Deep Augmentation for Electrode Shift Compensation in Transient-phase High-density Electromyographical Neural Interfaces: Towards Application in Neurorobotics

Tianyun Sun, Jacqueline Libby, *Member, IEEE*, JohnRoss Rizzo, S. Farokh Atashzar*, *Senior Member, IEEE*

***Abstract*—Going beyond the traditional sparse multi-channel peripheral human-machine interface that has been used widely in neurorobotics, high-density surface electromyography (HD-sEMG) has shown significant potential for decoding upper-limb motor control. We have recently proposed heterogeneous temporal dilation of LSTM-based deep neural networks for a very large number of gestures (>60), securing spatial resolution and fast convergence. However, several fundamental questions remain unanswered. One problem targeted explicitly in this paper is the issue of "electrode shift", which can happen specifically for high-density systems and during doffing and donning the sensor grid. Another real-world problem is the question of transient versus plateau classification, which connects to the temporal resolution of neural interfaces and seamless control. In this paper, for the first time, we implement gesture prediction on the transient phase of HD-sEMG data while robustifying the human-machine interface decoder to electrode shift. For this, we propose the concept of deep data augmentation for transient HD-sEMG. We show that without using the proposed augmentation, a slight shift of 10mm may drop the decoder's performance to as low as 20%. Combining the proposed data augmentation with a 3D Convolutional Neural Network (CNN), we recovered the performance to 84.6% while securing a high spatiotemporal resolution, robustifying to the electrode shift, and getting closer to large-scale adoption by the end-users, enhancing resiliency.**

***Keywords*—Human-centered Robotics, Neurorobotics, High-Density sEMG, Prosthetics**

I. Introduction

In the last decade, thanks to significant technical and technological developments, neural interfacing and neurorobotics have gone through major transformations, unleashing the potential for interfacing human cognition to machine intelligence. Focusing on peripheral interfacing in this paper, two major directions of research and development are (a) utilization of dense and flexible electronics and (b) implementation of deep biosignal processing, which can demystify the neural code of motor intention to be used for the control of neurorobots. Such advancements have been motivated by the concerning statistics regarding the growing population of people who (a) suffer from the loss of a biological limb and would benefit from intuitive control of prostheses [1], and (b) live with age-related neuromuscular disorders such as stroke, and can benefit from robotic orthoses that respond agilely to the motor intention of the user, such as exoskeletons and soft exosuits [2]. Motivated by the above-mentioned note, this paper focuses on solving one critical issue facing the translation of such technologies into large-scale practice, namely, electrode-shift-related degradation of information space collected by densely-located electronics.

Regarding the use of machine learning (ML) for peripheral human-machine interfacing, a large volume of literature has been conducted focusing on the processing of surface electromyogram (sEMG) signals for the control of upper limb prostheses and orthoses and through decoding the user's intended motor task. This field has gone through a rapid progression after the introduction of deep learning models (for example, [3]–[7]). Despite recent algorithmic progress related to deep learning (DL) models as well as mechatronic progress such as soft bionic hands [8] and soft exosuits [9], these technologies mainly work predictably in laboratory settings. One of the obstacles to real-world implementation is the need for extensive recalibration and the corresponding sensitivity of the trained model to various experimental conditions. Among the existing challenges is the topic of electrode shift, which is significantly more pronounced when signals are collected using miniaturized and densely-located electrodes. Electrode shift can be the result of doffing and donning the device, or it may happen due to skin stretches during intensive tasks. It is known that electrode shift would lead to degradation in model accuracy [10]–[13]. In order to address the issue of electrode displacement and misplacement, recalibration has been utilized in practice. However, it is a time-consuming process,

Tianyun Sun is with Facebook, USA and was with the Department of Electrical and Computer Engineering, New York University, New York, NY 11201, USA. Jacqueline Libby is with the NYU Center for Urban Science and Progress (CUSP), New York, NY USA. JohnRoss Rizzo with the NYU School of Medicine, New York NY, 10016. S. Farokh Atashzar is with the Department of Electrical and Computer Engineering, and with the Department of Mechanical and Aerospace Engineering, and with NYU WIRELESS, and also with the NYU Center for Urban Science and Progress (CUSP), New York University, New York, NY 11201 USA (e-mail: f.atashzar@nyu.edu). This material is based upon work supported by the US National Science Foundation under Grant No 2037878.
* Corresponding author: f.atashzar@nyu.edu.

complicated for non-expert users to conduct accurately, and cannot address several issues such as skin stretch. This has made the performance of such technologies limited and often increased the rejection rates of prosthetic systems. Thus, it can be summarized that there is an unmet need for the development of computational models and frameworks that minimize the need for recalibration and can generalize the performance over "possible electrode shift" configurations.

The challenge of the electrode shift has been discussed in the literature for over a decade, initially on multi-channel electrodes sparsely placed directly on the muscles of interest. For example, the electrode shift issue has been discussed in the context of gesture and motion prediction using traditional ML models, where manual feature engineering was conducted to enhance the performance. In this regard Young *et. al.* [14] proposed optimization of the electrode configuration, achieving 78% accuracy on seven classes. In addition, there exists some literature that suggests fixing the problem by collecting a larger, more representative dataset of multiple limb positions. Despite the concerns for generalizability and the limited number of gestures that were detected using classic algorithms, there are also concerns about the long data collection processes, which can be arduous for the end-user [15]–[18]. To avoid this, more recently, Ameri *et. al.* [19] proposed the use of transfer learning for sparse multi-channel recording to minimize the amount of new data that is needed each time the user doffs/dons the device, allowing for a faster recalibration process. While the proposed transfer learning is an effective later-stage processing step, the compounded performance would be limited by the functionality of the initial phase of processing. It should also be noted that in most of these examples, since the channels are treated independently, the spatial relationship between the channels is not fully exploited. This can be another limiting factor in considering electrode shift, which is inherently a spatial problem.

In order to improve the performance of peripheral neural interfaces, high-density sEMG (HD-sEMG) hardware has shown great potential providing 2D grids of densely-spaced electrodes, as shown in Fig. 1. This has significantly boosted the information rate captured from one muscle and has been (a) processed using computational methods (such as independent component analysis) to enable decomposition of the collected dense signal space into spike train activities deriving motor units in the muscles [20], and (b) more recently as a direct feed into deep neural network models to decode the motor intention of the user with ultimate spatiotemporal resolution [3]. In [3] we have recently proposed a novel deep recurrent neural network, one of the very first deep neural networks on transient-phase high-density EMG data, that can map the dynamic phase of high-density recording, captured using 128 channels from the upper limb (64 flexors and 64 extensors), into a prediction of over 60 classes of gestures. The proposed algorithm was able to make the convergence of the network 20 times faster when compared with conventional networks while providing high accuracy and sensitivity. In addition, our previous work showed that the transient phase of sEMG (which is classically discarded for ML-based processing) indeed has enough information to be used for augmenting the temporal resolution of human-machine interfaces. We also showed that the high-density information context significantly boosted up the power of such machines so that we could get over 80% performance for 65 gestures. The problem of electrode shift is significantly more pronounced for high-density recording of surface electromyography due to the size of the electrodes and since the existing computational models depend on the corresponding contextual meaning of electrode locations. This issue is discussed in [21] which sheds light on the effect of electrode shift on the process of decoding muscle synergies. The authors of [21] showed that when electrodes are shifted, as long as they are still recording from the same muscles, although the shift may not prevent the synergy estimation process, it would require a new model, which emphasizes the sensitivity of the problem to small electrode shifts and the existing central question regarding the generalizability of such models to various configurations.

In addition to the literature regarding the effect of electrode shift on decomposition and synergy, there exists some work showing the degradation of HD-sEMG-based decoders in the presence of small shifts of the sensor grid when using ML for gesture classification. Some have used classic ML algorithms and manual feature engineering. In this regard, Stango *et. al.* [22] used spatial correlation features fed into Support Vector Machines, achieving 80% accuracy on 9 classes. Pan *et. al.* [23] used Common Spacial Pattern features with LDA classifiers, achieving 80% accuracy on 11 classes. Also, He & Zhu [24] used Fourier domain features, achieving 85% accuracy on 11 classes. Also, Lv *et. al.* [25] built off of these traditonal machine learning approaches, using memory-free autoencoder design of multilayer perceptron, achieving 90% accuracy for 10 classes. In all the above-mentioned examples, the cost to address electrode shift was either the lower number of degrees of freedom (limiting the versatility of the human-robot interface) or the need for focusing on stabilized plateau signals [22]–[26]. However, these simplifications defeat the very purpose of neural interfaces, i.e., agility and dexterity in predicting human motor intention.

In order to further highlight the critical need for decoding the transient phase of muscle activation, it can be mentioned that the existing literature only looks at the plateau phase of the sEMG signals. However, a hand gesture involves a dynamic transient phase, which represents the dynamical recruitment of motor units in the first part of the task and contains critical temporal information. However, in the literature, this part of data is often discarded due to the complexity of modeling. Thus, it can be mentioned that

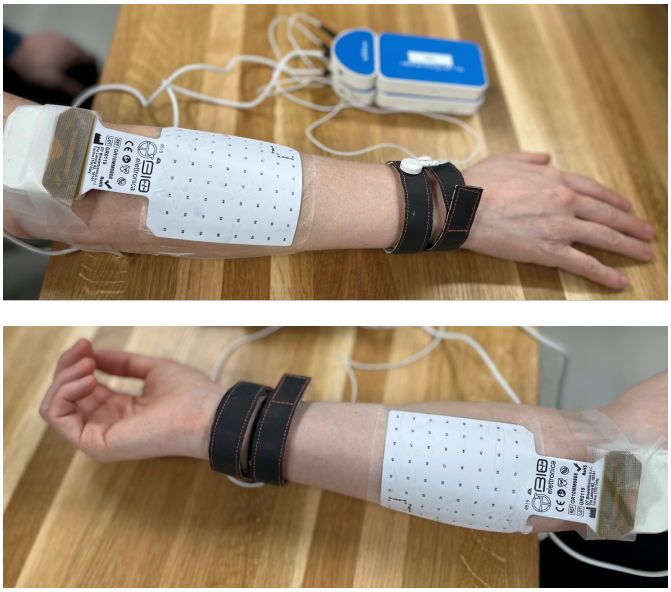

Fig. 1: Two wireless OT Bioelettronica high density sEMG 8×8 grids, attached to the extensors (top) and flexors (bottom) of the forearm for gesture prediction.

gesture prediction from the transient phase is an under-explored area of research. Addressing this would allow for detecting the motor intent at the very beginning of the muscle contraction and when the signals are not stabilized. Such an approach would enable faster response and better real-time control. In addition, it should be noted that stable plateau, stationary signals do not often exist when we design neurorobots for amputees and for people with neuromusculoskeletal disorders due to the damage to the neural and muscular structures. Thus, detecting abrupt transient phases of actions and translating that into motor intention is of high importance.

In this letter, for the first time, we address the problem of electrode shift for decoding the transient phase of high-density electromyography while securing a large number of gesture classes. We showed that the performance of the best deep learning algorithm could drop significantly (from %80 to %20) if the proposed deep data augmentation is not conducted. The specific augmentation proposed in this paper does not require extra data collection and, in conjunction with 3D CNN, has allowed us to develop a new neural interface that can robustly predict the gestures based on the transient behavior of EMG signals.

The rest of this paper is organized as follows. In Section II, we explain the material and method, including the nonlinear data normalization, transient phase labeling, and data preparation. In Section III, we explain the model structure and consider multiple derivations of the deep learning decoder. In Section IV, we provide details about our experiments and results, and we provide details about a comprehensive comparative analysis that was conducted to compare the performance of the proposed method with existing solutions.

II. Material and Methods

A. Data Acquisition

In this paper, we used a high-quality, high-density HD-sEMG database recently published in the scientific data of Nature [27]. The dataset contains 65 isometric hand gestures with different degrees of freedom (DoFs), including six 1-DoF finger and wrist gestures, forty-one 2-DoF compound gestures of fingers and wrist, and eight multi-DoF gestures of grasping, pointing, and pinching. Fig. 2 shows examples of these gestures: at rest, 1-Dof little finger bend, 1-Dof ring finger bend, and the 2-Dof combination of these two finger-bends. The HD-sEMG signals were recorded using a Quattrocento (OT Bioelettronica) biomedical amplifier system through two pads of 8×8 electrode grids (for a total of 128 channels), with a 10mm inter-electrode distance. A newer wireless version of this system is depicted in Fig. 1 to highlight the electrode locations. The signals are collected at a sampling rate of 2048 Hz. This system allows for collecting dense biosignal space with information in both space and time. The two electrode grids are placed on both sides of the forearm, namely the dorsal (outer extensor muscles) and the volar (inner flexor muscles) of the upper forearm. In order to reduce common-mode noise, the recording was performed in a *differential manner*, where the signal of channel i in the final output signal is the difference between electrode $i+1$ and i in the original collection. Twenty healthy adults, fourteen males, and six females, with age range between 25 and 57 years old, provided the data for collection. In this paper, we only used data from 19 subjects because data of subject 5 is unavailable. Each subject was asked to perform the gestures for five repetitions before switching to the next one. Each of those repetitions would last for 5 seconds, followed by an equal-duration interval of rest. In this way, the muscles get less influenced by fatigue and therefore provide more consistent data.

Some examples of muscle-activity heatmaps are shown in Fig. 3 for the best performing subject to make the readers familiar with the heatmap representation of HD-sEMG. In this figure, due to the space constraints, we only show 16 out of the 65 gestures from the extensors and flexors grids, thus, a total of 32 heatmaps. We choose the 16 1-DoF gestures since they form the basis for the higher-DoF compound movements. For example, it can be observed that for move 13 “wrist: bend”, which is wrist flexion, more muscle activity is observed on the inner forearm (which contains the flexors), as expected. Alternatively, for move 1 “little finger: bend”, which is flexion of the little finger, we see more activity on the outer forearm (the extensors), because it takes work for the other fingers to stay straight while the little finger is bending.

Forces measured at each finger and the wrist are included in the dataset to provide the labeling mechanism for cutting

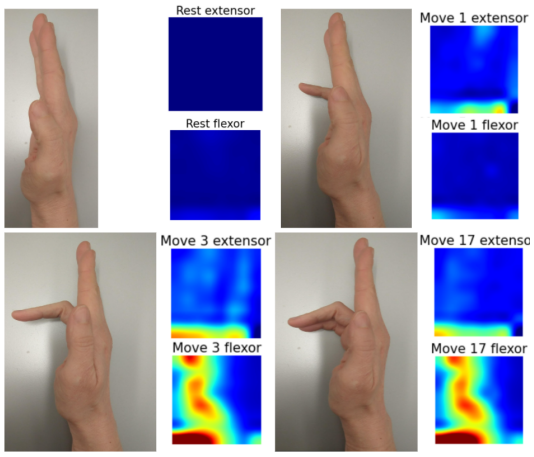

Fig. 2: Examples of four gestures. (a) At rest; (b) 1-Dof, gesture 1: little finger bend; (c) 1-Dof, gesture 3: ring finger bend; (d) 2-Dof, gesture 17: combining little finger bend and ring finger bend.

the transient and plateau phases of the data. The plateau phase is once the gesture has reached its final static position, and the transient phase is the movement of the body from rest to that final position, as can be seen in Fig. 5.

B. Data Preprocessing

As mentioned before, in this paper, we compare the performance of the proposed model on the transient phase with that of the model on the plateau phase. In this paper, we define these phases by averaging the force signal of each gesture (see Fig. 5). From this, we chose the first 0.5 seconds of data as the most dynamic portion of the sEMG time series, which encodes the transient phase. Force signals are shown in Fig. 5 for the same gestures shown in Fig. 2 (Movement 1: little finger bend, Movement 3: ring finger bend, and Movement 17: little finger bend plus ring finger bend.) The 0.5-second transient phase demarcation is shown with a dashed line. Force indices 0-8 denote measurements on the index finger, middle finger, ring finger, little finger, thumb finger flexion/extension, thumb finger abduction/adduction, wrist-flexion/extension, wrist-pronation/supination, and wrist-radial/ulnar, respectively.

In this paper, in order to normalize the signal space of HD-sEMG, we propose to use the -law decoding after applying a Min-Max normalization on the signals. -law transformation enhances the discriminability among information channels and therefore helps the model capture more information.

$$F(x_t) = sign(x_t)\frac{ln(1+\mu|x_t|)}{ln(1+\mu)}. \quad (1)$$

In (1), x_t denotes each data scalar and μ is a hyperparameter and selected to be 2048. Following the real-time implementation standards in myoelectric control, [28]–[32], we choose to conduct our experiments using a sliding window size of 200ms with a step size of 10ms. On each timestamp, one data point is generated, with the dimension of (sampling rate$\times$window size)$\times 8\times 8\times 2$, where 8×8 is the size of each of the two grids. To help understand, Fig. 4 visualizes the data collection and sliding window process.

C. Data Augmentation

In order to augment the input space without requiring to collect a very large size of data, instead of using the full set of $8\times 8\times 2$ (which is the complete neurophysiological window of activation with the size of $8\times 8 attwoelectrodes), weusesubsetsof 6\times 6\times 2$ channels (with a smaller size, named as the window of observation). This would leave out two extra electrodes which can be leveraged for data augmentation. In other words, the task will be to predict the gesture based on the window of observation, which may shift to right/left/top/down (or a dual combination) resembling electrode shift and augmenting the input space, which will force the neural network to learn the common underlying patterns of information which can be decoded in all shifted input space. In terms of the numbers, the aforementioned techniques would result in multiple choices of 6×6 observations window of observation within the 8×8 neurophysiological activation window. This data augmentation process is visually show in Fig. 6. Thus, two alternative spatial shifts can be considered, one includes only one step shifting, and one includes two-step shifting. In the experiments where we use the 1-step shifts, our dataset is multiplied by a factor of 5 (one 'standard' position of no shift + four '1-step' shifts including shift to the right, left, top and down). In the experiments where we use both the 1 and 2 step shifts, our dataset is multiplied by a factor of 9 (one 'standard' position + four '1 step' shifts + four '2 step' shifts).

III. Model Structure

Motivated by the literature on deep video processing [33]–[36],we treat the 2D signal at each timestamp as a frame of a neurophysiological video capturing distributed muscle activation. In our research, we specifically focus on the use of 3D and 2D CNN deep neural networks, and we also evaluate a hybrid solution using recurrent neural network. This is to conduct a comprehensive comparative analysis on the performance of the system taking into account complexity and performance.

A. CNN-RNN Hybrid Models

Our 2D CNN-RNN hybrid model is depicted in Fig. 7a. This is the most straightforward as it is a combination of CNN and RNN sections which individually take responsibility for decoding the spatial and temporal components of the signal, respectively. The CNN section is composed of

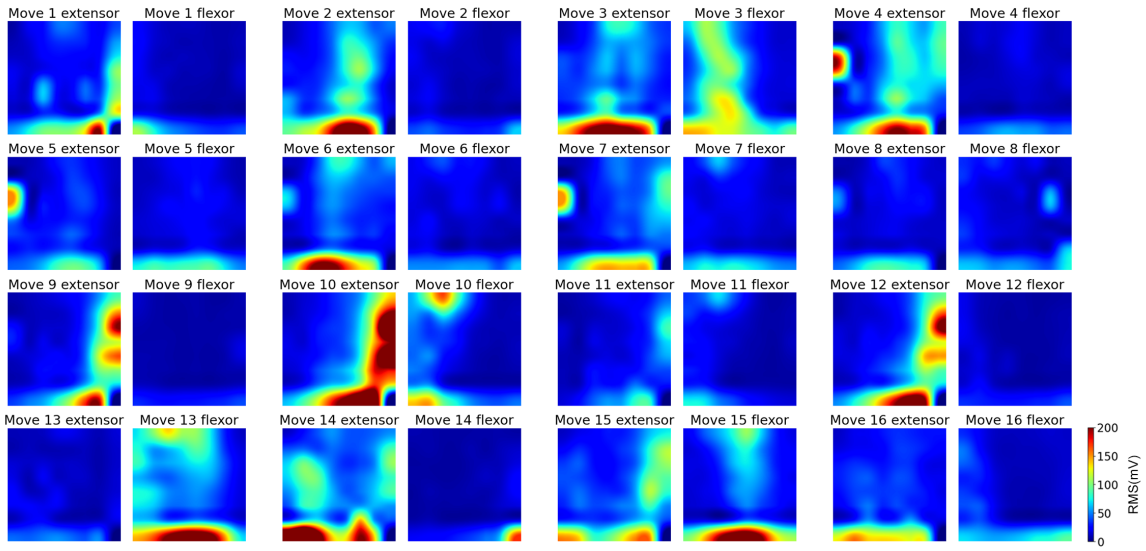

Fig. 3: 32 muscle-activity heatmaps associated with 16 1-DoF movements from the best-performing subject (#17). Each gesture has two heatmaps (forearm extensor and flexor). Each heatmap is an 8×8 grid, consisting 64 electrodes.

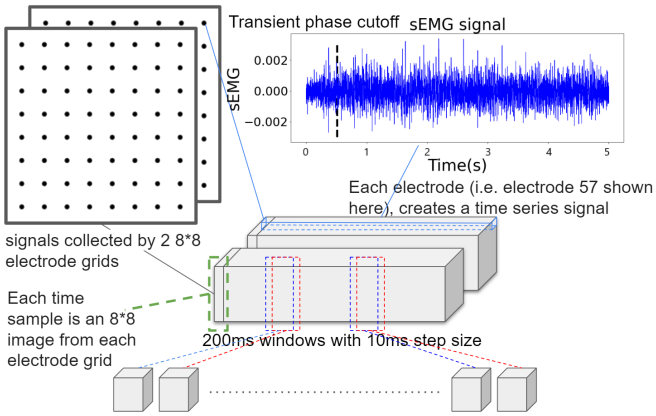

Fig. 4: Data acquisition and sliding window process.

a series of Conv2D layers with 2×2 kernel size and Relu activation function for nonlinearity, each with 8, 16, 32, 64 and 128 filters to gradually parse the images to a vector of 128 channels. It parses the (number of samples)×6×6×2-sized input signal to a shape of (number of samples)×72. Afterwards, it is fed into a 4-layer LSTM network with 128 hidden parameters and 400 LSTM nodes on each layer. Finally, the output of the LSTM layers is fed into 3 fully connected layers with 65 nodes in the last layer, corresponding to the number of classes for the classification task. We also train a 1D CNN-RNN hybrid model as a benchmark. For this, instead of taking the 6×6 signal through a 2D CNN, the signals are reshaped into a 36×1 vector and then consider 1D convolution.

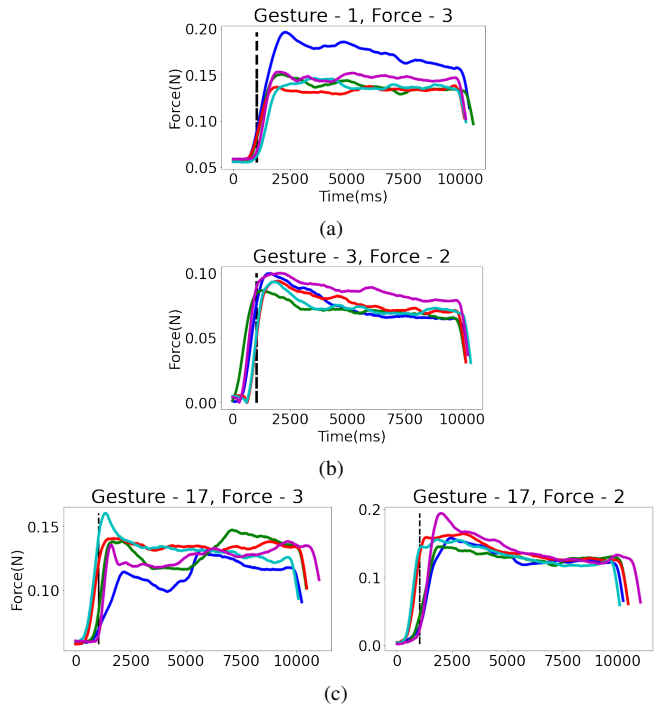

Fig. 5: The corresponding forces of the same three gestures shown in Fig. 2 on each repetition. The dashed lines indicate the end (0.5 seconds) of transient phases. Line colors denote five different repetitions. (a) Little finger force of little finger bend gesture; (b) Ring finger force of ring finger bend gesture; (c) Little finger and ring finger forces of Little finger: bend+Ring finger bend gesture.

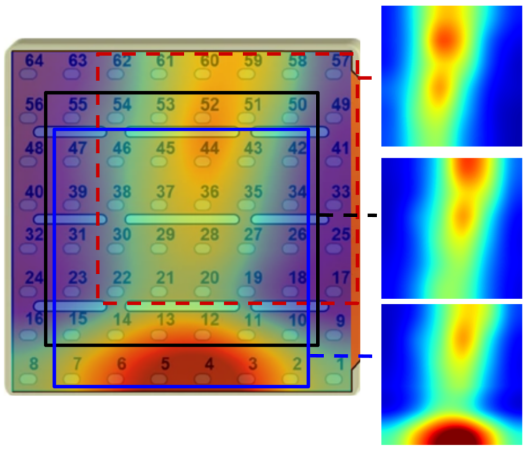

Fig. 6: The heatmap with the 8×8 electrode grids are shown. The black box shows the 'standard' position, the blue box shows a shift '1 step away', and the dashed red line shows a shift '2 steps away'. The corresponding heatmap of the boxes are also shown on the right side.

B. 3D CNN Model

In the case of 3D CNN models as shown in Fig. 7b, we directly take the whole input data through a bunch of Conv3D layers to let the models decompose spatial and temporal information at the same time which is critical to consider spatial and temporal interactions between various segments of the signal. In detail, the structure consists of 4 layers of $(100, 2, 2)$ kernel size, each with 8, 16, 32 and 64 filters, followed by 1 layer of $(4, 2, 2)$ kernel and 128 filters. Similar to the 2D case, the output was flattened and fed into three fully connected layers for classification.

IV. Experiments and Results

A. Experiment Models

We first compared the accuracy of the 3 aforementioned CNN models using data with '1 step away' augmentation. We tested using either transient phase or stable phase data. In both cases, the 3D CNN-based structure consistently gave best performance. We then trained the 3D CNN-based model using both '1 step away' and '2 steps away' augmentations. In addition, to demonstrate the need of this augmentation, we also show the results of a small experiment we call 'non-aug' where the best performing model was trained in the 'standard' position and tested in one of the '1 step away' shifts to simulate the case of misplacement. The experimental models and the corresponding model ID's are listed in table I. The model ID's act as a reference for a later box plot. The median model accuracy across all experiment subjects of every experiment is listed in table II.

B. Results and Statistical Analysis

To avoid bias and enforce generalization of the model performances, we utilized k-fold (k=5) cross validation when

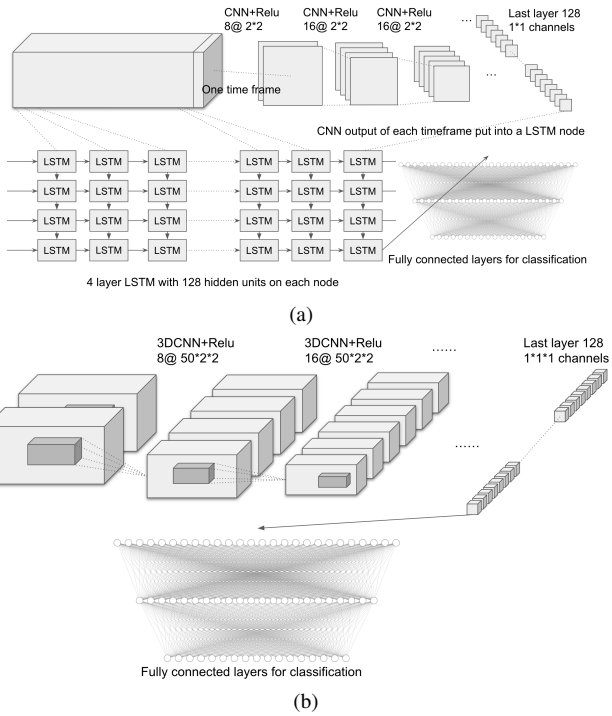

Fig. 7: (a) Hybrid model using a combination of 2D CNN and multilayer LSTM; (b) Model using 3D CNN

evaluating the model accuracy. We hold out one out of the five repetitions for testing and the rest for training purpose. In Fig. 8, we show the box plots of the classification accuracy of the aforementioned models using either the transient phase signal or the stable phase signal. In the box plot, the model ID corresponds to what is listed in table I, the black horizontal lines indicate the medians, and the green triangles indicate the means. In order to understand the significance of effect from our proposed augmentation methods, statistical analysis is performed on all the aforementioned models across all 19 subjects. We first performed Shapiro for normality test and found out that all but the distribution for the experiment result of model 1 passed normality. Because of that, we then performed Wilcoxon's test which determines whether the mean value differences between the results of each model is significant. correction was applied to the observed p-values to reduce the chance of false positives. The corrected p-value ranges are denoted with markers at the top of Fig 8. The marker symbols are defined as follows: (a) The ns marker (for not significant) represents 0.05 to 1; (b) * represents 0.01 to 0.05; (c) ** represents 0.001 to 0.01; (d) *** represents 0.0001 to 0.001; and (e) **** represents smaller than 0.0001.

adjustbox

As we expected, the model performance gradually increases when using 1D, 2D and 3D Convolution layers. Surprisingly, the models using Conv1D and Conv2D achieved similar performance under both cases using either tran-

TABLE I: Model ID

	Conv1d	Conv2d	Conv3d	Conv3d 2step
Transient	1	2	3	4
Stable	5	6	7	8

TABLE II: Median model accuracy

	Non-aug	Conv1d	Conv2d	Conv3d	Conv3d 2step
Transient	18.54%	69.01%	67.89%	78.50%	**80.22%**
Stable	20.03%	73.61%	72.63%	**84.60%**	84.45%

sient or stable phase data. In contrast, the Conv3D model achieved significant performance gain compared to the prior models. The Conv3D structure considers both the spatial and temporal components of the signal at the same time. This can be particularly important when handling electrode shift for the dynamic phase, since the spatial component is so critical for handling the electrode shift issue. We also observe that by adding data augmentations that are '2 steps way' from the 'standard position', the performance gained slight improvement.

V. Conclusion

The accuracy, agility, and robustness of myolectric gesture prediction is imperative for scaling up the practical translation of the upper limb neurorobotic systems, which can eventually improve the quality of life for those individuals that have suffered the loss of a biological limb or live with a neuromuscular disorder. With advances in device hardware technology, HD-sEMG sensor technology, and deep learning models, the tools exist to provide robustify the performance of such systems. In this paper, we have combined very rich HD-sEMG data, along with data augmentation and deep learning techniques to push this potential forward. We have used data augmentation to make our classifiers robust to electrode shift. Furthermore, we have applied our classification to the transient phase of the data, allowing for classification of gesture intention, before the plateau has been reached. Our techniques do not require any recalibration for the user, which allows for ease of use. Using 128 HD-sEMG channels with the proposed data augmentation technique, fed into a 3d CNN classifier, we are able to achieve 84.6% accuracy on the transient phase of the data, when the data that has been exposed to electrode shift. Combining the electrode shift challenge with the transient phase of the data, we have tackled two real-world problems at the same time. These results push the field one step closer to achieving intuitive and seamless human-machine interfaces for prosthetic and orthotic myolelectric control.

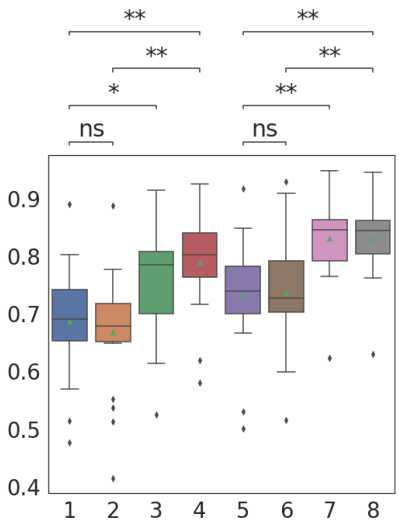

Fig. 8: Classification accuracy of the experiment models.

References

[1] F. Cordella, A. L. Ciancio, R. Sacchetti, A. Davalli, A. G. Cutti, E. Guglielmelli, and L. Zollo, "Literature review on needs of upper limb prosthesis users," *Frontiers in neuroscience*, vol. 10.

[2] M. J. Prince, F. Wu, Y. Guo, L. M. Gutierrez Robledo, M. O'Donnell, R. Sullivan, and S. Yusuf, "The burden of disease in older people and implications for health policy and practice," *The Lancet*, vol. 385, no. 9967, pp. 549–562, 2015.

[3] T. Sun, Q. Hu, J. Libby, and S. F. Atashzar, "Deep heterogeneous dilation of LSTM for transient-phase gesture prediction through high-density electromyography: Towards application in neurorobotics," *IEEE Robotics and Automation Letters*.

[4] T. Sun, Q. Hu, P. Gulati, and S. F. Atashzar, "Temporal dilation of deep LSTM for agile decoding of sEMG: Application in prediction of upper-limb motor intention in neurorobotics," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6212–6219, jun 2021.

[5] P. Gulati, Q. Hu, and S. F. Atashzar, "Toward deep generalization of peripheral EMG-based human-robot interfacing: A hybrid explainable solution for NeuroRobotic systems," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2650–2657, Apr. 2021.

[6] E. Rahimian, S. Zabihi, A. Asif, D. Farina, S. F. Atashzar, and A. Mohammadi, "Fs-hgr: Few-shot learning for hand gesture recognition via electromyography," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1004–1015, 2021.

[7] J. L. Betthauser, J. T. Krall, S. G. Bannowsky, G. Lévay, R. R. Kaliki, M. S. Fifer, and N. V. Thakor, "Stable responsive EMG sequence prediction and adaptive reinforcement with temporal convolutional networks," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 6, pp. 1707–1717, sep 2019.

[8] R. Deimel and O. Brock, "A novel type of compliant and underactuated robotic hand for dexterous grasping," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 161–185, oct 2016.

[9] L. Cappello, J. T. Meyer, K. C. Galloway, J. D. Peisner, R. Granberry, D. A. Wagner, S. Engelhardt, S. Paganoni, and C. J. Walsh, "Assisting hand function after spinal cord injury with a fabric-based soft robotic glove," *Journal of neuroengineering and rehabilitation*, vol. 15, no. 1, pp. 1–10, dec 2018.

[10] "Examining the adverse effects of limb position on pattern recognition based myoelectric control."

[11] A. J. Young, L. J. Hargrove, and T. A. Kuiken, "The effects of electrode size and orientation on the sensitivity of myoelectric pattern recognition systems to electrode shift," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 9, pp. 2537–2544, jun 2011.

[12] Y. Geng, P. Zhou, and G. Li, "Toward attenuating the impact of arm positions on electromyography pattern-recognition based motion classification in transradial amputees," *Journal of neuroengineering and rehabilitation*, vol. 9, no. 1, pp. 1–11, dec 2012.

[13] Y. Gu, D. Yang, Q. Huang, W. Yang, and H. Liu, "Robust EMG pattern recognition in the presence of confounding factors: features, classifiers and adaptive learning," *Expert Systems with Applications*, vol. 96.

[14] A. J. Young, L. J. Hargrove, and T. A. Kuiken, "Improving myoelectric pattern recognition robustness to electrode shift by changing inter-electrode distance and electrode configuration," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 3, pp. 645–652, nov 2011.

[15] L. Hargrove, K. Englehart, and B. Hudgins, "The effect of electrode displacements on pattern recognition based myoelectric control," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, aug 2006.

[16] ——, "A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control," *Biomedical signal processing and control*, vol. 3, no. 2, pp. 175–180, apr 2008.

[17] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *Journal of rehabilitation research and development*, vol. 48, no. 6, sep.

[18] A. Fougner, E. Scheme, A. D. Chan, K. Englehart, and Ø. Stavdahl, "Resolving the limb position effect in myoelectric pattern recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 6, pp. 644–651, aug 2011.

[19] A. Ameri, M. A. Akhaee, E. Scheme, and K. Englehart, "A deep transfer learning approach to reducing the effect of electrode shift in EMG pattern recognition-based control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 2, pp. 370–379, dec 2019.

[20] D. Farina, I. Vujaklija, M. Sartori, T. Kapelner, F. Negro, N. Jiang, K. Bergmeister, A. Andalib, J. Principe, and O. C. Aszmann, "Man/machine interface based on the discharge timings of spinal motor neurons after targeted muscle reinnervation," *Nature biomedical engineering*, vol. 1, no. 2, pp. 1–12, 2017.

[21] S. Muceli, N. Jiang, and D. Farina, "Extracting signals robust to electrode number and shift for online simultaneous and proportional myoelectric control by factorization algorithms," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 3, pp. 623–633, oct 2013.

[22] A. Stango, F. Negro, and D. Farina, "Spatial correlation of high density EMG signals provides features robust to electrode number and shift in pattern recognition for myocontrol," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 2, pp. 189–198, nov 2014.

[23] L. Pan, D. Zhang, N. Jiang, X. Sheng, and X. Zhu, "Improving robustness against electrode shift of high density EMG for myoelectric control through common spatial patterns," *Journal of neuroengineering and rehabilitation*, vol. 12, no. 2, pp. 1–16, dec 2015.

[24] J. He and X. Zhu, "Combining improved gray-level co-occurrence matrix with high density grid for myoelectric control robustness to electrode shift," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 9, pp. 1539–1548, dec 2016.

[25] B. Lv, X. Sheng, and X. Zhu, "Improving myoelectric pattern recognition robustness to electrode shift by autoencoder," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC*, jul 2018.

[26] L. Wu, X. Zhang, K. Wang, X. Chen, and X. Chen, "Improved high-density myoelectric pattern recognition control against electrode shift using data augmentation and dilated convolutional neural network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 12, pp. 2637–2646, oct 2020.

[27] N. Malešević, A. Olsson, P. Sager, E. Andersson, C. Cipriani, M. Controzzi, A. Björkman, and C. Antfolk, "A database of high-density surface electromyogram signals comprising 65 isometric hand gestures," *Sci Data*, vol. 8, no. 1, p. 63, Feb. 2021.

[28] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 848–854, Jul. 2003.

[29] H. F. Hassan, S. J. Abou-Loukh, and I. K. Ibraheem, "Teleoperated robotic arm movement using electromyography signal with wearable myo armband," *Journal of King Saud University - Engineering Sciences*, vol. 32, no. 6, pp. 378–387, Sep. 2020.

[30] W.-T. Shi, Z.-J. Lyu, S.-T. Tang, T.-L. Chia, and C.-Y. Yang, "A bionic hand controlled by hand gesture recognition based on surface EMG signals: A preliminary study," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 1, pp. 126–135, Jan. 2018.

[31] A. Aranceta-Garza and B. A. Conway, "Differentiating variations in thumb position from recordings of the surface electromyogram in adults performing static grips, a proof of concept study," *Front Bioeng Biotechnol*, vol. 7, p. 123, May 2019.

[32] M. F. Wahid, R. Tafreshi, and R. Langari, "A Multi-Window majority voting strategy to improve hand gesture recognition accuracies using electromyography signal," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 427–436, Feb. 2020.

[33] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*.

[34] Y. Fan, X. Lu, D. Li, and Y. Liu, "Video-based emotion recognition using CNN-RNN and C3D hybrid networks," in *Proceedings of the 18th ACM international conference on multimodal interaction*, pct 2016.

[35] C. Lea, A. Reiter, R. Vidal, and G. D. Hager, "Segmental spatiotemporal CNNs for fine-grained action segmentation," in *European Conference on Computer Vision*, oct 2016.

[36] A. Diba, M. Fayyaz, V. Sharma, A. H. Karami, M. M. Arzani, R. Yousefzadeh, and L. Van Gool, "Temporal 3d convnets: New architecture and transfer learning for video classification," *arXiv preprint arXiv:1711.08200*.