



# Integral Real-time Locomotion Mode Recognition Based on GA-CNN for Lower Limb Exoskeleton

Jiaqi Wang<sup>1</sup> · Dongmei Wu<sup>1</sup> · Yongzhuo Gao<sup>1</sup> · Xinrui Wang<sup>2</sup> · Xiaoqi Li<sup>2</sup> · Guoqiang Xu<sup>2</sup> · Wei Dong<sup>1</sup> 

Received: 18 March 2022 / Revised: 2 June 2022 / Accepted: 7 June 2022 / Published online: 18 July 2022  
© The Author(s) 2022

## Abstract

The wearable lower limb exoskeleton is a typical human-in-loop human–robot coupled system, which conducts natural and close cooperation with the human by recognizing human locomotion timely. Requiring subject-specific training is the main challenge of the existing approaches, and most methods have the problem of insufficient recognition. This paper proposes an integral subject-adaptive real-time Locomotion Mode Recognition (LMR) method based on GA-CNN for a lower limb exoskeleton system. The LMR method is a combination of Convolutional Neural Networks (CNN) and Genetic Algorithm (GA)-based multi-sensor information selection. To improve network performance, the hyper-parameters are optimized by Bayesian optimization. An exoskeleton prototype system with multi-type sensors and novel sensing-shoes is used to verify the proposed method. Twelve locomotion modes, which composed an integral locomotion system for the daily application of the exoskeleton, can be recognized by the proposed method. According to a series of experiments, the recognizer shows strong comprehensive abilities including high accuracy, low delay, and sufficient adaption to different subjects.

**Keywords** Locomotion mode recognition · Gait mode detection · Lower limb exoskeleton · Convolutional neural network · Genetic algorithm · Bionic design

## 1 Introduction

Technology advancements have led to the development of numerous wearable exoskeleton robots for military and medical purposes through abilities augment and physical assistance [1, 2]. Developing appropriate strategies to predict the intentions of wearers is one of the challenging issues. The availability of an exoskeleton system is extremely attractive for many applications, e.g., power supporting during daily living activities by automatically recognizing the physical activities of a human subject.

The locomotion for the lower limb can be classified into various modes according to the ground conditions (e.g., flat, stairs, ramp, etc.), and the kinematic as well as kinetic characteristics are different under different locomotion modes. Human can intuitively achieve these locomotion modes by

the innate motion system that employs different strategies. While as to exoskeleton, the complexity of locomotion must be considered in the control algorithms to assist the human movement more efficiently and naturally [3].

In recent years, many studies have proposed different approaches for the recognition of lower limb exoskeleton locomotion modes [4, 5]. The offline recognition method [6–8] can analyze human locomotion or generate trajectories. In order for the exoskeleton to effectively assist, the subject's motion intention must be decoded online. For sensing signals, some methods recognize human locomotion based on electromyography (EMG) [9–12] and electroencephalograph (EEG) [13–16] signals in recent years. However, EMG is sensitive to environmental conditions in the application, such as electric noise, and skin conditions [17]. About EEG, the unmaturing technology and the complexity of the signal process restrict the use of this method in portable devices. Therefore, since EMG and EEG suffer from poor repeatability, more approaches prefer the mechanical sensors network to increase the system's robustness and dependability [18–21].

As regards the recognition method, the rule-based strategy such as fuzzy-logic-based method [3, 20, 22, 23] is an

✉ Wei Dong  
dongwei@hit.edu.cn

<sup>1</sup> State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150001, China

<sup>2</sup> Weapon Equipment Research Institute, China South Industries Group Corporation, Beijing 102202, China

intuitive approach and has been widely used. The principle of the rule-based method is concise and explicit, but rule formulating is intractable and inexact when faced with too many modes [24, 25]. Most importantly, the expandability of rule-based methods may also be a tough one because re-making the rules are needed. Locomotion Mode Recognition (LMR) processed by machine learning is the most common way nowadays. For exoskeletons, the locomotion mode is widely recognized by Linear Discriminant Analysis (LDA) [3, 26] and Artificial Neural Network (ANN) classifier [3, 9]. A classifier based on a Hidden Markov Model (HMM) was employed in walking and jogging discrimination by Mannini and Sabatini [27]. Varol et al. [28] described an intent recognition framework that uses Gaussian Mixture Models (GMM) for the supervisory control. A Support Vector Machine (SVM) machine learning model is established and trained by Gao et al. [29] to classify terrain in lower limb rehabilitation systems. The features extracted by machine learning were more discriminative than the hand-crafted rules. However, it is difficult to search for a good classification model when with large data quantity and poor feature extraction. Seeking more intelligent feature extraction methods and improving the efficiency is still the focus of recognition.

Deep learning models simultaneously integrate and optimize feature extraction and predictive modeling. Those models have grown in popularity and practicality for complex tasks such as recognition and prediction of human motion [30]. Huang et al. [31] used a Recurrent Neural Network (RNN) for the prediction of real-time intended knee joint motion. Long short-term memory (LSTM) neural networks, a special case of RNN, was adopted to identify the wearer's locomotion mode [32, 33]. Gautam et al. [34] introduced a transfer learning based long-term recurrent convolution network (LRCN) for the classification of lower limb movements.

Regarding the locomotion modes of the existing LMR methods, Park et al. [35] proposed a spatio-spectral convolutional neural networks (CNN) having 83.4% accuracy on two gait motion modes recognition. [30, 31] both recognized three basic modes, standing, walking, and sitting. Almost all the recognition methods are for 2–7 locomotion modes [14, 15, 32, 34], and 5 dynamic gait motions are the most common including level walking, stairs ascent, stairs descent, ramp ascent, and ramp descent [3, 20, 31, 36–38]. However, for the exoskeleton daily living assistance, the recognition of all the most common static and dynamic motions is needed.

In this paper, a novel subject-adaptive real-time LMR method based on GA-CNN is proposed for the lower limb exoskeleton, which is a combination of deep CNN and multi-sensor information selection implemented by Genetic Algorithm (GA). The proposed method is an integral locomotion recognition for 12 modes which contain all common

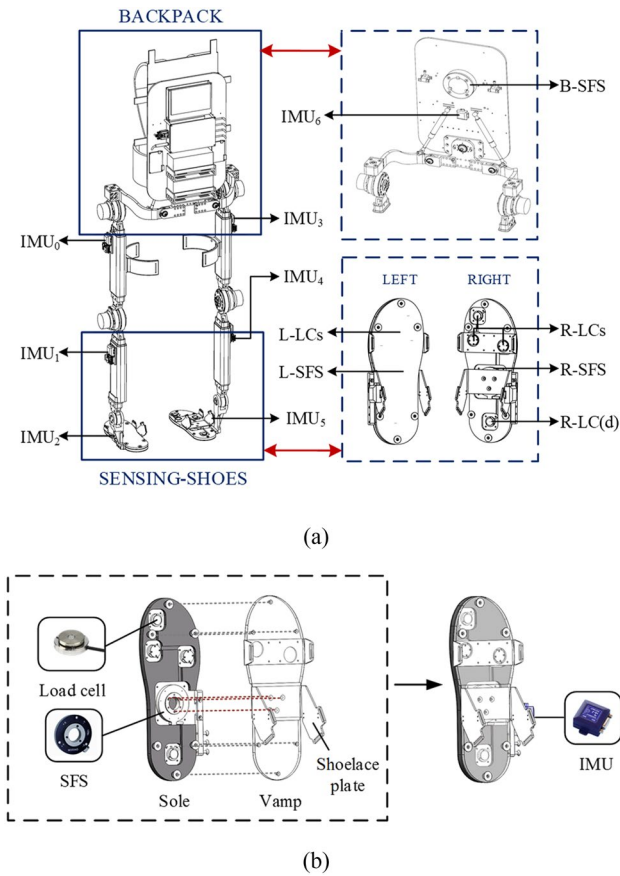
locomotion in daily life. Meanwhile, the method is dedicated to realizing strong comprehensive abilities including high accuracy, low delay, low computational complexity, and sufficient adaption to different subjects. CNN is adopted as the basic framework. Its extraordinary performance in feature extraction is able to reveal the inner principle of different motion data [39–41]. Besides, the weight sharing mechanism of CNN reduces the number of free parameters and computational complexity. To extract the characteristics of human motion more comprehensively and efficiently, the input of CNN must be sufficient and the most simplified. A multi-sensor system can perceive human multi-motion better in all respects, but considering the over-fitting problem and computational complexity the redundant data should be elaborately eliminated. Selecting more representative sensor data can be considered as a searching optimization problem with a high computational cost. For the sake of high efficiency, heuristic methods are usually used [42]. GA is a kind of heuristic method whose encoding method is very friendly to a binary searching problem, so GA is introduced to conduct multi-sensor information selection. Furthermore, Bayesian optimization is adopted to optimize the hyper-parameter of the deep network for better performance.

The comprehensive abilities of the proposed method are evaluated via online experiments. The proposed GA-CNN-based method can achieve an accuracy average of 97.91%. About the generalization, the average recognition accuracy is still up to 93.97% for new subjects. In addition, the real-time ability can be guaranteed with the maximum recognition delay within 60 ms. The proposed method also is compared with two classic strategies (LDA, and ANN), and the necessity of each presented part has also been confirmed by comparative experiments.

## 2 Hardware and Data Acquisition

### 2.1 Exoskeleton Hardware

The designed exoskeleton system called HEXO is shown in Fig. 1a, which is used to validate the proposed LMR method. The detailed mechanical structure and actuation system of the system was described in previous research [43, 44]. About the sensor system related to this paper, the exoskeleton is equipped with adequate sensors distributed throughout the body for observing motion state. To acquire the lower limb motion, four Inertial Measurement Units (IMUs) (LPMS-CU2, ALUBI, China) are mounted on the thighs and shank carbon-fiber limbs. Besides, a Six-axis Force Sensor (SFS) (M3554E, SRI, China) is installed at the back between the user and the exoskeleton to perceive the human–robot torso interaction force.



**Fig. 1** The exoskeleton hardware. **a** The exoskeleton sensor system. The left shoe shows the appearance, the right shoe displays the internal structure. **b** The new-type sensing-shoes. The dotted lines in the figure indicate where the physical connections are. Gray dotted lines indicate where the vamp and sole are connected. The SFS itself is fixed to the sole, and it is connected to the shoelace plate (fixed foot) through the red dotted line

Most importantly, a pair of wearable sensing-shoes is developed, which is shown in Fig. 1b. There are four load cells (AT8106, AUTODA, China) and one SFS for each shoe. Four load cells are installed in sensing-shoes at the hallux, the first metatarsophalangeal joint, the fourth metatarsophalangeal joint and the heel. The SFS in the middle of shoes is installed between lace and sole (Fig. 1b). The foot will be tied to the laces and the exoskeleton shank is attached to the shoe sole, and the SFS can be seen as the only connection between the human foot and the exoskeleton. Therefore, when the foot is lifted the SFS can still detect the human–robot interaction force, which is also the resistance of the exoskeleton. Conventional sensing-shoes [21, 45] usually cannot work in the swing phase. Hence, these new-type sensing-shoes can utilize the time and information when the leg is in the swinging phase. The IMU on the side of the shoe can also help further measure the attitude of the foot.

**Table 1** Detailed information of the nine subjects

Number	Gender	Age (years)	Height (cm)	Weight (kg)	Walking speed (m/s)
1# (Huang)	Male	23	184	57	1–2
2# (Hu)	Male	23	180	58	1–2
3# (Wang)	Male	23	187	70	1
4# (Guo)	Male	24	172	66	1–2
5# (Gao)	Male	29	180	69	1–3
6# (Du)	Male	24	176	72.5	1–3
7# (Li)	Male	26	188	71	1–3
8# (Geng)	Female	24	169	60.5	1
9# (Jiang)	Male	45	179	78	1–2

To sum up, the HEXO is equipped with 18 sensors that include 47 channels of sensor data in total. The whole sensor data set contains left thigh IMU<sub>0</sub>, left shank IMU<sub>1</sub>, left shoe IMU<sub>2</sub>, right thigh IMU<sub>3</sub>, right shank IMU<sub>4</sub>, right shoe IMU<sub>5</sub>, back IMU<sub>6</sub>, back SFS (B-SFS), left shoe SFS (L-SFS) and load cells (L-LCs), right shoe SFS (R-SFS) and load cells (R-LCs). The numbers of load cells from toe to heel are a, b, c, and d. Furthermore, each IMU has three axes of Euler angles, and SFS contains six axes of forces and torques. Therefore, the whole sensor data set can be described below.

$$\left\{ \begin{array}{l} \{ \theta_x, \theta_y, \theta_z \}_{IMU_0, IMU_1, IMU_2, IMU_3, IMU_4, IMU_5, IMU_6} \\ \{ F_x, F_y, F_z, T_x, T_y, T_z \}_{B-SFS, L-SFS, R-SFS} \\ \{ F_a, F_b, F_c, F_d \}_{L-LCs, R-LCs} \end{array} \right\} \quad (1)$$

### 2.2 Data Acquisition Protocol

The proposed LMR method can recognize five dynamic motions, three static states and four transition states. For dynamic motions, there are level walking (LW), stairs ascent (SA), stairs descent (SD), ramp ascent (RA), and ramp descent (RD). Static states include standing (ST), sitting (SI), and squatting (SQ). Transition states denote the transition process between the three static states, i.e., stand to sit (STSI), sit to stand (SIST), stand to squat (STSQ), and squat to stand (SQST). To sum up, a total of 12 locomotion modes are studied in the LMR method. Six healthy subjects, whose characteristics are listed in Table 1, participated in the data collection experiments. The remaining three are employed in the online experiments. Corresponding mimic scenarios are used to evaluate the recognition performance. The staircase is 12 cm in height, 80 cm in width, and 30 cm in-depth, and the ramp slope is 15°.

In this experiment, all six subjects are asked to perform 12 locomotion modes when wearing the exoskeleton with

zero force (walking with HEXO, but no torque assistance is enabled). All modes are repeated five times. There are 20–30 steps for dynamic motions every time. Static states and transition states continuously perform 5–10 every time. To evaluate the adaptability of different subjects of the proposed method, all subjects are asked to perform locomotion with their natural gaits (e.g., different step lengths and different step speeds).

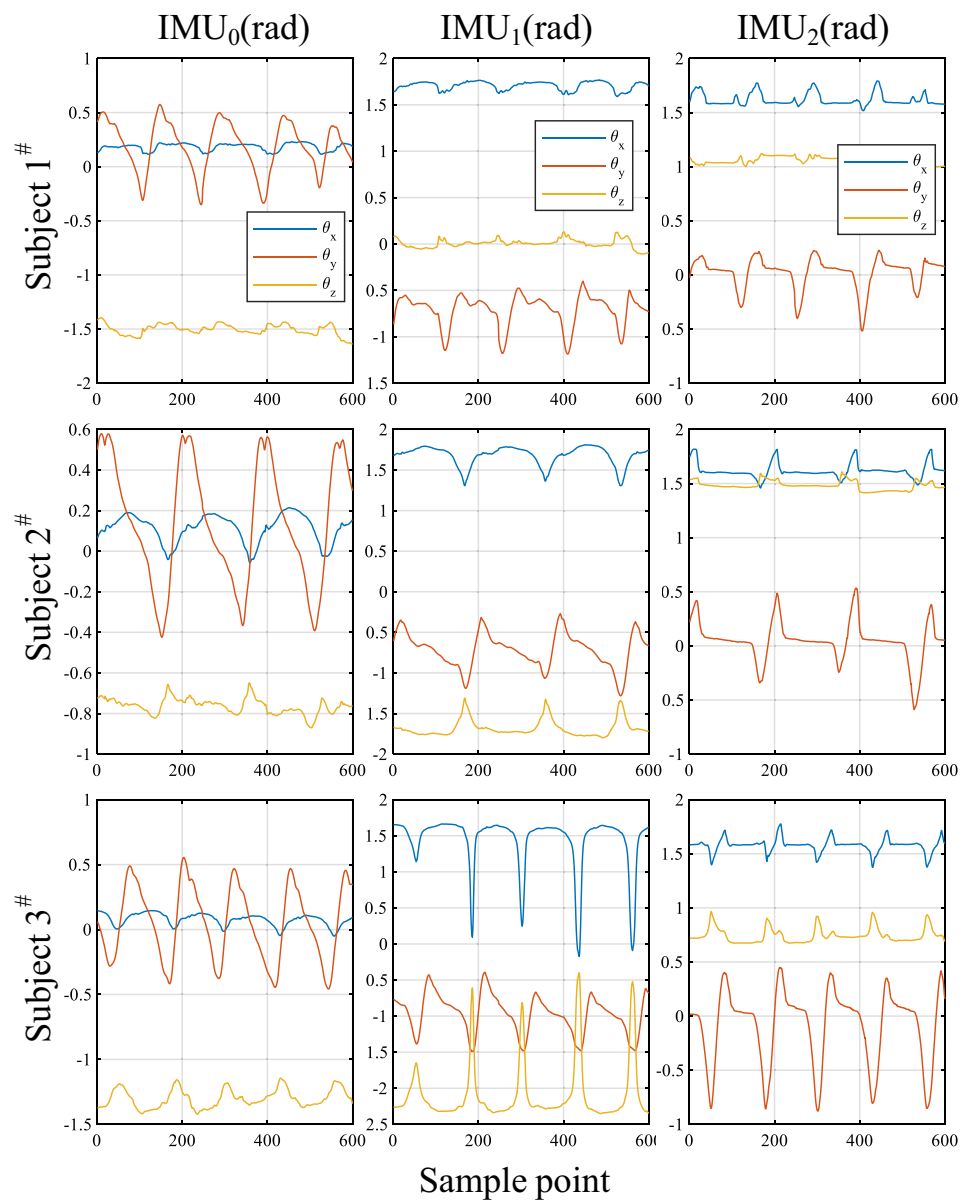
When using the acquired data for network training, the data amount of each mode is similar so that the unbalanced training issue can be avoided. Besides, 80% of the dataset of six subjects is chosen to be the training set, and the remaining is the validation set. The training set is used to optimize

the parameters of the model, and the validation set is used to optimize hyper-parameters.

As is seen in Fig. 2, the data curve of the same mode performed by different subjects can be very different. It can be derived that the detailed shape of the curves varies from person to person although the approximate shapes are similar. This is the reason why CNN is introduced to extract core and features from deformed and stretched sensor data.

Moreover, to eliminate the great magnitude differences of input data and avoid prediction errors, data normalization is employed. Before the sensor data is used as the input of the CNN, all the data are uniformly normalized to [0, 1]. The corresponding conversion is formulated as below,

**Fig. 2** A part of the sensor data of three subjects performing LW. Displayed sensor including IMU<sub>0</sub>, IMU<sub>1</sub> on the left leg and IMU<sub>2</sub> on the left foot



$$x^* = \begin{cases} \frac{x-\text{MIN}}{\text{MAX}-\text{MIN}} & 0 \leq \frac{x-\text{MIN}}{\text{MAX}-\text{MIN}} \leq 1 \\ 1 & \frac{x-\text{MIN}}{\text{MAX}-\text{MIN}} > 1 \\ 0 & \frac{x-\text{MIN}}{\text{MAX}-\text{MIN}} < 0, \end{cases} \quad (2)$$

where  $x$  is the raw sensor data,  $x^*$  is the normalized sensor data, MAX is a parameter that is larger than the maximum value of  $x$  in the training dataset, MIN is a parameter that is smaller than the minimum value of  $x$  in the training dataset.

### 3 LMR Method Based on GA-CNN

The procedure of the LMR method based on GA-CNN is described in Fig. 3. Pre-collected locomotion data are initially used to train a deep CNN learning model offline, and the model is subsequently adopted to recognize the subject's locomotion online. This section first shows the architecture of CNN, and then is the determination of the input. The input is raw sensor data in a sliding time-window with a specific length and width. The length is determined by experimental trade-offs between efficiency and accuracy, and the width is the number of sensor channels after rational multi-sensor information selection using GA. Afterward, last but not least, hyper-parameter optimization is implemented to get a more efficient model.

#### 3.1 CNN Architecture

The architecture of the CNN model is illustrated in Fig. 4. The input of CNN is a  $l_1 \times h_1$  time-window data frame. Except for the input and output layer, the entire CNN architecture contains five hidden learning layers: two convolutional layers, two pooling layers and one fully connected layer.

As is shown in Fig. 4,  $t$  is the current time and  $T$  is the time of the time-window data frame. The length of the input layer

$l_1$  which is the length of the data frame equals  $T$  divided by  $T_s$  (sampling period) and the width of the input layer  $h_1$  is the number of the remaining sensor channels after GA optimizes sensor selection. The first convolutional layer filters the  $l_1 \times h_1$  input data frame with  $w_1$  kernels of size  $f_{11} \times f_{12}$  with a stride of 1 data point. The feature map size of the first convolutional layer is  $l_{C1} \times h_{C1}$  the same as the previous input  $l_1 \times h_1$ . The purpose of convolution is to extract the different characteristics of the input. The max-pooling layer follows the convolutional layer, used for dimensional compression and selecting more obvious features. The first pooling layer pools the  $l_{C1} \times h_{C1}$  input data frame with  $w_1$  kernels (the number of pooling layer kernels is equal to the previous layer) of size  $p_{11} \times p_{12}$  with a stride of 2 data points. The feature map size of the first pooling layer is  $l_{L1} \times h_{L1}$ , which can be calculated in Eq. (3),

$$\begin{cases} h_{L1} = \frac{h_{C1} + 2 \cdot P - p_{11}}{S} + 1 \\ l_{L1} = \frac{l_{C1} + 2 \cdot P - p_{12}}{S} + 1, \end{cases} \quad (3)$$

where  $P=0$  (padding pattern is VAILD) and  $S=2$  (stride is 2). The convolutional layer 2 and the pooling layer 2 are in the same situation as the above description. The CNN becomes deeper through repeatedly stacking convolutional-pooling layer, and more global features and higher dimensional features can be extracted.  $n_F$  denotes the number of neurons in the fully connected layer.

Algebraically, the convolutional layers can be expressed as,

$$a_j^l = \sigma \left( \sum_{i=1}^{M_{l-1}} (a_i^{l-1} * W_j^l) + b_j^l \right), \quad (4)$$

where  $a_j^l$  represents the output feature map of the  $j$ th channel (kernel) in layer  $l$ ,  $a_i^{l-1}$  is the output of the  $i$ th channel of the previous layer,  $W_j^l$  is the convolution kernel matrix, and  $*$  denotes the convolution operator,  $M_{l-1}$  represents the total channel of the previous layer, and  $b_j^l$  represents the bias.  $\sigma(\cdot)$  is the activation function, and the leaky Rectified

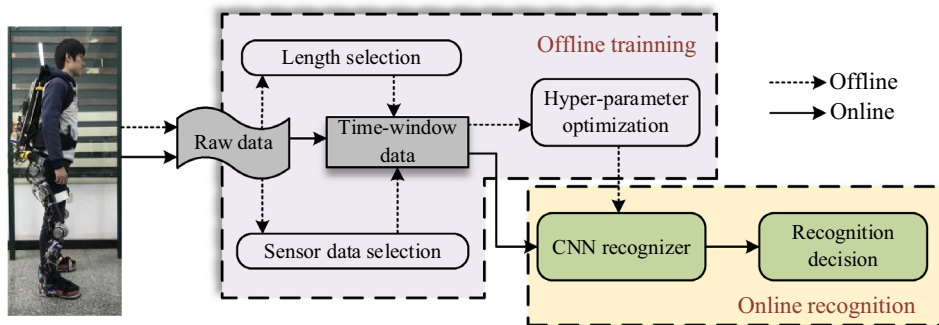


Fig. 3 The procedure of the LMR method based on GA-CNN



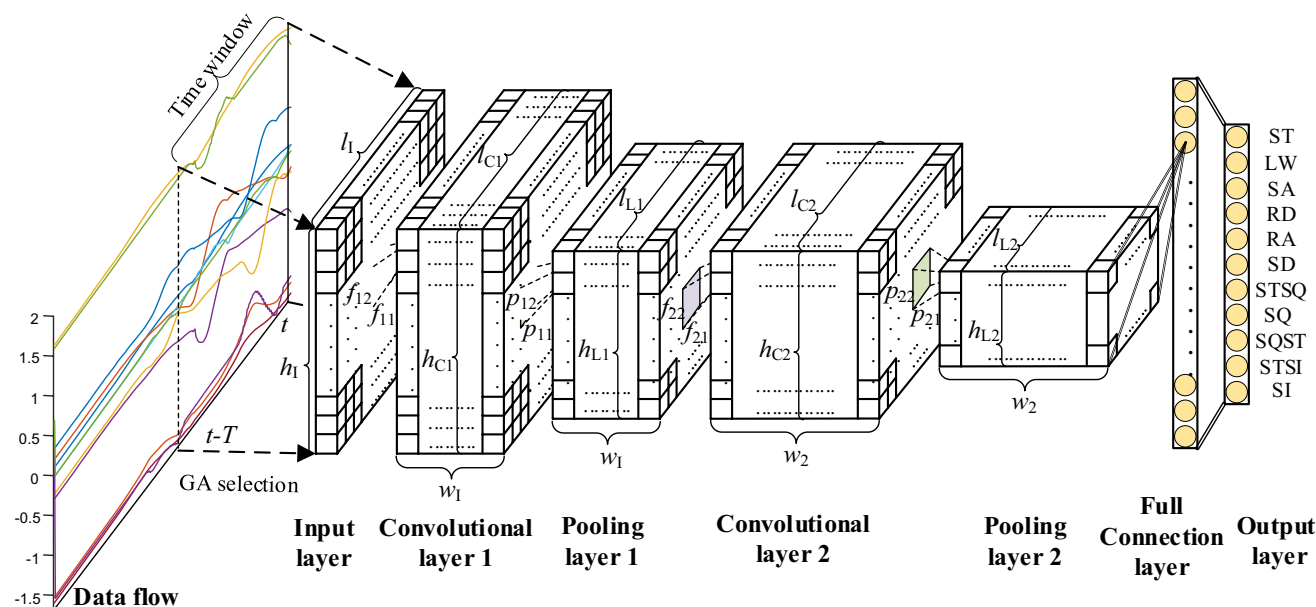


Fig. 4 Architecture of the deep CNN model.  $t$  is the current time.  $T$  is the time of time-window data

Linear Unit (ReLU) is applied in this case. The formula of ReLU activation function is expressed as

$$\sigma(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0. \end{cases} \tag{5}$$

The algebraic expression of the pooling layers can be written as

$$a_j^l = \max(a_j^{l-1}), \tag{6}$$

where  $a_j^l$  is the output feature maps of the  $j$ th channel in pooling layer  $l$ , and  $\max(\cdot)$  represents the max-pooling operation. The output of the fully connected layer can be calculated through the following equation,

$$a^l = \sigma(\alpha \cdot a^{l-1} + b^l), \tag{7}$$

where  $a^l$  is the output of the fully connected layer  $l$ ,  $\alpha$  is the weight coefficient of the fully connected layer,  $a^{l-1}$  denotes the output feature graphs of the upper layer, and  $b^l$  is the bias. Dropout with a probability of 0.5 is also implemented in the fully connected layer to reduce over-fitting. The 12-channel softmax function in the last fully connected layer produces a distribution of 12 class labels corresponding to 12 locomotion modes,

$$y_k = \frac{e^{a^k}}{\sum_{c=1}^N e^{a^c}}, \tag{8}$$

where  $y_k$  is the  $k$ th output of the softmax,  $a^c$  and  $a^k$  are the outputs of the fully connected layer, and  $N$  is the total number of the locomotion modes.

The loss function is cross entropy,

$$L = \frac{1}{M} \sum_{i=1}^M \left( - \sum_{k=1}^N \hat{y}_k \log y_{ik} \right), \tag{9}$$

where  $M$  is training batch size,  $i$  denotes each sample in the training batch,  $N$  is the total number of the locomotion mode, and  $k$  denotes each mode.  $\hat{y}_{ik}$  is the ground truth (0 or 1) of the  $k$ th mode, and  $y_{ik}$  is the calculated softmax output of the  $i$ th sample for the  $k$ th mode. Furthermore, the Adaptive Moment Estimation (Adam) optimizer is employed to optimize the weights and biases in CNN.

### 3.2 Time-window Length

For time-window flow data, the window size is essential to the system's efficiency. Hence, it is necessary to make a tradeoff between recognition efficiency and computational consumption for time-window length. In this paper, time-window length  $l_1$  is determined by cross entropy in the validation set and calculation consumption. All multi-sensor channel data are adopted in this section. A set of empirical values are adopted as the hyper-parameters in experiments before the optimization process.

As can be seen from Fig. 5, the calculation consumption constantly increases with the increase of  $l_1$ . While to cross entropy, it is distinctly decreasing with the increase

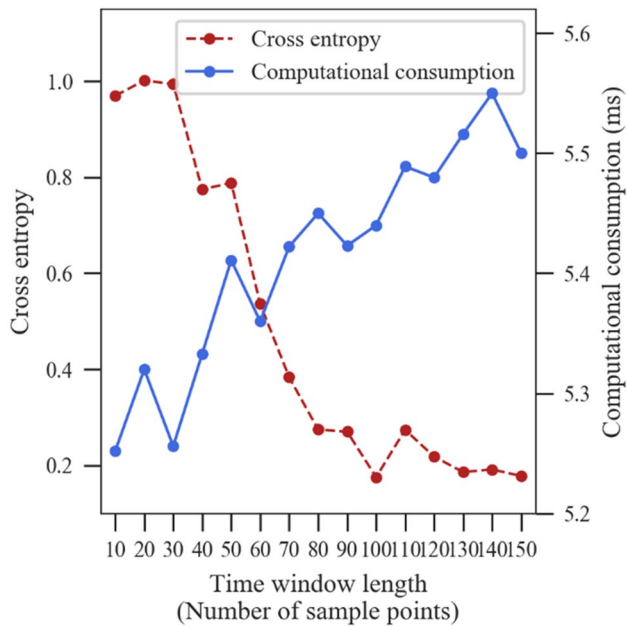


Fig. 5 Cross entropy and computational consumption by different time-window length

of  $l_1$  until  $l_1$  reaches 100. Consequently, the length of the time-window  $l_1$  is chosen as 100 in this work. In addition, the sliding step (window increment) is 10 ms as same as the sampling period.

### 3.3 GA-Based Multi-sensor Information Selection

To get a more representative but simplified input frame, the multi-sensor information selection is implemented. The operation of the multi-sensor information selection consists of two steps. First, the sensor data with apparently poor performance were manually removed. By observing the raw data, it is found that drifting always exists in the  $\theta_z$  channel of the IMU, which is irregular and uncorrelated to the motion. Hence, all of them are removed permanently. After manual sensor selection, 40 channels remain. The second step is the automatic sensor selection. The GA simulates the evolution process of the artificial population. The sensor signals are combined as different candidate individuals of the population. Through the mechanisms of Selection, Crossover and Mutation, a set of sensor candidate individuals is retained in each iteration. The sensor candidate set will be used to conduct a recognition mission. The GA aims to find the best set with minimal cross entropy in the recognition result. After the population evolves during several iterations, the optimal combination of sensor data is elected.

The combination of different sensor data is encoded in binary, and each bit represents a sensor channel. “1” means the sensor channel is selected and “0” means it is excluded.

Hence, the chromosome size of GA is 40. There is a total of 30 individuals in the population, and the crossover rate and mutation rate are 0.75 and 5/40 respectively. In the process of cross entropy descending in evolution as shown in Fig. 6, the cross entropy begins to converge to around 0.15 from the 22nd generation. The minimum cross entropy reaches 0.14396, and the corresponding individual is the optimal data set listed as follows.

$$\left. \begin{array}{l} \{\theta_x\}_{IMU_0,IMU_1,IMU_3,IMU_5,IMU_6} \\ \{\theta_y\}_{IMU_0,IMU_1,IMU_2,IMU_3,IMU_4,IMU_5,IMU_6} \\ \{F_x\}_{L-SFS,R-SFS} \\ \{F_y\}_{B-SFS,L-SFS,R-SFS} \\ \{T_x\}_{L-SFS} \\ \{T_y\}_{L-SFS,R-SFS} \\ \{T_z\}_{B-SFS,L-SFS} \\ \{F_a\}_{L-LCs,R-LCs} \end{array} \right\} \quad (10)$$

Finally, a total of 24 data channels are left. Therefore, for the proposed recognition, only 24 data channels are adopted all the time.

### 3.4 Hyper-parameter Optimization

Hyper-parameters have a great impact on network performance. Unlike grid search and random search, Bayesian optimization makes full use of the previous information when evaluating the next one, which can speed up the searching process. Thus, Bayesian optimization is adopted in this paper to minimize the cross entropy on the validation

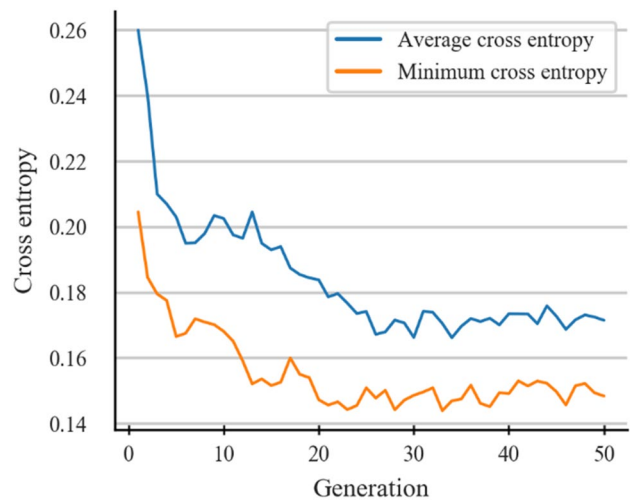


Fig. 6 Evolution of GA-based multi-sensor information selection

set. The hyper-parameters to be optimized and their bounds can be found in Table 2. Learning rate  $\mu_1$  is initialized and searched in logarithmic space.

The optimization procedure is as follows. In the beginning, five sets of random hyper-parameters  $x_1-x_5$  are generated by Latin Hypercube Sampling (LHS) [46]. LHS can form uniform sets that benefit initialization. The cross entropy of five networks on the validation set is used to formalize a prior distribution of Gaussian Process (GP),

$$L^n \sim N(0, k(x_n)), \tag{11}$$

where  $x_n$  is the  $n$ th hyper-parameters set, and  $L^n$  is the corresponding cross entropy. In this case, the mean function is zero and the covariance function  $k(x_n)$  is Matérn 5/2.

Considering  $L(x)$  is obtained from a GP prior, a posterior over acquisition function  $\alpha(x)$  is induced. Among the available acquisition functions, expected improvement (EI) has shown good performance in hyper-parameter tuning [47], which can be expressed as,

$$\alpha(x) = (L^{\text{optimal}} - \mu(x))\Phi\left(\frac{L^{\text{optimal}} - \mu(x)}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{L^{\text{optimal}} - \mu(x)}{\sigma(x)}\right), \tag{12}$$

where  $L^{\text{optimal}}$  is the optimal value currently,  $\mu(x)$  and  $\sigma(x)$  are the mean function and variance function of the posterior, respectively, and  $\Phi(\cdot)$  and  $\phi(\cdot)$  are the normal cumulative distribution function and probability density function, respectively.

Therefore, the next set is evaluated by  $x_{\text{next}} = \text{argmax}\alpha(x)$ . Once the next hyper-parameter set  $x_{\text{next}}$  is evaluated, it is then considered as new prior knowledge and a new prior GP distribution is obtained.

**Table 2** Details of the optimized hyper-parameters

Hyper-param-eters	Data type	Bounds	Optimized value
$f_{11}$	Int	[2, 15]	14
$f_{12}$	Int	[2, 15]	7
$w_1$	Int	[5, 25]	20
$p_{11}$	Int	[3, 8]	6
$p_{12}$	Int	[3, 8]	3
$f_{21}$	Int	[2, 15]	12
$f_{22}$	Int	[2, 15]	12
$w_2$	Int	[6, 40]	15
$p_{21}$	Int	[3, 7]	5
$p_{22}$	Int	[3, 7]	3
$n_F$	Int	[100, 5000]	4576
$\sigma_d$	Float	[0, 1]	0.9336
$\mu_1$	Float	$[10^{-5}, 10^{-1}]$	$10^{-3.9536}$

This processes 65 iterations for optimization. Figure 7 demonstrates the whole optimization process. The optimal set is obtained at the 27th iteration, and the cross entropy is 0.115. The optimization results are listed in Table 2. The fully connected layer is composed of 4576 neurons. The related final structural parameters of CNN are given in Table 3. The neural network has 7,614,055 parameters and 63,216 neurons in all.

## 4 Experiment and Discussion

### 4.1 Evaluation of GA-CNN

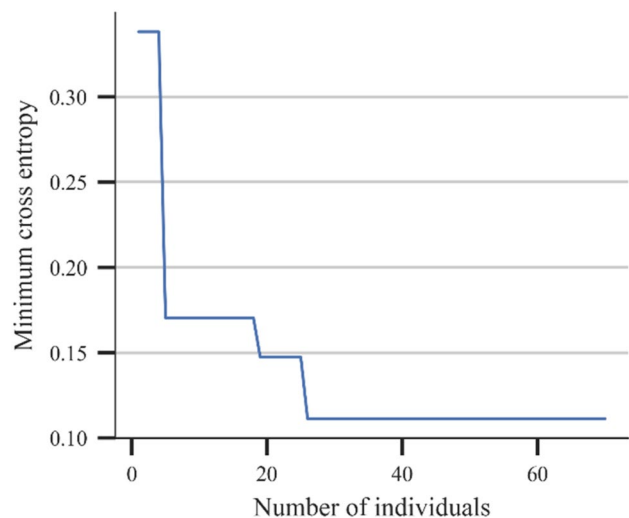
To evaluate the proposed LMR method, online experiments are conducted. Two metrics are used to measure the result of the experiments: recognition effects and recognition delay. The previously trained subjects 1<sup>#</sup>, 2<sup>#</sup>, 3<sup>#</sup> (locomotion data are trained by the network) as well as three new subjects 7<sup>#</sup>, 8<sup>#</sup>, 9<sup>#</sup> are adopted in this experiment.

#### 4.1.1 Recognition Effects

The LMR performance is evaluated in terms of the accuracy rate, which is calculated using the following equation,

$$S_k = \frac{N_s}{N_t}, \tag{13}$$

where  $k \in [1, 12]$  denotes the locomotion modes,  $N_s$  represents the number of successfully recognized data points in



**Fig. 7** The hyper-parameter optimization process



**Table 3** Parameters of the CNN architecture

Layer	Kernel	Stride	Num	Feature maps
Convolutional 1	$f_{11} \times f_{12}$	14 × 7	1	$l_{C1} \times h_{C1}$
Pooling 2	$p_{11} \times p_{12}$	6 × 3	2	$l_{L1} \times h_{L1}$
Convolutional 1	$f_{21} \times f_{22}$	12 × 12	1	$l_{C2} \times h_{C2}$
Pooling 2	$p_{21} \times p_{22}$	5 × 3	2	$l_{L2} \times h_{L2}$
Fully connected	4576			
Output	12			
Learning rate	$10^{-3.9536}$			
Dropout rate	0.9336			

**Table 4** Accuracy rates of the LMR method

	Dynamic motions					Static states			Transition states				All (%)
	LW (%)	SA (%)	SD (%)	RA (%)	RD (%)	ST (%)	SI (%)	SQ (%)	STSI (%)	SIST (%)	STSQ (%)	SQST (%)	
1 <sup>#a</sup>	99.80	98.56	98.22	98.24	99.13	100.00	97.68	96.70	96.65	97.91	96.19	96.88	98.00
2 <sup>#</sup>	99.76	97.89	98.01	98.44	99.23	100.00	97.03	97.00	96.22	97.44	96.17	96.81	97.83
3 <sup>#</sup>	99.70	98.19	98.15	98.78	99.00	100.00	97.43	96.69	96.29	97.71	96.17	96.83	97.91
7 <sup>#</sup>	96.08	95.15	93.61	95.35	94.88	100.00	89.91	88.91	90.24	90.79	90.66	89.89	92.96
8 <sup>#</sup>	98.67	97.45	96.23	97.65	97.00	100.00	92.11	91.31	93.09	93.73	93.33	93.04	95.30
9 <sup>#</sup>	96.92	95.38	94.12	95.81	95.62	100.00	90.59	89.99	91.57	91.57	91.84	90.55	93.66

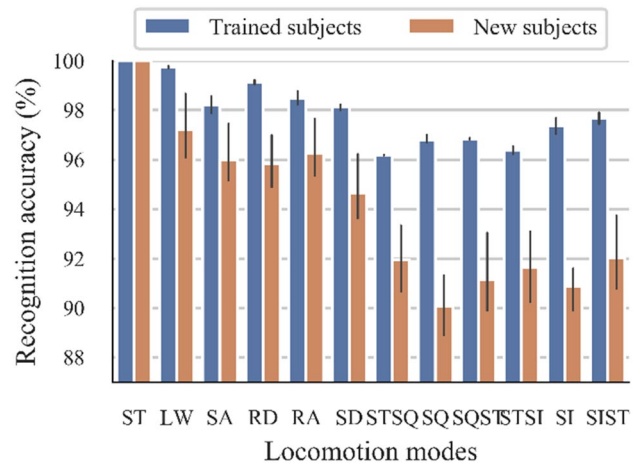
<sup>a</sup>1<sup>#</sup>, 2<sup>#</sup>, 3<sup>#</sup> is the number of trained subjects whose motion data is trained by the network, and 7<sup>#</sup>, 8<sup>#</sup>, 9<sup>#</sup> is the number of new subjects

the locomotion mode  $k$ , and  $N_k$  is the total number of sample data in the locomotion mode  $k$ .

The accuracy rates of the LMR method are listed in Table 4. The columns in the table are all locomotion modes and the rows are subjects. The average accuracy rate is 97.91% for the trained subjects and 93.97% for new subjects. As can be seen from the rightmost column, the accuracy rate of the three trained subjects are very close within the range of 97.91% ± 0.09, and the same close for each mode. The lowest rate is 96.17% (STSQ) and the highest is 99.80% (LW) except 100% for all ST. For new subjects, all rates of ST are still 100%, the second-highest rate is also in LW 98.67% and the lowest is in SQ 88.91%. In general, the accuracy of the network is great for the trained subjects, and for new subjects it is still high enough.

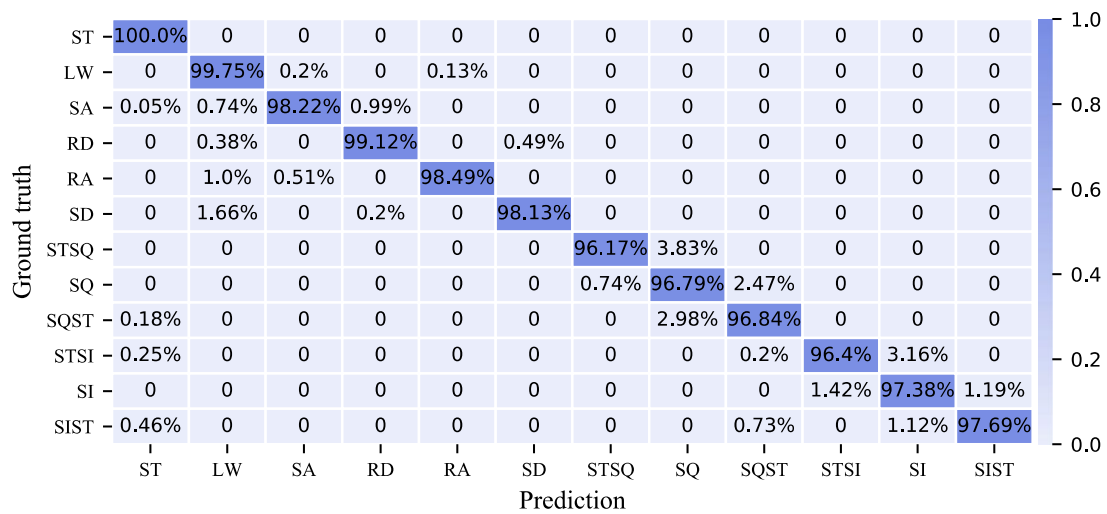
For locomotion mode, the recognition accuracy rates for the trained subjects and new subjects are shown in Fig. 8. The accuracy rate is relatively contrasting for different motion modes. The accuracy of ST still maintains 100%. Dynamic motions have an overall higher accuracy rate, compared with static states (except ST) and transitions. SQ and SI have the worst performance.

The accuracy rate is very straightforward but cannot display the classification details, and the specific recognition situation can be visually demonstrated by the confusion matrix. There are four types of prediction results in the matrix, including True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN).

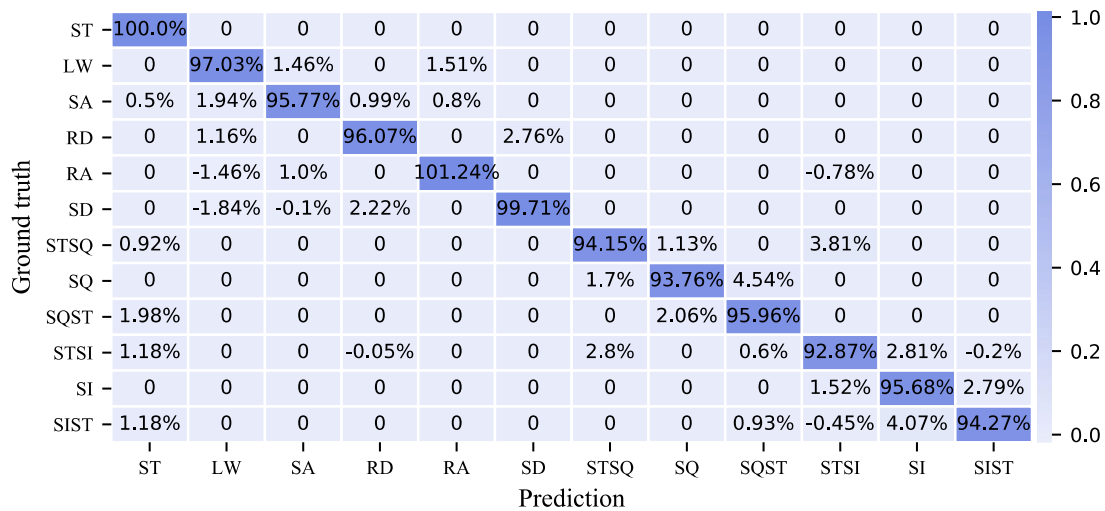


**Fig. 8** Accuracy rates of LMR for 12 locomotion modes

Figure 9 demonstrated the normalized confusion matrix for (a) the trained subjects and (b) the new subjects, respectively. The misclassification of (a) and (b) is roughly same. From the upper left corner to the lower right corner, first, mistakes are relatively easy to be made when dealing with similar locomotion modes like RA and SA, RD and SD, because the motion characteristics of the limbs are very similar when these movements are performed. Furthermore, SA, RA, SD, and RD all have the possibility to be wrongly classified as LW. Moreover, it can be



(a)



(b)

**Fig. 9** Confusion matrix of the LMR method. **a** Confusion matrix of the trained subjects. **b** Confusion matrix of the new subjects

deduced from the lower right part of Fig. 9 in the process of STSQ-SQ-SQST and STSI-SI-SIST that locomotion modes are easily wrongly classified into adjacent modes. In the middle of transition motion, the misclassification of SQ and SI is more. Additionally, there are more random misrecognitions of new subjects.

To further confirm the model's performance and avoid misleading results derived from the accuracy rates metrics, the  $F_1$ -score is adopted. Equation (14) represents the  $f_1$ -score under each motion, and Macro- $F_1$  (Eq. (15)) indicates the overall performance.

$$f_1_k = 2 \cdot \text{precision}_k \cdot \text{recall}_k / (\text{precision}_k + \text{recall}_k), \quad (14)$$

$$M-F_1 = \frac{1}{n} \sum_{k=1}^n f_1_k, \quad (15)$$

$$\text{precision}_k = \text{TP}_k / (\text{TP}_k + \text{FP}_k), \quad (16)$$

$$\text{recall}_k = \text{TP}_k / (\text{TP}_k + \text{FN}_k), \quad (17)$$

Table 5 shows the  $f_1$ -score of 12 locomotion modes and the Macro- $F_1$  of an integral model for the trained subjects and new subjects, respectively. The closer the value of the  $f_1$ -score is to 100%, the better the model behaves. The model still performs well and is stable for new subjects with 93.97% Macro- $F_1$ .

As can be seen from the accuracy rate and  $F_1$ -score that the proposed method performed well in the LMR task. It is reasonable that the performance of new subjects is not as good as the trained subjects, but it is still high enough in the application. Therefore, the proposed LMR method shows sufficient adaptation to new subjects.

### 4.1.2 Recognition Delay

As the other critical metric, the delay of the LMR process is presented in this section. The delay was defined as the time elapsed from the beginning of certain task execution to the time of its recognition, and this measure can be applied both for motions at the beginning of the classification processor and for transitions between different motions. For most gait modes recognition, the response time was defined as the time elapsed between the time when the terrain change was recognized and the critical time [12]. The critical time is defined as the moment when the leading leg touches the ground [20]. This standard is widely used and practical for recognizing LW, SA, and SD, etc., but it is not appropriate for this case because recognition in this paper includes varies modes like SI, and SQ, etc., in which there is no foot strike. In this case, the critical time adopts the exact moment when the wearer begins to execute or change the mode. The recognition delay is expressed using the following indexes,

$$D = T_r - T_c, \tag{18}$$

where  $T_r$  is the recognition moment, and  $T_c$  is the critical moment.

The delay of the LMR is given in Table 6. The average values of all possible transitions are listed in the table. For example, the delay of LW represents the average delay of ST-LW, RA-LW, RD-LW, SA-LW, and SD-LW, and the delay of ST denotes the average delay of SIST-ST and SQST-ST. In addition, the remaining only represents one transition, like SA represented LW-SA, and SQ represented STSQ-SQ. As can be seen from Table 6, the delay is generally between 10 and 60 ms with an average of 20–30 ms, and different modes vary greatly. For dynamic motions, the delay of tens of milliseconds is much shorter than the human gait period, which means that the proposed method is able to successfully recognize modes before the leading leg touches the ground. The delay of static states and transitions is generally less than dynamic motions.

The maximum recognition delay is  $29.92 \pm 5.86$  ms for the trained subjects and  $53.33 \pm 14.39$  ms for the new subjects. It is reasonable that the performance of new subjects is not as good as the trained subjects, but the recognition performance for new subjects is still satisfactory and applicable. Compared with most machine learning methods that need hundreds of milliseconds of delay, tens of milliseconds indicate good performance. The fuzzy-logic classifier proposed by Parri et al. [23] is always one step delay. The maximum delay is  $47.69\% \pm 28.77\%$  and  $17.4\% \pm 1.2\%$  of a gait cycle for [31] and [34], respectively. Assuming one gait cycle takes 500 ms,  $17.4\% \pm 1.2\%$  delay time reaches 85 ms of the deep neural networks proposed by Hua et al.

**Table 5**  $F_1$ -score of the LMR method for all locomotion modes

	Dynamic motions					Static states			Transition states				$M-F_1$
	LW	SA	SD	RA	RD	ST	SI	SQ	STSI	SIST	STSQ	SQST	
$f_1^{Ta}$	97.98%	98.75%	98.81%	99.17%	98.97%	99.53%	97.68%	95.08%	98.81%	95.08%	97.68%	96.73%	97.91%
$f_1^N$	94.85%	96.42%	95.77%	96.99%	96.51%	97.64%	91.23%	89.78%	90.90%	93.85%	91.91%	91.77%	93.97%

<sup>a</sup> $f_1^T$  refers to the average  $f_1$  of the trained subjects, and  $f_1^N$  refers to the average  $f_1$  of the new subjects

**Table 6** Delay of the LMR method

	LW	SA	SD	RA	RD	ST
$D^{Ta}$ (ms)	$22.42 \pm 3.10$	$27.78 \pm 3.12$	$22.85 \pm 2.88$	$28.89 \pm 4.01$	$29.92 \pm 5.86$	$10.22 \pm 4.87$
$D^N$ (ms)	$35.14 \pm 9.73$	$41.51 \pm 13.15$	$38.27 \pm 8.56$	$45.17 \pm 11.94$	$53.33 \pm 14.39$	$18.23 \pm 5.56$
	SI	SQ	STSI	SIST	STSQ	SQST
$D^T$ (ms)	$23.70 \pm 5.57$	$20.26 \pm 3.23$	$15.38 \pm 2.10$	$13.26 \pm 1.59$	$12.16 \pm 1.51$	$20.02 \pm 2.27$
$D^N$ (ms)	$29.20 \pm 10.27$	$25.06 \pm 9.01$	$17.88 \pm 5.55$	$25.68 \pm 6.21$	$18.07 \pm 5.61$	$35.36 \pm 9.33$

<sup>a</sup> $D^T$  refers to the average delay of the trained subjects, and  $D^N$  refers to the average delay of the new subjects

[34]. For the proposed method with a maximum delay of  $53.33 \pm 14.39$  ms delay, the real-time ability can be confidently guaranteed and can fully meet the needs of the control system.

Regarding the computational complexity, the weight sharing mechanism of CNN can reduce the number of free parameters and computational complexity. In addition, the embedded controller of the exoskeleton platform is based on TI AM437x ARM Cortex-A9, which possesses excellent computing ability. The average computational consumption to complete an online recognition via the proposed method is 5.41 ms with a standard deviation of 0.89 ms according to 10,000 experiments. Considering that the sampling period is 10 ms, the real-time capability can be guaranteed. Computational consumption has also been included in the delay time.

## 4.2 Comparative Study

The comparative experiment is reported in this section. LDA and ANN that represent statistical and neural classifiers are adopted as comparing methods, and then two variants of the proposed model are listed following as comparing objects to confirm the necessity of each presented part.

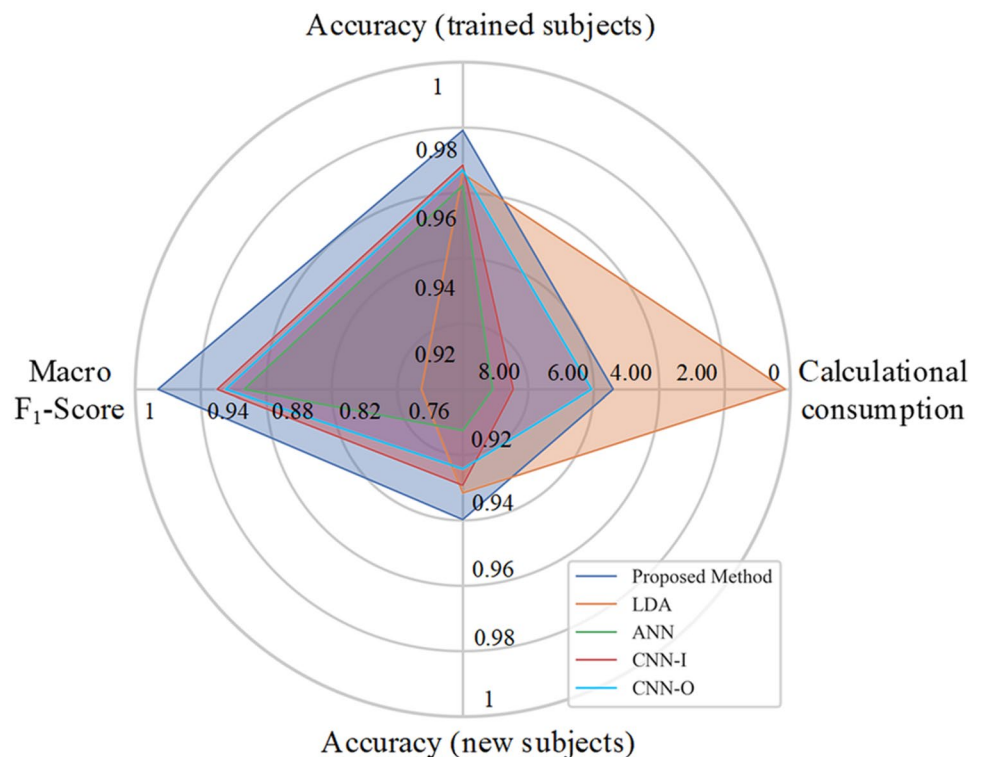
1. LDA: One LDA classifier was constructed to classify locomotion modes. While LDA is a simple classifier, it has been shown in other pattern recognition literatures

in which LDA has similar performance to non-linear and more complex classifiers [12]. The classifier is trained and tested in the same way as the proposed method.

2. ANN: In this method, a 6-layer ANN with dropout layers is adopted as a learning network. The input data is the same as the proposed method, which is the optimal data set selected by GA in the form of time-window flow. The same hyper-parameter optimization is executed in this method as well. The loss function still employs cross entropy, and the network is also optimized by Adam optimizer.
3. CNN-I: This variant employed the same architecture as the proposed method, but adopts all acquired data as input without removing any sensor channels. The purpose of establishing this comparison item is to evaluate the influence of the multi-sensor information selection process.
4. CNN-O: As a variant of the proposed method, CNN-O is implemented without auto hyper-parameter optimization, which means they are all tuned manually. Meanwhile, the other processes are the same as the proposed method.

The five methods including the proposed one are compared from four aspects: accuracy for the trained subject, Macro  $F_1$ -score, computational consumption and accuracy for the new subject. A radar plot shown in Fig. 10 makes

**Fig. 10** Radar plot of the comparative experiment



a clear comparison. Apparently, the proposed model significantly outperforms other methods in terms of accuracy and Macro  $F_1$ -score. For calculational consumption, LDA's performance is much ahead of all the other methods, which is followed by the proposed one. Thus, the proposed method has the best performance by trading all four aspects off. In addition, although with the best calculational consumption, LDA's Macro  $F_1$ -score is abnormally the lowest among all the methods, and its accuracy rate is also lower than the average level. The performance of the ANN method is also not well and remarkable on four standard axes. About CNN-I, the shape of its radar plot is very similar to the proposed method, and the area is smaller. This phenomenon is very reasonable, indicating that multi-sensor information selection has improved the algorithm in all aspects, especially for calculational consumption. The input sensor channels of CNN-O are the same as the proposed method, so the calculational consumption of them is very close. Due to the under-optimized hyper-parameters in CNN-O, its performance except for calculational consumption is mediocre. The comparative experiment verified the effectiveness of the method and the necessity of each part.

## 5 Conclusion

In this paper, an integral subject-adaptive real-time LMR method is proposed based on GA-CNN. A large and deep CNN is trained to recognize exoskeleton locomotion combined with GA multi-sensor information selection. A time-window with current and historical sensor data is employed to identify the current locomotion mode, and the appropriate length of the window is determined by balancing the cross entropy and computational consumption. An exoskeleton prototype system with multi-type sensors is established. The multi-sensor data as the input of CNN is further elaborately selected by GA. Besides, Bayesian optimization optimizes a set of best-performing hyper-parameters of the network. The proposed LMR method is capable of recognizing 12 modes that contain all common locomotion in daily life, with an average accuracy of 97.91% for the trained subjects and 93.97% for the new subjects. About the recognition delay, the maximum is  $29.92 \pm 5.86$  ms for the trained subjects and  $53.33 \pm 14.39$  ms for the new subjects. The average computational consumption for online recognition using the proposed method is  $5.41 \pm 0.89$  ms. The real-time ability can be confidently guaranteed and can fully meet the needs of the control system. The experiments validated the method possesses strong comprehensive abilities including high accuracy, low delay, low computational complexity, and sufficient adaption to different subjects. In the future,

this locomotion recognition part should be integrated with the appropriate control strategy to complete the exoskeleton functional hierarchical control scheme. It is at least equally important for a controller to solve more issues, i.e., eliminating external interference and in concert with the intentions of the wearer. In this way, the final performance of the assisted exoskeleton will be demonstrated.

## 6 Data Availability Statement

The datasets generated and analyzed during the current study are not publicly available as the data also forms part of an ongoing study but are available from the corresponding author on reasonable request.

**Acknowledgements** We would like to thank the editor and the reviewers for their useful feedback that improved this paper. The authors are grateful to all the participants for their work in the experiments, and to Dr. Yongchen Guo for his help with this work.

## Declarations

**Conflict of interest** This work was supported by the Pre-research project in the manned space field, Project number 020202, China. The authors certify that there are no conflicts relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Mo, F. H., Zhang, Q., Zhang, H. T., Long, J. J., Wang, Y. L., Chen, G., & Ye, J. (2021). A simulation-based framework with a proprioceptive musculoskeletal model for evaluating the rehabilitation exoskeleton system. *Computer Methods and Programs in Biomedicine*, 208, 106270.
2. Lerner, Z. F., Damiano, D. L., & Bulea, T. C. (2017). A lower-extremity exoskeleton improves knee extension in children with crouch gait from cerebral palsy. *Science Translational Medicine*, 9(404), 9145.
3. Huo, W., Mohammed, S., Amirat, Y., & Kong, K. (2018). Fast gait mode detection and assistive torque control of an exoskeletal robotic orthosis for walking assistance. *IEEE Transactions on Robotics*, 34(4), 1035–1052.



4. Chen, C. F., Du, Z. J., He, L., Shi, Y. J., Wang, J. Q., & Dong, W. (2021). A novel gait pattern recognition method based on LSTM-CNN for lower limb exoskeleton. *Journal of Bionic Engineering*, *18*(5), 1059–1072.
5. Zheng, T., Zhu, Y., Zhang, Z., Zhao, S., Chen, J., & Zhao, J. (2018). Parametric gait online generation of a lower-limb exoskeleton for individuals with paraplegia. *Journal of Bionic Engineering*, *15*(6), 941–949.
6. Martinez-Hernandez, U., Mahmood, I., & Dehghani-Sani, A. A. (2017). Simultaneous Bayesian recognition of locomotion and gait phases with wearable sensors. *IEEE Sensors Journal*, *18*(3), 1282–1290.
7. Wang, J. S., Lin, C. W., Yang, Y. T. C., & Ho, Y. G. (2012). Walking pattern classification and walking distance estimation algorithms using gait phase information. *IEEE Transactions on Biomedical Engineering*, *59*(10), 2884–2892.
8. Wang, Z. L., Zhao, C., & Qiu, S. (2014). A system of human vital signs monitoring and activity recognition based on body sensor network. *Sensor Review*, *34*(1), 42–50.
9. Au, S., Berniker, M., & Herr, H. (2008). Powered ankle-foot prosthesis to assist level-ground and stair-descent gaits. *Neural Networks*, *21*(4), 654–666.
10. Chen, B. J., Wang, Q. N., & Wang, L. (2014). Adaptive slope walking with a robotic transtibial prosthesis based on volitional EMG control. *IEEE/ASME Transactions on Mechatronics*, *20*(5), 2146–2157.
11. Huang, H., Kuiken, T. A., & Lipschutz, R. D. (2008). A strategy for identifying locomotion modes using surface electromyography. *IEEE Transactions on Biomedical Engineering*, *56*(1), 65–73.
12. Young, A. J., Kuiken, T. A., & Hargrove, L. J. (2014). Analysis of using EMG and mechanical sensors to enhance intent recognition in powered lower limb prostheses. *Journal of Neural Engineering*, *11*(5), 056021.
13. Kilicarslan, A., Prasad, S., Grossman, R. G., & Contreras-Vidal, J. L. (2013). High accuracy decoding of user intentions using EEG to control a lower-body exoskeleton. In *2013 35th Annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 5606–5609).
14. Bradford, J. C., Lukos, J. R., & Ferris, D. P. (2016). Electro-cortical activity distinguishes between uphill and level walking in humans. *Journal of Neurophysiology*, *115*(2), 958–966.
15. Luu, T. P., Brantley, J. A., Nakagome, S., Zhu, F., & Contreras-Vidal, J. L. (2017). Electro-cortical correlates of human level-ground, slope, and stair walking. *PLoS One*, *12*(11), e0188500.
16. Goh, S. K., Abbass, H. A., Tan, K. C., Al-Mamun, A., Thakor, N., Bezerianos, A., & Li, J. H. (2018). Spatio-spectral representation learning for electroencephalographic gait-pattern classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(9), 1858–1867.
17. Huo, W. G., Mohammed, S., Moreno, J. C., & Amirat, Y. (2014). Lower limb wearable robots for assistance and rehabilitation: A state of the art. *IEEE Systems Journal*, *10*(3), 1068–1081.
18. Goršič, M., Kamnik, R., Ambrožič, L., Vitiello, N., Lefebvre, D., Pasquini, G., & Munih, M. (2014). Online phase detection using wearable sensors for walking with a robotic prosthesis. *Sensors*, *14*(2), 2776–2794.
19. Ambrozic, L., Gorsic, M., Geeroms, J., Flynn, L., Lova, R. M., Kamnik, R., Munih, M., & Vitiello, N. (2014). CYBERLEGS: A user-oriented robotic transfemoral prosthesis with whole-body awareness control. *IEEE Robotics & Automation Magazine*, *21*(4), 82–93.
20. Yuan, K. B., Wang, Q. N., & Wang, L. (2014). Fuzzy-logic-based terrain identification with multisensor fusion for transtibial amputees. *IEEE/ASME Transactions on Mechatronics*, *20*(2), 618–630.
21. Chen, B. J., Wang, X. G., Huang, Y., Wei, K. L., & Wang, Q. N. (2015). A foot-wearable interface for locomotion mode recognition based on discrete contact force distribution. *Mechatronics*, *32*, 12–21.
22. Kong, K., & Tomizuka, M. (2009). A gait monitoring system based on air pressure sensors embedded in a shoe. *IEEE/ASME Transactions on Mechatronics*, *14*(3), 358–370.
23. Parri, A., Yuan, K. B., Marconi, D., Yan, T. F., Crea, S., Munih, M., Lova, R. M., Vitiello, N., & Wang, Q. N. (2017). Real-time hybrid locomotion mode recognition for lower limb wearable robots. *IEEE/ASME Transactions on Mechatronics*, *22*(6), 2480–2491.
24. Kale, A., Sundaresan, A., Rajagopalan, A., Cuntoor, N. P., Roy-Chowdhury, A. K., Kruger, V., & Chellappa, R. (2004). Identification of humans using gait. *IEEE Transactions on Image Processing*, *13*(9), 1163–1173.
25. Wang, L., Tan, T. N., Ning, H. Z., & Hu, W. M. (2003). Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(12), 1505–1517.
26. Hargrove, L. J., Simon, A. M., Lipschutz, R., Finucane, S. B., & Kuiken, T. A. (2013). Non-weight-bearing neural control of a powered transfemoral prosthesis. *Journal of Neuroengineering and Rehabilitation*, *10*(1), 62.
27. Mannini, A., & Sabatini, A. M. (2012). Gait phase detection and discrimination between walking-jogging activities using hidden Markov models applied to foot motion data from a gyroscope. *Gait & Posture*, *36*(4), 657–661.
28. Varol, H. A., Sup, F., & Goldfarb, M. (2009). Multiclass real-time intent recognition of a powered lower limb prosthesis. *IEEE Transactions on Biomedical Engineering*, *57*(3), 542–551.
29. Gao, S., Wang, Y. X., Fang, C. M., & Xu, L. J. (2020). A smart terrain identification technique based on electromyography, ground reaction force and machine learning for lower limb rehabilitation. *Applied Sciences*, *10*(8), 2638.
30. Hua, Y. X., Fan, J. Z., Liu, G. F., Zhang, X. H., Lai, M. Z., Li, M., Zheng, T. J., Zhang, G. A., Zhao, J., & Zhu, Y. H. (2019). A novel weight-bearing lower limb exoskeleton based on motion intention prediction and locomotion state identification. *IEEE Access*, *7*, 37620–37638.
31. Huang, Y. C., He, Z. X., Liu, Y. X., Yang, R. Y., Zhang, X. F., Chen, G., & Yi, J. G. (2019). Real-time intended knee joint motion prediction by deep-recurrent neural networks. *IEEE Sensors Journal*, *19*(23), 11503–11509.
32. Wang, C., Wu, X. Y., Ma, Y., Wu, G. Z., & Luo, Y. H. (2018). A flexible lower extremity exoskeleton robot with deep locomotion mode identification. *Complexity*, 2018.
33. Song, J. Y., Zhu, A. B., Tu, Y., Huang, H., Arif, M. A., Shen, Z. T., Zhang, X. D., & Cao, G. Z. (2020). Effects of different feature parameters of sEMG on human motion pattern recognition using multilayer perceptrons and LSTM neural networks. *Applied Sciences*, *10*(10), 3358.
34. Gautam, A., Panwar, M., Biswas, D., & Acharyya, A. (2020). MyoNet: A transfer-learning-based LRCN for lower limb movement recognition and knee joint angle prediction for remote monitoring of rehabilitation progress from sEMG. *IEEE Journal of Translational Engineering in Health and Medicine*, *8*, 1–10.
35. Park, S., Park, F. C., Choi, J., & Kim, H. (2019). EEG-based gait state and gait intention recognition using spatio-spectral convolutional neural network. In *2019 7th International winter conference on brain-computer interface (BCI)* (pp. 1–3).
36. Gao, F., Liu, G. Y., Liang, F. Y., & Liao, W. (2020). IMU-based locomotion mode identification for transtibial prostheses, orthoses, and exoskeletons. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(6), 1334–1343.
37. Kim, H., Shin, Y. J., & Kim, J. (2017). Kinematic-based locomotion mode recognition for power augmentation exoskeleton.

- International Journal of Advanced Robotic Systems*, 14(5), 1729881417730321.
38. Liu, M., Wang, D., & Huang, H. (2015). Development of an environment-aware locomotion mode recognition system for powered lower limb prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(4), 434–443.
  39. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
  40. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117.
  41. Kwak, N. S., Müller, K. R., & Lee, S. W. (2017). A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PLoS One*, 12(2), e0172578.
  42. Gao, K. Z., Cao, Z. G., Zhang, L., Chen, Z. H., Han, Y. Y., & Pan, Q. K. (2019). A review on swarm intelligence and evolutionary algorithms for solving flexible job shop scheduling problems. *IEEE/CAA Journal of Automatica Sinica*, 6(4), 904–916.
  43. Chen, C. F., Du, Z. J., He, L., Wang, J. Q., Wu, D. M., & Dong, W. (2019). Active disturbance rejection with fast terminal sliding mode control for a lower limb exoskeleton in swing phase. *IEEE Access*, 7, 343–357.
  44. Chen, C. F., Du, Z. J., He, L., Shi, Y. J., Wang, J. Q., Xu, G. Q., Zhang, Y., Wu, D. M., & Dong, W. (2019). Development and hybrid control of an electrically actuated lower limb exoskeleton for motion assistance. *IEEE Access*, 7, 107–122.
  45. Howell, A. M., Kobayashi, T., Hayes, H. A., Foreman, K. B., & Bamberg, S. J. M. (2013). Kinetic gait analysis using a low-cost insole. *IEEE Transactions on Biomedical Engineering*, 60(12), 3284–3290.
  46. Muro-De-La-Herran, A., Garcia-Zapirain, B., & Mendez-Zorrilla, A. (2014). Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors*, 14(2), 3362–3394.
  47. Shahriari, B., Swersky, K., Wang, Z. Y., Adams, R. P., & Freitas, N. D. (2015). Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1), 148–175.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.