Journal of Sustainable Engineering and Technology 2(1) 2025 64 – 72

JOURNAL OF SUSTAINABLE ENGINEERING AND TECHNOLOGY

Journal homepage: www.joset.com.ng

Long Short-Term Memory Neural Network Knee-Joint Angle Estimation using sEMG Signal and Time-Advanced Feature

Ahmed Usman[1], Bashir Muhammad[1], Hafiz Ibrahim Hassan[2], Rabiu Aliyu Abdulkadir[1]

[1]Department of Electrical Engineering, Faculty of Engineering, Aliko Dangote University of Science and Technology, P.M.B. 3244, Wudil, Kano, Nigeria

[2] Department of Mechatronics Engineering, Ahmadu Bello University Zaria 810211, Kaduna Nigeria, Kaduna State

ARTICLE INFO

Article History
Received: 24 Nov 2024
Accepted: 27 July 2025
Published: 31 July 2025

Corresponding Author:
Bashir Muhammad
Aliko Dangote University of Science and Technology, Wudil
muhdbashir1423@gmail.com

ABSTRACT

The ability of wearable robots to work based on the wearer's intension is a major challenge in the control and design of the robots. Previous work integrated a wearable reboot with Long short-term memory (LSTM) model Using Root mean square (RMS) feature to estimate the knee joint angle for rehabilitation purpose. However, other Time domain (TD) and Time advanced features (TAD) such as Mean average value (MAV), Variance (VAR) and Standard deviation were not considered. This can improve the continuously estimation of the knee-joint angle. In this research, Long Short-Term Memory (LSTM) Neural Network is utilized to continuously estimate the knee-joint angle from surface electromyography sEMG data so as to determine the amount of compensation needed, with Back Propagation Neural Network (BPNN) as comparison model. The sEMG signals TD and TAD features were extracted with the later taking into consideration that the production of sEMG signal starts 20-200 ms before the corresponding muscle action. RMS, MAV, VAR and STD were extraction from each domain and were used separately to train the LSTM model as well as the BPNN model. The performance of each was evaluated based on Root-Mean-Squared Error (RMSE). The extracted features of the sEMG signal were used as input and the knee-joint angle as output. Overall results show that LSTM using TAD across all features extracted shows improved performance (lower RMSE) by an average of 3.2566^0 when compared to BPNN model with an average lower RMSE of 6.1945^0. This implies that using other TAD features improves model performance.

Keywords: *sEMG · LSTM · BPNN · TAD · RMSE*

1. Introduction

Aging population is one of the most critical challenges facing the society in the past decades [1]. Aging sometimes leads to many disorders, one of which is arthritis which is one of the major causes of disability among adults. Arthritis is the swelling of the knee-joint that causes pain and limits movement of the affected area. Wearable robots known as exoskeleton robots (active mechanical device) are anthropomorphic in nature, worn by patient and fits closely to the body to work in concert with the wearers movement [2]. Wearable robots are designed for the purpose of power augmentation, rehabilitation and assistance. No matter the design purpose, the key issue when dealing with wearable robot especially for medical purpose is the ability of the robot to operate based on the wearers' intention. The control aspect determines the purpose of the design of the robot. The control mechanism or architecture integrates three domains depending on the application for example the rehabilitation architecture is mostly applied to medical field to restore physical, mental and sensory capabilities that were lost due to injuries or disease. The assistance control deals with our daily activities of walking and standing while the augmentation control augments our physical and normal activities like raising and carrying objects.

Surface electromyography (sEMG) signals, which reflect neuromuscular activity, are widely used as control inputs for wearable robots due to their higher stability and resistance to interference compared to signals like EEG, making them effective for joint angle estimation. These signals, representing neuronal force from the nervous system, have a near one-to-one relationship with muscle activity when electrodes are correctly placed. sEMG can be leveraged in advanced robotic control systems using deep learning techniques [3]. However, signal consistency is difficult even for repeated motions by the same person due to the non-linear and complex behaviour of muscles. AI-based control systems are therefore employed to handle this variability and uncertainty [2]. Furthermore, since sEMG signals are generated 20–200 ms before muscle contraction, they provide a useful lead time for predicting motion compared to physical measures like joint angles and angular velocity.

Many studies have been carried out for joint angle prediction using AI models. Triwiyanto *et. al.* [3] uses pattern based recognition with Kalman's filter and zero-crossing feature to estimate the elbow joint angle. Song *et al.* [4] uses LSTM and Multilayer Perception (MLP) for pattern based classification to estimate the joint angle using both frequency and time domain features. Yang *et al*. [5] used BPNN generalized regression neural network and Least-Square Support Vector Machine (SVM) regression as classifiers on three different features. Ma *et al.* [6] uses LSTM and Artificial Neural Network (ANN) to estimate the knee-joint angle on five healthy individuals with only RMS feature. Li *et al.* [7] used Random Forest (RF) with Principle Component Analysis (PCA) on data from six subject to train a new method of combining RF and PCA. Michieletto *et al.* [8] . Zhang *et al.* [9] used LSTM and transfer learning to predict lower limb torque with low prediction error in intra-subject scenarios. Truong *et al.* [10] used LSTM to estimate lower limb joint-angle movements with the model showing bias for angle movement when compared to knee. Based on previous studies, only Ma et al. [7] utilized time-domain (TAD) features and achieved high model performance using the RMS feature. J. Wang *et al*[13] applies a novel "closed-loop" sEMG-driven state-space model to estimate the elbow joint angle in the human upper limb during continuous motion. H. Teng *et al*. [14] Combines Graph Convolutional Neural Networks (GCN) with Long Short-Term Memory (LSTM) networks to predict joint angles. However, other TAD features such as Mean Absolute Value (MAV), Variance (VAR), and Standard Deviation (STD) were not explored. In this study, MAV, VAR, and STD are extracted as TAD features to evaluate the performance of an LSTM model, using RMSE as the evaluation metric.

2. Methodology

2.1 Data Acquisition Processing, normalization and Feature Extraction

The data used was acquired from an EMG based Neural Network (NN) approximation for knee-joint movement to assist osteoarthritis patient, were sEMG was obtained from fourteen healthy individuals all male with only the swing session of the right leg considered [11]. The EMG signals contain unwanted signal which needs to be eliminated. The raw EMG signals were rectified to cancel all the negative signals that is, the negative signals were moved up and added to the positive signal (i.e. full wave rectification equation).

The data was normalized using min-max normalization as shown in equation (1). The range is restricted to $|01|$ as all the EMGi data values range from 0 to 1. This will prevent the function from saturating. To have the error (deviation) in degree, the data of the knee-joint angle was not normalized because during the experiment, the knee joint angle sensor was calibrated from 0-180° with 90° being the neutral position.

Normalizing the data of the sEMGi signal makes the NN training more effective and prevents the function from saturation; it improves the training speed and increases the gradient value of the training dataset.

$$EMG_n = min' + \frac{(EMG_i - min)(max' - min')}{max - min} \tag{1}$$

Where EMG_n represent the normalized data, EMG_i represent the rectified EMG data, $max$and min represent the maximum and minimum values of EMG data, $max'$and min'represent the max and min range of normalization.

Each of the subjects contains two input data with a size of 300 samples (2X300) and a single output data of (1x300). The data of each subject was trained so as to obtain a network with least RMSE. The feature extraction from rectified EMG Signals is necessary because the rectified EMG signals are still not suitable as input to machine because minimization of the non-reproducible nature of the sEMG is still needed for application to machine learning. Features extraction methods of the sEMG can be divided into three: time domain, time-frequency domain and frequency domain. The time domain is used here because they can easily be extracted without any transformation. The RMS, MAV, VAR, and STD features were extracted for both TD and TAD using the sliding window technique or function with the window width set to 10.

The generation of the sEMG is 20-200ms earlier than the corresponding muscle action but the data for both sEMG and angle are collected simultaneously. This implies that the angle data at that time corresponds to the sEMG of previous movement. The original sEMG signal sequence and angle sequence are moved backward, and then the moved Sequences were delayed relative to the raw sEMG signal and raw angle signal. Therefore, the raw sEMG signal sequence was regarded as a time-advanced sEMG signal sequence, which corresponded to the delayed angle signal sequence [6].

Equation (*2*) – (*5*) [12] depicts the four features extracted for TD while (*6*) – (9) [12] depicts features for TAD. Moreover, equation (*10*) [12] represents the joint angle sequence moving backwards or lagging behind the raw sEMG and equation (11) [12] represents the rectification equation of the rEMG (rEMG) signals.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} EMGi^2} \tag{2}$$

$$MAV = \frac{1}{N}\sum_{i=1}^{N} EMGi - m(x) \vee \tag{3}$$

$$VAR = \frac{1}{N-1}\sum_{i=1}^{N}\left(EMGi - m(x)\right)^2 \tag{4}$$

$$STD = \sqrt{VAR} \quad . \tag{5}$$

The features extracted for TAD are:

$$RMS(k) = \sqrt{\frac{1}{N}\sum_{i=1}^{N+k} EMGi^2} \quad (6)$$

$$MAV(k) = \sum_{i=1}^{N+k} EMGi - m(x) \vee \quad (7\}$$

$$VAR(k) = \sum_{i=1}^{N+k}\left(EMGi - m(x)\right)^2 \quad (8)$$

$$STD(k) = \sqrt{VAR(k)} \quad (9)$$

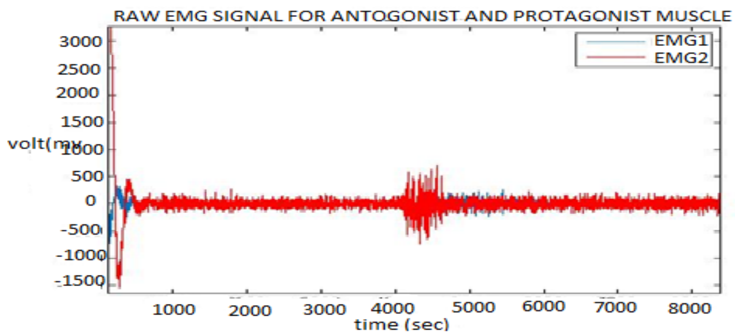
$$Angle(h) = \sum_{(h-1)+l}^{h} angle(i) \quad (10)$$

The rEMG were rectified using the full wave rectification formula as:

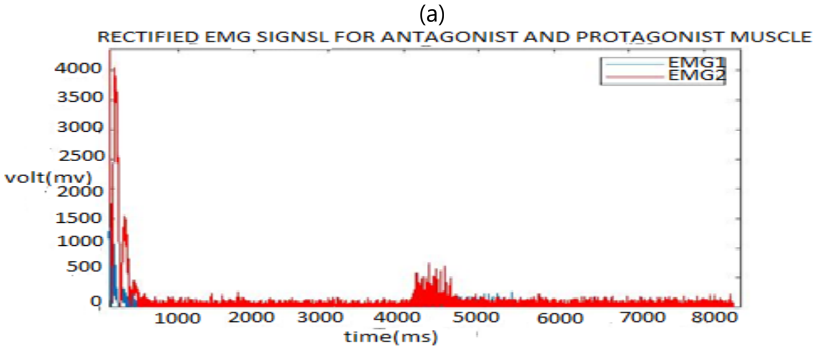
$$EMGi = rEMG \mid \quad (11)$$

Where EMGi represent the rectified rEMG signal, N represent the width of the sliding windows (a technique or method used to efficiently solve problems that involve defining a window or range in the input data arrays or strings and then moving that window across the data to perform some operation within the window) which was set to 10 in this feature extraction process, m(*x*) represents average mean of the data set. For the TAD, k represents the time advanced feature extracted from the EMGi signals that is the number of points the sEMG signal is moved backwards and i represents the number of points the angle sequence moves backwards.

Figure 1 depicts the rEMG and EMGi of subject 4 (the healthy individual tagged as number 4 during data collection) respectively.



(a)



(b)

Figure 1 EMG signals (a) Raw (b) Rectified

2.2 BPNN Architecture

The extracted features from the EMG signals were used as inputs, and the knee joint angle served as the output. The hyper-parameters tuned to optimize the accuracy of the BPNN included the number of hidden layers, number of hidden neurons, activation function(s), and optimization algorithm. For the training, the BPNN was structured with single hidden layer and 10 numbers hidden neurons. With hyperbolic tan-sigmoid (tansig) activation function used for the input and linear (purelin) activation function used for the output. The accuracy of the BPNN depends on the updates of the parametric weight of the network by propagating the errors back through the system causing an update of the weights for the next iteration to minimize the Root mean square error (RMSE) values and achieves a regression of one or almost one. The Levenberg-Marquardt training algorithm which from literature performs on function problems was deployed to optimize MSE error cost function by updating the weight and biases of the neural networks. Figure 2 depicts the BPNN architecture for the training.

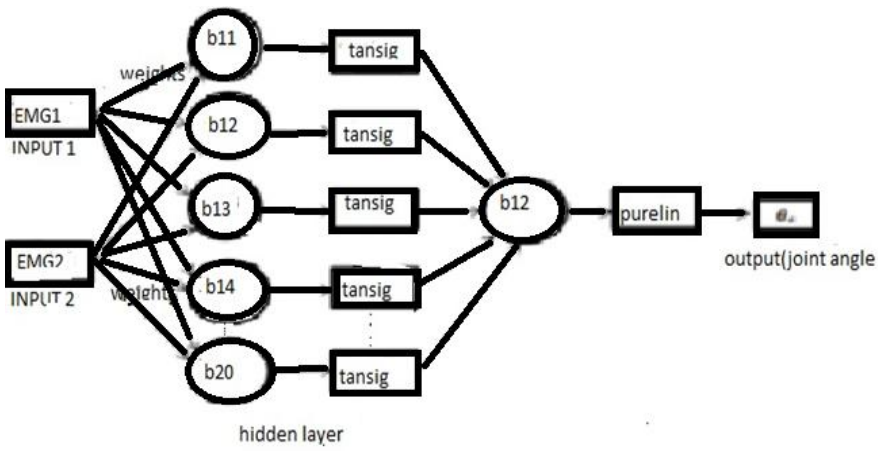


Figure 2. The BPNN Structure for the Training [14]

2.3 LSTM Architecture

A 100-layer LSTM network was used for training with two input layers, 200 hidden nodes and one hidden node in the output layer that contains the knee-joint angle. The network is trained using supervised learning and the target vector is the knee-joint angle collected simultaneously with the sEMG. The Adam optimization algorithm with 250 epoch of training in a mini batch size of 50 to minimize the loss function and the learning rate was set to 0.02. The dropout method was used as a regularization method and the drop probability was set to 0.1. The amount of data in the training set and test set accounted for 80% and 20% of the dataset, respectively. Figure 3 shows the LSTM architecture for the training.

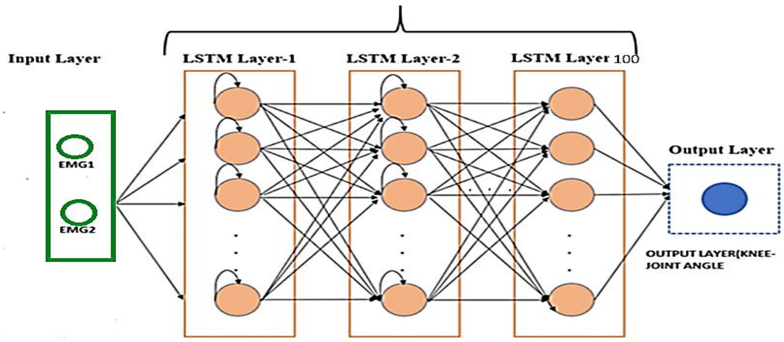


Figure 3. LSTM Structure for Training [6]

3. Results

Table 1 and 2 shows each subject's RMSE obtained with the TD and TAD for the four different features and their average values for the nine subjects. From Table 2, the average RMSE values obtained by

LSTM using TAD is 3.2566^0 for RMS feature, 3.8553^0 for MAV feature, 3.9919^0 for the VAR feature and 3.5742^0 for the STD feature. While from Table 1, that of BPNN using TAD were 6.1945^0 for RMS feature, 6.5285^0 for MAV feature, 6.5647 for VAR feature and 6.4883^0 for STD feature.

Table 1. BPNN RMSE for both Time Domain and Time-Advanced for Nine Subjects

SUBJECTS	**RMS**		**MAV**		**VAR**		**STD**	
	TD	**TAD**	**TD**	**TAD**	**TD**	**TAD**	**TD**	**TAD**
1	8.8632	8.0973	9.2015	9.7468	10.4881	9.9490	9.2195	6.1644
2	4.5088	3.2660	4.7958	4.1231	5.6569	5.0992	5.1962	4.5453
3	9.1652	7.5717	8.5633	6.4550	9.8488	6.3852	8.7367	6.0711
4	8.0746	2.4034	8.4774	2.6458	8.1853	3.1092	8.1240	5.7515
5	11.3871	6.6383	11.4891	7.5939	10.2470	4.2426	10.2956	5.2809
6	9.1942	6.3245	10.8166	7.5895	10.1651	8.6718	10.4243	8.7750
7	7.9372	6.2183	8.5047	7.7244	9.2195	8.6410	9.2103	6.4807
8	9.5304	8.9067	9.8656	9.0921	9.4163	8.0416	9.6610	8.4693
9	9.1287	6.3245	10.1324	3.7855	10.0499	4.9430	9.2736	4.8988
Average	8.6454±1.8402	6.1945±2.1229	9.0940±1.9324	6.5285±1.5180	9.2530±2.4752	6.5647±2.3592	8.9046±1.5627	6.2708±1.4709

Table 2. LSTM RMSE for both Time Domain and Time-Advanced for Nine Subjects

SUBJECTS	**RMS**		**MAV**		**VAR**		**STD**	
	TD	**TAD**	**TA**	**TAD**	**TD**	**TAD**	**TD**	**TAD**
1	2.9005	2.4321	4.0056	3.9498	4.8649	3.5233	3.5918	3.4610
2	4.7022	4.5051	4.4130	4.0907	4.3875	4.2221	4.2533	3.6908
3	4.1612	3.5989	4.4995	4.3304	4.2894	4.0669	3.9458	3.6172
4	3.7700	3.5494	4.9396	4.7954	4.1556	3.7637	4.2136	3.1779
5	3.0867	2.4969	3.9291	3.4384	4.0087	3.7282	3.6816	3.0535
6	3.4442	3.5217	3.1496	2.7487	3.1094	2.9449	4.1163	3.8621
7	4.3394	4.1213	4.2985	3.8502	4.4983	4.1967	3.5937	3.5797
8	2.7486	2.3485	3.1440	2.5185	6.5826	6.2437	3.3801	2.8784
9	2.9443	2.7357	5.5822	4.9755	3.3156	3.2379	5.0077	4.8474
Average	3.5663±0.7095	3.2566±0.7852	4.2179±0.7849	3.8553±0.8377	4.3569±1.0025	3.9919±0.9472	4.0315±0.4942	3.5742±0.5749

From Table 2, LSTM using TAD RMS versus LSTM using TD RMS, LSTM using TAD MAV versus LSTM using TD MAV, LSTM using TAD VAR versus LSTM using TD VAR and LSTM using TAD STD versus LSTM using TD STD, the average values were reduced by 8.68%, 8.60%, 8.38% and 11.34% respectively using the formula in Equation 12.

$$Percentage\ error = \frac{|TADa - TDa|}{TDa} \times 100 \qquad (12)$$

Where TADa is the average of the TAD RMSE for RMS, MAV, VAR and STD features respectively and TDa is the average time domain RMSE for RMS, MAV, VAR and STD feature respectively.

4. Discussion

This work evaluated the performance of TD and TAD features of lower limb EMG signal during normal gait. The feature having the least RMSE value was evaluated for both TD and TAD features. The features were used on BPNN network and LSTM to estimate the amount of knee-joint angle assistance needed by osteoarthritis patient.

The comparison of LSTM and BPNN models using TAD and TD features across the four features RMS, MAV, VAR, and STD shows consistent performance improvements with TAD features. Specifically, LSTM models using TAD features outperformed those using TD features with average RMSE reductions of 8.68% (RMS), 8.60% (MAV), 8.38% (VAR), and 11.34% (STD). When comparing LSTM and BPNN using both TAD and TD features, LSTM consistently yielded lower RMSE, with reductions reaching up to 58.75%. Moreover, BPNN models also showed better performance with TAD over TD features, with RMSE improvements ranging from approximately 28% to 30%, depending on the feature used. Overall, the results highlight the superiority of TAD features and the LSTM model in reducing prediction errors across all features extracted not only RMS. The features extracted in TAD are smoother than the features extracted in TD as depicted in Figure 4.

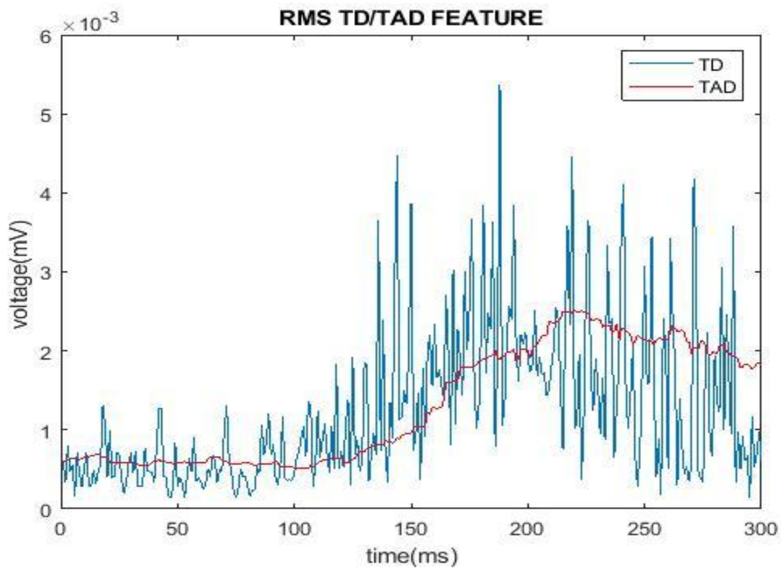


Figure 4. Sampled Signal for RMS TD/TAD Feature

5. Conclusion

The TAD features depict lower RMSE values across all features used when compared to TD there by improving the performance of the knee-joint angle estimation. The low RMSE values for TAD features was as a result of the smoothness of the features extracted compared to the TD features. Results shows improved performance with TAD feature for both models BPNN and LSTM across all features extracted. With more significant reduction in RMSE using LSTM model. Though there is need in improving an algorithm that can adapt to human differences between models training, this will unify performance of all subjects across all training models. This will prevent some of the discrepancies seen in Table 1 and 2 where the performance of one subject in a particular training model, may differ in another training model or among different features in same training model. Nevertheless, this work proves that TAD features extracted improve the accuracy of knee-joint angle estimation for osteoarthritis patients with RMSE error of 3.2566° with RMS feature. Also, among the two models, LSTM has an improved performance when estimating knee-joint angle from sEMG signal. In the future, other performance evaluation metrics can be used apart from the RMSE. There is need in improving an algorithm that can adapt to human differences between models training so as to avoid discrepancies in subject performance between different training models; also features can be combine to evaluate performance.

REFERENCES

[1] H. Shan, C. Jiang, Y. Mao, and X. Wang, "Design and control of a wearable active knee orthosis for walking assistance," *2016 IEEE 14th Int. Work. Adv. Motion Control. AMC 2016*, pp. 51–56, 2016, doi: 10.1109/AMC.2016.7496327.

[2] K.-I. Lee, C.-G. Lee, W.-K. Hong, and M. Kim, "Analysis of the Characteristics of Ground Reaction Force According to the Level of Knee Osteoarthritis During Gait," *Korean J. Sport Biomech.*, vol. 25, no. 4, pp. 393–399, 2015.

[3] T. Triwiyanto, O. Wahyunggoro, H. A. Nugroho, and H. Herianto, "Evaluating the Performance of Kalman Filter on Elbow Joint Angle Prediction Based on Electromyography," vol. 18, no. 12, pp. 1739–1748, 2017, doi: 10.1007/s12541-017-0202-5.

[4] J. Song *et al.*, "Effects of different feature parameters of sEMG on human motion pattern recognition using multilayer perceptrons and LSTM neural networks," *Appl. Sci.*, vol. 10, no. 10, pp. 1–18, 2020, doi: 10.3390/APP10103358.

[5] C. Yang, X. Xi, S. Chen, S. M. Miran, X. Hua, and Z. Luo, "SEMG-based multifeatures and predictive model for knee-joint-angle estimation," *AIP Adv.*, vol. 9, no. 9, 2019, doi: 10.1063/1.5120470.

[6] 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 (Switzerland)*, vol. 20, no. 17, pp. 1–18, 2020, doi: 10.3390/s20174966.

[7] K. Li, J. Zhang, L. Wang, M. Zhang, J. Li, and S. Bao, "A review of the key technologies for sEMG-based human-robot interaction systems," *Biomedical Signal Processing and Control*, vol. 62. Elsevier Ltd, Sep. 01, 2020, doi: 10.1016/j.bspc.2020.102074.

[8] S. Michieletto *et al.*, "GMM-based single-joint angle estimation using EMG signals," in *Advances in Intelligent Systems and Computing*, 2016, vol. 302, pp. 1173–1184, doi: 10.1007/978-3-319-08338-4_85.

[9] L. Zhang, D. Soselia, R. Wang, and E. M. Gutierrez-Farewik, "Lower-Limb Joint Torque Prediction Using LSTM Neural Networks and Transfer Learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 600–609, 2022, doi: 10.1109/TNSRE.2022.3156786.

[10] M. T. N. Truong, A. E. A. Ali, D. Owaki, and M. Hayashibe, “EMG-Based Estimation of Lower Limb Joint Angles and Moments Using Long Short-Term Memory Network,” *Sensors*, vol. 23, no. 6, 2023, doi: 10.3390/s23063331.

[11] hafiz ibrahim hassan, “EMG-based neural network estimator for knee joint movement,” International Islamic University Malaysia, 2016.

[12] H. I. Hassan, A. Usman, A. Y. Zubair, H. Ndagi, and I. L. Garba, “Evaluating the Performance of Emg Based Back Propagation Neural Network (Bpnn) Time Domain Features and Time Advance Features for Knee Joint Angle Estimation,” *Niger. J. Trop. Eng.*, vol. 16, no. 1, 2022, doi: 10.59081/njte.16.1.010.

[13]J. Wang, D. Cao, G. Liang and Y. Wu, "A Surface Electromyography-Driven State-Space Model for Joint Angle Estimation," in IEEE Sensors Journal, vol. 25, no. 2, pp. 3048-3060, 15 Jan.15, 2025, doi: 10.1109/JSEN.2024.3503689.

[14] H. Teng, G. Zhang, H. Zhou, S. M. A. Rahman and A. Frisoli, "Estimation of Elbow Joint Angle from EMG and IMU measurements based on Graph Convolution Neural Network," 2024 30th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), Leeds, United Kingdom, 2024, pp. 1-6, doi: 10.1109/M2VIP62491.2024.10746173.