A Bi-directional LSTM Network for Estimating continuous upper limb movement from Surface Electromyography

Chenfei Ma[1], Chuang Lin[2], Oluwarotimi Williams Samuel[2], Weiyu Guo[2], Hang Zhang[2], Steve Greenwald[3], Lisheng Xu[1], and Guanglin Li[2]

***Abstract*—In human-machine interaction systems, continuous movement estimation methods occupy an important position because they are more natural and intuitive than pattern-recognition methods. Essentially, arm position is decided by the shoulder and elbow joint angles. However, the various deformations of muscles around the shoulder and elbow often lead to difficulties in sensor fixation, which results in a loss of synchronization between the surface electromyography (sEMG) signals and joint angles. In order to accurately estimate movement angles using sEMG in situations where the sEMG is not synchronized with joint angles, we utilized a bi-directional long short-term memory (Bi-LSTM) network rather than other deep learning methods to estimate non-dominant arm movements, based on the sEMG signal from the dominant arm. This estimation protocol was designed to avoid a multiplicity of sensors and to simulate more complicated loss of synchronization problems). The performance of the Bi-LSTM was compared with multilayer perceptrons (MLPs), convolutional neural networks (CNNs), and a long short-term memory network (LSTM). The Pearson correlation coefficient (cc) between the estimated and target joint angle sequences was calculated to evaluate the performance of each neural network. The Wilcoxon signed-rank results showed that the Bi-LSTM model significantly outperformed the MLP, CNN, and LSTM models (tested with completely untrained newly recorded free movements).**

Manuscript received April 26, 2021; Accepted June 29, 2021. This paper was recommended for publication by Editor Markus Vincze upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the National Natural Science Foundation of China Grants (#U1613222, #61773110, #82050410452, #81850410557, #81927804, #U1913601), the Shenzhen Basic Research Grants (#SGLH20180625142402055, #JCYJ20170413152804728, #JCYJ20180507182508857), the Fundamental Research Funds for the Central Universities (#N181906001) and Hong Kong ITF Guangdong-Hong Kong Technology Cooperation Funding Scheme (GHP/055/18SZ).

[1]Chenfei Ma and Lisheng Xu are with College of Medicine and Biological Information Engineering, Northeastern University, Shenyang Liaoning 110819, China (e-mail: googlcfma@gmail.com; xuls@bmie.neu.edu.cn).

[2]Chuang Lin, Oluwarotimi Williams Samuel, Weiyu Guo, Hang Zhang, and Guanglin Li are with the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology (SIAT), Chinese Academy of Sciences (CAS), and the SIAT Branch, Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518055, China (e-mail: linchuang_78@126.com; samuel@siat.ac.cn; hang.zhang1@siat.ac.cn; guoweiyu@126.com; gl.li@siat.ac.cn).

[3]Steve Greenwald is with Queen Mary University of London. (e-mail: s.e.greenwald@qmul.ac.uk)

Chenfei Ma and Chuang Lin equally contributed to this work. (Corresponding author: Lisheng Xu and Guanglin Li).

Digital Object Identifier (DOI): see top of this page.

***Index Terms*—Machine Learning for Robot Control, Continuous movement estimation, surface electromyography, free movements test.**

I. Introduction

In recent years, as the use of robots has increased, the coordination between human and machine has attracted much attention, especially in the manufacturing, military, service delivery, and medical fields [1]–[4]. In this regard, making a machine "aware" of human movements in its vicinity [5] is particularly important. Hence, movement estimation strategies driven by computer vision-based methods [6], inertial device-based methods [7], and surface electromyogram (sEMG) based methods [8], have been proposed and widely adopted for some years. Technically, computer vision-based methods often require carefully specified camera positioning and lighting, but they do not provide force information. Similarly, inertial device-based methods lack force information and are further limited by the time delay between the occurrence of the moment and the sensor output. However, sEMG-based methods can provide information about movement intention a few instants before the movement occurs and with no requirement for an unobstructed field of view. Thus, the sEMG approach lends itself well to a wide range of practical applications and is therefore the focus of this study. Since the sEMG signal contains rich neuromuscular information from which useful kinematic characteristics can be obtained, it has been used in human-machine interaction control systems for decades [9]. However, most widely adopted sEMG based control methods rely mainly on myoelectric-pattern recognition (MPR) techniques [10], which are only able to provide discrete movement classification. Therefore, simultaneous and proportional control (SPC) methods have been proposed with the aim of providing continuous movement estimation, this being closer to the natural dynamics of the human arm in accomplishing day-to-day activities.

Towards enabling intuitive and natural control of robotic arms, Qing et al. [11] proposed a Hill-based muscle model that provided continuous estimation of elbow joint angles. Lin et al. [12] employed a degree of freedom (DoF)-wise non-negative matrix factorization (NMF) method to estimate neural control

information from multichannel sEMG, based on the concept of muscle synergy [13]. Muceli et al. [14] used an artificial neural networks (ANN) to estimate the kinematics of wrist and hand movements from high-density sEMG recordings obtained from six healthy subjects. In practical terms, musculoskeletal -model-based methods usually require a relatively complex computational procedure, thus often limiting their wide adoption in practical applications. The NMF based approach is primarily used in wrist movement estimation, and it provides a control matrix rather than joint angles [15]. The training protocols of most machine learning-based methods involve the collection of signals and estimation of joint angle on the same limb, which leads to instabilities when sensors become unsynchronized, a problem we encountered in a previous study [16]. Moreover, the usual experimental paradigm, involving partitioning the recorded data into training and test sets (which share similar patterns of joint angle movement) often does not perform well in practical applications especially when untrained movements (including multiple new pattern movements) are presented to the built model.

Thus, there is a need to find a novel, effective solution to continuously estimate joint angles using sEMG sensors alone which can overcome in part, the loss of synchronization problem even in the presence of untrained joint angle movements. In this study the elbow and shoulder joint angles were chosen, because they are the major joints of the upper limb and are involved in most of its movements. To simulate the loss of synchronization as it might arise in most practical settings, we arranged to simultaneously record sEMG data from the dominant arm and joint angle data from the non-dominant arm, while moving both arms in the same way. Also, in an attempt to provide a more realistic simulation, we used for testing purposes, only newly recorded untrained free movements rather than a selection from pre-recorded movement data. Another challenge is that the testing dataset often contains many untrained pattern movements. Therefore, we investigated the effectiveness of commonly adopted deep learning networks. Bi-directional long short-term memory (Bi-LSTM) [17] has proven its ability in various fields, such as speech recognition [18] and sentiment analysis [19]. It has also been suggested that Bi-LSTM is capable of dealing with data in which the input and target do not exactly correspond point-by-point. Therefore, we chose to apply Bi-LSTM, and to compare it with the commonly adopted deep learning-based methods (MLP [20], CNN [21] and LSTM [22]).

II. Method

A. Subjects

The dataset utilized in this study was obtained from eight able-bodied individuals (6 males and 2 females, ages: 22-40 years, all right-handed). None of the participants has any known neurological disorders. 7 of the subjects has never participated in any myoelectric control experiment or used any myoelectric interface device. All the subjects gave their written informed consent and agreed to the publication of their anonymized data for scientific and educational purposes. The experimental protocols were approved by the Research Ethics Board of the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences.

B. Experimental Protocol

As mentioned in the introduction, in an attempt to simulate the synchronization loss problem, the sEMG measurements were obtained from the dominant arm while simultaneous angle measurements were obtained from the other arm. Throughout the time that data were recorded, subjects were asked to move both arms in the same way, thus producing a 'built-in' loss of synchronization. This procedure also helped to avoid the overcrowding of sensors which would have resulted had they been on the same shoulder and arm.

During the measurement sessions, each subject was asked to stand in front of a full-length mirror with their arms resting naturally along the body and the palms facing forward. The sEMG signal sensors were placed on the biceps brachii, triceps brachii, anterior deltoid, posterior deltoid, and pectoralis major and the joint angle sensors were placed on shoulder and elbow as shown in Fig. 1.

For target movements, four single DoF movements and three multiple-DoF simultaneous movements were selected as being

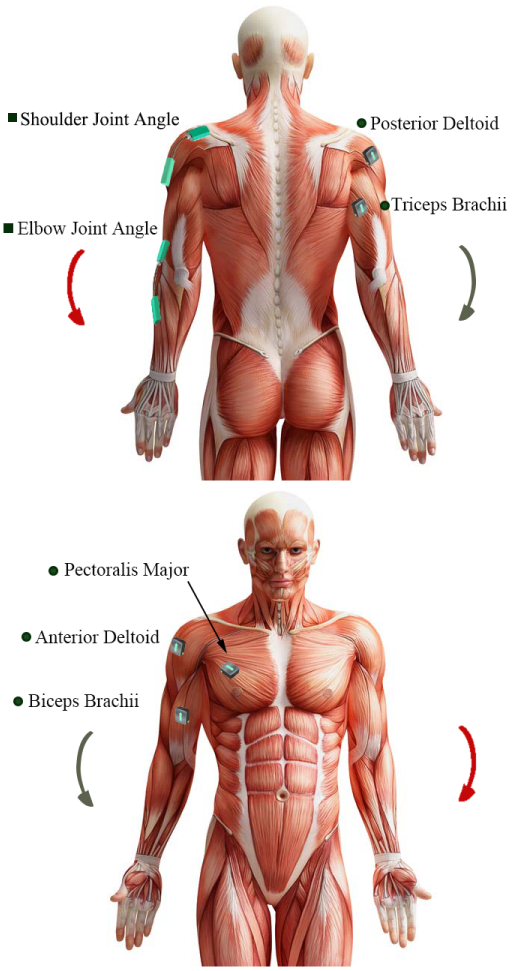


Fig. 1. Sensor placement on a right-handed subject, showing joint angle sensors (square markers) on the non-dominant side and sEMG sensors (circular markers) on the dominant side.

functionally relevant and representative of most shoulder-elbow flexions, rotations and translations occurring in day-to-day life. These are listed in TABLE I, and can be seen as images in [16]. The subjects were told to sequentially perform the seven types of movements on the non-dominant arm and to imitate these movements with the dominant arm, starting from the natural state with the arms extended, motionless and palms facing forwards. Subjects were prompted to start and end each movement in the sequence by a message on a computer screen. The participants were instructed to perform each set of movements simultaneously on both arms for approximately 90s at a rate of no more than one movement per second. Finally, the subjects were instructed to perform unrestricted simultaneous free arm movements for around 90s, again, while attempting to perform identical movements with each arm. After familiarization with the experimental protocol (including a basic warm-up exercise to allow the subjects to practice simultaneous and closely similar movements on both arms), the subjects performed the target movements sequentially with a rest period, while seated, of at least ten minutes between movements to ensure acquisition of quality sEMG signals that were free from muscle fatigue.

TABLE I
MOVEMENT PATTERNS INVESTIGATED

Movement index	Movements	Time	USAGE
1	Elbow joint flexion/extension (in the sagittal plane)	90s	Training
2	Shoulder joint flexion/extension (in the sagittal plane)	90s	
3	Shoulder joint horizontal adduction/abduction (in the coronal plane)	90s	
4	Shoulder joint vertical adduction/abduction (in the transverse plane)	90s	
5	Shoulder elbow simultaneous circular movement in the sagittal plane	90s	
6	Shoulder elbow simultaneous circular movement in the transverse plane	90s	
7	Shoulder elbow simultaneous circular movement in the coronal plane	90s	
8	Unlimited free movement	90s	Testing

TABLE II
RECORDED JOINT ANGLES

Abbreviations	Implication
EBH	Joint angle between the forearm and upper arm in the horizontal plane
EBV	Joint angle between the forearm and upper arm in the vertical plane
SDH	Joint angle between the upper arm and shoulder in the horizontal plane
SDV	Joint angle between the upper arm and shoulder in the vertical plane

The sEMG signals and joint angle signals were recorded synchronously using the Trigno™ wireless EMG acquisition system (Delsys, Inc, Natick, Massachusetts, USA) at a sampling frequency of 1926 Hz. To measure the joint angle, a Trigno™ Goniometer Adapter was used. These sensors provide 2 channels of joint angle data, from the horizontal and vertical planes respectively, as shown in Fig. 2. The sampling rate of each channel is 148 Hz. We recorded all 4 joint angles in the experiment. To ensure sample point synchrony between both sensor types, the joint angle signal sample rate was raised by bilinear interpolation to the same value (1926 Hz) as that used by the sEMG sensors. The labelling of each channel is described in TABLE II.

The joint angle sensor was triggered by the Delsys recording

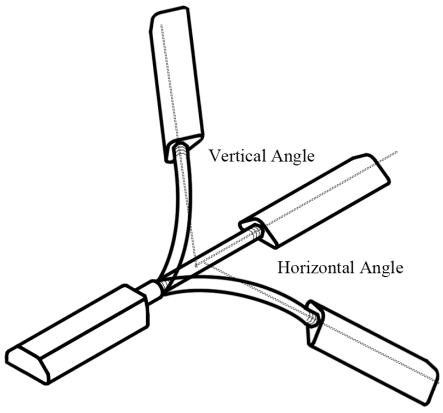


Fig. 2. The angle sensors detected joint angles in two orthogonal planes.

software, to ensure that the joint angle and sEMG were synchronized. Z-score normalization and power frequency filtering were applied to each channel of the recorded signals. The Z score normalization process is expressed by (1).

$$Z = \frac{(x-\mu)}{\rho} \tag{1}$$

where x stands for the original data vector, μ is the arithmetic mean of the given sequence and ρ is the standard deviation of the given data. Here, the sEMG normalization parameters were set to $\mu = 0$, $\rho = 0.5$. The root mean square value (RMS) was employed as the feature extraction approach due to its abundant information content and ease of computation. The RMS feature extraction used a window width of 100ms and a stride length of 0.5ms. The RMS [23] of an analysis window d can be calculated using the following expression:

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

where N is the analysis window length of window d, n_i stands

for the data values in window d and the $\bar{n}$ is the mean value of the data in window d. A comparison between the sEMG signal before and after applying the RMS feature extraction is shown in Fig. 3. When applying a neural network to bio-signal processing, this feature extraction scheme is effective in removing high frequency noise and revealing the meaningful information, as shown in Fig. 7 in the results section which compares the model performance operating on the raw sEMG input and RMS input.

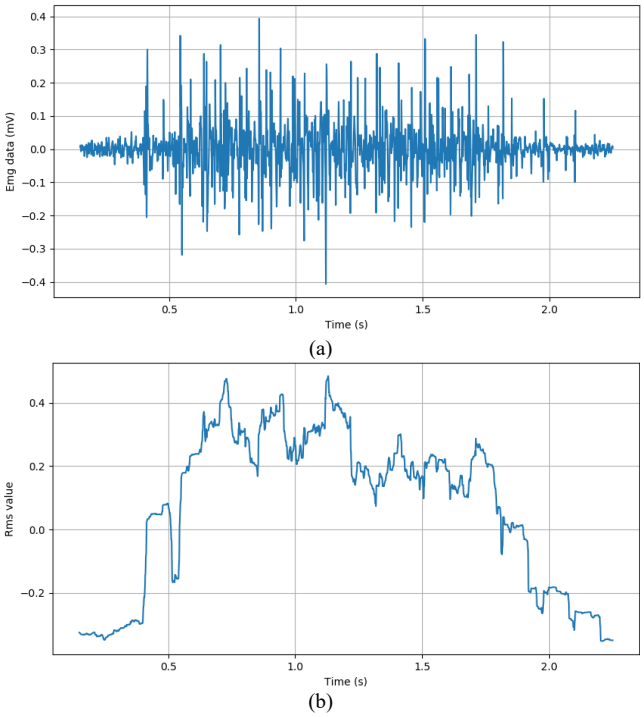


Fig. 3. Comparison between (a) the raw sEMG signal (after Z-normalization) and (b) the corresponding RMS values (after Z-normalization).

C. Model Development

1) Bi-directional LSTM

The long short term memory (LSTM) network, a type of recurrent neural network (RNN), was first proposed in 1997 by Hochreiter and Schmidhuber [24]. RNNs are more efficient than other kinds of neural network when dealing with time series data. However, an RNN can only provide a short time memory and therefore often fails to perform satisfactorily when applied to signals embodying relatively long-term dependencies. With the aim of dealing with data containing more persistent time information, the long short-term memory network was introduced. LSTM networks have proved to be more suitable for regression-based problems especially with input data which contains long term time information.

The normal LSTM model passes the hidden state in the forward direction. In other words, the information in the previous moments is considered in the processing but the information in future moments is not. LSTM-based networks are limited when it comes to complicated tasks, which prompted the development of bidirectional LSTM [25] [26], employing both the past and future features in one timestep. It is constructed from the same cell as LSTM but in both forward and backward directions, making it bi-directional. The normal output of a bi-directional LSTM usually only involves the last time-step of hidden layers. This is logical but not stable and is thus not able to estimate a continuous joint angle sequence effectively. Therefore, we have modified the Bi-LSTM to process the sEMG signals.

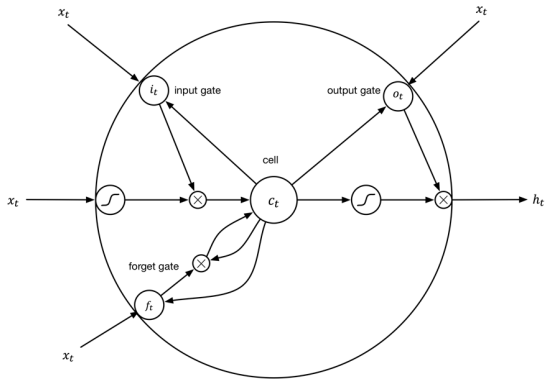


Fig. 4. The internal structure and workflow of a long short-term memory cell.

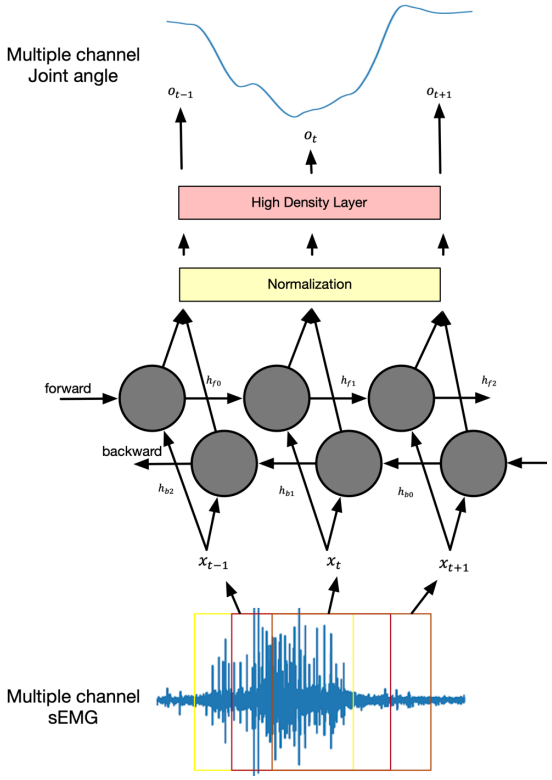


Fig. 5. Functional elements of the Bi-LSTM. The model includes both forward and backward timesteps, enabling the Bi-LSTM to utilize global features embedded in the sEMG signal where h_{f0}, h_{f1}, h_{f2} represent the hidden states output in the forward direction from moments $t-1$ to $t+1$ and h_{b0}, h_{b1}, h_{b2} stand for the hidden states output in the backward direction from moment $t+1$ to $t-1$.

Essentially, each LSTM unit consists of an input gate, an output gate, a forget gate and a memory cell [27], with weights and bias applied between any two gates. Fig. 4 illustrates the basic memory cell structure which can be represented by the following mathematical model (3)-(7):

$$i_t = \sigma(W_{x_i}x_t + W_{h_i}h_{t-1} + W_{c_i}c_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (4)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (5)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (6)$$

$$h_t = o_t tanh(c_t)h_{b0} \quad (7)$$

where x is the input, σ represents the activation function, i, f, o and c stand for the input gate, forget gate, output gate and cell state vectors respectively, all of which have the same size as the hidden vector h. All the vector subscripts (containing t or $t-1$) stand for the vector state at time t or $t-1$. The weight matrix subscripts, as the name suggests represent the magnitudes of the weight matrix. For instance, W_{hc} is the hidden-cell gate matrix while W_{xo} is the input-output gate matrix.

The Bi-LSTM contains both forward and backward pathways rather than only the forward, as shown in the Fig. 5. The information flowing in two directions fundamentally enlarges the receptive field of the Bi-LSTM. In the forward direction, h_{f0}, h_{f1}, h_{f2} stand for the cell outputs at times $t-1$, t, $t+1$, respectively. Similarly, in the backward direction h_{b0}, h_{b1}, h_{b2} represent the cell outputs at times $t+1$, t, $t-1$. Therefore, the outputs of Bi-LSTM should be $\{[h_{f0}, h_{b2}], [h_{f1}, h_{b1}], [h_{f2}, h_{b0}]\}$ rather than $\{h_{f0}, h_{f1}, h_{f2}\}$, which extends and enhances the information content of the network. Because the aim of our study was to estimate the continuous joint angle curve associated with the upper limb kinematics, the outputs should be strongly related. Therefore, all the outputs were normalized and fed into the same high-density layer during each timestep, which tightly bounded the outputs.

A backpropagation [28] technique was adopted as the training strategy. Both the conventional LSTM and Bi-LSTM were set to 5 layers containing 32 hidden units in each layer. In a preliminary experiment, it was found that the performance of the network improved as the number of layers increased up to and including 5 but rapidly deteriorated after 6. Again, following a preliminary experiment, the dropout [29] ratio was set to 0.3 to prevent overfitting. It was found that setting it to 0 would lead to overfitting and setting it to 0.5 reduced the convergence speed. The dimension of the high-density layer was set to double the size of hidden layer in order to accommodate the bidirectional information. During the training process, the data was shuffled when applied to the Bi-LSTM and LSTM models. Although this increases the convergence time, it was found to improve the performance of the model.

2) Comparison of the Models

The multilayer perceptron (MLP), considered to be the most basic artificial neural network [30], [31] has been used many times in investigating SPC systems [14], [32]. Therefore, we employed the MLP as the baseline model for comparison in this study. The MLP has been proved to have the ability to fit or approach any continuous curve or sequence. In this study, as a control group, three layers of non-fully connected networks were set as the hidden layers. Before sending the data into the networks, it was flattened into a 1-dimensional form (except for the batch dimension). The first layer expanded the data to 512 points. And the last two layers compressed the expanded data into 4 channels (to predict the joint angles). Again, in order to prevent the models from overfitting, the spatial dropout at 0.3 was added between every layer (only applied to the training step).

Convolutional neural networks have been extraordinarily successful in computer vision [33], intention recognition [34] etc. Compared with MLPs, CNNs have higher feature extraction ability and are more robust because their structure includes convolutional layers and pooling layers. In order to balance the quantity of trainable parameters, the hidden units were set to the following sizes 32, 64, 128, 256, 128, 64. Specifically, the convolutional operation was one-dimensional. The channels were independently convolved with a kernel of size 7. The convolution stride was 1 and the padding was set to 3 to avoid the change of window size. Additionally, the max pooling size was 2 for the feature size reduction, which prevented overfitting problems. Following this, all the output features were compressed into 4 channels with a dense layer as shown in Fig. 6. Again, the dropout rate was set to 0.3 as explained above. All the models were optimized with the Adam optimizer [35] and all input data was shuffled in the training process.

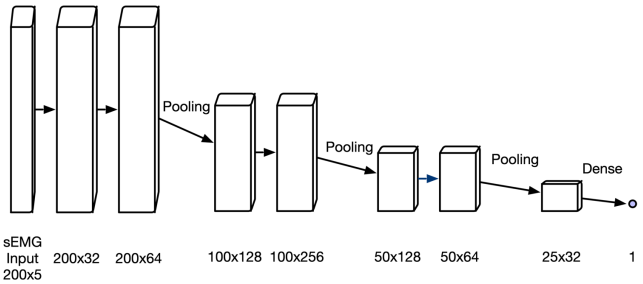


Fig. 6. The structure of CNN model employed in this study.

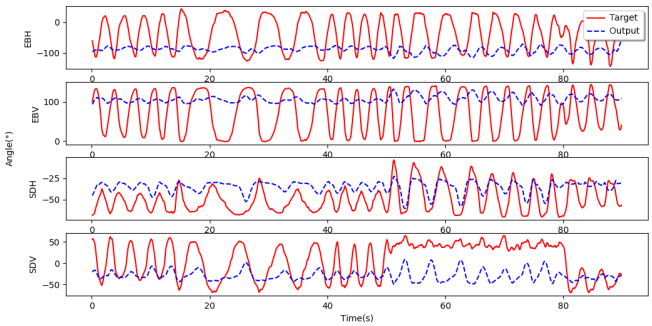
3) Statistical Analysis

Using the Kolmogorov-Smirnov test, it was found that the data were not normally distributed so non-parametric tests were used. The Friedman test was adopted, to assess significant differences among the subjects. If the Friedman test revealed a statistically significant effect, post-hoc comparisons were performed using the Wilcoxon sign-rank test, with the p values adjusted by the Bonferroni correction for multiple comparisons. p values < 0.05 were taken as the threshold of statistical significance.

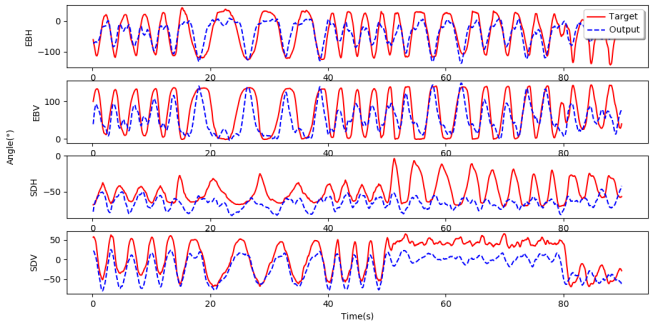
III. Results

Before a detailed analysis of the proposed and comparison networks, a performance comparison was made between them with the raw sEMG signal and the extracted feature (RMS) as inputs. (See Fig. 7, in which only the comparison for the CNN is presented due to space limitations).

It can be seen that the performance of the CNN was greatly improved by adopting feature extraction. Because of its unique convolutional structure, the CNN model was the only one that could handle the unprocessed sEMG signal whereas the MLP and LSTM could barely decode the raw sEMG signal without



(a)



(b)

Fig. 7. Performance comparison between (a) CNN model with raw input (b) CNN model with RMS feature input (See table II for an explanation of the y-axis labels. The red line represents the actual joint angles and the blue line indicates the predicted joint angles).

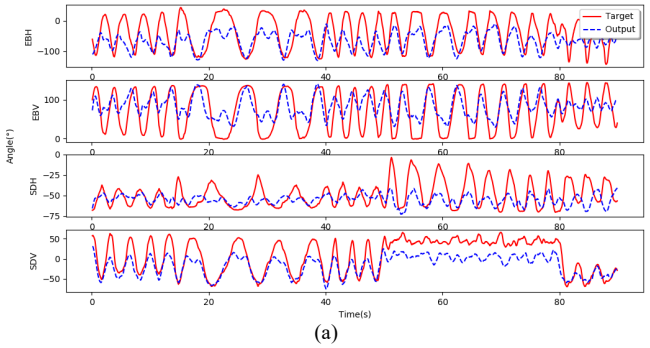
TABLE III
INDIVIDUAL SUBJECT CORRELATION COEFFICIENTS BETWEEN DIRECTLY MEASURED AND ESTIMATED JOINT ANGLES FOR EACH ANALYSIS MODEL (MOVEMENTS 1-7 FOR TRAINING ONLY 8 FOR TESTING)

Subject	MLP	CNN	LSTM	Bi-LSTM
1	0.751±0.188	0.785±0.111	0.835±0.102	0.863±0.052
2	0.659±0.190	0.691±0.125	0.739±0.046	0.787±0.038
3	0.660±0.185	0.761±0.085	0.842±0.023	0.872±0.075
4	0.769±0.090	0.769±0.109	0.700±0.290	0.834±0.069
5	0.687±0.127	0.712±0.106	0.798±0.118	0.839±0.025
6	0.701±0.159	0.711±0.147	0.745±0.340	0.753±0.258
7	0.717±0.095	0.789±0.055	0.848±0.072	0.871±0.086
8	0.694±0.159	0.664±0.096	0.857±0.151	0.881±0.144

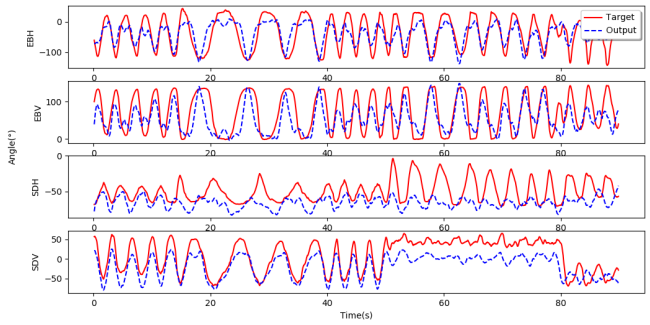
feature extraction. However, after having applied feature extraction, all models were able to learn the mapping relationship between sEMG features and joint angles.

After the comparison of raw data input and RMS input, all the models' hidden layer parameter sizes were kept below 100,000. For equitable comparison, all the four models were developed in Pytorch [36]. All the comparisons between the performance of the models were made on the free movement scenario (i.e., No.8 in Table I). As an example, the performance of each model on subject 2 is compared in Fig. 8. This shows the variation of measured joint angles with time in red and the predicted values in blue, (labelled “target” and “output”, respectively).

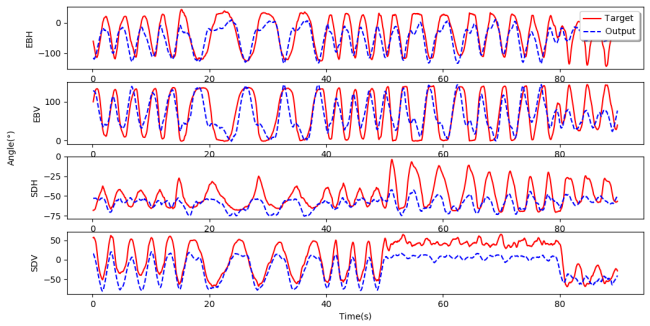
Panel (a) of the figure shows that the MLP model is barely able to match the pattern of any of the four measured joint angles and performed particularly badly when estimating the SDH angle. Relatively, the CNN model performed better, but it



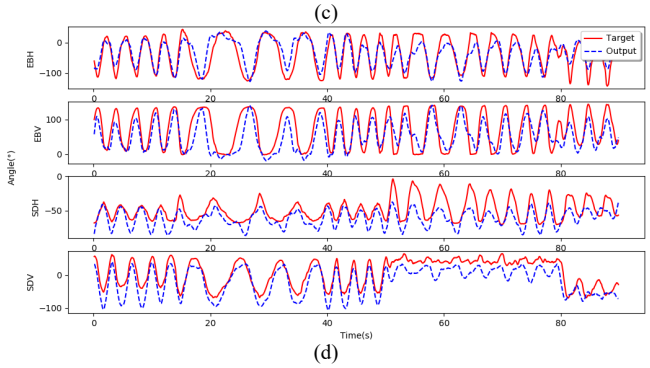
(a)



(b)



(c)



(d)

Fig. 8. Estimated outputs compared to the actual measured joint angles based on, a) MLP model (b) CNN model (c) LSTM model (d) Bi-LSTM model.

became unstable whenever a large change in the joint-angle occurred. The LSTM performed markedly better for the estimation of the elbow joint angle planes although its performance for the SDH was inferior. However, the Bi-LSTM compensated for the shortcomings of the LSTM and CNN, and achieved the best performance when compared to the other models. The Pearson correlation coefficient (cc) was used to provide a quantitative measure of the association between the joint angles estimated from the sEMG signals and those obtained by direct measurement for the 4 methods as applied to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LRA.2021.3097272, IEEE Robotics and Automation Letters

each subject, as shown in Table III. For conciseness, the mean and standard deviation of channels (EBH, EBV, SDH, SDV) were calculated for each subject and each model.

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 Bi-LSTM model ($cc = 0.8374 \pm 0.0455$) significantly outperformed the MLP ($cc = 0.7048 \pm 0.0395, p = 0.036$), the CNN ($cc = 0.735 \pm 0.0468, p = 0.036$) as well as the LSTM ($cc = 0.7952 \pm 0.0598, p = 0.036$). The superiority of the proposed method over the existing methods is further seen by the cc values of each subject, as shown in Fig. 9. (Error bars represent standard deviations of the mean values shown).

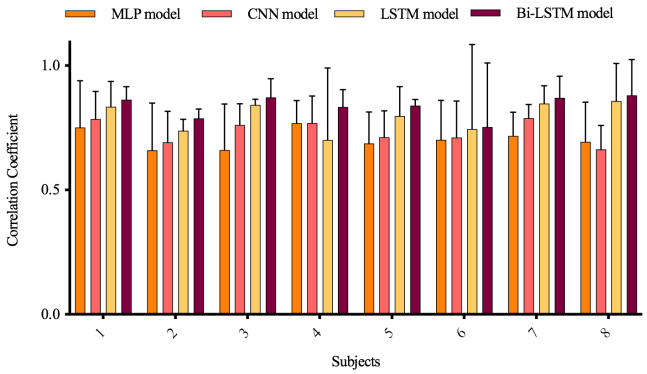


Fig. 9. Individual correlation coefficient results for each subject.

IV. Discussion

In this study, the Bi-LSTM was used to obtain indirectly estimated joint angles based on noninvasively collected sEMG signals and the results compared with those derived from the MLP, CNN, and normal LSTM models. The experimental results on untrained random movement patterns proved that the Bi-LSTM model was better able to exploit the kinematic information embedded in the sEMG signal than the MLP, CNN, and LSTM methods and that, in each case, these differences were statistically significant.

Taken together the overall results show that, due to the adoption of the training protocol, in which the joint angles were not synchronous with sEMG signals, it was a challenge for the existing models to estimate the joint angles adequately. For instance, the MLP model was neither capable of predicting the cyclical movement pattern nor the amplitude of the joint angle precisely because of the information loss caused by the flattening operation at the start of the data pathway. The CNN model was clearly more adaptive in exploiting the structural features of the input due to the inherent convolution operation, but it resulted in some unwanted noise. The LSTM provided smoother and more accurate joint angle curves when compared to the MLP and CNN models, since it was able to combine a priori information in its memory cell. Nevertheless, the decoding of the SDH joint angle remains a challenge. The Bi-LSTM compensated for this shortcoming due to its bidirectional structure and, overall, yielded the best estimation results. According to Fig. 9, the average CC of Bi-LSTM exceeded those of the MLP, CNN and LSTM models in every subject. However, from Fig. 9 it can be seen that the LSTM standard deviation is markedly higher in two subjects (4 and 6). For these subjects the performance of RNN-type networks was poorer than MLP and CNN on the SDV channel. We found that there are some problems with the data quality, leading to the high standard deviations. But compared to the uni-directional LSTM, the Bi-LSTM effectively suppressed the noise. Having analyzed the detailed results for each channel, we speculate that the abnormal performance of the RNN-based neural network is inevitable especially when data quality is low. A similar effect was reported in [37]. To tackle this problem in future work, one might consider estimating joint angles in specific channels (e.g., SDV) with an independent neural network.

Notably, in our previous upper limb joint angle estimation experiment, we compared the Bi-LSTM and LSTM approaches using data from the same limb. In that case the Bi-LSTM did not significantly outperform the normal LSTM. We speculate that, because the joint angles and sEMG were recorded from the same limb, both signals were closely synchronized. The joint angles were mainly associated with the sEMG signals before the current moment, and could, therefore be modeled by the LSTM; thus, performing as effectively as the Bi-LSTM. However, in this study, the Bi-LSTM significantly outperformed LSTM because, due to its bidirectional nature it is able to detect the association between the sEMG and joint angle signals even though they are only poorly synchronized.

According to the results of the Wilcoxon signed-rank test, the superior performance of the Bi-LSTM when compared to that of MLP, CNN and normal LSTM was statistically significant. This shows that when simulating the frequently occurring lack of synchronization between joint angles and their associated sEMG signals, the Bi-LSTM approach has proved to be the most robust solution to this lack of synchrony. Because of its ability to tolerate errors due to loss of synchronization, the technique proposed here lends itself well to practical applications which require the continuous estimation of limb movement. For instance, currently we are not able to collect accurate joint angle and sEMG signals simultaneously from amputees. With the technique described here, amputees could follow predetermined movement instructions while their sEMG signals are recorded. Then, the proposed algorithm could build relationships between the non-synchronized sEMG signals and joint angles.

We also note that all the measurements described here were carried out with the subjects standing and thus with gravity acting vertically through them. In future work we will investigate this limitation by exploring how changes in posture affect the timing and amplitude of the sEMG signals and their relationship with the measured joint angles.

V. Conclusions

Based on the weakly synchronized measurements, the kinematic information extracted from the sEMG signals of the muscles in the dominant arm can be used to simultaneously and

proportionally estimate the joint angles of the corresponding movements on the contralateral, non-dominant side. The results show that the Bi-LSTM model achieved a significantly better performance than the normal LSTM as well as the MLP and CNN models. In comparison to the commonly used network models, its time and memory requirements are modest and it is more accurate. Entirely untrained free movements were used for testing, with the aim of simulating the unpredictable movement patterns that occur in practice.

Whereas for non-medical applications, large amounts of training data are relatively easy to obtain, for ethical and practical reasons this is much harder for sEMG data. And this lack of data exacerbates the overfitting problem and makes it difficult to generalize. In future work, we will attempt to combine conventional machine learning methods and deep learning algorithms to focus on more effective information extraction and thus realize an effective remote movement estimation of a practical robotic arm.

REFERENCES

[1] C.-M. Kuo, L.-C. Chen, and C.-Y. Tseng, "Investigating an innovative service with hospitality robots," *International Journal of Contemporary Hospitality Management*, vol. 29, no. 5, pp. 1305–1321, 2017.

[2] G. E. Marchant *et al.*, "International governance of autonomous military robots," *Handbook of Unmanned Aerial Vehicles*, pp. 2879–2910, 2015.

[3] M. M. Veloso, J. Biswas, B. Coltin, and S. Rosenthal, "CoBots: Robust Symbiotic Autonomous Mobile Service Robots.," in *IJCAI*, 2015, p. 4423.

[4] T. K. Morimoto, E. W. Hawkes, and A. M. Okamura, "Design of a compact actuation and control system for flexible medical robots," *IEEE robotics and automation letters*, vol. 2, no. 3, pp. 1579–1585, 2017.

[5] J. Meng, S. Zhang, A. Bekyo, J. Olsoe, B. Baxter, and B. He, "Noninvasive electroencephalogram based control of a robotic arm for reach and grasp tasks," *Scientific Reports*, vol. 6, p. 38565, 2016.

[6] R. Szabó and A. Gontean, "Controlling a robotic arm in the 3D space with stereo vision," in *2013 21st Telecommunications Forum Telfor (TELFOR)*, 2013, pp. 916–919.

[7] P. Kopniak and M. Kamiński, "Natural interface for robotic arm controlling based on inertial motion capture," in *2016 9th International Conference on Human System Interactions (HSI)*, 2016, pp. 110–116.

[8] D. Farina *et al.*, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797–809, 2014.

[9] T. J. Zebo, "Myoelectric control of the Rancho electric arm," 1968.

[10] J.-S. Han, W.-K. Song, J.-S. Kim, W.-C. Bang, H. Lee, and Z. Bien, "New EMG pattern recognition based on soft computing techniques and its application to control of a rehabilitation robotic arm," in *Proc. of 6th International Conference on Soft Computing (IIZUKA2000)*, 2000, pp. 890–897.

[11] Q. C. Ding, A. B. Xiong, X. G. Zhao, and J. D. Han, "A novel EMG-driven state space model for the estimation of continuous joint movements," in *2011 IEEE International Conference on Systems, Man, and Cybernetics*, 2011, pp. 2891–2897.

[12] C. Lin, B. Wang, N. Jiang, and D. Farina, "Robust extraction of basis functions for simultaneous and proportional myoelectric control via sparse non-negative matrix factorization," *Journal of Neural Engineering*, vol. 15, no. 2, p. 26017, 2018.

[13] M. C. Tresch, P. Saltiel, and E. Bizzi, "The construction of movement by the spinal cord," *Nature neuroscience*, vol. 2, no. 2, p. 162, 1999.

[14] S. Muceli and D. Farina, "Simultaneous and proportional estimation of hand kinematics from EMG during mirrored movements at multiple degrees-of-freedom," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, pp. 371–378, 2012.

[15] C. Lin, B. Wang, N. Jiang, and D. Farina, "Robust extraction of basis functions for simultaneous and proportional myoelectric control via sparse non-negative matrix factorization," *Journal of neural engineering*, vol. 15, no. 2, p. 26017, 2018.

[16] C. Ma, C. Lin, O. W. Samuel, L. Xu, and G. Li, "Continuous estimation of upper limb joint angle from sEMG signals based on SCA-LSTM deep learning approach," *Biomedical Signal Processing and Control*, vol. 61, p. 102024, 2020.

[17] 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, 2005.

[18] N. J. and A. M. Alex Graves, "HYBRID SPEECH RECOGNITION WITH DEEP BIDIRECTIONAL LSTM," pp. 273–278, 2013.

[19] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN," *Expert Systems with Applications*, vol. 72, pp. 221–230, 2017.

[20] N. Jiang, K. B. Englehart, and P. A. Parker, "Extracting simultaneous and proportional neural control information for multiple-dof prostheses from the surface electromyographic signal," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1070–1080, 2009.

[21] N. A. Shrirao, N. P. Reddy, and D. R. Kosuri, "Neural network committees for finger joint angle estimation from surface EMG signals," *BioMedical Engineering Online*, vol. 8, pp. 1–11, 2009.

[22] C. Wang, W. Guo, H. Zhang, L. Guo, C. Huang, and C. Lin, "sEMG-based continuous estimation of grasp movements by long-short term memory network," *Biomedical Signal Processing and Control*, vol. 59, p. 101774, 2020.

[23] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate research*, vol. 30, no. 1, pp. 79–82, 2005.

[24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[25] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, no. 6, pp. 6645–6649, 2013.

[26] A. Graves, S. Fernández, and J. Schmidhuber, "Bidirectional LSTM networks for improved phoneme classification and recognition," in *International Conference on Artificial Neural Networks*, 2005, pp. 799–804.

[27] F. Beaufays, H. Sak, and A. Senior, "Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling Has," *Interspeech*, no. September, pp. 338–342, 2014.

[28] M. Boden, "A guide to recurrent neural networks and backpropagation," *the Dallas project*, 2002.

[29] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[30] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[31] K. Hornik, M. Stinchcombe, H. White, M. W. Gardner, and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmospheric environment*, vol. 32, no. 14–15, pp. 2627–2636, 1998.

[32] J.-U. Chu, I. Moon, and M.-S. Mun, "A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand," *IEEE Transactions on biomedical engineering*, vol. 53, no. 11, pp. 2232–2239, 2006.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[34] X. Ji, Q. Zhao, J. Cheng, and C. Ma, "Exploiting spatio-temporal representation for 3D human action recognition from depth map sequences," *Knowledge-Based Systems*, vol. 227, p. 107040, 2021.

[35] J. Lehman and K. O. Stanley, "Revising the evolutionary computation abstraction," in *Proceedings of the 12th annual conference on Genetic and evolutionary computation - GECCO '10*, 2010, p. 103.

[36] N. Ketkar, "Introduction to PyTorch," in *Deep Learning with Python*, N. Ketkar, Ed. Berkeley, CA: Apress, 2017, pp. 195–208.

[37] F. Quivira, T. Koike-Akino, Y. Wang, and D. Erdogmus, "Translating sEMG signals to continuous hand poses using recurrent neural networks," in *2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018*, 2018, vol. 2018-Janua, no. March, pp. 166–169.