3-STAGE NEURAL NETWORK TRAINING PROTOCOL FOR GENERALISABLE MYOELECTRIC CONTROL

Chenfei Ma, Xinyu Jiang, Kianoush Nazarpour

School of Informatics, The University of Edinburgh, United Kingdom

ABSTRACT

Myoelectric control methods have undergone rapid evolution since the pre-1960s era. However, a longstanding challenge has been the variability of myoelectric signals across individuals, which impedes the development of universally applicable myoelectric control models. Researchers and companies in the field have been active in exploring various aspects such as different control strategies, pattern recognition methods, signal processing, and decoding. For instance, Meta recently reported a common model for a database of 6700 able-bodied participants. Development of such datasets with people with limb difference, in the higher education sector is unrealistic. But what we believe could be helpful is a scheme to guide researchers in addressing different stages of the process, with the aim of collectively developing a general-purpose, pre-trained, and generalisable myoelectric model. In this paper, we propose a 3-stage neural network training paradigm. Experiments were conducted with able-bodied participants to demonstrate the significance and necessity of each stage in the proposed scheme. Work is in progress to further enhance and verify the method. We aim to share this approach at MEC to receive feedback and invite collaborations for standardising data collection and pulling together our resources.

INTRODUCTION

The increasing popularity of technology-enabled human-machine interfacing research and commercialisation has been significantly strengthened by the emergence of a wide array of wearables [1]. This surge is primarily driven by the demand for devices that prioritise intuitiveness, efficiency, portability, and wearability, thus placing myoelectric signals at the forefront of attention. Notably, the advancement of methods for myoelectric control has accelerated rapidly, particularly with the advent of deep learning techniques and the remarkable growth in computational power. Various deep learning architectures such as convolutional neural networks [2], recurrent neural networks [3], and deep belief networks [4] have been applied extensively in both discrete movement classification and continuous regression tasks with varying success.

While many studies have demonstrated outstanding performance in movement estimation, their direct application in practical settings remains challenging. This difficulty arises primarily from end-to-end training methodologies, which often results in overfitting. Additionally, factors such as privacy concerns, substantial individual differences among users, and the inherent complexity of human movement further contribute to these challenges. Furthermore, the opaque nature of neural networks poses additional hurdles, as it complicates efforts to calibrate or adapt models when accommodating new users who may not share similar data distributions with the original training set. These limitations curtail the applicability and scalability of neural networks in real-world contexts, highlighting the need for further refinement and innovative approaches in addressing these constraints.

This paper introduces a novel 3-stage neural network training scheme, with blocked referred to as Pretraining, Localisation, and Self-calibration. Each stage employs the simplest method to provide a clear and comprehensive explanation while demonstrating the viability of the proposed protocol. Through a series of multi-stage experiments conducted over a 2-day period with 28 participants, the need for each stage is demonstrated, thus validating the efficacy of the proposed approach.

We intend to present this approach at MEC with the objective of soliciting feedback and fostering collaborations aimed at standardising data collection practices and pooling our resources.

METHOD

Pretraining

As extensively utilised in deep learning methodologies, pretraining [5] serves as an efficient means to extract and organise prior knowledge from existing data. This approach facilitates the development of robust models and therefore is an integral component of the increasingly prevalent transfer learning paradigms. As previously mentioned, the scarcity of training data in myoelectric control scenarios underscores the critical importance of carefully selecting neural network structures for the pre-training stage of the proposed paradigm.

The Temporal Convolutional Network (TCN) structure [6] has proven to be effective in processing time-sequence data, as demonstrated in various applications such as action segmentation [6] and network traffic prediction [7]. In this paper, we employ a single shallow TCN structured as depicted in Figure 1. The utilisation of dilated and causal convolutions within the TCN significantly expands the receptive field of the network. This modification directs the network's focus towards information preceding the current time step, contrasting with the typical convolutional neural network approach, which tends to distribute attention across the entire input. These characteristics align well with the requirements of processing and classifying myoelectric signals. For further details, please refer to [6].

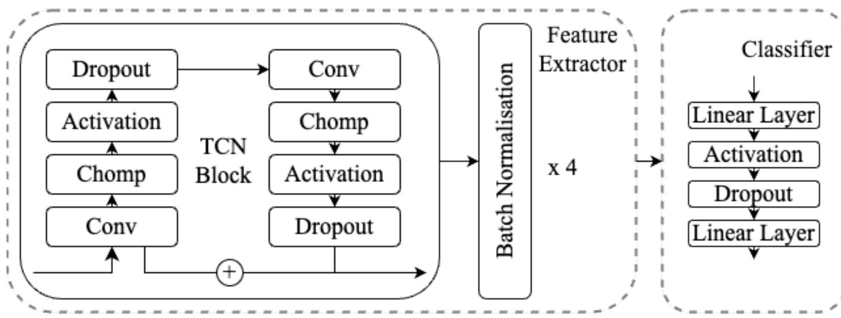


Figure 1: The TCN structure employed in the pretraining part, which is unified across all stages

Localisation

While the pretrained model demonstrates robust convergence on the pretraining dataset, the significant individual diversity poses challenges, occasionally resulting in complete failures when the myoelectric signal distribution from a new user diverges from the pretraining dataset. However, the limited size of the data collected from the new subject prohibits the establishment of a fair distribution or comprehensive representation within the model. Consequently, achieving proper calibration at this stage is impractical. Instead, a localisation approach will be implemented, which involves using a minimal amount of data to adapt the pretrained model to the new user. While one trial per movement of data may not provide sufficient information for precise adjustment of the pretrained model to the new user, experimental results demonstrate its efficacy in reducing total failures.

To localise the pretrained model, we employ fine-tuning [8]. We relax the weights for each layer of the TCN network and utilise the Adam optimizer [9] to decrease the gradient, which is calculated based on cross-entropy loss [10]. This approach is facilitated by the availability of one trial of data and its corresponding label, both of which are recorded during data collection from the new subject.

Self-calibration

Following the localisation process, the neural network begins to adapt to the distribution of the new subject's data. However, the second challenge mentioned earlier persists: the continuously evolving patterns over time within a user. This leads us to the third stage: self-calibration. Unlike the previous two stages, self-calibration utilises unlabelled myoelectric data. Its objective is to ensure the model remains adaptable to the ongoing changes in the distribution of myoelectric signals from the user. To achieve this, we propose a clustering-based semi-supervised learning approach, illustrated in Figure 2.

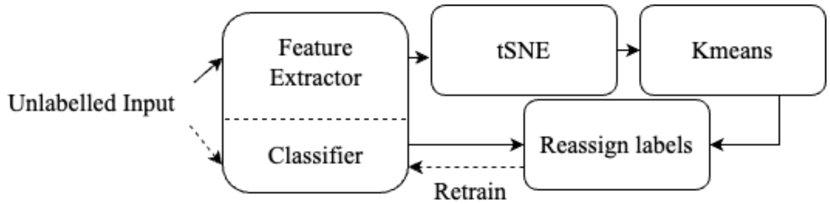


Figure 2: Schematic diagram for the proposed clustering-based pseudo labelling method

The TCN model, as described above, consists of two components: the feature extractor (FE) and the classifier. The feature extraction stage extracts high-level features from the input and passes them to the classifier. During the self-calibration process, the FE remains frozen, and its output is subjected to t-distributed Stochastic Neighbour Embedding (t-SNE) [11] for unsupervised non-linear dimensionality reduction. Subsequently, the low-dimensional embedding undergoes clustering via K-means. Although initial labels are generated, they may not align perfectly with the pre-set values. Therefore, the output of the classifier, representing predictions from the localised model, is utilised to reassign labels through comparison. Specifically, the majority label from the classifier for each sample cluster in the K-means output is reassigned to match the classifier's prediction. This process ensures uniform labelling, with the final step involving the combination of the new labels and input data to retrain the localised model. It is important to note that this self-calibration process occurs each time a certain number of samples is obtained for each label, thereby ensuring the continuous adaptability of the model to users.

Experiment design

All participants signed an informed consent form approved by the local ethics committee at the University of Edinburgh (reference number: 2019/89177), in accordance with the Declaration of Helsinki. The experiment comprised six movements: power, lateral, tripod, pointer, open, and rest. A total of 28 participants aged between 21 and 43 years, including 13 males and 15 females, were recruited. Upon informed consent, first, each participant performed one trial per movement, during which 15-channel Delsys electrodes were placed around the forearm near the elbow to collect data. Following this data collection phase, participants completed 10 blocks of tests consisting of five randomly ordered trials for each movement. During each trial, participants were instructed to mimic a gesture displayed on a computer screen for 2 seconds, with data and labels recorded during the latter 1-second interval to account for reaction time. No feedback was provided to participants to prevent bias in user behaviour. The test of the proposed scheme is illustrated in Figure 3, with each part involving the model from the preceding section. All tests were conducted using the last two trials of each block to ensure fair comparison. Furthermore, all three examinations were conducted as leave-one-out tests, wherein the same process was repeated 28 times, with each subject serving as the test subject while pretraining was conducted on the remaining subjects.

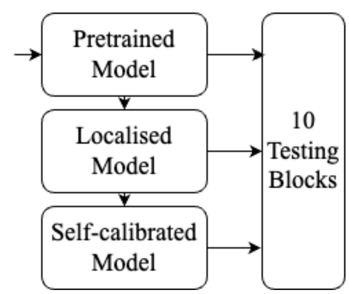


Figure 3: 3 Testing protocol

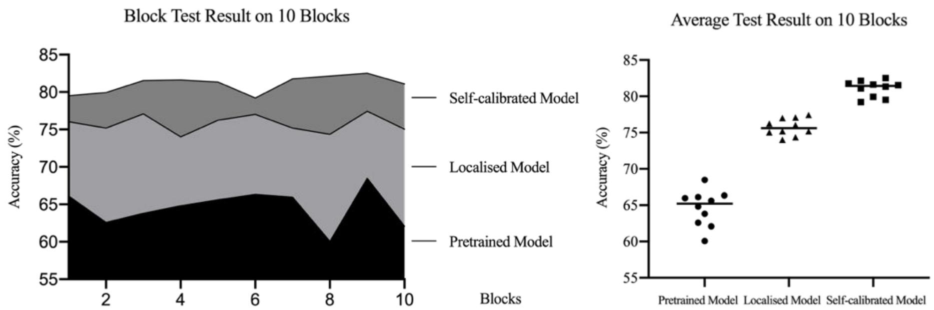
Feature extraction and processing

We extracted time-domain (waveform length, log variance, zero crossing, slope sign changes, and skewness) and frequency-domain (mean frequency, peak frequency, and variance of central frequency) features. The feature set was validated during the pretraining phase as the most reliable, yielding optimal performance. Additionally, data augmentation was performed by calculating averages between neighbouring channels to create virtual channels in between. Empirical analysis showed that this augmentation improved model performance by two percentage points.

RESULTS

The evaluation of model performance involved comparing the accuracy of the model outputs to the ground truth labels. As depicted in Figure 4(a), across all 10 blocks, the localised model consistently outperformed the pretrained model in terms of accuracy, and the self-calibrated model exhibited superior performance compared to the localised model. Although both the pretrained and localised models experienced a decline in performance during the second block, the localised model ultimately demonstrated better performance than the pretrained model, as illustrated in

Figure 4(b). Notably, the performance of the pretrained model became significantly more unstable during the last three blocks, possibly due to users forgetting the correct gesture movements. In contrast, the self-calibrated model's performance remained stable throughout all 10 blocks, beginning with a satisfactory accuracy of 79% and consistently maintaining an accuracy above 80% for most of the duration.



(a) (b)

Figure 4: (a) The test accuracy results on each block (b) The average test accuracy results on 10 blocks

DISCUSSION

The proposed training method offers a versatile approach that can be applied to any modern neural network architecture. Through three stages [pretraining, localisation, and self-calibration], it extracts information from existing datasets, personalises the network for new users, and continuously updates the model to accommodate changing myoelectric behavior. The methods chosen for each stage are intentionally simple and straightforward to facilitate clear communication of ideas and concepts. It is undeniable that machine learning based methods will increasingly dominate conventional approaches in myoelectric control. However, the way we train the neural networks in this field is equally important. The protocol outlined in this paper represents a starting point for segmenting the myoelectric signal processing pipeline and moving away from end-to-end training. In future research, we plan to explore transfer learning, aiming to apply it with a modest amount of data to fully leverage information from existing databases.

REFERENCES

[1] J. Cannan and H. Hu, "Human-Machine Interaction (HMI): A Survey," *Technical Report: CES-508*, 2011.

[2] A. Ameri, M. A. Akhaee, E. Scheme, and K. Englehart, "Real-time, simultaneous myoelectric control using a convolutional neural network," *PLoS One*, vol. 13, no. 9, 2018, doi: 10.1371/journal.pone.0203835.

[3] C. Ma *et al.*, "A Bi-Directional LSTM Network for Estimating Continuous Upper Limb Movement from Surface Electromyography," *IEEE Robot Autom Lett*, vol. 6, no. 4, 2021, doi: 10.1109/LRA.2021.3097272.

[4] J. Zhang, C. Ling, and S. Li, "EMG Signals based Human Action Recognition via Deep Belief Networks," in *IFAC-PapersOnLine*, 2019. doi: 10.1016/j.ifacol.2019.12.108.

[5] D. Hendrycks, K. Lee, and M. Mazeika, "Using pre-training can improve model robustness and uncertainty," in *36th International Conference on Machine Learning, ICML 2019*, 2019.

[6] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017. doi: 10.1109/CVPR.2017.113.

[7] J. Bi, X. Zhang, H. Yuan, J. Zhang, and M. C. Zhou, "A Hybrid Prediction Method for Realistic Network Traffic With Temporal Convolutional Network and LSTM," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, 2022, doi: 10.1109/TASE.2021.3077537.

[8] K. He, R. Girshick, and P. Dollar, "Rethinking imageNet pre-training," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019. doi: 10.1109/ICCV.2019.00502.

[9] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[10] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, no. 4, 1948, doi: 10.1002/j.1538-7305.1948.tb00917.x.

[11] G. Hinton and S. Roweis, "Stochastic Neighbor Embedding," in *NIPS 2002: Proceedings of the 15th International Conference on Neural Information Processing Systems*, 2002. doi: 10.1007/978-3-031-10602-6_16.