EMG-Based Hand Gesture Classification with Long Short-Term Memory Deep Recurrent Neural Networks

Milad Jabbari, Rami N. Khushaba, *Senior Member*, *IEEE*, and
Kianoush Nazarpour, Senior *Member, IEEE*

***Abstract*— Electromyogram (EMG) pattern recognition has been utilized with the traditional machine and deep learning architectures as a control strategy for upper-limb prostheses. However, most of these learning architectures, including those in convolutional neural networks, focus the spatial correlations only; but muscle contractions have a strong temporal dependency. Our primary aim in this paper is to investigate the effectiveness of recurrent deep learning networks in EMG classification as they can learn long-term and non-linear dynamics of time series. We used a Long Short-Term Memory (LSTM-based) neural network to perform multiclass classification with six grip gestures at three different force levels (low, medium, and high) generated by nine amputees. Four different feature sets were extracted from the raw signals and fed to LSTM. Moreover, to investigate a generalization of the proposed method, three different training approaches were tested including 1) training the network with feature extracted from one specific force level and testing it with the same force level, 2) training the network with one specific force level and testing it with two remained force levels, and 3) training the network with all of the force levels and testing it with a single force level. Our results show that LSTM-based neural network can provide reliable performance with average classification errors of around 9% across all nine amputees and force levels. We demonstrate the applicability of deep learning for upper-limb prosthesis control.**

***Index Terms*— Electromyography signal, LSTM, prosthesis**

I. Introduction

People with limb difference face extreme challenges in performing daily tasks. Upper-limb myoelectric prostheses are used to facilitate one's functions. Commercial devices mostly use an on-off control strategy, which lacks the required functionality [1]. An alternative approach to the on-off control strategy for controlling the upper-limb prosthesis employs the multi-class electromyogram (EMG) pattern recognition to map muscle activation patterns to prosthesis functions. EMG classification has been used for the decoding of wrist and elbow movements [2], [3], grasp gesture [4], [5], and finger movements [6] with highly accurate performances (accuracy > 90%). Despite the considerable accuracies achieved, more efforts are needed to translate laboratory-based outcomes into clinical applications. A number of EMG classification methods for controlling powered prostheses were reported to lack the stability required for clinical prostheses performance [1], with reports showing that 88% of abandon the prosthesis because using them make users so tired and establishment of the prosthesis is difficult [7]. Hence, a reliable classifier and appropriate feature extraction method is essential.

Several researches have been performed to access high classification accuracy with promising results for clinical applications [5], [8]. Confounding factors, including the effect of the numbers of electrodes/sensors, force level variation, training strategies during the experiment and clinical settings, selecting the highly accurate classifier and appropriate feature extraction method have been considered. Khushaba et al. [9] proposed the time-domain power spectral descriptors (TD-PSD) could achieve classification accuracy up to 91%, when the classifier was trained with the EMG recorded at multiple forearm orientations with medium muscular contractions. Krasoulis et al. [8] proposed three discriminant analysis based classifier, LDA, quadratic discriminant analysis (QDA), and regularized discriminant analysis (RDA), to realize real-time control of the hand prosthetic. Al-Timemy et al. [10] introduced a novel feature extraction method to achieve high classification accuracy with six grip movements and three different force level produced by the subject. The results show that using LDA method with presented feature extraction algorithm, improvements of around 6% to 9% have been obtained in the average classification error among all amputees and for the three force levels. Other works have been performed on the data set generated in [10] and different methods have been evaluated to improve the average classification errors. Al-Ani et al [11] designed a dynamic channel selection algorithm for improvement of the classification errors in EMG data set generated in [10] and the promising results achieved. Also, a novel feature extraction method, fusion of Time Domain Descriptors (fTDD) was used in [12] to perform an accurate classification in same data set. They show improvement of ~4% compared to other typical feature extraction methods.

Recently deep neural networks have been proposed in EMG pattern classification systems [13]-[16]. Mukhopadhyay et al. [17] proposed a deep neural network (DNN) to recognize different EMG patterns. The presented DNN consists of an

This work is supported by Engineering and Physical Science Research Council (EPSRC), UK (grant number: EP/R004242/1).

Milad Jabbari and Kianoush Nazarpour are both with Newcastle University, UK; e-mail: jabbari_milad@yahoo.com, Kianoush.Nazarpour@newcastle.ac.uk).

Rami N. Khushaba is with the University of Sydney, Australian Centre for Field Robotics, Australia. (e-mail: Rami.Khushaba@sydney.edu.au).

input layer, several feed-forward hidden layers, and a softmax layer in the output structure to provide a probabilistic decision function. The average classification accuracy achieved by using the presented DNN was 98.88%. Atzori et al. [18] performed classification of 50 hand movements for both intact and trans-radial subjects using a Convolutional Neural Network (CNN) and achieved comparable results to the other methods like SVM, and k-NN. However, CNNs based EMG pattern recognition models focus only on the spatial correlations; disregarding the temporal information.

We propose a recurrent deep neural network structure. Specifically we use a Long-Short Term Memory (LSTM) network for EMG pattern recognition because of its capability to learn and capture long term temporal dependencies [19]–[21]. Crucially, unlike conventional recurrent neural networks, the LSTM structure does not suffer from the vanishing and exploding gradient problems [20]. Our hypothesis is that by using an LSTM-based NN, we can improve the classification.

II. Materials and Methods

A. Experimental Protocol and Data Description

The EMG datasets utilized in this study were collected from nine transradial amputees. Eight pair of Ag/AgCl electrodes (Tyco, healthcare, Germany) acquired the EMG data and signal recording procedure was performed with a custom-build, 8 channel acquisition system (Fig. 1). To facilitate the generation of the needed force level, a Virtual Instrument established in LABVIEW (National Instrument, USA) was utilized. Subjects performed six different hand gestures [10], including: 1-Thumb flexion, 2-Index flexion, 3-Fine pinch, 4-Tripod grip, 5-Hook grip, and 6-Spherical grip (power). Three force levels were generated per each gesture (low, medium, and high), while using their intact-hand to help them imagine the needed movement with the required force level. Five to eight trials were collected per force level. A total of 8 EMG sensors were utilized for data collection with a sampling frequency of 2kHz. Ethical approvals were collected for the original research [10].

B. Data Analysis

All data processing was performed with MATLAB 2018a (Mathworks, USA). The acceptable controller delay is ranging around 0-300 ms [22]. All the process should be performed in less than 300 ms, around 150 ms for data collection and the remaining time for feature extraction/reduction and classification. Therefore, to segment the EMG signals, a window of 150 ms with 50 ms overlap was used. Four different feature sets were used to train the NN. These features were:

1) Fusion of Time Domain Descriptors (fTDD).
A fusion operator is applied to the Time Domain Descriptors (TDD) features such as feature samples extracted from the current window is multiplied by the features extracted from the n^{th} previous window [12].
2) A combination of time domain and Autoregressive model parameters, with an AR model order of 5 (AR+RMS).
3) Hudgins time-domain (TD) descriptors [23].
4) The energy of the wavelet coefficient at each node of Symmlet-8 family tree with file levels decompositions.

Figure 1. An example of the surface electrode locations for subject CG1 [10].

To test generalization of the presented NN, three different training approaches were considered:

- Training_Approach_1: Training the NN with feature extracted from one specific force level and testing it with the same force level.
- Training_Approach_2: Training the NN with feature extracted from one specific force level and testing it with two remained force levels.
- Training_Approach_3: Training the NN with feature extracted from all of the force levels and testing it with a single force level.

In our evaluation, the NN was trained subject-specifically. The first three trials have been used for training and remained trials have been utilized for testing.

C. Classification Method

The proposed NN comprises a multilayer LSTM with a concatenated softmax layer. LSTM learns the long term dependency [19]–[21] and its output is classified with the softmax [24]. A schematic representation of an LSTM layer is illustrated in Fig.2. This structure of LSTM is known as vanilla version and has three gates, a block input, a block output, and a memory cell. Every gate has a specific role which realizes the long-term dependency learning content. The amount of new information allowed into the cell is specified by the input gate. The forget gate is responsible to determine when to forget the content related to the internal state, and the output gate is responsible to control which information are allowed to go to the output. The mathematical description underlying the LSTM is formulated as:

$$\mathbf{i}_t = \sigma(\mathbf{x}_t\mathbf{W}_{xi} + \mathbf{y}_{t-1}\mathbf{W}_{yi} + \mathbf{b}_i) \quad (1)$$

$$\mathbf{f}_t = \sigma(\mathbf{x}_t\mathbf{W}_{xf} + \mathbf{y}_{t-1}\mathbf{W}_{yf} + \mathbf{b}_f) \quad (2)$$

$$\mathbf{z}_t = h(\mathbf{x}_t\mathbf{W}_{xz} + \mathbf{y}_{t-1}\mathbf{W}_{yz} + \mathbf{b}_z) \quad (3)$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \mathbf{z}_t \quad (4)$$

$$\mathbf{o}_t = \sigma(\mathbf{x}_t\mathbf{W}_{xo} + \mathbf{y}_{t-1}\mathbf{W}_{yo} + \mathbf{b}_o) \quad (5)$$

$$\mathbf{y}_t = \mathbf{o}_t \odot h(\mathbf{c}_t) \quad (6)$$

where $\mathbf{x_t}$ is the input vector that is entered to the LSTM block, $\mathbf{W}_{xi}$, $\mathbf{W}_{xz}$, $\mathbf{W}_{xf}$, and $\mathbf{W}_{xo}$ are the weighted connection matrices from input to the input gate, block input, forget gate, and output gate, respectively. The recurrent weighted connections from the output to the input gate, block input, forget gate, and output gate are illustrated with $\mathbf{W}_{yi}$, $\mathbf{W}_{yz}$, $\mathbf{W}_{yf}$, and $\mathbf{W}_{yo}$ respectively. The bias weights for the input gate, forget gate, block input, and output gate are represented by the $\mathbf{b}_i$, $\mathbf{b}_f$, $\mathbf{b}_z$, and $\mathbf{b}_o$ vectors respectively. The sigmoid function, σ, is chosen as the activation function for the input gate, forget gate, and output gate, respectively, and the hyperbolic tangent, *h,* is selected for the block input. The number of LSTM blocks is set to 2 and the number of units in each LSTM block is selected as 15 heuristically to get the best accuracy. By using a segmentation scheme, four different feature sets (fTDD, AR+RMS, TD, Wavelet) were extracted from the raw EMG signal and given to the multilayer LSTM block. The output of the LSTM block is connected to the softmax which has 6 units. The output of the every unit in the softmax layer is the probability of each class.

The softmax layer estimates the probability of inputs corresponding to each class based on a probabilistic function. The number of units in the softmax layer is equal to the number of classes (6). The output of every unit in the softmax layer is the probability of each class. The structure of proposed softmax layer is and overall structure of proposed LSTM- based NN is illustrated in Fig. 3. Four feature sets are extracted from the raw EMG signals and given to the multilayer LSTM block independently for evaluation. The output of the LSTM block is connected to the softmax layer. Training the NN was performed with stochastic gradient descent (SGD) algorithm with learning rate of 0.001 and batch size of 500 samples. The selection of the learning rate value and batch size is performed heuristically to achieve the best classification accuracy.

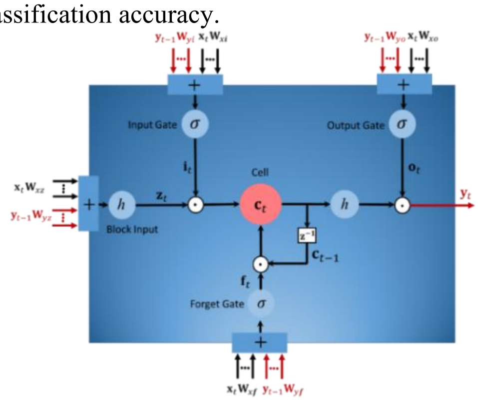

Figure 2. Detailed structure of a Long Short-Term Memory (LSTM).

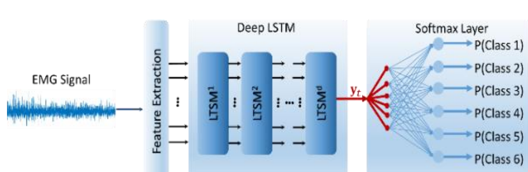

Figure 3. Overall architecture of the proposed classification algorithm.

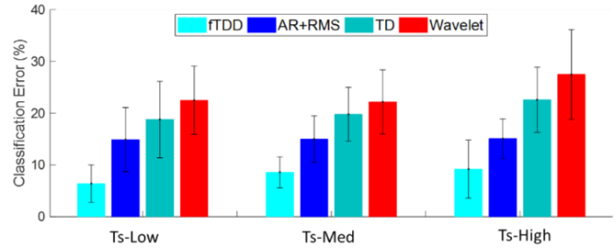

Figure 4. Average classification errors among all amputees for Training_Approach_1. *Ts* is represented for testing with a specific force level.

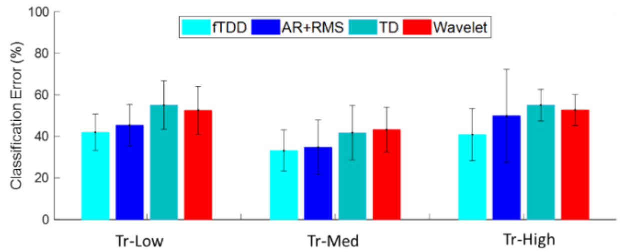

Figure 5. Average classification errors among all amputees for Training_Approach_2 with four feature sets (fTDD, AR+RMS, TD, and Wavelet). Tr is represented for training with a specific force level.

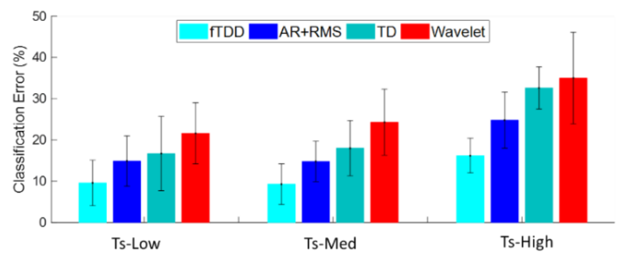

Figure 6. Average classification errors among all amputees for Training_Approach_3 with four feature sets (fTDD, AR+RMS, TD, and Wavelet). *Ts* is represented for testing with a specific force level.

III. Results

To evaluate our introduced LSTM-based NN, the classification accuracy was calculated for three predefined training approaches and average accuracies on the all subjects were reported. Fig. 4 illustrates the average classification errors for nine amputees when training and testing the NN with same force level. Results show that fTDD feature set have the best performance for all the force levels among all the feature sets with an average of 6.4±3.3%, 8.6±3.0%, 9.2±5.6% for low, medium and high force level testing, respectively. Moreover, it can be seen that training and testing with low force level leads to better classification accuracy compared to the med and high force level. The average classification errors for Training_Approach_2, is shown in Fig. 5. Similar to the Training_Approach_1 results, fTDD feature set have achieved the most accuracy performance compared to the AR+RMS, TD, and Wavelet feature sets. The results achieved from the Training_Approach_2 shows that training with medium force level and testing with two remained force levels (low, and high) has better performance compared to training with low or high level force. It is worth mentioning that, in this training approach because of training the NN was done with a specific force level and testing it with unseen force levels, highest classification errors have been achieved compared to the Training_Approach_1 and Training_Approach_3 and it is acceptable. Moreover, it is obvious that training the NN with med force level and testing it with low and high force levels, compared to training the NN with low and high force levels achieves less average classification errors. Figure 6 shows average classification errors for nine amputees when training and testing the NN with all the force levels and testing it with a single force level. Similar to the results achieved in Training_Approach_1, and Training_Approach_2, fTDD has the best performance for all the force levels. Moreover, the training and testing with low force level has achieved better accuracy compared to the medium and high levels.

IV. Discussions and Conclusion

We presented an LSTM-based NN for EMG pattern recognition. Our results show that proposed NN can achieve highly accurate classification results, which is in-line with the errors reported in usable systems (classification error < 10%). However, LSTM does not much higher accuracy than LDA [10], [16]. For Training_Approch_3, the average errors of LDA with fTDD were 5.4$\pm$3.4%, 5.3$\pm$3.2%, and 13.3$\pm$7.0% for low, medium and high force level testing, respectively, whereas for LSTM-based NN were 9.6$\pm$5.5%, 9.3$\pm$4.9%, and 16.2$\pm$4.2% for low, medium and high force level testing, respectively. If a modification is considered in LSTM block or changing the softmax with a discriminant analysis algorithm, e.g. LDA, we anticipate an improvement in accuracy. Although classification accuracy achieved in Training_Approach_1 and Training_Approach_2 promises feasibility of using proposed LSTM-based NN, the performance of the method in Training_Approach_2 indicates some serious limitations and challenges in situations where the method is trained with a specific force level and tested with an unseen level. This issue can occur in clinical settings. Therefore, to improve the performance this challenge should be addressed in the future.

References

[1] I. Kyranou, A. Krasoulis, M. S. Erden, K. Nazarpour, and S. Vijayakumar, "Real-time classification of multi-modal sensory data for prosthetic hand control," in *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2016, pp. 536–541, doi: 10.1109/BIOROB.2016.7523681.

[2] T. Pistohl, C. Cipriani, A. Jackson, and K. Nazarpour, "Abstract and proportional myoelectric control for multi-fingered hand prostheses," *Ann. Biomed. Eng.*, vol. 41, no. 12, pp. 2687–2698, Dec. 2013, doi: 10.1007/s10439-013-0876-5.

[3] J. He, D. Zhang, N. Jiang, X. Sheng, D. Farina, and X. Zhu, "User adaptation in long-term, open-loop myoelectric training: implications for EMG pattern recognition in prosthesis control," *J. Neural Eng.*, vol. 12, no. 4, p. 46005, Jun. 2015, doi: 10.1088/1741-2560/12/4/046005.

[4] C. Martelloni, J. Carpaneto, and S. Micera, "Classification of upper arm EMG signals during object-specific grasp," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008, pp. 5061–5064, doi: 10.1109/IEMBS.2008.4650351.

[5] C. Castellini and P. van der Smagt, "Surface EMG in advanced hand prosthetics," *Biol. Cybern.*, vol. 100, no. 1, pp. 35–47, Jan. 2009, doi: 10.1007/s00422-008-0278-1.

[6] F. V. G. Tenore, A. Ramos, A. Fahmy, S. Acharya, R. Etienne-Cummings, and N. V. Thakor, "Decoding of individuated finger movements using surface electromyography," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp. 1427–1434, May 2009, doi: 10.1109/TBME.2008.2005485.

[7] E. A. Biddiss and T. T. Chau, "Upper limb prosthesis use and abandonment: A survey of the last 25 years," *Prosthet. Orthot. Int.*, vol. 31, no. 3, pp. 236–257, Sep. 2007, doi: 10.1080/03093640600994581.

[8] A. Krasoulis, S. Vijayakumar, and K. Nazarpour, "Multi-grip classification-based prosthesis control with two EMG-IMU sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 508–518, 2020, doi: 10.1109/TNSRE.2019.2959243.

[9] R. N. Khushaba, A. Al-Timemy, S. Kodagoda, and K. Nazarpour, "Combined influence of forearm orientation and muscular contraction on EMG pattern recognition," *Expert Syst. Appl.*, vol. 61, pp. 154–161, Nov. 2016, doi: 10.1016/j.eswa.2016.05.031.

[10] A. H. Al-Timemy, R. N. Khushaba, G. Bugmann, and J. Escudero, "Improving the performance against force variation of EMG controlled multifunctional upper-limb prostheses for transradial amputees," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 6, pp. 650–661, Jun. 2016, doi: 10.1109/TNSRE.2015.2445634.

[11] A. Al-Ani, I. Koprinska, G. R. Naik, and R. N. Khushaba, "A dynamic channel selection algorithm for the classification of EEG and EMG Data," *International Joint Conference on Neural Networks*, pp. 4076–4081, Jul. 2016, doi: 10.1109/IJCNN.2016.7727730.

[12] R. N. Khushaba, A. Al-Ani, A. Al-Timemy, and A. Al-Jumaily, "A fusion of time-domain descriptors for improved myoelectric hand control," in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2016, pp. 1–6, doi: 10.1109/SSCI.2016.7850064.

[13] P. Angkoon, and E. Scheme, "EMG pattern recognition in the era of big data and deep learning," in *Big Data and Cognitive Computing*, 2018, p 21, doi: 10.3390/bdcc2030021.

[14] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen and B. Scodras, "Deep learning in EMG-based gesture recognition," *PhyCS*, 2018, pp. 107–114.

[15] P. Xia, J. Hu, and Y. Peng, "EMG-based estimation of limb movement using deep learning with recurrent convolutional neural networks," *Artificial Oegans,* vol. 42, no. 5, pp. 67–77, 2018, doi: https://doi.org/10.1111/aor.13004 .

[16] Zia ur Rehman, *et al.*, "Multiday EMG-based classification of hand motions with deep learning techniques," *Artificial Oegans,* vol. 18, no. 8, p. 2497, 2018, doi: 10.3390/s18082497.

[17] A. K. Mukhopadhyay and S. Samui, "An experimental study on upper limb position invariant EMG signal classification based on deep neural network," *Biomed. Signal Process. Control*, vol. 55, p. 101669, Jan. 2020, doi: 10.1016/j.bspc.2019.101669.

[18] M. Atzori, M. Cognolato, and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: a resource for the classification of movements for prosthetic hands," *Front. Neurorobotics*, vol. 10, pp. 9–9, Sep. 2016, doi: 10.3389/fnbot.2016.00009.

[19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[20] A. Murad and J.-Y. Pyun, "Deep recurrent neural networks for human activity recognition," *Sensors*, vol. 17, no. 11, p. 2556, Nov. 2017, doi: 10.3390/s17112556.

[21] A. Samadani, "Gated recurrent neural networks for EMG-based hand gesture classification. A comparative study," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 1–4, doi: 10.1109/EMBC.2018.8512531.

[22] T. R. Farrel, and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Trans. Neural syst. Rehabil Eng.*, vol. 15, no. 1, pp. 111–118, 2007, doi: 10.1109/TNSRE.2007.891391.

[23] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 1, pp. 82–94, Jan. 1993, doi: 10.1109/10.204774.

[24] H. Dong, A. Supratak, W. Pan, C. Wu, P. M. Matthews, and Y. Guo, "Mixed neural network approach for temporal sleep stage classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 324–333, Feb. 2018, doi: 10.1109/TNSRE.2017.2733220.