

A Cyclic Adaptation-Generalization Framework with Uncertainty-Guided Self-Paced Learning for Long-Term Brain-Machine Interfaces

Jiyu Wei^{1,2}, Di Hong^{1,2}, Zhanjie Zhang¹, Dazhong Rong^{*,1,2}, Qinming He^{1,2}, Yueming Wang^{1,2}

¹College of Computer Science and Technology, Zhejiang University, Hangzhou, China

²Nanhu Brain-Computer Interface Institute, Hangzhou, China

{weijiyu, hongd, cszsj, rdz98, hqm, ymingwang}@zju.edu.cn

Abstract—Brain-Machine Interfaces (BMIs), which link the brain to external devices, hold great potential in rehabilitation, human performance augmentation, and human-centered robotics. However, invasive BMIs face a critical challenge for long-term deployment due to neural drift, which degrades decoding performance over time and necessitates frequent recalibration. Existing methods designed to mitigate neural drift typically rely on either domain adaptation (DA) or domain generalization (DG) alone and often fail to capture fine-grained distribution shifts across neural subdomains, resulting in limited performance. To overcome these limitations, we propose Uncertainty-guided Self-paced Cycling (UnSPC), a robust framework that synergizes DA and DG for target domain refining under an Uncertainty-guided Self-paced Pseudo-labeling (UnSPL) mechanism. To handle subdomain neural drift across domains, UNSPL is proposed to iteratively mine reliable pseudo-labeled samples with a noise-robust ranking strategy for further fine-tuning. Leveraging these high-quality samples, we introduce a novel Cycling Adaptation and Generalization (CycAG) strategy, which integrates DA and DG within an iterative cycle to progressively mitigate both global and subdomain drift. This cyclic process enables effective alignment to evolving target distributions while preserving robust and transferable representations, thereby mitigating performance degradation under long-term neural drifts. Extensive experiments on multiple neural decoding datasets demonstrate the effectiveness and robustness of UnSPC. To our knowledge, our proposed UnSPC is the first to cyclically integrate DA and DG with pseudo-labeling, paving the way toward stable long-term BMI controls.

I. INTRODUCTION

As a promising paradigm for direct brain-device interaction, Brain-Machine Interfaces (BMIs) have attracted growing attention due to their potential in neurorehabilitation, assistive technologies, human performance augmentation, and human-robot interaction [1], [2], [3], [4], [5]. Among various BMI modalities, invasive motor BMIs (iBMIs) directly record neural activity from motor cortical areas, providing high-fidelity signals that enable accurate decoding of motor intentions for effective control of prosthetic devices and robotic systems [6], [7], [8]. Over the past decade, substantial progresses in neural population recording technologies [9] and neural decoding algorithms [10], [11], [12], [13] have significantly improved decoding accuracy and expanded the

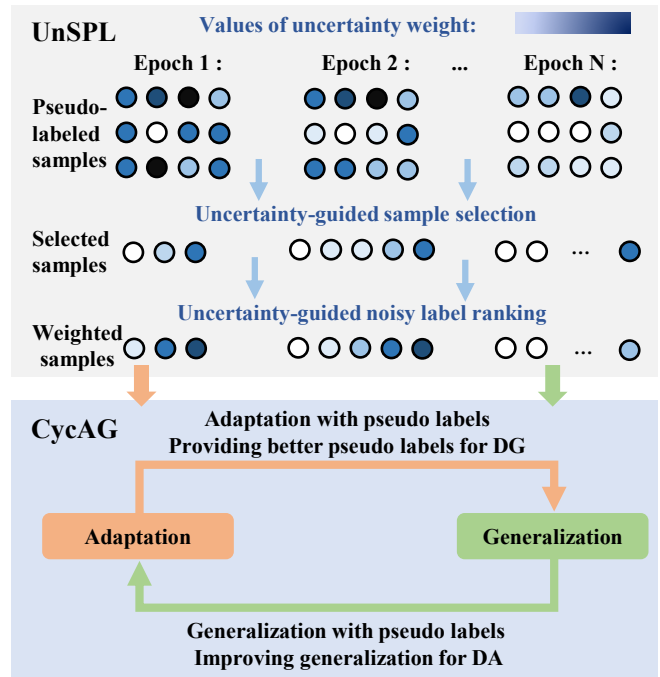


Fig. 1. UnSPC consists of two key modules: (i) UnSPL mines reliable pseudo-labeled samples; (ii) CycAG cyclically alternates DA and DG.

practical applicability of iBMIs. However, achieving reliable long-term operation of iBMIs remains a fundamental challenge. In real-world deployments, iBMI systems often suffer from neural drift, which may be caused by electrode shifts, neuronal turnover, and neural plasticity [14], [15]. As a result, the statistical relationship between neural activity and intended movements gradually changes over time. Consequently, decoders trained on labeled day₀ data often generalize poorly to later sessions day_k, leading to progressive performance degradation and necessitating periodic recalibration.

Numerous approaches have been proposed to address neural non-stationarity, typically categorized into Domain Adaptation (DA) and Domain Generalization (DG) based methods. DA based methods maintain performance by aligning the decoder with day_k statistics [16], [17], [18], [19], yet they risk negative transfer if driven by noisy features or se-

*Corresponding author: Dazhong Rong.

vere distribution shifts. Conversely, DG based methods [20], [21], [22], [23] emphasize learning invariant representations across sessions but are often overly conservative, failing to capture session-specific neural drifts by ignoring target-domain statistics. Crucially, existing methods typically treat DA and DG as isolated paradigms. In this paper, we argue that DA and DG are complementary where robust long-term neural decoding requires both adaptability to evolving neural distributions and generalization across unseen variations. Specifically, DG provides a stable initialization that shields DA from noise-induced overfitting, while DA leverages unlabeled target data to provide supervision signals that guide DG [24], [25], [26].

Furthermore, prior methods largely focus on global marginal distributions, often neglecting subdomain shifts, which are critical for fine-grained neural decoding [27], [28], [29]. Pseudo-labeling [30], [31], [32] offers a promising way to address this by assigning tentative labels to unlabeled target data for self-training. However, its application to iBMIs faces two unique hurdles: (i) the difficulty of confidence estimation in continuous regression (unlike discrete classification), and (ii) the risk of error propagation due to the inherent noise in neural signals. Consequently, a robust, regression-based pseudo-labeling framework for iBMIs remains an open challenge.

To this end, we propose a novel Uncertainty-guided Self-paced Cycling (UnSPC) framework, which unifies DA and DG in a single iterative learning paradigm with self-paced pseudo-labeling strategy through an uncertainty-guided mechanism. First, a Robust Pre-train Decoder (RPD) is learned on the source session through hybrid neural decoding and consistency constraints, which improve robustness to neural noise and electrode instability. To address the dual challenges of subdomain drift and label scarcity on unlabeled target domain, we introduce Uncertainty-guided Self-paced Pseudo-labeling (UnSPL) which orchestrates a robust curriculum that progressively mines high-confidence pseudo-labels, while effectively attenuating noise propagation through a rank-guided soft weighting mechanism.

Intuitively, these high-fidelity samples could serve as reliable supervision for the target domain refining, enabling a cyclic refinement process of DA and DG. Building on this, we introduce Cyclic Adaptation-Generalization (CycAG), a novel iterative optimization scheme which creates a synergistic cycle between DA and DG, progressively improving decoding performance on reliable data while extending robustness to the remaining uncertain samples. At each iteration, reliable pseudo-labeled samples are mined and weighted based on UnSPL, and further used to alternately perform DA and DG by CycAG. Note that, this alternating optimization between DA and DG enables a complementary interplay: adaptation improves target-specific performance, and generalization enhances robustness to distribution shifts, leading to a more stable and reliable decoding model under neural drift. Extensive experiments on multiple iBMI datasets demonstrate UnSPC significantly outperforms state-of-the-art DA and DG methods, achieving superior cross-session

decoding performance without requiring additional labeled data from target sessions.

Our main contributions are summarized as follows:

- We propose UnSPC, which to our knowledge is the first unified framework integrating DA and DG with pseudo-labeling for regression-based neural decoding, effectively bridging the gap between DA and DG.
- We introduce UnSPL to mine reliable pseudo-labeled samples from unlabeled target sessions, preventing error propagation during target domain fine-tuning.
- We design CycAG, a cyclic optimization scheme that alternates between DA and DG, enabling complementary learning of target-specific and invariant representations.
- Extensive validations across multiple datasets demonstrate that UnSPC significantly outperforms state-of-the-art baselines, establishing a stable computational methodology for long-term iBMI decoding.

II. RELATED WORK

Domain Adaptation (DA) based Methods. Supervised recalibration can effectively mitigate neural drift but is often impractical in longitudinal BMI settings due to the limited availability of labeled data [33]. As a result, Un-supervised Domain Adaptation (UDA) has become a widely adopted strategy. Representative UDA based approaches include adversarial learning methods, such as ADAN [16] and UAN [17], which minimize the distribution discrepancy between the source (day₀) and target (day_k) domain, as well as divergence-based alignment methods, such as DA-DCF [34] and WDGRL [18], which align latent feature distributions across sessions. In addition, SeSA [19] introduces semi-supervised DA, which leverages structural constraints to conditionally align kinematic features, thereby preserving the intrinsic manifold structure during domain adaptation.

Domain Generalization (DG) based Methods. Unlike DA, DG aims to learn representations that remain robust to distribution shifts using only source-domain data. In neural decoding, DG based methods typically seek invariant neural representations that generalize across recording sessions. Some studies focus on representation learning to obtain stable neural embeddings. For example, KRSRL [11] and CDNG [22] employ similarity-based objectives to enforce consistent latent representations for neural activities associated with similar movements. Some studies improve robustness through temporal modeling or ensemble strategies. For example, DyEnsemble [13] aggregates multiple dynamic decoders to capture diverse temporal dependencies in neural signals. Some studies attempt to model invariant latent dynamics of neural populations. For example, CAPTIVATE [21] and SABLE [35] learn structured latent representations to stabilize neural dynamics across sessions, while LFDA [20] combines neural data augmentation with contrastive learning to learn latent factors robust to recording condition changes.

Although the above DA and DG based methods achieve competitive performance, they overlook the benefits that integrating DA and DG could bring, limiting their performance.

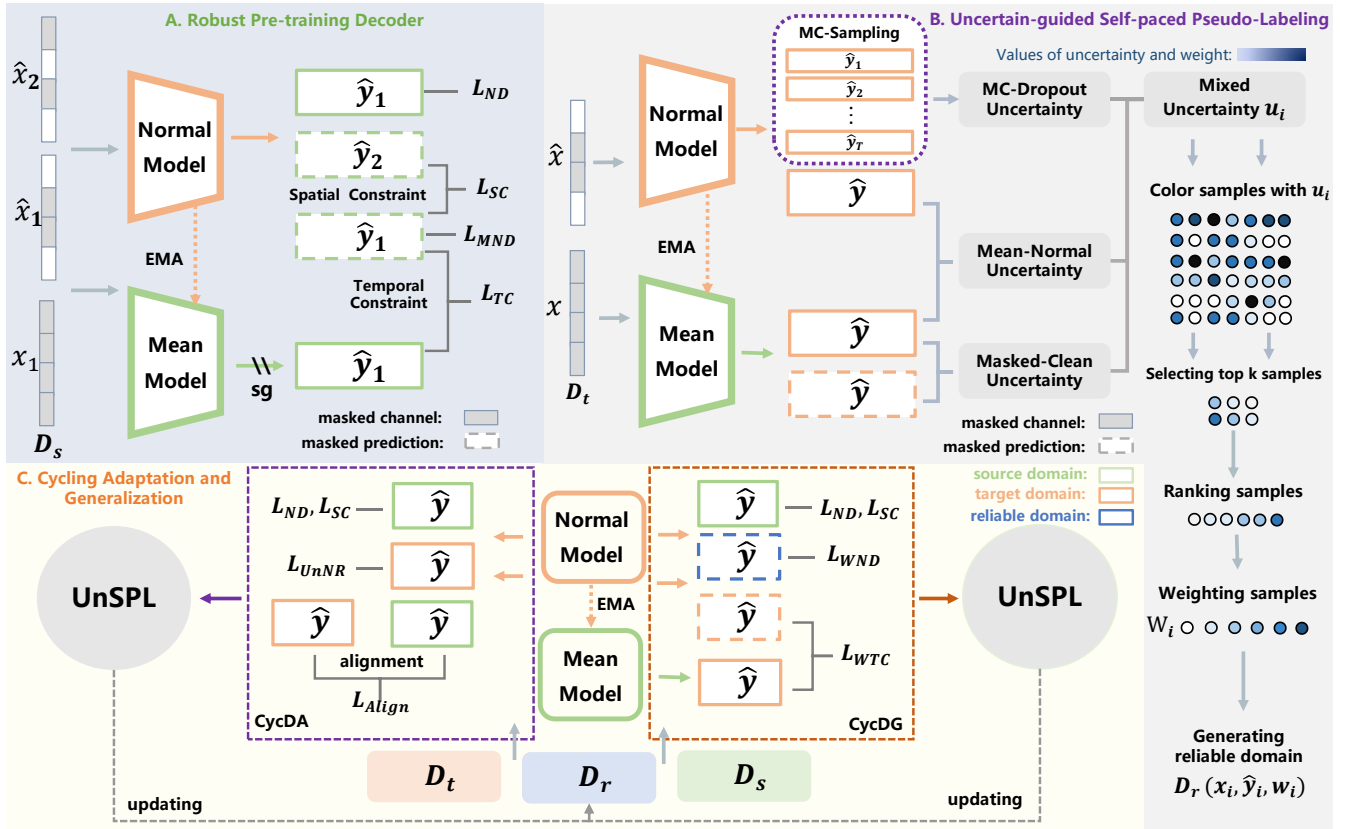


Fig. 2. The overall framework of our UnSPC. A: Robust Pre-training Decoder (RPD) pre-trained on the source domain for initialization. B: Uncertainty-guided Self-paced Pseudo-labeling (UnSPL) selecting and weighing reliable samples on the target domain to form a reliable domain. C: Cyclic Adaptation-Generalization (CycAG), cyclically alternating between adaptation and generalization using reliable pseudo-labeled samples.

III. METHOD

In this work, we propose Uncertainty-guided Self-paced Cycling (UnSPC), a framework designed to mitigate performance degradation under long-term neural drift in BCIs, as depicted in Fig. 1 and Fig. 2. The key idea of UnSPC is to progressively adapt the decoding model to evolving neural distributions while preserving robust and generalizable representations. We first introduce Robust Pre-training Decoder (RPD) as a stable backbone, which is pre-trained on the source domain via hybrid neural decoding and consistency constraints. To handle subdomain drift, we propose Uncertainty-guided Self-paced Pseudo-labeling (UnSPL), which quantifies predictive uncertainty to progressively mine high-fidelity pseudo-labeled samples from the unlabeled target domain. With these reliable samples, we introduce Cycling Adaptation-Generalization (CycAG), an iterative scheme that alternates between adaptation and generalization for target session refinement. In the following, we first formulate the neural decoding problem under session-wise distribution shifts and then introduce the proposed UnSPC and its key components.

A. Problem Formulation

We consider a neural decoding task where neural signals are used to predict corresponding behavioral variables, such as movement trajectories or intended actions. In iBMIs,

neural recordings collected across different sessions often exhibit substantial distribution shifts due to neural drift and recording instability. Let $\mathcal{S} = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$ and $\mathcal{T} = \{(x_i^t, y_i^t)\}_{i=1}^{N_t}$ denote the source-session (day_0) dataset and a subsequent target-session (day_k) dataset, respectively. Here, x_i and y_i denote neural signals and corresponding behavioral labels respectively. The objective is to learn a neural decoding model composed of a feature extractor $f_\theta(\cdot)$ and a regression head $h_\theta(\cdot)$, i.e., $h_\theta(f_\theta(\cdot))$, that maintains high decoding accuracy on \mathcal{T} despite the distribution shifts, where θ denotes the model parameters. Note that since target-session labels are unavailable during training, we assume access only to labeled day_0 data and unlabeled day_k data.

B. UnSPC Framework Overview

On the whole, our proposed UnSPC involves two stages as follows. (i) **Source-session Pre-training:** we first train a Robust Pre-training Decoding (RPD) backbone on the labeled source session \mathcal{S} as a robust initialization. (ii) **Target-session Refining:** given an unlabeled target session \mathcal{T} , we further refine the model through an iterative procedure; at each iteration w , the model first constructs a reliable pseudo-labeled set \mathcal{R}_w with soft weights via UnSPL, and then \mathcal{R}_w is utilized by CycDG to update the model by alternating DA and DG across successive iterations.

To stabilize training against inherent neural noise, we

employ a teacher-student paradigm [36] throughout both the pre-training stage and the refining stage. The teacher parameters are updated without gradients as an exponential moving average (EMA) of the student parameters:

$$\hat{\theta}_w = \begin{cases} \theta_w, & \text{if } w \leq w_0 \\ \frac{w-w_0}{w-w_0+1} \cdot \hat{\theta}_{w-1} + \frac{1}{w-w_0+1} \cdot \theta_w & \text{otherwise} \end{cases}, \quad (1)$$

where $\hat{\theta}_w$ and θ_w denote the parameters of the teacher and student models at iteration w , respectively, and w_0 denotes the warm-up step before the EMA update starts. Since the teacher model maintains a temporally smoothed version of the student parameters, during inference the teacher model is used as the finally obtained decoder for prediction.

C. Robust Pre-training Decoder

For initialization, we propose a Robust Pre-training Decoder (RPD) backbone using labeled data during the source-session pre-training. To improve robustness against electrode instability and neural noise, we adopt a hybrid neural decoding strategy together with consistency constraints.

1) **Hybrid Neural Decoding:** Neural recordings are often affected by electrode instability or channel dropouts across recording sessions. To improve robustness to such perturbations, we train the decoder using hybrid neural inputs consisting of both clean and masked observations. The loss for clean neural decoding is defined as:

$$\mathcal{L}_{\text{CND}} = \frac{1}{|\mathcal{S}|} \sum_{(x_i^s, y_i^s) \in \mathcal{S}} \|h_{\theta}(f_{\theta}(x_i^s)) - y_i^s\|_2^2. \quad (2)$$

To simulate electrode instability, we randomly mask c neural channels in x_i^s to generate a corrupted observation \hat{x}_i^s . The masked neural decoding loss is defined as:

$$\mathcal{L}_{\text{MND}} = \frac{1}{|\mathcal{S}|} \sum_{(x_i^s, y_i^s) \in \mathcal{S}} \|h_{\theta}(f_{\theta}(\hat{x}_i^s)) - y_i^s\|_2^2. \quad (3)$$

2) **Consistency Constraints:** In addition to hybrid neural decoding losses, we introduce consistency constraints to stabilize representation learning and improve cross-session generalization. Neural representations corresponding to similar behavioral states should exhibit similar geometric relationships. To preserve this structure, we enforce alignment between pairwise distances in the neural feature space and those in the behavioral space. The feature-label consistency loss is defined as:

$$\mathcal{L}_{\text{FLC}} = \frac{1}{|\mathcal{S}|^2} \sum_{(x_i^s, y_i^s) \in \mathcal{S}} \sum_{(x_j^s, y_j^s) \in \mathcal{S}} \delta_{ij}, \quad (4)$$

and

$$\delta_{ij} = \left(\text{BN}(\|f_{\theta}(x_i^s) - f_{\theta}(x_j^s)\|_2^2) - \text{BN}(\|y_i^s - y_j^s\|_2^2) \right)^2, \quad (5)$$

where $\text{BN}(\cdot)$ denotes the batch normalization operator applied to stabilize the scale of pairwise distances. Besides, to improve robustness to corrupted inputs, we enforce consistency between predictions of the teacher model on clean

signals and predictions of the student model on masked signals. The teacher-student consistency loss is defined as:

$$\mathcal{L}_{\text{TSC}} = \frac{1}{|\mathcal{S}|} \sum_{(x_i^s, y_i^s) \in \mathcal{S}} \|h_{\hat{\theta}}(f_{\hat{\theta}}(x_i^s)) - h_{\theta}(f_{\theta}(\hat{x}_i^s))\|_2^2. \quad (6)$$

3) **Overall Loss:** The final objective for source-session pre-training combines the hybrid decoding losses and the consistency constraints:

$$\mathcal{L}_{\text{PRE}} = \alpha \cdot (\mathcal{L}_{\text{CND}} + \mathcal{L}_{\text{FLC}}) + \mathcal{L}_{\text{MND}} + \mathcal{L}_{\text{TSC}}, \quad (7)$$

where α controls the trade-off between the losses involving both clean and masked neural signals and the losses involving only clean neural signals. By jointly optimizing these objectives, the decoder learns noise-robust neural representations that preserve the geometric relationship between neural activity and behavioral states, providing a reliable initialization for subsequent target-session refining.

D. Uncertainty-guided Self-paced Pseudo-labeling

Despite the above initialization, significant distribution shifts between \mathcal{S} and \mathcal{T} necessitate further fine-tuning. While conventional UDA aligns global feature distributions [34], [19], it often fails to enforce the fine-grained manifold structure required for high-precision regression. Pseudo-labeling [30] offers instance-level supervision but faces two impediments in iBMIs: the lack of a natural confidence metric for regression, and the risk of error propagation from noisy neural signals. To surmount these, we propose Uncertainty-guided Self-paced Pseudo-labeling (Fig. 2-B) to select and weigh pseudo-labeled samples on \mathcal{T} .

Given the unlabeled target-session data, we first perform Reliable Sample Selection (RSS), estimating the reliability of model predictions and constructing a set of pseudo-labeled samples. Specifically, for each target sample x_i^t , we obtain its pseudo-label \hat{y}_i^t using the teacher model:

$$\hat{y}_i^t = h_{\hat{\theta}}(f_{\hat{\theta}}(x_i^t)). \quad (8)$$

To assess the reliability of pseudo-labels, we estimate the prediction uncertainty from two perspectives. First, we compute the Monte-Carlo (MC) Dropout uncertainty by performing M stochastic forward passes with dropout activated at inference time:

$$u_i^{\text{MC}} = \frac{1}{M} \sum_{j=1}^M \|\hat{y}_{i,j}^t - \mathbb{E}[\hat{y}_i^t]\|_2^2. \quad (9)$$

Second, we measure the inconsistency (IC) between different model predictions:

$$u_i^{\text{IC}} = \|\hat{y}_i^t - h_{\hat{\theta}}(f_{\hat{\theta}}(\hat{x}_i^t))\|_2^2 + \|\hat{y}_i^t - h_{\theta}(f_{\theta}(x_i^t))\|_2^2, \quad (10)$$

where \hat{x}_i^t denotes the masked version of x_i^t . The final uncertainty is computed as:

$$u_i = \text{BN}(u_i^{\text{MC}}) + \text{BN}(u_i^{\text{IC}}). \quad (11)$$

Based on the uncertainty values, we adopt a self-paced strategy to progressively select reliable samples. We first define a candidate reliable set parameterized by a threshold:

$$\mathcal{R}(\tau) = \{(x_i^t, \hat{y}_i^t) \mid u_i \leq \tau\}. \quad (12)$$

At iteration w , the threshold τ_w is then chosen as the smallest value that ensures a desired proportion of selected samples:

$$\tau_w = \min\{\tau \mid \frac{|\mathcal{R}(\tau)|}{|\mathcal{T}|} \geq \gamma_0 + w \cdot \Delta\gamma\}, \quad (13)$$

where γ_0 controls the initial sampling ratio and $\Delta\gamma$ defines the incremental increase at each iteration.

Accordingly, the reliable set at iteration w is given by $\mathcal{R}_w = \mathcal{R}(\tau_w)$. This formulation ensures that the model first focuses on highly reliable samples and gradually incorporates more uncertain ones as training progresses. Samples in \mathcal{R}_w are ranked in ascending order based on their uncertainty values. Let r_i denote the rank of sample x_i^t . To further mitigate noise in pseudo-labels, we perform Reliable Sample Weighing (RSW), assigning each sample a weight based on its uncertainty ranking as follows:

$$z_i = \exp(-\beta \cdot \frac{r_i}{|\mathcal{R}_w|}), \quad (14)$$

This weighing mechanism allows more reliable samples to contribute more during training.

E. Cyclic Adaptation-Generalization for Refinement

By treating these reliable samples as newly labeled data, we move beyond considering adaptation and generalization as independent processes, and instead leverage them to jointly enhance both. This motivates the proposed Cyclic Adaptation-Generalization (CycAG), which alternates between DA and DG in a cyclic manner (Fig. 2-C).

1) **Domain Adaptation:** In this step, we leverage the reliable pseudo-labeled set \mathcal{R}_w to align the feature distributions between the source and target domains while improving prediction accuracy. First, we adopt a feature alignment loss based on Maximum Mean Discrepancy (MMD):

$$\mathcal{L}_{\text{MMD}} = \left\| \frac{1}{|\mathcal{S}|} \sum_{(x_i^s, y_i^s) \in \mathcal{S}} f_\theta(x_i^s) - \frac{1}{|\mathcal{T}|} \sum_{(x_i^t, y_i^t) \in \mathcal{T}} f_\theta(x_i^t) \right\|_{\mathcal{H}}^2, \quad (15)$$

where \mathcal{H} denotes a reproducing kernel Hilbert space. Then, we utilize the pseudo-labels and the assigned weights for weighted supervised learning (WSL) on the reliable set:

$$\mathcal{L}_{\text{WSL}} = \frac{1}{|\mathcal{R}_w|} \sum_{(x_i^t, \hat{y}_i^t) \in \mathcal{R}_w} z_i \cdot \|h_\theta(f_\theta(x_i^t)) - \hat{y}_i^t\|_2^2. \quad (16)$$

The overall loss for the DA step is:

$$\mathcal{L}_{\text{DA}} = \alpha \cdot (\mathcal{L}_{\text{CND}} + \mathcal{L}_{\text{FLC}}) + \beta \cdot \mathcal{L}_{\text{WSL}} + \eta \cdot \mathcal{L}_{\text{MMD}}. \quad (17)$$

This step focuses on improving target-domain performance by aligning global distributions and refining predictions using reliable pseudo-labels.

2) **Domain Generalization:** Complementary to DA, this step aims to improve the robustness of the model to unseen variations in neural signals. Instead of directly fitting clean

inputs, we encourage the model to learn invariant representations under corrupted inputs. We first introduce a teacher-student consistency loss over the target domain (TSCT):

$$\mathcal{L}_{\text{TSCT}} = \frac{1}{|\mathcal{T}|} \sum_{(x_i^t, y_i^t) \in \mathcal{T}} \|h_{\hat{\theta}}(f_{\hat{\theta}}(x_i^t)) - h_\theta(f_\theta(\hat{x}_i^t))\|_2^2. \quad (18)$$

Then, we define a weighted masked neural decoding (WMND) loss on the reliable samples:

$$\mathcal{L}_{\text{WMND}} = \frac{1}{|\mathcal{R}_w|} \sum_{(x_i^t, \hat{y}_i^t) \in \mathcal{R}_w} z_i \cdot \|h_\theta(f_\theta(\hat{x}_i^t)) - \hat{y}_i^t\|_2^2. \quad (19)$$

The overall loss for the DG step is:

$$\mathcal{L}_{\text{DG}} = \alpha \cdot (\mathcal{L}_{\text{CND}} + \mathcal{L}_{\text{FLC}}) + \beta \cdot \mathcal{L}_{\text{WMND}} + \phi \cdot \mathcal{L}_{\text{TSCT}}. \quad (20)$$

By enforcing consistency under input perturbations and leveraging weighted pseudo-label supervision, this step enhances the model's generalizability to unseen distributions.

IV. EXPERIMENTS

In this section, we first describe the experimental settings, and then present comparisons with state-of-the-art methods across datasets, followed by ablation and robustness studies to comprehensively evaluate UnSPC.

A. Experimental Settings

1) **Datasets:** We validate UnSPC on three public NHP datasets (Chewie, Mihili, and Jango) [17]. Subjects performed upper-limb tasks including center-out (CO), random-target (RT), and isometric wrist movements (ISO), with neural activity recorded from the primary motor cortex (M1) using 96-channel Utah arrays. Multiunit threshold crossings were extracted, randomly split, and aligned to the central 10% to 90% of each trial (from "go cue" to "trial end"). Firing rates were computed using 50 ms binning, followed by Gaussian smoothing ($\sigma = 100$ ms) and z-score normalization. For long-term stability evaluation, the first session was used as the labeled source session, and subsequent sessions as unlabeled target sessions.

2) **Implementation Details:** The feature extractor is a lightweight 3-layer MLP with ReLU activations and $p = 0.25$ dropout, followed by a linear regression head. We optimize the model using Adam with $learning_rate = 5 \times 10^{-3}$, $beta = (0.9, 0.999)$, and $weight_decay = 5 \times 10^{-4}$. The batch size is 128. Hyperparameters are fixed at $\beta = 0.1$, $\eta = 0.2$, and $\phi = 0.5$ across all datasets. To account for subject variability, the source preservation weight is set to $\alpha = 1.0$ for Chewie and Mihili, and $\alpha = 0.1$ for Jango. Results are reported as the mean of five independent runs with different random seeds.

3) **Metrics:** Decoding performance is assessed using three metrics: Pearson Correlation Coefficient (CC), Coefficient of Determination (R^2), and Mean Squared Error (MSE).

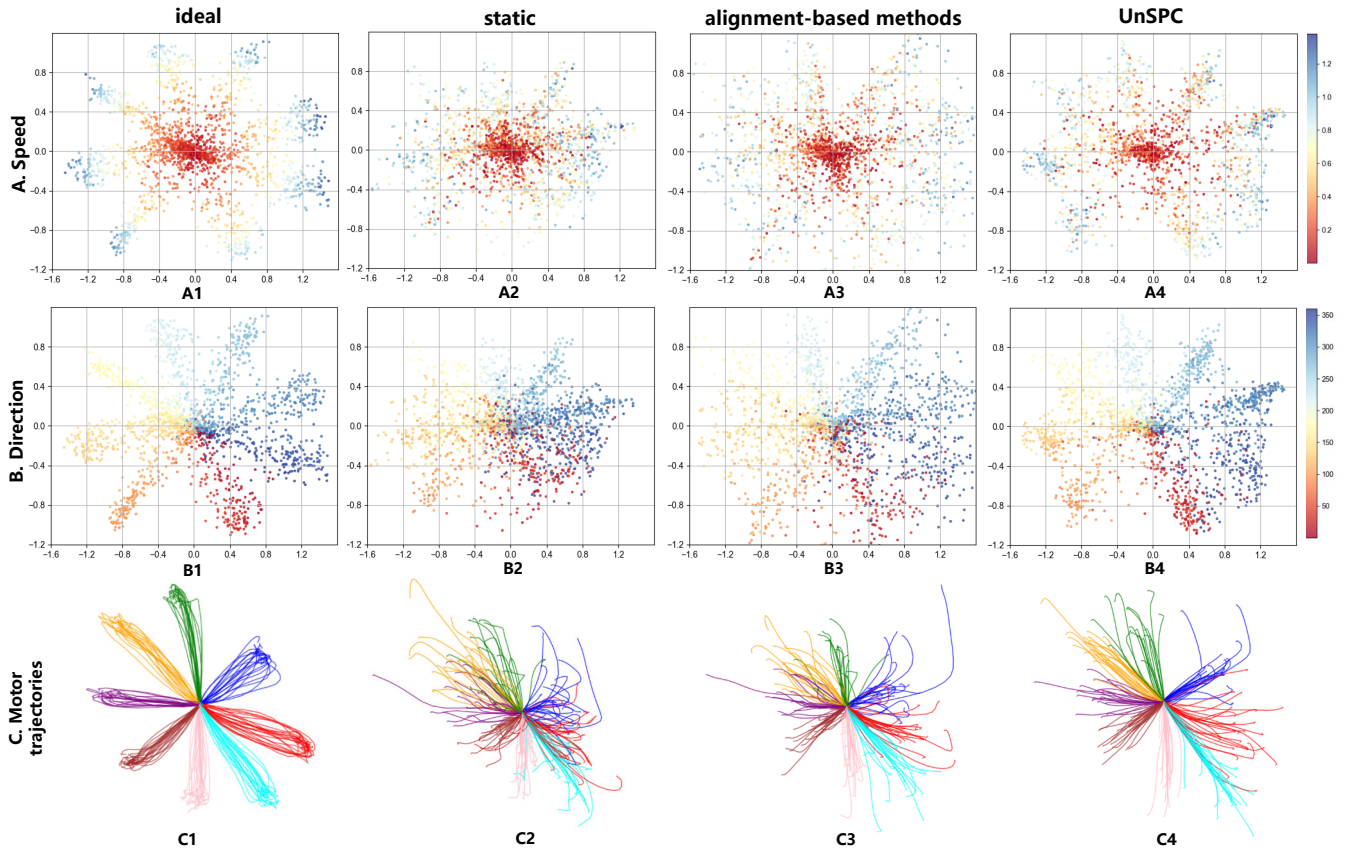


Fig. 3. Feature distributions and motor trajectories of ideal (supervised), static (uncalibrated), alignment-based method (WDGRL) and our UnSPC on C0–C1. A and B: Feature distributions under different alignments colored by values of speed and direction. C: Motor trajectories colored by trial direction.

TABLE I
RECALIBRATION PERFORMANCE COMPARISON (BOLD FOR THE BEST AND UNDERLINE FOR THE SECOND-BEST)

Dataset	Method													
	Static		LFDA		DA-DCF*		UAN		WDGRL*		SeSA		UnSPC	
	<i>CC</i>	<i>R</i> ²	<i>CC</i>	<i>R</i> ²	<i>CC</i>	<i>R</i> ²	<i>CC</i>	<i>R</i> ²	<i>CC</i>	<i>R</i> ²	<i>CC</i>	<i>R</i> ²	<i>CC</i>	<i>R</i> ²
C1	84.78	59.24	87.91	72.35	87.35	62.03	84.35	64.79	<u>86.92</u>	<u>72.45</u>	84.39	69.43	88.70	78.40
C2	65.94	40.55	74.58	55.12	75.12	51.64	77.91	55.57	78.82	65.92	<u>84.59</u>	<u>66.30</u>	88.91	78.77
C3	57.43	30.23	68.14	41.21	64.89	40.11	70.98	46.28	80.09	60.03	<u>82.15</u>	<u>65.35</u>	84.94	71.64
J1	81.89	66.41	<u>84.39</u>	<u>70.66</u>	83.09	67.62	81.19	65.04	83.26	68.05	83.14	70.66	85.42	72.08
J2	74.71	56.24	82.28	<u>65.30</u>	77.75	59.01	79.74	63.61	78.32	64.98	<u>82.45</u>	<u>65.75</u>	83.21	68.13
J3	77.48	60.95	80.86	58.43	75.84	58.55	77.82	59.63	78.15	62.33	<u>81.51</u>	<u>67.73</u>	83.38	68.33
M1	77.26	64.82	82.22	64.77	77.89	65.61	78.21	65.72	78.52	65.49	<u>78.09</u>	<u>66.17</u>	82.25	67.49
M2	67.66	49.17	74.10	54.89	67.84	50.15	67.94	51.12	71.28	55.58	<u>75.10</u>	<u>60.78</u>	79.66	62.99
M3	58.45	30.12	65.24	42.30	42.31	30.43	49.51	36.79	55.21	40.56	71.57	54.24	<u>68.61</u>	<u>46.99</u>

B. Recalibration Performance Comparison

We evaluate the recalibration performance through qualitative and quantitative experiments across multiple datasets. Data from four days per subject were used: C0–C3 (Chewie), J0–J3 (Jango), and M0–M3 (Mihili). C0/J0/M0 serve as source domains, and the remaining sessions correspond to short (~1 day), medium (~1 month), and long (~3 months) time spans. “*” denotes our reproduced results.

1) **Qualitative Results:** From the qualitative comparisons of feature distributions (Fig. 3-A and B) we can reveal that: (i) An ideal feature distribution exhibits a radial shape and

maintains label-feature consistency. (ii) The feature distributions of the static decoder show significant deviations. (iii) Previous methods perform feature alignment, which reduces the deviations to some extent. (iv) Our UnSPC achieves nearly ideal feature distributions. A similar conclusion can be drawn from the motor trajectories in Fig. 3-C, where the trajectories decoded by our method are closer to the ideal motion trajectories, demonstrating the effectiveness and superiority of UnSPC from a qualitative perspective.

2) **Quantitative results:** As shown in Table I, quantitative results show that the static Decoder degrades significantly

TABLE II
ABLATION RESULTS OF UNSPC

Configuration	CC	R^2	MSE
Baseline	84.78	59.24	.03668
Baseline+RPD	87.22	73.04	.02504
UnSPC w/o RSS	88.56	77.46	.02256
UnSPC w/o RSW	88.37	<u>77.80</u>	<u>.02225</u>
UnSPC w/o UnSPL	86.78	73.30	.02688
UnSPC w/o CycDA	<u>88.57</u>	76.36	.02379
UnSPC w/o CycDG	87.94	77.01	.02305
UnSPC w/o Cyc	88.52	77.69	.02263
UnSPC	88.70	78.40	.02168

without recalibration. DA based methods (DA-DCF, UAN, WDGR) partially mitigate this via feature alignment, and the DG based method (LFDA) performs comparably, yet all remain limited by the absence of labels. The semi-supervised method SeSA improves over most unsupervised methods but requires a few labeled target data. In contrast, UnSPC substantially outperforms the static decoder baseline and achieves state-of-the-art performance among DA and DG based methods across all datasets, surpassing SeSA on most datasets despite being fully unsupervised.

C. Ablation Study

To validate the effectiveness of each component in UnSPC, we conduct ablation experiments with C0 as the source domain and C1 as the target domain (Table II).

1) **RPD is a strong backbone:** Directly applying a source-trained neural decoder to the target domain results in a significant performance drop, highlighting neural signal shifts across days. Incorporating RPD substantially mitigates this drop, confirming its role as a strong backbone.

2) **UnSPL is effective and critical:** Removing either RSS or RSW leads to performance degradation, demonstrating their significance. Without UnSPL (i.e., applying pseudo-labels directly) performance drops severely (lower than Baseline+RPD), confirming that pseudo-label noise can be detrimental and UnSPL is essential for effective recalibration.

3) **CycAG can improve performance:** When UnSPC is trained without CycDA and CycDG, relying only on RPD and UnSPL, performance drops compared to using CycDA or CycDG individually, confirming that CycDA improves adaptation on unlabeled data while CycDG enhances generalization. Combining DA and DG without cycling yields competitive results, but the full CycAG framework achieves the best performance, demonstrating the effectiveness of iterative cycling. These findings highlight that the synergistic integration of UnSPL and CycAG leads to optimal recalibration, establishing UnSPC as a state-of-the-art method for neural decoding.

D. Robustness Analysis

1) **Sensitivity Analysis of Hyperparameters:** We assess the sensitivity of α , β , η , and ϕ , keeping other parameters

TABLE III
SENSITIVITY OF HYPERPARAMETERS ON C1

Parameter	Tested Values	R^2
α	0.1, 0.5, 1.0, 2.0	73.0, 76.5, 78.4, 78.2
β	0.05, 0.1, 0.2, 0.5	77.5, 78.4, 77.6, 75.8
η	0.05, 0.2, 0.5	78.0, 78.4, 78.3
ϕ	0.1, 0.5, 1.0	77.9, 78.4, 78.5

TABLE IV
ROBUSTNESS ON HUMAN DATASET

Target Domain	Static	UAN	WDGR	UnSPC
01-08	10.43	12.01	<u>12.71</u>	14.58
01-20	9.14	11.16	<u>12.39</u>	14.17
02-09	8.76	10.98	<u>11.74</u>	13.91

at default values. As shown in Table III, α should be sufficiently large to preserve baseline performance, β moderate to leverage pseudo-labels without amplifying noise, while η and ϕ have minor effects but can stabilize training. Overall, within a reasonably wide range, UnSPC demonstrates low sensitivity to these hyperparameters, indicating robust performance across different settings.

2) **Robustness on Human Dataset:** We further evaluated the robustness of UnSPC on a human robotic effector dataset from FALCON [37], where the source domain consists of recordings on Date 01-01 and the participant performed a virtual arm reaching task. Table IV reports performance across three subsequent target days. Compared with the static decoder and SOTA methods, UnSPC consistently achieves the highest performance, demonstrating its effectiveness and robustness in real-world human BMI scenarios. These results indicate that our method can reliably handle cross-day neural variability and maintain stable decoding without access to labeled target data.

V. CONCLUSION

In this paper, we propose UnSPC, a novel framework for long-term iBMIs that effectively integrates DA and DG with pseudo-labeling to address subdomain neural drift. Extensive experiments across multiple datasets demonstrate that UnSPC achieves superior performance compared to existing methods, highlighting its robustness and practical potential for human-robot interaction and control in robotic systems.

VI. ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant (U25D9015, 62336007), in part by the Starry Night Science Fund of the Zhejiang University Shanghai Institute for Advanced Study under Grant SN-ZJU-SIAS-002, in part by the Fundamental Research Funds for the Central Universities. The use of Gemini are acknowledged by authors for text polishing for this manuscript.

REFERENCES

- [1] K. Shi, R. Huang, F. Mu, Z. Peng, K. Huang, Y. Qin, X. Yang, and H. Cheng, "A novel multimodal human-exoskeleton interface based on EEG and semg activity for rehabilitation training," in *2022 International Conference on Robotics and Automation, 2022, PA, USA, 2022*. IEEE, 2022, pp. 8076–8082.
- [2] H. J. Choi, S. Das, S. Peng, R. Bajcsy, and N. Figueroa, "On the feasibility of eeg-based motor intention detection for real-time robot assistive control," in *IEEE International Conference on Robotics and Automation, 2024, Yokohama, Japan, May 13-17, 2024*. IEEE, 2024, pp. 5592–5599.
- [3] X. Gao, D. Rong, and Q. He, "Cross-modal alignment between visual stimuli and neural responses in the visual cortex," in *Journal of Physics: Conference Series*, vol. 3147, no. 1. IOP Publishing, 2025, p. 012026.
- [4] D. Rong, H. Dong, X. Gao, J. Wei, D. Hong, Y. Hao, Q. He, and Y. Wang, "Improving unsupervised task-driven models of ventral visual stream via relative position predictivity," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 47, 2025.
- [5] A. Thomas, J. Chen, A. Hella-Szabo, M. A. Kelly, and T. Carlson, "High stimuli virtual reality training for a brain controlled robotic wheelchair," in *IEEE International Conference on Robotics and Automation, 2024, Yokohama, Japan, 2024, 2024*, pp. 11 305–11 311.
- [6] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.
- [7] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. Van Der Smagt, et al., "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, 2012.
- [8] C. Pandarinath, P. Nuyujukian, C. H. Blabe, B. L. Sorice, J. Saab, F. R. Willett, L. R. Hochberg, K. V. Shenoy, and J. M. Henderson, "High performance communication by people with paralysis using an intracortical brain-computer interface," *elife*, vol. 6, p. e18554, 2017.
- [9] J. K. Chapin, "Using multi-neuron population recordings for neural prosthetics," *Nature neuroscience*, vol. 7, no. 5, pp. 452–455, 2004.
- [10] P. G. Patil, J. M. Carmena, M. A. Nicolelis, and D. A. Turner, "Ensemble recordings of human subcortical neurons as a source of motor control signals for a brain-machine interface," *Neurosurgery*, vol. 55, no. 1, pp. 27–38, 2004.
- [11] Y. Li, Y. Qi, Y. Wang, Y. Wang, K. Xu, and G. Pan, "Robust neural decoding by kernel regression with siamese representation learning," *Journal of Neural Engineering*, vol. 18, no. 5, p. 056062, 2021.
- [12] Y. Qi, X. Zhu, K. Xu, F. Ren, H. Jiang, J. Zhu, J. Zhang, G. Pan, and Y. Wang, "Dynamic ensemble bayesian filter for robust control of a human brain-machine interface," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 12, pp. 3825–3835, 2022.
- [13] X. Zhu, Y. Qi, G. Pan, and Y. Wang, "Tracking functional changes in nonstationary signals with evolutionary ensemble bayesian model for robust neural decoding," *Advances in Neural Information Processing Systems*, vol. 35, pp. 22 576–22 588, 2022.
- [14] A. Buttfield, P. W. Ferrez, and J. R. Millan, "Towards a robust bci: error potentials and online learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 164–168, 2006.
- [15] H. Duffau, "Brain plasticity and reorganization before, during, and after glioma resection," *Glioblastoma*, pp. 225–236, 2016.
- [16] A. Farshchian, J. A. Gallego, J. P. Cohen, Y. Bengio, L. E. Miller, and S. A. Solla, "Adversarial domain adaptation for stable brain-machine interfaces," in *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019.
- [17] X. Ma, F. Rizzoglio, K. L. Bodkin, E. Perreault, L. E. Miller, and A. Kennedy, "Using adversarial networks to extend brain computer interface decoding accuracy over time," *elife*, vol. 12, p. e84296, 2023.
- [18] Y. Qi, L. Ding, Y. Wang, and G. Pan, "Learning robust features from nonstationary brain signals by multiscale domain adaptation networks for seizure prediction," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 3, pp. 1208–1216, 2021.
- [19] J. Wei, D. Rong, X. Zhu, Q. He, and Y. Wang, "Speed-enhanced subdomain alignment for long-term stable neural decoding in brain-computer interfaces," in *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2024, pp. 3832–3837.
- [20] S.-H. Yang, C.-J. Huang, and J.-S. Huang, "Increasing robustness of intracortical brain-computer interfaces for recording condition changes via data augmentation," *Computer Methods and Programs in Biomedicine*, vol. 251, p. 108208, 2024.
- [21] J. Jude, M. G. Perich, L. E. Miller, and M. H. Hennig, "Capturing cross-session neural population variability through self-supervised identification of consistent neuron ensembles," in *NeurIPS Workshop on Symmetry and Geometry in Neural Representations*. PMLR, 2023, pp. 234–257.
- [22] J. Wei, D. Rong, D. Hong, Z. Zhang, X. Zhu, Q. He, and Y. Wang, "Complementary-disentangled neural generalization: a robust framework for stable brain-computer interfaces," in *2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2025, pp. 4230–4235.
- [23] S. Schneider, J. H. Lee, and M. W. Mathis, "Learnable latent embeddings for joint behavioural and neural analysis," *Nature*, vol. 617, no. 7960, pp. 360–368, 2023.
- [24] W. Cho, J. Park, and T. Kim, "Complementary domain adaptation and generalization for unsupervised continual domain shift learning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 11 442–11 452.
- [25] K. Zhou, Y. Yang, Y. Qiao, and T. Xiang, "Mixstyle neural networks for domain generalization and adaptation," *International Journal of Computer Vision*, vol. 132, no. 3, pp. 822–836, 2024.
- [26] M. Ghifary, D. Balduzzi, W. B. Kleijn, and M. Zhang, "Scatter component analysis: A unified framework for domain adaptation and domain generalization," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 7, pp. 1414–1430, 2016.
- [27] Y. Zhu, F. Zhuang, J. Wang, G. Ke, J. Chen, J. Bian, H. Xiong, and Q. He, "Deep subdomain adaptation network for image classification," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 4, pp. 1713–1722, 2020.
- [28] Z. Zhong, D. Wang, Q. Zhou, and Y. Lan, "Deep joint subdomain alignment for unsupervised domain adaptation," *Expert Systems with Applications*, vol. 262, p. 125602, 2025.
- [29] P. Wei, Y. Ke, X. Qu, and T.-Y. Leong, "Subdomain adaptation with manifolds discrepancy alignment," *IEEE Transactions on Cybernetics*, vol. 52, no. 11, pp. 11 698–11 708, 2021.
- [30] D.-H. Lee et al., "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Workshop on challenges in representation learning, ICML*, vol. 3, no. 2. Atlanta, 2013, p. 896.
- [31] P. Cascante-Bonilla, F. Tan, Y. Qi, and V. Ordonez, "Curriculum labeling: Revisiting pseudo-labeling for semi-supervised learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, no. 8, 2021, pp. 6912–6920.
- [32] E. Arazo, D. Ortego, P. Albert, N. E. O'Connor, and K. McGuinness, "Pseudo-labeling and confirmation bias in deep semi-supervised learning," in *2020 International joint conference on neural networks (IJCNN)*. IEEE, 2020, pp. 1–8.
- [33] A. B. Ajiboye, F. R. Willett, D. R. Young, W. D. Mernberg, B. A. Murphy, J. P. Miller, B. L. Walter, J. A. Sweet, H. A. Hoyen, M. W. Keith, et al., "Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration," *The Lancet*, vol. 389, no. 10081, pp. 1821–1830, 2017.
- [34] Y. Dong, D. Hu, S. Wang, and J. He, "Decoder calibration framework for intracortical brain-computer interface system via domain adaptation," *Biomedical Signal Processing and Control*, vol. 81, p. 104453, 2023.
- [35] J. Jude, M. Perich, L. Miller, and M. Hennig, "Robust alignment of cross-session recordings of neural population activity by behaviour via unsupervised domain adaptation," in *International Conference on Machine Learning*. PMLR, 2022, pp. 10 462–10 475.
- [36] A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," *Advances in neural information processing systems*, vol. 30, 2017.
- [37] B. M. Karpowicz, J. Ye, C. Fan, P. Tostado-Marcos, F. Rizzoglio, C. Washington, T. Scodeler, D. de Lucena, S. R. Nason-Tomaszewski, M. J. Mender, et al., "Few-shot algorithms for consistent neural decoding (falcon) benchmark," *Advances in Neural Information Processing Systems*, vol. 37, pp. 76 578–76 615, 2024.