Multi-Source Transfer Learning for EEG Classification Based on Domain Adversarial Neural Network

Dezheng Liu, Jia Zhang, *Member, IEEE*, Hanrui Wu, Siwei Liu, and Jinyi Long

***Abstract*— Electroencephalogram (EEG) classification has attracted great attention in recent years, and many models have been presented for this task. Nevertheless, EEG data vary from subject to subject, which may lead to the performance of a classifier degrades due to individual differences. To collect enough labeled data to model would address the issue, but it is often time-consuming and labor-intensive. In this paper, we propose a new multi-source transfer learning method based on domain adversarial neural network for EEG classification. Specifically, we design a domain adversarial neural network, which includes a feature extractor, a classifier, and a domain discriminator, and therefore reduce the domain shift to achieve the purpose. In addition, a unified multi-source optimization framework is constructed to further improve the performance, and the result for EEG classification is induced by the weighted combination of the predictions from multiple source domains. Experiments on three publicly available EEG datasets validate the advantages of the proposed method.**

***Index Terms*— Brain-computer interfaces, multi-source fusion, adversarial learning, electroencephalogram, transfer learning.**

I. Introduction

BRAIN-COMPUTER interface (BCI) provides a direct communication pathway between the human brain and an external device [1]. It works by receiving signals from the brain and converting the signals into recognizable instructions or emotional states, etc. Typically, BCI input neural signals can be divided into the following categories, i.e., electroencephalogram (EEG), magnetoencephalography (MEG), electrocorticography (ECoG), functional magnetic resonance imaging (fMRI), and near-infrared spectroscopy (NIRS) [2], [3], [4], [5], [6]. Among these signals, EEG recordings have received particular attention due to their low price, non-invasiveness and high temporal resolution compared with other neuro-physiological modalities. With this, EEG based BCI systems inspire a wide range of applications, including communication, movement, and rehabilitation for disabilities, as well as entertainment for the able-bodied [1].

To implement the applications mentioned above, leveraging machine learning techniques to conduct EEG signal analysis is a crucial step. For this purpose, EEG signal preprocessing, feature extraction, and classification methods have been intensively investigated. However, there still exist some major challenges. First, different subjects have different neural responses to the same stimulus, and even the same subject can have different neural responses to the same stimulus at different time/locations [7]. Individual differences are notorious to tackle for EEG signal analysis, leading to the difficulty of effectively recognizing a new arrival [8], [9]. Second, model training usually requires abundant labeled training data. Nevertheless, the acquisition of EEG signals is not easy as it is time-consuming and expensive [10], [11]. This fact gives rise to another problem, that is, only a small number of samples can be used for training.

To face with these challenges, transfer learning strategy has become the prime focus, and many applications of transfer learning in BCIs have been put forward. For example, Kang and Choi [12] proposed a Bayesian CSP model for motor imagery tasks. Lan et al. [13] modeled the discrepancy from different subjects using maximum independence domain adaptation and transfer component analysis. Li et al. [14] considered to select informative sources, and then applied style transfer mapping to emotion recognition. All of them attempt to apply the knowledge learned from a specific domain to train classifiers or extract significant information, thus achieving the successful learning on a related but different domain. Furthermore, due to the excellent performance of deep neural networks in complex information processing, recently some studies have investigated transfer learning based on deep

Manuscript received 26 June 2022; revised 29 September 2022; accepted 1 November 2022. Date of publication 4 November 2022; date of current version 31 January 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62206111, Grant 62276115, and Grant 62106084; in part by the Outstanding Youth Project of Guangdong Natural Science Foundation of China under Grant 2021B1515020076; in part by the National Natural Science Foundation of Guangdong, China, under Grant 2019A1515012175 and Grant 2022A1515010468; in part by the Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization under Grant 2021B1212040007; and in part by the Fundamental Research Funds for the Central Universities, Jinan University, under Grant 21621026 and Grant 21622326. *(Dezheng Liu and Jia Zhang contributed equally to this work.) (Corresponding author: Jinyi Long.)*

Dezheng Liu, Jia Zhang, Hanrui Wu, and Siwei Liu are with the College of Information Science and Technology, Jinan University, Guangzhou 510632, China.

Jinyi Long is with the College of Information Science and Technology, Jinan University, Guangzhou 510632, China, also with the Guangdong Key Laboratory of Traditional Chinese Medicine Information Technology, Guangzhou 510632, China, and also with the Pazhou Laboratory, Guangzhou 510335, China (e-mail: jinyil@jnu.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2022.3219418

neural networks for EEG classification, specially, generative adversarial network (GAN) transfer is one of representative methods [15], [16], [17], [18], [19]. In general, GAN is composed of a discriminative network (discriminator) and a generative network (generator), the generator learns the features of the source and target domains and sends them to the discriminator, which verdicts the source of the features and feeds the result back to the generators until they are indistinguishable. This type of transfer learning is conducive to facilitating EEG classification, but it has the potential problem like the collapse of the training process, which is also uncontrollable for obtaining a stable and effective result [20], [21].

In addition, transferring knowledge from multiple source domains has shown its superiority in providing an extensive view of the target domain, such as [22], [23], and [24]. For the EEG classification problems, we can easily obtain signals from multiple subjects and adopt different subjects to improve the classification performance of the target subject. However, since the individual differences among different subjects are varied, equally treating the source domain may lead to negative transfer [25]. In consequence, how to estimate the contributions of the source domains remains an open and challenging question.

In this paper, we propose a Multi-source Transfer learning method based on Domain Adversarial Neural Network, namely MTDANN, for EEG classification. The key strategies of MTDANN mainly include the training of the domain adversarial neural network (DANN) and the multi-source weighted fusion. To be specific, the DANN model is constructed to achieve feature extraction and EEG recognition based on EEGNet [26], in the meanwhile, it leverages a domain discriminant network to reduce the shift between the source and target domains. Besides, a data fusion method is further served to obtain a more reliable recognition result. In detail, we train a DANN model based on each source domain and the specified target domain, and then design an optimization framework. Under this framework, the assignment of each source weight and the weighted combination prediction are iteratively updated to achieve multi-source transfer learning. Extensive experiments on three widely used EEG MI datasets reveal the feasibility and effectiveness of the proposed method. The contributions of this paper are summarized as follows:

- We propose an adversarial inference approach to conduct EEG classification. To further enhance the performance, we attain the goal by integrating multiple subjects with weighted fusion.
- We present an improved adversarial neural network, which consists of EEGNet and a domain discriminant network, to learn the invariant representation between the source and target domains.
- Experiments on several datasets show the advantages of the proposed method.

The remainder of this article is organized as follows. Section II gives a brief review of the existing methods for EEG signal analysis. In Section III, we describe the proposed method in detail, and explain the experimental result in Section IV. Finally, in Section V, the conclusion is given.

II. Related Work

Towards EEG signal analysis, many strategies have emerged with insufficient data for subject-to-subject transfer. A popular and effective strategy is to extract similar features from different subjects, and a well-known example is the method of common spatial patterns (CSP) [27], which is one early work to learn a spatial filter that maximizes (or minimizes) the ratio of the filtered variance between different classes of EEG signals. To improve the adaptability of CSP to the problem of insufficient labels and large feature space differences, some refined methods have been put forward in recent years [28], [29]. For example, Dai et al. [28] considered to match the distribution of source and target subjects directly. For achieving this, they proposed a transfer kernel CSP (TKCSP) approach to learn a domain-invariant kernel. Azab et al. [29] used the Kullback-Leibler divergence to measure the similarity of common spatial patterns for specific subjects. In addition, many methods have been presented to reduce inter-domain discrepancies for generating domain-invariant features [13], [14]. Following this principle, Lan et al. [13] demonstrated the effectiveness of Maximum Independence Domain Adaptation (MIDA) and Transfer component analysis (TCA) in coping with domain discrepancy. Li et al. [14] proposed a multi-source transfer learning model for emotion recognition based on source selection and style transfer mapping. Furthermore, there are some other well-established methods to learn a subspace that makes feature distributions similar [11], [30], and [31]. For example, Zanini et al. [30] transformed the covariance matrices of every session/subject, and applied standard minimum distance to the classifier for cross-subject motor imagery tasks. Wu et al. [11] aligned the covariance matrices of the EEG trials in the Riemannian manifold, and then extracted features in the tangent space.

By extracting similar features, generating domain-invariant features or the same distribution features from different subjects, the early methods have gained promising results in the task of EEG single analysis. Recently, leveraging deep neural networks begins to be the mainstream due to their powerful learning capability. Deep neural networks directly extract and classify features from raw EEG signals by imitating the structure and connections of neurons, which showed an advantage over shallow models by learning a deep representation [32], [33], [34]. To face with the challenges in EEG single analysis, i.e., individual difference and insufficient information, exploiting deep neural networks to conduct transfer learning has received extensive attention, and commonly used methods include fine-tuning and deep neural network adaptation. In which fine-tuning is widely applied to a pre-trained model when there is no significant discrepancy in the source and target domains [35], [36], [37], [38]. For example, Daoud et al. [35] proposed a deep convolutional autoencoder architecture extracting the significant spatial features from different scalp positions for epileptic seizure prediction, and then fine-tuned the model by pre-trained parameters. Furthermore, several successful network structures, such as AlexNet, ResNet, and VGG-16 [36], [37], [38], have been used directly as pre-training models in fine-tuning. In general, the training time can be reduced and the learning accuracy can be improved

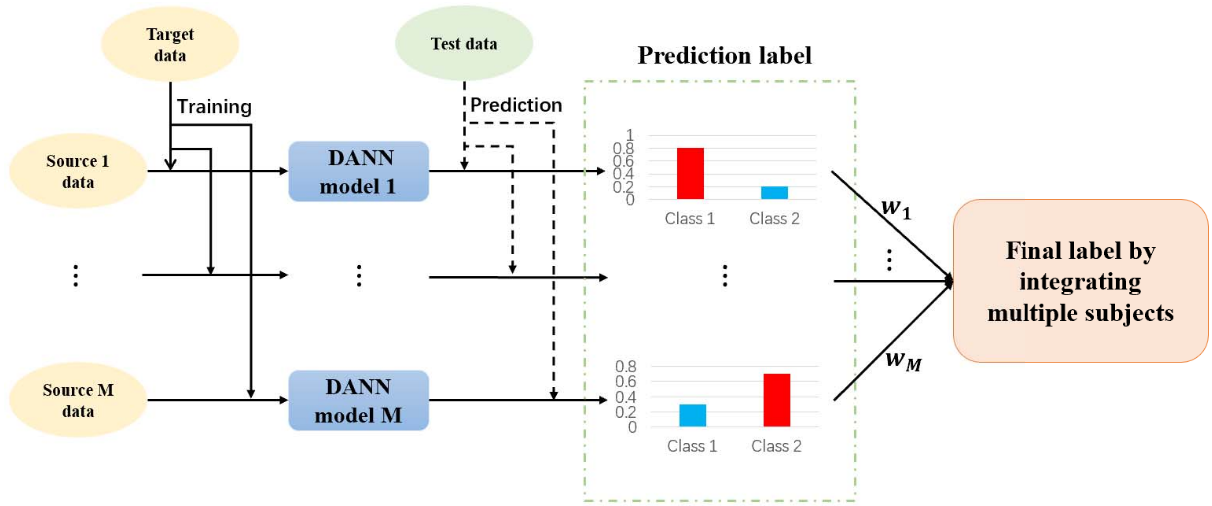

Fig. 1. Paradigm of the proposed method MTDANN.

via fine-tuning, but the disadvantage of this type of method is that it less effective when the distribution of the source and target domains are different. In this case, researchers tried to add the adaptive layer for deep neural networks, so that the distribution of the source and target domains can be closer [39], [40], [41]. Following this way using deep neural network adaptation, Zhang et al. [40] proposed a cross-subject emotion recognition method based on convolutional neural network and deep domain confusion to reduce the difference of feature distributions from different domains. Tan et al. [41] designed a deep transfer learning framework using AlexNet with an adversarial network to extract the general features, and therefore detected the difference and transferability of different domains.

Generative adversarial network (GAN) transfer can be regarded as a typical method of deep neural network adaptation. The principle of GAN has been widely used in transfer learning in BCI field [15], [16], [17], [18], [19] because of the excellent performance of GAN in the fields of image, video, and so on [42], [43], [44]. GAN was first proposed in 2014 [21], which works as follow: The fake samples are generated by the generator according to the given data and estimated by the discriminator to distinguish their source. By using GAN with application to EEG study, Li et al. [15] proposed an effective joint distributed adaptive method, which used the connection between the adversarial adaptive strategy and the functional layer of the neural network to achieve cross-subject and cross-session transfer. Zhao et al. [16] introduced a novel end-to-end deep domain adaptation method to improve the classification performance on a single subject (target domain) by taking the useful information from multiple subjects (source domain) into consideration. Ko et al. [17] presented a semi-supervised deep adversarial learning framework. This method utilized generated artificial samples along with labeled and unlabeled real samples in discovering class-discriminative features to boost the robustness of a classifier. Panwar et al. [18] considered to employ the Wasserstein generative adversarial network with gradient penalty to synthesize EEG data. This network addressed several modeling challenges of simulating time-series EEG data including frequency artifacts and training instability. To improve the generality of learning model, Li et al. [19] proposed a bi-hemisphere domain adversarial neural network to reduce the possible domain differences in each hemisphere between the source and target domains.

We also adopt GAN to design our learning framework in this paper. Different from the above GAN based methods, we leverage EEGNet to generate the effective and nonhandcrafted deep representation with adversarial learning, thereby achieving cross-subject EEG classification. For avoiding the problem of the adversarial network (like model collapse) in the process of training, we further build a multi-source weighted fusion mechanism to achieve the purpose. In the future, we will continue to focus on cross-subject EEG classification to obtain more stable results.

III. Methods

A. Problem Formulation

Formally, we suppose that there exist M source domains. In each source domain, we use X_s and Y_s to denote the EEG data and the corresponding label, respectively. Let $X = \{X_s, X_t\} = \{x_i\}_{i=1}^{N}$, $x_i \in R^{C\times T}$, where X_t is the part of data from the target domain, N is the sample size, C is the number of electrode channels, and T is the number of time points. In addition, $Y_s = \{y_i\}_{i=1}^{n_s}$, $y_i \in \{0, 1, \ldots, n_c\}$, where n_s is the number of source samples, and n_c denotes the number of classes.

To predict the label Y^* of test data, we use the labeled data from multiple subjects as the source domain and the target domain data X_t without labeling information to model. For implementation, a multi-source transfer learning method MTDANN is proposed, and the overall architecture is shown in Fig. 1. Explanatorily, we first design a domain adversarial neural network namely DANN, which consists of three components, including a feature extractor, a classifier, and a domain discriminator. Owing to the advantage of the domain adversarial architecture, the distribution of features from the

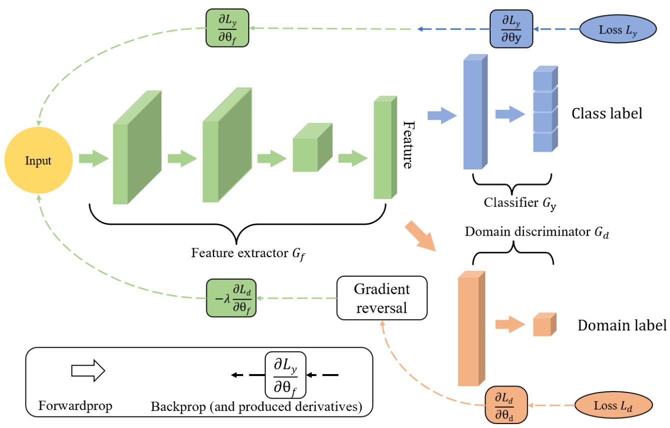

Fig. 2. The structure of the proposed domain adversarial neural network.

source domain and the target domain can be similar, so that the model is able to relieve the problem of individual differences. By using the designed DANN, we can obtain different predicted results based on M source domains respectively. Here, we denote the generated M predicted results as $\{\hat{Y}_i\}_{i=1}^{M}$, and therefore fuse them to achieve the final prediction for Y^*. In consideration of that the fusion result should be close to the prediction from the reliable source, we generalize our model as follow:

$$\begin{aligned} \min_{W,Y^*} \quad & \sum_{m=1}^{M} \sum_{\substack{y_i^* \in Y^* \\ \hat{y}_i \in \hat{Y}_m}} w_m (y_i^* - \hat{y}_i)^2 \\ s.t. \quad & \delta(W) = 1, w_m > 0. \end{aligned} \tag{1}$$

In this optimization problem, the distribution of the weights of M source domains are represented by $W = \{w_1, w_2, \ldots, w_M\}$, where w_m denotes the weight of the m-th source domain, and $\delta(\cdot)$ is the regularization function. By minimizing Eq. (1), we attempt to obtain the final prediction result Y^* by incorporating all source data into the unified optimization framework. Thus, Y^* is closer to the prediction of the source domain with the larger weight, vice versa.

B. Single-Source Transfer Based on DANN

As shown in Eq. (1), for predicting the final result Y^* of the test set, we should generate label prediction based on DANN in advance. To attain the goal, an adversarial inference approach with deep neural network is introduced from single-source domain. We illustrate the network architecture in Fig. 2, and give the description in the following subsection.

1) Architecture: Convolutional neural network (CNN) is able to decoding EEG signals in an end-to-end manner, and many related methods have received an outstanding performance by the CNN architecture in MI classification task [45], [46], [47]. Thus, we employ a popular improved CNN model, i.e., EEGNet [26], as the feature extractor. We use $G_f(\cdot;\theta_f)$ to represent the model, where θ_f is the parameters of G_f. In addition, a shallow yet efficient classifier $G_y(\cdot;\theta_y)$ with two fully connected layers is applied to EEG classification, where θ_y is the parameters of G_y. Based on this setting, the EEG signals are mapped into a discriminative representation via the feature extractor, formally written as $f_s = G_f(X_s;\theta_f)$. Then, through the classifier G_y, the features f_s can be used to generate the label $\hat{Y}$:

$$\hat{Y} = G_y(f_s;\theta_y) \tag{2}$$

Next, to make the DANN as correctly as possible to predict the label of the EEG data, a loss function is adopted to minimize the difference between the predicted label $\hat{Y}$ and corresponding ground truth Y_s. The optimization formula is given as follow:

$$\begin{aligned} E_y(\theta_f,\theta_y) &= L_y(G_y(G_f(X_s;\theta_f);\theta_y), Y_s) \\ &= L_y(G_y(f_s;\theta_y), Y_s) \\ &= L_y(\hat{Y}, Y_s) \end{aligned} \tag{3}$$

where L_y is defined as the cross entropy function due to its simplicity and effectiveness.

However, the feature transferability is significantly degraded in the fully connected layer when the domain discrepancy is enlarged. Moreover, the learned features are different between the source and target domains due to the shift in the marginal distribution caused by individual differences. Therefore, training a model only utilizing the source data easily causes the overfitting to the source distribution. To solve the above problem, a domain discriminator $G_d(\cdot;\theta_d)$, where θ_d is the parameters of G_d, is introduced to reduce the domain shift. Note that this component is essential in our network, which offers the ability for the transfer. By fooling the discriminator in the training, the features generated by the extractor G_f can be domain confusing. The optimization function is defined as follow:

$$\begin{aligned} E_d(\theta_f,\theta_d) &= L_d(G_d(G_f(X;\theta_f);\theta_d), D) \\ &= L_d(G_d(f;\theta_d), D) \\ &= L_d(\hat{D}, D) \end{aligned} \tag{4}$$

where L_d is also defined as the cross entropy function. D is the truth domain label, $\hat{D}$ is the predicted domain label, and f is the features obtained from both source domain and target domain by the extractor G_f.

2) Training Detail and Prediction: As the above stated, the model is not only expected to learn the distinguished features from the source domain, but also reduce the negative influence caused by individual differences. To this end, one approach to meet both these criteria is to minimize the loss function:

$$E(\theta_f,\theta_y,\theta_d) = E_y(\theta_f,\theta_y) - \lambda E_d(\theta_f,\theta_d) \tag{5}$$

where λ is the hyperparameter trading off the two terms. Moreover, the update rule of the domain adversarial neural network for all the parameters is designed as follows:

$$\begin{aligned} (\hat{\theta_f},\hat{\theta_y}) &= \arg\min_{\theta_f,\theta_y} E(\theta_f,\theta_y,\hat{\theta_d}) \\ \hat{\theta_d} &= \arg\max_{\theta_d} E(\hat{\theta_f},\hat{\theta_y},\theta_d) \end{aligned} \tag{6}$$

where the parameters θ_f, θ_y, θ_d deliver the saddle points of Eq. (5). It can be observed that the loss of G_y is minimized and the loss of G_d is maximized during the process of training.

TABLE I
MODEL PARAMETERS OF THE FEATURE EXTRACTOR, THE CLASSIFIER AND THE DOMAIN DISCRIMINATOR

	Layer	Filters x (Kernel Size)	Activation
Feature Extractor	Input EEG		
	Conv2D	8 x (1, 32)	Linear
	BatchNorm		
	Depthwise Conv2D	2 x (64, 1)	Linear
	BatchNorm		ELU
	Mean Pooling	(1, 4)	
	Dropout		
	Separable Conv2D	16 x (1, 16)	Linear
	BatchNorm		ELU
	Mean Pooling	(1, 8)	
	Dropout		
	flatten		
	Layer	**Kernel Size**	**Activation**
Classifier	Fully connected	25	Relu
	Dropout		
	Fully connected	2	softmax
Domain Discriminator	Fully connected	100	Relu
	Dropout		
	Fully connected	2	softmax

With this, the final model can help to extract domain-invariant class features for EEG classification. Finally, for giving a clear demonstration, we give the details of the DANN, including the information of layer, number of filters, kernel size, activation function, and options, as shown in Table I.

C. Multi-Source Transfer Framework

Based on the designed domain adversarial neural network, we consider to integrate multiple source domains by optimizing Eq. (1). The proposed objective function in Eq. (1) involves two unknown variables, that is, the weight distribution W and the final fusion result Y^*. Here, we adopt an alternating iterative strategy to develop a practical method suitable for the source weight estimation, thus achieving multi-source transfer learning. In each iteration, one of the variables is updated while the others are fixed. Therefore, the whole problem can be reduced to several simpler subproblems. In the following subsections, we give the update procedures for W and Y^*.

1) Learning the Weight Distribution W: We fix Y^* to obtain the optimal distribution of source weights W. For achieving this purpose, we let $w_m = s_m^2$, and then define the regularization function as $\delta(W) = \sum_{m=1}^{M} s_m$. By this way, the optimization for W is transformed into a constrained quadratic programming problem.

$$\begin{aligned} \min_{W} \quad & \sum_{m=1}^{M} \sum_{\substack{y_i^* \in Y^* \\ \hat{y}_i \in \hat{Y}_m}} s_m^2 \left(y_i^* - \hat{y}_i\right)^2 \\ s.t. \quad & \sum_{m=1}^{M} s_m = 1, s_m > 0 \end{aligned} \tag{7}$$

We use the Lagrange multiplier method to solve the above problem, and the definition of the Lagrangian function is as follow:

$$L(W,\mu) = \sum_{m=1}^{M} \sum_{\substack{y_i^* \in Y^* \\ \hat{y}_i \in \hat{Y}_m}} s_m^2 \left(y_i^* - \hat{y}_i\right)^2 + \mu(\sum_{m=1}^{M} s_m - 1) \tag{8}$$

By setting the partial derivatives of Lagrange multipliers μ and s_m to zero, we can derive the following formulas:

$$\begin{cases} \dfrac{\partial L(W,\mu)}{\partial s_m} = 2 \displaystyle\sum_{\substack{y_i^* \in Y^* \\ \hat{y}_i \in \hat{Y}_m}} s_m \left(y_i^* - \hat{y}_i\right)^2 + \mu = 0, \\ \qquad \times (m = 1, 2, \ldots, M) \\ \dfrac{\partial L(W,\mu)}{\partial \mu} = \displaystyle\sum_{m=1}^{M} s_m - 1 = 0 \end{cases} \tag{9}$$

In Eq. (9), $m = 1, 2, \ldots, M$, therefore, there are $M+1$ equations involving $M+1$ independent variables, namely $\{s_m\}_{m=1}^{M}$ and μ. Thus, the formula used to calculate the optimal solution for W is shown as follow:

$$s_m = \sum_{\substack{y_i^* \in Y^* \\ \hat{y}_i \in \hat{Y}_m}} \frac{1}{\left(y_i^* - \hat{y}_i\right)^2} \left(\sum_{m'=1}^{M} \frac{1}{\left(y_i^* - \hat{y}_i^{m'}\right)^2} \right)^{-1} \tag{10}$$

where $\hat{y}_i^{m'}$ is the label of the i-th test sample which is predicted based on the m-th source domain. It can be seen in Eq. (10) that the weight of a source is inversely proportional to the average deviation of its prediction for the test set, the smaller of the value $(y_i^* - \hat{y}_i)^2$, the greater weight of the m-th source domain, vice versa.

2) Learning the Fusion Result Y^*: In this step, we fix the source weight distribution W to calculate Y^*. Since each weighted fusion result $y_i^* \in Y^*$ can be updated independently, the original objective function Eq. (1) can become as follow:

$$\min_{y_i^*} \sum_{m=1}^{M} s_m^2 \left(y_i^* - \hat{y}_i\right)^2, \quad \hat{y}_i \in \hat{Y}_m \tag{11}$$

Setting the gradient of Eq. (11) to zero, the fusion result $y_i^* \in Y^*$ can be solved as follow:

$$y_i^* = \frac{\sum_{m=1}^{M} s_m^2 \hat{y}_i}{\sum_{m=1}^{M} s_m^2} \tag{12}$$

In Eq. (12), we can find out that the final fusion result Y^* is closely relevant to the source weights and the predicted results of different sources, which tends to be assigned with a value based on the prediction from the reliable source.

Notably, some methods, e.g., Major Voting and Averaging, are effective ways to the initialization of iterative optimization [48]. We initialize the value of weight distribution W using the Averaging method. It means that the initialization of the proposed optimization framework is set as $\{s_m\}_{m=1}^{M} = \frac{1}{M}$. After that, we calculate the fusion result Y^* by Eq. (12). Then, in each iteration, we update the variables, i.e., W and Y^*, by Eq. (10) and Eq. (12), respectively. Thus, the source domains which have greater predictive ability are noticed with the larger weight, so that the final result is closer to the prediction from informative source domains. In order to provide a clear description, we display the pseudo-code of the proposal in Algorithm 1.

Algorithm 1 The MTDANN Method

Input: The training set $\{X_s, X_t, Y_s\}$.
Output: The predicted label Y^* for test set X_{test}.

Initialize the parameters θ_f,θ_y,θ_d, and the value of source weights $W = \{w_1, w_2, \ldots, w_M\}$.
for $m = 1$ to M **do**
  Train the domain adversarial neural network G^m by $\{X_s, X_t, Y_s\}$, where $\{X_s, Y_s\}$ is from the m-th subject; //G^m means the domain adversarial neural network trained based on the m-th source domain.
end for
for $m = 1$ to M **do**
  Predict the label for test set: $\hat{Y}_m = G^m(X_{test})$;
end for
repeat
  for $i = 1$ to length(X_{test}) **do**
    Update $y_i^* \in Y^*$ using Eq. (12);
  end for
  Compute $\{s_m\}_{m=1}^{M}$ using Eq. (10) to update W;
until Convergence
return the final predicted label set Y^*.

IV. Experiment and Results

A. Data Description

We evaluate on three publicly available EEG datasets of motor imagery, i.e., data set Iva and data set IIIb of the BCI Competition III, and data set 1 of the BCI Competition IV. Their statistics are summarized in Table II.

(1) Dataset Iva of BCI Competition III [49]: The dataset[1] (labeled as MI1) contains 118-channel EEG signals recorded at 100 Hz from from five healthy subjects (termed as a1, a2, …, a5). The dataset is collected on two MI tasks (right hand and right foot). For each subject, 280 trials are collected.

(2) Dataset IIIb of BCI Competition III [50]: The dataset[2] (labeled as MI2) contains 2 bipolar channel EEG signals recorded at 125 Hz from three healthy subjects (termed as a1, a2, a3). The dataset is collected on two MI tasks (left and right). The number of trials for a1, a2, and a3 are 320, 1080 and 1080, respectively.

(3) Dataset 1 of BCI Competition IV [51]: The dataset[3] (labeled as MI3) contains 59-channel EEG signals recorded at 100 Hz from seven healthy subjects (termed as a1, a2, …, a7). The dataset is collected on two MI tasks (left hand and right hand), and each subject is with 200 trials.

EEG segments between [0.5, 3.5], [1, 5], and [0.5, 3.5] seconds of MI1, MI2, and MI3 are extracted after the cue onsets, respectively. In addition, a causal 50-order 8-30Hz finite impulse response (FIR) band-pass filter is used to remove muscle artifacts and direct current drift.

[1]https://bbci.de/competition/iii/desc_IVa.html
[2]https://www.bbci.de/competition/iii/desc_IIIb.pdf
[3]https://www.bbci.de/competition/iv/desc_1.html

TABLE II
The Statistics of the Three MI Datasets

Dataset	Number of Subjects	Number of Channels	Number of Time Samples	Trails per Subject	Class-Imbalance
MI1	5	118	350	280	No
MI2	3	2	625	640,1080,1080	No
MI3	7	59	300	200	No

B. Experimental Settings

There are two transfers evaluation methods in our experiment, i.e., single-source to single-target (S→S) and multi-source to single-target (M→S):

(1) S→S: S→S is designed to verify the performance of the proposed domain adversarial network in MTDANN. Here, each subject is selected as the target subject, and the rest ones as the source subject, respectively. Let k denotes the number of subjects in a dataset. So, there are $k(k-1)$ group experiments. We compute the average of $k-1$ experiments as the predicted result for the corresponding target subject.

(2) M→S: M→S is designed to verify the superiority of our MTDANN method in the case of multiple sources. In M→S, we sequentially select one subject as the target subject, and the rest ones as source subjects for training, hence there are k different M→S tasks as the number of subjects is set as k. It is worth mentioning that in M→S, most transfer learning methods concatenate all source domains into one domain. Different to them, the proposed MTDANN method analyzes the relationship between each source domain and the target domain individually, and assigns the weights to measure the importance of different source domains, based on their own predictive ability for the target data.

For example, there are five subjects (a1, a2, …, a5) in MI1 dataset, so we have $5 \times 4 = 20$ S→S tasks, e.g., a1→a2 (Subject a1 as the source domain, and subject a2 as the target domain), and five M→S tasks, e.g., a1, a2, a3, a4→a5. In this paper, the balanced classification accuracy (BCA) is used as the performance measure metric, which has been widely used in EEG classification [11], [52], [53]. The evaluation metric is defined as follow:

$$BCA = \frac{1}{l}\sum_{c=1}^{l}\frac{tP_c}{n_c} \tag{13}$$

where tP_c and n_c are the number of true positives and the number of samples in class c, respectively, and l denotes the number of classes.

C. Comparative Studies

We compare our MTDANN method with several state-of-the-art baseline algorithms for EEG classification, as follows:

- EEGNet [26]: We train EEGNet model with only the annotated source data, and directly evaluate the performance of the model on the target domain data.

- 1D-CNN (One-dimensional Convolution Neural Network) [54]: The method designs a 10-layer one-dimensional convolutional neural network (1D-CNN) for EEG classification. Then, the late layers of the source domain model are further trained using the target domain data to make the model applicable to the target domain data.
- CSP-LDA (Common Spatial Pattern-Linear Discriminant Analysis) [27]: This algorithm maximizes the variance of one class while minimizing the variance of the other class, to obtain the most distinguishable feature vector. The generated features input into a LDA (Linear Discriminant Analysis) [55] to EEG classification.
- EA-CSP-LDA: EA (Euclidean-space Alignment) [56] aligns the EEG trials from different subjects in the Euclidean space to make them more similar. Then CSP features are extracted and input to the LDA classifier for EEG classification.
- CA (Centroid Alignment) [11]: CA aligns the covariance matrices of the EEG trials in the Riemannian manifold to reduce the marginal probability distribution shift of different domains. The aligned covariance matrix is then used as features for EEG classification.
- CA-CORAL [57]: A improved method for CA, in which CORAL (CORrelation ALignment) is an unsupervised domain adaption method that minimizes domain shifts by aligning the second-order statistics of the source and target distributions.
- CA-JDA [58]: JDA (Joint Distribution Adaptation) adapts to both marginal and conditional distributions in a principled dimensionality reduction process and constructs new feature representations that are efficient and robust to a large number of distributional differences.
- CA-JGSA [59]: JGSA (Joint Geometrical and Statistical Alignment) is an unsupervised domain adaptation method that projects data from the source and target domains into low-dimensional subspaces by learning two coupled projections, where the geometric and distributional shifts are reduced simultaneously.

The above CA-CORAL, CA-JDA, and CA-JGSA algorithms are used to reduce the geometric and distribution shifts of the features after CA, and all of them use SVM [60] as the classifier. The parameter of each selected comparing method (if any) is set as the corresponding literature suggested. For the configuration of MTDANN, the Adam optimization method [61] (with learning rate $1e-3$) is adopted, and we use the default parameters in Adam: $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10e-8$. The batch size is 16, and we search for the parameter λ between 0.2 and 1 at pace 0.2, and we find that the parameter $\lambda = 1$ is suitable for both MI datasets. Therefore, for all experiments, the parameter λ is set to 1. Besides, we conduct the experiment on TensorFlow libraries.

In addition to the above comparing methods, we evaluate the effectiveness of MTDANN with its variants MTDANN-c, MTDANN-a, and MTDANN-v, thus further highlighting the superiority in multi-source fusion. Among them, MTDANN-c concatenates all source domains for generating a larger feature representation as input for EEG classification. MTDANN-a and MTDANN-v stand for the average result and voting result generated by MTDANN based on multiple sources, respectively. The average method denotes that the weight of each source is the same during the prediction, while the voting method selects the class with the most votes as the final result. In the light of that there are only two subjects used as the source domain in the MI2 dataset, therefore the result of voting on this dataset is not reported.

TABLE III
MEAN (%) AND STANDARD DEVIATION (%; IN PARENTHESIS) OF THE BCAS IN S→S TRANSFERS

	MI1	MI2	MI3	Avg
EEGNet	64.77(14.38)	62.39(7.08)	56.32(6.04)	61.16
1D-CNN	67.23(4.08)	65.68(3.41)	61.49(5.23)	64.80
CSP-LDA	55.80(8.50)	49.97(0.81)	57.17(9.57)	54.31
EA-CSP-LDA	65.95(8.88)	64.23(20.39)	67.06(9.81)	65.75
CA	70.79(13.12)	57.01(10.76)	65.75(9.54)	64.51
CA-JGSA	61.55(17.04)	50.41(1.21)	69.40(12.46)	60.45
CA-JDA	66.48(12.92)	57.29(10.35)	68.02(10.50)	63.93
CA-CORAL	74.30(12.25)	56.63(9.42)	66.68(9.85)	65.87
MTDANN	**76.43 (9.38)**	**66.22(8.78)**	**71.90(15.41)**	**71.52**

Moreover, all the datasets in method comparison are class-balance. Thus, in the following these literatures [16], [62], [63], we randomly select 70% data from the target domain set and combine it with the source domain set for training, while the remaining data in the target domain set are used to validate the proposed method. Note that the labeling information of the target domain data is not involved in training. For ensuring the quality of the classification results, we repeat the process ten runs to obtain different partitions, and finally calculate the average result of ten results to make a comparison.

D. Experimental Results Analysis

Table III and Table IV show the experimental result with S→S and M→S transfers, respectively, where the best result among all the methods on each dataset is highlighted in boldface. According to Table III and Table IV, we have a couple of observations:

In Table III, it can be seen from that some well-established transfer learning methods can obtain a good performance on the S→S transfer, such as CA-CORAL on the MI1 dataset, 1D-CNN on the MI2 dataset, and CA-JGSA on the MI3 dataset. However, they are inferior to our proposed method, and our proposed method perform well on both datasets. It indicates that MTDANN can learn the better domain invariant representation from EEG data. This is mainly due to that the three comparing methods perform feature extraction and classification by optimizing different objective functions, and the extracted features may not be optimized for classification. In contrast, our method jointly learns discriminative features and the classifier with an end-to-end optimization strategy, which can leverage adversarial learning to learn crucial information from other domains to achieve optimal result. Similarly, the proposed MTDANN method can compare favorably with the other methods, such as EEGNet, CSP-LDA,

TABLE IV
MEAN (%) AND STANDARD DEVIATION (%; IN PARENTHESIS) OF THE BCAS IN M→S TRANSFERS

	MI1	MI2	MI3	Avg
EEGNet	76.29(15.13)	68.99(8.34)	67.71(14.71)	71.00
1D-CNN	71.79(4.26)	66.82(6.63)	62.86(6.20)	67.16
CSP-LDA	63.64(16.12)	60.09(13.39)	59.93(12.84)	61.22
EA-CSP-LDA	79.57(10.83)	56.18(7.59)	79.79(6.57)	71.85
CA	75.50(11.20)	63.28 (18.43)	72.36(7.62)	70.38
CA-JGSA	65.43(20.76)	51.65(2.86)	63.64(16.43)	60.24
CA-JDA	72.79(10.79)	63.28(18.43)	77.14(9.23)	71.07
CA-CORAL	75.21(7.95)	63.16(18.65)	74.36(10.15)	70.91
MTDANN	**84.05(8.31)**	**70.32(8.02)**	**84.29(11.62)**	**79.55**
MTDANN-c	80.24(8.85)	69.10(7.23)	79.86(10.57)	76.40
MTDANN-a	78.33(9.27)	66.90(5.37)	81.90(14.29)	75.71
MTDANN-v	74.52(10.08)	\	73.33(16.83)	\

EA-CSP-LDA, CA, and CA-JDA. In addition, our proposed method also performs the best on the M→S transfer, as shown in Table IV.

To be specific, MTDANN achieves the best results on the balanced classification accuracy, i.e., 84.05,70.32 and 84.29 on the MI1, MI2 and MI3 datasets, respectively. It can be concluded that the proposed method not only outperforms all the selected comparing methods, but also compares better against MTDANN-c, MTDANN-a, and MTDANN-v. Thus, we can come to that the multi-source fusion mechanism of out method is effective, which is helpful to fuse multiple source domains for EEG classification.

Moreover, we employ the paired t-test to identify whether the performance improvements of the proposed MTDANN method are statistically significant. Since the number of subjects in the three datasets is too small to obtain a convincing result, we combine all the results on MI1, MI2 and MI3 for the paired t-test. Before each t-test, the Lilliefors test [64] is performed to verify that the null hypothesis, that is, the data come from a normal distribution, cannot be rejected. Then, we set the fix significance level α as 0.05 and perform false discovery rate corrections by a linear-step up procedure [65]. Table V shows the false discovery rate adjusted p-values (q-values). From Table V, we can find out that MTDANN significantly outperforms all baselines in S→S and M→S transfers on the three selected datasets. The result further validates the effectiveness of our method.

TABLE V
FALSE DISCOVERY RATE ADJUSTED p-VALUES IN PAIRED t-TEST (SIGNIFICANCE LEVEL $\alpha = 0.05$)

	MTDANN vs	**MI1+MI2+MI3**
S→S	EEGNet	**.000**
	1D-CNN	**.000**
	CSP-LDA	**.000**
	EA-CSP-LDA	**.000**
	CA	**.000**
	CA-CORAL	**.002**
	CA-JDA	**.000**
	CA-JGSA	**.002**
M→S	EEGNet	**.003**
	1D-CNN	**.000**
	CSP-LDA	**.000**
	EA-CSP-LDA	**.035**
	CA	**.001**
	CA-CORAL	**.001**
	CA-JDA	**.009**
	CA-JGSA	**.001**
	MTDANN-c	**.026**
	MTDANN-a	**.006**

TABLE VI
MUTI-SOURCE TRANSFER LEARNING ON THE MI1 DATASET

Target \ **Source**	Subject a1	Subject a2	Subject a3	Subject a4	Subject a5	Fusion Result
Subject a1	\	76.19	59.52	77.38	64.29	**78.57**
Subject a2	89.29	\	73.81	91.67	90.48	**92.86**
Subject a3	71.43	67.86	\	73.81	**75.00**	**75.00**
Subject a4	85.71	89.29	72.62	\	69.05	**92.86**
Subject a5	73.81	78.57	70.24	78.57	\	**80.96**

TABLE VII
MUTI-SOURCE TRANSFER LEARNING ON THE MI2 DATASET

Target \ **Source**	Subject a1	Subject a2	Subject a3	Fusion Result
Subject a1	\	77.42	75.27	**79.56**
Subject a2	57.84	\	64.05	**66.30**
Subject a3	57.63	**65.11**	\	**65.11**

TABLE VIII
MUTI-SOURCE TRANSFER LEARNING ON THE MI3 DATASET

Target \ **Source**	Subject a1	Subject a2	Subject a3	Subject a4	Subject a5	Subject a6	Subject a7	Fusion Result
Subject a1	\	81.67	71.67	76.67	76.67	76.67	78.33	**85.00**
Subject a2	43.33	\	53.33	60.00	**65.00**	56.67	56.67	**65.00**
Subject a3	**75.00**	45.00	\	50.00	70.00	60.00	73.33	**75.00**
Subject a4	81.67	85.00	41.67	\	**95.00**	56.67	**95.00**	**95.00**
Subject a5	46.67	88.33	46.67	95.00	\	91.67	68.33	**98.33**
Subject a6	80.00	76.67	76.67	66.67	**81.67**	\	63.33	**81.67**
Subject a7	83.33	86.67	85.00	**90.00**	85.00	83.33	\	**90.00**

E. The Effectiveness of Muti-Source Transfer Learning

The effectiveness of muti-source transfer learning is validated by comparing the proposal with some other methods, as shown in Section IV-D. We further verify that it can work well in the M→S transfer via data source comparison. Table VI, Table VII and Table VIII summarize the fusion result on the MI1, MI2 and MI3 datasets respectively. From the above above tables, we can observe that MTDANN always obtains the best performance by combining all source domains, and the predicted result based on single source domain compares unfavorably with the fusion result. For example, in Table VI, when subject a1 is set as the target domain, we can see that the balanced classification accuracy is only 59.52 when a3 is set as the source domain, while the best result of the balanced classification accuracy is 77.38 by using a4 as the source domain. By fusing subjects a2-a5 using MTDANN, the performance can be improved and reach to 78.57. It is worth noting that the performance of the model trained by a3 is much worse than the model by a4. This is mainly

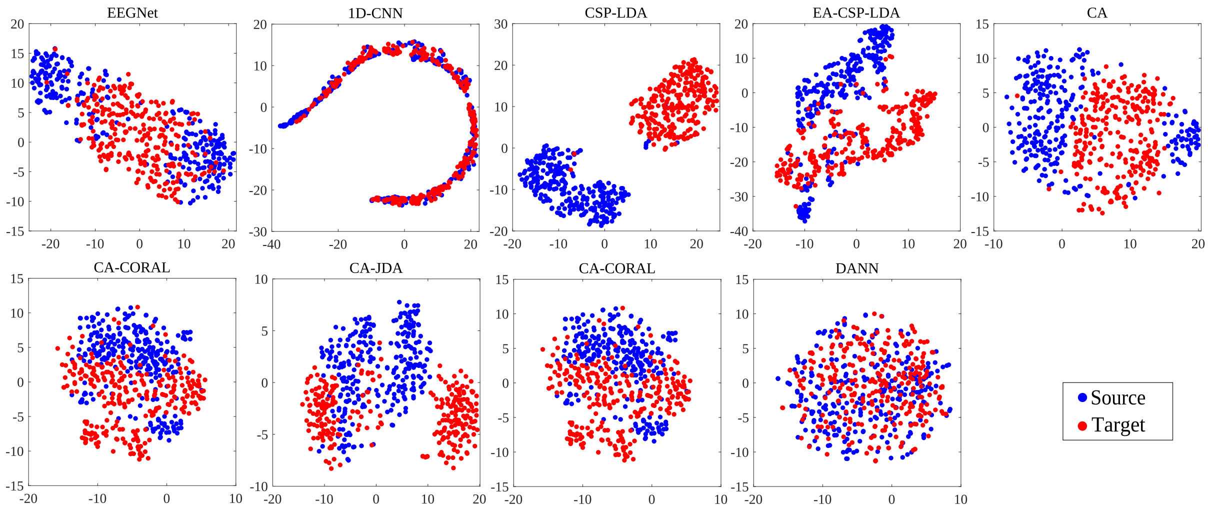

Fig. 3. *t*-SNE visualization of the data distributions with different transfer learning approaches, when transferring subject 1's data (source) to subject 5 (target) on MI1.

due to the potential problem, i.e., training collapse of GAN. The proposed unified multi-source optimization framework can assign larger weight to the informative sources. Following this, the final weighted fusion result is closer to the prediction from data sources with greater performance, thus facilitating a stable and effective result. A similar phenomenon occurs on the other cases. In a nutshell, the combination of different source domains dose help to improve the performance.

F. Visualization

In order to demonstrate the effect of the proposed method in transfer intuitively, we use *t*-SNE [66] to reduce the dimensionality of EEG feature data, and visualize whether the designed method can reduce individual differences effectively. Fig. 3 shows the results on transferring Subject a1's data to Subject a5. Subjects a1 and a5 are come from the MI1 dataset. As shown in Fig. 3, comparing with the other comparing methods, the proposed method can clearly relieve the domain shift. In addition, we can observe that the feature distribution discrepancies are large by using EEGNet, CSP-LDA, EA-CSP-LDA, CA, and CA-JGSA. Accordingly, the adaptability of the features, generated by the above methods, is relatively weak. The cross-subject features from 1D-CNN, CA-JDA and CA-CORAL are closer than those from the other comparing methods, but we can still find out that the features from the source and target domains follow different distributions. Thus, we conclude that the proposed method can make the overall feature distribution consistent, which can give the profit to EEG classification.

In addition, we give an example to illustrate the variation of source weights and the corresponding classification result in terms of each iteration, as shown in Fig. 4. In Fig. 4, we can observe that the initial weight of each source is 0.25 and the corresponding fusion result is 71.43. By iterative optimization, the weights of informative source domains, i.e., a2 and a4, are constantly increased, and vice versa. Simultaneously, the fusion result is continuously optimized with the learned

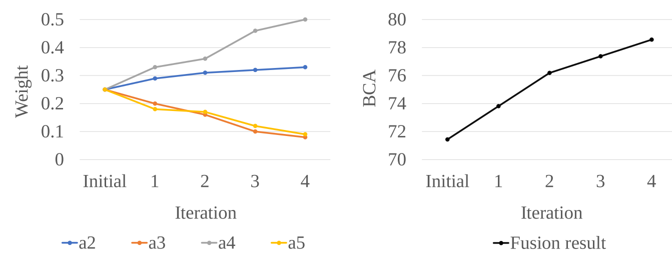

Fig. 4. The iteration of weights and the corresponding fusion result when subject a1 as the target domain and the others as the source domain in dataset MI1.

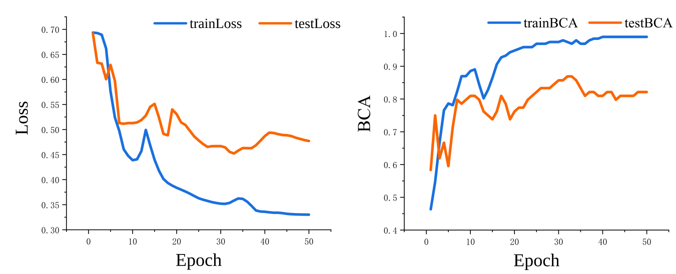

Fig. 5. Convergence visualization of loss and balanced classification accuracy (BCA) on MI1.

source weights. Finally, the most informative source domain, namely a4, gets the maximum weight of 0.5, while the final fusion result is improved to 78.57. Obviously, the proposed iterative optimization framework can inform important sources to obtain a more effective fusion result.

G. Convergence Analysis

We design a domain adversarial network for EEG classification, the objective loss function is shown in Eq. (5), which contains three unknow variable sets, that is, θ_f, θ_y and θ_d, it is the crucial link to study the convergence rate. Once one of the unknown variable sets converges, and the balanced classification accuracy reaches stable.

To show the convergence of the designed domain adversarial network in MTDANN, we conduct the experiment on the MI1 dataset. For demonstration, subject a1 is used as the target domain and subject a2 as the source domain. Fig. 5 illustrates the change of the loss and the balanced classification accuracy with different epochs. Obviously, we can see in Fig. 5 that the value of the loss (the BCA) first decreases (increases) dramatically and then remains stable. Therefore, the proposed method is guaranteed to converge.

V. Conclusion

For EEG classification, we proposed a new multi-source transfer learning method based on domain adversarial neural network in this paper. One main contribution is to propose an improved domain adversarial neural network to extract the domain-invariant features, thus reducing individual differences for EEG classification. More importantly, our method can easily scale to the problem of integrating multiple source domains to achieve the purpose. For that, a weighted fusion method was presented to obtain a more effective and reliable result. Extensive experiments on three EEG MI datasets manifested the advantages of the proposed method. In the future, it is interesting to further investigate the stable EEG classification mechanism with deep adversarial learning.

References

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

[2] J. Xu, S. Mitra, C. Van Hoof, R. F. Yazicioglu, and K. A. A. Makinwa, "Active electrodes for wearable EEG acquisition: Review and electronics design methodology," *IEEE Rev. Biomed. Eng.*, vol. 10, pp. 187–198, 2017.

[3] J. Zhang, G. Sudre, X. Li, W. Wang, D. J. Weber, and A. Bagic, "Clustering linear discriminant analysis for MEG-based brain computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 3, pp. 221–231, Jun. 2011.

[4] G. Schalk and E. C. Leuthardt, "Brain-computer interfaces using electrocorticographic signals," *IEEE Rev. Biomed. Eng.*, vol. 4, pp. 140–154, 2011.

[5] R. Sitaram, N. Weiskopf, A. Caria, R. Veit, M. Erb, and N. Birbaumer, "FMRI brain-computer interfaces," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 95–106, Dec. 2008.

[6] K. Watanabe, H. Tanaka, K. Takahashi, Y. Niimura, K. Watanabe, and Y. Kurihara, "NIRS-based language learning BCI system," *IEEE Sensors J.*, vol. 16, no. 8, pp. 2726–2734, Apr. 2016.

[7] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 4, no. 2, p. R1, Jan. 2007.

[8] V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, and M. Grosse-Wentrup, "Transfer learning in brain-computer interfaces," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 20–31, Feb. 2016.

[9] W.-L. Zheng and B.-L. Lu, "Personalizing EEG-based affective models with transfer learning," in *Proc. 25th Int. Joint Conf. Artif. Intell.*, Jul. 2016, pp. 2732–2738.

[10] M. A. M. Joadder, J. J. Myszewski, M. H. Rahman, and I. Wang, "A performance based feature selection technique for subject independent MI based BCI," *Health Inf. Sci. Syst.*, vol. 7, no. 1, pp. 1–10, Aug. 2019.

[11] W. Zhang and D. Wu, "Manifold embedded knowledge transfer for brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1117–1127, May 2020.

[12] H. Kang and S. Choi, "Bayesian common spatial patterns for multi-subject EEG classification," *Neural Netw.*, vol. 57, pp. 39–50, Sep. 2014.

[13] Z. Lan, O. Sourina, L. Wang, R. Scherer, and G. R. Müller-Putz, "Domain adaptation techniques for EEG-based emotion recognition: A comparative study on two public datasets," *IEEE Trans. Cognit. Develop. Syst.*, vol. 11, no. 1, pp. 85–94, Mar. 2019.

[14] J. Li, S. Qiu, Y.-Y. Shen, C.-L. Liu, and H. He, "Multisource transfer learning for cross-subject EEG emotion recognition," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3281–3293, Jul. 2020.

[15] J. Li, S. Qiu, C. Du, Y. Wang, and H. He, "Domain adaptation for EEG emotion recognition based on latent representation similarity," *IEEE Trans. Cognit. Develop. Syst.*, vol. 12, no. 2, pp. 344–353, Jun. 2020.

[16] H. Zhao, Q. Zheng, K. Ma, H. Li, and Y. Zheng, "Deep representation-based domain adaptation for nonstationary EEG classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 535–545, Feb. 2021.

[17] W. Ko, E. Jeon, J. Lee, and H.-I. Suk, "Semi-supervised deep adversarial learning for brain-computer interface," in *Proc. 7th Int. Winter Conf. Brain-Comput. Interface (BCI)*, Feb. 2019, pp. 1–4.

[18] S. Panwar, P. Rad, T.-P. Jung, and Y. Huang, "Modeling EEG data distribution with a Wasserstein generative adversarial network to predict RSVP events," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 8, pp. 1720–1730, Aug. 2020.

[19] Y. Li, W. Zheng, Y. Zong, Z. Cui, T. Zhang, and X. Zhou, "A bi-hemisphere domain adversarial neural network model for EEG emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 12, no. 2, pp. 494–504, Apr. 2021.

[20] T. Hinz, M. Fisher, O. Wang, and S. Wermter, "Improved techniques for training single-image GANs," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 1300–1309.

[21] I. J. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, Red Hook, NY, USA, 2014, pp. 2672–2680.

[22] H. Wu and M. K. Ng, "Multiple graphs and low-rank embedding for multi-source heterogeneous domain adaptation," *ACM Trans. Knowl. Discovery Data*, vol. 16, no. 4, pp. 1–25, Aug. 2022.

[23] H. Wu, Y. Yan, G. Lin, M. Yang, M. K.-P. Ng, and Q. Wu, "Iterative refinement for multi-source visual domain adaptation," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 6, pp. 2810–2823, Jun. 2022.

[24] H. Wu, Y. Yan, M. K. Ng, and Q. Wu, "Domain-attention conditional Wasserstein distance for multi-source domain adaptation," *ACM Trans. Intell. Syst. Technol.*, vol. 11, no. 4, pp. 1–19, Aug. 2020.

[25] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Dec. 2009.

[26] 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.

[27] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*, vol. 4. Berlin, Germany: Springer, 2006.

[28] M. Dai, D. Zheng, S. Liu, and P. Zhang, "Transfer kernel common spatial patterns for motor imagery brain-computer interface classification," *Comput. Math. Methods Med.*, vol. 2018, pp. 1–9, Mar. 2018.

[29] A. M. Azab, L. Mihaylova, K. K. Ang, and M. Arvaneh, "Weighted transfer learning for improving motor imagery-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 7, pp. 1352–1359, Jul. 2019.

[30] P. Zanini, M. Congedo, C. Jutten, S. Said, and Y. Berthoumieu, "Transfer learning: A Riemannian geometry framework with applications to brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 5, pp. 1107–1116, May 2018.

[31] C. H. Nguyen and P. Artemiadis, "EEG feature descriptors and discriminant analysis under Riemannian manifold perspective," *Neurocomputing*, vol. 275, pp. 1871–1883, Jan. 2018.

[32] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, "Deep learning for motor imagery EEG-based classification: A review," *Biomed. Signal Process. Control*, vol. 63, Jan. 2021, Art. no. 102172.

[33] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, Apr. 2019, Art. no. 031001.

[34] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using EEG signals: A survey," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 374–393, Jul. 2019.

[35] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 5, pp. 804–813, Oct. 2019.

[36] A. M. Azab, H. Ahmadi, L. Mihaylova, and M. Arvaneh, "Dynamic time warping-based transfer learning for improving common spatial patterns in brain–computer interface," *J. Neural Eng.*, vol. 17, no. 1, Feb. 2020, Art. no. 016061.

[37] S. Siddharth, T.-P. Jung, and T. J. Sejnowski, "Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 96–107, Jan. 2022.

[38] G. Xu et al., "A deep transfer convolutional neural network framework for EEG signal classification," *IEEE Access*, vol. 7, pp. 112767–112776, 2019.

[39] X.-Z. Zhang, W.-L. Zheng, and B.-L. Lu, "EEG-based sleep quality evaluation with deep transfer learning," in *Proc. Int. Conf. Neural Inf. Process.* Cham, Switzerland: Springer, 2017, pp. 543–552.

[40] W. Zhang, F. Wang, Y. Jiang, Z. Xu, S. Wu, and Y. Zhang, "Cross-subject EEG-based emotion recognition with deep domain confusion," in *Proc. Int. Conf. Intell. Robot. Appl.* Cham, Switzerland: Springer, Aug. 2019, pp. 558–570.

[41] C. Tan, F. Sun, and W. Zhang, "Deep transfer learning for EEG-based brain computer interface," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 916–920.

[42] U. Demir and G. Unal, "Patch-based image inpainting with generative adversarial networks," 2018, *arXiv:1803.07422*.

[43] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5505–5514.

[44] X. Liang, L. Lee, W. Dai, and E. P. Xing, "Dual motion GAN for future-flow embedded video prediction," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1744–1752.

[45] 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, Aug. 2017.

[46] G. Dai, J. Zhou, J. Huang, and N. Wang, "HS-CNN: A CNN with hybrid convolution scale for EEG motor imagery classification," *J. Neural Eng.*, vol. 17, no. 1, Jan. 2020, Art. no. 016025.

[47] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "EEG classification of motor imagery using a novel deep learning framework," *Sensors*, vol. 19, no. 3, p. 551, 2019.

[48] V. S. Sheng, J. Zhang, B. Gu, and X. Wu, "Majority voting and pairing with multiple noisy labeling," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 7, pp. 1355–1368, Jul. 2019.

[49] 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.

[50] A. Schlögl, "Dataset IIIB: Non-stationary 2-class BCI data," BCI Competition III, Technische Universität Berlin, Berlin, Germany, Tech. Rep., 2005.

[51] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive Berlin brain–computer interface: Fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.

[52] L. Zhu et al., "Multi-source fusion domain adaptation using resting-state knowledge for motor imagery classification tasks," *IEEE Sensors J.*, vol. 21, no. 19, pp. 21772–21781, Aug. 2021.

[53] Y. Wang and D. Wu, "Deep learning for sleep stage classification," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2018, pp. 3833–3838.

[54] F. Mattioli, C. Porcaro, and G. Baldassarre, "A 1D CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface," *J. Neural Eng.*, vol. 18, no. 6, Jan. 2022, Art. no. 066053.

[55] A. J. Izenman, "Linear discriminant analysis," in *Modern Multivariate Statistical Techniques*. Berlin, Germany: Springer, 2013, pp. 237–280.

[56] H. He and D. Wu, "Transfer learning for brain–computer interfaces: A Euclidean space data alignment approach," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 2, pp. 399–410, Apr. 2019.

[57] B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, Mar. 2016, vol. 30, no. 1, pp. 1–8.

[58] M. Long, J. Wang, G. Ding, J. Sun, and P. S. Yu, "Transfer feature learning with joint distribution adaptation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep. 2013, pp. 2200–2207.

[59] J. Zhang, W. Li, and P. Ogunbona, "Joint geometrical and statistical alignment for visual domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 1859–1867.

[60] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, 2011.

[61] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[62] Z. Feng et al., "Design a novel BCI for neurorehabilitation using concurrent LFP and EEG features: A case study," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 5, pp. 1554–1563, Apr. 2022.

[63] K. Zhang, J. Wang, C. W. de Silva, and C. Fu, "Unsupervised cross-subject adaptation for predicting human locomotion intent," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 3, pp. 646–657, Mar. 2020.

[64] H. W. Lilliefors, "On the Kolmogorov–Smirnov test for normality with mean and variance unknown," *J. Amer. Stat. Assoc.*, vol. 62, no. 318, pp. 399–402, 1967.

[65] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Royal Stat. Soc., Ser. B*, vol. 57, no. 1, pp. 289–300, Aug. 1995.

[66] L. V. D. Maaten, "Visualizing datausing t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.