Hand-movement Prediction from EMG with LSTM-Recurrent Neural Networks

M. A. García-Vellisca, A. Matran-Fernandez✉, R. Poli and L. Citi

BCI-NE group, School of Computer Science and Electronic Engineering

University of Essex

Wivenhoe Park, Colchester CO4 3SQ, UK

Email: amatra@essex.ac.uk

Abstract—In this paper we present an approach based on shallow recurrent long short-term memory neural networks for the prediction of hand kinematics for hand-prosthesis control from data acquired via high-density surface electromyography (HD-sEMG). We used 134-channel HD-sEMG recordings from seven participants while performing multiple repetitions of 13 hand movements. A CyberGlove II was used to simultaneously record 18 degrees of freedom (joint angles) used as ground truth for predicting the hand movements. Traditional features were calculated over 100 ms windows and fed to the network. Specifically we used: Mean Absolute Value (MAV), variance, and number of zero-crossings. Our results indicate that: (a) a small number of channels is sufficient to make accurate predictions, (b) many features are redundant, and MAV is sufficient for the job, (c) the simple neural network architecture we propose is effective in this task. These findings have important implications in terms of computational efficiency and memory storage, which are important considerations in relation to implementability in the typically very low-power and low-resources computers on-board of hand prostheses.

I. INTRODUCTION

Pattern recognition (PR) has been used intensively during the last decades to improve the control of prostheses [1], [21]. In a typical PR setting, the raw EMG is segmented in time-windows and, from these, a set of features is extracted and classified using standard classification algorithms [15], [16], [20]. In spite of this significant research effort, PR-based control approaches have had limited application in commercial prostheses. One main reason is that they do not allow simultaneous and proportional control (SPC) of two or more degrees of freedom (DoFs) [8], [2], [17].

Particularly, ANNs have shown to be very useful for time series prediction, including the possibility of using them for SPC of multiple degrees of freedom [4], [13].

Some publications have shown promising results using the muscle EMG activity for decoding hand movement intention combining the computation of many features with early prediction systems [18], [9], [10]. Our approach uses long short-term memory recurrent neural networks (LSTM-RNN) as a regression method to predict the kinematics of the hand from the muscle activations captured through surface EMG. In the following we test whether sufficient accuracy can be attained using only one time-domain feature from a reduced set of electrodes in an attempt to reduce the computational cost of the model. This approach might help reduce the response time and the battery consumption when implemented on an embedded system, like those typically found within a prosthetic-hand, such as the one developed within the Horizon 2020 "DeTOP" project.

II. METHODOLOGY

In this study, we use EMG data and joint angles from the first 7 participants in the SEEDS database [12]. Eight monopolar electrodes are used to record the signals of the extensor muscles of the forearm, while a set of 126 electrodes arranged in an high-density EMG (HD-sEMG) array of 9 rows by 14 columns is used for recording the signals from the flexor muscles. Therefore, we have access to 134 channels. Hand kinematics were simultaneously recorded through a CyberGlove II with 18 sensors placed in the locations shown in Fig. 1. Due to space constraints, and considering the lack of substantial differences across the 7 participants, for brevity in the following sections we present figures from the first one only.

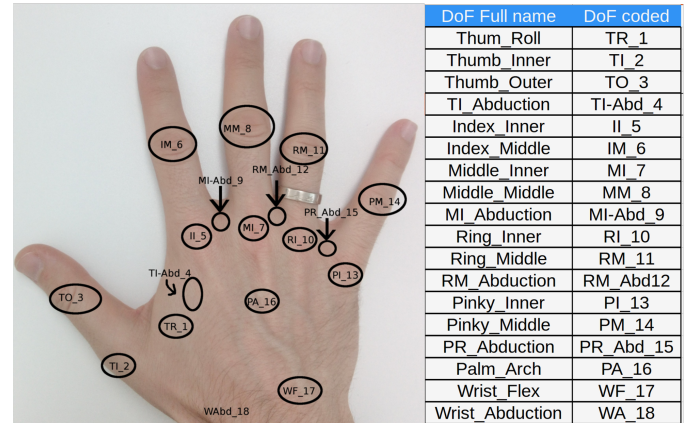| DoF Full name	DoF coded
Thum_Roll	TR_1
Thumb_Inner	TI_2
Thumb_Outer	TO_3
TI_Abduction	TI-Abd_4
Index_Inner	II_5
Index_Middle	IM_6
Middle_Inner	MI_7
Middle_Middle	MM_8
MI_Abduction	MI-Abd_9
Ring_Inner	RI_10
Ring_Middle	RM_11
RM_Abduction	RM_Abd12
Pinky_Inner	PI_13
Pinky_Middle	PM_14
PR_Abduction	PR_Abd_15
Palm_Arch	PA_16
Wrist_Flex	WF_17
Wrist_Abduction	WA_18

Fig. 1. (Left) DoF location for recording hand movements. (Right) Correspondence between the full name of the CyberGlove sensors and our coded names [12].

A. Channel Reduction

We started by reducing the number of available electrodes in a first attempt to reduce the computational load to train our neural network. The set of electrodes used in this work was comprised of:

- The 8 monopolar electrodes placed on the extensor muscles;

- The 9 electrodes in the seventh (i.e., central) column of the HD-sEMG array, in order to have a transverse scan of the anterior part of the forearm, to gather information from the flexor muscles.

This reduced the number of electrodes from 134 to 17.

B. Decoding and Feature Selection

For decoding myoelectric information we used adjacent windowing. We used non-overlapping windows of 100 ms, since previous studies have shown this to be a good trade-off between responsiveness and accuracy [5].

Within the literature, a wide range of features have been proposed in the time domain, in the frequency domain or in mixed domains (frequency-time and scale-time) [14]. In particular, the methods for extracting features in the time domain can be divided into four groups: 1) energy and complexity methods, 2) frequency information methods, 3) prediction-model methods, and 4) time-dependence methods. According to [15], the features that belong to the first and second groups give better results. As a consequence, we initially chose three computed from the time domain, which fall into the first two groups:

- Mean Absolute Value (MAV) that essentially captures the energy and complexity of the signal:

$$\text{MAV} = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{1}$$

- Variance (VAR) that is related to the power of the force developed by the muscle [21]

$$\text{VAR} = \frac{1}{N-1} \sum_{i=1}^{N} x_i^2 \tag{2}$$

- The number of Zero-Crossings (ZC) in the window, which is a time-domain measurement that captures frequency information.

$$\text{ZC} = \sum_{i=1}^{N} \text{neg}(x_i x_{i+1}), \tag{3}$$

where

$$\text{neg}(a) = \begin{cases} 1, & \text{if } a < 0, \\ 0, & \text{otherwise,} \end{cases}$$

and x_i is the i-th sample of the EMG signal band-pass filtered between 10 and 500 Hz.

C. Recurrent Neural Network

We use a form of recurrent neural networks (RNNs) to predict a hand movement based on the EMG records taken from the forearm. Due to the nature of the movement, the order in which the temporal sequence of data is given is extremely important. RNNs are particularly well suited to deal with this situation, and different studies have shown good results using RNNs [19], [6], [10].

A problem affecting traditional RNNs is the so-called vanishing-gradient problem (the weight updates rapidly becoming so small that the network can hardly learn) which make them difficult to train with gradient-based learning methods and backpropagation [7]. One solution to this problem is offered by Long Short-Term Memory (LSTM) RNNs, where the gradient hardly vanishes [11], [3]. In this study we designed and compared the performance of two LSTM-RNNs:

- One with 17 inputs (which correspond to the VAR feature computed for each of the 17 electrodes used), an LSTM hidden layer with 9 neurons, and 18 outputs (one for each DoF);
- One with 51 inputs (where we extracted MAV, VAR and ZC from each of the 17 electrodes), an LSTM hidden layer with 9 neurons, and 18 outputs (as above).

We trained individual models for each of the first 7 participants in the SEEDS dataset [12], using the first two sessions for training (80%) and validation (20%), and the third one for testing. Each of the sessions includes 6 repetitions at different speeds of 13 different movements, performed in random order.

For each model, we computed a time-series prediction for the six repetitions from each movement from the test set. To assess the error between target and prediction we used two metrics: mean absolute error (MAE) which can be calculated as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2 \tag{4}$$

and the coefficient of determination (R^2) defined as the relative improvement over a baseline constant predictor:

$$R^2 = \frac{(\text{MSE}_0 - \text{MSE})}{\text{MSE}_0} \tag{5}$$

with

$$\text{MSE}_0 = \frac{1}{n} \sum_{i=1}^{n} (z - x_i)^2 \tag{6}$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2 \tag{7}$$

where z is the mean target on traning set, y_i is the prediction, x_i the target for time-window i and n is the total number of time windows.

Obviously, not all DoFs are involved in meaningful ways in all movements. To make our performance evaluation more meaningful we used knowledge of the anatomy and function of the hand to identify the subset of the DoFs that contribute to each individual movement. These will be used when we report results in the next section.

We used R^2 obtained on the test set to compare the performance of our two LSTM-RNNs by considering the null hypothesis H_0 that the LSTM-RNN with 17 inputs is as

accurate as the one with 51 inputs. Due to the non-Gaussian distribution of the R^2 values computed, we used a Wilcoxon-test with a confidence level of $\alpha = 0.05$.

III. RESULTS

A. LSTM-RNN with 51 inputs

Let us start with the case of the LSTM-RNN with 51 inputs, i.e., where we rely on three EMG features: MAV, VAR and ZC. As an example of the typical test-set performance achieved, in Fig. 2 we show 6 repetitions (3 fast and 3 slow) of the fist movement as measured on the Index-Inner CyberGlove sensor (II_5 in Fig. 1). As one can visually appreciate (and verify through the low MAE and a R^2 value close to one, reported at the top of the figure), this LSTM-RNN predicts with a good degree of accuracy.
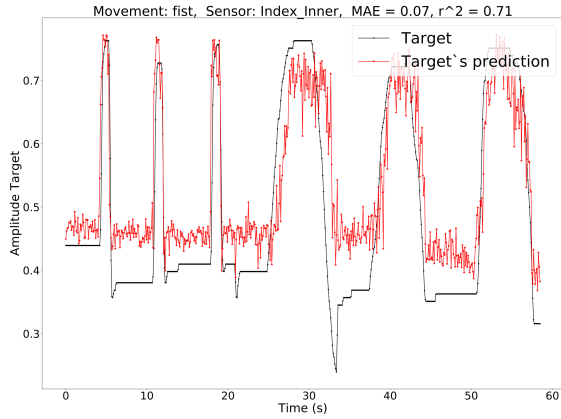

Fig. 2. LSTM-RNN with 51 inputs: time series for the recorded joint angle and the network's prediction for the 'fist' movement, from the Index-Inner (II_5) DoF.

Fig. 3 reports the R^2 value obtained in the test set separately for each of the 18 DoF and 13 movements, giving a total of 234 combinations. We have marked with a black dot the cells for which the DoF is meaningful for a given movement (based on knowledge about the hand function). From a prediction point of view, one can see many cells dark blue which means the time-series prediction performs better than a horizontal line predicting the mean on the training set ($R^2 = 0$).

B. LSTM-RNN with 17 inputs

The LSTM-RNN with 17 inputs relies only on one feature, namely, the MAV, extracted from each electrode. As an example of the performance achieved on the test set and for comparison with the 51-input model, Fig. 4 shows the prediction on the same movement ("fist") and DoF as that of Fig. 2. It can be visually appreciated taht the predictions are very close to the ground truth time series. Also, comparing Figs. 2 and 4 we see that both qualitatively and quantitatively (MAE and R^2) results are very similar.
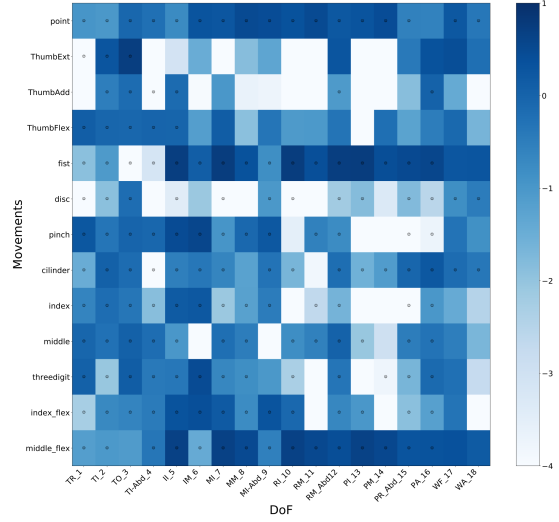

Fig. 3. LSTM-RNN with 51 inputs: R^2 value for all possible combinations between DoF and movements. Dots in a cell represent that the DoF is meaningful for the movement.
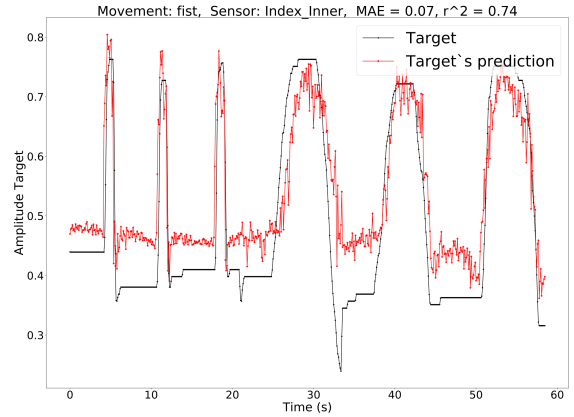

Fig. 4. LSTM-RNN with 17 inputs: Time series with both joint angle and its prediction for the 'fist' movement, from the Index-Inner (II_5) DoF.

The test-set R^2 obtained for the 234 combinations of 18 DoF and 13 movements by the LSTM-RNN with 17 inputs is reported in Fig. 5 as a colour map. Once again, for each movement we have marked which DoFs are meaningful (i.e., based on knowledge of the hand function, which DoFs one would expect to detect variation for that movement). Comparing Fig. 3 and Fig. 5, there are some slight differences based on R^2 due to the network with 51 inputs is trained with more information than in the case with 17 inputs.

A visual comparison of Figs. 3 and 5 shows that essentially the two architectures provide comparable level of performance. This has also been verified by applying the Wilcoxon-
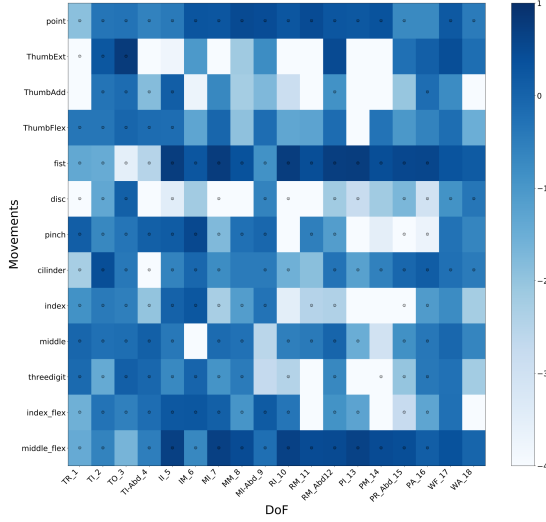
Fig. 5. LSTM-RNN with 17 inputs: R^2 value for all possible combinations between DoF and movements. Dots in a cell represent that the DoF is meaningful for the movement

test to the R^2 computed across the 234 combinations, which indicates that the distributions are *not* statistically different (i.e., the null hypothesis H_0 formulated at the end of section II cannot be rejected at the chosen level of significance).
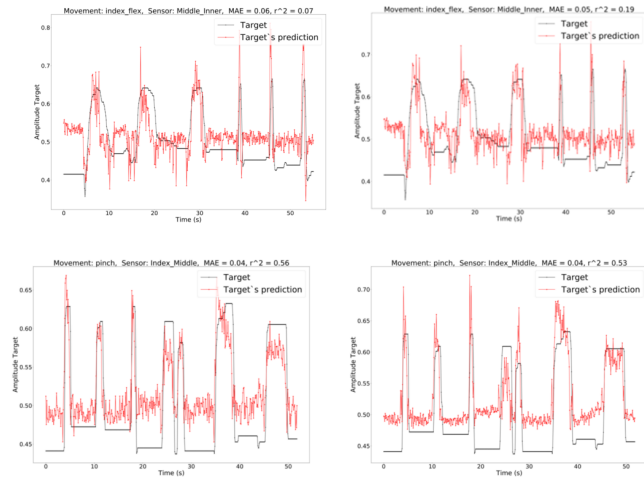

Fig. 6. Time series with both joint angle and its prediction for two different movements: Index flexion (top) and pinch (bottom), both seen from different DoF. Results computed for two different cases: (From left to right) using 3 features, using only MAV.

IV. CONCLUSION

In this paper we presented an approach to predicting hand kinematics for hand-prosthesis control from data acquired via high-density EMG (HD-EMG), from seven subjects, using

LSTM-RNNs. The original HD-EMG recordings included 134 channels and related to 13 hand movements during which 18 DoFs were recorded. Because handling such a large number of channels is expensive with the typical processors on board available on hand prostheses, we decided to verify whether the number of channels could be reduced to 17 while maintaining a comparable performance.

For the same reasons, we proposed and tested two relatively small LSTM-RNN topologies: one relying on three features (MAV, VAR and ZC) computed every 100 ms for each channel, and a more efficient one using only one feature, MAV, per channel. Both topologies had one hidden layer with only 9 neurons and one output layer with 18 neurons (as many as sensors recorded).

Our results show that, despite the extreme parsimony of these architectures, both can predict to a sufficient degree the trajectories of these joints over 13 key hand-movements. Furthermore, we found that the two architectures provide almost indistinguishable results, indicating that one input feature, namely MAV, is sufficient for the task at hand.

While more research is clearly necessary to see whether even smaller architectures (e.g., fewer hidden neurons and fewer input channels) would be sufficient, we believe that both are implementable in the low-power hardware available in the hand-prosthesis developed in the DeTOP project.

REFERENCES

[1] A. L. Ciancio, F. Cordella, R. Barone, R. A. Romeo, A. D. Bellingegni, R. Sacchetti, A. Davalli, G. Di Pino, F. Ranieri, V. Di Lazzaro, E. Guglielmelli, and L. Zollo, "Control of prosthetic hands via the peripheral nervous system," *Frontiers in neuroscience*, vol. 10, pp. 116–116, Apr 2016.

[2] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "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, July 2014.

[3] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018.

[4] J. M. Hahne, F. Bießmann, N. Jiang, H. Rehbaum, D. Farina, F. C. Meinecke, K. . Müller, and L. C. Parra, "Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 269–279, March 2014.

[5] M. Hakonen, H. Piitulainen, and A. Visala, "Current state of digital signal processing in myoelectric interfaces and related applications," *Biomedical Signal Processing and Control*, vol. 18, pp. 334 – 359, 2015.

[6] M. Hioki and H. Kawasaki, "Estimation of finger joint angles from semg using a recurrent neural network with time-delayed input vectors," in *2009 IEEE International Conference on Rehabilitation Robotics*, June 2009, pp. 289–294.

[7] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, pp. 107–116, 04 1998.

[8] N. Jiang, S. Dosen, K. Muller, and D. Farina, "Myoelectric control of artificial limbs—is there a need to change focus? [in the spotlight]," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 152–150, Sep. 2012.

[9] P. Koch, H. Phan, M. Maass, F. Katzberg, and A. Mertins, "Early prediction of future hand movements using semg data," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2017, pp. 54–57.

[10] ——, "Recurrent neural network based early prediction of future hand movements," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2018, pp. 4710–4713.

[11] J. Li, A. Mohamed, G. Zweig, and Y. Gong, "Lstm time and frequency recurrence for automatic speech recognition," in *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, Dec 2015, pp. 187–191.

[12] A. Matran-Fernandez, I. J. Rodríguez Martínez, R. Poli, C. Cipriani, and L. Citi, "Seeds, simultaneous recordings of high-density emg and finger joint angles during multiple hand movements," *Scientific Data*, vol. 6, no. 1, p. 186, 2019.

[13] 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, May 2012.

[14] M. A. Oskoei and H. Hu, "Ga-based feature subset selection for myoelectric classification," in *2006 IEEE International Conference on Robotics and Biomimetics*, Dec 2006, pp. 1465–1470.

[15] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for emg signal classification," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420–7431, 2012.

[16] J. Rafiee, M. A. Rafiee, F. Yavari, and M. P. Schoen, "Feature extraction of forearm emg signals for prosthetics," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4058–4067, 2011.

[17] A. D. Roche, H. Rehbaum, D. Farina, and O. C. Aszmann, "Prosthetic myoelectric control strategies: A clinical perspective," *Current Surgery Reports*, vol. 2, no. 3, p. 44, Jan 2014.

[18] T. Teban, R. Precup, E. Lunca, A. Albu, C. Bojan-Dragos, and E. M. Petriu, *Recurrent Neural Network Models for Myoelectricbased Control of a Prosthetic Hand*, Oct 2018.

[19] P. Xia, J. Hu, and Y. Peng, "Emg-based estimation of limb movement using deep learning with recurrent convolutional neural networks," *Artificial Organs*, vol. 42, no. 5, pp. E67–E77, 2018.

[20] M. Zardoshti-Kermani, B. C. Wheeler, K. Badie, and R. M. Hashemi, "Emg feature evaluation for movement control of upper extremity prostheses," *IEEE Transactions on Rehabilitation Engineering*, vol. 3, no. 4, pp. 324–333, 1995.

[21] M. Zecca, S. Micera, M. C. Carrozza, and P. Dario, "Control of multifunctional prosthetic hands by processing the electromyographic signal," *Critical Reviews™ in Biomedical Engineering*, vol. 30, no. 4-6, pp. 459–485, 2002.