Continuous Estimation of Human Knee Joint Angles by Fusing Kinematic and Myoelectric Signals

Ning Sun, Menglin Cao, Yuhao Chen, Yanni Chen, Jue Wang, *Member, IEEE*, Qiong Wang, Xiaocong Chen, and Tian Liu

***Abstract*— Exoskeleton robot is an essential tool in active rehabilitation training for patients with lower limb motor dysfunctions. Accurate and real-time recognition in human motion intention is a great challenge in exoskeleton robot, which can be implemented by continues estimation of human joint angles. In this study, we innovatively proposed a novel feature-based convolutional neural network-bi-directional long-short term memory network (CNN-BiLSTM) model to predict the knee joint angles more accurately and in real time. We validated our method on a public dataset, including surface electromyography(sEMG) and inertial measurement unit (IMU) data of 10 healthy subjects during normal walking. Initially, features extraction from each modal data achieved feature-level fusion. Then the importance of each sEMG and IMU signal feature for knee joint angle prediction was quantified by ensemble feature scorer (EFS) and the number of features required for prediction while ensuring accuracy was simplified by profile likelihood maximization (PLM) algorithm. Finally, the CNN-BiLSTM model was created by using the determined simplest features to further fuse the spatio-temporal correlation of signals. The results indicated that the EFS and PLM algorithm could remove the feature redundancy perfectly and estimation performance would become better when bi-modal gait data were fused. For the estimation performance, the average root mean square error (RMSE), adjusted R^2 and pearson correlation coefficient (CC) of our algorithm were 4.07, 0.95, and 0.98, respectively, which was better than CNN, BiLSTM and other three traditional machine learning methods. In addition, the model test time was 62.47 ± 0.29 ms, which was less than the predicted horizon of 100 ms. The real-time performance and accuracy are satisfactory. Compared with previous works, our method has great advantages in feature selection and model design, which further improves the prediction accuracy. These promising results demonstrate that the proposed method has considerable potential to be applied to exoskeleton robot control.**

***Index Terms*— Surface electromyography (sEMG), inertial measurement unit (IMU), ensemble feature scorer (EFS), CNN-BiLSTM, prediction.**

I. Introduction

FOR individuals suffering from spinal cord injury, children with developmental disabilities, and adults with neuromotor impairment, it is increasingly crucial to identify lower limb activities in order to implement active training in the rehabilitation process [1], [2], [3]. Exoskeleton robots instead of artificial rehabilitation exercising have achieved initial results [4], [5]. The key technology of exoskeleton robot is to realize human-like control, which requires the ability to recognize human motion intention in real time and formulates reasonable control strategies. The commonly used control signal sources are biomechanical signals and bioelectric signals. The biomechanical signals obviously lag behind the human movement, which cannot realize the regular and natural human-computer interaction. Surface electromyography (sEMG) is generated by the contraction of the muscles on the surface of the human body. It is directly related to the movement of the limbs and can be collected from the superficial muscles through electrodes. Owing to the rich kinematic information and mature non-invasive acquisition technology of sEMG, it has been widely used in human motion intention recognition in recent years [6], [7].

Many researchers focused on the classification problem of gait patterns, such as literatures [8], [9], [10], [11]. They employed different classification algorithms to automatically identify pathological or normal gait patterns based on myoelectric, kinematic, or kinetic data. Although these studies had excellent classification results, discrete motion classification is essentially a qualitative analysis method. It is the lack of ability to monitor continuous changes in movement. When the complex motion types appear, they will not recognize the human's intention. The discrete movements classification have limited application prospects [12]. In recent years, several studies have designed new methods for joint trajectory

Manuscript received 4 June 2022; revised 26 July 2022; accepted 7 August 2022. Date of publication 22 August 2022; date of current version 1 September 2022. This work was supported in part by the National Key Research and Development Program Project under Grant 2021YFC2400203 and in part by the National Natural Science Foundation of China under Grant 61503295. *(Ning Sun and Menglin Cao contributed equally to this work.) (Corresponding author: Tian Liu.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Northwestern University Institutional Review Board and performed in line with the Declaration of Helsinki.

Ning Sun, Menglin Cao, Yuhao Chen, Jue Wang, and Tian Liu are with the Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Institute of Health and Rehabilitation Science, Xi'an Jiaotong University, Xi'an 710049, China, also with the National Engineering Research Center for Healthcare Devices, Guangzhou 510500, China, and also with the Key Laboratory of Neuro-Informatics and Rehabilitation Engineering of Ministry of Civil Affairs, Xi'an 710049, China (e-mail: ningsun@stu.xjtu.edu.cn; menglincao@stu.xjtu.edu.cn; andychen@stu.xjtu.edu.cn; juewang@mail.xjtu.edu.cn; tianliu@xjtu.edu.cn).

Yanni Chen, Qiong Wang, and Xiaocong Chen are with the Xi'an Children's Hospital, Xi'an 710049, China (e-mail: chenyannichil@163.com; 2270817519@qq.com; 684455975@qq.com).

Digital Object Identifier 10.1109/TNSRE.2022.3200485

estimation [13], [14]. However, joint kinematic regression based on sEMG is more valuable, especially in the field of rehabilitation robots. The realization of human-computer interaction based on sEMG signal will help to improve the comprehensive performance of rehabilitation robots, thereby improving the life quality of patients with lower limb movement disorders. Some researchers tried to use machine learning methods to estimate the joint angles of lower limbs based on sEMG signal [15], [16]. However, these studies usually reduced the feature redundancy phenomenon by principal component analysis (PCA), which cannot better retain the original information of the data and leads the lower prediction accuracy.

Deep learning algorithms are widely used in various classification and regression tasks because of the ability to obtain the high-level abstract features from large-scale data. Liu *et al.* [17] proposed to use CNN to realize the knee joint angles prediction. They proved that feature-based CNN had higher estimation accuracy than original data-based CNN. A large number of recent studies have shown that the LSTM has achieved good results in terms of movement intention [18], [19]. Therefore, to fully consider the spatio-temporal correlation of gait data, some researchers used the CNN-LSTM model for gait pattern recognition [20], [21], [22]. Among them, LSTM extracted the temporal features of the data, and CNN extracted the spatial features of the data. But these studies were usually calculated based on only a single modal data source, which cannot fully reflect the human's movement intention. In addition, the normal LSTM only considers the information in the previous moment, but does not consider the information of the future moment. At present, the regression of lower limb joint dynamics based on deep learning is still immature, and the prediction accuracy needs to be improved.

For addressing the limitations of the above methods, we proposed the CNN-BiLSTM model with ensemble feature scorer (EFS) based on IMU and sEMG data. Initially, features extraction from each modal data achieved feature-level fusion. Then the importance of each sEMG and IMU signal feature for knee joint angle prediction was quantified by EFS and the number of features required for prediction was simplified by profile likelihood maximization (PLM) algorithm. Finally, the CNN-BiLSTM which can fully learn the spatial and temporal features of signals was created for knee angle prediction. Compared with the CNN, BiLSTM and other traditional machine learning algorithms for prediction, our model had a greater advantage in prediction accuracy. It will provide a theoretical basis for the efficient use of sEMG as a stable and good signal source to control the lower limbs exoskeleton robots.

The main contributions of this study include:

1) We used the fused feature signal of sEMG and IMU as the input of the model. The fusion signals have highly feature complementary and avoid the limitations of single mode data. Most previous studies have focused on a single sEMG signal or IMU signal.

2)We innovatively used the feature selection method combining PLM and EFS to better remove feature redundancy. The importance of each sEMG and IMU signal feature for knee joint angle prediction was quantified by EFS and the number of features required for prediction was simplified by PLM algorithm.

3) CNN-BiLSTM model was built to fully consider the spatio-temporal correlation of gait data. Compared with the CNN, BiLSTM and other three traditional machine learning algorithms for prediction, our model had a greater advantage in prediction accuracy.

TABLE I
THE CLINICAL INFORMATION OF 10 HEALTHY SUBJECTS

Participants	Height (cm)	Weight (kg)	Age
Subject 1	193	77	26
Subject 2	181	75	28
Subject 3	163	54	24
Subject 4	185	87	25
Subject 5	178	66	24
Subject 6	160	54	23
Subject 7	173	54	26
Subject 8	170	75	23
Subject 9	185	91	27
Subject 10	160	61	29
Average ± std	174.8 ± 11.5	69.4 ± 13.7	25.5 ± 2.1

II. METHOD

A block diagram of the proposed method is illustrated in Fig. 1. We conducted feature selection with EFS and PLM on the preprocessed kinematic data and sEMG, obtaining more valuable features used to train the CNN-BiLSTM model. The trained model was used to predict knee joint angle.

A. Dataset

To validate the proposed framework, we employed a public Benchmark Dataset [23]. The dataset consists of kinematic data and sEMG collected from 10 healthy subjects in the task of walking on flat ground. Specific information of subjects is shown in Table I. Fourteen channels of sEMG signals were collected at 1000 Hz by a bipolar surface electrodes (DE2.1; Delsys, Boston, MA, USA). The monitored muscles involved the tibialis anterior (TA), medial gastrocnemius (MG), soleus (SOL), vastus lateralis (VL), rectus femoris (RF), biceps femoris (BF), and semitendinosus (ST) of both lower limbs. At the same time, inertial measurement units (IMUs) were placed bilaterally on the subjects' thigh (below rectus femoris) and shank (next to anterior tibial) and sampled at 500 Hz ((MPU-9250; Invensense, San Jose, CA, USA), which can record acceleration (ACC) and gyroscope (GYRO) in three axes. In this study, the kinematic data were resampled to 1000 Hz. Joint kinematic signals (sagittal plane only) were recorded using electrogoniometers (SG150; Biometrics Ltd., Newport, UK) placed on the knee. The data with first 120 segments of walking in every subject were selected as the experimental dataset. The description of the dataset is detailed in the study [23].

B. Data Preprocessing and Feature Extraction

sEMG is mostly concentrated in the frequency range of 20∼500 Hz. So raw sEMG data were first bandpass filtered (20∼500 Hz) through a 6th order butterworth filter, then processed by 6th order butterworth notch filter (f_c = 60, 180, 300 Hz) to remove motion artifacts and environmental

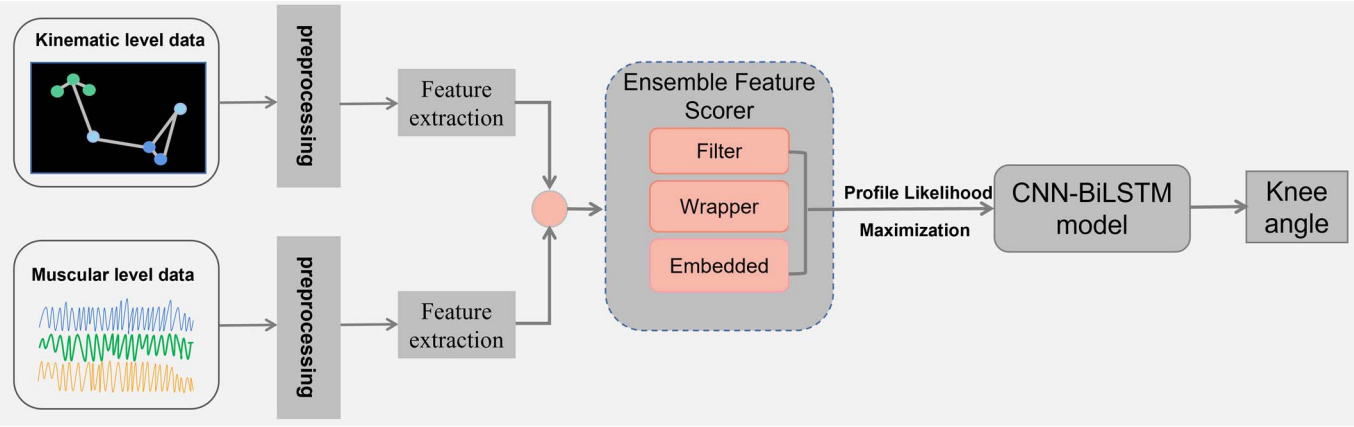

Fig. 1. Block diagram of the proposed method for knee angle prediction.

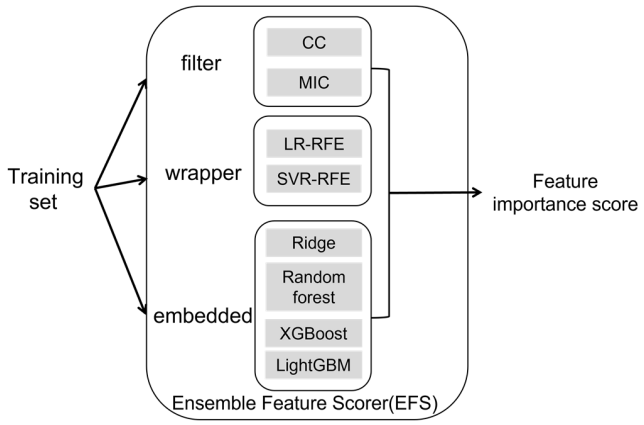

Fig. 2. Proposed ensemble feature scorer (EFS) structure.

interference. The IMU data were low-pass filtered at 20 Hz to remove high-frequency noise. In this way, the data can meet the research requirements.

Consulting the previous research theories [24], [25], we extracted seven time-domain features and one time-frequency domain feature. Time domain features included mean absolute value (MAV), root mean square (RMS), slope sign change (SSC), wave length (WL), log variance (Logvar), willison amplitude (WAMP), and sEMG Profile (PF). Time-frequency domain feature was the mean absolute value of the DB7 wavelet decomposition coefficient (DB7-MAV). For IMU signals, supposed that three-axis ACC and GYRO signals as a_x, a_y, a_z, b_x, b_y, b_z, respectively, and the instantaneous resultant ACC and GYRO signals can be computed according to (1) and (2). Then we extracted the maximum value (max), the minimum value (min), the mean value (mean) and the standard deviation (std) as the IMU features. Specifically, the above all used the moving average method with 100 ms window length and 10ms sliding step to extract features.

$$ACC = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{1}$$

$$GYRO = \sqrt{b_x^2 + b_y^2 + b_z^2} \tag{2}$$

C. Ensemble Feature Scorer (EFS)

In this paper, we constructed an EFS as shown in Fig. 2. It integrated three feature selection methods to score the extracted features' importance, and then profile likelihood maximization (PLM) [26] simplified the feature dimension required for prediction while ensuring accuracy.

1) Filter Method: We employed pearson correlation coefficient (CC) and maximal information coefficient (MIC) to test the dependence between two variables. The value is between 0 and 1, and 1 means the feature has a strong correlation with the result. The feature scoring steps of the filtering method are as follows. For the given regression problem, the target sequence is T, and the feature sequence is $F = [f_1, f_2, f_3, \ldots, f_k]$, where f_j is the j^{th} feature. The influence of the feature on the result can be obtained by calculating $CC\ (f_j, T)$ and $MIC\ (f_j, T)$ [27].

$$CC\left(f_j, T\right) = \frac{cov(f_j, T)}{\sqrt{var(f_j) * var(T)}} \tag{3}$$

$$I\left(f_j, T\right) = \sum\nolimits_{f_j, T} p\left(f_j, T\right) \log_2 \frac{p\left(f_j, T\right)}{p\left(f_j\right) p\left(T\right)} \tag{4}$$

$$MIC\left(f_j, T\right) = \frac{I\left(f_j, T\right)}{\log_2(\min(|f_j|, |T|))} \tag{5}$$

where $I(f_j, T)$ is mutual information between f_j and T; $\mathrm{cov}(f_j, T)$ is covariance between feature f_j and target sequence T; $\mathrm{var}(f_j)$ is variance of feature f_j; $\mathrm{var}(T)$ is variance of target sequence T; $p(f_j)$ is the mass probability; $p(f_j, T)$ is the approximate probability density distribution.

2) Wrapper Method: For wrapper method, we used support vector regression (SVR) and linear regression (LR) based on recursive feature elimination (RFE) to select features. Feature score process was divided into three steps. Firstly, SVR or LR was trained on a training dataset. Then features were ranked using weights from model results and the feature with the smallest weight will be removed. Finally, the process was repeated on the training set until the feature set was empty.

3) Embedded Method: The embedded method is to complete the feature selection during the model training process. We used machine learning algorithms such as ridge regression, random forest (RF), eXtreme gradient boosting (XGBoost), and LightGBM to obtain the weight of each feature.

In ridge regression, the coefficients of some features were forced to be reduced to zero in the training process by adding regularization, and only important features were retained. For

given feature vectors X, the optimization objective was defined as:

$$\min_{\omega} ||\omega X - y||_2^2 + \alpha||\omega||_2^2 \tag{6}$$

where $\alpha \geq 0$ is the complexity coefficient. We took the default value of 0.058 for α.

RF achieves features selection by calculating the average reduction of impurity. From the origin training data set, the bootstrap method was used to randomly choose a certain number of samples. For each sample, features were extracted randomly at each node. Then a regression tree was built, and the top features among them were selected for node splitting. Finally, the importance of the features has recorded on each node.

XGBoost computes feature importance score via information gain. The larger the gain, the more important the feature is [28]. LightBGM is an optimization algorithm for XGBboost which greatly reduces storage space and computational cost through histogram-based decision tree algorithm. The sum information gains of all nodes are the importance of the feature.

Our EFS integrated all the above feature selection methods. The process can be summarized as the following steps. Firstly, we used all of the above methods to obtain feature scores based on a single subject. The feature scores were then normalized to [0, 1]. Finally, the average value of the feature importance based on all subjects was calculated as the final feature score. The grades of all features were sorted in descending order to facilitate the subsequent calculation of the inflection point.

4) Profile Likelihood Maximization (PLM): After the features were scored by EFS, we used the PLM to determine the simplest features. PLM was first proposed by Zhu and Ghodsi [26] when using principal component analysis for data dimensionality reduction. Next, we will introduce this method in details.

Given the data points $Q = \{q_1, q_2, \ldots, q_n\}$, the purpose of PLM is to assume that $1 \leq k \leq n$ exists, such that $Q_1 = \{q_1, q_2, \ldots, q_k\}$ and $Q_2 = \{q_{k+1}, q_{k+2}, \ldots, q_n\}$ fit two different probability density distribution functions $f\ (q, \theta_1)$ and $f\ (q, \theta_2)$. Based on this assumption, the algorithm can obtain the value of k that maximizes the log-likelihood function:

$$L(k, \theta_1, \theta_2) = \sum_{i=1}^{k} \log f(q;\theta_1) + \sum_{j=k+1}^{n} \log f\ (q;\theta_2) \tag{7}$$

The maximum likelihood estimates of θ_1 and θ_2 can be derived from data point sets Q_1 and Q_2, respectively, as a function of k. Thus, a profile likelihood for k can be calculated as:

$$L_k(k) = \sum_{i=1}^{k} \log f(q; \hat{\theta}_1(q)) + \sum_{j=k+1}^{n} \log f\left(q; \hat{\theta}_2(q)\right) \tag{8}$$

We assumed that the two datasets Q_1 and Q_2 belong to different gaussian models, and obtained the optimal value of k by calculating the maximum profile likelihood estimation:

$$\hat{k} = \underset{d=1,2,3,\ldots,n}{argmix}\ L_k(d) \tag{9}$$

D. CNN-BiLSTM Model Structure

1) CNN Module: CNN is a kind of deep neural network, which is mainly composed of convolution layer, pooling layer and full connection layer. Among them, convolution layer is the most important part of CNN. Each convolution layer has multiple convolution kernels, which performs convolution operations on the input information to capture the hidden features of the data and then forms a feature map. Finally, the output is completed by the nonlinear activation function. The output of convolution layer is as follows:

$$h_i = f\ (\omega_i * x_i + b_i) \tag{10}$$

where x_i denotes the input of convolution layer, h_i is the i^{th} output feature map, ω_i is a weight matrix, b_i is the bias vector, and $f\ (\cdot)$ represents the activation function. The rectified linear unit (ReLU) function is widely chosen as the activation function. The formula of ReLU can be expressed as:

$$h_i = f\ (c_i) = \max\ (0, c_i) \tag{11}$$

where c_i is the results of convolutional operations.

One-dimensional CNN (1D CNN) is essentially the same as CNN, but its convolution kernel is one-dimensional, which can well find the spatial correlation between one-dimensional sequence and convolution kernel. In recent years, it has been gradually applied to the field of sequential signal processing.

2) LSTM Module: The LSTM is good at processing time series data and has been proved to be suitable for regression problems in recent years. There are three gate units in each LSTM units, called forget gate, memory gate and output gate, respectively. These gates protect and control the state of a cell, then complete computing in four steps. First, the forget gate combines the state vector h_{t-1} of the previous hidden layer and the input vector x_t at the current moment, and determines the forgotten information through the sigmoid function:

$$f_t = \sigma\ \left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{12}$$

Second, the memory gate and tanh function determine the information to be memorized, and obtain the candidate cell state vector $\tilde{C}$:

$$i_t = \sigma\ \left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{13}$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \tag{14}$$

Then, the cell state vector C_t at the current moment is calculated as:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{15}$$

where f_t is the forget gate vector. i_t is the memory gate vector. Finally, the hidden layer state vector h_t and the output gate vector o_t at the current moment are calculated as:

$$o_t = \sigma\ (W_0\ [h_t, x_t] + b_0) \tag{16}$$

$$h_t = o_t * \tanh(C_t) \tag{17}$$

However, the normal LSTM only considers the information in the previous moment, but does not consider the information of the future moment. Therefore, the normal LSTM has limitations when dealing with complex tasks [29]. We employed

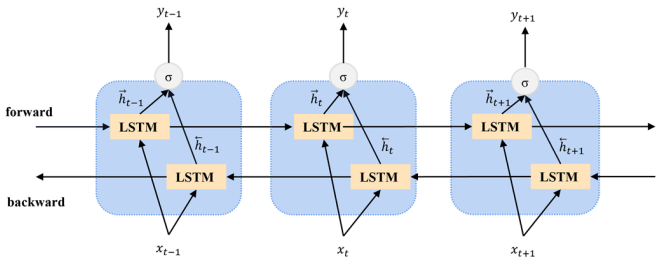

Fig. 3. Functional elements of the BiLSTM. The model includes both forward and backward timesteps.

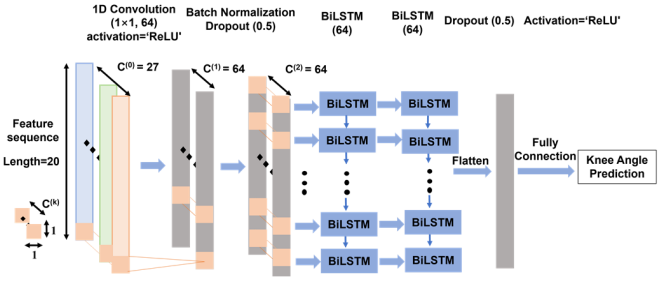

Fig. 4. Proposed CNN-BiLSTM model architecture.

the bidirectional long short-term memory network (BiLSTM), which consists of forward and backward LSTM networks. Fig. 3 shows the structure of the expansion in BiLSTM network model along the time axis at t-1, t, t+1. x_t is the model input and y_t is the output. Bidirectional architecture could extract the temporal correlations of signals from both directions at the same time with forward hidden layers and backward hidden layers. In Fig. 3, $\vec{h}_t$ and $\overleftarrow{h}_t$ are the hidden layer state of the forward LSTM network and the hidden layer state of the backward LSTM at time t, respectively. They are calculated by the standard LSTM. The BiLSTM layer yields the output at time t, which can be computed according to (18)

$$y_t = \sigma\left(\vec{h}_t, \overleftarrow{h}_t\right) \tag{18}$$

where y_t is the output of BiLSTM, σ is the tanh activation function.

3) CNN-BiLSTM: Considering the spatio-temporal correlation of gait data, we proposed a novel CNN-BiLSTM model for spatio-temporal modeling of kinematic data and sEMG to achieve knee joint angles prediction. A graphical representation of the model architecture was provided in Fig. 4. The model constructed in this paper mainly consisted of two parts: 1D CNN network and BiLSTM network.

Firstly, after EFS feature selection, we combined the fused signal features extracted from 20 consecutive sliding time windows into a feature time series X as the input of the deep learning model. The input X can be expressed by (19). 1-D convolutional operators slid 64 filters with the same window size of 1 over input sequences to capture the low-level implicit features from the raw fused feature sequences. Then, a batch normalization layer (dropout = 0.5) was utilized to prevent the overfitting of the CNN. ReLU was used as the nonlinear activation function of this layer. Next, the feature sequences output by CNN were respectively input into the BiLSTM layers. Each BiLSTM with 64 cell units not only learned the knowledge from the preceding term of current time point but also obtained the knowledge from succeeding term. The features extracted from the BiLSTM were fed to the last fully connected layer whose activation function was ReLU. Finally, the knee joint angle prediction value was output from the fully connected layer.

$$X = \begin{bmatrix} x_1^1 & x_1^2 & \dots & x_1^W \\ x_2^1 & x_2^2 & \dots & x_2^W \\ \vdots & \vdots & \vdots & \vdots \\ x_L^1 & x_L^2 & \dots & x_L^W \end{bmatrix} \tag{19}$$

where the set $x_L = (x_L^1, x_L^2, \dots, x_L^W)$ is the attribute values of W features at the Lth sliding time window. In this paper, $L = 20$, $W = 27$. The row vector represents the 27 sEMG and IMU fused feature values in the optimal feature subset, and the column vector represents the temporal feature sequence composed of fused feature values extracted from 20 consecutive sliding time windows.

E. Dataset Partitioning Strategy

The training strategy of the multi-modal framework can be summarized in the following steps. Take a single subject as example, the bimodal data of level walking was divided into the training and test set (training set: test set = 5:1). Then the training set was utilized to build EFS through ten-fold cross-validation. In order to obtain a reliable and stable model, the training set needed to be split into two non-overlapping parts (90% for training and 10% for validation) for constructing CNN-BiLSTM. Finally, the test set was used to test the model performance.

F. Evaluation

The evaluation module mainly had three parts. On the one hand, one-way analysis of variance (one-way ANOVA) was used to evaluate the effect of EFS feature selection. It was designed to test whether there was a significant difference in the mean prediction results before and after feature selection.

On the other hand, pearson correlation coefficient (CC), root mean square error (RMSE) and adjusted R^2 was used to evaluate the model. RMSE can measure the deviation between predicted angles and real angles while CC can measure the similarity between signals. For eliminating the interference of the sample size on the results, we use adjusted R^2 as a supplement. The calculation formulas are shown as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum(\theta_{pre} - \theta_{real})^2} \tag{20}$$

$$CC = \frac{cov(\theta_{pre}, \theta_{real})}{\sigma_{pre}\sigma_{real}} \tag{21}$$

$$R^2 = 1 - \frac{\sum(\theta_{pre} - \theta_{real})^2}{\sum(\theta_{real} - mean(\theta_{real}))} \tag{22}$$

$$adjusted\, R^2 = 1 - \frac{\left(1 - R^2\right)(N-1)}{N - p - 1} \tag{23}$$

where θ_{pre} is the predicted angle, while θ_{real} is the real angle. p is the number of features. N is the number of samples.

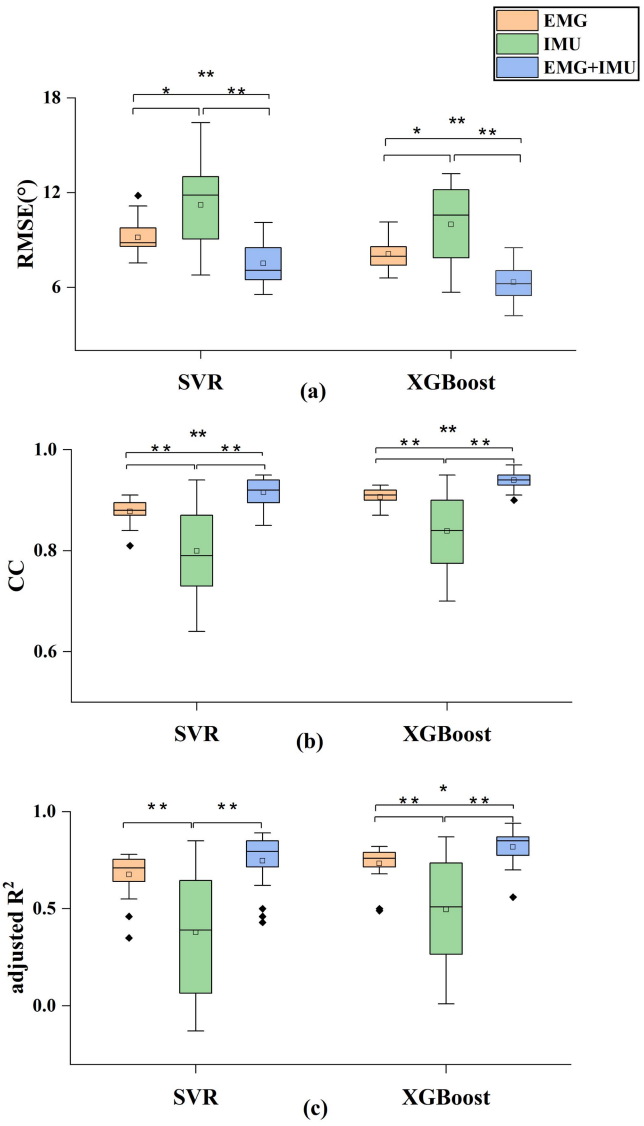

Fig. 5. Prediction effect of sEMG signal features, IMU signal features and fusion signal feature as model input, respectively: (a): RMSE; (b): CC; (c) adjusted R^2.

Furthermore, we designed a python program to calculate the model prediction time based on the trained CNN-BiLSTM model to preliminarily evaluate the real-time performance of the algorithm. If the algorithm delay time is less than the time to predict the knee joint angles in advance (100 ms), it means that the real-time performance is better.

III. Results

A. Influence of Fusion Features on Prediction Results

To study the effect of fusion features on the prediction accuracy, we took the selected sEMG features alone, kinematic signals features alone and fusion features respectively as the input of SVR and XGBoost model, and mapped to the knee joint angles at 100 ms after the end of the time window. For performance evaluation, we adopted CC, RMSE and adjusted R^2 value as evaluation metrics. The average value of each metric after ten-fold cross-validation was used as the final calculation result. The prediction performance of the two

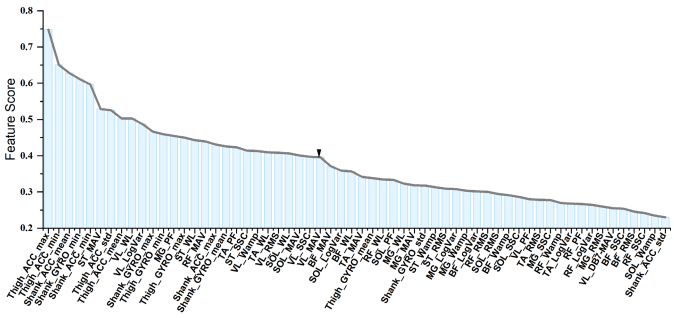

Fig. 6. Weighted average feature score curve.

TABLE II
KINEMATIC FEATURES SELECTED BY EFS

	Features
Thigh	ACCmax, ACCmin, ACCmean, ACCstd GYROmax, GYROmin
Shank	ACCmax, ACCmin, ACCmean GYROmax, GYROmin, GYROmean

models with the selected three feature types is described in Fig. 5. Fig. 5 (a)-(c) show the RMSE, CC, and adjusted R^2 value corresponding to three different inputs, respectively. It can be seen that no matter SVR model or XGBoost, the results of the fusion signals are better than the single EMG and IMU signal. The * in the figure represents one-way variance analysis $P < 0.05$, and ** represents one-way variance analysis $P < 0.01$, which both have statistically significant differences. Therefore, we chosen the fusion features of sEMG and IMU signals as the input of the model.

B. The Result of EFS

The high dimensional data probably deteriorate the generalization performance of the algorithms when the number of experimental samples is not very large [30]. To this end, the EFS was used for each modal data to reduce redundant information and extract effective signal features. After creating the EFS for each subject, we add and average the scores to get a list of weighted average scores based on all subjects. Fig. 6 intuitively reflects the feature importance scores based on all subjects. The horizontal axis is most sEMG and IMU signal feature types, and the ordinate is the importance scores of the features. As is shown in Fig. 6, we arrange the feature importance scores in descending order to form a screen-plot curve, then the inflection point is obtained by PLM method. The first inflection point is at the 27th feature, which is located at the arrow position in Fig. 6 and so we determined 27 best and simplest features for knee angles prediction finally.

The muscles and features used for the prediction of knee joint angles of human lower limbs are shown in Table II and Table III, including 12 kinematic features and 15 sEMG signal features. It can be seen that the most muscles are extracted the WL and MAV features of sEMG, and the muscles with the most extracted features is vastus lateralis. The thigh and shank have the same kinematic features, including the maximum, minimum and average values of acceleration and the maximum and minimum values of angular velocity.

TABLE III
sEMG FEATURES SELECTED BY EFS

Lower extremity muscles	Features
Vastus Lateralis（VL）	MAV, WL, Wamp, Logvar, RMS, SSC
Semitendinosus（ST）	MAV, WL, SSC
Soleus（SOL）	MAV, WL
Tibialis Anterior（TA）	MAV, PF
Medil Gastrocnemius（MG）	PF
Rectus Femoris (RF)	MAV

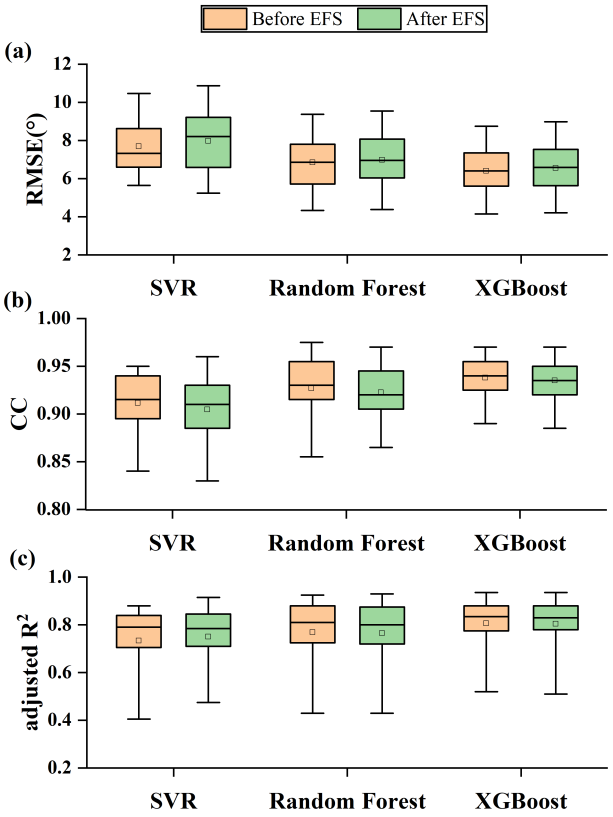

Fig. 7. Prediction effect of sEMG signal features, IMU signal features and fusion signal feature as model input, respectively: (a): RMSE; (b): CC; (c) adjusted R^2.

C. Influence of EFS on Prediction Results

To illustrate the effectiveness of EFS, we used the fused features before and after feature screening as the input of SVR, XGBoost and random forest models to forecast knee joint angles. The evaluation metrics are the same as in Part A. We used one-way ANOVA to analyze and compare the evaluation results of different models. The results are shown in Fig. 7. There is no statistical difference between the prediction results before and after feature screening ($P>0.05$), indicating that EFS reduced the input feature dimension while ensuring that the prediction accuracy does not decrease.

D. The Performance of CNN-BiLSTM

Table IV shows the results of knee joint angles prediction based on different models. Models were evaluated using RMSE, adjusted R^2 and CC on a separate test set. As is shown in Table IV, compared with other methods, the prediction performance of the CNN-BiLSTM model is the best. The

TABLE IV
RESULTS OF JOINT ANGLE PREDICTION BASED ON DIFFERENT MODELS

Method	RMSE (°)	adjusted R^2	CC
SVR	11.67	0.64	0.87
XGBoost	9.40	0.64	0.88
ELM	9.16	0.69	0.88
CNN	5.06	0.91	0.96
BiLSTM	4.41	0.94	0.97
CNN-BiLSTM	**4.07**	**0.95**	**0.98**

ELM: Extreme Learning Machine algorithm

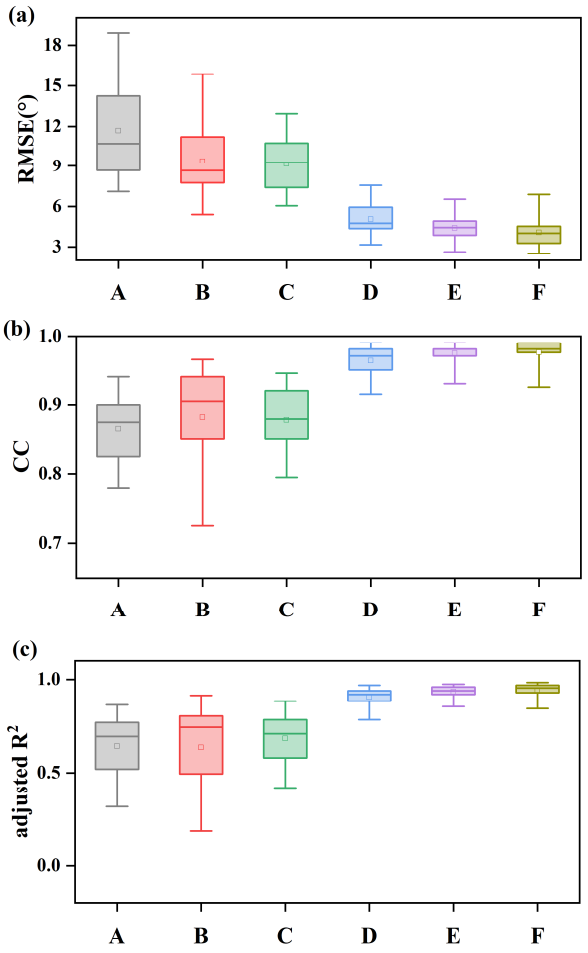

Fig. 8. Comparison of prediction performance based on different models; A: SVR, B: XGBoost, C: ELM, D: CNN, E: BiLSTM, F: CNN-BiLSTM.

average of RMSE, adjusted R^2 and CC for all 10 healthy subjects are 4.07, 0.95 and 0.98. Fig. 8 shows the comparison of prediction results in different models. (a)-(c) represent the RMSE, CC and adjusted R^2 of different models, respectively. The prediction results of BiLSTM are close to our proposed model, but our model has lower RMSE and higher adjusted R^2.

Further, we selected a segment random data in 001∼010 subject as the test set to visualize the performance of multiple models. Due to the performance of traditional machine learning methods is worse, we do not visualize here. Fig. 9 shows the visual comparison in 001∼005 between the actual knee joint angles and the predicted knee joint angles based on the

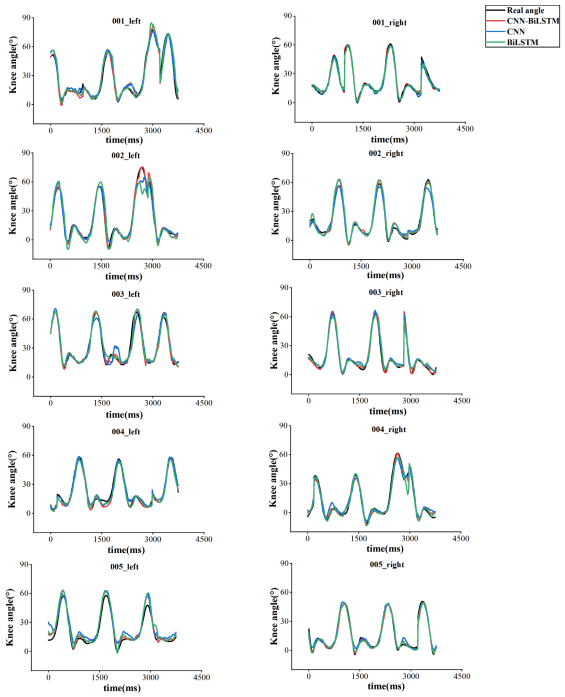

Fig. 9. Visual comparison of prediction results of 001∼005 based on CNN-BiLSTM, CNN and BiLSTM.

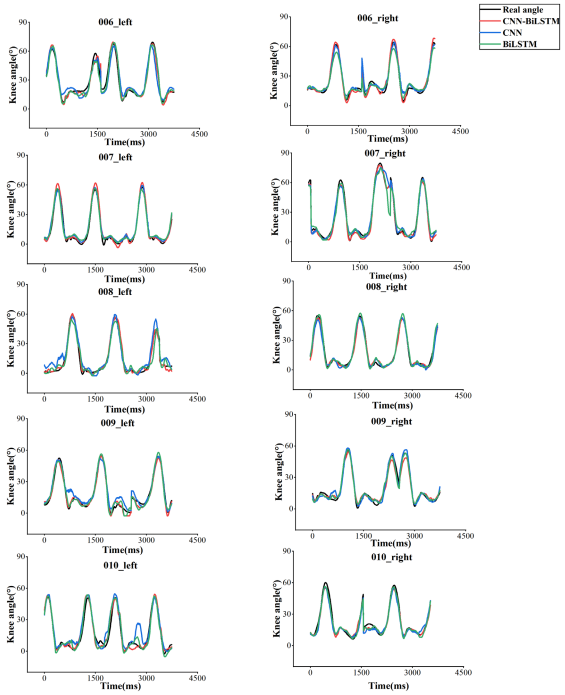

Fig. 10. Visual comparison of prediction results of 006∼010 based on CNN-BiLSTM, CNN and BiLSTM.

CNN-BiLSTM, BiLSTM and CNN model. The visualization results of other subjects are in Fig.10. The actual knee angle is represented by the black curve, and the predicted knee angles of CNN-BiLSTM, BiLSTM and CNN are represented by the red curve, green curve and blue curve respectively. The CNN model has the worst prediction effect, and the BiLSTM and our proposed model have better prediction effect. A closer look at the 002 left leg, the 004 right leg and 007 right leg reveals that the prediction results of our proposed CNN-BiLSTM model are smoother and more precise. In addition, we preliminarily verified the real-time performance of the algorithm. The twenty experiments results showed that the model average testing time was 62.47 $\pm$ 0.29 ms, and the algorithm delay time was less than the time to predict the knee joint angles in advance. In conclusion, our algorithm meets the requirements in terms of real-time performance and accuracy.

IV. Discussion

In this study, we proposed an CNN-BiLSTM model to combine the information from sEMG and IMU for knee angles prediction, which compensated for the one-sidedness of single modality prediction methods that only learn gait changes in a single measurement dimension. Meanwhile, we introduced the EFS to perform feature screening and remove feature redundancy. The CNN-BiLSTM, a spatio-temporal neural network structure, can better mine the characteristics of bimodal data. After testing on a separate test set, the average RMSE of our algorithm was 4.07. In addition, the visualization results show that it has the potential to realize the smooth control of exoskeleton robot.

We used fusion information from multiple sensors to achieve knee angles prediction. To compare multi-information and single information, we carried out comparative experiments of different input signals to the machine learning model. The results of statistical analysis showed that there was significant difference in prediction accuracy under different signal fusion (sEMG, IMU, sEMG + IMU). As is shown in Fig. 5, when IMU is used as input, the result of prediction is better than that of sEMG alone. This may be because the sEMG signal is weak, random and complex, and the traditional machine learning algorithms are not enough to learn deeper features. The prediction accuracy of sEMG + IMU is significantly higher than that of IMU. The movement of the human body is accomplished by the contraction and stretching of muscle groups to pull the corresponding bones. The sEMG has potential to identify earlier data response compared to IMU data and reflects muscle strength. The feedback of IMU on motion state and direction is more intuitive than sEMG and reflects the amount of exercise [3]. The fusion signals have highly feature complementary. Therefore, the method of multi-source signal fusion greatly improves the prediction accuracy.

In our experiment, we hope to get the contribution of each extracted feature to the result by EFS. The results showed that the contribution of kinematic features was higher than that of sEMG features. We speculate that the main reason is sEMG has a low amplitude, which is susceptible to noise interference, and it is difficult to learn deep-seated information of sEMG features for traditional machine learning models. From the final sEMG feature selection results, the main features involved Logvar, WL, WAMP, RMS, SSC, PF and MAV, among which the muscle feature combination of MAV and WL appeared the most. Previous studies have shown that these time-domain signal features are very effective in motion pattern recognition of lower limbs [3], [24], [31], [32], [33], [34]. MAV is often used to extract sEMG features due to its better performance [35], which reflects the overall activation of muscles. WL is a measure of the complexity of

the sEMG. They represent the amplitude and power level of the signal, respectively. From the final muscle combination, vastus lateralis and semitendinosus contained more sEMG features, followed by soleus, tibialis anterior, medial gastrocnemius and rectus femoris. This is mainly related to the function of the muscles, corresponding to the main muscle activity in each phase of the gait. As a knee extensor, the vastus lateralis controls and restricts the swing of the calf through eccentric contraction, thus making it possible to swing the lower limb forward [36]. It is the main muscle that produces power in the gait cycle. However, rectus femoris, which is also a knee extensor muscle, had only one feature preserved. We speculate that this is because both rectus femoris and vastus lateralis are components of the quadriceps muscle, the sEMG signals influence and the features intersect each other. This also reflects that EFS removes muscle feature redundancy well. Semitendinosus has the function of flexion and internal rotation of knee joint. Soleus and medial gastrocnemius are components of the calf triceps. In the middle stage of stance phase, the medial gastrocnemius contracts strongly and provides the driving force during walking, so the sEMG signal is very obvious. Soleus immobilizes the ankle and knee joints to prevent the body from leaning forward during the stance phase of the gait. Moreover, biceps femoris, as a knee flexor, was not present in the final muscle assemblage, possibly due to the thicker thigh fat and weak sEMG, which cannot reflect the actual muscle contraction well.

By utilizing feature selection done by EFS and PLM, feature redundancy information was removed, and robust prediction results were also guaranteed while improving computational performance. It has been confirmed that ensemble feature selection is a stable feature selection method, which can combine the advantages of different methods to generate feature subsets and address the limitations of a single algorithm [37], [38], [39]. This is consistent with our results. In addition, most of the previous integration methods were aimed at one or two of filtering, packaging and embedded selection. We properly combined the three methods to concentrate the advantages of more methods and improved the effect of feature selection. However, we can find that the prediction results are still slightly different from those before feature screening through careful observation, but there is no statistical difference. We speculate that this may be related to the inflection point selection of PLM algorithm.

Although previous studies have confirmed that end-to-end deep learning models can extract representative features from sEMG signals [40], [41], [42], after all, deep learning models are like a black box with poor interpretability. In our experiment, we therefore first manually selected a large number of features, filtered them through EFS, and then fed them into a deep learning model for training. This enables the model to extract deeper detailed descriptors from representative features, increasing thephysiological interpretation. In addition, the combination of CNN and BiLSTM can significantly improve the regression accuracy and robustness. On the one hand, the convolution operation enables the CNN to extract the spatial characteristics of sEMG signals [43], thereby representing the pattern of muscle activation. On the other hand, the CNN essentially ignores the temporal dependence of the signals. The introduction of BiLSTM can mine the historical information of deep feature vectors, which is superior to LSTM. The bidirectional nature can effectively detect the correlation between bimodal signals and joint angle signals [29]. Therefore, CNN-BiLSTM further improves the prediction accuracy by effectively extracting the spatio-temporal correlations of the fused signals.

During the normal gait, the knee joint will produce a large flexion angle, which varies from 0 to 60° in the sagittal plane. So, the joint angles prediction requires higher real-time performance. In this study, we used the characteristic that the sEMG signal is generated 30∼150 ms before the action to predict the knee joint angle 100 ms in advance. The algorithm delay is only about 62.47 ms. Our model is basically in line with the satisfaction of the demand for real-time applications, and can leave some time for other mechanical delays.

Nonetheless, there are a few limitations in our work. First, knee, hip and ankle joint all place the key role in walking rehabilitation. We will do multi joint prediction to achieve more comprehensive rehabilitation robot control in future research. Secondly, bi-modal data were fused in this study. More modal data can further be fused to improve the accuracy and real-time of prediction. Thirdly, our model has not been tested in the actual control system. In the future, we will extend our model to online testing and deeply explore its performance in the real exoskeleton control system.

V. Conclusion

Accurate and real-time prediction of knee joint angles plays an important role in the rehabilitation training of patients with abnormal lower limb motor function. In the study, we presented a feature-based CNN-BiLSTM model to predict the lower limb knee angles after 100 ms in advance for the subjects, using an EFS-based ensemble algorithm for efficient feature selection of both sEMG features and IMU features. The results indicated that the EFS and PLM algorithm could remove the feature redundancy perfectly and estimation performance would become better when bi-modal gait data were fused. For the estimation performance, our model had smaller estimation error and higher correlation coefficient than well-established models based on the same subjects. The model test time was 62.47 ± 0.29 ms, less than 100 ms. This provesmm that our model has the potential to be applied to exoskeleton robot control, and thereby improve the efficiency of rehabilitation training for patients.

References

[1] R. S. Calabrò *et al.*, "Robotic gait rehabilitation and substitution devices in neurological disorders: Where are we now?" *Neurol. Sci.*, vol. 37, no. 4, pp. 503–514, Apr. 2016.

[2] A. Gonzalez, L. Garcia, J. Kilby, and P. McNair, "Robotic devices for paediatric rehabilitation: A review of design features," *Biomed. Eng. OnLine*, vol. 20, no. 1, p. 89, Dec. 2021.

[3] B. Zhou *et al.*, "Accurate recognition of lower limb ambulation mode based on surface electromyography and motion data using machine learning," *Comput. Methods Programs Biomed.*, vol. 193, Sep. 2020, Art. no. 105486.

[4] W. Meng, Q. Liu, Z. D. Zhou, Q. S. Ai, B. Sheng, and S. Q. Xie, "Recent development of mechanisms and control strategies for robot-assisted lower limb rehabilitation," *Mechatronics*, vol. 31, pp. 132–145, Oct. 2015.

[5] G. Morone *et al.*, "Robot-assisted gait training for stroke patients: Current state of the art and perspectives of robotics," *Neuropsychiatric Disease Treatment*, vol. 13, pp. 1303–1311, May 2017.

[6] T. M. Bittibssi, A. H. Zekry, M. A. Genedy, and S. A. Maged, "SEMG pattern recognition based on recurrent neural network," *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 103048.

[7] F. Zhang *et al.*, "SEMG-based continuous estimation of joint angles of human legs by using BP neural network," *Neurocomputing*, vol. 78, no. 1, pp. 139–148, Feb. 2012.

[8] P. Patil, K. S. Kumar, N. Gaud, and V. B. Semwal, "Clinical human gait classification: Extreme learning machine approach," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–6.

[9] V. B. Semwal, N. Gaud, P. Lalwani, V. Bijalwan, and A. K. Alok, "Pattern identification of different human joints for different human walking styles using inertial measurement unit (IMU) sensor," *Artif. Intell. Rev.*, vol. 55, no. 2, pp. 1149–1169, Feb. 2022.

[10] F. Gao, T. Tian, T. Yao, and Q. Zhang, "Human gait recognition based on multiple feature combination and parameter optimization algorithms," *Comput. Intell. Neurosci.*, vol. 2021, Feb. 2021, Art. no. 6693206.

[11] A. Gupta and V. B. Semwal, "Multiple task human gait analysis and identification: Ensemble learning approach," in *Emotion and Information Processing*, S. N. Mohanty, Ed. Berlin, Germany: Springer, 2020, pp. 185–197.

[12] C. Lin, C. Hao, W. Jue, and L. Tian, "Advance in human motion intention recognition based on surface electromyography (review)," *Chin. J. Rehabil. Theory Pract.*, vol. 27, no. 5, pp. 595–603, 2021.

[13] M. Raj, V. B. Semwal, and G. C. Nandi, "Bidirectional association of joint angle trajectories for humanoid locomotion: The restricted Boltzmann machine approach," *Neural Comput. Appl.*, vol. 30, no. 6, pp. 1747–1755, Sep. 2018.

[14] A. Gupta and V. B. Semwal, "Occluded gait reconstruction in multi person gait environment using different numerical methods," *Multimedia Tools Appl.*, vol. 81, no. 16, pp. 23421–23448, Jul. 2022.

[15] Y. Deng, F. Gao, and H. Chen, "Angle estimation for knee joint movement based on PCA-RELM algorithm," *Symmetry*, vol. 12, no. 1, p. 130, Jan. 2020.

[16] Z. Li, X. Guan, K. Zou, and C. Xu, "Estimation of knee movement from surface EMG using random forest with principal component analysis," *Electronics*, vol. 9, no. 1, p. 43, Dec. 2019.

[17] G. Liu, L. Zhang, B. Han, T. Zhang, Z. Wang, and P. Wei, "SEMG-based continuous estimation of knee joint angle using deep learning with convolutional neural network," in *Proc. IEEE 15th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2019, pp. 140–145.

[18] A. Olsson, N. Malesevic, A. Bjorkman, and C. Antfolk, "Exploiting the intertemporal structure of the upper-limb sEMG: Comparisons between an LSTM network and cross-sectional myoelectric pattern recognition methods," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 6611–6615.

[19] C. Wang, W. Guo, H. Zhang, L. Guo, C. Huang, and C. Lin, "SEMG-based continuous estimation of grasp movements by long-short term memory network," *Biomed. Signal Process. Control*, vol. 59, May 2020, Art. no. 101774.

[20] V. B. Semwal, A. Gupta, and P. Lalwani, "An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition," *J. Supercomput.*, vol. 77, pp. 12256–12279, Apr. 2021.

[21] R. Jain, V. B. Semwal, and P. Kaushik, "Deep ensemble learning approach for lower extremity activities recognition using wearable sensors," *Expert Syst.*, vol. 39, no. 6, Jul. 2022, Art. no. e12743.

[22] C.-F. Chen, Z.-J. Du, L. He, Y.-J. Shi, J.-Q. Wang, and W. Dong, "A novel gait pattern recognition method based on LSTM-CNN for lower limb exoskeleton," *J. Bionic Eng.*, vol. 18, no. 5, pp. 1059–1072, Sep. 2021.

[23] B. Hu, E. Rouse, and L. Hargrove, "Benchmark datasets for bilateral lower-limb neuromechanical signals from wearable sensors during unassisted locomotion in able-bodied individuals," *Frontiers Robot. AI*, vol. 5, p. 14, Feb. 2018.

[24] G. Hao, "Research on human lower limb motion intention detection and prediction based on EMG," M.S. thesis, Dept. Mach. Eng., Harbin Inst. Technol., Harbin, China, 2019.

[25] R. Merletti and L. R. Lo Conte, "Advances in processing of surface myoelectric signals: Part 1," *Med. Biol. Eng. Comput.*, vol. 33, no. 3, pp. 362–372, May 1995.

[26] M. Zhu and A. Ghodsi, "Automatic dimensionality selection from the scree plot via the use of profile likelihood," *Comput. Statist. Data Anal.*, vol. 51, no. 2, pp. 918–930, Nov. 2006.

[27] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Elect. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014.

[28] C. Chen *et al.*, "DNN-DTIs: Improved drug-target interactions prediction using XGBoost feature selection and deep neural network," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104676.

[29] C. Ma *et al.*, "A bi-directional LSTM network for estimating continuous upper limb movement from surface electromyography," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 7217–7224, Oct. 2021.

[30] S. J. Raudys and A. K. Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners and open problems," in *Proc. 10th Int. Conf. Pattern Recognit.*, Atlantic City, NJ, USA, 1990, pp. 417–423.

[31] H. Huang, F. Zhang, L. J. Hargrove, Z. Dou, D. R. Rogers, and K. B. Englehart, "Continuous locomotion-mode identification for prosthetic legs based on neuromuscular-mechanical fusion," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2867–2875, Oct. 2011.

[32] A. Phinyomark and E. Scheme, "A feature extraction issue for myoelectric control based on wearable EMG sensors," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Seoul, South Korea, Mar. 2018, pp. 1–6.

[33] A. J. Young, L. H. Smith, E. J. Rouse, and L. J. Hargrove, "Classification of simultaneous movements using surface EMG pattern recognition," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 5, pp. 1250–1258, May 2013.

[34] S. Kyeong, J. Feng, J. K. Ryu, J. J. Park, K. H. Lee, and J. Kim, "Surface electromyography characteristics for motion intention recognition and implementation issues in lower-limb exoskeletons," *Int. J. Control, Autom. Syst.*, vol. 20, no. 3, pp. 1018–1028, Mar. 2022.

[35] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Syst. Appl.*, vol. 39, pp. 7420–7431, Jun. 2012.

[36] J. Wang, L. Wang, X. Xi, S. M. Miran, and A. Xue, "Estimation and correlation analysis of lower limb joint angles based on surface electromyography," *Electronics*, vol. 9, no. 4, p. 556, Mar. 2020.

[37] T. Abeel, T. Helleputte, Y. Van de Peer, P. Dupont, and Y. Saeys, "Robust biomarker identification for cancer diagnosis with ensemble feature selection methods," *Bioinformatics*, vol. 26, no. 3, pp. 392–398, 2010.

[38] Y. Li, S. Gao, and S. Chen, "Ensemble feature weighting based on local learning and diversity," in *Proc. 26th AAAI Conf. Artif. Intell.*, vol. 2, 2012, pp. 1019–1025.

[39] Y. Li, T. Li, and H. Liu, "Recent advances in feature selection and its applications," *Knowl. Inf. Syst.*, vol. 53, no. 3, pp. 551–577, Dec. 2017.

[40] C. Xie, D. Wang, H. Wu, and L. Gao, "A long short-term memory neural network model for knee joint acceleration estimation using mechanomyography signals," *Int. J. Adv. Robotic Syst.*, vol. 17, no. 6, Nov. 2020, Art. no. 172988142096870.

[41] A. L. Delgado, A. F. Da Rocha, A. S. Leon, A. Ruiz-Olaya, K. R. Montero, and A. L. Delis, "Estimation of joint angle from sEMG and inertial measurements based on deep learning approach," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 700–703.

[42] U. Cātāllard *et al.*, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Oct. 2019.

[43] T. Bao, S. A. R. Zaidi, S. Xie, P. Yang, and Z.-Q. Zhang, "A CNN-LSTM hybrid model for wrist kinematics estimation using surface electromyography," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.