



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/240959/>

Version: Preprint

Preprint:

Touko, N., Ellis, M.O.A., Capone, C. et al. (Submitted: 2026) Lightweight test-time adaptation for EMG-based gesture recognition. [Preprint - arXiv] (Submitted)

<https://doi.org/10.48550/arXiv.2601.04181>

© 2026 The Author(s). For reuse permissions, please contact the Author(s).

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Lightweight Test-Time Adaptation for EMG-Based Gesture Recognition

Nia Touko*, Matthew O A Ellis*, Cristiano Capone†, Alessio Burrello‡, Elisa Donati§¶, Luca Manneschi*¶

*School of Computer Science, University of Sheffield, Sheffield, S10 2TN, United Kingdom †Natl. Center for Radiation Protection and Computational Physics, Istituto Superiore di Sanità, 00161 Rome, Italy ‡Politecnico di Torino, Italy §Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland ¶Joint senior authorship

Abstract—Reliable long-term decoding of surface electromyography (EMG) is hindered by signal drift caused by electrode shifts, muscle fatigue, and posture changes. While state-of-the-art models achieve high intra-session accuracy, their performance often degrades sharply. Existing solutions typically demand large datasets or high-compute pipelines that are impractical for energy-efficient wearables.

We propose a lightweight framework for Test-Time Adaptation (TTA) using a Temporal Convolutional Network (TCN) backbone. We introduce three deployment-ready strategies: (i) causal adaptive batch normalization for real-time statistical alignment; (ii) a Gaussian Mixture Model (GMM) alignment with experience replay to prevent forgetting; and (iii) meta-learning for rapid, few-shot calibration.

Evaluated on the NinaPro DB6 multi-session dataset, our framework significantly bridges the inter-session accuracy gap with minimal overhead. Our results show that experience-replay updates yield superior stability under limited data, while meta-learning achieves competitive performance in one- and two-shot regimes using only a fraction of the data required by current benchmarks. This work establishes a path toward robust, "plug-and-play" myoelectric control for long-term prosthetic use.

Index Terms—Human-Machine Interaction, temporal convolutional networks, meta-learning, continual learning, adaptation, long-term stability

I. INTRODUCTION

Reliable long-term decoding of neuromuscular signals is the primary bottleneck in developing practical myoelectric interfaces. While surface electromyography (EMG) classifiers achieve high accuracy in controlled settings [1], their performance degrades significantly across recording sessions [2]. This decline is driven by the inherent non-stationarity of EMG signals, where electrode displacement, muscle fatigue, and changes in skin impedance cause severe distribution shifts, often referred to as signal drift [3, 4]. In real-world applications, where daily re-donning of wearable devices is required, models trained on static data often fail to generalize to the target domain.

Early solutions to EMG non-stationarity utilized multi-session training or data augmentation to model variability like electrode displacement [2, 4]. However, the data intensity and calibration frequency of these methods are incompatible with resource-constrained wearables. While Transfer Learning (TL) and Domain Adaptation (DA) improve cross-session robustness by updating parameter subsets [5, 6] or through statistical

alignment, such as Canonical Correlation Analysis (CCA) [7], they typically assume *a priori* access to target data.

Several paradigms offer more flexible alternatives. Test-Time Training (TTT) [8] enables online adaptation at deployment by updating model parameters using only incoming (typically unlabeled) data. Depending on the objective, this can range from computationally more demanding self-supervised [9, 10] training to lightweight variants that adapt only normalization statistics, such as Adaptive Batch Normalization (AdaBN) [11, 12]. To handle non-stationary data streams while mitigating catastrophic forgetting, Continual Learning (CL) proposes a range of strategies, including regularization-based updates and replay-based methods. In particular, replay mechanisms such as Latent Replay (LR) [13] or Dark Experience Replay (DER) [14] can maintain performance on prior distributions within tiny memory buffers. Complementarily, meta-learning optimizes for rapid, few-shot generalization during brief calibration phases. Despite these theoretical frameworks, their application to noisy, high-dimensional EMG signals remains largely underexplored.

Real-world EMG applications require robust performance across unseen recording sessions. However, the field often relies on intra-session validation, which artificially inflates performance metrics and obscures the impact of temporal signal drift. This study systematically evaluates intra- and inter-session generalization using a TCN decoder [15] tailored for physiological time-series. Benchmarking on the NinaPro DB6 dataset [16], comprising multi-day recordings of seven grasping gestures, reveals that while the baseline achieves state-of-the-art intra-session accuracy, performance significantly degrades in inter-session scenarios. This discrepancy necessitates the development of lightweight, test-time adaptation strategies for field deployment.

To bridge this performance gap, we evaluate three complementary test-time adaptation schemes. The first, causal Batch Normalization (BN) adaptation, is an unsupervised method that updates normalization statistics online to enable rapid alignment during inference with negligible computational overhead. The second, replay-regularized statistical alignment, updates a low-rank parameter subset while utilizing a replay buffer to preserve source-domain representations and mitigate catastrophic forgetting. Finally, we employ meta-learning as an optimization strategy to leverage cross-session variability for high-efficiency, few-shot calibration [17].

These strategies integrate perspectives from TTT [18], CL

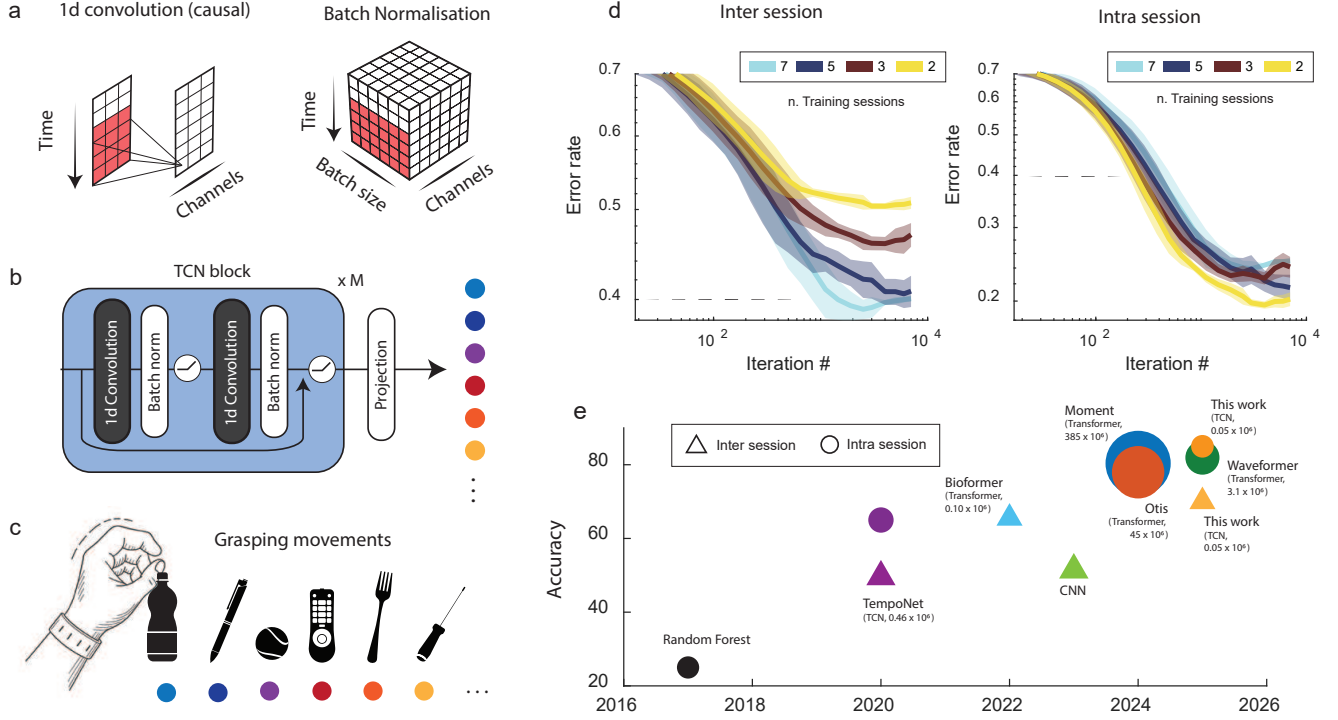


Fig. 1: Overview of the proposed framework for inter-session adaptation in surface EMG decoding. (a) 1D causal convolution and batch normalization applied over time, channels, and batch dimensions. (b) Temporal Convolutional Network (TCN) architecture with M residual causal-convolution blocks and a linear projection layer. (c) NinaPro DB6 setup: seven grasping gestures recorded using 14 double-differential EMG electrodes around the forearm. (d) Inter- vs. intra-session performance across training sessions $k \in \{2, 3, 5, 7\}$, showing persistent cross-day domain shift. (e) Benchmark on NinaPro DB6 comparing accuracy and model size, highlighting the competitiveness of the lightweight TCN (0.05M parameters).

[19], and meta-learning [20] into a single EMG decoding protocol. Our findings demonstrate that minimal test-time intervention substantially improves inter-session robustness, narrowing the accuracy gap and advancing the feasibility of energy-efficient, long-term myoelectric control for wearable applications.

II. MATERIAL AND METHODS

The proposed framework for inter-session adaptation in surface EMG decoding is illustrated in Fig. 1. The study focuses on classifying seven grasping movements from the NinaPro DB6 dataset (Fig. 1c) using a TCN architecture composed of stacked causal convolutional blocks with batch normalization and residual connections (Fig. 1a–b). The model was trained and evaluated under both intra-session and inter-session conditions to quantify generalization across recording days (Fig. 1d). The network, both inter- and intra-, was benchmarked with previous literature (Fig. 1e). This experimental setup provides a controlled benchmark for assessing adaptation mechanisms that aim to improve long-term robustness in EMG-based gesture decoding.

A. Temporal Convolutional Network Architecture

The proposed decoder is based on a TCN, which processes temporal sequences through one-dimensional causal convo-

lutions, effectively capturing correlations between adjacent time samples. This approach is particularly suited for EMG signals, imposing temporal translational equivariance across the different network operations.

As shown in Fig. 1(a–b), the network consists of multiple residual blocks, each containing two causal convolutional layers followed by normalization and non-linear activation. The use of dilation allows each subsequent block to cover a wider temporal context, progressively expanding the receptive field while maintaining computational efficiency. We employ batch normalization [21] within each network block to stabilize optimization and accelerate convergence. As discussed in Section II-D1, we then introduce a causal variant that adaptively updates normalization statistics online for test-time adaptation. The outputs of the convolutional blocks are projected point-wise through a linear readout layer, generating a response for every input data point. This residual and dilated design preserves the temporal resolution of the input while enabling hierarchical extraction of temporal dependencies.

As illustrated in Fig. 1d, the TCN architecture achieves stable intra-session accuracy, confirming its ability to model short-term temporal structure effectively. However, the same figure highlights a substantial drop in inter-session performance due to cross-day variability and electrode repositioning. This discrepancy motivates the adaptation mechanisms introduced

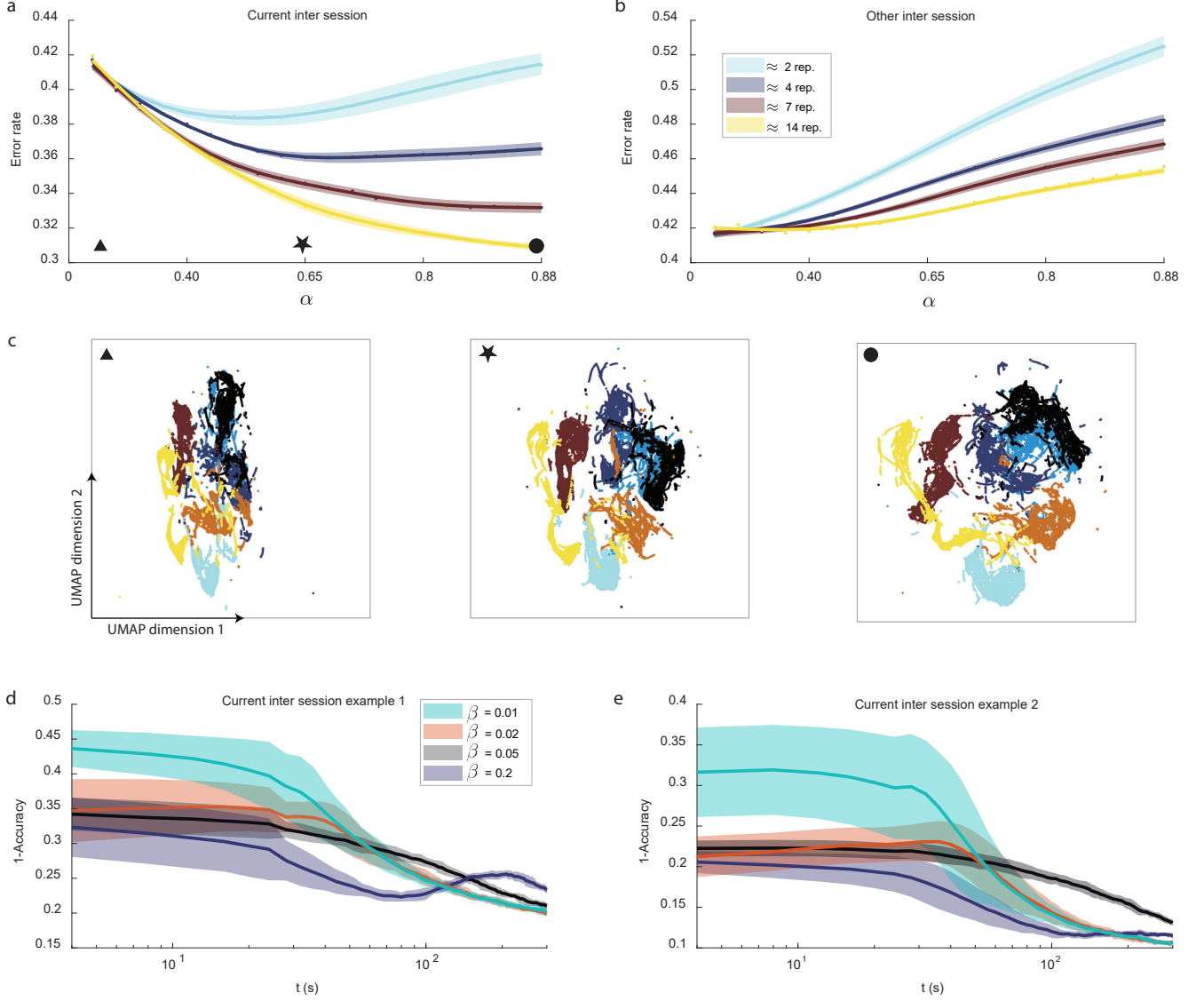


Fig. 2: Performance of the causal variant of batch normalization. (a,b) Batch-normalization adaptation: evolution of error rate on the current session (a) and on all other sessions (b) as a function of the mixing coefficient between old and new statistics. Results are reported for 2, 4, 7, and 14 gesture repetitions used for adaptation. (c) illustrates the shift of the UMAP projections across the adaptation strengths. (d,e) Online adaptation performance: evolution of the error rate over time for two representative subjects with different colors representing different adaptation speeds β .

later in the work, which aim to preserve the strengths of the TCN while improving its robustness across recording sessions.

B. Dataset

Experiments were conducted using the NinaPro DB6 dataset [16], created to investigate the consistency of surface EMG-based hand grasp recognition across multiple recording sessions. The experimental setup is illustrated in Fig. 1c. The dataset comprises data from ten healthy participants who performed seven distinct grasp types. Each grasp was repeated twelve times per session, with two sessions recorded per day over a period of five consecutive days, yielding ten sessions for each subject. EMG activity was measured using fourteen Delsys Trigno double-differential wireless electrodes distributed

uniformly around the forearm. The upper eight electrodes were positioned close to the radio-humeral joint, and the remaining six were placed farther down the forearm. Signals were sampled at 2 kHz. Each trial consisted of a four-second grasp followed by a four-second rest period. The acquisition protocol enables systematic analysis of inter-session variability and temporal adaptation in EMG-based gesture decoding.

Fig. 1e highlights the trade-off between accuracy and model size in NinaPro DB6 gesture recognition. Classical methods such as Random Forests [16] are extremely lightweight but exhibit large inter-session drops, whereas TempoNet [22] (<100k parameters) improves both intra- and inter-session performance with minimal computational cost. Larger deep models, including uncertainty-aware Convolutional Neural

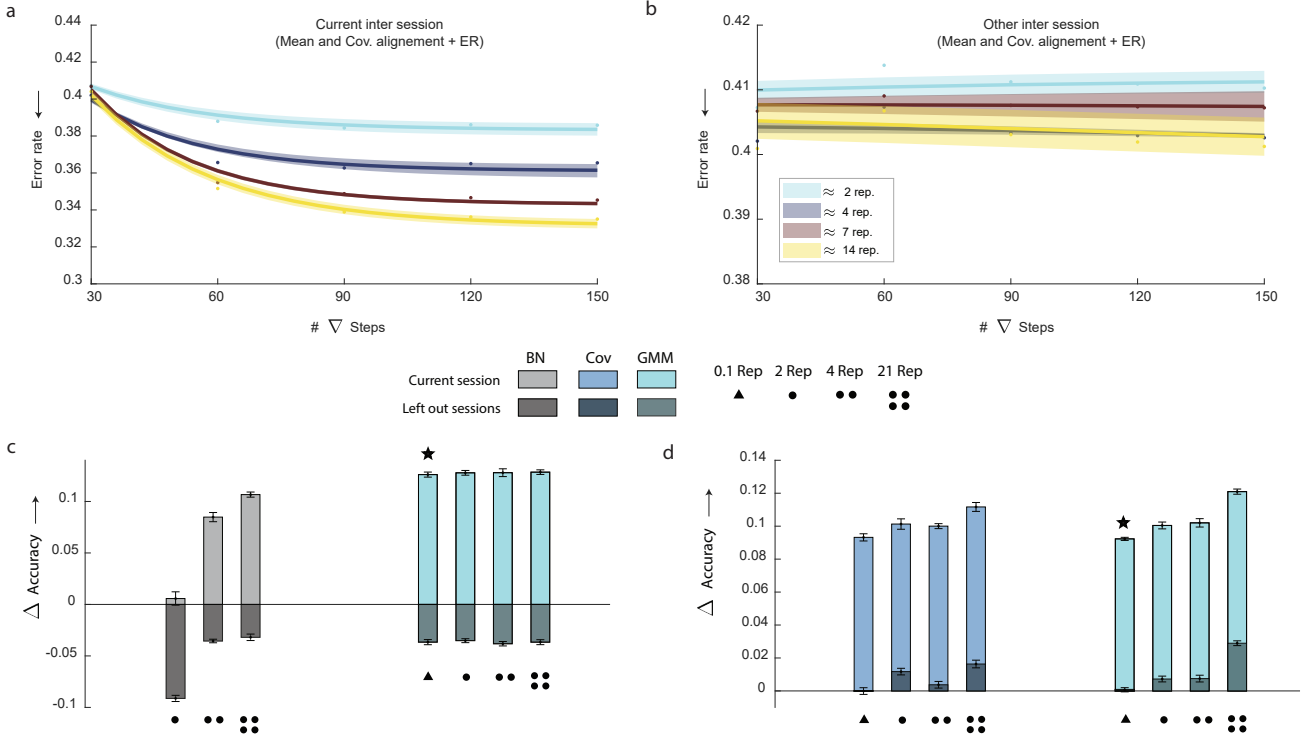


Fig. 3: Performance of the statistical alignment strategy across sessions. (a,b) GMM statistical loss with DER alignment scheme’s evolution of error rate on the current session (a) and on the other sessions (b) with respect to how big a buffer is required. (c) shows the performance of the batch-normalization adaptation and GMM adaptation without a DER component, while (d) compares the performance of the Covariance and GMM approaches with a DER component.

Networks (CNNs) [23], lightweight transformers [24, 25], and high-capacity architectures like OTIS [26] and MOMENT [27], achieve strong intra-session accuracy but require hundreds of thousands to millions of parameters and remain sensitive to cross-day variability. In contrast, the proposed TCN ($\approx 50k$ parameters) offers a superior accuracy–complexity balance, matching larger models intra-session while providing competitive or improved inter-session robustness at substantially lower cost.

C. Validation Methodology

The literature on the NinaPro DB6 dataset uses vastly different pre-processing methodologies to measure model performance and efficiency, greatly complicating direct comparison. To strictly and fairly evaluate each model and simulate a realistic deployment, we adopted a causal testing methodology to evaluate our models. The ten sessions available for each subject in NinaPro DB6 were partitioned strictly by time. The first five sessions constitute the Source Domain (D^{src}); specifically, 70% of this data was used for training, while the remaining 30% was reserved for validating the intra-session performance. The remaining five sessions form the Target Domain (D^{tgt}) to test inter-session generalization. This split ensures that future sessions are never seen during training, mimicking realistic deployment conditions.

For test-time adaptation, we define two strategies for constructing the adaptation buffer from the incoming target stream:

- **Balanced Sequences:** A control setting where the buffer contains an equal number of repetitions for each gesture class.
- **Unbalanced Sequences:** A realistic "streaming" setting where the buffer is populated by a random sequence of gestures. This mimics the unpredictable nature of daily usage, where the data gathered would be from a block of unknown data from standard use.

Notably, the literature often omits transient periods from the training and evaluation of the model, which improves model performance but results in a model that cannot be implemented for effective prosthetic control. Hence, this work opts to include all of the data (including the transient and rest periods) within the model and test-time buffers to better simulate real-world performance. Additionally, to prevent metric inflation due to the prevalence and relative ease of classifying of the "Rest" state, we report the active gesture accuracy, excluding the rest class, to calculate solely on the active movements. Furthermore, to emulate the post-processing smoothing typically seen in applied prosthetic control, we report Majority Voted Accuracy, computed over a sliding window of the predictions.

D. Adaptation Strategies

We consider session-based EMG data, where each session s is a sequence $\mathcal{D}_s = \{(x_t^{(s)}, y_t^{(s)})\}_{t=1}^{T_s}$, with $x_t^{(s)} \in \mathbb{R}^d$ and optional labels $y_t^{(s)}$. Sessions are divided into a *source* set $\mathcal{D}^{src}$ for training and a *target* set $\mathcal{D}^{tgt}$ for evaluation.

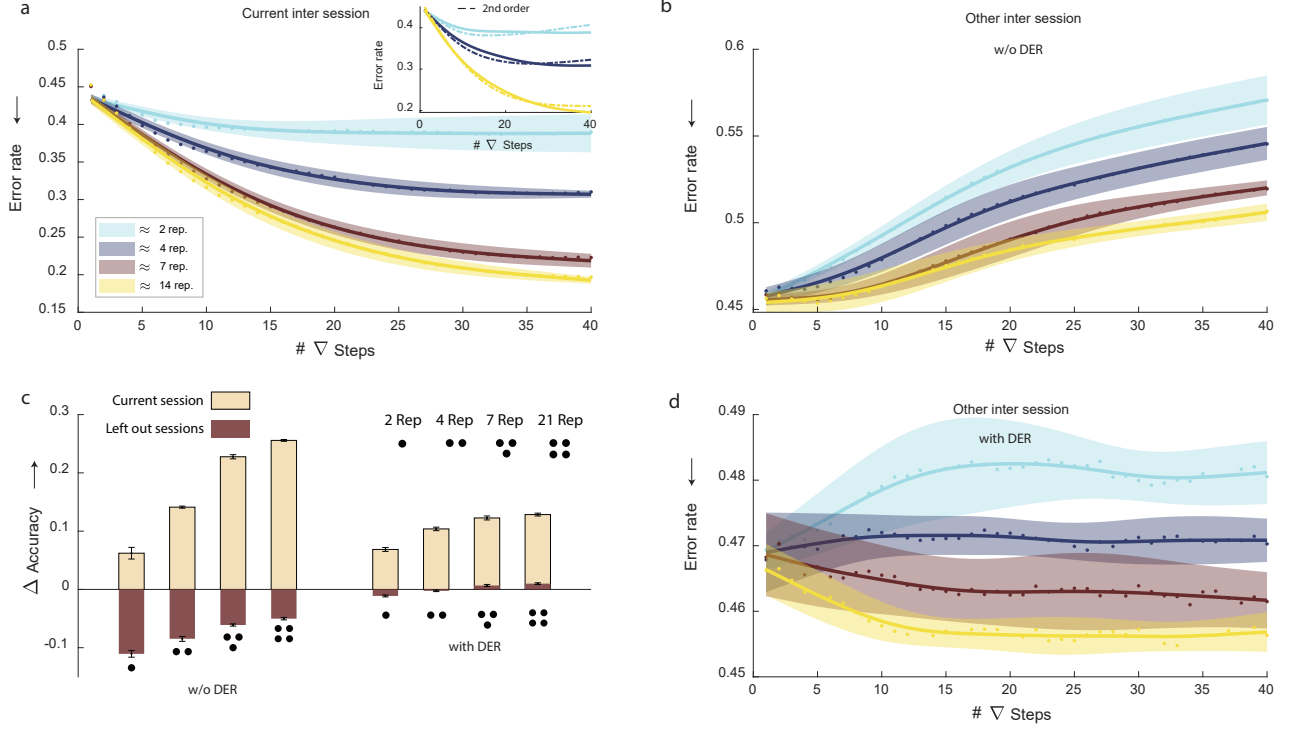


Fig. 4: Performance of Meta-learning across sessions. (a,b) illustrate the evolution of the error rate of the current session (a) and other sessions (b) across adaptation steps (x-axis) without DER and with respect to buffer size. The same data points are used for the different steps of the adaptation, as in few-shot learning settings. A comparative bar plot is made in (c) to display the effect of buffer size and DER on the final accuracy. Finally, (d) showcases the left-out session error rate when using the DER.

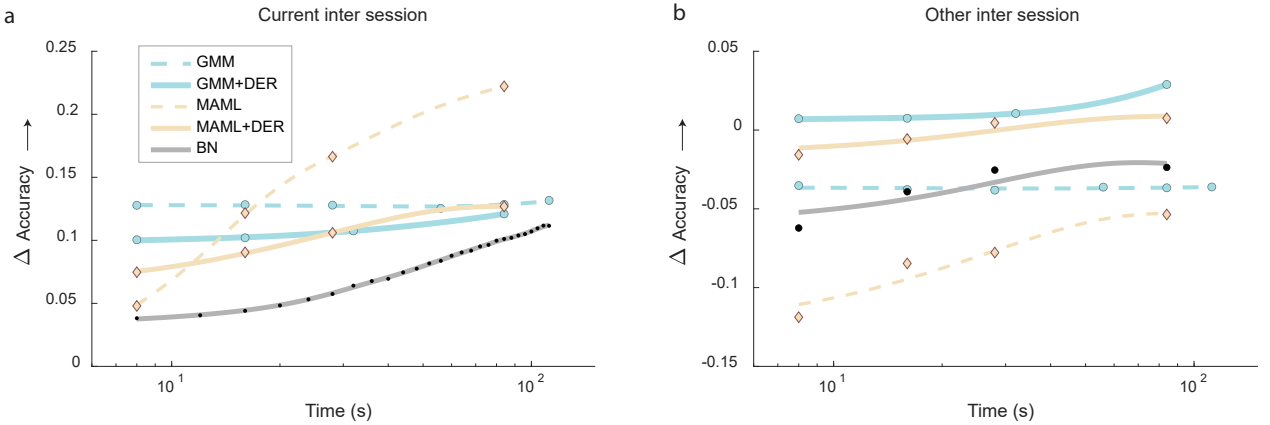


Fig. 5: Relative accuracy improvement of each test-time adaptation approach over the baseline with respect to the data buffer size, measured in seconds.

At test-time, the model processes a target session $\mathcal{D}_\tau$ in a streaming fashion, having access only to the prefix $\mathcal{D}_\tau^{\leq t}$ containing all samples observed up to time t . This setting naturally motivates TTT, where the model adapts as new data arrive.

We study three levels of test-time data availability:

- 1) **Unsupervised TTT:** Only the unlabeled prefix $\mathcal{D}_\tau^{\leq t, \text{unlab}}$ is available, and no source data can be accessed at test-time.
- 2) **Resource-aware adaptation:** The model additionally

uses a small exemplar buffer $\mathcal{D}_{\text{er}} \subset \mathcal{D}^{\text{src}}$ under strict memory constraints.

- 3) **Few-shot learning:** The unlabeled prefix is augmented with a sparse set of labels $\mathcal{D}_\tau^{\leq t, \text{lab}}$, simulating minimal user-assisted calibration.

Relevant to settings (1)–(2), a key distinction from standard TTT benchmarks is that the target EMG arrives as a continuous and highly noisy stream, with limited data available for adaptation. In this regime, adaptation cannot assume consistently reliable gesture segmentation, making

it difficult to form semantically coherent mini-batches or to define label-preserving (or, in other words, physiologically plausible) augmentations. Consequently, contrastive self-supervised objectives (e.g., SimCLR-style learning) are often ill-suited: naïve positive/negative pair construction may mix heterogeneous activations, and common augmentations can distort physiologically meaningful structure. Moreover, individual EMG samples are weakly informative in isolation and require temporal context for stable updates. For these reasons, we focus on adaptation via statistical alignment between source and target feature distributions, which remains well defined under streaming, noisy, and low-data constraints. For few-shot learning (setting (3)) under supervised calibration phases, we instead rely on meta-learning, which explicitly optimizes the model for rapid adaptation from a small number of labeled examples.

1) Adaptive Batch Normalization: This section addresses the first regime in Section II-D, i.e., TTT, where only unlabeled target data are available. To enable adaptation under this constraint, we introduce an online, causal variant of BN that can be readily adopted for deep neural networks.

Classical BN computes per-channel statistics over a mini-batch of B sequences of length U . Denoting $z_{b,c,u}^{(l)}$ as the activation at layer l , channel c , sample b , and time u in batch τ , the batch mean, and variance are

$$\mu_c^{\text{batch}}(\tau) = \frac{1}{BU} \sum_{b=1}^B \sum_{u=1}^U z_{b,c,u}^{(l)}, \quad (1)$$

$$(\sigma_c^{\text{batch}}(\tau))^2 = \frac{1}{BU} \sum_{b=1}^B \sum_{u=1}^U (z_{b,c,u}^{(l)} - \mu_c^{\text{batch}}(\tau))^2. \quad (2)$$

These are used to normalize activations:

$$\hat{z}_{b,c,u}^{(l)} = \frac{z_{b,c,u}^{(l)} - \mu_c^{\text{batch}}}{\sqrt{(\sigma_c^{\text{batch}})^2 + \varepsilon}}, \quad y_{b,c,u}^{(l)} = \gamma_c^{(l)} \hat{z}_{b,c,u}^{(l)} + \beta_c^{(l)}, \quad (3)$$

while long-term population statistics are tracked via Exponential Moving Averages (EMAs).

In the streaming test-time setting, we process a single sequence ($B = 1$) and compute online statistics over the prefix $\{z_c(i)\}_{i=1}^t$. At time t :

$$\mu_c(t) = \frac{1}{t} \sum_{i=1}^t z_c(i), \quad \sigma_c^2(t) = \frac{1}{t} \sum_{i=1}^t (z_c(i) - \mu_c(t))^2, \quad (4)$$

which can be updated incrementally using Welford's method:

$$\mu_c(t) = \mu_c(t-1) + \frac{z_c(t) - \mu_c(t-1)}{t}, \quad (5)$$

$$S_c(t) = S_c(t-1) + (z_c(t) - \mu_c(t-1))(z_c(t) - \mu_c(t)), \quad (6)$$

$$\sigma_c^2(t) = \frac{S_c(t)}{t}. \quad (7)$$

To retain information from the source domain, we blend the source EMA statistics with the online test-time estimates using a mixing coefficient $\alpha \in [0, 1]$:

$$\mu_c^{\text{use}}(t) = (1 - \alpha) \mu_c^{\text{EMA}} + \alpha \mu_c(t), \quad (8)$$

$$(\sigma_c^{\text{use}}(t))^2 = (1 - \alpha) (\sigma_c^{\text{EMA}})^2 + \alpha \sigma_c^2(t). \quad (9)$$

In the fully online setting, α should reflect the reliability of the target estimates, which increases as more samples are observed. We therefore adopt an adaptive, causal schedule

$$\alpha(t) = \min\left(1, \frac{\beta}{N} t\right), \quad (10)$$

where t is the number of target samples seen so far, N denotes the (maximum) number of samples in the target session, and β is a scalar controlling the overall adaptation strength. This schedule gradually shifts the normalization from the source EMA statistics towards the target online estimates, while remaining causal and well-defined at every time step.

In our analysis, we study the effect of this adaptive causal BN by explicitly controlling t and α (either fixing α or using Eq. (10) to tie it to the observed prefix length). During deployment, the model updates $\alpha(t)$ online as new target samples arrive and uses the blended statistics $\{\mu_c^{\text{use}}(t), \sigma_c^{\text{use}}(t)\}$ in place of the static EMA statistics in Eq. (3).

2) Statistical Alignment: This section corresponds to the second regime in Section II-D, where a small exemplar buffer $\mathcal{D}_{\text{er}}$ is available at test-time. We perform resource-efficient online adaptation by updating only low-rank adaptation parameters through LoRA [28], while keeping the backbone weights frozen. Adaptation is driven by statistical alignment in latent feature space, and the exemplar buffer is used as a regularizer to stabilize updates and mitigate forgetting.

Let $z^{(l)} \in \mathbb{R}^{d_l}$ denote the latent representation at layer l . At test-time, we maintain an online buffer $Z_\tau = \{z_i\}_{i=1}^M$ of recent target features (extracted with the current model), from which the target statistics required by the chosen alignment objective are computed. We consider three alignment objectives:

Global moment matching. We summarize the source feature distribution at layer l with its mean and covariance,

$$\mu_{\text{src}} = \mathbb{E}_{\mathcal{D}^{\text{src}}} [z], \quad \Sigma_{\text{src}} = \mathbb{E}_{\mathcal{D}^{\text{src}}} [(z - \mu_{\text{src}})(z - \mu_{\text{src}})^T], \quad (11)$$

and compute empirical target moments (μ_τ, Σ_τ) from Z_τ . Adaptation is driven by

$$\mathcal{L}_{\text{align}}^{\text{mom}} = \lambda_\mu \|\mu_\tau - \mu_{\text{src}}\|_2^2 + \lambda_\Sigma \|\Sigma_\tau - \Sigma_{\text{src}}\|_F^2. \quad (12)$$

Class-conditional moment matching. We partition the source features into rest (R) and gesture (G) states, and estimate class-specific moments $(\mu_{\text{src},k}, \Sigma_{\text{src},k})$ for $k \in \{R, G\}$. At test-time, pseudo-labels assign each feature in Z_τ to $k \in \{R, G\}$, producing subsets $Z_{\tau,k}$ and corresponding empirical moments $(\mu_{\tau,k}, \Sigma_{\tau,k})$. We then align moments per state:

$$\mathcal{L}_{\text{align}}^{\text{cc}} = \sum_{k \in \{R, G\}} \lambda_\mu \|\mu_{\tau,k} - \mu_{\text{src},k}\|_2^2 + \lambda_\Sigma \|\Sigma_{\tau,k} - \Sigma_{\text{src},k}\|_F^2. \quad (13)$$

Distribution matching via Gaussian Mixture Model (GMM). Instead of restricting to low-order moments, we fit a GMM $\mathcal{G}_{\text{src}}$ to source features at layer l . During adaptation, we draw samples $\hat{Z}_{\mathcal{G}_{\text{src}}}$ from $\mathcal{G}_{\text{src}}$ and match their distribution to the target buffer Z_τ using the Sliced Wasserstein Discrepancy (SWD):

$$\mathcal{L}_{\text{align}}^{\text{gmm}} = \text{SWD}(\hat{Z}_{\mathcal{G}_{\text{src}}}, Z_\tau). \quad (14)$$

To mitigate catastrophic forgetting while updating θ_{LoRA} , we regularize adaptation using exemplars $(x_{\text{er}}, y_{\text{er}}) \in \mathcal{D}_{\text{er}}$ following DER. For each exemplar, we store the source model output $\hat{y}_{\text{er}}$ (e.g., logits) and penalize deviations under the adapted model, optionally combining this with supervised replay:

$$\mathcal{L}_{\text{er}} = \|\hat{y}(x_{\text{er}}) - \hat{y}_{\text{er}}\|_2^2 + \mathcal{L}_{\text{CE}}(\hat{p}(x_{\text{er}}), y_{\text{er}}), \quad (15)$$

where $\hat{y}(\cdot)$ denotes current model outputs, $\hat{p}(\cdot)$ the corresponding class probabilities, and $\mathcal{L}_{\text{CE}}$ is the cross-entropy loss.

To control the magnitude of test-time updates induced by statistical alignment, we restrict adaptation to Low-Rank Adaptation (LoRA) parameters inserted in the convolutional layers. Denoting by ϕ the collection of trainable LoRA weights, we optimize only

$$\mathcal{L}(\phi) = \alpha \mathcal{L}_{\text{align}} + \beta \mathcal{L}_{\text{er}}, \quad (16)$$

where $\mathcal{L}_{\text{align}} \in \{\mathcal{L}_{\text{align}}^{\text{mom}}, \mathcal{L}_{\text{align}}^{\text{cc}}, \mathcal{L}_{\text{align}}^{\text{gmm}}\}$ specifies the chosen alignment objective.

3) Meta-learning: This section corresponds to the third regime in Section II-D, where a small number of supervisory signals is available at test-time, i.e., few-shot settings. This regime quantifies the performance gains achievable through a short calibration phase, during which the user repeats a small set of gestures to personalize the decoder.

To exploit limited supervision efficiently, we adopt gradient-based meta-learning in the model-agnostic meta-learning (MAML) framework, tailored to sessioned EMG data. We treat each session as a task $\mathcal{T}_s$ and learn an initialization that can be adapted to a new session using only a few labeled calibration examples.

To reduce the adaptation dimensionality, we proceed similarly to the statistical alignment methods and adopt LoRA. We parameterize the model as $f(x; \theta, \phi)$, where θ are *frozen* backbone parameters and ϕ are the LoRA parameters. Task-specific adaptation operates only on ϕ , while θ remains fixed. Therefore, the meta-learned initialization consists of a shared LoRA starting point ϕ (together with the fixed θ learned on $\mathcal{D}^{\text{src}}$).

For each task/session s , we split labeled samples into a support (calibration) set $\mathcal{D}_s^{\text{sup}}$ and a query set $\mathcal{D}_s^{\text{qry}}$. Let $\mathcal{L}(\theta, \phi; \mathcal{D})$ denote the supervised loss (cross entropy). Starting from the meta-initialization ϕ , the inner-loop adaptation performs K gradient steps on the support set. For notational simplicity, we omit the explicit dependence of $\mathcal{L}$ on the support data $\mathcal{D}_s^{\text{sup}}$ (i.e., $\mathcal{L}(\theta, \phi; \mathcal{D}_s^{\text{sup}})$):

$$\phi_s^{(0)} = \phi, \quad (17)$$

$$\phi_s^{(k)} = \phi_s^{(k-1)} - \eta_i \nabla_{\phi_s^{(k-1)}} \mathcal{L}(\theta, \phi_s^{(k-1)}), \quad k = 1, \dots, K, \quad (18)$$

yielding adapted LoRA parameters $\phi'_s \equiv \phi_s^{(K)}$, where η_i is the inner-loop learning rate.

The outer-loop meta-objective then updates the shared initialization (θ, ϕ) such that the adapted parameters ϕ'_s generalize on the corresponding query sets:

$$\min_{\theta, \phi} \sum_{\mathcal{D}_s \in \mathcal{D}^{\text{src}}} \mathcal{L}(\theta, \phi'_s), \quad (19)$$

where ϕ'_s is obtained via Eq. (18). Using a meta-learning rate η_o , we perform the meta-update

$$(\theta, \phi) \leftarrow (\theta, \phi) - \eta_o \nabla_{(\theta, \phi)} \sum_{\mathcal{D}_s \in \mathcal{D}^{\text{src}}} \mathcal{L}(\theta, \phi'_s; \mathcal{D}_s^{\text{qry}}). \quad (20)$$

The difference between first- and second-order meta-learning lies in how the meta-gradient in Eq. (20) is computed. In second-order MAML, the gradient $\nabla_{(\theta, \phi)} \mathcal{L}(\theta, \phi'_s; \mathcal{D}_s^{\text{qry}})$ backpropagates through the K inner-loop updates in Eq. (18), and therefore includes second-order terms arising from differentiating through the inner-loop gradients (i.e., Hessian-vector products). In first-order MAML, these higher-order terms are neglected by treating ϕ'_s as a constant with respect to (θ, ϕ) when computing the meta-update. Equivalently, second-order MAML can be viewed as learning an initialization by optimizing through the unrolled K -step adaptation dynamics, analogous to backpropagation through time in a recurrent system.

To mitigate forgetting arising from rapid adaptation, we also evaluate a second variant that combines meta-learning with replay-based regularization. Specifically, we incorporate a DER-style term using a small exemplar memory analogously to the replay-regularized statistical alignment methods.

III. RESULTS

We first evaluated the baseline performance of the proposed TCN model on the NinaPro DB6 dataset to assess its ability to generalize across recording sessions. The network was trained to classify seven grasping gestures from multi-day EMG recordings, with ten sessions available for each subject. Two evaluation schemes were considered: intra-session, where training and testing data belong to the same session, and inter-session, where testing is performed on unseen sessions to evaluate cross-day generalization.

Fig. 1d reports the classification accuracy for both intra-session and inter-session evaluations based on the number of sessions used for training ($k \in 2, 3, 5, 7$). In the intra-session case, where training and testing samples belong to the same session, performance remains stable at approximately 80% across all training configurations, indicating that the network reliably captures temporal patterns under stationary conditions. In the inter-session case, where the model is evaluated on entirely new sessions, accuracy initially drops to less than 60%, reflecting the substantial domain shift caused by electrode repositioning and day-to-day physiological variability. As more sessions are included in training, inter-session performance progressively improves, demonstrating that exposure to multiple recording conditions enhances generalization to unseen data.

A. Adaptive Batch Normalization

We next evaluated the effectiveness of adaptive batch normalization using the inter-session setup described above (Fig. 2). Each session was treated as a target domain unseen during training, and adaptation was performed by blending the target and source batch-normalization moments as described in Section II-D1.

Figs. 2a,b report performance as we vary the blending coefficient α across four adaptation set sizes, corresponding to 2, 4, 7, and 14 gesture repetitions. For each target session, repetitions were selected by sampling an unbalanced sequence of gestures to mimic a realistic stream at test-time, yielding adaptation prefixes spanning approximately 8–112 s of data (depending on the number of repetitions). Performance is then computed on the remaining samples from the same session, highlighting the *transductive* benefit of adapting on an unlabeled prefix of the target stream. The left panel shows the evolution of the error rate on the current session as the model progressively replaces the source BN statistics with the target estimates. When very limited adaptation data are available (two or four repetitions), excessive weighting of the new statistics can degrade performance, since early estimates poorly represent the full gesture distribution. This sensitivity is amplified by the intrinsically *aggressive* nature of BN adaptation: normalization affects activations at every layer, and even small errors in the estimated moments propagate through the network by re-centering and re-scaling features (towards zero mean and unit variance). In the following, we show how to make this adaptation more robust under limited test-time data. With larger adaptation sets (seven or fourteen repetitions), the new statistics become more representative, allowing the model to recover to about 70% accuracy. The right panel reports the error rate on the *other* test sessions (i.e., the remaining sessions not currently used for adaptation), evaluated as the model adapts to the current target session. It reveals a slight increase in error on these held-out sessions as the model specializes to the current one, an expected side effect of session-specific adaptation.

Figs. 2c,d illustrate how class structure in the latent space evolves as adaptation progresses. We report the case with 14 repetitions and highlight three increasing adaptation strengths (triangle, star, and circle, respectively). Latent features are visualized with UMAP, showing progressively tighter clusters and improved class separability as adaptation strengthens.

To enable a consistent comparison across time steps and adaptation strengths, we compute a shared UMAP embedding and use it to project features extracted at different adaptation points (rather than fitting UMAP independently for each plot). We also regularize the visualization by discouraging large shifts of the latent representations across time, so that changes reflect genuine improvements in separability rather than arbitrary embedding drift. Finally, we note that perfect separation is not expected: the EMG stream transitions continuously between gestures, and boundary segments naturally exhibit overlapping activations.

Finally, Figs. 2d,e show online batch-normalization adaptation. Under our causal test-time formulation, we can track the error rate as the model adapts continuously and $\alpha(t)$ increases with the amount of target data observed (Eq. 10). Results are shown for two representative subjects; for each subject, we average the online error over ~ 100 test-time runs obtained by randomizing the gesture order within the available target session. While this shuffling may disrupt temporal dependencies, the model remains robust, and performance typically improves over time for all reported values of β .

As expected, β controls the adaptation timescale: larger

β yields faster but less stable updates, whereas smaller β produces slower yet smoother improvement. Consistent with this interpretation, the most aggressive setting (dark blue) exhibits noticeable oscillations in panel d, indicating that early statistics are noisy and can temporarily over-correct the normalization before converging as more target data accumulate. In contrast, the lighter blue trend demonstrates slower dynamics, continuing to adapt even after the streamed test-time data used for adaptation has ended.

B. Statistical Alignment

We now turn to statistical alignment methods, analyzing how adaptation granularity, from matching low-order moments in feature space to fitting a GMM to the full distribution, affects performance, and how a replay buffer mitigates forgetting. During preliminary exploration, we inserted LoRA modules either in the first four TCN blocks or only in the fourth block. We found that injecting a rank-4 LoRA at a single intermediate block is sufficient to achieve the best overall performance, while keeping the number of trainable parameters and test-time computational cost minimal. Accordingly, unless otherwise stated, we report results for this single-block LoRA configuration in the following.

Figs. 3a,b present metrics analogous to Figs. 2a,b, focusing on GMM and covariance-based feature alignment combined with the DER replay buffer. The trends indicate effective adaptation to the current session, even with a limited number of gesture repetitions. Compared to adaptive batch normalization, statistical alignment is less sensitive to early, noisy target estimates, since it matches feature statistics through a constrained subset of parameters (via LoRA) rather than updating normalization at every layer. Moreover, the replay buffer anchors the model to source exemplars during adaptation, improving stability and reducing over-specialization to the current session. This is visible in panel b, where performance on the remaining test sessions (i.e., those not used for adaptation) is largely preserved and, in some cases, improves as more repetitions are observed.

While some retention is expected from a strong regularizer, the fact that non-adapted sessions can also benefit suggests that the update is not purely session-specific. Instead, aligning latent statistics while replaying source exemplars appears to push the model toward feature representations that are more robust to local session-to-session variability, rather than merely fitting idiosyncrasies of the current target stream. More broadly, these results highlight the complementarity between test-time self-supervision (via unlabeled statistical alignment) and continual-learning mechanisms (via replay): together, they can incorporate useful information from the target stream into the model while maintaining, and in some cases also improving, generalization and robustness.

Fig. 3c summarizes, via a bar plot, the change in accuracy relative to the base model after adaptation *without* replay regularization, and contrasts it with adaptive batch normalization (which likewise does not use a buffer). This comparison isolates the effect of the alignment objective itself. Lighter colors report performance on the target session used for adaptation, whereas darker colors report performance on the remaining test sessions.

With limited adaptation data, batch normalization tends to adapt too aggressively: gains on the current session are modest, while performance on the remaining sessions drops noticeably, suggesting over-specialization to idiosyncratic statistics of the observed target prefix. In contrast, explicit statistical alignment through LoRA already provides consistent improvements on the adapted session under the same low-data regime. Nevertheless, all methods exhibit some degradation in the remaining sessions when replay is removed, highlighting the stabilizing role of exemplar-based regularization.

Fig. 3d reports results obtained by covariance alignment (blue) and GMM-based distribution matching (green), both combined with the replay buffer. Overall, the GMM approach achieves higher accuracy than the purely moment-based variant, consistent with its ability to capture finer-grained structure in the feature distribution beyond low-order statistics. Notably, performance on the remaining sessions also improves as more gesture repetitions are used for adaptation, suggesting that replay-regularized distribution matching can incorporate useful target information while preserving, and in some cases strengthening, generalization across sessions.

C. Meta-learning

Finally, we turn to meta-learning. Its purpose is twofold: to quantify the gains achievable through fast *supervised* adaptation, thereby mimicking a short user calibration phase; to provide a supervised reference point for interpreting the limits of the self-supervised test-time adaptation strategies studied above. We perform these experiments with the MAML [29] algorithm, using four inner-loop gradient steps during meta-training.

Figure 4 reports adaptation trends analogous to the previous sections. As expected, the availability of labels leads to substantially faster convergence and higher accuracy on the adapted session (Fig. 4a). The inset in Fig. 4a compares first-order (solid) and second-order (dashed) MAML variants: both benefit from few-shot calibration, while second-order MAML typically converges faster. This speedup of second-order MAML, however, can come with increased sensitivity to the number of adaptation steps, consistent with a higher tendency to overfit, an effect that can be attributed to the fact that second-order gradients effectively shape the update dynamics, making the optimization process itself part of what is learned. In addition, because supervised calibration explicitly optimizes performance on the target session, meta-learning can become more session-specific, which manifests as a performance drop on the remaining test sessions (Fig. 4b). As in prior experiments, increasing the number of calibration repetitions mitigates this effect, yielding less over-specialization and improved retention.

To further reduce forgetting, we also augment meta-learning with an experience replay buffer. This stabilizes the calibration updates and preserves performance on sessions not used for adaptation (Fig. 4d), albeit typically at the cost of lower accuracy on the adapted session. Figure 4c summarizes these results: left bars correspond to standard meta-learning, while right bars report the replay-regularized variant.

D. Compute/Memory Trade-offs

Table I compares our models with prior work, focusing on inter-session accuracy. For consistency with the literature, we report performance on the *current* adaptation session and omit variants that use the DER buffer. Unless otherwise stated, our results are computed on the full continuous stream, i.e., without excluding transient segments.

Despite its simplicity, the base TCN already improves upon earlier temporal baselines, and unlabeled test-time adaptation yields further gains. AdaBN achieves strong improvements with negligible additional memory and MMAC overhead, since it only updates normalization moments. The GMM-based alignment also improves performance while remaining lightweight: in our accounting, it requires only ~ 68 kB of additional memory and ~ 1.2 MMAC extra compute beyond the forward pass. These compute figures are reported for an example batch size of 32 and a sequence length of 1 to match common reporting in the literature.

We note, however, that this unit-length accounting does not correspond to a full adaptation update in practice. Both BN statistics and GMM/moment estimates must accumulate over longer prefixes (e.g., a full gesture repetition) to become reliable. Consequently, a complete adaptation step over an entire gesture incurs a proportionally larger cost (e.g., ~ 10.1 MMAC), but this increase is comparable to processing the same time horizon with a standard forward pass. Finally, the compute gap between GMM-based alignment and meta-learning stems from where adaptation is applied: statistical alignment updates a constrained set of LoRA parameters at an intermediate representation (after four blocks), whereas meta-learning requires backpropagation through the full network during the inner-loop gradient steps.

IV. DISCUSSIONS

Achieving stable and generalizable decoding of EMG signals across sessions is essential for the reliable use of myoelectric interfaces in real-world, long-term applications. Our experiments highlight distinct adaptation behaviors across the proposed strategies in inter-session EMG decoding. The task consists of classifying seven grasping movements from the NinaPro DB6 dataset, which is challenging due to the similarity between gestures and the strong temporal variability of EMG. Without adaptation, the baseline TCN achieves roughly 80% intra-session accuracy but drops to less than 60% inter-session, confirming the limited generalization of EMG-based decoders.

The first method, based on causal test-time updates of batch-normalization statistics, provides a lightweight, unsupervised mechanism to recalibrate feature distributions without retraining, making it well suited for edge or wearable deployment. Its performance, however, depends strongly on the amount (and balance) of available adaptation data: with only a few repetitions, target normalization statistics are poorly estimated, which can cause a transient accuracy drop below the baseline and mild over-specialization to the current session. Because updates are performed online, we can track performance continuously as target statistics accumulate; in practice, reliable recovery typically requires a sufficiently long horizon at test-time (on the order of tens of seconds) to obtain stable estimates.

TABLE I: Performance and Resource Usage Comparison with State-Of-the-Art Methods

* indicates results including rest state classification in the accuracy. MMAC are reported per time step (sequence length 1) with an example batch size 32, separating the base model and adaptation cost (in parenthesis).

Model	Accuracy Inter (Intra)	σ (Acc) Inter (Intra)	Parameters	Memory	MMAC
Previous Works					
TempoNet* [30]	49.6 (54.5)	—	320k	1.8 MB	16
ECNN-A* [23]	51.44	11.6	—	—	—
This work (Unlabeled adaptation)					
Base Model, TCN	56.61 (85.12)	0.39 (0.58)	47k	200 kB	1.5
Adaptive BN	68.93	3.12	47k	200 kB	1.5
GMM	69.78	0.56	48.7k	268.2 kB (4 Gestures)	1.5 + (1.2)
This work (Supervised adaptation) Meta	80.30	0.37	48.7k	558.4 kB (21 Gestures)	1.5 + (1.5)

This approach remains computationally inexpensive, as BN incurs negligible additional MAC operations relative to the full network. Evaluating performance on other test sessions while adapting to the current one nevertheless reveals a small deterioration, consistent with the model becoming increasingly session-specific. From this perspective, BN adaptation can be viewed as a form of statistical alignment, where features are re-centered and re-scaled (towards zero mean and unit variance) using target-domain estimates.

Turning to explicit statistical alignment with a tiny experience replay buffer in the spirit of dark experience replay (DER), we observe a complementary behavior: adaptation to the current session remains effective while performance on sessions not used for adaptation is largely preserved, and with more gesture repetitions can even improve. Replay anchors updates to previously observed exemplars, mitigating over-specialization and making adaptation more stable under limited or unbalanced target data. Notably, statistical alignment yields meaningful gains with very little target data (down to a small fraction of the available repetitions), suggesting that part of session variability can be compensated by adapting representations through statistics that are less sensitive to gesture-specific idiosyncrasies. Ablations that remove the replay buffer further isolate this effect: without replay, alignment still improves the adapted session with limited data, but retention on non-adapted sessions degrades, underscoring the role of replay in preventing forgetting. The price for this stability is higher computational cost than BN, since adaptation requires gradient-based updates (albeit restricted to a small parameter subset).

Finally, meta-learning provides the expected upper bound when sparse labels are available at test-time. With only a small number of calibration repetitions, MAML rapidly improves performance and can approach the accuracy of a fully supervised pipeline. This benefit, however, comes with a trade-off: adaptation tends to produce more session-specific solutions, leading to a drop in performance on other sessions. Augmenting MAML with an experience replay buffer mitigates this deterioration, albeit typically at the cost of reduced adaptation strength on the current session. For both MAML and statistical alignment, we use LoRA to constrain the adaptation space for efficiency; moreover, updating only low-rank parameters makes it straightforward to revert the model

to its base configuration when adaptation is not desired.

Figure 5 summarizes the comparative trade-offs among the adaptation strategies. AdaBN recovers performance on a slower timescale (Fig. 5a), for values of the adaptation speed β that avoid overly aggressive updates; the black curve reports the average trend over the set of plausible β choices of Fig. 2, panels d and e. Statistical alignment methods are particularly effective in the low-data regime, in some cases surpassing MAML when the number of repetitions is smaller than the number of gesture classes, whereas MAML becomes more advantageous as additional labeled repetitions are provided. Figure 5b further reports performance on sessions held out from the current adaptation, showing that only replay-regularized methods (i.e., approaches that incorporate continual-learning principles) can consistently preserve, and even improve when considering more data, accuracy on non-adapted sessions, especially as more repetitions are observed.

Overall, all approaches substantially improve cross-session robustness compared to the non-adapted baseline, but with complementary strengths: BN provides lightweight recalibration directly on the target stream; replay-regularized statistical alignment promotes stability and mitigates forgetting across sessions; and meta-learning further improves accuracy when supervisory signals are available during calibration. These results support the conclusion that lightweight test-time adaptation can significantly enhance inter-session generalization, an aspect that is often obscured by evaluations dominated by intra-session validation.

Although session shifts in EMG are often attributable to practical factors such as sensor displacement and day-to-day re-donning, it remains valuable to equip the model with mechanisms that can autonomously detect and track distributional changes. The behavior of our causal BN variant across different values of β , with no single setting performing best at all times, suggests that adaptation over multiple timescales could further improve robustness. Moreover, discrepancies between stored source statistics and test-time estimates computed at different timescales may provide a simple signal for monitoring out-of-distribution (OOD) conditions. Such a signal could be used to drive adaptive control of the deployment pipeline, for example, deciding when to increase or slow down adaptation, trigger a stronger replay-regularized update or a brief user calibration, or

fall back to the base trained configuration when online updates appear unreliable.

In this regard, our results indeed suggest that different adaptation mechanisms could be combined to balance responsiveness and stability during deployment. In a practical wearable or embedded implementation, BN adaptation could run continuously on-device to provide immediate, low-cost recalibration to changing conditions. In parallel, replay-regularized fine-tuning, via statistical alignment or via meta-learning when a brief user calibration is available, could be performed intermittently (e.g., off-device) to consolidate learning and update model parameters more robustly. Over time, such updates could leverage an expanding user-specific history of recordings, enabling progressively better personalization.

This hybrid pipeline would exploit complementary strengths, fast on-device adaptation and longer-term robustness, supporting scalable and energy-efficient EMG decoding in real-world applications. As future work, this direction could be integrated with privacy, preserving personalization schemes (e.g., federated learning), where model updates are aggregated across users without centralizing raw EMG data.

V. CONCLUSIONS

We investigated lightweight test-time adaptation for inter-session EMG-based gesture recognition. While the baseline achieved strong intra-session accuracy, performance degraded markedly on unseen sessions, reflecting the pervasive out-of-distribution (OOD) shifts encountered in real-world neuromuscular sensing.

More broadly, these findings support the view that robustness in neural networks need not rely solely on training-time coverage of all operating conditions, an assumption that can be particularly unrealistic for medical data and costly for energy- and memory-constrained wearables. Instead, modest adaptation at deployment can correct session-specific shifts using only the data observed in operation.

We evaluated three complementary strategies: causal AdaBN updates of normalization statistics, replay-regularized statistical alignment with LoRA parameters, and few-shot meta-learning to mimic supervised calibration. All methods improved inter-session robustness over the non-adapted baseline. AdaBN provides the lowest-cost recalibration but benefits from longer test-time horizons, whereas replay-regularized alignment yields the most stable behavior under limited data by mitigating over-specialization and forgetting; meta-learning delivers the fastest gains when sparse labels are available. Overall, minimal test-time adaptation offers a practical route to more reliable long-term EMG decoding for wearable and prosthetic control.

REFERENCES

- [1] Simone Benatti et al. “EMG Acquisition and Processing for Hand Movement Decoding on Embedded Systems: State of the Art and Challenges”. *Proceedings of the IEEE* (2025).
- [2] B. Milosevic, E. Farella, and S. Benatti. “Exploring Arm Posture and Temporal Variability in Myoelectric Hand Gesture Recognition”. *2018 7th IEEE International Conference on Biomedical Robotics and Biomechanics (Biorob)*. 2018.
- [3] Aaron J Young, Levi J Hargrove, and Todd 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* 58 (2011).
- [4] P. Kaufmann et al. “Fluctuating EMG signals: Investigating long-term effects of pattern matching algorithms”. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE. 2010.
- [5] U. Côté-Allard et al. “Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning”. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27 (2019).
- [6] Fuzhen Zhuang et al. “A Comprehensive Survey on Transfer Learning”. *Proceedings of the IEEE* 109 (2021).
- [7] Elisa Donati et al. “Long-term stable electromyography classification using canonical correlation analysis”. *2023 11th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE. 2023.
- [8] Yuejiang Liu et al. “Ttt++: When does self-supervised test-time training fail or thrive?” *Advances in Neural Information Processing Systems* 34 (2021).
- [9] Yufan He et al. “Autoencoder based self-supervised test-time adaptation for medical image analysis”. *Medical image analysis* 72 (2021).
- [10] Yossi Gandelsman et al. “Test-time training with masked autoencoders”. *Advances in Neural Information Processing Systems* 35 (2022).
- [11] Yanghao Li et al. “Revisiting batch normalization for practical domain adaptation”. *arXiv preprint arXiv:1603.04779* (2016).
- [12] Sergey Ioffe and Christian Szegedy. “Batch normalization: Accelerating deep network training by reducing internal covariate shift”. *International conference on machine learning*. pmlr. 2015.
- [13] Leonardo Ravaglia et al. “Memory-Latency-Accuracy Trade-Offs for Continual Learning on a RISC-V Extreme-Edge Node”. *2020 IEEE Workshop on Signal Processing Systems (SiPS)*. 2020.
- [14] Pietro Buzzega et al. *Dark Experience for General Continual Learning: a Strong, Simple Baseline*. 2020. arXiv: 2004.07211 [stat.ML].
- [15] M. Zanghieri et al. “sEMG-based Regression of Hand Kinematics with Temporal Convolutional Networks on a Low-Power Edge Microcontroller”. *IEEE COINS 2021*. 2021.
- [16] Francesca Palermo et al. “Repeatability of grasp recognition for robotic hand prosthesis control based on sEMG data”. *2017 International Conference on Rehabilitation Robotics (ICORR)*. IEEE. 2017.
- [17] Adam Santoro et al. “Meta-learning with memory-augmented neural networks”. *International conference on machine learning*. PMLR. 2016.

- [18] Jian Liang, Ran He, and Tieniu Tan. “A comprehensive survey on test-time adaptation under distribution shifts”. *International Journal of Computer Vision* 133 (2025).
- [19] German I Parisi et al. “Continual lifelong learning with neural networks: A review”. *Neural networks* 113 (2019).
- [20] Mike Huisman, Jan N Van Rijn, and Aske Plaat. “A survey of deep meta-learning”. *Artificial Intelligence Review* 54 (2021).
- [21] Sergey Ioffe. “Batch normalization: Accelerating deep network training by reducing internal covariate shift”. *arXiv preprint arXiv:1502.03167* (2015).
- [22] M Zanghieri et al. “Robust Real-Time Embedded EMG Recognition Framework using Temporal Convolutional Networks on a Multicore IoT Processor”. *IEEE TBIOCAS* 14 (2019).
- [23] Yuzhou Lin et al. “Robust long-term hand grasp recognition with raw electromyographic signals using multi-dimensional uncertainty-aware models”. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 31 (2023).
- [24] Alessio Burrello et al. “Bioformers: Embedding transformers for ultra-low power sEMG-based gesture recognition”. *2022 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE. 2022.
- [25] Yanlong Chen et al. “WaveFormer: A Lightweight Transformer Model for sEMG-based Gesture Recognition”. *arXiv preprint arXiv:2506.11168* (2025).
- [26] Özgün Turgut et al. “Towards generalisable time series understanding across domains”. *arXiv preprint arXiv:2410.07299* (2024).
- [27] Mononito Goswami et al. “Moment: A family of open time-series foundation models”. *arXiv preprint arXiv:2402.03885* (2024).
- [28] Edward J. Hu et al. *LoRA: Low-Rank Adaptation of Large Language Models*. 2021. arXiv: [2106.09685](https://arxiv.org/abs/2106.09685) [cs.CL].
- [29] Chelsea Finn, Pieter Abbeel, and Sergey Levine. “Model-agnostic meta-learning for fast adaptation of deep networks”. *International conference on machine learning*. PMLR. 2017.
- [30] Marcello Zanghieri et al. “sEMG-based regression of hand kinematics with temporal convolutional networks on a low-power edge microcontroller”. *2021 IEEE International Conference on Omni-Layer Intelligent Systems (COINS)*. IEEE. 2021.