

Contrastive Learning for Trial-Robust EMG Pattern Recognition

Rina Tazaki* and Akira Furui*

Graduate School of Advanced Science and Engineering, Hiroshima University, Higashi-Hiroshima, Japan

{rinatazaki, akirafurui}@hiroshima-u.ac.jp

*These authors contributed equally to this work.

Abstract—Surface electromyography (EMG)-based pattern recognition enables intuitive human-machine interaction but suffers from significant performance degradation across recording trials due to electrode repositioning, muscle fatigue, and user adaptation. While contrastive learning has been applied to biosignal analysis for addressing inter-subject variability, its potential for mitigating trial-level variability in EMG recognition remains unexplored. This paper presents a systematic investigation of supervised contrastive learning for trial-robust EMG pattern recognition. We evaluate the standard supervised contrastive loss (SupCon) and propose Cross-Trial-Contrast (CTC) loss, which explicitly leverages trial information by treating only cross-trial same-class samples as positive pairs. Experiments on three public datasets with diverse trial structures demonstrate that contrastive learning consistently improves both classification accuracy and representation quality compared to cross-entropy training. The proposed CTC loss achieves the highest accuracy on two of three datasets while consistently improving representation quality metrics. These results demonstrate that explicit trial-aware positive pair selection enhances both representation learning and classification performance. Our findings provide practical insights for applying contrastive learning to EMG-based pattern recognition systems.

Index Terms—Electromyography, pattern recognition, contrastive learning, trial variability, deep learning

I. INTRODUCTION

Surface electromyography (EMG) provides a non-invasive window into muscle activity and has been widely adopted for pattern recognition in applications such as prosthetic control, human-computer interaction, and rehabilitation [1]. Recent advances in deep learning have enabled high classification accuracy under controlled laboratory conditions, with neural networks effectively capturing the complex spatiotemporal patterns embedded in EMG signals. However, translating these achievements to practical deployment remains challenging, as real-world use involves variations in recording conditions that substantially degrade recognition performance.

A particularly problematic source of variability arises when the same user performs gestures across different recording trials or sessions. Even with the same electrode placement protocol, EMG signals exhibit substantial variations due to electrode repositioning, muscle fatigue, and changes in motor execution strategies [2], [3]. This trial-to-trial variability causes models trained on initial trials to perform poorly on subsequent recordings, presenting a significant barrier to the

deployment of EMG-based systems in daily use scenarios where consistent long-term performance is essential.

Several approaches have been proposed to address this challenge, including calibration-based methods that require additional data collection and domain adaptation techniques such as Domain-Adversarial Neural Networks (DANN) [4], [5]. Meanwhile, contrastive learning has emerged as a powerful paradigm for learning robust representations by encouraging models to capture semantically meaningful features invariant to nuisance factors [6]. Recent biosignal studies have applied contrastive learning to address inter-subject variability—for instance, CLISA [7] learns subject-invariant electroencephalogram (EEG) representations, and STCNet-SAC [8] incorporates subject-aware contrastive learning for EMG. However, these approaches primarily target inter-subject invariance rather than intra-subject trial-to-trial variability, and the potential of contrastive learning for mitigating the trial-to-trial variability remains unexplored.

Supervised contrastive learning pulls together samples of the same class while pushing apart samples of different classes, producing representations that are robust to intra-class variations. We posit that this property can help mitigate trial-induced variability: by requiring same-class samples from different trials to have similar representations, the encoder is encouraged to focus on gesture-relevant features rather than trial-specific characteristics. Furthermore, we investigate whether explicitly incorporating trial information into the contrastive objective, by treating only cross-trial same-class pairs as positives, yields additional benefits.

In this paper, we examine the effectiveness of contrastive learning for trial-robust EMG pattern recognition. We compare standard supervised contrastive learning with a trial-aware variant and analyze the learned representations through cross-trial evaluation. The main contributions are as follows:

- We systematically investigate supervised contrastive learning for addressing trial variability in EMG pattern recognition, demonstrating its effectiveness across datasets with diverse trial structures.
- We propose Cross-Trial-Contrast (CTC) loss, which explicitly incorporates trial information by treating only cross-trial same-class pairs as positives, encouraging trial-invariant representations.
- We quantitatively analyze the trial invariance of learned representations using cross-trial consistency and variance

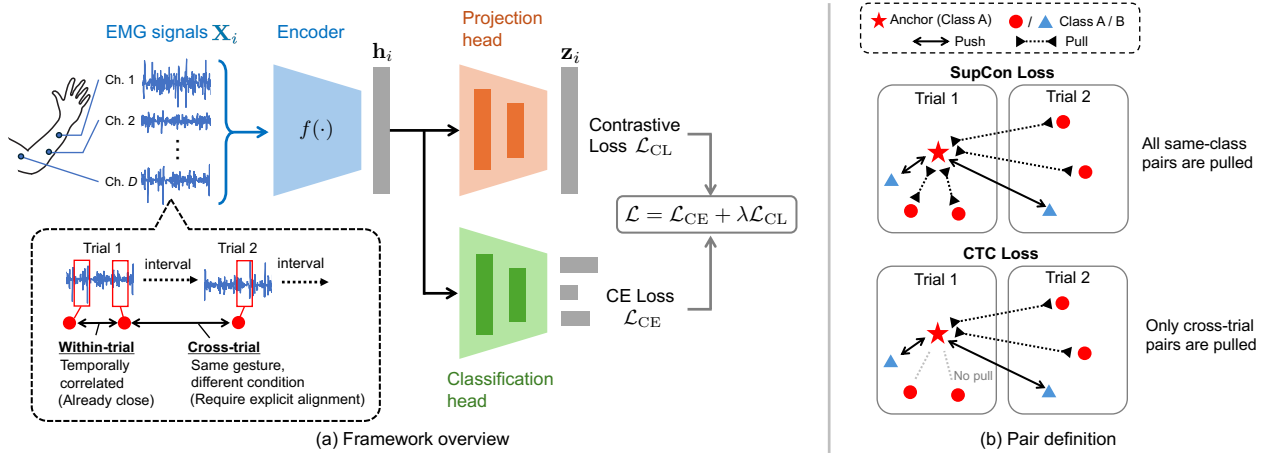


Fig. 1. Overview of the proposed framework. (a) The encoder processes multi-channel EMG signals and produces representations fed into two branches: a projection head for contrastive learning and a classification head for classification. (b) Comparison of positive pair definitions. SupCon treats all same-class pairs as positives regardless of trial membership. CTC loss restricts positives to cross-trial same-class pairs, explicitly encouraging trial-invariant representations.

ratio metrics.

II. METHODS

A. Problem Setting

Let $\mathbf{X}_i \in \mathbb{R}^{C \times T}$ denote the i -th windowed EMG signal, where C is the number of channels and T is the number of time samples per window, with gesture label $y_i \in \{1, \dots, K\}$ and trial identifier $r_i \in \{1, \dots, R\}$. In a typical EMG data collection protocol, subjects perform each of K gestures repeatedly across R trials, resulting in trial-specific variations.

Our goal is to learn an encoder $f: \mathbb{R}^{C \times T} \rightarrow \mathbb{R}^M$ that produces M -dimensional representations $\mathbf{h}_i = f(\mathbf{X}_i)$ which are trial-invariant yet class-discriminative. The learned representations should cluster samples by gesture class regardless of which trial they originate from. To this end, we employ a joint training framework that combines contrastive learning with classification, as shown in Fig. 1(a).

B. Supervised Contrastive Learning

Supervised contrastive learning [6] extends self-supervised contrastive methods by leveraging label information to define positive pairs. Given a mini-batch of N samples, the method learns representations by pulling together samples of the same class while pushing apart samples of different classes.

For each anchor sample i , the positive set is defined as all other samples sharing the same class label:

$$\mathcal{P}(i) = \{j : y_j = y_i, j \neq i\}. \quad (1)$$

The supervised contrastive loss (SupCon) is given by

$$\mathcal{L}_{\text{SupCon}} = - \sum_{i=1}^N \frac{1}{|\mathcal{P}(i)|} \sum_{p \in \mathcal{P}(i)} \log \frac{\exp(\mathbf{z}_i \cdot \mathbf{z}_p / \tau)}{\sum_{k \neq i} \exp(\mathbf{z}_i \cdot \mathbf{z}_k / \tau)}, \quad (2)$$

where $\mathbf{z}_i = g(\mathbf{h}_i) / \|g(\mathbf{h}_i)\|$ is the ℓ_2 -normalized output of a projection head g , and $\tau > 0$ is a temperature parameter.

C. Cross-Trial-Contrast Loss

When applying SupCon to EMG data collected across multiple trials, same-class samples from the same trial are expected to exhibit high similarity. This arises from temporal autocorrelation within continuous recordings and consistent electrode-skin interface conditions. We hypothesize that these same-trial same-class pairs may constitute “easy positives” that can be matched through trial-specific signal characteristics (e.g., baseline amplitude, noise patterns, and electrode impedance) rather than gesture-discriminative features. Such easy positives may dominate the contrastive objective, reducing the incentive for the encoder to learn features that generalize across trials.

To address this issue, we propose Cross-Trial-Contrast (CTC) loss, which explicitly incorporates trial information into the positive set definition. Specifically, we redefine the positive set to include only samples that share the same class label *and* originate from a different trial:

$$\mathcal{P}_{\text{CTC}}(i) = \{j : y_j = y_i \wedge r_j \neq r_i, j \neq i\}. \quad (3)$$

By excluding same-trial pairs from the positive set, CTC loss forces the encoder to identify features that are consistent across trials for each gesture class. The loss function is given by

$$\mathcal{L}_{\text{CTC}} = - \sum_{i=1}^N \frac{1}{|\mathcal{P}_{\text{CTC}}(i)|} \sum_{p \in \mathcal{P}_{\text{CTC}}(i)} \log \frac{\exp(\mathbf{z}_i \cdot \mathbf{z}_p / \tau)}{\sum_{k \neq i} \exp(\mathbf{z}_i \cdot \mathbf{z}_k / \tau)}, \quad (4)$$

where the denominator follows the standard InfoNCE formulation, summing over all non-self samples regardless of their class or trial membership. Fig. 1(b) illustrates the difference between the positive pair definitions in SupCon and CTC loss.

D. Training Objective

We adopt a joint training strategy that combines cross-entropy (CE) loss for classification with a contrastive loss for representation learning. The total training objective is given by

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda \mathcal{L}_{\text{CL}}, \quad (5)$$

where $\mathcal{L}_{CE}$ is the cross-entropy loss computed from $\mathbf{h}_i$ via a classification head, $\mathcal{L}_{CL} \in \{\mathcal{L}_{SupCon}, \mathcal{L}_{CTC}\}$ is the contrastive loss computed from the projected embeddings $\mathbf{z}_i$, and $\lambda > 0$ controls the relative weight.

III. EXPERIMENTAL SETUP

A. Datasets and Evaluation Protocol

We evaluate the proposed method on three publicly available EMG datasets representing diverse trial structures and recording conditions: Khushaba 2013 [9] (8 subjects, 14 gestures, 4 trials, 8 channels sampled at 4 kHz), NinaPro DB4 [10] (10 subjects, 40 gestures from Exercises B and C, 6 trials, 12 channels, 2 kHz), and Kanoga 2020 [11] (5 subjects, 8 gestures, 4 trials/day over 30 days, totaling 120 trials with daily electrode repositioning, 8 channels, 200 Hz). All signals are segmented into 200 ms windows with 100 ms overlap. Each EMG channel is independently normalized by dividing by its maximum absolute value computed from the training data; the same normalization factors are applied to both training and test data.

We adopt a cross-trial evaluation protocol where later trials are held out for testing: for Khushaba, trials 1–3 for training and trial 4 for testing; for NinaPro DB4, trials 1–4 for training and trials 5–6 for testing; for Kanoga, the first 5 days (trials 1–20) for training and the remaining 25 days (trials 21–120) for testing. This setup simulates practical deployment scenarios where models must generalize to future sessions. All experiments are conducted within-subject, with final metrics averaged across subjects.

B. Compared Methods

We compare the following methods:

- **CE**: Standard cross-entropy training (baseline), where the encoder is trained solely to minimize classification loss.
- **DANN**: Domain-adversarial training [4] treating each trial as a domain. A domain discriminator predicts trial identity, and gradient reversal encourages trial-invariant features. Unlike standard DANN, we do not use unlabeled target data, as future trial data is unavailable at training time in practical deployment; the discriminator is trained to distinguish among source trials only.
- **SupCon + CE**: Joint training with supervised contrastive loss (Eq. 2) and cross-entropy loss. All same-class pairs are treated as positives regardless of trial membership.
- **CTC + CE**: Joint training with the proposed Cross-Trial-Contrast loss (Eq. 4) and cross-entropy loss. Only cross-trial same-class pairs are treated as positives.

C. Evaluation Metrics

We evaluate the models using the following metrics. **Classification accuracy** measures the proportion of correctly classified test samples. **Cross-trial class consistency** quantifies the average cosine similarity between same-class, different-trial sample pairs in the encoder feature space:

$$\bar{s}_{\text{cross}} = \frac{1}{K} \sum_{c=1}^K \mathbb{E}_{i,j: y_i=y_j=c, r_i \neq r_j} [\hat{\mathbf{h}}_i \cdot \hat{\mathbf{h}}_j], \quad (6)$$

where $\hat{\mathbf{h}}_i = \mathbf{h}_i / \|\mathbf{h}_i\|$. Higher values indicate better alignment of same-class representations across different trials. **Variance ratio** measures class separability as the ratio of inter-class to intra-class variance in the encoder feature space $\{\mathbf{h}_i\}$:

$$\text{VR} = \frac{\text{tr}(\mathbf{S}_B)}{\text{tr}(\mathbf{S}_W)}, \quad (7)$$

where $\mathbf{S}_B = \sum_{c=1}^K n_c (\boldsymbol{\mu}_c - \boldsymbol{\mu})(\boldsymbol{\mu}_c - \boldsymbol{\mu})^\top$ and $\mathbf{S}_W = \sum_{c=1}^K \sum_{i: y_i=c} (\mathbf{h}_i - \boldsymbol{\mu}_c)(\mathbf{h}_i - \boldsymbol{\mu}_c)^\top$ are the between-class and within-class scatter matrices, with n_c , $\boldsymbol{\mu}_c$, and $\boldsymbol{\mu}$ denoting class- c sample count, class centroid, and global centroid, respectively. Higher values indicate greater class separability.

We assess statistical significance using a stratified permutation test with each dataset as a stratum. Per-subject paired differences are sign-flipped within each stratum, and the test statistic is computed as the sum of per-dataset means. We use 100,000 iterations with a two-sided test (significance level: 0.05), applying Holm correction within each metric for the six pairwise comparisons among methods.

D. Implementation Details

We use a 4-layer 1D convolutional neural network (CNN) encoder with channel dimensions [16, 32, 64, 128], batch normalization after each convolutional layer, and adaptive average pooling, producing 128-dimensional features. The projection head maps features to 128-dimensional ℓ_2 -normalized embeddings, and the classification head maps to class logits; both are two-layer multilayer perceptrons (MLPs). For DANN, the domain discriminator uses the same MLP architecture, with the adversarial weight scheduled as in [4]. For both contrastive methods, we set $\tau = 0.2$ and $\lambda = 0.5$ with a batch size of 128. For NinaPro DB4, we reduce λ to 0.1 as the larger number of classes ($K = 40$) yields fewer positive pairs per mini-batch, making the contrastive loss noisier. Models are trained with AdamW (learning rate 10^{-3}) for 100 epochs.

We apply stochastic data augmentation to all methods, generating two augmented views per sample. Each view is created by randomly applying Gaussian noise ($\sigma = 0.01$), temporal shift (up to 10% of window length), and amplitude scaling (factor in $[0.9, 1.1]$), each with probability 0.5.

IV. RESULTS AND ANALYSIS

A. Effectiveness of Contrastive Learning

Table I summarizes the cross-trial evaluation results across the three datasets. Both contrastive learning methods (SupCon + CE and CTC + CE) improve accuracy over the CE baseline on all three datasets. A stratified permutation test confirms significance for CTC + CE ($p = 0.028$), while SupCon + CE shows a consistent but statistically non-significant improvement ($p = 0.244$). These improvements across diverse data collection protocols suggest that contrastive learning enhances robustness to trial variability.

CTC + CE achieves the highest accuracy on Khushaba (+1.02 pp) and NinaPro DB4 (+1.67 pp), while SupCon + CE shows marginally higher accuracy on Kanoga. In contrast, DANN shows inconsistent results ($p = 0.274$ vs. CE):

TABLE I
CROSS-TRIAL EVALUATION RESULTS ACROSS THREE EMG DATASETS

Dataset	Method	Classification			Representation	
		Accuracy (%)	Δ (pp)	#Wins	Trial Cons.	Variance Ratio
Khushaba 2013 [9]	CE	92.07 ± 8.03	—	—	0.959 ± 0.007	8.16 ± 1.13
	DANN	92.50 ± 7.26	+0.43	4/8	0.970 ± 0.003	8.00 ± 1.15
	SupCon + CE	92.42 ± 7.40	+0.35	3/8	0.964 ± 0.004	<u>9.07 ± 1.23</u>
	CTC + CE	93.09 ± 6.49	+1.02	5/8	<u>0.967 ± 0.006</u>	9.42 ± 1.54
NinaPro DB4 [10]	CE	51.46 ± 5.01	—	—	0.899 ± 0.015	0.86 ± 0.18
	DANN	50.24 ± 5.61	−1.22	4/10	0.925 ± 0.011	0.88 ± 0.18
	SupCon + CE	<u>52.27 ± 5.89</u>	+0.81	7/10	0.904 ± 0.014	<u>0.89 ± 0.18</u>
	CTC + CE	53.13 ± 6.51	+1.67	8/10	<u>0.905 ± 0.014</u>	0.90 ± 0.18
Kanoga 2020 [11]	CE	90.43 ± 6.21	—	—	0.911 ± 0.013	5.50 ± 1.44
	DANN	89.86 ± 5.74	−0.57	1/5	0.901 ± 0.017	4.49 ± 1.29
	SupCon + CE	90.69 ± 6.49	+0.26	4/5	<u>0.921 ± 0.014</u>	<u>6.23 ± 1.34</u>
	CTC + CE	<u>90.58 ± 6.35</u>	+0.15	3/5	0.924 ± 0.014	6.56 ± 1.81

Bold: best, underline: second best. Δ : difference from CE baseline. #Wins: number of subjects where the method outperforms CE. Values are mean $\pm$ std across subjects. Classification: test data. Representation: training data.

marginal improvement on Khushaba (+0.43 pp) but degradation on NinaPro DB4 (−1.22 pp) and Kanoga (−0.57 pp).

Subject-wise analysis reveals consistent improvements for contrastive methods: SupCon outperforms CE in 14/23 subjects (61%), and CTC in 16/23 subjects (70%), confirming that the benefits of contrastive learning are not driven by a few outlier subjects.

B. Effect of Trial-Aware Design

To investigate whether explicit trial information provides additional benefits, we compare the representation quality metrics in Table I and examine their relationship with classification accuracy. Both contrastive methods improve representation quality over the CE baseline, and CTC + CE consistently achieves the highest variance ratio across all three datasets, indicating superior class separability. The stratified permutation test confirms that CTC + CE significantly outperforms CE in both variance ratio ($p < 0.001$) and cross-trial consistency ($p < 0.001$), and also outperforms SupCon in cross-trial consistency ($p = 0.007$) with a marginally significant improvement in variance ratio ($p = 0.056$).

Interestingly, DANN achieves the highest cross-trial consistency on Khushaba (0.970) and NinaPro DB4 (0.925), yet this does not translate to accuracy gains. This indicates that maximizing trial-invariance alone is insufficient; the discriminative power of features must be preserved.

In contrast to DANN, the improvements in representation quality generally translate to higher classification accuracy. On Khushaba and NinaPro DB4, CTC + CE achieves the highest variance ratio and accuracy among all methods. Only on the Kanoga dataset does SupCon + CE achieve marginally higher accuracy (90.69% vs. 90.58%) despite CTC + CE showing better representation metrics. We further analyze this relationship in Section IV-D.

C. Representation Analysis

Figure 2 visualizes the learned embedding spaces using t -SNE. Both CE and CTC + CE produce well-separated class

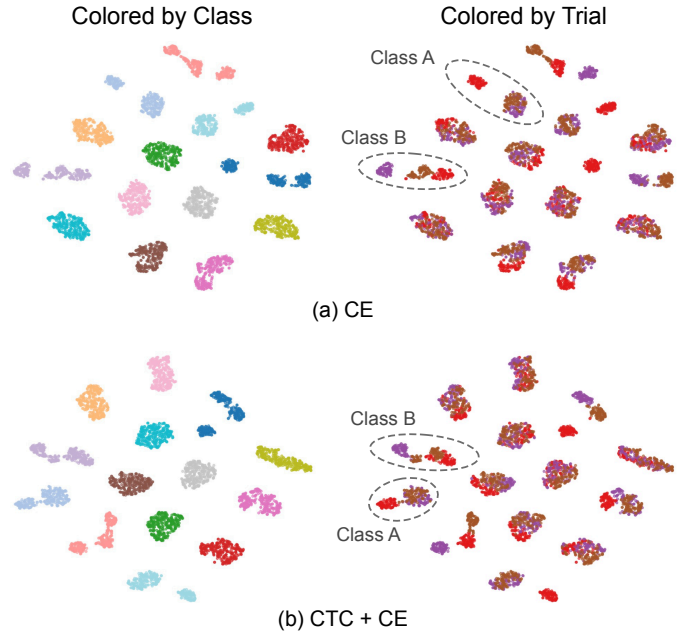


Fig. 2. t -SNE visualization of learned representations on Khushaba dataset (Subject 4, training data). Left: colored by gesture class. Right: colored by trial. (a) CE baseline. (b) CTC + CE. Dashed circles highlight two representative clusters (classes A and B).

clusters (left panels), confirming that the addition of contrastive training does not compromise class discriminability. In the trial-colored views (right panels), the CE baseline exhibits trial-dependent substructure within certain clusters: for example, classes A and B (dashed circles) contain visible sub-groupings by trial identity. In the corresponding clusters of CTC + CE, trial colors are more evenly mixed, indicating that the learned representations are less dependent on trial-specific variations. This visual pattern is consistent with the higher cross-trial consistency reported in Table I.

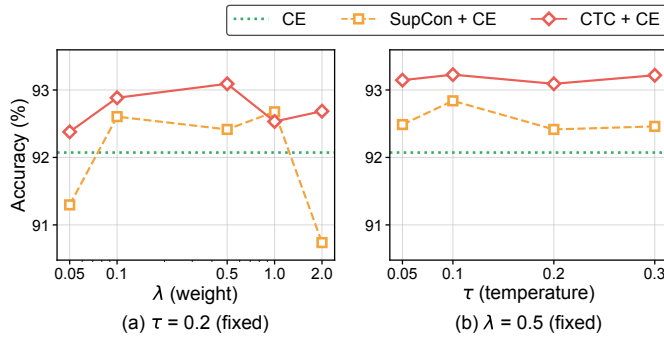


Fig. 3. Sensitivity analysis of hyperparameters on Khushaba dataset. (a) Effect of contrastive loss weight λ with fixed $\tau = 0.2$. (b) Effect of temperature τ with fixed $\lambda = 0.5$. Dotted line indicates CE baseline accuracy.

D. Discussion

Effect of dataset characteristics. The accuracy gain of CTC + CE over the CE baseline varies across datasets: +1.02pp on Khushaba and +1.67pp on NinaPro DB4, but only +0.15pp on Kanoga where SupCon achieves a comparable +0.26pp. We attribute this to the fact that Kanoga uses 20 training trials across 5 days, providing abundant cross-trial same-class pairs even without explicit cross-trial constraints; standard SupCon batches naturally contain diverse trial combinations, implicitly encouraging trial-invariant representations. In contrast, Khushaba (3 training trials) and NinaPro DB4 (4 training trials) offer fewer trial combinations per batch, making explicit cross-trial positive pair selection more beneficial. These results suggest that trial-aware contrastive learning is particularly effective when training trials are limited.

Comparison with domain adversarial training. DANN achieves high cross-trial consistency yet fails to consistently improve accuracy. Unlike the cross-subject or cross-sensor domain shifts for which adversarial training learns invariant features, trial-to-trial variation arises from subtle condition changes such as impedance drift and positional shifts; enforcing complete invariance to such fine-grained variation risks discarding class-discriminative information. Contrastive learning instead provides a softer constraint that encourages trial-invariant representations while preserving class separability, as reflected in the significantly higher variance ratios ($p < 0.001$ for CTC + CE vs. both CE and DANN).

Hyperparameter sensitivity. Figure 3 shows the sensitivity of contrastive methods to the loss weight λ and temperature τ on the Khushaba dataset. CTC + CE maintains accuracy above the CE baseline across all tested λ values (0.05–2.0), whereas SupCon + CE drops below the baseline at extreme values ($\lambda = 0.05$ and $\lambda = 2.0$), suggesting that CTC + CE is more robust to the choice of λ . For the temperature τ , both methods show stable behavior across the tested range (0.05–0.3), consistently maintaining performance above the baseline. These results suggest that while λ requires moderate tuning—particularly for SupCon— τ is relatively robust within typical ranges used in prior work [6], [12]. Similar trends were observed on the NinaPro DB4 and Kanoga datasets.

V. CONCLUSION

This paper presented a systematic investigation of supervised contrastive learning for trial-robust EMG pattern recognition. Our experiments demonstrated that contrastive learning consistently improves cross-trial generalization over standard cross-entropy training. The proposed CTC loss achieved the highest variance ratio and accuracy on two of three datasets (Khushaba: +1.02pp, NinaPro DB4: +1.67pp), demonstrating that explicit trial-aware positive pair selection enhances both representation learning and classification.

From a practical perspective, contrastive learning offers a promising approach for EMG systems requiring cross-trial robustness, with CTC providing additional benefits when the number of training trials is limited and trial labels are available.

Future work will investigate the relationship between trial variability characteristics and optimal contrastive strategies, and extend this framework to cross-subject generalization scenarios with larger-scale datasets.

REFERENCES

- [1] D. Farina, F. Negro, S. Muceli, and R. M. Enoka, “Principles of motor unit physiology evolve with advances in technology,” *Physiology*, vol. 31, no. 2, pp. 83–94, 2016.
- [2] I. Kyranou, S. Vijayakumar, and M. S. Erden, “Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses,” *Frontiers in Neurorobotics*, vol. 12, p. 58, 2018.
- [3] 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, pp. 643–659, 2011.
- [4] Y. Ganin and V. Lempitsky, “Unsupervised domain adaptation by backpropagation,” in *Proceedings of the 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 37. PMLR, 2015, pp. 1180–1189.
- [5] U. Côté-Allard, G. Gagnon-Turcotte, A. Phinyomark, K. Glette, E. Scheme, F. Laviolette, and B. Gosselin, “A transferable adaptive domain adversarial neural network for virtual reality augmented EMG-based gesture recognition,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 546–555, 2021.
- [6] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, “Supervised contrastive learning,” in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 18 661–18 673.
- [7] X. Shen, X. Liu, X. Hu, D. Zhang, and S. Song, “Contrastive learning of subject-invariant EEG representations for cross-subject emotion recognition,” *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 2496–2511, 2023.
- [8] J. Yang, D. Cha, D.-G. Lee, and S. Ahn, “STCNet: Spatio-temporal cross network with subject-aware contrastive learning for hand gesture recognition in surface EMG,” *Computers in Biology and Medicine*, vol. 185, p. 109525, Feb. 2025.
- [9] R. N. Khushaba, S. Kodagoda, D. Liu, and G. Dissanayake, “Muscle computer interfaces for driver distraction reduction,” *Computer Methods and Programs in Biomedicine*, vol. 110, no. 2, pp. 137–149, 2013.
- [10] S. Pizzolato, L. Tagliapietra, M. Cognolato, M. Reggiani, H. Müller, and M. Atzori, “Comparison of six electromyography acquisition setups on hand movement classification tasks,” *PLoS ONE*, vol. 12, no. 10, p. e0186132, 2017.
- [11] S. Kanoga, A. Kanemura, and H. Asoh, “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.
- [12] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A simple framework for contrastive learning of visual representations,” in *Proceedings of the 37th International Conference on Machine Learning*, 2020, pp. 1597–1607.