

Centroid-Guided Pseudo-Label Refinement and Dual-Level Contrastive Learning for Motor-Execution-to-Motor-Imagery EEG Adaptation

Cailong Yi, Roneel V. Sharan, and Junhua Li*, *Senior Member, IEEE*

Abstract— Electroencephalography (EEG)-based motor imagery (MI) decoding typically relies on labeled MI data for subject-specific calibration, while MI task compliance is difficult to verify objectively, raising concerns about label reliability. In contrast, motor execution (ME) can be directly observed, providing more reliable supervision. However, differences in neural activity patterns and data distributions between ME and MI lead to a cross-task distribution shift and suboptimal direct-transfer performance. To address this issue, we propose Centroid Refinement and Dual-Level Contrastive Learning with a Multi-Scale Depthwise Convolutional Network (CRDCL-MSDCNet), a sequential two-stage framework for subject-specific ME-to-MI adaptation. In Stage 1, labeled ME data from multiple subjects are pooled to train the multi-scale depthwise convolutional network. In Stage 2, subject-specific adaptation uses labeled ME data as the source domain and unlabeled MI data as the target domain. Centroid-Guided pseudo-label refinement iteratively updates class centroids and target assignments, while dual-level contrastive learning enhances the intra-class compactness and inter-class separability of target MI representations through center-level and instance-level constraints. Experimental results show that CRDCL-MSDCNet achieves accuracies of $80.52 \pm 11.55\%$ and $79.56 \pm 8.52\%$ on EEGMMIDB and the lower-limb dataset, respectively. It outperforms the best compared method (i.e., the Joint Adaptation Network) by 0.43 and 0.67 percentage points on the two datasets, respectively. The ablation study demonstrates that performance was reduced by 3.48 and 7.98 percentage points without Stage 2 adaptation and by 2.40 and 2.20 percentage points when only source cross-entropy (Source CE) was used. These results demonstrate effective transfer of discriminative knowledge from labeled ME to unlabeled MI without using MI labels for training or adaptation.

Clinical Relevance— In this study, we proposed the transfer from ME to MI, enabling the trained model using ME labeled data to classify unlabeled MI data. It provides model transferring to facilitate the brain-computer interface (BCI) systems for clinical assistance and rehabilitation applications.

I. INTRODUCTION

Electroencephalography (EEG) signals exhibit substantial inter-subject variability. In practical applications, motor imagery (MI)-based brain-computer interfaces

(MI-BCIs) typically require labeled MI data from the current user for subject-specific model calibration, which increases data acquisition and system preparation time and imposes additional training and calibration burdens, whereas motor execution (ME) generally involves lower subjective task difficulty and fatigue than MI, making it a more practical source of movement-class supervision [1]. Moreover, whether participants consistently perform the instructed MI as required is difficult to verify objectively, whereas ME produces observable movements, allowing experimenters to verify task execution more directly and thus providing more reliable supervisory information [2].

ME and MI share partially similar motor-related neural activity, providing a neurophysiological basis for transferring movement-discriminative information across the two tasks [1]. For users able to perform the corresponding movements, we consider training a subject-specific model on labeled ME data and adapting it with unlabeled MI data to reduce reliance on labeled MI calibration. However, differences in neural activity patterns and data distributions between ME and MI can lead to cross-task distribution shifts [3]. Therefore, effective ME-to-MI transfer requires preserving the supervisory information learned from ME while adapting the model to the target MI distribution.

To address the discrepancy between ME and MI and improve cross-task transfer, existing studies have explored several strategies, which can be broadly categorized into target-label-assisted learning, direct cross-task inference, and explicit domain adaptation. In target-label-assisted learning, Lee et al. proposed a relation-network-based transfer learning method that uses ME data to support MI decoding [4]. Miao et al. further combined ME pretraining, ME-MI alignment, and fine-tuning with partial MI data, demonstrating the feasibility of ME-to-MI transfer on the OpenBMI and GIST datasets [3]. Although these studies confirm that ME data can benefit MI decoding, they still use labeled MI data during model training or fine-tuning.

Another line of research applies ME-trained models directly to MI without target adaptation. Using ME, motor observation, and MI data from 28 subjects, Gwon and Ahn reported that ME-trained models achieved overall performance statistically comparable to within-MI training [1]. Pérez-Velasco et al. further evaluated ME-to-MI inter-task transfer on EEGMMIDB without MI fine-tuning [2]. Although these approaches avoid target MI labels, they cannot exploit the distributional and class-structural information in the current subject's unlabeled MI data.

Explicit domain adaptation provides another strategy. The domain-aware domain-class adaptation network (DDCA Net) reduces cross-task shift through shared feature extraction, domain-specific feature reweighting, maximum mean discrepancy alignment, and dual-classifier adversarial

C. Yi is with the Laboratory for Brain-Bionic Intelligence and Computational Neuroscience, Wuyi University, Jiangmen 529020, China.

R. V. Sharan is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, UK.

J. Li is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, UK, and also with the Laboratory for Brain-Bionic Intelligence and Computational Neuroscience, Wuyi University, Jiangmen 529020, China.

* indicates the corresponding author (Junhua Li, junhua.li@essex.ac.uk).

learning [5]. However, it still requires some ground-truth target MI labels during training. Overall, existing methods either rely on labeled MI data for target adaptation or avoid target labels but do not adapt to the current subject's unlabeled MI data. Therefore, an effective approach should preserve supervision from labeled ME data while exploiting the class structure of unlabeled MI data and improving pseudo-label reliability.

To address this problem, we propose Centroid Refinement and Dual-Level Contrastive Learning (CRDCL) for subject-specific ME-to-MI adaptation. CRDCL is integrated with a Multi-Scale Depthwise Convolutional Network (MSDCNet) to form a sequential two-stage framework (CRDCL-MSDCNet). In Stage 1, MSDCNet is pretrained on labeled ME data pooled from all retained subjects. In Stage 2, the model is adapted to each target subject using that subject's labeled ME data as the source domain and unlabeled MI data as the target domain. To reduce the influence of erroneous initial predictions, probability-weighted class centroids are first used to obtain initial target assignments and are subsequently re-estimated to refine the pseudo-labels. Based on the refined pseudo-labels, dual-level contrastive learning (DCL), comprising a center-level contrastive loss (CCL) and an instance-level contrastive loss (ICL), further exploits local-center and instance-level relationships to enhance the class structure of target MI representations.

The main contributions of this work are summarized as follows:

(1) We propose a sequential two-stage framework for subject-specific ME-to-MI adaptation. The model is first pretrained on labeled ME data pooled from all retained subjects and then adapted to each target subject using subject-specific labeled ME data and unlabeled MI data, without access to ground-truth target MI labels during adaptation.

(2) We introduce CRDCL, which combines centroid-guided pseudo-label refinement with dual-level contrastive learning. CRDCL iteratively re-estimates class centroids to refine target pseudo-labels, while CCL and ICL jointly enhance the intra-class compactness and inter-class separability of target MI representations at the center and instance levels, respectively.

(3) We evaluate CRDCL-MSDCNet on EEGMMIDB and the lower-limb dataset. The proposed method achieves mean classification accuracies of $80.52 \pm 11.55\%$ and $79.56 \pm 8.52\%$, respectively, and consistently outperforms the compared methods and ablation variants on both datasets.

II. METHOD

A. Problem Definition and Overall Framework

For a given subject, let the labeled ME trials constitute the source domain $\mathcal{D}^s$ and the unlabeled MI trials constitute the target domain $\mathcal{D}^t$. Both domains share the binary left/right label space $\mathcal{Y} = \{0,1\}$:

$$\mathcal{D}^s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}, \quad \mathcal{D}^t = \{x_j^t\}_{j=1}^{N_t}, \quad (1)$$

where x_i^s and y_i^s denote the i -th source trial and its label, x_j^t denotes the j -th target trial, and N_s and N_t denote the numbers of source and target trials, respectively.

Although ME and MI share the same class semantics, they correspond to different task states. We characterize their cross-task distribution discrepancy as:

$$P_s(X, Y) \neq P_t(X, Y), \quad (2)$$

where $P_s(X, Y)$ and $P_t(X, Y)$ denote the source-domain and target-domain joint distributions, respectively.

As illustrated in Fig. 1, CRDCL-MSDCNet adopts a sequential two-stage strategy. In Stage 1, MSDCNet is pretrained on pooled labeled ME data from all retained subjects to learn an initial left/right movement representation, and the learned parameters are subsequently used to initialize Stage 2. In Stage 2, subject-specific adaptation is performed using the current subject's labeled ME data for source cross-entropy (Source CE) supervision and unlabeled MI data to construct the target feature bank. Centroid-guided pseudo-label refinement is then applied to obtain reliable target assignments, which are further exploited by CCL and ICL to jointly form the DCL loss. Source and target samples are processed by the same encoder and classifier. Through joint optimization, the framework preserves discriminative supervision from labeled ME data while exploiting the internal class structure of unlabeled MI data to adapt the target representation.

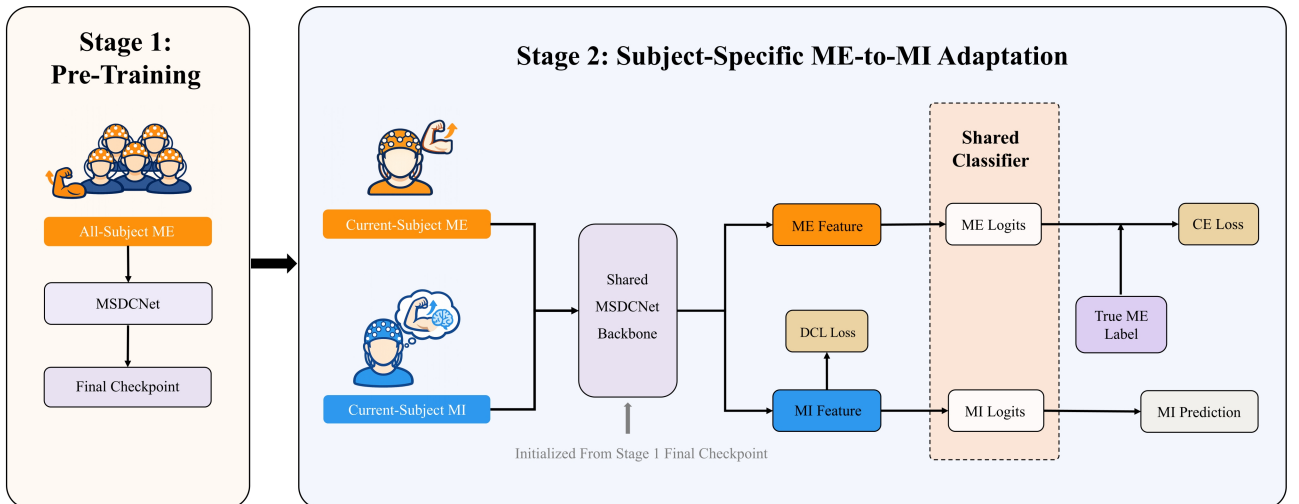


Fig. 1. Two-stage pretraining and subject-specific ME-to-MI adaptation framework of CRDCL-MSDCNet.

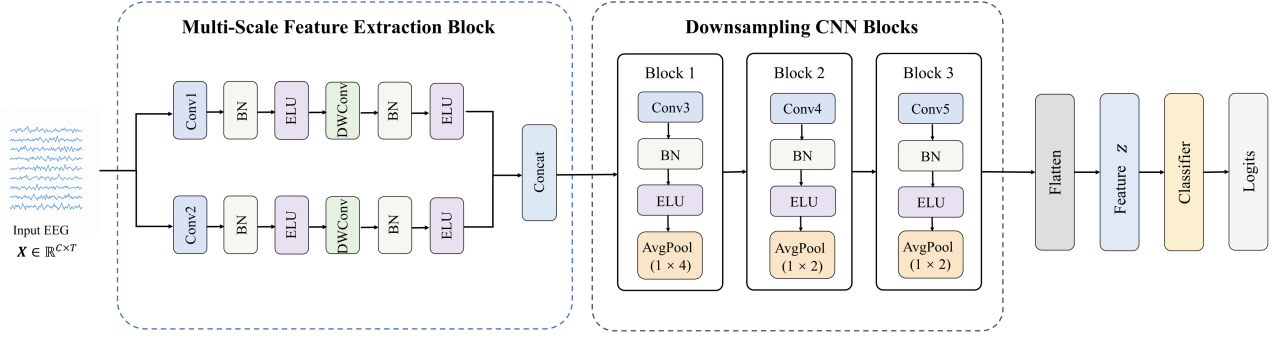


Fig. 2. Architecture of MSDCNet.

B. Multi-Scale Depthwise Convolutional Encoder

Each EEG trial is represented as $X \in \mathbb{R}^{C \times T}$, where C and T denote the number of retained EEG channels and temporal samples per trial, respectively.

MSDCNet employs two parallel temporal-spatial branches at the front end. Temporal convolutions with different kernel sizes are first applied to capture multi-scale temporal patterns, followed by depthwise spatial convolutions spanning all C EEG channels to learn cross-channel representations. The outputs of the two branches are concatenated and subsequently processed by three convolution-pooling blocks for hierarchical feature extraction and progressive downsampling. The resulting features are then flattened and passed to a linear classifier to produce binary classification logits:

$$h = f_{\theta}(X), \quad o = g_{\phi}(h), \quad (3)$$

where $f_{\theta}(\cdot)$, $g_{\phi}(\cdot)$, h , and o denote the feature encoder, linear classifier, latent representation, and binary classification logits, respectively. The network is first optimized using labeled ME data in Stage 1 and is subsequently adapted in Stage 2 using Source CE together with the DCL loss based on the refined target pseudo-labels. The overall architecture of MSDCNet is shown in Fig. 2, and the detailed backbone configuration is summarized in Table I. In Table I, κ_1 and κ_2 denote the front-end temporal kernel lengths, while κ_3 , κ_4 , and κ_5 denote the temporal kernel lengths of Blocks 1–3, respectively. The depthwise spatial kernel height is determined by the number of retained EEG channels C .

TABLE I. OVERVIEW OF MSDCNET BACKBONE

Module	Structure/Operation	Output Ch.	Kernel	Pooling
Front-End Branch 1	Temporal Conv + Depthwise Spatial Conv	16	$1 \times \kappa_1, C \times 1$	—
Front-End Branch 2	Temporal Conv + Depthwise Spatial Conv	16	$1 \times \kappa_2, C \times 1$	—
Multi-Scale Fusion	Channel Concatenation	32	—	—
Block 1	Temporal Conv + BN + ELU + Dropout	16	$1 \times \kappa_3$	1×4
Block 2	Temporal Conv + BN + ELU + Dropout	8	$1 \times \kappa_4$	1×2
Block 3	Temporal Conv + BN + ELU + Dropout	4	$1 \times \kappa_5$	1×2
Output Layer	Flatten + Linear	2	—	—

C. Centroid-Guided Pseudo-Label Refinement

Following the Stage 2 warm-up, the current model performs a forward pass over all unlabeled MI trials of the target subject at the beginning of each adaptation epoch to construct a global target feature bank. Let h_i denote the latent representation of the i -th target sample, and define its L_2 -normalized feature as $z_i = \text{Norm}(h_i)$. Let o_i denote the corresponding classification logits. The class probability $p_{i,c}$ is obtained by applying softmax to o_i , and the maximum predicted probability q_i is defined as the prediction confidence:

$$p_{i,c} = \text{Softmax}(o_i)_c, \quad q_i = \max_{c \in \mathcal{Y}} p_{i,c}. \quad (4)$$

Centroid-based pseudo-label refinement has been used to improve target pseudo-labels in unsupervised domain adaptation [6]. As illustrated in Fig. 3, our refinement procedure first initializes class centroids using probability-weighted target features and obtains initial pseudo-label assignments. For each class c , $p_{i,c}$ is used to weight the contribution of each target sample, and the aggregated feature is normalized to obtain the initial class centroid $\mu_c^{(0)}$.

$$\mu_c^{(0)} = \text{Norm} \left(\frac{\sum_{i=1}^{N_t} p_{i,c} z_i}{\sum_{i=1}^{N_t} p_{i,c}} \right). \quad (5)$$

Each target feature is then assigned to the class whose centroid has the highest cosine similarity, yielding the initial pseudo-label $\hat{y}_i^{(0)}$:

$$\hat{y}_i^{(0)} = \underset{c \in \mathcal{Y}}{\text{argmax}} z_i^{\top} \mu_c^{(0)}. \quad (6)$$

Based on these initial pseudo-label assignments, the class centroids are subsequently re-estimated. For each class c , the number of target samples initially assigned to that class is computed as

$$M_c = \sum_{i=1}^{N_t} \mathbb{I}(\hat{y}_i^{(0)} = c), \quad (7)$$

and the corresponding target features are averaged and normalized to obtain the refined centroid $\mu_c^{(1)}$:

$$\mu_c^{(1)} = \text{Norm} \left(\frac{\sum_{i=1}^{N_t} \mathbb{I}(\hat{y}_i^{(0)} = c) z_i}{M_c} \right). \quad (8)$$

Each target sample is then reassigned to the class with the highest cosine similarity to the refined centroids, producing the final pseudo-label $\hat{y}_i$:

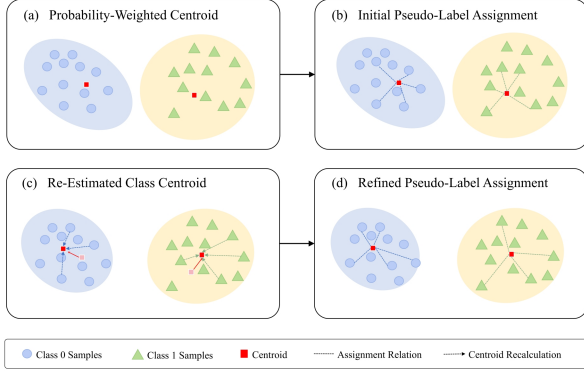


Fig. 3. Centroid-Guided pseudo-label refinement procedure.

$$\hat{y}_i = \operatorname{argmax}_{c \in \mathcal{Y}} z_i^T \mu_c^{(l)}, \quad (9)$$

where $i \in \{1, \dots, N_t\}$ indexes target MI samples. N_t is the target-sample count. $c \in \mathcal{Y}$, with $\mathcal{Y} = \{0, 1\}$, specifies the class index. $p_{i,c}$ and q_i are the class probability and prediction confidence, respectively. $\mu_c^{(0)}$ and $\mu_c^{(l)}$ are the

probability-weighted and re-estimated class centroids, respectively. $\hat{y}_i^{(0)}$ and $\hat{y}_i$ are the initial and refined pseudo-labels, respectively. M_c is the number of samples initially assigned to class c , $\mathbb{I}(\cdot)$ is the indicator function, and $\text{Norm}(\cdot)$ denotes L_2 normalization. If $M_c = 0$, we set $\mu_c^{(l)} = \mu_c^{(0)}$.

Because both the target features and class centroids are L_2 -normalized, the inner products in (6) and (9) are equivalent to cosine similarities. Rather than directly taking the argmax of the classifier softmax outputs, the proposed refinement first constructs probability-weighted class centroids and then re-estimates them based on the target feature geometry after the initial assignment. The subsequent reassignment therefore incorporates both classifier predictions and the intrinsic structure of the target feature space. This refinement procedure is repeated at every adaptation epoch, allowing the class centroids and pseudo-label assignments to evolve with the target feature space. The resulting refined pseudo-label $\hat{y}_i$ is subsequently used for same-pseudo-class neighbor retrieval and for constructing center-level and instance-level contrastive relationships in Section II-D.

D. Dual-Level Contrastive Learning

After pseudo-label refinement, the DCL component constructs center-level and instance-level keys from the target feature bank. For each anchor z_i , up to K nearest neighbors sharing the same refined pseudo-label are retrieved according to cosine similarity:

$$\mathcal{N}_i = \operatorname{TopK}_{j \neq i, \hat{y}_j = \hat{y}_i} (z_i^T z_j). \quad (10)$$

The normalized mean of these neighbor features is used as the local neighborhood center c_i when at least one same-pseudo-class neighbor is available:

$$c_i = \text{Norm} \left(\frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} z_j \right), \quad K_i > 0, \quad (11)$$

where $\mathcal{N}_i$ denotes the index set returned by TopK in (10), containing same-pseudo-class nearest neighbors of anchor z_i , $K_i = |\mathcal{N}_i|$, and K is the maximum number of retrieved neighbors. When $K_i > 0$, c_i is the L_2 -normalized mean of features in $\mathcal{N}_i$. When $K_i = 0$, no local neighborhood center

is constructed for anchor z_i , and the anchor is excluded from CCL.

The set of valid local neighborhood centers forms the center-level key set $\mathcal{K}_C$, while the original target features constitute the instance-level key set $\mathcal{K}_I$. The former captures local same-pseudo-class neighborhood structure, whereas the latter preserves sample-level information. These two key sets are subsequently used by CCL and ICL, respectively.

$$\mathcal{K}_C = \{c_j: K_j > 0\}, \quad \mathcal{K}_I = \{z_j\}_{j=1}^{N_t}. \quad (12)$$

The contrastive objective adopts a standard multi-positive formulation [7], with positive sets defined by refined pseudo-labels for both views. CCL operates on the valid center-level key set $\mathcal{K}_C$, whereas ICL uses the original target features in $\mathcal{K}_I$ as keys. Only anchors with at least one valid same-pseudo-label neighbor are included in CCL. For ICL, the anchor itself is excluded from both the positive set and the normalization denominator. The key probability and the multi-positive contrastive loss for a single anchor are defined as:

$$\rho_{i,j}^{(v)} = \frac{\exp(z_i^T k_j^{(v)}/\tau)}{\sum_{a \in \mathcal{A}_i^{(v)}} \exp(z_i^T k_a^{(v)}/\tau)}, \quad (13)$$

$$\ell_i^{(v)} = - \sum_{j \in \mathcal{P}_i^{(v)}} \frac{q_j}{\sum_{p \in \mathcal{P}_i^{(v)}} q_p} \log \rho_{i,j}^{(v)}, \quad (14)$$

where $v \in \{\text{center, instance}\}$ denotes the key view. $k_j^{(v)}$ denotes the j -th key in view v . $\mathcal{A}_i^{(v)}$ is the denominator key set used for softmax normalization of anchor i . $\mathcal{P}_i^{(v)} \subseteq \mathcal{A}_i^{(v)}$ is the positive-key subset sharing the same refined pseudo-label as the anchor. τ is the temperature coefficient. $\rho_{i,j}^{(v)}$ is the normalized similarity probability from anchor z_i to key $k_j^{(v)}$. $\ell_i^{(v)}$ is the corresponding single-anchor multi-positive contrastive loss. q_j is the prediction confidence associated with target sample j . For a center key c_j , q_j and $\hat{y}_j$ are inherited from the target sample indexed by j for loss weighting and positive-set construction, respectively.

Further, let $\mathcal{V}^{(v)}$ denote the set of anchors with valid positive keys in view v . The single-anchor losses are aggregated using prediction confidence q_i :

$$\mathcal{L}_v = \frac{\sum_{i \in \mathcal{V}^{(v)}} q_i \ell_i^{(v)}}{\sum_{i \in \mathcal{V}^{(v)}} q_i}, \quad (15)$$

where q_i is the prediction confidence of anchor i , and $\mathcal{L}_v$ is the confidence-weighted contrastive loss for view v . CCL and ICL are summed to obtain the DCL loss:

$$\mathcal{L}_{\text{DCL}} = \mathcal{L}_{\text{CCL}} + \mathcal{L}_{\text{ICL}}. \quad (16)$$

CCL captures the aggregated structure of local neighborhood centers, whereas ICL preserves fine-grained relationships among the original target instances. Key and anchor prediction confidences weight their respective contrastive contributions. By integrating these complementary constraints, the DCL loss jointly enhances neighborhood-level structure and instance-level discriminability.

E. Overall Optimization Objective

Stage 2 begins with a supervised source-domain warm-up. During epochs 1–20, only the labeled ME data from the current subject are used to optimize the label-smoothed cross-entropy loss, referred to as Source CE. For the label space $\mathcal{Y}$, the smoothed source label is defined as:

$$\tilde{y}_{i,c}^s = (1 - \varepsilon)y_{i,c}^s + \frac{\varepsilon}{|\mathcal{Y}| - 1}(1 - y_{i,c}^s), \quad (17)$$

where $\tilde{y}_{i,c}^s$ is the smoothed source label, $y_{i,c}^s$ is the corresponding one-hot label, and ε is the label-smoothing coefficient, set to 0.05 in all experiments. Source CE is computed between the smoothed source-label distribution and the source-class probabilities obtained from the classifier softmax outputs.

After the warm-up stage, the DCL loss is introduced from epoch 21 onward and jointly optimized with Source CE. The overall Stage 2 objective is defined as:

$$\mathcal{L}_{\text{total}}(e) = \mathcal{L}_{\text{SourceCE}} + \lambda_t(e)\mathcal{L}_{\text{DCL}}, \quad (18)$$

where $\lambda_t(e)$ denotes the epoch-dependent weight of the DCL loss at epoch e . During warm-up, $\lambda_t(e) = 0$. After warm-up, the weight is gradually increased using a cosine ramp:

$$\lambda_t(e) = \begin{cases} 0, & e \leq E_w, \\ \frac{1}{2}\lambda_{\text{max}}[1 - \cos(\pi r(e))], & e > E_w. \end{cases} \quad (19)$$

where $r(e) = \min\left(\frac{e - E_w}{E_r}, 1\right)$ is the normalized post-warm-up progress. E_w , E_r , and λ_{max} denote the warm-up length, ramp duration, and maximum weight of the DCL loss, respectively.

III. EXPERIMENTS

A. Datasets and Preprocessing

EEGMMIDB: The upper-limb experiments are conducted on the PhysioNet EEG Motor Movement/Imagery Dataset (EEGMMIDB) [8]–[10], which contains recordings from 109 subjects collected using the BCI2000 system with 64 EEG channels sampled at 160 Hz. Only the binary left-fist/right-fist tasks are retained in this study. Left- and right-fist motor execution trials are used as the ME source domain, while the corresponding motor imagery trials constitute the MI target domain. A 4-s task interval is extracted from each trial, resulting in 640 temporal samples. Subjects are further screened according to predefined criteria on trial length, trial count, and data completeness. After screening, 103 subjects are retained, with 45 ME and 45 MI trials per subject.

Lower-limb dataset: The lower-limb experiments use the dataset described in [11], with the experimental protocol detailed in [12], which contains recordings from 28 subjects performing six lower-limb actions. In this study, only the left/right real movement and motor imagery conditions are considered. For each subject, RM-L and RM-R are treated as the ME source domain, with 144 trials in total, while MI-L and MI-R constitute the MI target domain, also with 144 trials. A total of 62 EEG channels are retained. For each trial, a 6-s task interval is extracted, resulting in 1500 temporal samples at a sampling rate of 250 Hz. The data collection protocol was reviewed and approved by the Institutional Review Board of the National University of

Singapore and the Humanities, Science and Health, or Social Science Ethics Sub-Committee at the University of Essex. Written informed consent was obtained from all participants.

The same preprocessing pipeline is applied to both datasets. The retained EEG channels are first average-referenced, followed by zero-phase finite impulse response (FIR) band-pass filtering between 8 and 30 Hz. The signals are then segmented according to the predefined task intervals. Finally, each ME and MI trial is independently z-score normalized along the temporal dimension for each EEG channel.

B. Implementation Details and Evaluation Protocol

The MSDCNet configuration follows Table I. For EEGMMIDB, the temporal kernel lengths are set to $(\kappa_1, \kappa_2) = (40, 20)$ and $(\kappa_3, \kappa_4, \kappa_5) = (10, 5, 3)$, whereas the corresponding values for the lower-limb dataset are $(64, 32)$ and $(16, 8, 4)$. Pooling factors are fixed at $(4, 2, 2)$ for both datasets. All experiments use PyTorch and Adam. Stage 1 uses an initial learning rate of 1×10^{-3} , while Stage 2 uses 1×10^{-4} . Both use a weight decay of 3×10^{-4} , $\beta_1 = 0.9$, $\beta_2 = 0.99$, label smoothing $\varepsilon = 0.05$, gradient clipping of 5.0, and seed 42. Stage 2 loads the final Stage 1 checkpoint at epoch 30, with batch sizes of 15 and 32 for EEGMMIDB and the lower-limb dataset, respectively. Following the schedule in (19), we set $E_w = 20$, $E_r = 50$, and $\lambda_{\text{max}} = 0.25$. We set $K = 5$ and $\tau = 0.20$.

Ground-truth target MI labels are not used for training, checkpoint selection, or early stopping. The reported accuracies are obtained after model training for 300 epochs on EEGMMIDB and 200 epochs on the lower-limb dataset, with target labels used only for final evaluation. Results are summarized as mean $\pm$ standard deviation across subjects.

IV. RESULTS AND DISCUSSION

A. Overall Results

To evaluate the effectiveness of the proposed method, five representative domain adaptation methods are selected for comparison: Joint Adaptation Network (JAN) [13], Local Maximum Mean Discrepancy (LMMD) [14], Prototype-oriented Conditional Transport (PCT) [15], Maximum Mean Discrepancy (MMD) [16], and Deep CORAL [17]. These methods represent joint-distribution alignment, class-aware subdomain alignment, prototype-based conditional alignment, global distribution alignment, and covariance alignment, respectively. For a fair comparison, all methods use MSDCNet as the shared backbone under identical preprocessing, optimization, and evaluation settings. The classification results of the proposed method and the compared methods on the two datasets are summarized in Table II.

TABLE II. COMPARISON ON EEGMMIDB AND LOWER-LIMB DATASET

Method	EEGMMIDB Acc. (%)	Lower-Limb Dataset Acc. (%)
CRDCL-MSDCNet (Ours)	80.52 $\pm$ 11.55	79.56 $\pm$ 8.52
JAN	80.09 $\pm$ 11.35	78.89 $\pm$ 8.78
LMMD	79.31 $\pm$ 11.73	78.60 $\pm$ 8.63
PCT	79.20 $\pm$ 11.84	78.32 $\pm$ 8.32
MMD	78.86 $\pm$ 11.71	78.10 $\pm$ 8.58
Deep CORAL	78.06 $\pm$ 11.64	77.38 $\pm$ 8.22

As shown in Table II, CRDCL-MSDCNet achieves the highest mean classification accuracies on both datasets, reaching $80.52 \pm 11.55\%$ on EEGMMIDB and $79.56 \pm 8.52\%$ on the lower-limb dataset. Compared with JAN, the proposed method improves the mean accuracy by 0.43 and 0.67 percentage points, respectively, and also consistently outperforms LMMD, PCT, MMD, and Deep CORAL.

The results suggest that, for ME-to-MI adaptation, exploiting the class structure of unlabeled MI data through pseudo-label refinement and dual-level contrastive learning may be more effective than relying on alignment strategies alone.

B. Ablation Study

To assess the contribution of each component, four ablation settings are considered: w/o Pseudo-label Refinement, w/o CCL, w/o ICL, and w/o DCL. The first replaces the centroid-refined pseudo-labels with classifier argmax predictions, while w/o CCL and w/o ICL remove CCL and ICL, respectively. w/o DCL removes both CCL and ICL, leaving no target-domain loss. Thus, Stage 2 is optimized using Source CE only. In addition, Stage 1 Direct Inference is included as a reference and denotes the classification performance without Stage 2 adaptation. The corresponding results are summarized in Table III.

TABLE III. ABLATION RESULTS OF CRDCL-MSDCNet

Setting	EEGMMIDB Acc. (%)	Lower-Limb Dataset Acc. (%)
Stage 1 Direct Inference	77.04 ± 11.67	71.58 ± 8.22
CRDCL-MSDCNet (Full)	80.52 ± 11.55	79.56 ± 8.52
w/o Pseudo-Label Refinement	79.78 ± 11.12	79.44 ± 8.66
w/o CCL	80.09 ± 11.77	78.65 ± 8.45
w/o ICL	80.26 ± 11.45	79.19 ± 8.52
w/o DCL (Source CE Only)	78.12 ± 11.58	77.36 ± 8.26

Table III shows that Source CE only improves Stage 1 Direct Inference by 1.08 and 5.78 percentage points on EEGMMIDB and the lower-limb dataset, respectively. Adding the DCL loss based on refined target pseudo-labels further increases the accuracy by 2.40 and 2.20 points, yielding total gains of 3.48 and 7.98 points over Stage 1 Direct Inference. These results suggest that subject-specific ME supervision provides a useful basis, while target-domain structural learning further improves ME-to-MI adaptation.

Removing any component reduces performance on both datasets, suggesting that pseudo-label refinement, CCL, and ICL all contribute to the overall performance. The larger drops for pseudo-label refinement on EEGMMIDB and CCL on the lower-limb dataset indicate dataset-dependent effects.

V. CONCLUSION

This study proposes CRDCL-MSDCNet for subject-specific ME-to-MI adaptation with unlabeled target MI data. The framework first pretrains MSDCNet on pooled labeled ME data and then adapts it to each subject using subject-specific labeled ME and unlabeled MI data. Centroid-Guided pseudo-label refinement improves target assignments, while dual-level contrastive learning strengthens the class structure of target MI representations through center-level and instance-level constraints.

Experiments on EEGMMIDB and the lower-limb dataset, together with comparative and ablation results, support the effectiveness of the proposed components. By transferring supervision from observable ME to unlabeled MI, CRDCL-MSDCNet reduces reliance on labeled MI calibration. Future work will investigate adaptation under varying levels of target-data availability and further evaluate generalization across different motor tasks and EEG datasets.

REFERENCES

- [1] D. Gwon and M. Ahn, "Motor task-to-task transfer learning for motor imagery brain-computer interfaces," *NeuroImage*, vol. 302, Art. no. 120906, Nov. 2024.
- [2] S. Perez-Velasco, D. Marcos-Martinez, E. Santamaria-Vazquez, V. Martinez-Cagigal, and R. Hornero, "Bridging motor execution and motor imagery BCI paradigms: An inter-task transfer learning approach," *Biomed. Signal Process. Control*, vol. 107, Art. no. 107834, 2025.
- [3] M. Miao, Z. Yang, H. Zeng, W. Zhang, B. Xu, and W. Hu, "Explainable cross-task adaptive transfer learning for motor imagery EEG classification," *J. Neural Eng.*, vol. 20, no. 6, Art. no. 066021, Nov. 2023.
- [4] D.-Y. Lee, J.-H. Jeong, K.-H. Shim, and S.-W. Lee, "Decoding movement imagination and execution from EEG signals using BCI-transfer learning method based on relation network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, 2020, pp. 1354–1358.
- [5] J. Wang, G. Xu, C. Du, Z. Li, H. Li, S. Chen, C. Han, and S. Zhang, "Domain-aware domain-class adaptation network for motor execution to motor imagery EEG classification," *Front. Neurosci.*, vol. 20, Art. no. 1851006, Jun. 2026.
- [6] J. Liang, D. Hu, and J. Feng, "Do we really need to access the source data? Source hypothesis transfer for unsupervised domain adaptation," in *Proc. 37th Int. Conf. Mach. Learn. (ICML)*, vol. 119, pp. 6028–6039, 2020.
- [7] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, pp. 18661–18673, 2020.
- [8] G. Schalk, "EEG Motor Movement/Imagery Dataset," PhysioNet, version 1.0.0, 2009.
- [9] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, 2004.
- [10] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.
- [11] H. Yan, Z. Wang, and J. Li, "MSC-transformer-based 3D-attention with knowledge distillation for multi-action classification of separate lower limbs," *Neural Netw.*, vol. 191, Art. no. 107806, Nov. 2025.
- [12] Z. Wang, J. Li, I. Daly, and J. Li, "Machine learning for multi-action classification of lower limbs for BCI," in *Proc. 2022 Int. Conf. Comput., Electron. Commun. Eng. (iCCECE)*, 2022, pp. 84–89.
- [13] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Deep transfer learning with joint adaptation networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, vol. 70, pp. 2208–2217, 2017.
- [14] 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 Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1713–1722, Apr. 2021.
- [15] K. Tanwisuth, X. Fan, H. Zheng, S. Zhang, H. Zhang, B. Chen, and M. Zhou, "A prototype-oriented framework for unsupervised domain adaptation," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, pp. 17194–17208, 2021.
- [16] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)*, vol. 37, pp. 97–105, 2015.
- [17] B. Sun and K. Saenko, "Deep CORAL: Correlation alignment for deep domain adaptation," in *Proc. Eur. Conf. Comput. Vis. Workshops (ECCVW)*, 2016, pp. 443–450.