An Attention-based Bidirectional LSTM Model for Continuous Cross-subject Estimation of Knee Joint Angle during Running from sEMG Signals

Alireza Rezaie Zangene, Oluwarotimi Williams Samuel*, IEEE Senior Member, Ali Abbasi, Kianoush Nazarpour, IEEE Senior Member, Alistair A. McEwan, and Guanglin Li, IEEE Senior Member

Abstract— **Running is an essential locomotion activity that plays a critical role in everyday life and exercise activities, and may be impeded by joint disease and neurological impairments. Accurate and robust estimation of joint kinematics via surface electromyogram (sEMG) signals provides a human-machine interaction-based method that can be used to adequately control rehabilitation robots while performing complex movements such as running for motor function restoration in affected persons. To this end, this paper proposes a novel deep learning-based model (AM-BiLSTM) that integrates an attention mechanism (AM) and a bidirectional long short-term memory (BiLSTM) network. The proposed method was evaluated using knee joint kinematic and sEMG signals of fourteen subjects who performed running at 2 m/s speed. The proposed model's generalizability was tested for within- and cross-subject scenarios, and compared with standard LSTM and multi-layer perceptron (MLP) networks in terms of normalized root-mean-square error and correlation coefficient evaluation metrics. Based on the statistical tests, the proposed AM-BiLSTM model significantly outperformed the LSTM and MLP methods in both within- and cross-subject scenarios (p<0.05) and achieved state-of-the-art performance.**

***Clinical Relevance*— The promising results of this study suggest that the AM-BiLSTM model may be potential for continuous cross-subject estimation of lower limb kinematics during running in controlling sEMG-driven exoskeleton robots oriented to rehabilitation training.**

I. INTRODUCTION

Running is one of the most popular leisure exercise activities [1], which benefits both human physical and mental health [2]. Neurological impairments and joint diseases such as knee osteoarthritis can negatively affect performing important locomotion activities such as walking and running [3], [4] and, as a result, reduce the quality of life. Human-machine interaction (HMI) based exoskeleton robots can help patients to restore normal motor function and perform essential movements [5]. Surface electromyogram (sEMG) signal that contains active muscles' physiological information is essential for controlling HMI-based exoskeleton robots because it is recorded 20 to 200 milliseconds before the initiation of limb motion [6].

Fundamentally, two different control approaches can be used to realize the interaction between the patient and robot via sEMG signals, which are myoelectric-pattern recognition (MPR) based control and simultaneous and proportional control (SPC). In the MPR method, sEMG is used as a switch signal and different locomotion modes are recognized automatically from it using feature extraction techniques and classification algorithms. For instance, the work presented in [7] implemented the principal component analysis method on sEMG signals for feature extraction and classified gait phases with over 90% accuracy using the support vector machine algorithm. In another work conducted by Qin et al. (2020) using an effective feature extraction technique, the accuracy of the classification reached 96% [8]. However, the MPR method is discrete in nature, and through this approach, the smooth coordination between patient and robot as well as the fine-grained control are greatly affected, especially in complex locomotion tasks associated with the lower limbs [9]. Thereby, the MPR approach has limited application in real-world scenarios [10]. The SPC is the alternative approach, which can offer a more natural solution to control exoskeleton robots while performing complex activities, such as running, through continuous motion estimation.

In previous research, traditional machine learning algorithms such as multi-layer perceptron (MLP) [11], random forest [12], and support vector regression [13] are used to continuously estimate lower limbs' motions from sEMG signals. However, these classic models fail at capturing time series' long-term dependencies which leads to lower prediction accuracy compared to deep neural networks (DNNs) [14]. Biomechanical signals (including sEMG) are time series with a sequential structure, and DNNs can capture long-term dependencies and model the long-term relationships between the input and target data [15]. Among the most promising DNNs, the long short-term memory (LSTM) and gated recurrent unit (GRU) networks have been exploited to estimate knee joint angles via sEMG signals [16-18]. For instance, Ref. [16] employed an LSTM model for continuous estimation of knee joint angle during level-ground walking using sEMG signals. In Ref. [18], a hybrid framework incorporating a GRU network and a convolutional neural network was purposed for continuous estimation of knee joint angles via sEMG signals, and similar to Ref. [16], the prediction was carried out during the walking scenario. However, none of the previous studies focused on controlling exoskeleton robots during running, an essential movement which is more complex than walking, and plays an important role in maintaining physical and mental health. Furthermore, in both literature [16] and [18] the prediction of the knee joint angle is carried out only in the within-subject scenario, and the proposed model's generalization ability has not been investigated in the cross-subject scenario which is a more challenging testing level and could aid the practical deployment of the SPC scheme in real-world scenarios.

O.W. Samuel and A.A. McEwan (email: a.mcewan@derby.ac.uk) is with the School of Computing and Engineering, University of Derby, Derby, DE22 3AW, United Kingdom (Correspondence: e-mail: o.samuel@derby.ac.uk).
O.W. Samuel and G. Li (e-mail: gl.li@siat.ac.cn) are also with the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences (CAS), Shenzhen, Guangdong 518055, China.

A.R. Zangene and A. Abbasi are with the Department of Sports Biomechanics, Kharazmi University, Tehran, Iran (email: std_alireza.reyzan@khu.ac.ir; abbasi@khu.ac.ir).
K. Nazarpour is with the Edinburgh Neuroprosthetics Laboratory, School of Informatics, The University of Edinburgh, Edinburgh EH8 9AB, UK (email: kianoush.nazarpour@ed.ac.uk).

To address the aforementioned shortcomings of previous research, this study aimed to estimate knee joint angle via sEMG signals during complex locomotion activity such as running. A major concern related to the control of sEMG-driven exoskeleton robots is large EMG cross-subject variability due to subject-specific differences in muscles' physiological characteristics. Hence, we considered within-subject as well as cross-subject testing scenarios to predict the knee joint angles via sEMG signals. For accurate estimation in the presence of subject-specific differences, we propose a novel deep learning-based model (AM-BiLSTM) that integrates an attention mechanism (AM) [19] and a bidirectional LSTM (BiLSTM) network [20]. The BiLSTM network is an updated version of the standard LSTM, which in addition to processing the signal in the forward direction, also processes the signal in the reverse direction. This architecture leads to learning long-term dependencies more effectively, and as a result, will improve the performance of the network compared to standard (unidirectional) LSTM [21]. Moreover, the AM was adopted to further improve the performance of the BiLSTM network by assigning weights to the relevant features and highlighting critical information. For comparison purposes, previously used approaches including LSTM and MLP networks are also implemented in this study.

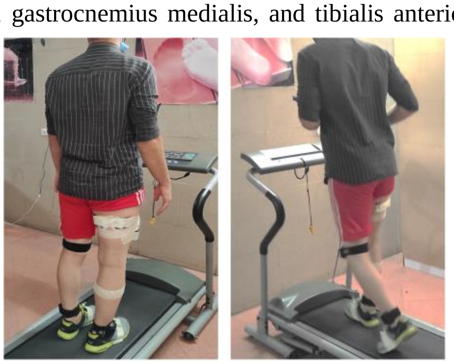
II. METHODS

A. Participants Information

A total of fourteen healthy male subjects (age: 24.42±9.01; mass: 82.31±12.10 kg; height: 178.28±4.46 cm; BMI: 25.85±3.09) with no lower-limb disabilities were recruited in this study. All subjects provided informed written consent before the experiment. The Institutional Review Board of Sport Sciences Research Institute of Iran approved the ethical review for this study with a reference identification number of *IR.SSRI.REC.1400.1200*.

B. Data Collection and Preprocessing

The dataset of this study was collected during treadmill running at 2 m/s speed (see Fig. 1). Each trial was repeated three times and a 10-minute interval was considered between each trial to prevent muscle fatigue. During each trial, sEMG signals from six muscles were recorded for one minute (after stabilization of running speed) with the NORAXON wireless EMG system (myoMUSCLE, NORAXON USA Inc., Scottsdale, Arizona). The sampling frequency of the myoMUSCLE system was 3000 Hz and vastus medialis, rectus femoris, semitendinosus, biceps femoris, gastrocnemius medialis, and tibialis anterior were the selected muscles. Simultaneously, the angle of the knee joint was recorded using the NORAXON myoMOTION motion analysis system (myoMOTION, NORAXON USA Inc., Scottsdale, Arizona) with a sampling rate of 200 Hz.



(a) (b)

Figure 1. Data collection from one of the subjects. (a) Subject in anatomical position; (b) Subject at running.

Subsequently, the raw knee joint angles were filtered using a low-pass 6 Hz filter (4th order zero-phase Butterworth filter) [22]. Also, a 20-500 Hz band-pass filter (4th order zero-phase Butterworth filter) was applied to the raw sEMG signals to remove other inherent noises [22].

C. Feature Extraction from sEMG Signals

sEMG is a non-stationary signal that has highly complex time and frequency characteristics and its nature change over time [23]. Therefore, to analyze sEMG signals, time-frequency feature extraction techniques such as wavelet transform and short-time Fourier transform (STFT) are more powerful methods compared to the classical time and frequency feature extraction methods. We used the discrete wavelet transform (DWT) for feature extraction from sEMG signals, which has a better performance compared to the STFT due to its variable-sized windowing technique. The DWT is a dual channel sub-band coding technique that decomposes the input signal into the approximation (low-frequency part) (cA) and detail (high-frequency part) (cD) subsets. The approximation coefficients that provide significant information about the original sEMG signal can be used as the predictor features of neural networks [24].

Using the DWT method, the approximation subset was calculated by passing the input signal ($x(k)$) through a low-pass filter with impulse response $g(n)$.

$$Z_{low} = \sum_{k=1}^{L} x(k)\, g(2n-k) \tag{1}$$

Simultaneously, the detail subset was calculated by passing the input signal through a high-pass filter with impulse response $h(n)$.

$$Z_{high} = \sum_{k=1}^{L} x(k)\, h(2n-k) \tag{2}$$

The decomposition of sEMG signal can be repeated iteratively until the desired approximation coefficients are obtained. In this paper, the de-noised sEMG signals were

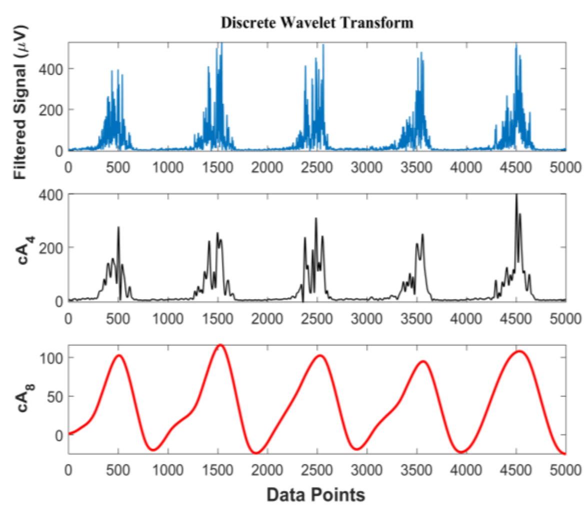


Figure 2. Feature extraction from sEMG signals using discrete wavelet transform.

decomposed into eight levels, and we extracted the desired time-frequency feature from the approximation subset of eighth-level DWT decomposition (cA8) (see Fig. 2).

D. Proposed AM-BiLSTM Model Description

1) Bidirectional LSTM Network: LSTM was first proposed in 1997 by Hochreiter and Schmidhuber to address the gradient explosion and gradient disappearance problems of classic recurrent neural networks [25]. In the BiLSTM network, which is the improved version of the standard LSTM, the input sequence is used twice in the training process, first using a forward LSTM ($\overrightarrow{LSTM}$) and then using a backward LSTM ($\overleftarrow{LSTM}$). The memorizing process of both $\overrightarrow{LSTM}$ and $\overleftarrow{LSTM}$ is adjusted by three special gates named forget gate (f_t), input gate (i_t), and output gate (o_t) which can be represented by the following equations (3)–(7): (In all of these equations, the subscripts *t* and *t-1* represent the present and previous time steps, respectively.)

$$f_t = sigmoid(\mathbf{W}_{\boldsymbol{f}}.[h_{t-1}, x_t] + b_f) \tag{3}$$

$$i_t = sigmoid(\mathbf{W}_{\boldsymbol{i}}.[h_{t-1}, x_t] + b_i) \tag{4}$$

$$o_t = sigmoid(\mathbf{W}_{\boldsymbol{o}}.[h_{t-1}, x_t] + b_o) \tag{5}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot (tanh(\mathbf{W}_{\boldsymbol{c}}.[h_{t-1}, x_t] + b_c)) \tag{6}$$

$$h_t = o_t \odot tanh(C_t) \tag{7}$$

Where *x* is the input sequence, *h* is the hidden state, and *C* is the cell state. Also, **W** and *b* stand for the weight matrix and bias values.

In the BiLSTM network, first, the forward hidden states ($\vec{h}_i$) are calculated (($\vec{h}_1$,…, $\vec{h}_N$) = $\overrightarrow{LSTM}(x_1,\ldots, x_N)$). Next, the input sequence is processed in the reverse form and the backward hidden states ($\overleftarrow{h}_i$) are calculated (($\overleftarrow{h}_N$,…, $\overleftarrow{h}_1$) =$\overleftarrow{LSTM}(x_N,\ldots, x_1)$). In the final stage, by concatenating the forward and backward hidden states, the final hidden state ($\tilde{h}_i$) is obtained as follows.

$$\tilde{h}_i = [\, \vec{h}_i , \overleftarrow{h}_i \,], i = 1, 2, 3, \ldots, N \tag{8}$$

To further enhance the performance of the BiLSTM network in predicting the target signals, we implemented and integrated a variant of the attention mechanism to the BiLSTM module, and the description is as follows:

2) Attention Mechanism: Attention is a new concept in deep learning which has become one of the most important topics in the field [26]. The *RNNsearch* was the first attention-based model, proposed by Bahdanau et al. (2014) for machine translation [19]. Since then, different attentive neural networks have achieved state-of-the-art performances in other fields of expertise such as predicting electrical load [28] and stock price [29], which is due to highlighting key information when handling vast amounts of data by AM. The AM can be represented by the following equations (9)–(11).

First, the final hidden states were applied to a one-layer perceptron.

$$u_i = tanh(\mathbf{W}\tilde{h}_i + b) \tag{9}$$

Where, **W** and *b* stand for the weight matrix and the bias values of the single-layer perceptron, and u_i is the hidden representation of final hidden states.

In the next step, the attention weight (α_i) is calculated by performing SoftMax normalization on the corresponding u_i .

$$\alpha_i = \frac{\exp(u_i^{\mathrm{T}} u_s)}{\sum_i \exp(u_i^{\mathrm{T}} u_s)} \tag{10}$$

Where u_s is the weight vector, which is randomly initialized.

Finally, the attention layer output (v) is calculated through the following equation.

$$v = \sum_{i=1}^{N} \alpha_i \tilde{h}_i \tag{11}$$

Where *N* is the length of the sequence.

E. Performance Indices and Testing Scenario

For a comprehensive investigation, the prediction accuracy of the proposed AM-BiLSTM model was tested in within- and cross-subject scenarios and compared with the performance of notable existing approaches (LSTM and MLP). In the within-subject scenario, each model (AM-BiLSTM, LSTM, and MLP) was trained using 75% of data corresponding to a specific subject and tested with 25% remaining data of the same subject. This process was repeated four times for each subject. In the cross-subject scenario, each model was trained with the corresponding data of thirteen subjects out of fourteen and tested for the unseen subject. This process was repeated fourteen times to determine the performance of the corresponding model for all subjects.

Herein, two standard evaluation indices, the normalized root-mean-square error (*NRMSE*) and correlation coefficient (CC), were used to quantify the performance of all models. The *NRMSE* and CC metrics are defined as:

$$NRMSE = \frac{\sqrt{\frac{\sum_{t=1}^{N}(\tilde{\theta}_t - \theta_t)^2}{N}}}{\theta_{max} - \theta_{min}} \times 100 \tag{12}$$

$$\mathrm{CC} = \frac{\frac{1}{N}\sum_{t=1}^{N}(\theta_t - \bar{\theta})(\tilde{\theta}_t - \bar{\tilde{\theta}})}{\sqrt{\frac{1}{N}\sum_{t=1}^{N}(\theta_t - \bar{\theta})^2}\sqrt{\frac{1}{N}\sum_{t=1}^{N}(\tilde{\theta}_t - \bar{\tilde{\theta}})^2}} \tag{13}$$

Where $\tilde{\theta}_t$ is the prediction at *t*, θ_t is the real knee joint angle at *t*, and *N* is the sequence length.

F. Statistical Analysis

To identify significant differences between the performance of the proposed method and previously used approaches including LSTM and MLP, Friedman test and Wilcoxon signed-rank test with Bonferroni correction for multiple comparisons were applied ($p<0.05$).

III. RESULTS AND DISCUSSION

The proposed AM-BiLSTM model was built in TensorFlow platform. Also, the LSTM and MLP networks were implemented in MATLAB (The MathWorks Inc., Natick, MA, USA). Using the grid search method, the optimal training option for each model was determined and the corresponding hyperparameters were tuned. Due to the faster convergence rate than other optimization algorithms, the Adam algorithm was selected to minimize the cost function of each model.

Fig. 3 shows the actual knee joint angles against the estimated knee joint angles when using each of the three methods. In the within-subject scenario, both AM-BiLSTM and LSTM models provided accurate trends (the proposed achieved better results) and obviously performed better than the MLP network (Fig. 3(a)). In the cross-subject scenario, the MLP network lost its fitting ability to a great extent and the output of this network is accompanied by a high amount of distortion (Fig. 3(b)). Also, the LSTM's performance decreased in cross-subject tests, especially in predicting the peaks of target signals (Fig. 3(b)). This decrease in LSTM's performance in cross-subject tests is due to large subject-specific differences in sEMG signals. However, despite the individual differences in cross-subject tests, the performance of the proposed AM-BiLSTM model is still high and is markedly better than both comparative networks (Fig. 3(b)), which indicates the high generalizability of the proposed model compared to other methods.

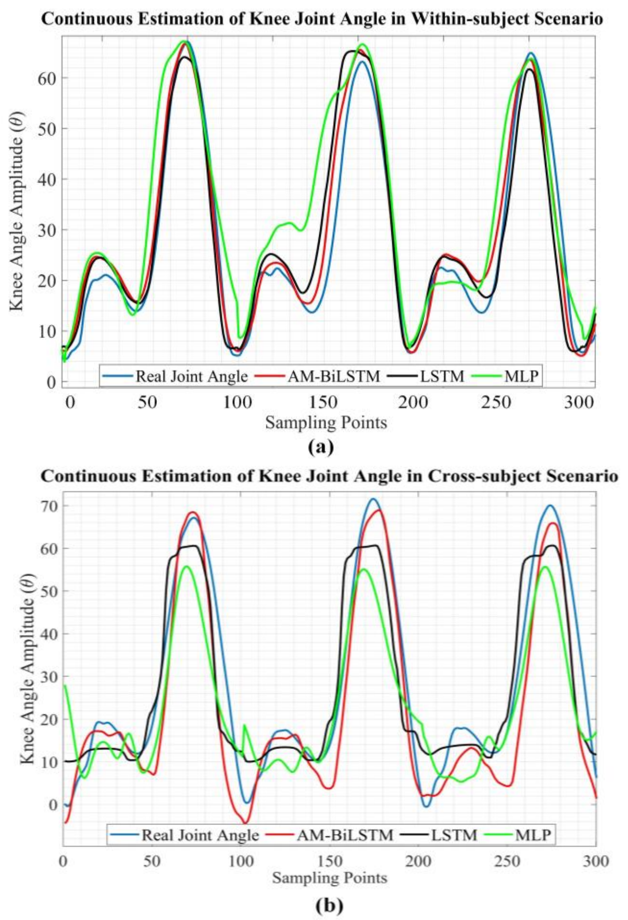


Figure 3. Estimated angles in (a) within-subject scenario and (b) cross-subject scenario. The solid black, red, blue, and green lines are the actual knee joint angles, and the outputs of the AM-BiLSTM, LSTM, and MLP networks, respectively.

The average values across all subjects corresponding to the *NRMSE* and CC metrics related to the within- and cross-subject tests are shown in Table I. According to Table I, the results of the within-subject evaluation scenario indicated strong correlations between the AM-BiLSTM outputs and the actual knee joint angles and relatively low prediction error (CC=0.984 and *NRMSE*=4.807%). The prediction accuracy of the proposed model decreased slightly in the cross-subject scenario (CC=0.928 and *NRMSE*=7.101%), nevertheless, in both evaluation protocols, the accuracy of the proposed model was significantly higher than LSTM and MLP networks ($p<0.05$).

Overall, our model significantly outperformed both LSTM and MLP networks in the within-subject testing scenario ($p<0.05$) as well as the cross-subject scenario ($p<0.05$) and achieved state-of-the-art performance, which indicates the high potential of this model for practical deployment in real-world scenarios. The prediction robustness and higher generalization ability of the proposed model in comparison to the standard LSTM and MLP networks can be due to the importance of highlighting pivotal information by the AM layer and incorporation of future information in the BiLSTM network's training process.

Models	***Within-subject***		***Cross-subject***	
	CC	***NRMSE***	**CC**	***NRMSE***
MLP	0.930	7.213	0.810	12.04
LSTM	0.951	6.320	0.887	9.131
AM-BiLSTM	**0.984***	4.807	**0.928***	7.101

Table I. Average CC and *NRMSE* values of all models across fourteen subjects for within- and cross-subject scenarios.

* Indicates a significant difference between the performance of the proposed model and LSTM and MLP networks based on the results of statistical tests.

IV. Conclusion

Running is a periodic and continuous movement for the human lower limb, which plays an important role in everyday life and exercise activities. Accurate estimation of human knee joint angle during running activity via sEMG signals can be used to intuitively control exoskeleton robots in the rehabilitation training for patients with neurological impairments and joint disease, conditions that, to the best of our knowledge, have not been considered to date. To this end, we proposed a novel deep learning-driven model that integrates attention mechanism and bidirectional LSTM network (AM-BiLSTM) for accurate prediction of target knee joint angle signals via muscles' electrical activity that is represented by sEMG signals. To investigate the proposed model's generalizability, within- and cross-subject testing scenarios were carried out, and the performance of the proposed AM-BiLSTM model was compared with LSTM and MLP networks in each scenario in terms of CC and *NRMSE* metrics. According to the statistical tests, the proposed model significantly outperformed both standard LSTM and MLP networks in each testing scenario ($p<0.05$) and achieved state-of-the-art performance. Findings from this study may aid the practical deployment of intuitive deep learning-based control schemes for lower limb rehabilitation robotic systems.

Furthermore, our future work will focus on testing the proposed method on running movement at various speeds and slopes, which is often performed in practice.

References

[1] W. van Mechelen, "Running injuries: A review of the epidemiological literature", *Sports Med.*, vol. 14, no. 5, pp. 320-335, Nov. 1992.

[2] F. Oswald, J. Campbell, C. Williamson, J. Richards, and P. Kelly, "A scoping review of the relationship between running and mental health," *Int. J. Environ. Res. Public Health*, vol. 17, no. 21, p. 8059, 2020.

[3] S. Kyeong, W. Shin, M. Yang, U. Heo, J. Feng, and J. Kim, “Recognition of walking environments and gait period by surface electromyography,” *Front. Inf. Technol. Electron. Eng.*, vol. 20, no. 3, pp. 342–352, 2019.
[4] 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, 2016.
[5] L. Wang, Z. Du, W. Dong, Y. Shen and G. Zhao, "Hierarchical human machine interaction learning for a lower extremity augmentation device", *Int. J. Social Robot.*, vol. 11, no. 1, pp. 123-139, Jan. 2019.
[6] T. Teramae, T. Noda and J. Morimoto, "EMG-based model predictive control for physical human–robot interaction: Application for assist-as-needed control", *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 210-217, Jan. 2018.
[7] D. C. Toledo-Pérez, M. A. Martinez-Prado, R. A. Gómez-Loenzo, W. J. Paredes-Garcia, and J. Rodriguez-Reséndiz, “A study of movement classification of the lower limb based on up to 4-EMG channels,” *Electronics*, vol. 8, no. 3, p. 259, 2019.
[8] P. Qin and X. Shi, “Evaluation of Feature Extraction and Classification for Lower Limb Motion Based on sEMG Signal,” *Entropy 2020, Vol. 22, Page 852*, vol. 22, no. 8, p. 852, Jul. 2020, doi: 10.3390/E22080852.
[9] J. Wang, L. Wang, S. M. Miran, X. Xi and A. Xue, "Surface electromyography based estimation of knee joint angle by using correlation dimension of wavelet coefficient", *IEEE Access*, vol. 7, pp. 60522-60531, 2019.
[10] 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.
[11] J. Chen, X. Zhang, Y. Cheng and N. Xi, "Surface EMG based continuous estimation of human lower limb joint angles by using deep belief networks", *Biomed. Signal Process. Control*, vol. 40, pp. 335-342, Feb. 2018.
[12] Z. Li, X. Guan, K. Zou and C. Xu, "Estimation of knee movement from surface emg using random forest with principal component analysis", *Electron*, vol. 9, no. 1, pp. 43-56, 2020.
[13] Q. L. Li, Y. Song and Z. G. Hou, "Estimation of lower limb periodic motions from sEMG using least squares support vector regression", *Neural Process. Lett.*, vol. 41, no. 3, pp. 371-388, 2015.
[14] Y. Chen *et al.*, “A Continuous Estimation Model of Upper Limb Joint Angles by Using Surface Electromyography and Deep Learning Method,” *IEEE Access*, vol. 7, pp. 174940–174950, 2019, doi: 10.1109/ACCESS.2019.2956951.
[15] R. Zhao, J. Wang, R. Yan and K. Mao, "Machine health monitoring with LSTM networks", *10th International Conference on Sensing Technology (ICST)*, pp. 1-6, 2016.
[16] X. Ma, Y. Liu, Q. Song, and C. Wang, “Continuous Estimation of Knee Joint Angle Based on Surface Electromyography Using a Long Short-Term Memory Neural Network and Time-Advanced Feature,” *Sensors 2020, Vol. 20, Page 4966*, vol. 20, no. 17, p. 4966, Sep. 2020, doi: 10.3390/S20174966.
[17] A. R. Zangene, A. Abbasi, and K. Nazarpour, “Estimation of Lower Limb Kinematics during Squat Task in Different Loading Using sEMG Activity and Deep Recurrent Neural Networks,” *Sensors 2021, Vol. 21, Page 7773*, vol. 21, no. 23, p. 7773, Nov. 2021, doi: 10.3390/S21237773.
[18] X. Wang et al., "SEMG-based consecutive estimation of human lower limb movement by using multi-branch neural network", *Biomed. Signal Process. Control*, vol. 68, Jul. 2021.
[19] D. Bahdanau, K. Cho and Y. Bengio, "Neural machine translation by jointly learning to align and translate", *arXiv:1409.0473*, 2014.
[20] A. Graves and J. Schmidhuber, "Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures," Neural Networks, vol. 18, no. 5-6, pp. 602-610, June/July 2005.
[21] S. Siami-Namini, N. Tavakoli and A. S. Namin, "The performance of LSTM and BiLSTM in forecasting time series", *Proc. IEEE Int. Conf. Big Data (Big Data)*, pp. 3285-3292, Dec. 2019.
[22] G. Robertson, G. Caldwell, J. Hamill, G. Kamen and S. Whittlesey, Research Methods in Biomechanics, Champaign, IL, USA:Human Kinetics, 2013.
[23] M. R. Canal, "Comparison of wavelet and short time Fourier transform methods in the analysis of EMG signals", *J. Med. Syst.*, vol. 34, no. 1, pp. 91-94, 2010.
[24] A. Phinyomark, C. Limsakul and P. Phukpattaranont, "Application of wavelet analysis in EMG feature extraction for pattern classification", *Meas. S ci. R ev.*, vol. 11, pp. 45-52, 2011.
[25] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997.
[26] Z. Niu, G. Zhong and H. Yu, "A review on the attention mechanism of deep learning", *Neurocomputing*, vol. 452, pp. 48-62, Sep. 2021.
[27] K. H. Wu, J. Wu, L. Feng, B. Yang, R. Liang, S. Q. Yang, et al., "An attention-based CNN-LSTM-BiLSTM model for short-term electric load forecasting in integrated energy system", *Int. Trans. Elect. Energy.*, vol. 31, no. 1, Jan. 2020.
[28] W. Lu, J. Li, J. Wang, and L. Qin, ‘‘A CNN-BiLSTM-AM method for stock price prediction,’’ Neural Comput. Appl., vol. 33, no. 10, pp. 4741–4753, May 2021.