An Attention-based CNN-LSTM Model with Limb Synergy for Gait Trajectory Prediction

Chang Zhu, Quan Liu, Wei Meng, *Member, IEEE*, Qingsong Ai, *Member, IEEE*, Sheng Q Xie, *Senior Member, IEEE*

Abstract— **Gait prediction is crucial in exoskeleton-assisted gait rehabilitation by recognizing the movement intention of patients, so as to realize the adaptive and transparent robotic assistance. Human locomotion has inherent synergies and coordination, and the dynamic mapping of the upper and lower limbs is beneficial to improve the prediction accuracy. Current prediction methods did not consider the correlation of gait data in time and space, resulting in a large amount of redundant data and low prediction accuracy. This paper proposes a gait trajectory prediction method based on convolutional neural network-long short-term memory (CNN-LSTM) model, which predicts the human knee/ankle joint trajectory based on upper and lower limb collaborative data. The attention mechanism is applied to determine which dimensions are essential in gait prediction, so the accuracy can be improved by adopting key elements. Results show that, within a predicted horizon of 50ms, the prediction RMSE is as low as 0.317 degrees.**

I. Introduction

According to World Health Organization (WHO), the proportion of the global aging population will increase from 12% to 22% between 2015 to 2050 [1]. Meanwhile, there are 24.72 million disabled people in China [2], and China ranks the first in the world for the incidence of stroke [3]. The elderly and stroke patients are accompanied by varying degrees of limb weakness. Lower-limb exoskeletons can provide gait rehabilitation for patients, which can effectively promote muscle strength and neural circuit remodeling. In existing exoskeletons control strategies, pre-defined gait trajectory control is usually used for gait rehabilitation [4], but this pre-defined gait trajectory usually does not take into account the characteristics of the patient’s gait and cannot meet the individualized rehabilitation needs. The patient's motion intention prediction model has been applied to gait trajectory prediction. However, the prediction accuracy of these models is still limited without fully extracting the characteristics of the data in time and space.

This project is supported by National Natural Science Foundation of China (Grant No. 52075398) and the Application Foundation Frontier Project of Wuhan Science and Technology Program (Grant No. 2020020601012220). (*Corresponding author: Wei Meng*)

C. Zhu is with School of Information Engineering, Wuhan University of Technology, 122 Luoshi Road, Wuhan, China (changzhu@whut.edu.cn).

Q. Liu is with the School of Information Engineering, Wuhan University of Technology, 122 Luoshi Road, Wuhan (quanliu@whut.edu.cn).

W. Meng is with the School of Information Engineering, Wuhan University of Technology, 122 Luoshi Road, Wuhan, China (email: weimeng@whut.edu.cn).

Q. Ai is with the School of Information Engineering, Wuhan University of Tecqhnology, 122 Luoshi Road, Wuhan (qingsongai@whut.edu.cn).

S. Q. Xie is with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, LS2 9JT, UK (email: s.q.xie@leeds.ac.uk).

Trajectory prediction of lower limb is currently the main focus, but the coordination of upper and lower limbs has a positive significance for gait prediction. Several studies have examined the coordination between the upper and lower limbs of the gait. Research has proven that the activities of the arms and legs remain coordinated in various human locomotion, and the arm and leg movement keep the frequency locked at 1/1 during walking [5]. Boudali et al. used Koopman operator to identify the dynamic mapping between the upper limb and its opposite lower limb in human motion, and the authors intended to design a new method of controlling the exoskeleton of the lower limbs based on the mapping principles. The wearer can control the movement of the damaged leg through the movement of the arm to realize more voluntary control [6]. In addition, no difference has been observed in any coordination measurements between healthy subjects and stroke subjects, indicating that the ability to coordinate arm and leg movements during walking is still maintained in stroke individuals [7].

Human motions contain complex neuromusculoskeletal coordination between the different joints of the upper and lower limbs. Therefore, the upper and lower limbs can be coordinated to more accurately predict the joint angle through this connection [8, 9]. Eslamy et al. developed an advanced controller which could map calf kinematics to ankle torque and angle. Meanwhile, it did not require speed determination, gait percentage recognition, etc. [10]. He et al. proposed a method for lower limb trajectory prediction based on LSTM model. This approach based on synergy theory utilized the joint angle of the previous swing process to generate the motion trajectory of the lower limb joint [11].

Machine learning methods have been used in gait prediction to achieve high prediction accuracy. Huang et al. built a model to predict knee angles in real time, which was created based on a combination of electromyography (EMG) signals and inertial data of the thigh and calf. The prediction model based on fusion signal achieve a balanced between prediction accuracy and computational complexity [12]. Compared with a single neural network, hybrid neural networks can achieve better performance. Xiong et al. used a small number of input variables selected through the elastic network as the input of the artificial neural network (ANN) to predict the joint torque, which could make predictions in daily life, and has good individual adaptability and environmental adaptability [13]. Gautam et al. developed an accurate hybrid deep learning model that could recognize lower limb movements and predict the joint angle information of the executed limb movements [14].

Due to the periodicity of gait data, long short-term memory (LSTM) can well capture the time characteristics of gait signals [15]. He et al. used a method based on the synergy theory to make the joint angle of the previous upper limb data can generate the subsequent lower limb joint trajectory [11]. Liu et al. used LSTM to take advantage of the synergy of the lower limbs, and predicted the motion trajectory of the knee joint based on the hip and ankle joints [16]. Convolutional neural network (CNN) model can predict gait data as the convolution kernel can learn the characteristics of gait data for a period of time, and make predictions based on historical gait data. Compared with LSTM, CNN can see more historical data. Although LSTM has a certain memory ability, this memory is only short-lived [17]. William et al. used five pre-trained CNN models to compare the multiple regression of the 3D ground reaction force and torque based on the marker-based motion capture, and each model had been tested for margins, which made it possible to accurately predict the force and torque outside the laboratory [18]. Gholami et al. chose the CNN model as the regression model to minimize the joint angle prediction error in the scene between participants, and the root mean square error (RMSE) of the CNN model's prediction error from the actual angle was less than 3.5° and 6.5° in intra- and inter-participant scenarios [19].

By fusing LSTM and CNN neural network, more spatial and temporal characteristics of gait data can be integrated, thereby improving the prediction accuracy [20, 21]. Zhen et al. proposed an algorithm based on long and short-term memory network and convolutional neural network (LCWSnet), which uses leg Euler angle information to diagnose and classify gait abnormalities, and can adaptively adjust feature-related parameters [20]. The attention mechanism can devote more attention to important areas to obtain more detailed information and suppress other useless information [22]. Chen et al. proposed an attention-based CNN-LSTM method for sleep awakening detection using heterogeneous sensor data, with a significant improvement from 5% to 46% [23].

In this paper, we propose a lower limb trajectory prediction framework for patients with certain mobility in the middle and late stage of rehabilitation, which combines the synergy and the attention mechanism to improve the prediction accuracy. The personalized predicted trajectory used to drive the exoskeleton robot is more conducive to assisting patients in rehabilitation. The remaining of this paper is organized as follows: Section Ⅱ introduces adopted model and the basic principles of CNN, LSTM and attention mechanism. Section Ⅲ describes the experimental environment and data preprocessing. The experimental results and their analysis are in Section Ⅳ. Section Ⅴ is the discussion and prospect of this article.

II. METHODS

A. Attention-based CNN-LSTM Model

CNN is suitable for spatial abstraction and generalization, while LSTM is fit for extending temporal features and processing sequential data. To construct a gait trajectory prediction model, it is necessary to consider not only the relationship of its spatial characteristics, but also the associated information in the time dimension. Therefore, this article combines the advantages of the two models to use the CNN-LSTM model. CNN mines the correlation between multi-dimensional data and removes noise and unstable components. LSTM uses the information processed by CNN for long sequence prediction. The attention mechanism assigns more weights to important features to improve the prediction accuracy.

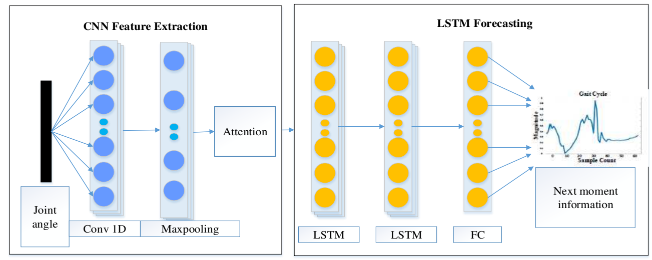


Figure1. The structure of attention-based CNN-LSTM model

The model constructed in this paper consists of three parts: CNN network, LSTM network and attention mechanism. First, in the CNN network, a one-dimensional convolution kernel is used to convolve the joint angle data to extract the characteristic components in the spatial structure. The MaxPooling layer reduces the number of model parameters and overfitting problem. Subsequently, the LSTM network performs sequence prediction based on the extracted feature components. Since the features extracted by the CNN network still have timing characteristics, it can be directly and effectively modeled by using LSTM. The position of the attention module is adjustable. The attention module multiplies and adds the output vectors of the hidden layer at different time points and the corresponding weights to give more weight to important feature as the final feature expression of the model. The attention module can make the model obtain more comprehensive and detailed feature information. The attention-based CNN-LSTM model (attention module before the LSTM) is shown in Figure 1.

B. CNN Networks

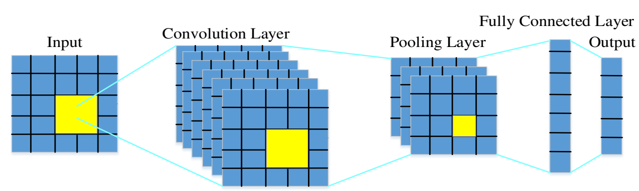


Figure2. The structure of CNN model

CNN model uses a sliding window to perform convolution operations on the original data to extract gait features over a period of time. The MaxPooling layer to extract more important features to form high-dimensional features. The fully connected layer is to reduce the feature dimensionality extracted by the MaxPooling layer, which interprets the high-dimensional features as low-dimensional outputs. Finally, the dense layer is used to correspond to the output dimensions. The CNN structure is shown in Figure 2

Spatiotemporal features can be easily extracted by the one-dimensional (1D) CNN (1D-CNN) from the model input. Let the given model input be $\mathrm{X} = [\mathrm{x}_1, \mathrm{x}_2, \ , \mathrm{x}_t]$, consisting of joint angle of upper and lower limbs. Firstly, the model input X is input to the 1D-CNN layer, and there has

$$l_t = \tanh(x_t \times k_t + b_l) \quad (1)$$

where x_t represents the input vector, k_t is the convolution kernel, b_l represents bias vector, and l_t is the output vector of the 1D-CNN layer. The output of the 1D-CNN layer is a spatiotemporal feature matrix $L = [l_1, l_2, \ , l_t]$.

C. LSTM Networks

As a special kind of RNN, the LSTM network is capable of learning long-term dependencies. It has the advantage of connecting previous information to the present task. Because of its special memory cell architecture, the LSTM network overcomes the defects of the traditional RNN, especially the problems of gradient disappearance and gradient explosion. The architecture of an LSTM memory cell is shown in Figure 3, where each cell has three "gate" structures, namely, the input gate, the forget gate, and the output gate. A chain of repeating cells forms the LSTM layer.

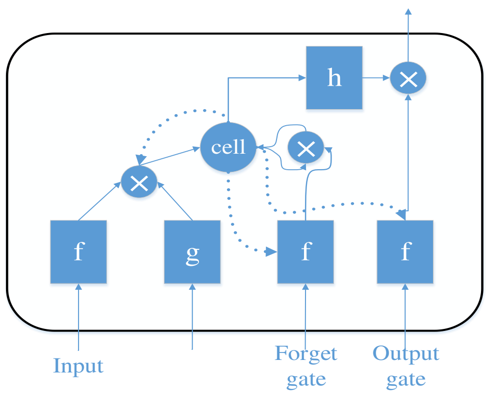


Figure3. Standard LSTM model

X_t represents the *t*th time series value fed in LSTM. c_t represents the memory cell, which is the core of LSTM. Memory cell can control the transformation of different time information. The input gate determines the information that the current time deliver to the next time. The forget gate indicates how much information of the previous time has been retained in the current time. The output gate determines the output of the current state to the next state. The equation of different cells in LSTM is shown as the following:

$$\begin{aligned} i_t &= \delta(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_t &= \delta(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ o_t &= \delta(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\ \tilde{c_t} &= \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ c_t &= f_t e c_{t-1} + i_t e \tilde{c_t} \\ h_t &= o_t e \tanh(c_t) \end{aligned} \tag{2}$$

where i_t, f_t, o_t represent the *t*th input gate, forget gate and output gate function. W_{xi}, W_{xf}, W_{xo}, W_{xc} represent weights of input gate, forget gate, input gate and memory cell. W_{hi}, W_{hf}, W_{ho}, W_{hc} represent weights from hidden layers to input gate, forget gate, input gate and the memory cell. b_i, b_f, b_o, b_c are the bias values of the input gate, forget gate, output gate and the memory cell.

D. Attention Layer

The attention mechanism performs a weighted summation calculation on the hidden layer vector expression output, where the size of the weight indicates the importance of the feature at each time point. The attention mechanism is shown in the Figure 4.

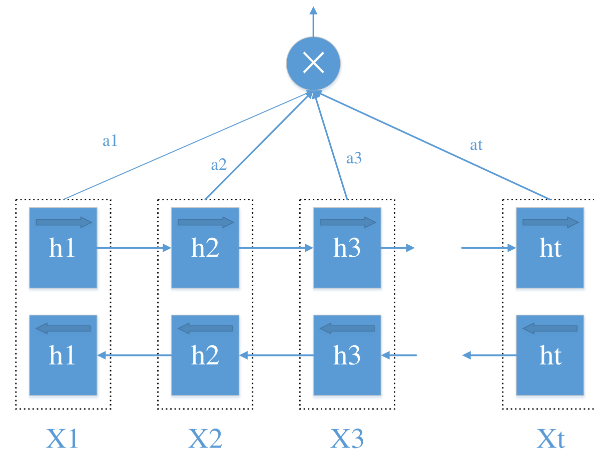


Figure4. Attention mechanism [24]

$$c_i = \sum_{i=1}^{k} a_i h_i \tag{3}$$

Suppose the input is k feature vectors h_i (Encoder hidden state), $i = 1,2, \ldots k$. Attention weight a_i is the weight added by the state. The content vector c_i is the weighted sum of all hidden states of the encoder and their corresponding attention weights.

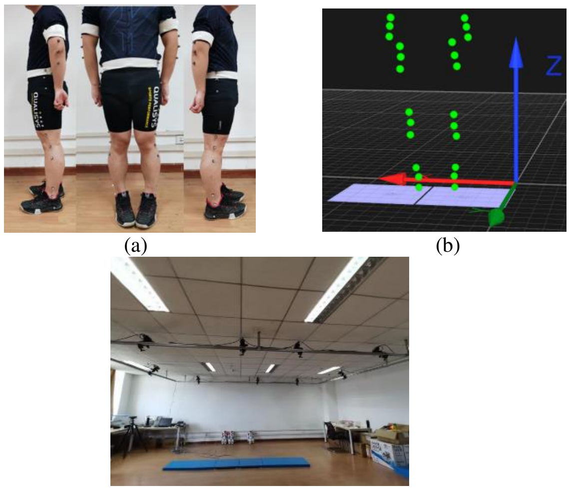
III. Experiments

A. Experimental Setup

The following experimental procedures are designed to develop the proposed algorithm and verify its performance. The subjects participated in the experiment and reached a consensus with the research team on the detailed procedures. The subject is healthy and has no history of neurological abnormalities.

In terms of experimental data, the subjects (male between 23 and 24 years old) with body weight ranging from 46 kg to 70 kg and height ranging from 158 cm to 177 cm were selected to collect joint data. The subject's legs or feet did not have any diseases that could affect normal walking.

We prepared a total of 24 gait feature data from the subject's upper and lower limbs, and the normal walking data was captured by the Qualisys system. Figure 5 (a, b) shows the marking points attached to the subject, Figure 5 (c) shows the camera position of the motion capture system. Our experiment is implemented on the Tensorflow framework, which is a popular deep learning framework. The sampling frequency of the Qualisys system is 150 Hz, and the step length is 15, so the joint angle after 0.1s can be predicted.



(a) (b)

(c)

Figure5. (a) Marked points on the subject, (b) Marked points on QTM software, (c) Motion capture using Qualisys system

B. Model Evaluation Metrics

There are two common metrics of neural network regression prediction.

1. Root Mean Square Error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{m}^{1}(y_i - \hat{y}_i)^2} \qquad (4)$$

2. Mean Absolute Error (MAE)

$$\mathrm{MAE} = \frac{1}{m}\sum_{m}^{1}|y_i - \hat{y}_i| \qquad (5)$$

where y_i is the forecasted value, $\hat{y}_i$ is the true value. From these equations, lower MSE and MAE represent better forecasting accuracy.

C. Results.

The proposed model is built by Python deep learning module TensorFlow 2.0. The loss function is set to MSE, optimizer is set to Adaptive moment estimation. Epochs are set to 300. The batch size is set to 128. Experiment platform is shown as follows. Operation system: Windows 10, CPU: Intel Xeon, Random Access Memory (RAM): 13 GB.

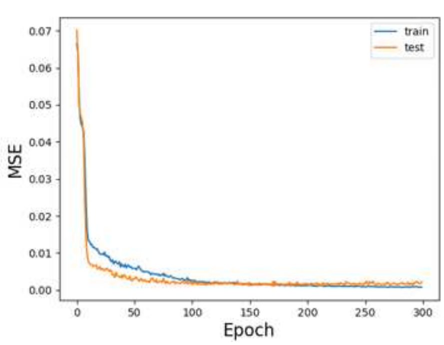


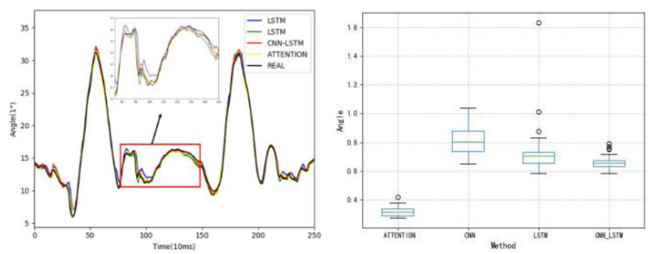
Figure6. Loss curve in the training process

As illustrated schematically in Figure 6, the loss function values of both the training set and the testing set decrease rapidly. The loss of the training set converges after 100 epochs, and the loss of the test set converges after 80 epochs.

TABLE I. PREDICTION ERROR RESULTS FOR ANGLE

Method	Metric	Right lower limb	Lower limb	Upper and lower limb
CNN	RMSE	1.369	0.995	0.810
	MAE	1.031	0.744	0.619
LSTM	RMSE	1.014	0.844	0.745
	MAE	0.770	0.636	0.562
CNN-LSTM	RMSE	0.823	0.741	0.665
	MAE	0.618	0.569	0.523
CNN-LSTM-attention	RMSE	0.391	0.373	0.317
	MAE	0.269	0.249	0.202

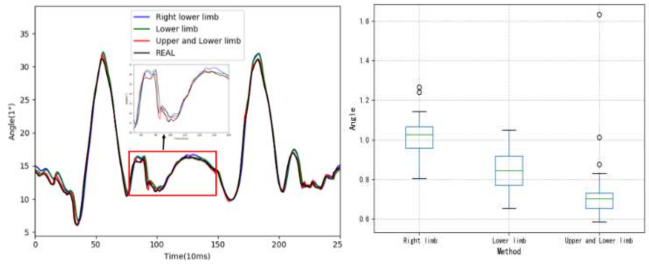
In order to ensure the effectiveness and stability of the proposed method, for the same sample dataset, the average value of predicted angles from thirty repeated experiments is calculated. Table I shows the MAE and RMSE of various methods and synergies in 30 repeated calculations. Some conclusions can be drawn by comparing the prediction angles of different algorithms and synergies.



(a) (b)

Figure7. Prediction results of different algorithms

According to Figure 7, CNN-LSTM has lower forecasting error than CNN and LSTM, indicating that the combined model of CNN and LSTM has better predictive ability. In addition, the model using CNN is better than the model without CNN, which shows that CNN can effectively extract features from the data to improve prediction performance. The performance of attention-based CNN-LSTM is better than the other three structures, indicating that assigning appropriate weights to different dimensions can effectively improve the prediction accuracy. It is worth mentioning the results of the proposed model are more stable than CNN-LSTM models in 30 calculations. When there are many input dimensions and the importance of different dimensions is varying, more attention can be paid to the important dimensions to achieve better prediction performance.



(a) (b)

Figure8. Prediction results of different synergies

This study tested the effect of upper and lower limb synergy on improving the accuracy of single joint prediction. Figure 8 shows the influence of different synergies on the prediction results of the knee joint angle. Figure 8 (b) shows that the error of lower limb is lower than that of the right lower limb, indicating that there is a synergy between the left and right legs. Similarly, there is also a synergy between the upper and lower limbs. A healthy gait movement is the result of complex neural coordination between different moving joints of the human body, and this synergy can be used to improve the accuracy of gait prediction.

TABLE II. PREDICTION ERROR OF MAIN ATTENTION ON ONE JOINT

Type	Ipsilateral ankle	Ipsilateral hip	Opposite knee
RMSE	0.390	0.358	0.338
MAE	0.256	0.253	0.250

It can be seen from Table Ⅱ that main attention on different joints has various prediction error when we predict the right knee joint trajectory. It shows that although human gait movement is the result of the limb synergy, the degree of synergy between different joints is different.

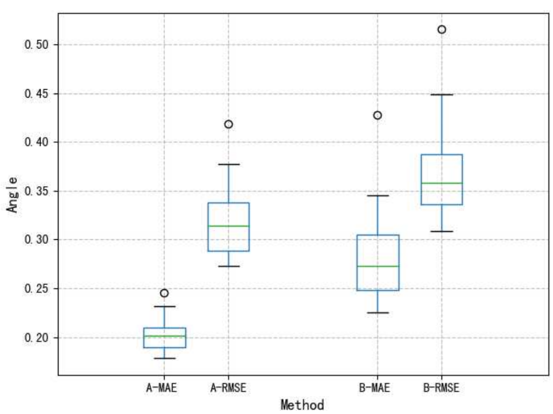


Figure9. Prediction errors of different attention mechanism location

Adding attention mechanism to different parts of the model has different meanings, and its structure will be different. The attention layer has an N-to-1 structure before LSTM, and the attention layer has an N-to-N structure after LSTM. We will transform the position of the attention layer and place them in the input layer (before the LSTM) and the output layer (after the LSTM) of the entire classification model to compare it to check whether the attention mechanism can capture key information in each place. From Figure 9, it shows that using the attention mechanism after the LSTM layer has lower errors than using the attention mechanism before the LSTM layer. When using the attention mechanism after LSTM, more features will be assigned attention, but these features are more abstract to explain.

Ⅳ. DISCUSSION

The experimental results show that the proposed model combines the advantages of CNN and LSTM. Theoretically, CNN model is more suitable for spatial expansion, extracting local features of data and combining them into high-level features. The gait information can reflect the current state of motion in a period of time, and its gait characteristics also conform to local association and weight sharing, that is, each neuron only needs to analyze the surrounding data. LSTM is more suitable for time extension, has long-term memory function, and is more suitable for processing time series. The CNN-LSTM model has the ability of spatiotemporal feature expression, and the prediction effect will be more accurate. The CNN-LSTM model with attention mechanism has better accuracy in joint prediction tasks, which proves that attention mechanism can adjust the network structure and solve the loss of important features of CNN-LSTM model.

The deep-RNN method collecting data from EMG and IMUs sensors is suitable for patients outside the lab [12], and the LSTM-based prediction method is suitable for the early stage of rehabilitation, generating lower limb intentions through previous swing process of upper limb [11]. By contrast, the proposed method is more suitable for patients in the middle and late stages of rehabilitation to perform specific exercise training in the lab. The gait information predicts the motion trajectory of the required joints, which has a wide range of applications in exoskeletons. When there are more input dimensions, more attention can be paid to important features to obtain higher accuracy. Therefore, it is a good option to combine the upper and lower extremity collaborative prediction with the attention mechanism.

Ⅴ. CONCLUSION

This paper establishes a neural network model, which uses the movements of the upper and lower limbs to predict the trajectory of the knee joint throughout the gait cycle. The results show that the upper and lower limbs have the better effect of synergy prediction. In addition, the results indicate that the degree of synergy between different joints is different. The upper and lower limb synergy and attention mechanism have the potential to be used in the control of exoskeleton robots. Correctly predicting angle can be used to achieve continuous and smooth control on exoskeletons, enabling the robot to perform human-like smooth movements, which is of great significance for human-robot compliance control. In future works, we intend to quantify the synergy of different joints in different movements and apply the method to gait recognition to control the exoskeleton more compliantly.

REFERENCES

[1] WHO. (2018.02.15). *Ageing and health.* Available: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

[2] *China Statistical Yearbook on the Work for Persons with Disabilities-2019,* C. S. Press, 2019.

[3] W. Longde, L. Jianmin, Y. Yi, P. Bin, and W. Yilong, "The Prevention and Treatment of Stroke Still Face Huge Challenges——Brief Report on Stroke Prevention and Treatment in China," *Chinese Circulation Journal,* vol. 34(02), pp. 105-119, 2019.

[4] T. Yan, M. Cempini, C. M. Oddo, and N. Vitiello, "Review of assistive strategies in powered lower-limb orthoses and exoskeletons," *Robotics and Autonomous Systems,* vol. 64, pp. 120-136, 2015.

[5] T. W. C. B. G. C. V. Dietz, "Arm to leg coordination in humans during walking, creeping and swimming activities," *Exp Brain Res,* vol. 141(3), pp. 375–379, 2001.

[6] A. Mounir Boudali, Peter J. Sinclair, Richard Smith, and I. R. Manchester., "Human locomotion analysis: Identifying a dynamic mapping between upper and lower limb joints using the Koopman operator," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 1889-1892.

[7] J. L. Stephenson, A. Lamontagne, and S. J. De Serres, "The coordination of upper and lower limb movements during gait in healthy and stroke individuals," *Gait Posture,* vol. 29, no. 1, pp. 11-6, Jan 2009.

[8] M. Baniasad, F. Farahmand, M. Arazpour, and H. Zohoor, "Coordinated activities of trunk and upper extremity muscles during walker-assisted paraplegic gait: A synergy study," *Hum Mov Sci,* vol. 62, pp. 184-193, Dec 2018.

[9] B. Gialanella and R. Santoro, "Prediction of functional outcomes in stroke patients: the role of motor patterns according to limb synergies," *Aging Clin Exp Res,* vol. 27, no. 5, pp. 637-45, Oct 2015.

[10] M. Eslamy and K. Alipour, "Synergy-based Gaussian Process Estimation of Ankle Angle and Torque: Conceptualization for High level Controlling of Active Foot Prostheses/Orthoses," *J Biomech Eng,* Oct 22 2018.

[11] J. He, Z. Guo, Z. Shao, J. Zhao, and G. Dan, "An LSTM-Based Prediction Method for Lower Limb Intention Perception by Integrative Analysis of Kinect Visual Signal," *J Healthc Eng,* vol. 2020, p. 8024789, 2020.

[12] Y. Huang *et al.*, "Real-Time Intended Knee Joint Motion Prediction by Deep-Recurrent Neural Networks," *IEEE Sensors Journal,* vol. 19, no. 23, pp. 11503-11509, 2019.

[13] B. Xiong *et al.*, "Intelligent Prediction of Human Lower Extremity Joint Moment: An Artificial Neural Network Approach," *IEEE Access,* vol. 7, pp. 29973-29980, 2019.

[14] A. Gautam, M. Panwar, D. Biswas, and A. Acharyya, "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 J Transl Eng Health Med,* vol. 8, p. 2100310, 2020.

[15] A. Zaroug, D. T. H. Lai, K. Mudie, and R. Begg, "Lower Limb Kinematics Trajectory Prediction Using Long Short-Term Memory Neural Networks," *Front Bioeng Biotechnol,* vol. 8, p. 362, 2020.

[16] D.-X. Liu, X. Wu, W. Du, C. Wang, C. Chen, and T. Xu, "Deep Spatial-Temporal Model for rehabilitation gait: optimal trajectory generation for knee joint of lower-limb exoskeleton," *Assembly Automation,* Article vol. 37, no. 3, pp. 369-378, 2017.

[17] X. Wang, J. Zhang, and W. Q. Yan, "Gait recognition using multichannel convolution neural networks," *Neural Computing and Applications,* vol. 32, no. 18, pp. 14275-14285, 2019.

[18] J. A. William R. Johnson, David Lloyd, and Ajmal Mian, "Predicting Athlete Ground Reaction Forces andMoments from Spatio-temporal DrivenCNN Models," *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING,* vol. 66, pp. 689-694, 2019.

[19] M. Gholami, C. Napier, and C. Menon, "Estimating Lower Extremity Running Gait Kinematics with a Single Accelerometer: A Deep Learning Approach," *Sensors (Basel),* vol. 20, no. 10, May 22 2020.

[20] J. Gao, P. Gu, Q. Ren, J. Zhang, and X. Song, "Abnormal Gait Recognition Algorithm Based on LSTM-CNN Fusion Network," *IEEE Access,* vol. 7, pp. 163180-163190, 2019.

[21] T. Zhen, L. Yan, and P. Yuan, "Walking Gait Phase Detection Based on Acceleration Signals Using LSTM-DNN Algorithm," *Algorithms,* vol. 12, no. 12, 2019.

[22] K. Wu *et al.*, "An attention-based CNN - LSTM - BiLSTM model for short - term electric load forecasting in integrated energy system," *International Transactions on Electrical Energy Systems,* vol. e12637, pp. 1-15, 2020.

[23] Z. Chen, M. Wu, W. Cui, C. Liu, and X. Li, "An Attention Based CNN-LSTM Approach for Sleep-Wake Detection with Heterogeneous Sensors," *IEEE Journal of Biomedical and Health Informatics,* pp. 1-9, 2020.

[24] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *ICLR*, 2014, pp. 1-15.