ENABLING MYOELECTRIC CONTROL TRAINING USING CONTINUOUS DATA THROUGH SELF-SUPERVISED REPRESENTATION LEARNING

Shriram Tallam Puranam Raghu[1], Dawn MacIsaac[1], and Erik Scheme[1]

[1]*Institute of Biomedical Engineering, University of New Brunswick, Canada*

ABSTRACT

In this work, we explore the potential of integrating continuous transition data into the training process for pattern recognition-based myoelectric control. We use a set of steady-state and continuous transition performance metrics to compare the performance of classifiers trained with continuous data versus the traditional ramp contraction approach. We further compare the performance of the popular LDA classifier with that of a deep gated recurrent unit (GRU) classifier capable of leveraging the temporal dynamics. We also introduce a novel self-supervised contrastive representation learning approach with augmentations that significantly improves the offline steady-state and transition performance. This work provides compelling early evidence of the potential for semi-supervised learning approaches to leverage temporal dynamics in continuous training data to improve the performance of pattern recognition-based myoelectric control.

INTRODUCTION

Pattern recognition (PR) based myoelectric control has been heavily explored due to its ability to provide intuitive control via learned patterns of surface electromyography (sEMG) signals from multiple channels [1]. Its susceptibility to various sources of noise and confounding factors, however, has encouraged researchers to continue to improve its robustness through various algorithmic and training approaches.

Previous works have shown that training PR classifiers with ramp data, where users increase their contraction intensity from rest, rather than static contractions provides improved online (usability) performance [2]. This is likely because ramp data offer a fuller representation of the motion classes through their inclusion of contraction dynamics. While these dynamics help the classifier to learn about how to transition from no movement to each motion class, they don't offer much help in learning about transitions from one motion class to another. This is an important area of exploration, since studies have indicated that errors during transition impact usability [3], [4]. It is conceivable that including examples of these 'continuous transitions' in training data may further inform PR models beyond what is achieved through ramp data alone, especially if they are able to better inform temporal models like long short-term memory networks (LSTMs) or gated recurrent units (GRUs), which are capable of exploiting time series dynamics. The purpose of this work was to explore this possibility.

Training with data that includes continuous transitions from one motion class to another, however, presents new challenges. Besides increased burden on users to collect more training data, supervised classifiers like the conventional LDA, LSTM, and GRU all require labeled training data, and it is not clear how to label regions of transition, or even identify region bounds. Until a robust EMG-based algorithm for marking transition bounds between active classes is established, a secondary source such as a motion sensor could be used. Even when the bounds are identified, a labeling scheme must be chosen; for example, a naive scheme could label all frames within a transition as the next motion class, hoping to drive the decision stream to the next class quickly. Alternatively, self-supervised learning approaches could offer a possible solution, eliminating the need for pre-established labels altogether.

Self-supervised approaches do not require labels to train the model, and instead rely on augmentations to learn the structure of the data [5]. The model learns to maximize similarity of two augmented views of the same sample; for example, during training, two scaled-amplitude versions of a frame would be tagged as matches so the model could be updated accordingly. This makes them robust to perturbations and particularly interesting for learning useful information from hard-to-label dynamic data. They construct a latent space that drives similar samples closer together, which may have the effect of 'clustering' motion classes in a way that includes transition frames appropriately. Once the feature vectors from the latent space are established, a strategy that focuses on those members of each cluster that are easily classed with their pre-established labels (i.e. a group of steady state members) can be used to establish the motion classes of latent space feature vectors so they can be properly classified during inference.

In this work, we explore the potential of leveraging such a semi-supervised representation learning approach to leverage dynamics in continuous transition training data to inform PR-based myoelectric control. We show that the proposed approach significantly improves offline steady-state and transition performance compared the traditional ramp approach using a set of steady-state and continuous transition performance metrics [6].

METHODS

A dataset collected from 43 able-bodied participants, fully described in [7] and approved by the UNB Research Ethics Board (REB #2021-116), was used in this study. Briefly, each participant completed 5 ramp trials and 6 continuous transition trials, following screen prompts of 6 motion classes: Wrist Flexion (WF), Wrist Extension (WE), Wrist Pronation (WP), Wrist Supination (WS), Chuck Grip (CG), and Hand Open (HO), and a No Motion (NM) class. Ramp contractions always began in the No Motion class and gradually increased in intensity to reach the steady state contraction of a motion class (about 3 s). Continuous transition trials started in the No Motion class and then randomly transitioned between classes continuously, holding the steady between each transition for about 3 s, until all transitions between each class and every other (including No Motion) were complete. All 42 transitions were captured in a random order each trial.

All data were collected using six channels of sEMG sampled at a rate of 2kHz using Delsys Trigno ®electrodes spaced equidistantly around the circumference of the participant's right forearm. Each trial was segmented into overlapping frames with a length of $162\,\mathrm{ms}$ and an increment of $13.5\,\mathrm{ms}$. Then, the Low-Sampling Frequency 4 (LSF4) feature set [8] was extracted from each frame, as this feature set has been shown to robust and generalizable. Kinematic data were also collected with a Leap Motion Controller (LMC) infrared sensor positioned below the hand so that transitions between motion classes could be identified through movement.

Two classifiers – a Linear Discriminant Analysis (LDA) classifier, and a deep Gated Recurrent Unit (GRU) classifier – were evaluated across the two training conditions (ramp and continuous). Both conditions were *evaluated* on continuous transition data using the following offline performance metrics: steady-state Active Error Rate (SS-AER), steady-state Total Error Rate (SS-TER), and steady-state Instability (SS-INS), and transition: Offset Delay (T_{OFF}), Onset Delay (T_{ON}), Transition Duration (T_{TD}), Instability (INS), Tertiary Class Error (TER), and Percent No Movement (PNM) [6].

All five trials of *the ramp data* were used to train the LDA classifer, which was evaluated using all six trials of the continuous transition data. When using the GRU, one random trial was used for validation and the remaining four trials were used for training. When training with *the continuous data*, a leave-one-trial-out approach was used for evaluation. For each participant, after removing one trial for evaluation, the remaining 5 were used to train the LDA. Again, for the GRU, one of the five trials was held out for validation, and the remaining four were used for training. This process was repeated six times so that trained models could be evaluated on all six trials of the continuous transition data. To establish the labels required for supervised learning of the continuous data, the frames were re-labelled according to a modified version of the prompts. The start of each transition was identified as the first frame after a prompt change that coincided with the initiation of movement (as identified by the LMC). All the frames after this frame were labeled as belonging to the prompted motion class.

A conventional cross-entropy loss function was used to train the *supervised* GRUs described above. Additionally, a novel *self-supervised* contrastive loss function called VICReg [9] was investigated to avoid the need for the LMC-dependent re-labeling approach when using the continuous transition data. Although many augmentations have been proposed in the literature for use in self-supervised learning [10], random feature scaling, additive white Gaussian noise, and time shifting were used in this work. Two augmented views of the frames in the training data were generated, and these augmented views were then used train the self-supervised GRU. Once the model was trained, the prompt-labeled latent space feature vectors were used to determine the centroid of each motion class, so that a Nearest Centroid (NC) classifier could be applied on the test trials during evaluation.

RESULTS

Figure 1 show box plots of the steady-state and transition metrics for each of the models. One-way ANOVA ($\alpha = 0.05$) comparing training conditions (ramp vs continuous transition) for *supervised* models showed significant differences in all three steady-state metrics ($p < 0.0001$) regardless of the classifier. For transition metrics, four of the six metrics showed significant differences ($p < 0.02$). The classifiers had significantly shorter transition duration when trained with continuous data, but at the expense of slight but significantly longer T_{OFF}. INS also showed a significant difference, but a post-hoc multiple comparison test with Šidák correction ($\alpha = 0.05$) revealed that only GRU[R] had statistically lower values ($p < 0.0001$).

One-way ANOVA ($\alpha = 0.05$) comparing the classifiers trained with continuous transition data showed significant differences in SS-INS ($p < 0.0001$). For transition metrics, T_{OFF}, INS, and TCE also showed significant differences. A post-hoc multiple comparison test with Šidák correction ($\alpha = 0.05$) revealed that VICReg had statistically lower values ($p < 0.0001$) for SS-INS, INS and TCE, but a statistically higher value for T_{OFF} ($p < 0.0001$).

Figure 2 shows an example of the decision and probability streams of nearest centroid classifiers trained using the original, unmodified feature space, and that of the VICReg latent space. Figure 3 shows an example of the latent space learnt by VICReg for the ramp and continuous training data compared to that of a continuous testing trial.

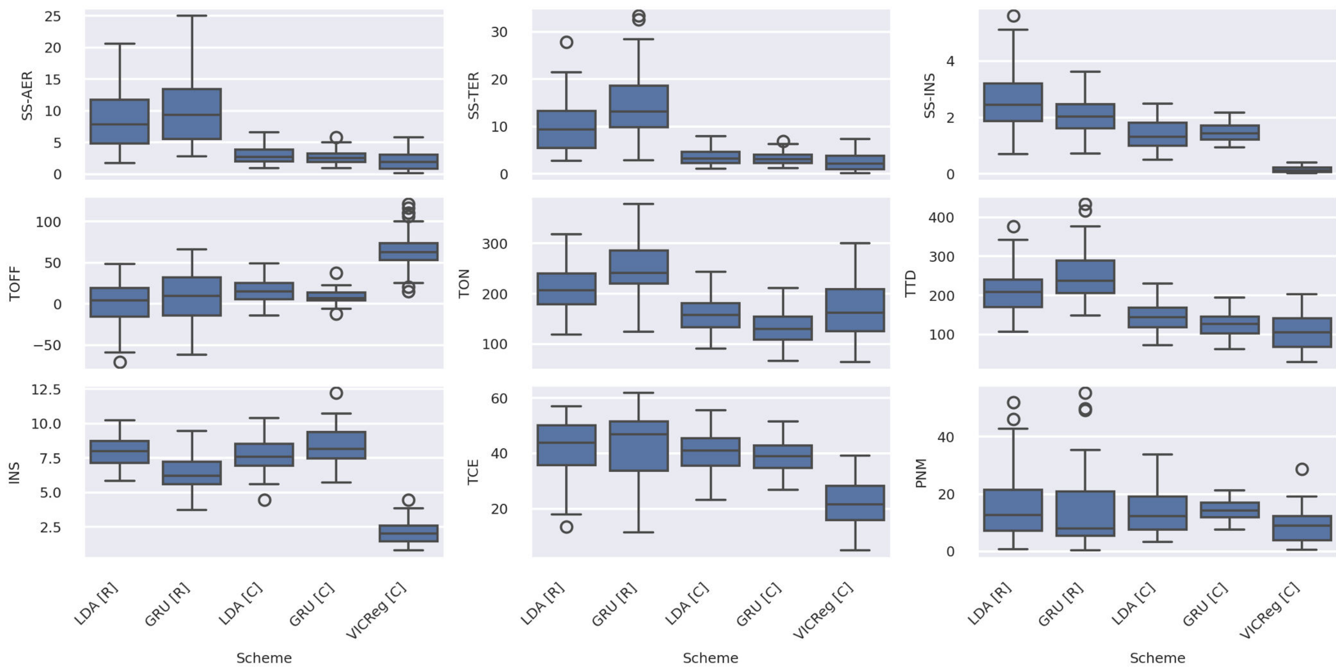

Fig. 1: Comparison of the performance of the different classifiers. The top row shows steady-state metrics, whereas the bottom two rows show transition metrics. Classifier schemes denoted as [R] were trained with Ramp data and [C] were trained with continuous data.

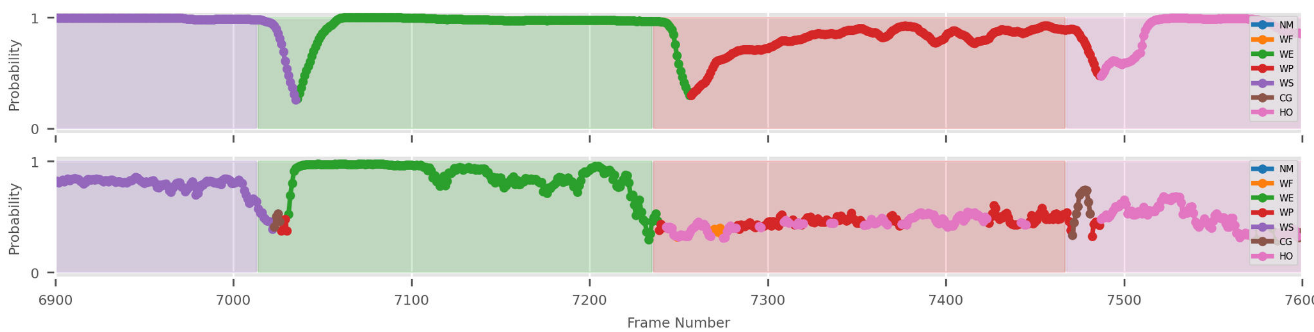

Fig. 2: An example of the nearest centroid classifier probability and decision streams when using the original, unmodified feature space (Bottom) and when using the VICReg latent space (Top). The color of the lines represents the classifier output, whereas the background color denotes the prompted class.

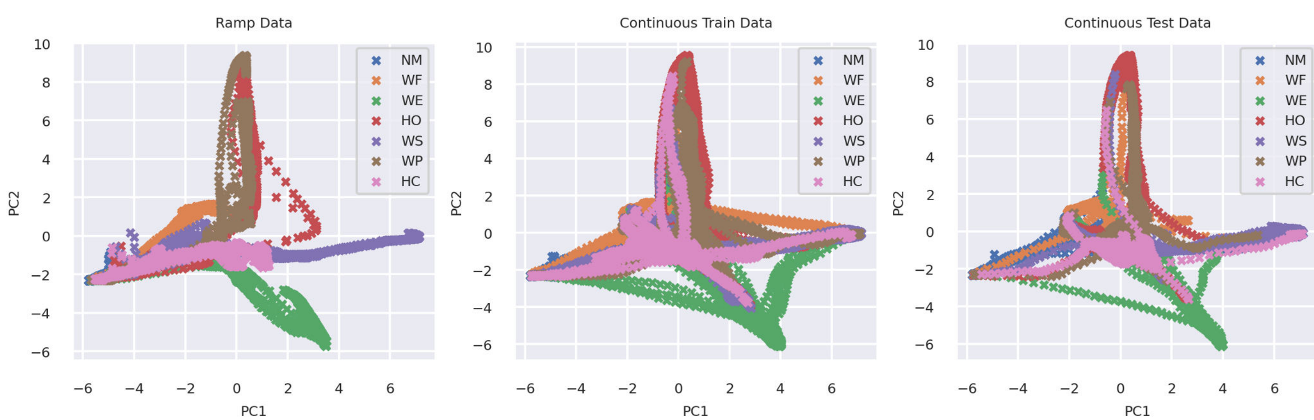

Fig. 3: A visualization of the VICReg latent space for an example participant, projected into a 2 dimensional space using PCA. Left: Ramp Training Data, Middle: Continuous Training Data, Right: Continuous Testing Data

DISCUSSION

Results indicate that classifiers trained with continuous data performed better than those trained with ramp data, according to most of the offline metrics. However, we observed that LDA and GRU performed on par in all metrics except T_{ON}. This was surprising given that the GRU is a temporal classifier and should be better at exploiting temporal information in the data compared to the LDA. It may be that the naive labeling strategy prevented the GRU from fully exploiting its capacity to learn dynamics. This is supported by the fact that VICReg, which was designed to avoid the need for labelling, showed significantly improved performance in terms of INS, SS-INS, and TCE. Regardless, to fully exploit training supervised classifiers with continuous transition data, more work is required regarding how best to label transitions. In particular, in a closed-set classification problem such as this, the data during transitions does not truly belong to any of the prescribed labels. Consequently, labelling them as belonging to an 'unknown' class or applying label smoothing [11] may be more appropriate. Also, requiring the use of an LMC device to mark the bounds of transition regions is undesirable and even infeasible for persons with limb differences. Furthermore, EMG activity precedes kinematic movement [12], which was not accounted for in this work. Consequently, a segmentation technique that works directly with the raw EMG data in real-time would be more more desirable, though more work is required to solve this challenging problem.

Resistant to these labeling issues, the self-supervised VICReg classifier also yielded considerably better performance across many of the metrics, particularly in instability and transition errors. The generated latent spaces shown in Figure 3 suggests that the embedding not only clusters points belonging to the same class together (even though the model had never seen any class labels), but that it also encodes transitions into narrow corridors travelling between those clusters in a seemingly repeatable way. The two figures with continuous data also show the better population of these transition regions. This behavior is further evidenced by the probability decision stream when using VICReg in Figure 2, whose probability is more consistent while in steady state and transitions directly from the previous class to the target class.

While the self supervised VICReg improves the stability of the decision stream, however, it appears to do so at the expense of a slight increase in onset lag. If this lag isn't perceptible to device users during online use, the improved stability and transitions may substantially improve usability. However, if this lag is perceivable, it may cause the control to feel sluggish. Subsequent work should continue to explore this effect and if other augmentations may overcome this trade-off. Ongoing work is testing these models in an online Fitt's law test and seeking to correlate the offline and online metrics.

Even with the noted benefits of using continuous transition training data, the increased burden on the users to collect these data is a critical disadvantage. Fortunately, techniques that rely on transfer learning and domain adaptation are emerging and may be leveraged to mitigate this issue [13]. Such techniques may be able to shift the burden of collecting most training data away from the end user while still providing most, if not all, of the observed benefits of training with continuous data. Further research into such strategies may be necessary to make training with continuous data practical.

REFERENCES

[1] M. Asghari Oskoei and H. Hu, "Myoelectric control systems—A survey," *Biomedical Signal Processing and Control*, vol. 2, no. 4, pp. 275–294, Oct. 2007.

[2] L. Hargrove, Y. Losier, B. Lock, K. Englehart, and B. Hudgins, "A Real-Time Pattern Recognition Based Myoelectric Control Usability Study Implemented in a Virtual Environment," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2007, pp. 4842–4845, iSSN: 1558-4615.

[3] J. W. Robertson, K. B. Englehart, and E. J. Scheme, "Effects of Confidence-Based Rejection on Usability and Error in Pattern Recognition-Based Myoelectric Control," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 5, pp. 2002–2008, Sep. 2019.

[4] A. M. Simon, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "A Decision-Based Velocity Ramp for Minimizing the Effect of Misclassifications During Real-Time Pattern Recognition Control," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 8, pp. 2360–2368, Aug. 2011.

[5] R. Balestriero, M. Ibrahim, V. Sobal, A. Morcos, S. Shekhar, T. Goldstein, F. Bordes, A. Bardes, G. Mialon, Y. Tian, A. Schwarzschild, A. G. Wilson, J. Geiping, Q. Garrido, P. Fernandez, A. Bar, H. Pirsiavash, Y. LeCun, and M. Goldblum, "A Cookbook of Self-Supervised Learning," Jun. 2023, arXiv:2304.12210 [cs].

[6] S. Tallam Puranam Raghu, D. MacIsaac, and E. Scheme, "Analyzing the impact of class transitions on the design of pattern recognition-based myoelectric control schemes," *Biomedical Signal Processing and Control*, vol. 71, p. 103134, Jan. 2022.

[7] S. T. P. Raghu, D. MacIsaac, and E. Scheme, "Decision-Change Informed Rejection Improves Robustness in Pattern Recognition-Based Myoelectric Control," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 12, pp. 6051–6061, Dec. 2023.

[8] A. Phinyomark, R. N. Khushaba, and E. Scheme, "Feature Extraction and Selection for Myoelectric Control Based on Wearable EMG Sensors," *Sensors*, vol. 18, no. 5, p. 1615, May 2018.

[9] A. Bardes, J. Ponce, and Y. LeCun, "VICReg: Variance-Invariance-Covariance Regularization for Self-Supervised Learning," Jan. 2022, arXiv:2105.04906 [cs].

[10] X. Zhang, Z. Zhao, T. Tsiligkaridis, and M. Zitnik, "Self-Supervised Contrastive Pre-Training For Time Series via Time-Frequency Consistency," *Advances in Neural Information Processing Systems*, vol. 35, pp. 3988–4003, Dec. 2022.

[11] R. Müller, S. Kornblith, and G. E. Hinton, "When does label smoothing help?" in *Advances in Neural Information Processing Systems*, vol. 32. Curran Associates, Inc., 2019.

[12] G. L. Gottlieb, "Muscle Activation Patterns During Two Types of Voluntary Single-Joint Movement," *Journal of Neurophysiology*, vol. 80, no. 4, pp. 1860–1867, Oct. 1998.

[13] X. Chen, Y. Li, R. Hu, X. Zhang, and X. Chen, "Hand Gesture Recognition based on Surface Electromyography using Convolutional Neural Network with Transfer Learning Method," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 1292–1304, Apr. 2021.