scientific reports

OPEN

A linear-attention based network for estimating continuous upper limb movement from surface electromyography

Chuang Lin[1✉], Chunxiao Zhao[1], Na Li[2] & Ming Dai[3✉]

Continuous kinematics estimation methods are important in human-machine interaction systems because they offer a more natural and intuitive result than pattern-recognition methods. The estimation of upper-limb movement is important, which involves the estimation of the movement of the elbow and shoulder joints. However, the deformations of muscles around the elbow and shoulder can cause the shifting of the angle sensors. Nevertheless, Vicon can measure joint angles using cameras that have no contact with muscles. In this paper, we propose a Linear-Attention-based model (LABD) rather than other deep learning models to estimate upper-limb movements. The sEMG signals were collected from eight sEMG sensors and angles were measured from Vicon. The experimental results of LABD were compared with multilayer perceptrons (MLP), Temporal Convolutional Network (TCN), long short-term memory network (LSTM), and a dot-product attention-based model (DABD). The Pearson correlation coefficients (PCC) between the target and estimated joint angle sequences were calculated to evaluate the performance of each model. The Wilcoxon signed-rank results showed that the LABD significantly outperformed the other models.

The use of robots has increased in recent years, this caused the booming attention attached by human-machine coordination in manufacturing, military, service delivery and other fields[1–3], in this field, making the robot understand the movement of human is particularly important[4,5]. The computer vision-based methods[6] and sEMG-based methods[7] have been proposed and widely adopted several years ago. Actually, the computer version-based models often need fixed cameras and a bright environment, and the block of the body will significantly influence the prediction result. Also, the method needs a lot of computing resources and has a delay between the predicted result and the real movement. However, the sEMG-based models can acquire the kinematics in a few instants before the movement occurs and do not need the conditions required by the computer version-based models. The sEMG signals contain rich information about muscles and have a correlation with movements. sEMG has been used in human-machine interaction control systems for decades[8–10], so the sEMG-based method was the focus of the study. However, the most widely used sEMG-based methods were myoelectric-pattern recognition techniques[11], which are only able to provide discrete movement classification. Therefore, to provide continuous kinematic estimation, the sEMG-based simultaneous and proportional control method was proposed; the method provided continuous movement estimation, and it is closer to common human activities.

To achieve the goal of intuitive and natural control of robotic arms, Lin *et al.*[12] employed wise non-negative matrix factorization (NMF) method based on the concept of muscle synergy to estimate kinematics from multichannel sEMG[13], Qing *et al.*[14] proposed a Hill-based muscle model that estimate continuous elbow angles from sEMG signals. Wang *et al.*.[15] used Long Short-Term Memory (LSTM) networks to acquire continuous hand kinematics of six movements from forearm sEMG signals. Actually, musculoskeletal base models need Complex calculations, so they are hard to be widely adopted in practical applications. The non-negative matrix factorization-based methods are usually used in whist movement estimation, and they provide a control matrix rather than joint angles. Most machine learning models use sensors to collect angles, which makes the collection of instabilities of the collected angles when muscles are moving. Moreover, the traditional models cannot analyze

[1]The School of Information Science and Technology, Dalian Maritime University, Lingshui Street, Dalian 116000, China. [2]The National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Changdu 610017, China. [3]School of Artificial Intelligence, Shenzhen Polytechnic University, Shenzhen 518055, China. ✉email: linchuang_78@126.com; daiming@szpu.edu.cn

Abbreviations	Implication
EBH	Joint angle between the forearm and upper arm
SBV	Joint angle between the upper arm and shoulder in the coronal plane
SDH	Joint angle between the upper arm and shoulder in the sagittal plane
SDV	Joint angle between the upper arm and shoulder in the transverse plane

Table 1. Recorded Joint Angles.

Index	Movements	repetitions
1	Elbow joint flexion/extension(in the sagittal plane)	6
2	Shoulder joint flexion/extension (in the sagittal plane)	6
3	Shoulder joint horizontal adduction/abduction (in the coronal plane)	6
4	Shoulder joint vertical adduction/abduction (in the transverse plane)	6
5	Shoulder elbow simultaneous circular movement (in the sagittal plane)	6
6	Shoulder elbow simultaneous circular movement (in the transverse plane)	6
7	Shoulder elbow simultaneous circular movement (in the coronal plane)	6
8	Move the cup	6

Table 2. Eight recorded movements.

the time features and the relationship between different sEMG channels simultaneously, which cannot achieve a high estimation accuracy.

Thus, there is a need to find a novel method that can collect joint angles and estimate them accurately. Traditional angle sensors should be fixed in muscles; the movement of muscles may cause instabilities. The Motion Capture technology can measure angles without sensors fixed in muscles, and can calculate angles from the key point fixed in human joints, which makes the collection of joint angles more accurate. So we used a Motion Capture device to measure joint angles. We used the designed software to synchronize sEMG signals and joint angles to ensure the Neural signals and kinematics are synchronized. To make the accuracy higher, we employed the Linear-Attention[16] in this study. Linear-Attention can extract the features of input data in 2 dimensions, so we chose to apply Linear-Attention and to compare it with the commonly adopted deep learning-based methods (MLP[17], TCN[18], LSTM[19], and dot-production attention[20]).

Method
Subjects
The dataset used in this study was obtained from eight full-body individuals (all male, ages 24–40, all right-handed). None of the participants has any known neurological disorders. All participants read and signed an informed consent form before proceeding with the experiment. The experimental protocol and all methods adhered to the principles outlined in the Declaration of Helsinki and were approved by the Research Ethics Committee of West China Hospital (#2022-505).

Experimental protocol
As introduced in the Introduction, to synchronize the sEMG signals and arm movements, the sEMG signals and angle measurements were obtained from the same arm, the sEMG signals were recorded by Noraxon (Noraxon Ultium EMG, China), and the angle measurements were obtained from Vicon(Vicon, England) and are synchronous by designed software.

During the measurement, each subject was asked to sit down in front of a laptop displaying a repeated movement video. The eight sEMG sensors were placed in the anterior deltoid, the middle deltoid, the posterior deltoid, the biceps, the posterior triceps, the medial triceps, the pectoralis major, and the musculi supraspinatus, as shown in Fig. 2. The joint angles were captured from Vicon, and four angles were recorded and are shown in Table 1.

For target movements, four single DoF movements, three simultaneous multiple DoF movements, and a free movement were selected; they are functionally relevant and representative of most shoulder-elbow flexions, translations, and rotations in common. These are listed in Table 2. The subjects were told to sequentially perform the eight types of movements on the dominant arm, starting from the natural state with the arms extended, motionless, and palms facing forward. Subjects were asked to start and end each movement in the sequence by a prompt on a laptop screen. The subjects were instructed to imitate the movements shown on the screen, each movement repeated six times, and had a rest for at least 90 seconds between each movement, which ensured the quality of the sEMG signals acquired from the subjects.

The Noraxon (Noraxon Ultium EMG, China) and Vicon (Vicon, England) were used to acquire sEMG signals and angles. The Noraxon applied sEMG signals for 8 channels at sampling frequency of 2000Hz, The Vicon can measure all known joint angles in the human body. In this measurement, four needed angles were recorded and are mentioned in the first paragraph in this section. The Vicon recorded angles at a sampling frequency of
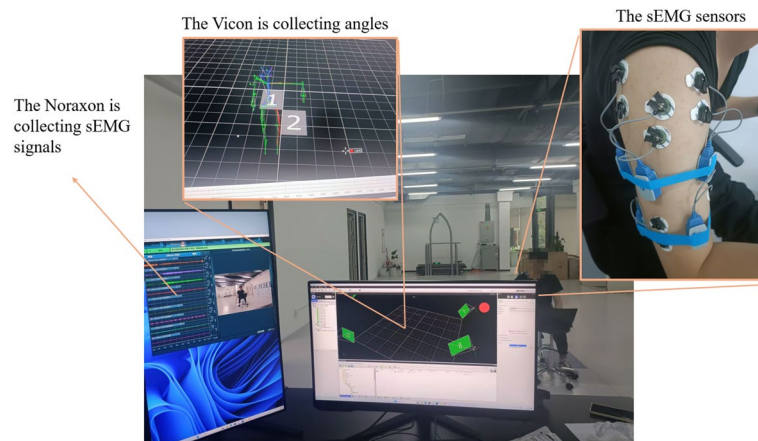
Fig. 1. The experiment setup, the subject sits on a chair imitating the movements on a laptop, with the sEMG sensor on the dominant arm recording sEMG signals, and Vicon recording the angles of the four joints of the subject.
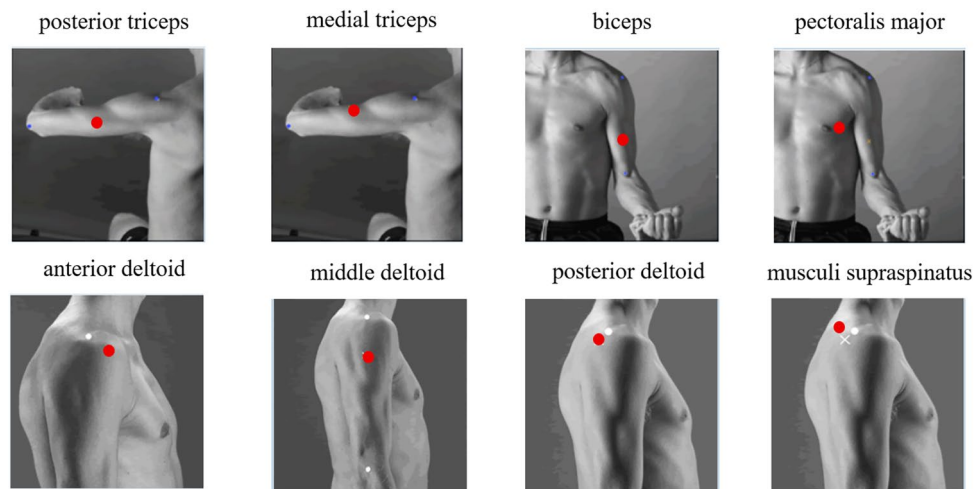

Fig. 2. sEMG sensor placement of subjects, the red dot represents the position of the sensors.

100Hz, and were resampled to 2000Hz after measurement. The synchrony of sEMG and angle sampling points was ensured by a designed software. 10Hz-500Hz Butterworth filter was applied to the recorded sEMG signals. The setup of the experiment is shown in Fig. 1.

The root mean square value (RMS) feature was chosen as the feature extract from sEMG signals for it contains a great amount of information. The RMS feature was extracted by a window of length 100ms, and a step at 0.5ms. The RMS can be calculated using the following equation:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(n_i - \overline{n})^2} \tag{1}$$

Where the N is the size of the window,n_i represents the $i-th$ sampling point of the sEMG in the sliding window, and the $\overline{n}$ represents the mean value of the data in the sliding window.

After the extract for the RMS, The $\mu - law$ normalization was concluded. The μ-law normalization[21] amplifies the input sEMG data close to 0 using a logarithmic approach, allowing them to contribute more to the predicted results. Normalization of the μ law has been shown to effectively improve the performance of the model to evaluate EMG signals. The μ-law normalization formula is as follows:

$$F(x) = sign(x_i)\frac{\ln(1 + \mu|x_i|)}{\ln(1 + \mu)} \tag{2}$$

where x_i is the t-th sampling point of the input, and the hyperparameter μ decides the range after normalization.

Model development

Dot-product attention

Attention based model:The dot-product attention attention mechanism[22] is introduced by Google cooperation in 2017. The mechanism has demonstrated exceptional performance in natural language processing. The mechanism is crucial in generating outputs based on the interdependence among the input vector sequences, which can extract the correlation from different sEMG channels and can generate a better angle. The dot-product attention is considered the most basic attention mechanism, so it was chosen to be in comparison as a baseline model in this study. The model computing formula has been displayed below.

Given the input matrix X with a dimension of $l \times c$, where l is the length of the input sequence and c is the number of sEMG channels in the experiment, l is set to 100ms. The attention layer generates the attention matrix with the same dimension as X, utilizing a set of parallel and independent heads. The attention layer contains three linear layers, and can generate the query matrix Q, key matrix K, and values matrix V, respectively. The corresponding formulas are defined as follows:

$$Q = W_q X + b_q \tag{3}$$

$$K = W_k X + b_k \tag{4}$$

$$Q = W_v X + b_v \tag{5}$$

$$\alpha = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \tag{6}$$

Where the W_q, W_k, W_v are three learnable matrices of size $R^{l \times l}$, b_q, b_k, b_v are in $R^{l \times c}$, d_k is a scalar value, which equals to the first dimension of W_q, W_k and W_v, the dot-product of two matrix Q and K are divided by $\sqrt{d_k}$. α represents the attention matrix and was normalized to [0,1].

Linear-attention based model

Linear- Attention mechanism was first provided by Angelos in 2020[16], different from the dot-product attention mechanism, Linear Attention used the structure similar to RNN, which can deal with the sequence data better then other models, also the advantage of dot-product attention remained. Linear-Attention mechanism combines the advantages of RNN and dot-product attention, which makes it perform better in sEMG continues estimation.

Dot-product attention mechanism can calculate the relationship between every sEMG channel, but cannot understand the time feature in the sEMG data. The Linear-Attention overcomes the disadvantage.

In linear-attention mechanism, the attention layer was formalized as a recurrent neural network, the resulting RNN has two hidden states, namely the attention memory s and the normalized memory z. The formula has been subscribed to donate the timestep in the recurrence:

$$s_0 = 0 \tag{7}$$

$$z_0 = 0 \tag{8}$$

$$s_i = s_{i-1} + \phi(x_i W_K)(x_i W_V)^T \tag{9}$$

$$z_i = z_{i-1} + \phi(x_i W_K) \tag{10}$$

$$y_i = f_l\left(\frac{\phi(x_i W_Q)^T s_i}{\phi(x_i W_Q)^T z_i)} + x_i\right) \tag{11}$$

$$\phi(x) = elu(x) + 1 \tag{12}$$

where x_i denotes the ith input and the y_i denotes the ith output for a specific transformer layer, the W_Q, W_K, W_V are three learnable matrices. The function $elu()$ denotes the exponential linear unit.

As show in Fig. 3, the Linear-Attention use the structure similar to RNN, the s_{t-1} and s_t represent the attention memory output at time $t-1$ and time t. The z_{t-1} and z_t represent the attention memory output at time $t-1$ and time t also. They recorded the feature in previous sample points, this allows Linear-Attention mechanism to understand the sEMG signals better. The x_t represents the eight channels of sEMG in time t, x_t was calculated by the Linear-Attention cells then calculated by an exponential linear unit, this allow the Linear-Attention mechanism to understand the relationship between different sEMG channels and perform better in angle estimation.

The linear-attention based model is displayed in Fig. 3, as shown in the figure The output of Linear-Attention cell then calculated by a exponential linear unit, and then add to the raw input to prevent gradient vanishing and explosion[23], then the output go though a full connect layer to get the output into a vector in 1×4 which related to four estimated angles.

Comparison of the models

The multilayer perceptron (MLP), considered to be the most basic neural network and has been used frequently in investigating human-machine systems[24,25]. Therefore, in this study, MLP has been chosen as the baseline model for comparison. The MLP has been shown to have the ability to fit a continuous curve or sequence value. In this study, three layers of non-fully connected networks were set as the hidden layers. The sEMG data was
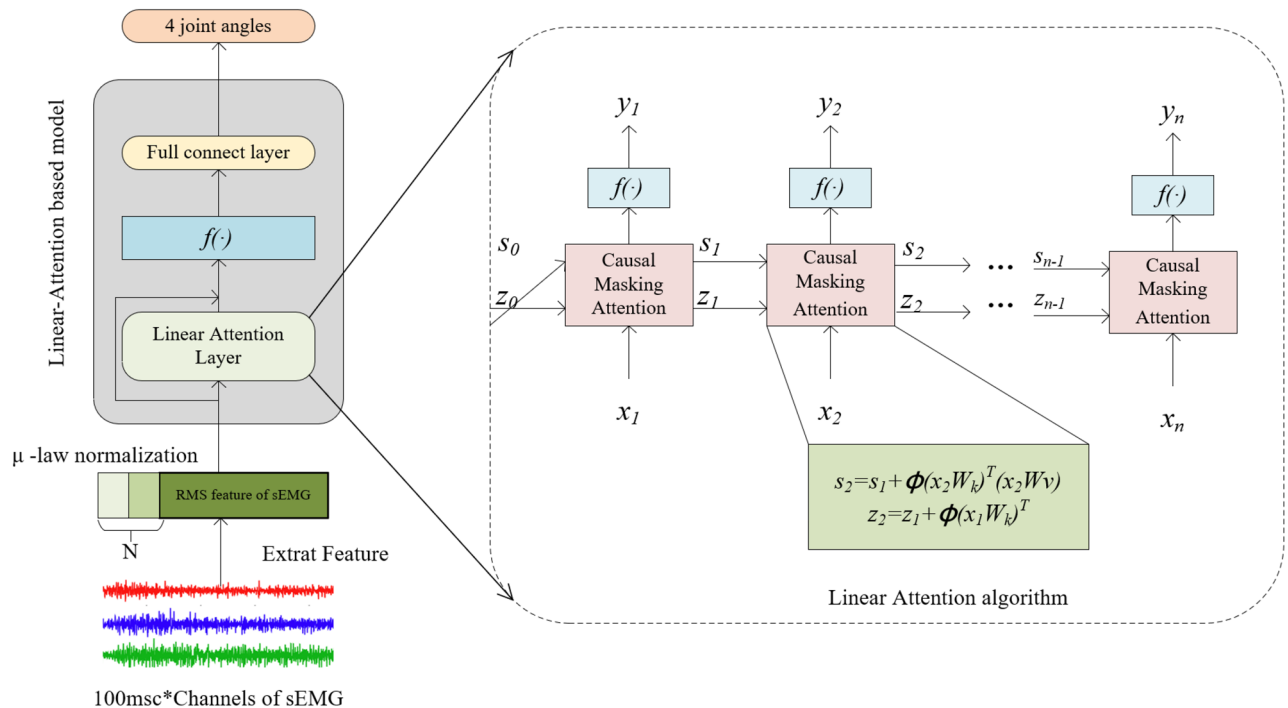
Fig. 3. The linear-attention based model(left) and linear-attention algorithm(right).

flattened into a 1-dimensional form before send into the neural network. The last two layers compressed the expanded data into 4 channels to predict the joint angles. Finally, to prevent the models from overfitting, the spatial dropout at 0.3 was added between each layer during training.

As a comparison model, we employ a Temporal Convolutional Network (TCN)[18]. TCN utilizes stacked one-dimensional convolutional layers with dilated kernels to model long-range temporal dependencies in sequential data. Critically, they enforce causality—ensuring predictions depend solely on past inputs while enabling parallel computation over entire sequences, offering computational advantages over recurrent architectures. TCN had been used in previous studies[26] and performed well in continuous angle estimation, so TCN was chosen as a comparison model.

Long Short-Term Memory network (LSTM)[19], has been used in previous study. LSTM is a recurrent neural network (RNN) variant specifically designed with gating mechanisms (input, forget, output gates) and a cell state to learn long-term temporal dependencies in sequential data. This architecture effectively addresses the vanishing gradient problem inherent in standard RNNs, making it well-suited for modeling physiological signals like sEMG.

Settings
The backpropagation algorithm was chosen to be the training approach in this study. Both the dot product attention and Linear-Attention set the layer into 1 The LSTM was set to 3. the dropout value was set to 0.3, The dropout value was set to 0.3 to prevent overfitting. All models were trained for 300 epochs, the learning rete was set to 0.001 and half after 100 epochs during training. In this study, all hyper-parameters were chosen best performed from the grouped values.

In this experiment, the Friedman's test[27] and Wilcoxon's sign rank test[28] were used to evaluate the significance of the results.

Results
Before a detailed analysis of the proposed and comparison networks, a performance comparison was made. The input data used the $\mu - law$ normalization and raw input data (as show in Table 5), the Table show the comparison of all models in this study.

As shown in Table 5 ,all models were greatly improved by adopting the $\mu - law$ normalization, because it amplifies the input sEMG data close to 0 using a logarithmic approach, allowing them to contribute more to the predicted results. It has been demonstrated that $\mu - law$ normalization can effectively improve he model's performance for evaluating EMG signals .

After the comparison of raw input and $\mu - law$ normalization, all models were developed in PyTorch[29]. All movements were combined, four repetitions of each movement were set as the training set and the remaining two repetitions were set as the testing set. All results are shown in Table 3. As an example, the performance of each model in subject 3 is compared in Fig. 4, which shows the variation of the measured joint angles and the estimated angles over time.

Subject	MLP	TCN	LSTM	DABD	LABD
1	0.7912	0.8511	0.8703	0.8913	0.9141
2	0.8417	0.8411	0.8978	0.8847	0.9054
3	0.8266	0.8710	0.9144	0.9511	0.9608
4	0.8271	0.8195	0.9015	0.9010	0.9637
5	0.8038	0.8886	0.9027	0.9143	0.9239
6	0.8403	0.8917	0.9076	0.9368	0.9438
7	0.8267	0.8522	0.9177	0.9507	0.9706
8	0.8115	0.8450	0.9001	0.9172	0.9613

Table 3. Individual subject correlation coefficients between directly measure and estimated joint angles for each analysis model.

Fig. 4. Estimated outputs compared to the actual measured joint angles based on, (**a**) MLP model, (**b**) TCN model, (**c**) LSTM model, (**d**) DABD, (**e**) LABD.

Panel (a) of the figure shows that the MLP model is hardly able to match the pattern of any of the four measured joint angles and particularly performed badly when estimating the SDV angle. The TCN model performed better, when the real joint angle became stable, the estimated joint angle became unstable. The LSTM performed better for the estimation of the elbow joint angle but it performed unwell for the SDV angle. The
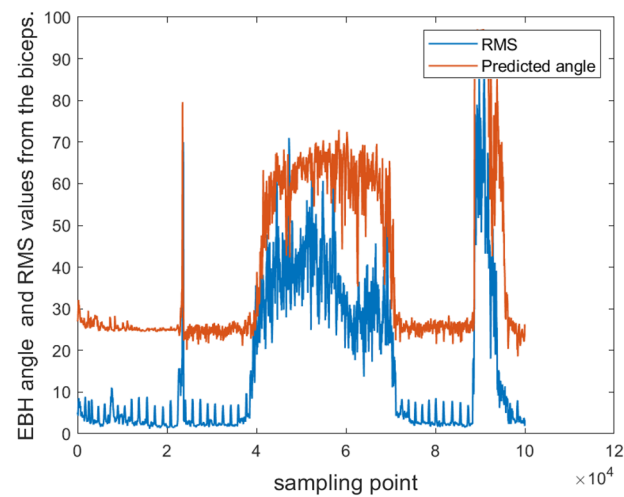
Fig. 5. Comparison of the EBH angle values and the RMS values extracted from the recorded sEMG data from the biceps.

	MLP	TCN	LSTM	DABD	LABD
CC	0.8204	0.8575	0.9015	0.9083	0.9430
RMSE	7.88	4.40	4.20	4.50	4.08
DTW	25354	12603	11805	11879	8863

Table 4. Average performance (CC, RMSE, DTW) for individual subject between directly measure and estimated joint angles for each analysis model. *RMSE* root mean square error, *DTW* dynamic time warping distance.

Model name	Raw input	$\mu - law$ input
MLP	0.7845	0.8204
TCN	0.8215	0.8575
LSTM	0.8648	0.9015
DABD	0.8847	0.9083
LABD	0.9017	0.9430

Table 5. Average performance (CC) comparison between all models with raw input and $\mu - law$ normalization input.

DABD performed better then the upper models but there was still somewhere inaccuracy in SBV and SDV angle. However, the LABD compensated for the shortcomings of the upper models, and achieved the best performance when compared. The Pearson correlation coefficient (cc) was used to provide a quantitative measure of the association between the estimated joint angles and real joint angles of each subject, as shown in Table 3.

To analyze the relevance of the features, we plot a line chart of predictions vs. the single input feature. The EBH angle depends mostly on the biceps, so we compared the angle with the RMS value extracted from the recorded sEMG data from the biceps, as show in Fig. 5, the rest angel and RMS figures were put in a github link (click the "github link" to view them), they have a similar trend, so all models performed well in EBH angle as shown in Fig. 4, While the rest angels depend on several muscles, so they do not have the similar trends with a single RMS channel. Our model can efficiently acquire information from different sEMG (RMS) channels, that is the reason our model performed well in the rest angles' estimation.

We also reported the average CC value, Root Mean Square Error (RMSE) and Dynamic Time Warping distance (DWT) in Table 4. Our method achieved the best performance in RMSE and DTW, while MLP performed the worse, which showed the complete advantage of our model LABD.

To evaluate the generalization ability of our method, we conducted the leave-one-subject-out tests, that is train the model on other 7 subjects and test on the left single subject, between our model and other models and reported the average performance (CC, RMSE and DTW) in Table 6, from the Table, our model LABD continues performed well in CC value, and performed almost the same as the model DABD and LSTM in RMSE and DTW,

	MLP	TCN	LSTM	DABD	LABD
CC	0.5011	0.6710	0.7717	0.7648	0.8014
RMSE	15.47	13.82	12.44	12.02	12.06
DTW	16667	21533	18945	15657	15929

Table 6. The average correlation coefficients, RMSE, and DTW between directly measured and estimated joint angles for models in cross-subjects senario. *RMSE* root mean square error, *DTW* dynamic time warping distance.

Model name	CPU (ms)	GPU (ms)	ARM (ms)	Training (s)	Params (MB)
MLP	4.00	1.00	20.52	1.49	1.14
TCN	7.59	1.02	33.50	1.95	2.19
LSTM	23.59	9.02	121.50	5.64	5.83
DABD	10.46	1.01	57.14	4.64	1.48
LABD	8.36	0.99	22.97	2.43	0.68

Table 7. The comparisons of inference time, training time, model params of different models on 4 joint angles on I9-139000H(CPU), NVIDIA RTX4060(GPU), and Raspberry Pi 4B(ARM). The used device is CPU: I9-139000H, GPU: NVIDIA RTX4060, ARM: Raspberry Pi 4B and training: NVIDIA RTX4060

which means our model can be generalized in cross-subject tasks and has a better performance than the baseline models (TCN, MLP, and LSTM), this showed the potential of our model in cross-subject tasks.

The different methods showed significantly different performance on subjects according to the Friedman test. And the Wilcoxon signed-rank test result after Bonferroni correction indicated that the LABD(CC=0.9430) significantly outperformed the MLP (CC=0.8204, p=0.0419), the TCN (CC=0.8575, p=0.0374), the LSTM (CC=0.9015, p=0.0714) and the DABD (CC=0.9083, p=0.0799).

To evaluate the model from a deployment perspective, we tested the inference time of different models on different devices: CPU, GPU, and Raspberry Pi 4B. The results were shown in Table 7. From the table, the MLP and TCN achieved shorter inference time than our model in three kinds of devices and training time, but they have poor performance in model accuracy. Our model LABD, performed better than the LSTM and DABD models in inference time and training time, and well within the real-time requirement for human-machine interaction, which showed our model is suitable for real-time deployment.

Discussion

In this study, a Linear-Attention based model (LABD) was proposed, the method was used to obtain estimated joint angles from non-invasively collected sEMG signals, and compared to MLP, LSTM, TCN and DABD. The result of the experiment showed that LABD outperformed MLP, LSTM, TCN, and DABD in exploiting the kinematic information from sEMG signals in each case, and the differences were statistically significant.

The experiment show that, there is a great challenge in estimating the joint angles adequately using existing models. The MLP model cannot estimate the joint angles precisely because of its structure, the input data must be flattened at first, and the step makes the sEMG signal lose its information in the time dimension and channel dimension. TCN is a Convolutional Neural Network designed for time-series data. It can deal with sEMG signals better than traditional CNN models, but the Convolutional operation caused some unwanted noise. Compare to MLP and TCN model, the LSTM model can generate smoother and more accurate joint angles because its memory cell can combine prior information and current information, but the cell is too simple to deal with the sEMG information and cannot extract features from sEMG channels, which makes the joint angles decided by more than one channel badly estimated. Different from LSTM, the DABD can not only combine prior information but also can combine future information and the model structure is Sufficiently complex in dealing the sEMG signals, which make it outperformed than LSTM, but the DABD cannot extract features from sEMG channels as well as LSTM model. The proposed method: LABD, solved the problem due to its structure and yielded the best estimation results.

Conclusion

The proposed LABD model achieved a mean Pearson correlation coefficient of 0.94 (±0.02) and an RMSE of 3.75, outperforming MLP(0.82±0.2, p=0.0419), TCN (0.86±0.05, p=0.0374), LSTM (0.90±0.01, p=0.0714) and DBAD (0.93±0.04, p=0.0799), while our model parameters (0.68MB) and inference time (CPU: 8.36ms, GPU: 0.99ms, ARM: 22.97ms) are outperformed than the well-performed models (DABD and LSTM).

Though the synchronized measurements, the kinematic information acquired from sEMG signals can estimate the four joint angles of the upper-limb, the results show that the Linear-Attention based model achieved a significantly better accuracy than MLP, TCN, LSTM and dot-production models, the structure of LABD makes it own modest time and memory, the eight movements were used to ensure the model can estimate enough and broad movements.

The model was trained for every single subject, absolutely, the training strategy is complicated and makes it difficult to generalize. Cross-subject style can train a model across subjects and adapt to multiple subjects[30]. Transfer learning can use a small number of training epochs to reduce model training time by preserving similar parts between subjects. Our future work will focus on applying our methods to better performed cross-subject study, transfer learning and synthetic EMG data to improve robustness in scenarios with small sample sizes, as demonstrated in recent generative AI studies on rare diseases[31]. In short, our method is a outperformed and suitable approach for a single subject.

Data availability

References

1. Kuo, C.-M., Chen, L.-C. & Tseng, C.-Y. Investigating an innovative service with hospitality robots. *Int. J. Contemp. Hosp. Manag.* **29**(5), 1305–1321. https://doi.org/10.1108/IJCHM-08-2015-0414 (2017).
2. Noorman, M. & Johnson, D. G. Negotiating autonomy and responsibility in military robots. *Ethics Inf. Technol.* **16**(1), 51–62. https://doi.org/10.1007/s10676-013-9335-0 (2014).
3. Veloso, M., Biswas, J., Coltin, B. & Rosenthal, S. Cobots: Robust symbiotic autonomous mobile service robots. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)* . 4423–4429 (2015)
4. Oliveira, D. S. et al. A direct spinal cord-computer interface enables the control of the paralysed hand in spinal cord injury. *Brain* **147**, 3583–3595 (2024).
5. Bi, L. Z., Feleke, A. G. & Guan, C. T. A review on emg-based motor intention prediction of continuous human upper limb motion for human-robot collaboration. *Biomed. Signal Process. Control* **51**, 113–127 (2019).
6. Szabo, R. & Gontean, A. Controlling a robotic arm in the 3D space with stereo vision. In *2013 21th Telecommunications Forum (TELFOR)*. 916–919 (2013)
7. Farina, D. et al. The extraction of neural information from the surface emg for the control of upper-limb prostheses: Emerging avenues and challenges. *IEEE Trans. Neural Syst. Rehabil. Eng.* **22**(4), 797–809. https://doi.org/10.1109/TNSRE.2014.2305111 (2014).
8. Cho, K., Merrienboer, B.V., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. & Bengio, Y.J.C.S. Learning phrase representations using rnn encoder-decoder for statistical machine translation (2014).
9. Montazerin, M. et al. Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density emg signals. *Sci. Rep.* https://doi.org/10.1038/s41598-023-36490-w (2023).
10. He, Y.N., Fukuda, O., Bu, N., Okumura, H. & Yamaguchi, N. Surface emg pattern recognition using long short-term memory combined with multilayer perceptron. In *40th Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society (EMBC)*. 5636–5639 (2018)
11. Bien, Z. New emg pattern recognition based on soft computing techniques and its application to control of a rehabilitation robotic arm. (2000). https://api.semanticscholar.org/CorpusID:5827523
12. Chuang, L., Binghui, W., Ning, J. & Farina, D.: Robust extraction of basis functions for simultaneous and proportional myoelectric control via sparse non-negative matrix factorization. *J. Neural Eng.* **15** (2018)
13. Tresch, P. et al. The construction of movement by the spinal cord. *Nat. Neurosci.* **2**(10), 162–167 (1999). https://doi.org/10.1038/5721.
14. Ding, Q.C., Xiong, A.B., Zhao, X.G. & Han, J.D. A novel emg-driven state space model for the estimation of continuous joint movements. In *2011 IEEE International Conference on Systems, Man, and Cybernetics*. 2891–2897 (2011). https://doi.org/10.1109/ICSMC.2011.6084104
15. Wang, C. et al. semg-based continuous estimation of grasp movements by long-short term memory network. *Biomed. Signal Process. Control* https://doi.org/10.1016/j.bspc.2019.101774 (2020).
16. Katharopoulos, A., Vyas, A., Pappas, N. & Fleuret, F. Transformers are rnns: Fast autoregressive transformers with linear attention. In *International Conference on Machine Learning* (Daume, H., Singh, A. eds.) . Vol. 119. (2020).
17. Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8 (1989).
18. Lea, C., Flynn, M.D., Vidal, R., Reiter, A. & Hager, G.D. Temporal convolutional networks for action segmentation and detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017)
19. He, Y., Fukuda, O., Bu, N., Okumura, H. & Yamaguchi, N. Surface emg pattern recognition using long short-term memory combined with multilayer perceptron. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 5636–5639 (2018)
20. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. & Polosukhin, I. Attention is All You Need (2023). arXiv:1706.03762.
21. Rahimian, E. et al. Fs-hgr: Few-shot learning for hand gesture recognition via electromyography. *IEEE Trans. Neural Syst. Rehabil. Eng.* **29**, 1004–1015. https://doi.org/10.1109/tnsre.2021.3077413 (2021).
22. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. & Polosukhin, I. Attention is all you need. In *31st Annual Conference on Neural Information Processing Systems (NIPS)*. Vol. 30 (2017)
23. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition (2015). arXiv:1512.03385.
24. Muceli, S. & Farina, D. Simultaneous and proportional estimation of hand kinematics from emg during mirrored movements at multiple degrees-of-freedom. *IEEE Trans. Neural Syst. Rehabil. Eng.* **20**(3), 371–378. https://doi.org/10.1109/TNSRE.2011.2178039 (2012).
25. Chu, J.-U., Moon, I. & Mun, M.-S. A real-time emg pattern recognition based on linear-nonlinear feature projection for multifunction myoelectric hand. In *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*. 295–298 (2005). https://doi.org/10.1109/ICORR.2005.1501105.
26. Chen, C. et al. SEMG-based continuous estimation of finger kinematics via large-scale temporal convolutional network. *Appl. Sci.-Basel* https://doi.org/10.3390/app11104678 (2021).
27. Azman, N., Syarif, A., Brahmia, M.-e.-A., Dollinger, J.-F., Ouchani, S. & Idoumghar, L. Performance analysis of rpl protocols in lln network using Friedman's test. In *2020 7th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)* (2020). https://doi.org/10.1109/IOTSMS52051.2020.9340203.

28. Sahoo, U.K., Panda, G. & Mulgrew, B. Sign-regressor Wilcoxon and sign-sign Wilcoxon. In *2010 International Conference on Advances in Recent Technologies in Communication and Computing*. 35–39 (2010). https://doi.org/10.1109/ARTCom.2010.37 .
29. Ketkar, N. Introduction to PyTorch. 195–208. (Apress, 2017). https://doi.org/10.1007/978-1-4842-2766-4_12 .
30. Chen, X. J., Guo, W. Y., Lin, C., Jiang, N. & Su, J. Y. Cross-subject lifelong learning for continuous estimation from surface electromyographic signal. *IEEE Trans. Neural Syst. Rehabil. Eng.* **32**, 1965–1973 (2024).
31. Trabassi, D. et al. Optimizing rare disease gait classification through data balancing and generative AI: Insights from hereditary cerebellar ataxia. *Sensors* **24**(11), 24 (2024).

Acknowledgements

Author contributions

Chuang Lin, Ming Dai and Chunxiao Zhao proposed an experimental approach and designed a research plan. Chunxiao Zhao transformed their ideas into a practical experiment. Chuang Lin provides guidance on the writing approach of the article and checks the rigor of its expression. Na Li helped complete the measurement of the sEMG and angle signals. All authors reviewed the manuscript. Consent for publication was given by all participants.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

All procedures and were approved by the Institutional Review Board of the West China Hospital (#2022-505). Written informed consent was obtained from all participants prior to data collection.

Additional information

Correspondence and requests for materials should be addressed to C.L. or M.D.

Reprints and permissions information is available at www.nature.com/reprints.

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