A Semi-Supervised Progressive Learning Algorithm for Brain–Computer Interface

Yuxuan Wei, Jie Li, Hongfei Ji, Lingjing Jin, Lingyu Liu, Zhongfei Bai, and Chen Ye, *Member, IEEE*

***Abstract*— Brain-computer interface (BCI) usually suffers from the problem of low recognition accuracy and large calibration time, especially when identifying motor imagery tasks for subjects with indistinct features and classifying fine grained motion control tasks by electroencephalogram (EEG)-electromyogram (EMG) fusion analysis. To fill the research gap, this paper presents an end-to-end semi-supervised learning framework for EEG classification and EEG–EMG fusion analysis. Benefiting from the proposed metric learning based label estimation strategy, sampling criterion and progressive learning scheme, the proposed framework efficiently extracts distinctive feature embedding from the unlabeled EEG samples and achieves a 5.40% improvement on BCI Competition IV Dataset IIa with 80% unlabeled samples and an average 3.35% improvement on two public BCI datasets. By employing synchronous EMG features as pseudo labels for the unlabeled EEG samples, the proposed framework further extracts deep level features of the synergistic complementarity between the EEG signals and EMG features based on the deep encoders, which improves the performance of hybrid BCI (with a 5.53% improvement for the Upper Limb Motion Dataset and an average 4.34% improvement on two hybrid datasets). Moreover, the ablation experiments show that the proposed framework can substantially improve the performance of the deep encoders (with an average 5.53% improvement). The proposed framework not only largely improves the performance of deep networks in the BCI system, but also significantly reduces the calibration time for EEG-EMG fusion analysis, which shows great potential for building an efficient and high-performance hybrid BCI for the motor rehabilitation process.**

***Index Terms*— EEG classification, EEG-EMG fusion, progressive learning, semi-supervised learning, brain computer interface.**

Manuscript received 21 April 2022; revised 25 June 2022; accepted 15 July 2022. Date of publication 19 July 2022; date of current version 28 July 2022. This work was supported in part by the Shanghai Municipal Science and Technology Major Project under Grant 2021SHZDZX0100, in part by the Fundamental Research Funds for the Central Universities, and in part by the Science and Technology Innovation Action Plan of Shanghai Science and Technology Commission under Grant 19441908000. *(Corresponding authors: Jie Li; Hongfei Ji; Lingjing Jin.)*

Yuxuan Wei, Jie Li, Hongfei Ji, and Chen Ye are with the Department of Computer Science and Technology, Shanghai YangZhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Electronic and Information Engineering, Tongji University, Shanghai 201804, China (e-mail: 2030777@tongji.edu.cn; nijanice@163.com; jhf@tongji.edu.cn; yechen@tongji.edu.cn).

Lingjing Jin, Lingyu Liu, and Zhongfei Bai are with the Department of Neurorehabilitation, Shanghai YangZhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Medicine, Tongji University, Shanghai 201619, China (e-mail: lingjingjin@163.com; happyneurologist@163.com; bzf567@163.com).

Digital Object Identifier 10.1109/TNSRE.2022.3192448

I. Introduction

BRAIN Computer Interface (BCI) is a new type of intelligent sensing system that uses advanced machine learning and pattern recognition algorithms to recognize signals from different brain activities and translate them into control commands to directly control external devices [1]. BCI is of great research importance and application value in fields like medicine, industry, entertainment, and especially rehabilitation medicine [2].

Non-invasive scalp-based electroencephalogram (EEG) [3] signal is the most widely used signal in BCI. The motor imagery (MI) task is one of the most common tasks in spontaneous EEG-based BCI systems, which can activate the related motor execution (ME) cortex of the brain, by mentally sensing an action process (kinesthetic motor imagery) or by imagining a motor process in the mind (visual motor imagery) [4]. The MI-based BCI system recognizes the task of imagining the movement of different parts of the limb, and the recognition results are fed back to the subjects in real time through various forms of visual stimulation, vibratory stimulation, sensory stimulation, and exoskeleton. Patients with limb dysfunction due to illnesses such as stroke can reestablish a neural circuit from motor intention to motor execution to perceptual feedback by actively performing the MI-based BCI interaction process, which helps patients remodel their brain function and restore their control ability of limb motor [5]. Therefore, MI-based BCI systems are widely used in the motor rehabilitation of patients with limb dysfunction.

However, the generation of motor imagery features is influenced by the cognitive and motor-sensory abilities of the subject [6]. Due to the decline of motor sensory [7], motor imagery features from patients with limb dysfunction are usually less distinctive compared to healthy subjects, and their labels for the motor imagery task are also likely to be mismatched, resulting in low recognition accuracy of the BCI system. Moreover, these large numbers of samples with indistinct features and inaccurate labels are useless for supervised learning-based BCI systems, which can lead to long-term recognition errors and cause great frustration to patients, possibly leading to physical and mental exhaustion. Most of the existing EEG-based BCI systems do not perform well on semi-supervised learning tasks, and there is room for improvement. Therefore, how to build stable and reliable classification models for MI tasks with indistinct features and inaccurate labels to reduce the calibration time of the BCI system is a major problem.

In this paper, to better exploit the large number of unlabeled EEG samples with indistinct features, we propose a progressive semi-supervised learning strategy. State-of-the-art deep learning models for BCI systems are trained through three steps to achieve excellent performance. In the first step, the model uses labeled samples for supervised learning. In the second step, the pre-trained model generates pseudo labels [8] for the unlabeled samples based on the proposed label estimation strategy, and then some of the unlabeled samples with reliable pseudo labels are added to the training dataset based on the proposed sampling criterion. Then the pre-trained model continues training using the increased training dataset based on semi-supervised loss with label smoothing [9]. In the third step, all unlabeled samples with their pseudo labels are added to the training dataset, and the pre-trained model is trained with all the labeled and unlabeled samples in a semi-supervised fashion.

Besides, during the motor rehabilitation process, patients usually have very limited motor abilities in the early stages of rehabilitation, and motor imagery tasks can play an important role at this time [10]. With repeated exercises, the function of the patients' muscles can be partly improved. At this time, if the patients still rely on the BCI for motor imagery tasks instead of real motor control execution, it is not conducive to the further improvement of the patients' motor ability [5]. Many researchers have suggested to gradually reduce the dependence on motor imagery tasks in the BCI rehabilitation system according to the actual rehabilitation situation of patients, and encourage patients to perform real motor control execution at the same time [11]. Moreover, it is more meaningful to perceive and identify more refined and complex movements of the limb with dysfunction for BCI systems that concentrate on motor rehabilitation. However, the extremely low spatial resolution of EEG leads to the problem with the EEG-based BCI system's poor discrimination ability of refined limb movements. Although several research attempts have been made to decode the refined movements of limbs through decoding EEG signals [12]–[14], no satisfactory classification results have been achieved. Prior studies have shown that the primary motor sensor cortex of the human brain has a wide range of 15-35 Hz neural oscillatory activity [15], through which the motor nervous system transmits motor control information and causes synchronous oscillatory activity in motor units. This synchronous oscillatory activity reveals the coupled connectivity between the cerebral cortex and muscles during movement [16]. Therefore, in recent years, electromyography (EMG) [17], [18] have been added into the hybrid brain-computer interfaces [19]–[22] to acquire a more comprehensive and accurate perception of fine-grained motion control tasks. However, most of the existing hybrid BCI systems simply integrate the analysis results of EEG and EMG, which usually requires accurate labels and significant computational calibration time to calculate the correlation between each pair of EEG and EMG channels, while ignoring the synergistic complementarity between EEG and EMG.

The main contributions of this paper are as follows:

1) It proposes a novel progressive semi-supervised learning strategy for MI-based BCI systems to improve the utilization of unlabeled EEG samples with indistinct features. The proposed method adopts a stepwise training strategy that starts with labeled samples and gradually includes unlabeled samples. Through applying the label estimation strategy and sampling criterion to the semi-supervised learning task, models can efficiently learn effective feature representations of unlabeled EEG samples, and further improve their performance. In principle, the proposed training strategy can be applied to almost all deep learning based BCI systems.

2) It presents a new semi-supervised learning framework for EEG and EMG fusion analysis in hybrid brain-computer interfaces. By employing synchronous EMG features as pseudo labels for the unlabeled EEG samples, the BCI system simulates and analyzes the synchronous oscillatory activity between cortical neural oscillations and muscle motor units during motion execution, which significantly improved its ability to perceive fine-grained and complex movements. Compared to traditional EEG-EMG fusion analysis methods in hybrid BCI, the deep learning-based framework largely reduced the time cost for data preparation, calibration, and comparison.

3) It combines the majority of cutting-edge deep learning models with progressive learning strategies on semi-supervised learning tasks for hybrid brain-computer interfaces in the end-to-end framework. The proposed framework improves the performance of most deep networks on semi-supervised learning tasks for different BCI datasets, which shows great potential for building an efficient and high-performance hybrid BCI for the motor rehabilitation process of patients based on EEG–EMG fusion analysis.

The remainder of the paper is organized as follows. Section II introduces related work. Section III details the proposed methodology. Section IV presents the experimental results, and discusses the effectiveness as well as limitations of the model. Section V concludes the paper and suggests potential future studies.

II. Related Work

In this section, we briefly introduce some state-of-the-art methods related to our approach.

A. Supervised Learning Method

Traditional machine learning models in motor imagery tasks usually contain two stages: feature extraction and feature classification. By extracting spatial and temporal patterns of the EEG signal with models like Common Spatial Patterns (CSP) [23], Filter-Bank Common Spatial Pattern (FBCSP) [24], Extreme Learning Machine (ELM) [25], Broad Learning System (BLS) [26] and Internal Feature Selection Method [27], and classifying extracted features with models like Support Vector Machine (SVM) [28] and Linear Discriminant Analysis (LDA) [28], these two-stage models obtained superior accuracy on the motor imagery task. In recent years, deep learning models have brought huge improvements in the performance of MI-based BCI [29]. Deep learning models, as opposed to traditional machine learning models, perform end-to-end learning for raw EEG data, allowing them to extract intricate, high-dimensional features from

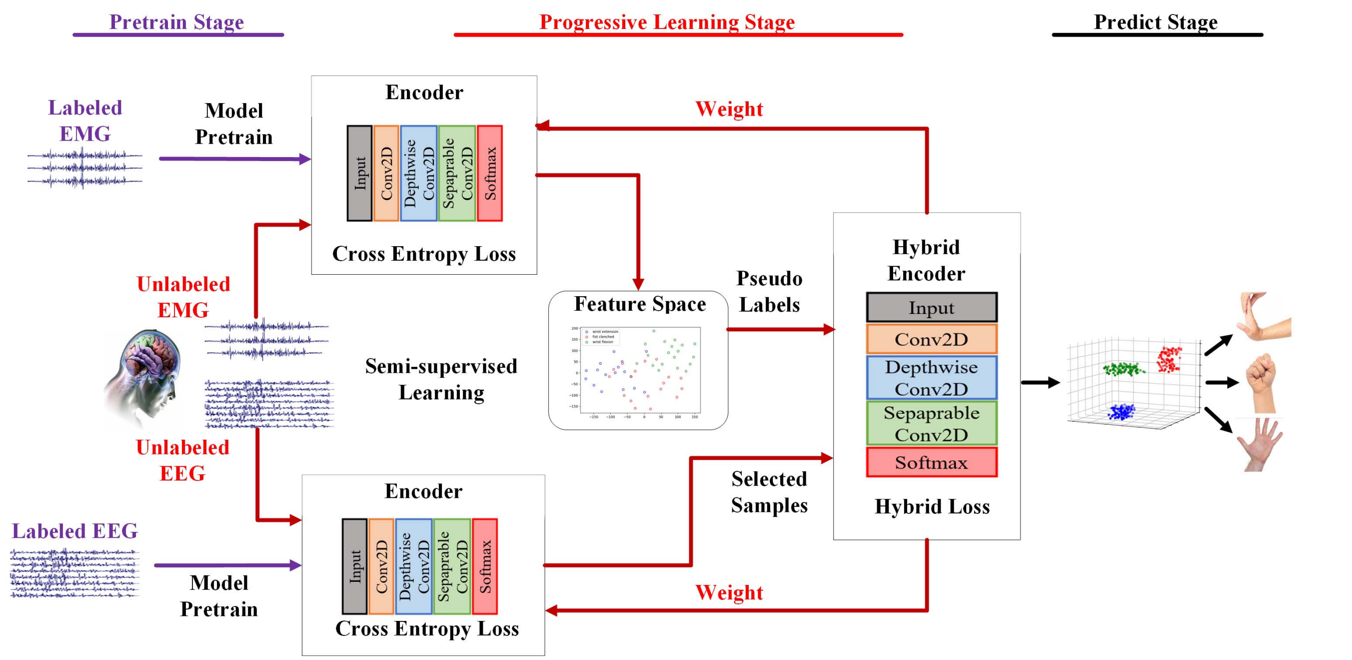


Fig. 1. Framework of the proposed method.

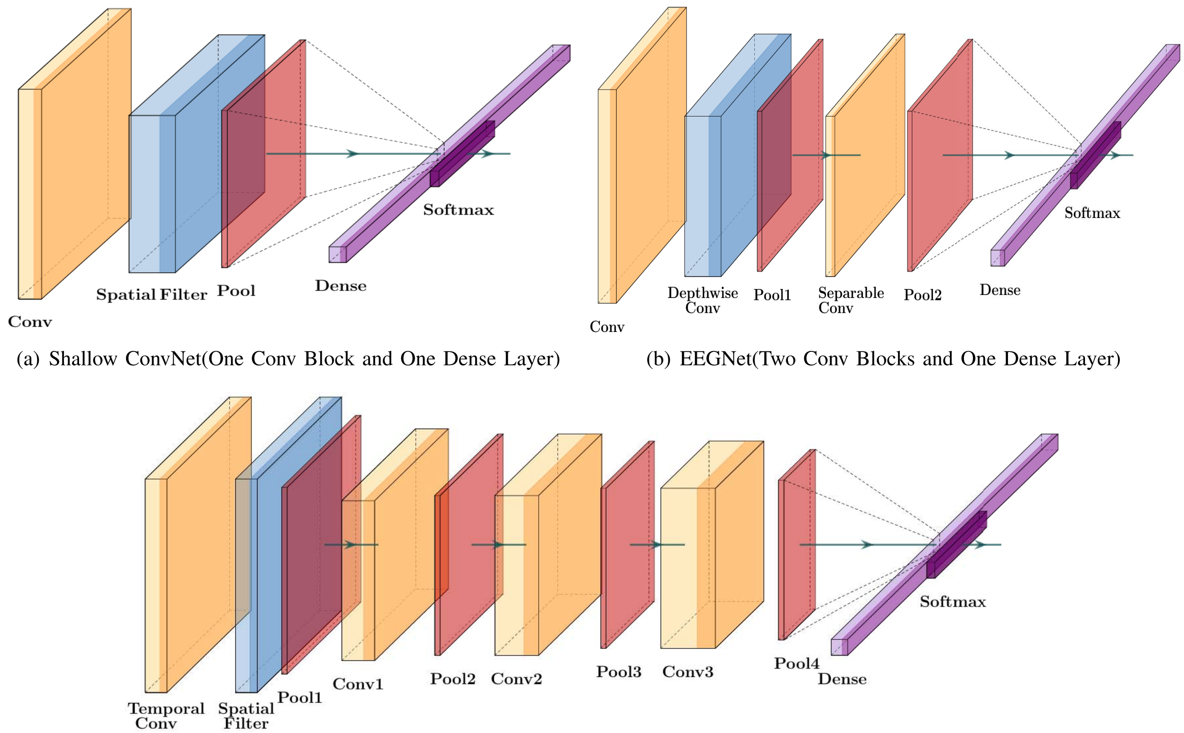
large amounts of samples. By designing appropriate temporal and spatial convolution layers, Convolutional Neural Networks (CNN) achieved excellent results on supervised learning tasks for motor imagery [30], [31]. However, supervised learning-based models require accurate labels for each train sample, while for large amounts of raw EEG data, performing signal sample segmentation and label acquisition is often difficult.

B. Semi-Supervised Learning Method

Semi-supervised learning aims to build suitable machine learning models to learn both labeled and unlabeled datasets, and to improve the performance of the model on both supervised and unsupervised learning tasks [32]. In the field of collaborative learning, the Squared-loss Mutual Information Regularization (SMIR) model [33], Semi-supervised Extreme Learning Machines (SS-ELM) [34] and Graph-based Semi-supervised Broad Learning System (GSS-BLS) [35] utilizes the differences and similarities in model features between the labeled and unlabeled datasets to select useful information. The collaborative representation-based semi-supervised extreme learning machine (CR-SSELM) [36] uses algorithms to reconstruct the unlabeled samples based on predictions of the ELM and define the risk degree of unlabeled samples, and then the unlabeled samples with a risk-based regularization term are further added into the training process of the model. Based on the idea of pseudo label, deep neural networks with contrastive learning and adversarial training strategies are applied to solve the problems of label uncertainty and label scarcity in MI-based BCI [37]. Based on the idea of meta learning, Model-Agnostic Meta-Learning (MAML) [38] focuses on training the model's initial parameters such that the model has maximal performance on a new task with a small amount of data. Besides, transformers and contrastive self-supervised learning tasks are used to build a pretrained Bert-inspired Neural Data Representations (BENDR) model [39], which is transferable to novel EEG recorded from unseen subjects, different hardware, and different tasks. However, it is still challenging for EEG-based BCI to extract features for fine-grained motion control tasks.

C. Fusion Analysis of EEG and EMG

In addition to EEG signals, EMG signals are the most direct physiological signals that can be used to assess the perception of motor performance during motor rehabilitation. EMG features can represent physiological information such as muscle fatigue, strength, and thickness [17] that can be used to control rehabilitation equipment, such as exoskeletons. EMG signals are one of the first electrophysiological signals to be used in motor rehabilitation and to assess the level of motor function [18]. By fusing analysis of EEG and EMG, hybrid brain-computer interfaces can obtain a more comprehensive perception of fine-grained motion control tasks. Most of the existing hybrid BCI systems simply integrate the analysis results of EEG and EMG based on strategies like correlation coefficient of power spectrum [19], time-frequency-domain copula-based granger causality [20], cross-association analysis of time series [21] and wavelet coherence [22], without considering the synergistic complementarity between EEG and EMG as motor intention and motor execution [16]. The fine-grained motion control information obtained by these fusion analysis strategies is very limited [40], while hybrid BCI systems that concentrate on motor rehabilitation require a more powerful fusion analysis strategy to extract deep level features of the synergistic complementarity between EEG and EMG.



(a) Shallow ConvNet(One Conv Block and One Dense Layer)

(b) EEGNet(Two Conv Blocks and One Dense Layer)

(c) Deep ConvNet(Four Conv Blocks and One Dense Layer)

Fig. 2. Network Structure of the Encoder.

III. Methodology

In this section, we first illustrate the whole framework of the proposed method and then describe every essential component of our method.

A. Framework Overview

The whole framework of our proposed approach is organized as Fig. 1. The framework uses state-of-the-art deep neural networks in the field of MI-based BCI as the encoder. The proposed approach first pre-trains the encoder with labeled samples, and then updates the encoder by the following two steps iteratively: 1. generates pseudo labels for the unlabeled samples based on a label estimation strategy, and then some of the unlabeled samples with reliable pseudo labels are added to the training dataset based on a sampling criterion. 2. continues the training process of the encoder using the increased training dataset based on the semi-supervised loss. Specifically, for the first few epochs of progressive learning, a fixed number of unlabeled samples with reliable pseudo labels are selected for semi-supervised learning in each epoch. In the remaining training epochs of progressive learning, all the unlabeled samples with pseudo labels are used for semi-supervised learning. In particular, the proposed label estimation strategy could be extended to synchronous collected EMG in hybrid BCI, which is effective for extracting deep level features in fine-grained motion control tasks. The proposed framework is a practical end-to-end model for both EEG-based BCI and hybrid BCI with combined EEG and EMG.

B. Pre-Training With Deep Encoder

By designing appropriate temporal and spatial convolution layers, deep neural networks can learn an effective representation of the complex EEG features with fixed dimensions, which is transferable to downstream tasks like EEG classification and Fusion Analysis of EEG with EMG. For the proposed framework, we chose three state-of-the-art CNN networks as the encoders, namely: Shallow ConvNet [30], Deep ConvNet [30] and EEGNet [31]. The network structure of Shallow ConvNet, EEGNet and Deep ConvNet is shown in Fig. 2 (a), Fig. 2 (b) and Fig. 2 (c).

Before performing the semi-supervised learning task based on progressive learning, pretrained encoders are needed to extract an effective representation of the input signals. The pretrained encoders were trained on the labeled samples with the cross-entropy loss L_{ce}:

$$L_{ce} = -\frac{1}{N_l} * \sum_{i=1}^{N_l} \sum_{c=1}^{M} y_{ic} log(p_{ic}) \tag{1}$$

where N_l is the number of labeled samples, M is the number of categories. y_{ic} is 1 when sample i belongs to category c and otherwise it is 0. p_{ic} is the probability of a sample i belonging to category c, which is predicted by the model.

C. Label Estimation Strategy and Sampling Criterion

In order to introduce unlabeled samples into the semi-supervised learning framework, we designed a label estimation strategy to generate reliable pseudo labels for the unlabeled samples. Inspired by the metric learning method, which focuses on feature transformation of samples through embedding learning and solves the learning task by learning how to distinguish between different classes of samples [41], pseudo labels are generated based on the distance between the unlabeled samples and labeled samples in the feature space of the deep encoder. More specifically, an unlabeled sample x_i will receive a pseudo label c if the distance between its feature vector and the center of the feature vectors for labeled class c samples is relatively small in the feature space, and the distance between its feature vector and the center of the feature vectors for all the other classes is relatively large. The confidence of the pseudo label for unlabeled samples can be further defined as:

$$Distance_{ik} = \lambda_1 ||\phi(\theta, x_i) - Mean_{j\in X^k}\phi(\theta, x_j)|| \quad (2)$$

$$Distance_{ic} = \lambda_2 ||\phi(\theta, x_i) - Mean_{j\in X^c}\phi(\theta, x_j)|| \quad (3)$$

$$D_{ic} = \sum_{k\in M, k\neq c} Distance_{ik} - Distance_{ic} \quad (4)$$

$$Pseudo_{ic} = \frac{exp(D_{ic})}{\sum_{k\in M} exp(D_{ik})} \quad (5)$$

where $Pseudo_{ic}$ represents the level of confidence for an unlabeled sample i belongs to category c. $||\cdot\cdot||$ is the Euclidean distance. $\phi(\theta, x_i)$ is the embedding of EEG sample x_i from the pretrained deep encoder with weights θ. M is the number of categories. X^k is the set of all the labeled samples which belongs to category k. λ_1 and λ_2 are weight parameters. Moreover, to further learn a more distinctive feature embedding that makes the feature representation of unlabeled samples closer to the center of their categories' embedding in the feature space and away from other categories, the metric loss L_{mt} is defined as:

$$L_{mt} = \frac{1}{N} * \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} y_{ij} Dis_{ij}^2 + (1 - y_{ij}) max(M - Dis_{ij}, 0)^2 \quad (6)$$

$$Dis_{ij} = \lambda_3 ||\phi(\theta, x_i) - \phi(\theta, x_j)|| \quad (7)$$

where N is the number of labeled and unlabeled samples. y_{ij} is 1 when sample i and sample j have the same (pseudo) label, otherwise it is 0. M and λ_3 are hyperparameters. The metric loss enables the model to learn to distinguish between the unlabeled samples, by maximizing the distance between pairs of the samples with different (pseudo) labels, and minimizing the distance between pairs of the samples with the same (pseudo) label.

To avoid model overfitting on unlabeled dataset with pseudo labels, we propose an efficient sampling criterion. For samples in the same category, the confidence of pseudo labels is measured by the distance between the unlabeled samples and the center of each category in the feature space, which reflects the feature similarity between the unlabeled samples and each category. Therefore, a high sampling threshold of confidence for pseudo labels will filter out unlabeled samples with distinct features, and we pick the sampling threshold of confidence as 95%. However, some subjects of MI-based BCI are unable to produce the distinct patterns after long-term training, resulting in ineffective control of BCI, i.e., the problem of BCI Illiteracy [42]. Based on the indistinct features of these subjects, the deep encoder will extract invalid feature representations, which lead to the errors of pseudo labels. To mitigate the problem of error labels, the proposed framework uses an improved label smoothing technique rather than cross-entropy to calculate the semi-supervised loss of pseudo labels from unlabeled samples. The label smoothing technique is a regularization method based on cross-entropy loss that can improve the generalization ability of deep networks [9]. The improved label smoothing loss for pseudo labels is defined as:

$$L_{ls} = -\frac{1}{N_s} * \sum_{i=1}^{N_s} \sum_{c=1}^{M} Pseudo_{ic} log(p_{ic}) \quad (8)$$

where N_s is the number of selected unlabeled samples, and $Pseudo_{ic}$ is calculated by equation 5. By softening the training targets based on the confidence of the pseudo label, the improved label smoothing technique prevents the encoder from overfitting, and makes the prediction probabilities of the encode more accurately represent the confidence of the pseudo label.

D. Progressive Learning Scheme

By iteratively learning a more accurate representation of the unlabeled samples, we progressively train the model based on the semi-supervised dataset. To achieve the goal of simultaneously extracting feature representations from the labeled dataset and distinguishing as well as extracting feature representations from the unlabeled dataset, the training loss of the entire semi-supervised learning framework is defined as:

$$L = \omega_1 L_{ce} + \omega_2 L_{mt} + \omega_3 L_{ls} \quad (9)$$

where L_{ce} is the cross-entropy loss for supervised learning task on labeled dataset from equation 1. L_{mt} is the metric loss for label estimation task on unlabeled dataset from equation 6. L_{ls} is the label smoothing loss for semi-supervised learning task on unlabeled dataset from equation 8. ω_1, ω_2 and ω_3 are weight parameters.

For the semi-supervised learning task, we progressively train the pre-trained deep encoder based on the label estimation strategy and sampling criterion. More precisely, the progressive learning scheme is summarized in Algorithm 1.

E. Fusion Analysis Strategy of EEG and EMG

Existing methods for EEG and EMG fusion analysis usually select pairs of EEG and EMG channels with high correlation in fixed frequency bands as the feature representation for the hybrid BCI tasks. However, the fusion analysis models from these methods, such as the correlation coefficient of power spectrum [19] and wavelet coherence [22], usually require accurate labels and significant computational calibration time to calculate the correlation between each pair of EEG and EMG channels at each frequency band, which is not suitable for the semi-supervised task. Inspired by the self-supervised learning method, which focuses on learning from unsupervised information from unlabeled samples [43], the feature

Algorithm 1 Progressive Learning Scheme for EEG-Based BCI.

Require: Labeled dataset D_L, unlabeled samples S_U, pre-trained deep encoder $\phi(\theta, \cdot)$, number of epochs for first learning stage E_1, number of epochs for second learning stage E_2, number of selected unlabeled samples N_s.
Ensure: The best deep encoder $\phi(\theta^*, \cdot)$.
  $\phi(\theta^*, \cdot) \leftarrow \phi(\theta, \cdot)$, $P \leftarrow 0$
  for $i = 0$ to E_1 **do**
    Estimate pseudo labels for S_U by equation 5
    Select N_s samples with pseudo labels from S_U based on sampling criterion $\rightarrow D_S$
    Update $\phi(\theta^*, \cdot)$ on D_L and D_S by equation 9 $\rightarrow \phi(\theta^i, \cdot)$
    Evaluate $\phi(\theta^i, \cdot)$ on the validation set for the task $\rightarrow$ performance P_i
    if $P_i > P$ **then**
      $P \leftarrow P_i$, $\phi(\theta^*, \cdot) \leftarrow \phi(\theta^i, \cdot)$
    end if
  end for
  for $j = 0$ to E_2 **do**
    Estimate pseudo labels for S_U by equation 5
    Select all samples with pseudo labels from $S_U \rightarrow D_U$
    Update $\phi(\theta^*, \cdot)$ on D_L and D_U by equation 9 $\rightarrow \phi(\theta^j, \cdot)$
    Evaluate $\phi(\theta^j, \cdot)$ on the validation set for the task $\rightarrow$ performance P_j
    if $P_j > P$ **then**
      $P \leftarrow P_j$, $\phi(\theta^*, \cdot) \leftarrow \phi(\theta^j, \cdot)$
    end if
  end for

representation of the EMG signals from the deep encoder is used as the pseudo labels in the proposed framework for the hybrid BCI system. More precisely, during the progressive learning process, the synchronous unlabeled EMG samples and unlabeled EEG samples are fed into two pre-trained deep encoders separately to obtain their feature embeddings, then the similarities of these feature embeddings are used to calculate the fusion loss of the model. The fusion loss of the unlabeled synchronous EMG and EEG samples is defined as:

$$L_{fu} = -\frac{1}{N_u} * \sum_{i=1}^{N_u} log \frac{exp(sim(\phi(\theta_1, x_i^{eeg}), \phi(\theta_2, x_i^{emg})))}{\sum_{j \neq i}^{N_u} exp(sim(\phi(\theta_1, x_i^{eeg}), \phi(\theta_2, x_j^{emg})))} \tag{10}$$

where N_u is the number of unlabeled synchronous EMG and EEG samples. $sim(v_1, v_2)$ represents the cosine similarity between vector v_1 and v_2. $\phi(\theta_1, x_i^{eeg})$ is the embedding of EEG sample x_i^{eeg} from the pretrained deep encoder with weights θ_1. $\phi(\theta_2, x_i^{emg})$ is the embedding of EMG sample x_i^{emg} from the pretrained deep encoder with weights θ_2. Moreover, the training loss of the entire semi-supervised learning framework for the hybird BCI system is defined as:

$$L_{hybrid} = \omega_1 L_{ce} + \omega_4 L_{fu} \tag{11}$$

where L_{ce} is the cross-entropy loss for supervised learning task on labeled synchronous EMG and EEG dataset from equation 1. L_{fu} is the fusion loss for semi-supervised learning task on unlabeled synchronous EMG and EEG dataset from equation 10. ω_1 and ω_4 are weight parameters.

IV. Experiments and Discussion

In this section, we demonstrate the effectiveness of the proposed framework by applying it to public BCI datasets and hybrid datasets with synchronous EMG and EEG.

A. EEG Datasets and Experimental Settings

BCI Competition IV dataset IIa [45] and BCI Competition III dataset IVa [46], as the public benchmark datasets for the motor imagery task in EEG, are used to evaluate the performance of the proposed framework. The BCI Competition IV dataset IIa comprises EEG measurements from nine subjects with four classes of MI tasks, namely, left hand imagery, right hand imagery, feet imagery, and tongue imagery. During the sessions, 22 channels of EEG data were recorded at 250 Hz and bandpass-filtered between 0.5 Hz and 100 Hz. Each session is comprised of 288 trials, with 72 trials for each of the four categories. The BCI Competition III dataset IVa collects EEG data for the motor imagery task from five healthy subjects. For each subject, 280 trials of data are collected. There are two classes of MI tasks: right hand movement imagery and foot movement imagery. The EEG data is collected on 118 channels at 1000Hz and bandpass-filtered between 0.05 Hz and 200 Hz.

In the semi-supervised learning experiments, we used three different deep learning models, namely, Shallow ConvNet [30], Deep ConvNet [30] and EEGNet [31], as the encoders of the proposed framework. For the loss equation 9 and equation 11, the weight parameters ω_1, ω_2, ω_3 and ω_4 is empirically set to 1.0, 1.0, 3.0 and 3.0, as recommended in [8]. For the classification tasks, the samples from the training session are used as the training dataset, while the samples from the evaluation session are used as the evaluation dataset, and the performance is evaluated in terms of the mean kappa value of 100 training procedures. During the semi-supervised training, we randomly selected some of the samples from the training dataset as the labeled dataset, while the remaining samples were treated as the unlabeled dataset.

B. Results on EEG Datasets

To evaluate the performance of the proposed framework, comparative experiments have been conducted between the proposed framework and other state-of-the-art semi-supervised learning models based on the EEG datasets. According to the experimental settings of most state-of-the-art models, we randomly selected 20%, 50% and 80% samples from the dataset as the labeled samples for the semi-supervised classification tasks on each subject, and also provided experiments with 100% labeled samples to compare with supervised learning models. As shown in Table I, the proposed framework achieved an average improvement of 3.10% compared to the

TABLE I
COMPARISON OF PROPOSED METHOD AND STATE-OF-THE-ART METHODS ON EEG DATASETS

	Four Categories Classification Mean Kappa Value BCI Competition IV Dataset IIa Ratio of Labeled Samples				Two Categories Classification Mean Kappa Value BCI Competition III Dataset IVa Ratio of Labeled Samples			
	20%	50%	80%	100%	20%	50%	80%	100%
FBCSP [23]				0.663				0.907
ELM [25]	0.438	0.462	0.558	0.563	0.545	0.603		
BLS [26]					0.508	0.558		
SS-ELM [34]		0.487	0.564					
EEGNet [31]	0.453	0.473	0.542	0.604	0.634	0.727	0.768	0.886
SMIR [33]					0.469	0.568		
GSS-BLS [35]					0.589	0.692		
CR-SSELM [36]		0.503	0.571					
MAML [38]					0.605	0.605	0.605	0.605
BENDR [39]	0.473	0.473	0.473	0.473				
Proposed Method with EEGNet	**0.507**	**0.538**	**0.595**	0.604	**0.674**	**0.771**	**0.792**	0.886

Missing values in the table are due to the absence of relevant experiments in the reference papers

TABLE II
COMPARISON OF PROPOSED METHOD AND STATE-OF-THE-ART METHODS ON HYBRID DATASETS FOR EEG-EMG FUSION ANALYSIS

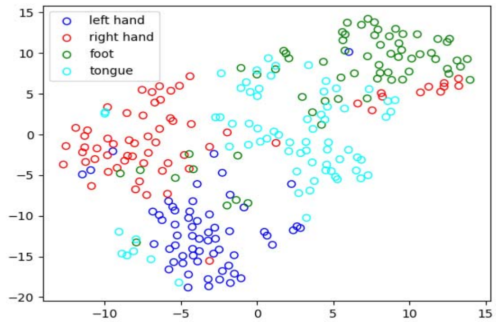
	Three Categories Classification Mean Accuracy Upper Limb Motion Dataset Ratio of Labeled Samples			Four Categories Classification Mean Kappa Value Two-finger Gameplay Motion Dataset Ratio of Labeled Samples		
	30%	60%	100%	30%	60%	100%
Correlation Coefficient of Power Spectrum [19]	68.45%	71.36%	75.74%	0.443	0.518	0.654
Cross-association Analysis of Time Series [21]	74.56%	79.73%	81.57%	0.422	0.486	0.632
CSP & ERP [44]				0.364	0.503	0.686
Coherence Analysis with Enhanced MSC [22]		83.50%	83.50%			
Proposed Method with EEGNet	**80.29%**	**88.83%**	**90.78%**	**0.476**	**0.548**	0.604

Missing values in the table are due to the absence of relevant experiments in the reference papers

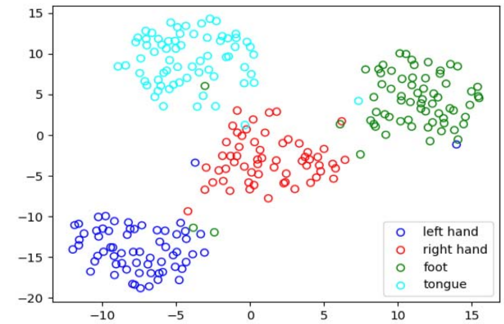
state-of-the-art models on the semi-supervised classification task for BCI Competition IV Dataset IIa. Moreover, the proposed framework achieved an average improvement of 3.60% compared to the state-of-the-art models on the semi-supervised classification task for BCI Competition III Dataset IVa. The proposed framework brings performance improvement in all these three experiments with different ratios of unlabeled samples. For the supervised learning tasks with 100% labeled samples, the model will not bring performance improvement to the deep encoders, since the proposed progressive learning scheme does not modify the structure of the deep encoders and the supervised learning task does not provide unlabeled samples for the proposed label estimation strategy and sampling criterion. More precisely, the proposed framework is ineffective for supervised learning tasks. A possible future optimization method is to improve the structure of the deep encoders based on metric learning methods.

C. Hybrid Datasets for EEG-EMG Fusion Analysis

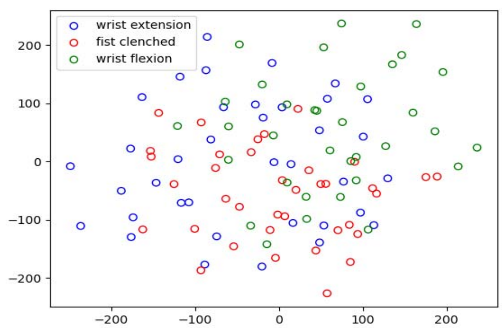
The hybrid datasets with synchronous EEG and EMG are collected from the previous study of enhanced EEG–EMG fusion analysis [22] and the impact of loss of control on movement BCIs [44]. For the Upper Limb Motion Dataset [22], the EEG data was collected from the C3 channel. The EMG data was collected on the subject's ulnar extensional wrist



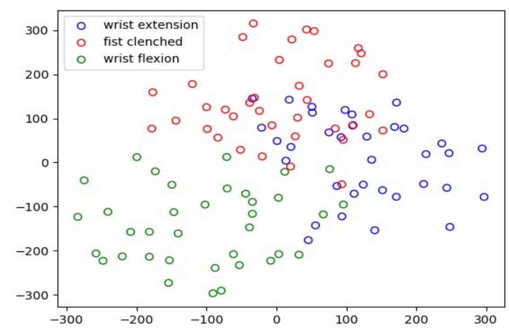
(a) Deep encoder without proposed framework for BCI Competition IV Dataset IIa subject 6



(b) Deep encoder with proposed framework for BCI Competition IV Dataset IIa subject 6



(c) Deep encoder without proposed framework for Upper Limb Motion Dataset



(d) Deep encoder with proposed framework for Upper Limb Motion Dataset

Fig. 3. Visualization of the feature space of deep encoder for different tasks on different datasets.

muscle (ECU), radial extension of the wrist muscle (ECR) and flexor digitorum (FD). Five subjects were asked to perform three fine-grained motion control movements, including hand fisting, wrist extension, and wrist flexion. Three sessions were recorded for each subject. Both EEG and EMG data were sampled at 512 Hz. For the Two-finger Gameplay Motion Dataset [44], the EEG data was collected from 32 channels. Two pairs of EMG signals were collected over the left and right flexor digitorum profundus. Nine subjects were asked to perform four fine-grained motion control movements, including key pressing with the left index finger and key pressing with the right index finger for normal condition and LOC condition. In the LOC condition, 15% of the keyboard input was ignored, and a visual lag was induced.

In the semi-supervised learning experiments, we keep the structure and parameter settings of the framework consistent with those in Section IV-A. For the classification tasks of hybrid datasets with synchronous EEG and EMG, the performance is evaluated in terms of the mean accuracy and mean kappa value using 100×10-fold cross-validation for all the sessions.

D. Results on Hybrid Datasets

To evaluate the performance of the proposed framework, comparative experiments have been conducted between the proposed framework and the state-of-the-art models on the hybrid datasets with EEG and EMG for fine-grained motion control. According to the experimental settings of most state-of-the-art models, we randomly selected 30% and 60% samples from the dataset as the labeled samples for the semi-supervised classification tasks on each subject, and also provided experiments with 100% labeled samples to compare with supervised learning models. As shown in Table II, the proposed framework achieved an average improvement of 5.53% compared to the state-of-the-art models on the semi-supervised classification task for Upper Limb Motion Dataset. Moreover, the proposed framework achieved an average improvement of 3.15% compared to the state-of-the-art models on the semi-supervised classification task for the Two-finger Gameplay Motion Dataset. The proposed framework brings performance improvement in these two experiments with different ratios of unlabeled samples.

E. Ablation Analysis

To analyze the effectiveness of the proposed framework, Fig. 3 uses t-stochastic neighbor embedding (t-SNE) [47], a high-dimensional feature visualization technique, to illustrate the feature space of the deep encoder with and without the proposed framework for both motor imagery classification tasks on the BCI Competition IV Dataset IIa and fine-grained upper limb motion classification tasks on the Upper Limb

TABLE III
ABLATION ANALYSIS OF DEEP ENCODERS

BCI Competition IV Dataset IIa with 50% labeled samples				
Shallow ConvNet	Deep ConvNet	EEGNet	without Framework	with Framework
✓			0.387	0.450
	✓		0.485	0.513
		✓	0.473	0.538
Upper Limb Motion Dataset with 60% labeled samples				
Shallow ConvNet	Deep ConvNet	EEGNet	without Framework	with Framework
✓			64.52%	69.35%
	✓		76.61%	83.30%
		✓	82.76%	88.83%

TABLE IV
ABLATION ANALYSIS OF THE PROPOSED FRAMEWORK

BCI Competition IV Dataset IIa with 50% labeled samples			
Label Estimation	Sampling Criterion	Semi Loss	Mean Kappa
×			0.473
✓			0.458
✓		✓	0.506
✓	✓		0.495
✓	✓	✓	0.538
Upper Limb Motion Dataset with 60% labeled samples			
Label Estimation	Sampling Criterion	Semi Loss	Mean Accuracy
×			82.76%
✓			79.83%
✓		✓	85.56%
✓	✓		84.69%
✓	✓	✓	88.83%

Motion Dataset with synchronous EMG and EEG. As shown in Fig. 3 (a), (b), (c) and (d), the samples from different categories are spread into more distinguishable and tighter clusters after applying the proposed framework for these two datasets, which means that the deep encoders have learned a more distinctive feature embedding on the semi-supervised learning tasks, further demonstrating the effectiveness of the proposed framework.

To further illustrate the effectiveness of the proposed framework, the ablation analysis results of deep encoders are shown in Table III. For the three deep encoders, the proposed framework achieved an average improvement of 5.20% for semi-supervised learning task on BCI Competition IV Dataset IIa with 50% labeled samples, and an average improvement of 5.86% for semi-supervised learning task on the Upper Limb Motion Dataset with 60% labeled samples. The proposed framework improved the profermance of these three deep encoders without changing their network structure. The proposed framework is suitable for most deep networks and can be applied to both EEG-based BCI and hybrid BCI without changing the architecture. Besides, ablation experiments have been conducted to illustrate the effectiveness of different components of the proposed framework. As shown in Table IV, while using the proposed label estimation strategy alone produces pseudo labels with low confidence and leads to performance degradation, when combined with the proposed sampling criterion, the effective sampling strategy would extract high-confidence pseudo labels, which improve the performance of the whole framework. Moreover, while using the proposed label estimation strategy alone produces potentially erroneous pseudo labels and leads to performance degradation, when combined with the proposed semi-supervised loss, the effective semi-supervised loss would perceive and reduce the metric loss of the pseudo labels during the training process, which improves the performance of the whole framework. Combining these three components together, the proposed framework obtained excellent results on semi-supervised learning tasks for both EEG-based BCI and hybrid BCI.

V. CONCLUSION

In this paper, we presented an end-to-end progressive learning framework for the semi-supervised learning task in the fields of both motor imagery EEG classification and EEG–EMG fusion analysis. Starting from the deficiencies of the related model, our method aims at alleviating the problems of low recognition accuracy on subjects with unlabeled samples and indistinct features and the huge calibration time for EEG-EMG fusion analysis in the fine-grained motion control task. Based on the effective label estimation strategy and sampling criterion, the proposed framework learned distinctive feature embedding from the unlabeled EEG samples and synchronous EMG samples, which significantly improved the recognition accuracy of subjects with unlabeled samples and indistinct features and substantially reduced the data calibration time for EEG-EMG fusion analysis. Extensive experiments have been conducted and the proposed framework obtained superiority results on semi-supervised learning tasks for both EEG-based BCI and hybrid BCI. Moreover, the proposed framework can be applied to almost all deep learning based BCI systems.

Our approach shows great potential in the development of both high-performance EEG-based BCI systems and high-performance hybrid BCI systems with strong generalization capability and practicality, which can be applied to build an efficient and practical BCI motor rehabilitation system, and we take this part as our future research work.

REFERENCES

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, “Brain-computer interfaces for communication and control,” *Clin. Neurophys.*, vol. 113, no. 6, pp. 767–791, 2002.

[2] X. Xue, H. Tu, Z. Deng, L. Zhou, N. Li, and X. Wang, "Effects of brain-computer interface training on upper limb function recovery in stroke patients: A protocol for systematic review and meta-analysis," *Medicine*, vol. 100, no. 23, 2021, Art. no. e26254.

[3] A. Chatterjee, A. Mandal, S. Roy, S. Sinha, A. Priya, and Y. Gupta, "A decadal walk on BCI technology: A walkthrough," in *Analyzing Data Through Probabilistic Modeling in Statistics*. Hershey, PA, USA: IGI Global, 2021, pp. 158–183.

[4] N. Sharma, J.-C. Baron, and J. B. Rowe, "Motor imagery after stroke: Relating outcome to motor network connectivity," *Ann. Neurol.*, vol. 66, no. 5, pp. 604–616, Nov. 2009.

[5] P. D. E. Baniqued *et al.*, "Brain–computer interface robotics for hand rehabilitation after stroke: A systematic review," *J. Neuroeng. Rehabil.*, vol. 18, no. 1, pp. 1–25, 2021.

[6] M. van der Meulen, G. Allali, S. W. Rieger, F. Assal, and P. Vuilleumier, "The influence of individual motor imagery ability on cerebral recruitment during gait imagery," *Hum. Brain Mapping*, vol. 35, no. 2, pp. 455–470, Feb. 2014.

[7] R. Scherer *et al.*, "Individually adapted imagery improves brain-computer interface performance in end-users with disability," *PLoS ONE*, vol. 10, no. 5, May 2015, Art. no. e0123727.

[8] D.-H. Lee *et al.*, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, vol. 3, no. 2, 2013, p. 896.

[9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2016, pp. 2818–2826.

[10] J. Sobierajewicz, S. Szarkiewicz, A. Przekoracka-Krawczyk, W. Jaśkowski, and R. van der Lubbe, "To what extent can motor imagery replace motor execution while learning a fine motor skill?" *Adv. Cogn. Psychol.*, vol. 12, no. 4, p. 179, 2016.

[11] L. Marchal-Crespo and D. J. Reinkensmeyer, "Review of control strategies for robotic movement training after neurologic injury," *J. Neuroeng. Rehabil.*, vol. 6, no. 1, pp. 1–15, 2009.

[12] A. Schwarz, P. Ofner, J. Pereira, A. I. Sburlea, and G. R. Müller-Putz, "Decoding natural reach-and-grasp actions from human EEG," *J. Neural Eng.*, vol. 15, no. 1, Feb. 2018, Art. no. 016005.

[13] V. K. Benzy, A. P. Vinod, R. Subasree, S. Alladi, and K. Raghavendra, "Motor imagery hand movement direction decoding using brain computer interface to aid stroke recovery and rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 3051–3062, Dec. 2020.

[14] H. Yuan, C. Perdoni, and B. He, "Relationship between speed and EEG activity during imagined and executed hand movements," *J. Neural Eng.*, vol. 7, no. 2, Apr. 2010, Art. no. 026001.

[15] J. N. Sanes and J. P. Donoghue, "Oscillations in local field potentials of the primate motor cortex during voluntary movement.," *Proc. Nat. Acad. Sci. USA*, vol. 90, no. 10, pp. 4470–4474, May 1993.

[16] J. Groß, P. Tass, S. Salenius, R. Hari, H. Freund, and A. Schnitzler, "Cortico-muscular synchronization during isometric muscle contraction in humans as revealed by magnetoencephalography," *J. Physiol.*, vol. 527, no. Pt3, pp. 623–631, 2000.

[17] D. G. Lloyd and T. F. Besier, "An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments *in vivo*," *J. Biomech.*, vol. 36, no. 6, pp. 765–776, Jun. 2003.

[18] Z. O. Khokhar, Z. G. Xiao, and C. Menon, "Surface EMG pattern recognition for real-time control of a wrist exoskeleton," *Biomed. Eng. OnLine*, vol. 9, no. 1, pp. 1–17, Dec. 2010.

[19] G. Bartur, H. Pratt, and N. Soroker, "Changes in mu and beta amplitude of the EEG during upper limb movement correlate with motor impairment and structural damage in subacute stroke," *Clin. Neurophysiol.*, vol. 130, no. 9, pp. 1644–1651, 2019.

[20] Q. She, H. Zheng, T. Tan, B. Zhang, Y. Fan, and Z. Luo, "Time-frequency-domain copula-based Granger causality and application to corticomuscular coupling in stroke," *Int. J. Humanoid Robot.*, vol. 16, no. 4, Aug. 2019, Art. no. 1950018.

[21] B. Kim, L. Kim, Y.-H. Kim, and S. K. Yoo, "Cross-association analysis of EEG and EMG signals according to movement intention state," *Cognit. Syst. Res.*, vol. 44, pp. 1–9, Aug. 2017.

[22] X. Xi *et al.*, "Enhanced EEG–EMG coherence analysis based on hand movements," *Biomed. Signal Process. Control*, vol. 56, Feb. 2020, Art. no. 101727.

[23] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selelction in motor imagery based brain-computer interface," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, pp. 5392–5395.

[24] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intell.)*, Jun. 2008, pp. 2390–2397.

[25] J. Cao, K. Zhang, M. Luo, C. Yin, and X. Lai, "Extreme learning machine and adaptive sparse representation for image classification," *Neural Netw.*, vol. 81, pp. 91–102, Sep. 2016.

[26] C. L. P. Chen and Z. L. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2017.

[27] J. Jin, R. Xiao, I. Daly, Y. Miao, X. Wang, and A. Cichocki, "Internal feature selection method of CSP based on L1-norm and Dempster–Shafer theory," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4814–4825, Nov. 2020.

[28] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Exp. Syst. Appl.*, vol. 37, no. 12, pp. 8659–8666, Dec. 2010.

[29] H. Altaheri *et al.*, "Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: A review," *Neural Comput. Appl.*, vol. 33, pp. 1–42, Aug. 2021, doi: 10.1007/s00521-021-06352-5.

[30] R. T. Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[31] V. Lawhern, A. Solon, N. Waytowich, S. M. Gordon, C. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, 2018, Art. no. 056013.

[32] J. E. Van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Mach. Learn.*, vol. 109, no. 2, pp. 373–440, 2020.

[33] G. Niu, W. Jitkrittum, B. Dai, H. Hachiya, and M. Sugiyama, "Squared-loss mutual information regularization: A novel information-theoretic approach to semi-supervised learning," in *Proc. 30th Int. Conf. Mach. Learn. (PMLR)*, 2013, pp. 10–18.

[34] G. Huang, S. Song, J. N. D. Gupta, and C. Wu, "Semi-supervised and unsupervised extreme learning machines," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2405–2417, Dec. 2014.

[35] Q. She, Y. Zhou, H. Gan, Y. Ma, and Z. Luo, "Decoding EEG in motor imagery tasks with graph semi-supervised broad learning," *Electronics*, vol. 8, no. 11, p. 1273, Nov. 2019.

[36] Q. She, J. Zou, Z. Luo, T. Nguyen, R. Li, and Y. Zhang, "Multi-class motor imagery EEG classification using collaborative representation-based semi-supervised extreme learning machine," *Med. Biol. Eng. Comput.*, vol. 58, no. 9, pp. 2119–2130, Sep. 2020.

[37] J. Han, X. Gu, and B. Lo, "Semi-supervised contrastive learning for generalizable motor imagery EEG classification," in *Proc. IEEE 17th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, Jul. 2021, pp. 1–4.

[38] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. 34th Int. Conf. Mach. Learn. (PMLR)*, in Machine Learning Research, vol. 70, D. Precup and Y. W. Teh, Eds. Aug. 2017, pp. 1126–1135.

[39] D. Kostas, S. Aroca-Ouellette, and F. Rudzicz, "BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data," *Frontiers Hum. Neurosci.*, vol. 15, p. 253, Jun. 2021.

[40] J. Liu, Y. Sheng, and H. Liu, "Corticomuscular coherence and its applications: A review," *Frontiers Hum. Neurosci.*, vol. 13, p. 100, Mar. 2019.

[41] D. Li and Y. Tian, "Survey and experimental study on metric learning methods," *Neural Netw.*, vol. 105, pp. 447–462, Sep. 2018.

[42] B. Blankertz *et al.*, "Predicting BCI performance to study BCI illiteracy," *BMC Neurosci.*, vol. 10, no. S1, p. 84, 2009.

[43] X. Liu *et al.*, "Self-supervised learning: Generative or contrastive," *IEEE Trans. Knowl. Data Eng.*, early access, Jun. 22, 2021, doi: 10.1109/TKDE.2021.3090866.

[44] B. Reuderink, M. Poel, and A. Nijholt, "The impact of loss of control on movement BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 6, pp. 628–637, Dec. 2011.

[45] M. Tangermann *et al.*, "Review of the BCI competition IV," *Frontiers Neurosci.*, vol. 6, p. 55, Jul. 2012.

[46] G. Dornhege, B. Blankertz, G. Curio, and K. R. Müller, "Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 993–1002, Jun. 2004.

[47] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 1–27, 2008.