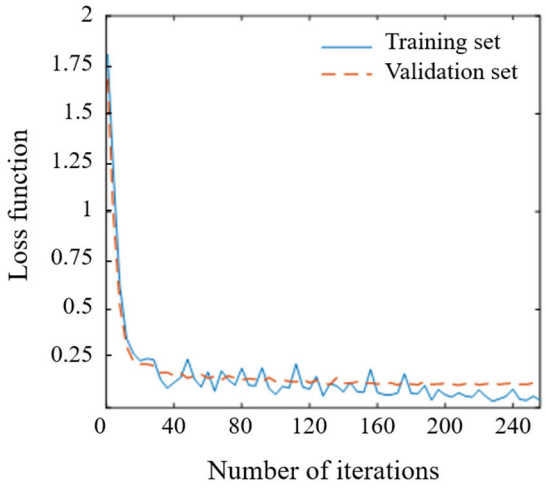


**FIGURE 15.** CNN training loss function.

Receiver Operating Characteristic (ROC) curves are commonly utilized to evaluate the performance of deep learning models. The horizontal axis of the ROC curve represents the false positive rate (FPR), while the vertical axis denotes the true positive rate (TPR). The closer the ROC curve approaches the top left corner, the better the model's performance. The Area Under the Curve (AUC) signifies the area beneath the ROC curve, with higher values indicating superior model performance. Fig.14 illustrates the application of three methods (CNN, CNN-RNN, and CNN-LSTM) for identifying human behavioral motion patterns in the dataset presented in this paper, with corresponding ROC curves plotted. The findings demonstrate that the CNN-LSTM algorithm utilized in this paper surpasses the other two algorithm models, positioning closer to the upper-left corner of the ROC plot, and yielding a larger AUC value compared to the other two algorithms. Thus, indicating the superiority of the algorithm proposed in this paper.

### B. GAIT RECOGNITION EXPERIMENT

Labels were added to foot data images of walking, climbing, and descending steps based on their gait phase. CNN under different movement modes was trained separately, and model

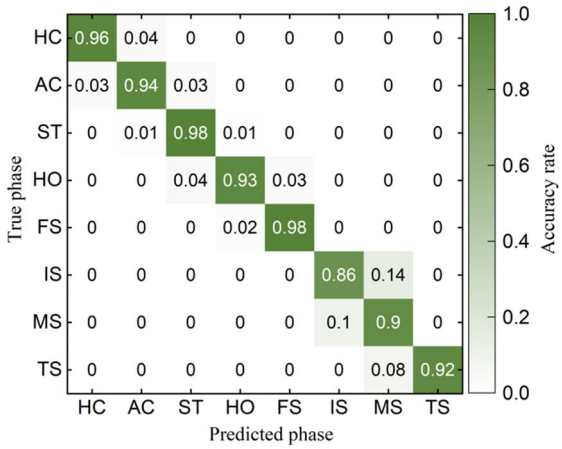


**FIGURE 16.** Flat walking recognition accuracy.

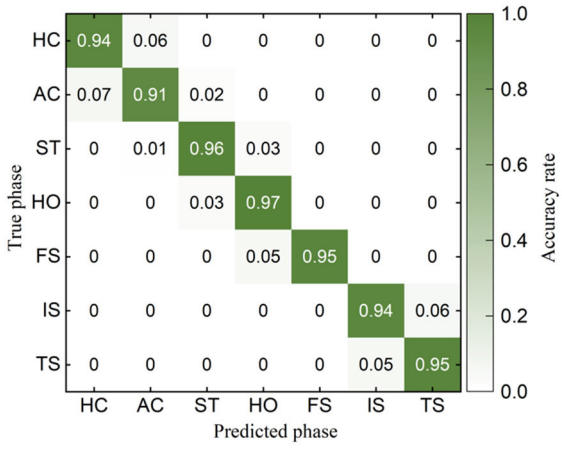


**FIGURE 17.** Up stairs recognition accuracy.

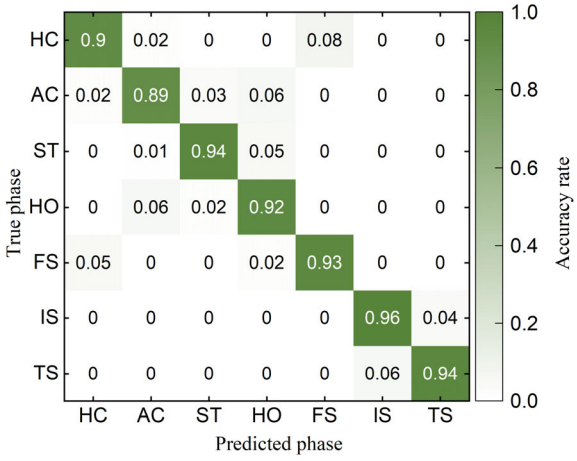


**FIGURE 18.** Down stairs recognition accuracy.

parameters were saved. It is found that when the learning rate is set to 0.005, the batch size is 64, and the training period is 20, the loss function of the gait recognition network is minimized, and the error between the recognition result and the data label is relatively small. At this point, the training loss function of the network is displayed in Fig.15.

**TABLE 3. Comparison of recognition accuracy for different recognition methods.**

| Method | BP | SVM | CNN |
|---|---|---|---|
| Flat walking accuracy | 89.85% | 91.41% | 93.38% |
| Up stairs accuracy | 90.31% | 92.62% | 94.57% |
| Down stairs accuracy | 90.67% | 91.93% | 92.57% |

Real-time gait phase recognition tests were performed, collecting ten gait cycles each for flat walking, up stairs, and down stairs. The trained CNN gait recognition model was utilized for testing, and the recognition results were depicted in a confusion matrix to assess confusion levels among different states. Fig.16-Fig.18 display the gait recognition accuracy for flat walking, up stairs, and down stairs, respectively.

Based on the aforementioned experimental findings, the overall intention recognition accuracy is as follows: the gait recognition accuracy for flat walking is 93.38%, while for walking up and down stairs, it is 94.57% and 92.57%, respectively. In Fig.16, recognition errors during flat walking primarily occur during the transition between adjacent gaits due to the similarity in foot movement data images. In Fig.17, the short duration of arch contact during up stairs leads to less prominent features in the plantar pressure image, resulting in a recognition accuracy of approximately 0.91. In contrast to flat walking, the swing phase gait recognition accuracy is higher during up stairs, attributed to the distinct characteristics of swing phase initiation and termination involving increased and decreased displacements. In Fig.18, during down stairs, pressure nephrograph characteristics of the toe and mid-contact states during ground contact resemble those of forefoot support and heel-off states during takeoff, primarily distinguished by foot movement characteristics, resulting in a recognition accuracy of approximately 0.89 for the foot touch phase.

To assess the effectiveness of the CNN network employed in this paper for gait phase recognition, the backpropagation (BP) neural network and support vector machine (SVM) was utilized to compare and identify the gait phases during flat walking, up stairs, and down stairs. In the BP neural network, the number of hidden layer neurons was set to 25, with the activation function tanh and transfer function purelin. The maximum number of iterations was 200, the learning rate was 0.01, and the training error target was 0.00001. For SVM, the Radial Basis Function (RBF) was selected as the kernel function, and optimal parameter values were determined through grid search. Ultimately, the optimal parameter values obtained were penalty factor $C = 2.3265$ and kernel function parameter gamma $= 0.0247$. The experimental results are presented in Table 3.

Upon inspection of Table 3, it is evident that the recognition accuracy achieved by the CNN in this paper is significantly superior to that of the other two methods. The average recognition accuracy of gait phases across the three motion modes is 93.51%, demonstrating the CNN network's capability to accurately identify human motion gait phases and achieve commendable outcomes in gait recognition and intention perception.

## VI. CONCLUSION

This paper focuses on investigating human behavioral intentions. It proposes a wearable system for acquiring foot movement data to capture high-resolution plantar pressure and foot movement characteristics. it analyzes common human movement patterns and gait phases. A CNN-LSTM network is constructed for recognizing human movement patterns, including identifying states such as flat walking and walking up and down stairs. Additionally, a CNN network is constructed to recognize gait phases during motion. Finally, the test of flat walking, walking up and down stairs was carried out. The results indicate that the proposed intention perception method effectively perceives intentions during various human movements with good outcomes. This method can offer fundamental information for the control system of exoskeleton robots and can also detect the motion state of the human.

## REFERENCES

[1] B. Chen, H. Ma, L.-Y. Qin, F. Gao, K.-M. Chan, S.-W. Law, L. Qin, and W.-H. Liao, "Recent developments and challenges of lower extremity exoskeletons," *J. Orthopaedic Transl.*, vol. 5, pp. 26–37, Apr. 2016.

[2] N. S. S. Sanjeevi, Y. Singh, and V. Vashista, "Recent advances in lower-extremity exoskeletons in promoting performance restoration," *Current Opinion Biomed. Eng.*, vol. 20, Dec. 2021, Art. no. 100338.

[3] R. Zheng, Z. Yu, H. Liu, J. Chen, Z. Zhao, and L. Jia, "End-to-end high-level control of lower-limb exoskeleton for human performance augmentation based on deep reinforcement learning," *IEEE Access*, vol. 11, pp. 102340–102351, 2023.

[4] Y. Long, Z. Du, L. Cong, W. Wang, Z. Zhang, and W. Dong, "Active disturbance rejection control based human gait tracking for lower extremity rehabilitation exoskeleton," *ISA Trans.*, vol. 67, pp. 389–397, Mar. 2017.

[5] J. Ji, W. Chen, W. Wang, and J. Xi, "Design and control of an omni-directional robotic Walker based on human–machine interaction," *IEEE Access*, vol. 9, pp. 111358–111367, 2021.

[6] Y. Sun, Y. Tang, J. Zheng, D. Dong, X. Chen, and L. Bai, "From sensing to control of lower limb exoskeleton: A systematic review," *Annu. Rev. Control*, vol. 53, pp. 83–96, 2022.

[7] A. C. Villa-Parra, D. Delisle-Rodríguez, A. López-Delis, T. Bastos-Filho, R. Sagaró, and A. Frizera-Neto, "Towards a robotic knee exoskeleton control based on human motion intention through EEG and SEMGsignals," *Proc. Manuf.*, vol. 3, pp. 1379–1386, Aug. 2015.

[8] Y. He, D. Eguren, J. M. Azorín, R. G. Grossman, T. P. Luu, and J. L. Contreras-Vidal, "Brain–machine interfaces for controlling lower-limb powered robotic systems," *J. Neural Eng.*, vol. 15, no. 2, Apr. 2018, Art. no. 021004.

[9] J. Ryu, B.-H. Lee, and D.-H. Kim, "SEMG signal-based lower limb human motion detection using a top and slope feature extraction algorithm," *IEEE Signal Process. Lett.*, vol. 24, no. 7, pp. 929–932, Jul. 2017.

[10] F. Peng, C. Zhang, B. Xu, J. Li, Z. Wang, and H. Su, "Locomotion prediction for lower limb prostheses in complex environments via sEMG and inertial sensors," *Complexity*, vol. 2020, pp. 1–12, Dec. 2020.

[11] S. Hazra, A. A. Pratap, D. Tripathy, and A. Nandy, "Novel data fusion strategy for human gait analysis using multiple Kinect sensors," *Biomed. Signal Process. Control*, vol. 67, May 2021, Art. no. 102512.

[12] A. S. M. H. Bari and M. L. Gavrilova, "Artificial neural network based gait recognition using Kinect sensor," *IEEE Access*, vol. 7, pp. 162708–162722, 2019.

[13] R. M. Kanko, E. K. Laende, E. M. Davis, W. S. Selbie, and K. J. Deluzio, "Concurrent assessment of gait kinematics using marker-based and markerless motion capture," *J. Biomechanics*, vol. 127, Oct. 2021, Art. no. 110665.
[14] S. K. Yadav, K. Tiwari, H. M. Pandey, and S. A. Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions," *Knowl.-Based Syst.*, vol. 223, Jul. 2021, Art. no. 106970.
[15] C. Virginia Anikwe, H. Friday Nweke, A. Chukwu Ikegwu, C. Adolphus Egwuonwu, F. Uchenna Onu, U. Rita Alo, and Y. Wah Teh, "Mobile and wearable sensors for data-driven health monitoring system: State-of-the-art and future prospect," *Expert Syst. Appl.*, vol. 202, Sep. 2022, Art. no. 117362.
[16] G. Bastas, J. J. Fleck, R. A. Peters, and K. E. Zelik, "IMU-based gait analysis in lower limb prosthesis users: Comparison of step demarcation algorithms," *Gait Posture*, vol. 64, pp. 30–37, Jul. 2018.
[17] D. H. Yoon, J.-H. Kim, K. Lee, J.-S. Cho, S.-H. Jang, and S.-U. Lee, "Inertial measurement unit sensor-based gait analysis in adults and older adults: A cross-sectional study," *Gait Posture*, vol. 107, pp. 212–217, Jan. 2024.
[18] M. Wang, X. Wang, Z. Fan, F. Chen, S. Zhang, and C. Peng, "Research on feature extraction algorithm for plantar pressure image and gait analysis in stroke patients," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 525–531, Jan. 2019.
[19] N. T. Beigh, F. T. Beigh, and D. Mallick, "Machine learning assisted hybrid transduction nanocomposite based flexible pressure sensor matrix for human gait analysis," *Nano Energy*, vol. 116, Nov. 2023, Art. no. 108824.
[20] M. F. Antwi-Afari, H. Li, S. Anwer, S. K. Yevu, Z. Wu, P. Antwi-Afari, and I. Kim, "Quantifying workers' gait patterns to identify safety hazards in construction using a wearable insole pressure system," *Saf. Sci.*, vol. 129, Sep. 2020, Art. no. 104855.
[21] W. Li, W. Lu, X. Sha, H. Xing, J. Lou, H. Sun, and Y. Zhao, "Wearable gait recognition systems based on MEMS pressure and inertial sensors: A review," *IEEE Sensors J.*, vol. 22, no. 2, pp. 1092–1104, Jan. 2022.
[22] F. Zhang, X. Xu, Q. Lin, D. Xian, K. Yao, N. Zhao, L. Zhao, and Z. Jiang, "The flexible and distributed pressure sensor with 64 units for on-line gait recognition analysis," *Measurement*, vol. 223, Dec. 2023, Art. no. 113726.
[23] S. Ding, X. Ouyang, T. Liu, Z. Li, and H. Yang, "Gait event detection of a lower extremity exoskeleton robot by an intelligent IMU," *IEEE Sensors J.*, vol. 18, no. 23, pp. 9728–9735, Dec. 2018.
[24] C. Lou, S. Wang, T. Liang, C. Pang, L. Huang, M. Run, and X. Liu, "A graphene-based flexible pressure sensor with applications to plantar pressure measurement and gait analysis," *Materials*, vol. 10, no. 9, p. 1068, Sep. 2017.
[25] X. Zhang, S. Chang, and Z. Tong, "Facile fabrication of a highly sensitive and robust flexible pressure sensor with batten microstructures," *Micromachines*, vol. 13, no. 8, p. 1164, Jul. 2022.
[26] M. Han, J. Lee, J. K. Kim, H. K. An, S.-W. Kang, and D. Jung, "Highly sensitive and flexible wearable pressure sensor with dielectric elastomer and carbon nanotube electrodes," *Sens. Actuators A, Phys.*, vol. 305, Apr. 2020, Art. no. 111941.
[27] R. Li, Q. Zhou, Y. Bi, S. Cao, X. Xia, A. Yang, S. Li, and X. Xiao, "Research progress of flexible capacitive pressure sensor for sensitivity enhancement approaches," *Sens. Actuators A, Phys.*, vol. 321, Apr. 2021, Art. no. 112425.
[28] A. Kumar, "Recent progress in the fabrication and applications of flexible capacitive and resistive pressure sensors," *Sens. Actuators A, Phys.*, vol. 344, Sep. 2022, Art. no. 113770.
[29] I. Papavasileiou, W. Zhang, and S. Han, "Real-time data-driven gait phase detection using ground contact force measurements: Algorithms, platform design and performance," *Smart Health*, vols. 1–2, pp. 34–49, Jun. 2017.
[30] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022.
[31] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," 2015, *arXiv:1511.08458*.
[32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[33] J. Narayan, S. Johri, and S. K. Dwivedy, "LSTM neural network-based classification of sensory signals for healthy and unhealthy gait assessment," in *Predictive Modeling in Biomedical Data Mining and Analysis*. New York, NY, USA: Academic, 2022, pp. 207–223.

**FANG YANG** was born in Hebei, China, in 1982. She received the M.S. degree in nursing from the North China University of Science and Technology, Tangshan, Hebei, China, in 2013. From 2014 to 2022, she was a Lecturer with the College of Nursing and Rehabilitation, North China University of Science and Technology, where she has been an Associate Professor with the College of Nursing and Rehabilitation, since 2023. Her research interests include nursing education, clinical nursing, and the design and development of rehabilitation robots.

**QIANQIAN ZHENG** was born in Hebei, China, in 1996. She received the master's degree from the North China University of Science and Technology, Tangshan, Hebei, China. Her research interests include nursing education and nursing informatics.

**LIANQING CHEN** received the B.S. degree from the Computer Department, Huaibei Coal Industry Teachers College, Huaibei, Anhui, China, in 2005, and the M.A. (Eng.) degree in computer technology engineering from the North China University of Science and Technology, Tangshan, Hebei, China, in 2015. He is currently an Engineer and the Office Director of the College of Nursing and Rehabilitation, North China University of Science and Technology. His research interests include medical information engineering and medical information management.

**XIONG TAN** was born in Hunan, China, in 1998. He is currently pursuing the master's degree with Hebei University of Technology. His research interests include gait recognition and rehabilitation robots.

**PENGCHENG CHE** was born in Hebei, China, in 1970. He received the Ph.D. degree in surgery from Hebei Medical University, Hebei. He is currently a Professor and the Dean of the College of Nursing and Rehabilitation, North China University of Science and Technology. He has published over 50 academic articles. His research interests include rehabilitation engineering and regenerative medicine. He was awarded a Key Project of the Joint Funds of the National Natural Science Foundation of China and a Young Scientists Fund of the National Natural Science Foundation of China. He won the Third Prize for Scientific and Technological Progress in the Chinese People's Armed Police Force.