

From Unstable Electrode Contacts to Reliable Control: A Deep Learning Approach for HD-sEMG in Neurorobotics

Eion Tyacke, Kunal Gupta, Jay Patel, Raghav Katoch, S. Farokh Atashzar, *Senior Member, IEEE*

Abstract—In the past decade, there has been significant advancement in designing wearable neural interfaces for controlling neurobotic systems, particularly bionic limbs. These interfaces function by decoding signals captured non-invasively from the skin’s surface. Portable high-density surface electromyography (HD-sEMG) modules combined with deep learning decoding have attracted interest by achieving excellent gesture prediction and myoelectric control of prosthetic systems and neurobots. However, factors like small electrode size and unstable electrode-skin contacts make HD-sEMG susceptible to pixel electrode drops. The sparse electrode-skin disconnections rooted in issues such as low adhesion, sweating, hair blockage, and skin stretch challenge the reliability and scalability of these modules as the perception unit for neurobotic systems. This paper proposes a novel deep-learning model providing resiliency for HD-sEMG modules, which can be used in the wearable interfaces of neurobots. The proposed 3D Dilated Efficient CapsNet model trains on an augmented input space to computationally ‘force’ the network to learn channel dropout variations and thus learn robustness to channel dropout. The proposed framework maintained high performance under a sensor dropout reliability study conducted. Results show conventional models’ performance significantly degrades with dropout and is recovered using the proposed architecture and the training paradigm.

I. INTRODUCTION

Portable high-density electromyography (HD-sEMG) signal recording has emerged as a promising technology for enhancing wearable neural interfaces and accurately predicting user motor intention, especially for applications like controlling powered prosthesis and bionic limbs [1]. HD-sEMG provides higher spatial resolution by using closely spaced electrodes over a targeted muscle area [2]. When combined with deep learning decoding architectures like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), the neural interfaces utilizing unperturbed and fixed position recordings have shown high accuracy (>95%) on gesture classification needed for the control of upper-limb prosthetic devices [3], [4].

Of the various deep learning models, Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) are most frequently used in this context [5]–[7]. While CNNs are typically used with a focus on spatial features, RNNs are motivated by encapsulating the temporal dynamics of the signal; both prove beneficial in improving gesture classification accuracy. The hybridization of CNN and RNN has also been explored by recent literature, leading to further improvement

in performance by drawing on the benefits of both [5], [7], [8].

Despite the ability of deep learning models to decode HD-sEMG signal space, their lack of robustness to real-world conditions has prevented large-scale adoption [9]. Current sEMG is inhibited from mass adoption due to poor spatial resolution of decoding, sensitivity to electrode shift, sensitivity to orientation of placement, unpredictable signal variability, and channel dropout [10]–[16]. While there have been some recent attempts at combating HD-sEMG’s shortcomings, notably electrode shift [13], addressing channel dropout is still an open and important problem.

This paper proposes a deep-learning model to process high-density electromyographic signals for controlling neurobotic systems. Our model aims to be resilient to common HD-sEMG challenges like sensor dropout and electrode-skin disconnections by effectively utilizing residual data to ensure optimal performance. The model is also designed to achieve high classification accuracy regardless of electrodes shifting over time. This is important because in real-world applications, EMG sensors gradually deviate from their initial calibrated skin positions, leading to substantial accuracy loss with classic decoding neural networks [17]. Creating a model that maintains performance despite electrode shifts would eliminate the need for frequent sensor reapplication or recalibration.

To capture vital features of HD-sEMG data, we propose the 3D Dilated Efficient CapsNet (3D CapsNet) neural network model, which we train using the specially augmented input space designed in this paper. CapsNet variants classify objects independent of scale, rotation, and other linear transformations while still drawing information from spacial relationships [18]. This makes the architecture ideal for addressing the altered spatial relationships but maintaining global context caused by channel dropout. We have recently shown that the 2D CapsNet significantly outperforms the standard CNN, MLP, and RNN-CNN hybrid model in gesture classification on transient sEMG [19]. The proposed input space augmentation forces the neural network to learn to adjust for the multitude of local variations induced by channel dropout. However, due to the logistical complexity of large-scale data collection from human subjects, it would not be feasible to create an augmented input space by conducting experiments using hundreds of HD-sEMG electrodes. Instead, we proposed generating a specially-designed synthetic data augmentation scheme through which the input space will be altered with disturbed synthetic examples made from undisturbed experimental data. This is to force the network

E. Tyacke is with Carnegie Mellon University (the work was conceptualized when Tyacke was at New York University). K. Gupta, J. Patel, R. Katoch, and S. Farokh Atashzar are with New York University. Atashzar is the corresponding author. f.atashzar@nyu.edu.

to approach better generalizability by being less reliant on the location and dependability of sensors. Augmenting the data will force the model to rely on global channel context to properly classify gestures.

The results showed that the proposed model, trained based on this specialized augmentation, can maintain significant performance under various complex conditions with signal dropout percentages ranging from 10% to 75%, demonstrating robustness and resiliency. Without implementation of the proposed diverse augmentation, a performance drop of $> 37\%$ with only 10% channels dropped was observed. These results highlight the reliance of conventional deep learning EMG decoding models on precise electrode position/location and that without training for robustness, performance degrades rapidly with dropout. We evaluated model performance on data from 19 subjects using a publicly available upper limb sEMG database for benchmarking.

To the best of our knowledge, this is the first work to propose a deep learning model that specifically investigated the feasibility of addressing HD-sEMG decoding robustness to channel dropout for gesture prediction. This technique could help improve adoption of HD-sEMG interfaces in real-world prosthetic and robotic systems. In addition to the above-mentioned main contribution, it should be noted that in this paper, to boost translational power, we focused on the transient sEMG, which is a more challenging problem to address compared to the steady-state signal which has been typically used in the literature [20], [21]. The use of transient sEMG would allow us to enhance the agility in response for real-world scenarios, reduce delay, and make the controls for prosthetic limbs more intuitive [22]. The rest of the paper is organized as follows: Section II provides background on the dataset, data preprocessing, data augmentations, and training strategies; Section III describes the proposed model architectures; Section IV presents the experimental setup, results, and discussion; and Section V concludes the paper.

II. MATERIALS AND METHOD

A. The Utilized Dataset:

The publicly-available HD-sEMG dataset (from [23]) utilized in this paper comprises 20 participants. Each participant performed 65 hand gestures over five repetitions. Each repetition is a single gesture performance, punctuated by a 5-second rest period to avoid muscle fatigue. Among these, 36 gestures were single-degree-of-freedom (DoF) movements, mainly involving flexion and extension of the fingers, wrist, and thumb. Another 25 gestures were two DoF movements. The remaining 5 gestures were complex movements (n-DoF), such as common hand gestures like grips or pinches. Due to data corruption, the information from subject #5 was excluded in this paper. Two 8×8 electrode grids positioned over the extensor and flexor forearm muscles recorded the HD-sEMG signals, sampled at 2048Hz using the Quattrocento biomedical amplifier system. The raw signals were hardware band-pass filtered at 10Hz to 900Hz, and the 50Hz power line frequency removed with a zero-phase 3rd order band-pass Butterworth filter with a 4Hz width.

B. Data Preprocessing:

In this paper, the time window of 1.2 seconds is included from each repetition, which primarily contains the transient and more dynamic phase of the contraction (more details in [13], [24]). This is done to enhance the system's agility in responding to user intentions. The HD-sEMG signals then undergo a three-part transformation before windowing. The data passes through a 4th-order Butterworth band-pass filter from 10Hz to 500Hz, further reducing possible noise or interference and ensuring the signal primarily contains relevant muscle activation information. Z-score normalization then scales the data to zero mean and unit standard deviation. Repetitions 1, 3, and 4 constitute the training data, while repetitions 2 and 5 are the testing data. The data is then rectified, converting all negative values to positive. A sliding prediction window with a length of 200ms and increment of 10ms (95% overlap) is applied.

C. Data Augmentation:

For data augmentation, randomized channel dropout is applied to accurately simulate real-world dropped signals. For each subject, when a channel is dropped, its value is set to zero along the entire recording. Channel dropout is defined by masks. A mask is a 6×6 Boolean array with each value corresponding to each electrode, with a true value indicating a channel should be dropped. Masks are generated according to a dropout rate. For example, if there are 36 electrodes and a mask is generated with a 0.5 dropout rate, 18 channels will be dropped. Each "ring" of the 6×6 mask has an equal proportion of the ring dropped out. For example, with a 0.75 dropout rate, the outer ring containing 20 electrodes will have 15 electrodes randomly dropped, while the second inner ring containing 12 electrodes will have 9 electrodes randomly dropped. In this paper, models are trained on 5 dropout rates: 0%, 10%, 25%, 50%, 75%. Figure 1 visualises the 5 different channel dropout augmentations. For each rate, 6 masks were generated and individually applied to the dataset, then combined with a duplicate of the clean data, effectively increasing the training data seven-fold. We also trained a model on all rates, combining the data of every augmented input space. Multiple masks were used because when trained on a single mask, the model only recognized the dropped pattern of that mask, not improving robustness. This augmentation markedly increased variety in the dataset and overall algorithm performance, mitigating overfitting and improving accuracy compared to models trained on a single mask. In addition, synthetic electrode shift is also applied to simulate slight misplacement of the grids without requiring extensive recollected data with varying placements. Beginning with the two full 8×8 grids, a kernel window of $1 \times 6 \times 6 \times 1$ is selected and the patches extracted with stride one. This procedure is synonymous with the combination of one-step and two-step shifting described by [13].

III. MODEL STRUCTURE

The validated architecture in this paper is named 3D Dilated Efficient CapsNet and is compared with the perfor-

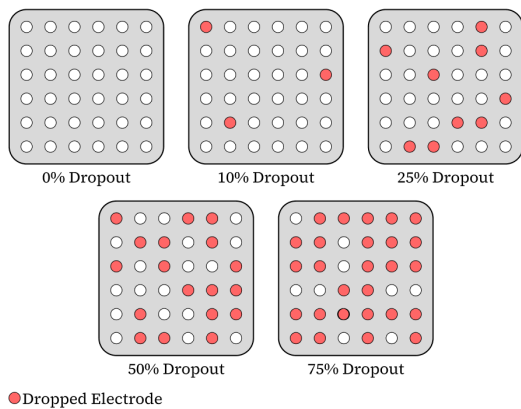


Fig. 1. 5 HD-sEMG modules demonstrating the effect of the channel drop augmentation. A white circle indicates a functional sensor while a red circle indicates a synthetically dropped sensor. The number of dropped sensors is in accordance to the labeled dropout rate.

mance of 3D CNN as the conventional approach. The 3D CNN is chosen as a baseline as it has been effectively used before [13] in the context of electrode shift. The combination of one-step and two-step shifts in [13] is similar to our setting of electrode-shifted data with 0% channel dropout.

A. 3D CNN

The 3D CNN model discussed in this paper follows the settings from [13]. The model consists of a block with five 3D convolutional layers and a classifier with two dense layers. The initial four convolutional layers have a kernel size of (100, 2, 2) and filter counts of 8, 16, 32, and 64, respectively. The last convolutional layer has a (4, 2, 2) kernel size with 128 filters. The dense classifier has hidden layers of sizes 128 and 96.

B. 3D Dilated Efficient CapsNet

The 3D CapsNet architecture (see Fig. 2), inspired by [19], enhances the original Dilated Eff-Caps model for the utilization of 3D data through the conversion of 2D convolutional layers to 3D convolutional layers. In general, in capsular networks, the model classifies objects through collections of neurons called capsules. Each capsule outputs a vector whose length describes the probability of a feature existing and whose orientation describes the feature’s pose. High-probability capsules then predict the output of capsules in the subsequent layer in a process known as routing by agreement. When clusters of capsules agree on a capsule’s output in the subsequent layer, that capsule is activated. The process is repeated on each layer using activated capsules until a steady state is reached. This architecture organizes a feature hierarchy not reliant on spatial information, making it excellent at classifying gestures despite augmented input spaces. Within the convolutional block of the model, the first two convolutional layers have 128 features each, and the last two have 256 features; each layer possesses a kernel size of (3, 2, 2) and is followed by a dropout layer with a dropout rate of 0.5 and a batch normalization layer. To increase the receptive field along the temporal dimension

for successive layers without increasing the convolutional kernel size, we employ an increasing dilation rate from convolutional layers two through four with rates of (4, 1, 1), (16, 1, 1), and (64, 1, 1). Within the capsule block, both the primary convolution and the capsule layer utilized a capsule dimension of eight. Additionally, the primary convolution layer has four channels, a kernel size of (3, 1, 1), and a stride of two.

IV. EXPERIMENT AND RESULTS

To evaluate the 3D CNN and CapsNet performances, each model was tested on the electrode shifted data at 0%, 10%, 25%, 50%, and 75% channel dropout rates. Each test assessed a model’s average performance across 30 randomly selected drop channel patterns. These randomly selected test masks are distinct from those used in the processes of generating the augmented input space for training and are consistent between each model evaluation. Fig. 3 and 5 display their respective base model and its variants. The x-axis shows rising dropout percentages, and the y-axis shows the 30-mask average accuracy of each model.

To better evaluate our models’ robustness, we used the two-sided Mann-Whitney-Wilcoxon test for statistical significance testing. This test compared the performance distributions between model variants tested with 0% channel dropout and each increasing level of channel dropout. The Bonferroni correction was applied to provide a more conservative estimate when conducting multiple tests, reducing the risk of making a Type I error. In this paper, significance markers ns, *, **, ***, and **** correspond to p-values of 1-0.05, 0.05-0.01, 0.01-0.001, 0.001-0.0001, and < 0.0001 respectively. The results of this testing can be seen in Tables II and IV.

A. Performance evaluation for the 3D CNN architecture

We compared a 3D CNN trained with 0% dropout (baseline) to the same model trained on five augmented input spaces. Results are shown in Fig. 3 and 4 and are summarized in Tables. I, and II. As illustrated in Fig. 3, the augmented input spaces (shown by different colors) are tested for channel dropout rates ranging from 10% to 75% (x-axis of the figure). The figure demonstrates that our augmented training approach improves the robustness of the gesture classifier. This enhancement is evident from the superior average accuracy rates achieved by all non-baseline models under dropout conditions. As the dropout rate in the augmented training set increases (represented by different color codes in Fig. 3 and 4) from 0% to 75%, the impact of randomized dropout intensity (represented by the x-axis of Fig. 3 and 4) reduces, resulting in a boost in the model’s average accuracy. The increasing trend of ‘no significant difference’ across the rows of Table II indicates enhanced consistency of model performance and thus robustness to dropout. Both Fig. 3 and 4 emphasize this, with non-baseline models displaying higher averages and smaller slopes (which means less variability caused by dropout). Notably, the model trained with the combined training mask stands out with the highest accuracy

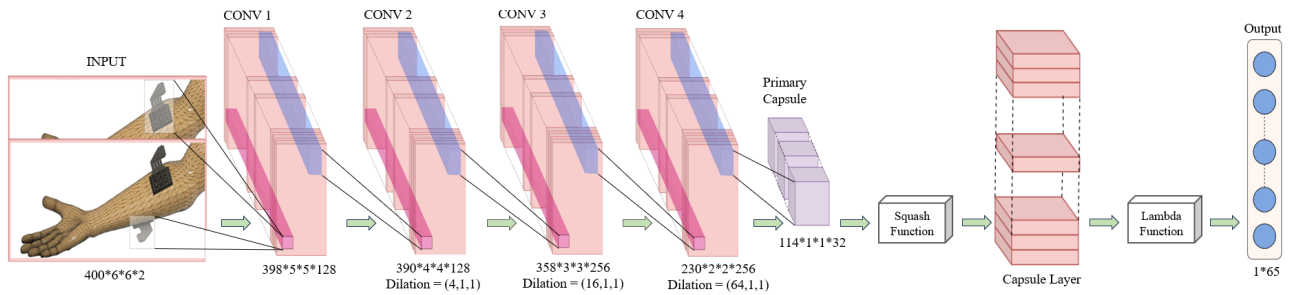


Fig. 2. An illustration of the proposed 3D Dilated Efficient CapsNet architecture. There are four 3D Convolutional layers (Conv). The Capsule block has a Primary Conv followed by a squash function, a Capsule layer to perform routing-by-agreement, and a Lambda Function to generate the output confidence for each of the 65 gestures.

and the highest robustness to randomized dropouts shown, underscoring the performance of the proposed augmentation.

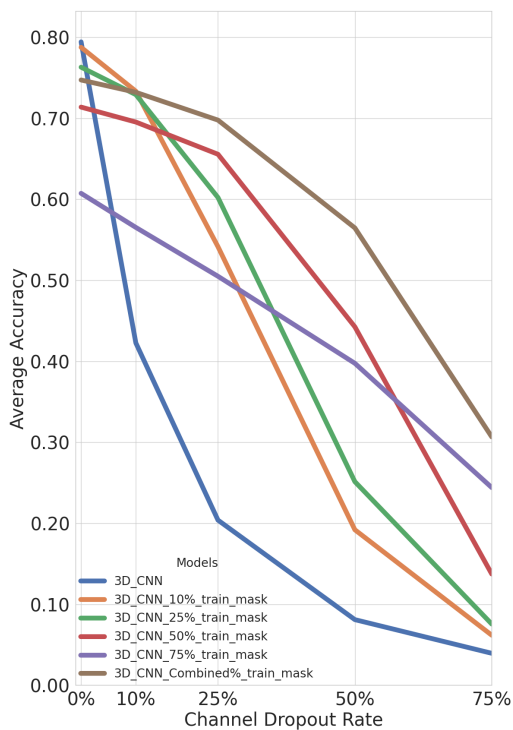


Fig. 3. 3D CNN average accuracy results for each level of augmentation evaluated on 30 randomly assigned dropped channel configurations at each dropout rate. X-axis is the dropout rate and different colors show different augmentation strategy.

TABLE I
RESULTS OF CHANNEL DROPOUT RATES ON 3D CNN
(AVERAGE ACCURACY)

Dropout Rate→	0%	10%	25%	50%	75%
0% Augmentation	0.794	0.423	0.204	0.081	0.040
10% Augmentation	0.787	0.733	0.541	0.192	0.062
25% Augmentation	0.763	0.729	0.602	0.251	0.076
50% Augmentation	0.714	0.695	0.655	0.442	0.137
75% Augmentation	0.607	0.565	0.505	0.397	0.244
Combined	0.747	0.732	0.698	0.565	0.307

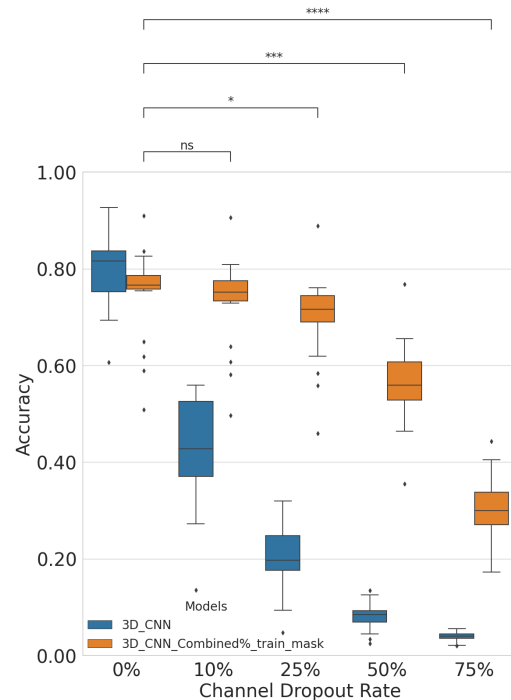


Fig. 4. 3D CNN results at each dropout level for training on the combination of all channel dropout data augmented strategies for each level

TABLE II
RESULTS OF CHANNEL DROPOUT RATES ON 3D CNN
(STATISTICAL SIGNIFICANCE)

Dropout Rate→	0%	10%	25%	50%	75%
0% Augmentation	-	****	****	****	****
10% Augmentation	-	*	****	****	****
25% Augmentation	-	ns	****	****	****
50% Augmentation	-	ns	ns	****	****
75% Augmentation	-	ns	**	****	****
Combined	-	ns	*	***	****

B. Performance evaluation for the 3D Dilated Efficient CapsNet architecture

Similarly to the 3D CNN, our proposed CapsNet model was tested against different channel dropout levels, with and without training on augmented datasets. The results, shown in Fig. 5 and 6, reveal that the CapsNet models boosted by the augmented training on channel dropout outperform the

baseline model (when no augmentation is conducted). When the proposed CapsNet is trained on the combination of all augmentation data (Table IV and Fig. 6), consistent model performance was observed with no statistically significant difference when dropping 25%, 10% and 0% of the channels; and the model robustly maintains the performance of 83%, resembling robustness, making it the best-performing model among all the CapsNets and 3D CNNs evaluated in this study. Conducting a stronger stress test, the performance of the proposed model remains above 70% even in the presence of 50% dropout, when the baseline model can only maintain the performance of $\approx 6\%$ (see Table III). The results highlight the resiliency of the proposed model to even severe cases of drop out highlighting the importance of the proposed synthetic input augmentation for practical application.

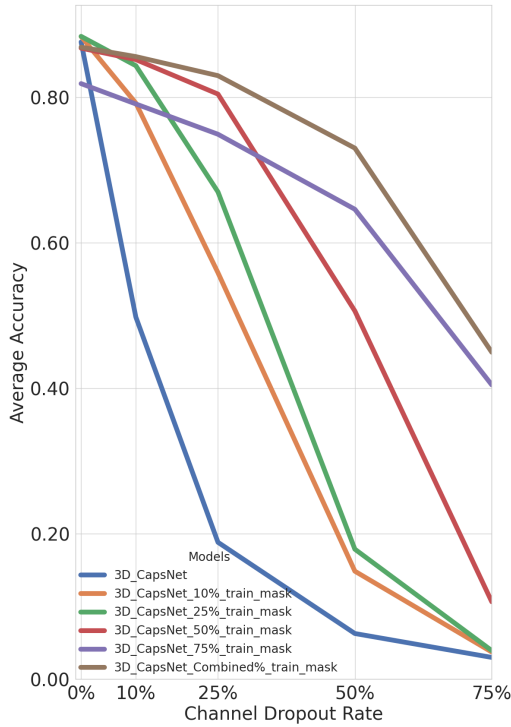


Fig. 5. 3D CapsNet average accuracy results for each level of augmentation evaluated on 30 randomly assigned dropped channel configurations at each dropout rate

TABLE III
RESULTS OF CHANNEL DROPOUT RATES ON 3D CAPSNET
(AVERAGE ACCURACY)

Dropout Rate →	0%	10%	25%	50%	75%
0% Augmentation	0.876	0.498	0.188	0.063	0.030
10% Augmentation	0.884	0.792	0.559	0.149	0.038
25% Augmentation	0.884	0.844	0.670	0.179	0.039
50% Augmentation	0.868	0.852	0.805	0.506	0.107
75% Augmentation	0.819	0.791	0.749	0.646	0.405
Combined	0.869	0.856	0.830	0.730	0.450

Figures 4 and 6 demonstrate that models trained on the combined augmented dataset maintain significantly high

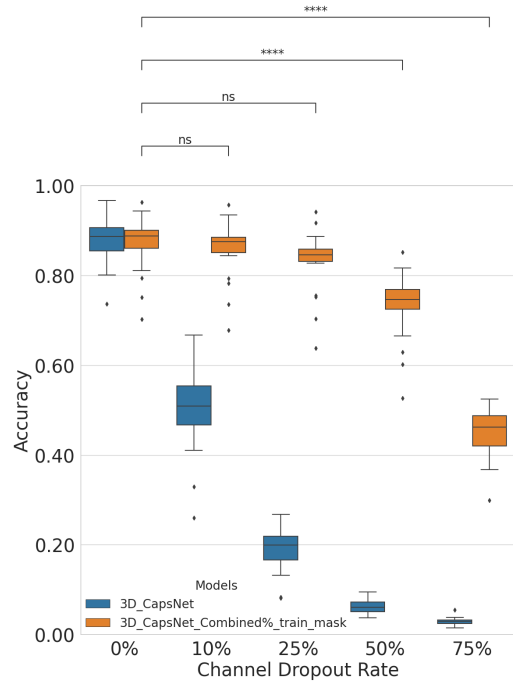


Fig. 6. 3D CapsNet results at each dropout level for training on the combination of channel dropout data augmented for each level

TABLE IV
RESULTS OF CHANNEL DROPOUT RATES ON 3D CAPSNET
(STATISTICAL SIGNIFICANCE)

Dropout Rate →	0%	10%	25%	50%	75%
0% Augmentation	-	****	****	****	****
10% Augmentation	-	***	****	****	****
25% Augmentation	-	ns	****	****	****
50% Augmentation	-	ns	**	****	****
75% Augmentation	-	ns	*	****	****
Combined	-	ns	ns	****	****

accuracy and exhibit smaller interquartile ranges up to a 25% experimental channel dropout rate, compared to their base model. The results in Fig. 5 and Table III suggest that models trained with augmentation by higher dropout rates tend to outperform those trained with augmentation by lower dropout rates, even when evaluated on data with a lower number of channels dropped. Notably, the 3D CapsNet model, when trained through augmentation by 50% dropout augmentation strategy, outperformed the model augmented by 10% dropout strategy and 25% dropout strategy for all experimental ratios of sensor dropout (including the low and high rates). As expected when combining all the augmentation strategies for training the CapsNet due to the large size of input space it requires much higher computational support for training. Thus based on the results, it seems that training the model using an augmentation strategy of 50% dropout emerges as an efficient and less resource-intensive option with high performance. The model trained by 50% dropout augmentation can simulate a range of dropout scenarios by learning diverse patterns between small and sparse clusters

of electrodes rather than relying on a limited number of large, dense patches. This highlights the significance of learning features at various scales, as the smaller and sparser features learned at the 50% augmentation level contribute to better generalization. It should be highlighted that training to combat higher dropout rates is a trade-off between computing resources and accuracy. As the channel dropout rate increases, the possible combinations of channels that remain active, or ‘live’, also increases. This means that to effectively train the model at higher dropout rates, the number of masks used to augment the training data should be increased. This would expose the model to a wider range of potential configurations of ‘live’ channels. However, it can increase the computational cost during the training.

V. ACKNOWLEDGEMENT

The authors would like to acknowledge the help of Ananya Shah in the project. The authors would also like to acknowledge the support from NYU IT High-Performance Computing resources, services, and staff.

VI. CONCLUSION

This paper presents a specialized data augmentation methodology to increase the performance of high-density sEMG-based decoding with ultimate application in neuro-robotics. The goal is to recover robustness in the presence of sensor pixel dropouts, which can significantly challenge the performance of the interfaces in practice. Applying the proposed data augmentation methodology on the majority transient phase of HD-sEMG signal and training the proposed Capsular Neural Network model has shown the viability of the sustained performance of such systems in a degraded setting, achieving an accuracy of 83% when 0-25% channel dropout was present for 65 gesture prediction. The model kept the performance above 70% in the presence of severe dropout (i.e., 50%) when the baseline model could only provide $\approx 6\%$ performance due to the complexity of the task. In other words using the proposed framework we have shown that high-density electromyography provides a very rich source of information that can secure high performance if the model is trained in a way that do not rely on spatial locations of the information source across the targeted muscle. Addressing the drop-out issue through the proposed robust system opens the door to wider acceptance of peripheral wearable devices for control of neuro-robots.

REFERENCES

- [1] Z. Chen *et al.*, “A review of myoelectric control for prosthetic hand manipulation,” *Biomimetics*, vol. 8, no. 3, p. 328, 2023.
- [2] X. Jiang *et al.*, “Optimization of hd-semg-based cross-day hand gesture classification by optimal feature extraction and data augmentation,” *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 6, pp. 1281–1291, 2022.
- [3] L. Alzubaidi *et al.*, “Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions,” *Journal of big Data*, vol. 8, pp. 1–74, 2021.
- [4] M. Montazerin *et al.*, “Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density emg signals,” *Scientific Reports*, vol. 13, no. 1, p. 11000, 2023.

- [5] Y. Hu *et al.*, “A novel attention-based hybrid cnn-rnn architecture for semg-based gesture recognition,” *PloS one*, vol. 13, no. 10, p. e0206049, 2018.
- [6] R. Antonius and H. Tjahyadi, “Electromyography gesture identification using cnn-rnn neural network for controlling quadcopters,” in *Journal of Physics: Conference Series*, vol. 1858, no. 1. IOP Publishing, 2021, p. 012075.
- [7] S. Li *et al.*, “Real-time semg pattern recognition of multiple-mode movements for artificial limbs based on cnn-rnn algorithm,” *Electronics*, vol. 12, no. 11, p. 2444, 2023.
- [8] A. T. Erözen, “A new cnn approach for hand gesture classification using semg data,” *Journal of Innovative Science and Engineering (JISE)*, vol. 4, no. 1, pp. 44–55, 2020.
- [9] P. Shi *et al.*, “Improving the robustness and adaptability of semg-based pattern recognition using deep domain adaptation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 11, pp. 5450–5460, 2022.
- [10] Z. Li *et al.*, “Electrode shifts estimation and adaptive correction for improving robustness of semg-based recognition,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 1101–1110, 2020.
- [11] S. E. Ovrur *et al.*, “A novel autonomous learning framework to enhance semg-based hand gesture recognition using depth information,” *Biomedical Signal Processing and Control*, vol. 66, p. 102444, 2021.
- [12] N. G. Nia *et al.*, “Reinforcement learning-based grasp pattern control of upper limb prosthetics in an ai platform,” in *2022 3rd International Informatics and Software Engineering Conference (IISEC)*. IEEE, 2022, pp. 1–4.
- [13] T. Sun *et al.*, “Deep augmentation for electrode shift compensation in transient high-density semg: Towards application in neuro-robotics,” in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 6148–6153.
- [14] Z. Fan *et al.*, “A canonical correlation analysis based emg classification algorithm for eliminating electrode shift effect,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 867–870.
- [15] P. Koch *et al.*, “Regression of hand movements from semg data with recurrent neural networks,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 3783–3787.
- [16] P. Wang *et al.*, “Prediction of emg signal on missing channel from signal captured from other related channels via deep neural network,” in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2018, pp. 1287–1291.
- [17] S. Kanoga *et al.*, “Are armband semg devices dense enough for long-term use?—sensor placement shifts cause significant reduction in recognition accuracy,” *Biomedical Signal Processing and Control*, vol. 60, p. 101981, 2020.
- [18] G. Chen *et al.*, “Two-dimensional discrete feature based spatial attention capsnet for semg signal recognition,” *Applied Intelligence*, vol. 50, pp. 3503–3520, 2020.
- [19] E. Tyacke *et al.*, “Hand gesture recognition via transient semg using transfer learning of dilated efficient capsnet: Towards generalization for neuro-robotics,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9216–9223, 2022.
- [20] F. Peng *et al.*, “Locomotion prediction for lower limb prostheses in complex environments via semg and inertial sensors,” *Complexity*, vol. 2020, pp. 1–12, 2020.
- [21] R. Ma *et al.*, “Grasping force prediction based on semg signals,” *Alexandria Engineering Journal*, vol. 59, no. 3, pp. 1135–1147, 2020.
- [22] A. Hagenruber *et al.*, “A new labeling approach for proportional electromyographic control,” *Sensors*, vol. 22, no. 4, p. 1368, 2022.
- [23] N. Malešević *et al.*, “A database of high-density surface electromyogram signals comprising 65 isometric hand gestures,” *Scientific Data*, vol. 8, no. 1, p. 63, 2021.
- [24] T. Sun *et al.*, “Deep heterogeneous dilation of lstm for transient-phase gesture prediction through high-density electromyography: Towards application in neuro-robotics,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2851–2858, 2022.