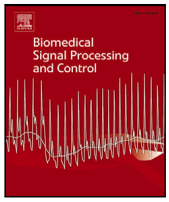
Biomedical Signal Processing and Control 86 (2023) 105138

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Learning a robust unified domain adaptation framework for cross-subject EEG-based emotion recognition☆

Magdiel Jiménez-Guarneros *, Gibran Fuentes-Pineda

Department of Computer Science, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Circuito Escolar s/n, Ciudad Universitaria, Coyoacán, CDMX, 04510, Mexico

ARTICLE INFO

Keywords:
Unsupervised domain adaptation
Deep learning
Emotion recognition
Electroencephalogram

ABSTRACT

Over the last few years, unsupervised domain adaptation (UDA) based on deep learning has emerged as a solution to build cross-subject emotion recognition models from Electroencephalogram (EEG) signals, aligning the subject distributions within a latent feature space. However, most reported works have a common intrinsic limitation: the subject distribution alignment is coarse-grained, but not all of the feature space is shared between subjects. In this paper, we propose a robust unified domain adaptation framework, named Multi-source Feature Alignment and Label Rectification (MFA-LR), which performs a fine-grained domain alignment at subject and class levels, while inter-class separation and robustness against input perturbations are encouraged in coarse grain. As a complementary step, a pseudo-labeling correction procedure is used to rectify mislabeled target samples. Our proposal was assessed over two public datasets, SEED and SEED-IV, on each of the three available sessions, using leave-one-subject-out cross-validation. Experimental results show an accuracy performance of up to 89.11 ± 07.72% and 74.99 ± 12.10% for the best session on SEED and SEED-IV, as well as an average accuracy of 85.27% and 69.58% on all three sessions, outperforming state-of-the-art results.

1. Introduction

Human emotion is an important component in daily life, affecting activities such as communication, decision making or perception [1,2]. Emotion recognition refers to characterizing and identifying the responses of human emotions through machines using voice, body postures, facial expressions, and physiological signals. Recently, emotion recognition based on physiological signals, represented by electroencephalography (EEG), has attracted attention due to its objectivity and sensitivity to the emotional reactivity [3]. Several studies have reported successful results for emotion recognition using EEG signals, but most of these studies are focused on designing subject-dependent classifiers, i.e., they assume that data maintain the same distribution along training and testing phases. Consequently, these classifiers tend to under-perform with new users due to EEG signals vary across subjects and maintain a high non-stationarity [4–6]. To overcome this issue, subject-independent or cross-subject classifiers have been proposed in the literature [7], leveraging on labeled data from known subjects to classify unknown data from a new subject. Thus, this research focuses on building cross-subject classifiers to enable the applicability of existing emotion recognition systems to new users.

Existing works have used two main approaches to address cross-subject EEG-based signal recognition: (1) generic and (2) unsupervised domain adaptation (UDA) methods [5,8]. The first one assumes that a set of invariant features can be extracted across different subjects, and then, a generic classifier can be trained to recognize samples belonging to new subjects. Moreover, domain adaptation aims to transfer useful knowledge from known subjects to a new subject. This knowledge transfer consists in reducing the dataset shift problem, that is, the divergence of the joint distributions of EEG data between subjects [4,9,10]. For this, UDA assumes the availability of a labeled source domain (data from known subjects) and an unlabeled target domain (data from a new subject) during a calibration process [11]. Indeed, generic classifiers have been designed to completely suppress the calibration effort time, while UDA methods aim to reduce it [5]. Despite this, domain adaptation has played an important role in building cross-subject classifiers, as generic classifiers are still unfeasible due to high structural and functional variability between subjects and the non-stationarity of EEG signals [5,12,13]. Recently, some UDA methods have been reported in Brain–Computer Interfaces (BCI) [14–17] to significantly reduce the calibration effort time, using a reduced set of unlabeled samples from a new subject, and leveraging on pre-trained

☆ This work was partially supported by the National Council of Science and Technology in Mexico (CONACYT).
* Corresponding author.
E-mail address: mjmnzg@gmail.com (M. Jiménez-Guarneros).

https://doi.org/10.1016/j.bspc.2023.105138
Received 5 January 2023; Received in revised form 11 May 2023; Accepted 8 June 2023
Available online 22 June 2023

models over known subjects. However, the reported methods still suffer from overfitting, and results show high variability. Particularly in EEG-based emotion recognition, UDA solutions [18–21] have mainly focused on addressing the dataset shift problem to improve the knowledge transfer between subjects. Nevertheless, reported UDA solutions have not yet shown to be effective and robust, even using all available data from known subjects and a new subject during the calibration process. Thus, the presented research is focused on how to perform a better knowledge transfer between subjects, such that effective and robust classifiers can be built on new subjects in EEG-based emotion recognition.

Over the last few years, with the rapid progress of artificial intelligence, UDA methods based on deep learning (D-UDA) have been under the spotlight due to their high capacity to learn transferable features between subjects [11,22,23]. In this sense, a first group of works [4,22,24] adopted divergence measures to reduce the divergence between the marginal distributions of labeled data from known subjects and unlabeled data from a new subject. Indeed, a change in the joint distributions suggests differences in marginal and conditional distributions. Thus, some works have been reported to encourage a similarity in marginal and conditional distributions among subjects by learning a feature representation with similar statistics [21,25,26]. Despite these advances, several studies have been focused on designing specialized deep neural network architectures to learn relationships between different EEG channels, while only marginal distributions between subject data are reduced via adversarial training [18,27–30]. Likewise, attention mechanisms have been adopted to transfer information between brain regions [19]. However, while conventional UDA aims to transfer knowledge from a labeled source domain to an unlabeled target domain, the distributions to be aligned come from multiple known subjects, which requires exploiting fine-grained structures. In this sense, multi-source domain adaptation (MDA) arises as a practical and useful setting in which knowledge is transferred from multiple distinct source domains to a single unlabeled target domain.

For multi-source domain adaptation, domain-specific distributions between each source and target domains are aligned to learn domain-invariant representations [31]. Indeed, some efforts have been reported in emotion recognition that assume multiple source distributions from different subjects. For example, remarkable methods were proposed in [32,33] to reduce the distributions between the target subject and the most closely related source subjects. Nevertheless, labels of the target subject are considered during the domain adaptation process, which is not allowed in UDA. Following the unsupervised learning paradigm, authors in [34] used regularization based on a correlation metric to align the multiple subject distributions. Reported works in [35,36] designed domain-specific feature extractors to transfer specific knowledge from each specific subject. Although the proposed methods are only focused on distribution alignment between each pair of subjects, other levels of fine-grained alignment are not considered, as not all of the feature space is shared between subjects. Likewise, an appropriate strategy is not used to encourage an inter-class separation during the domain alignment process, since such separation is assumed. As consequence, MDA methods still present a low performance in comparison with methods that assume a single-source domain. Thus, there is still a gap for building robust and accurate cross-subject classifiers in EEG-based emotion recognition, as a fine-grained transfer has not been exploited effectively.

In this paper, we propose a robust unified domain adaptation framework, named *Multi-source Feature Alignment and Label Rectification* (MFA-LR), which is motivated by the observation that not all of the feature space is shared between subjects. Firstly, we perform a fine-grained distribution alignment based on a weighted alignment loss, which adopts Central Moment Discrepancy [37] and Semantic loss [38] to perform a domain alignment at subject and class levels. At the same time, a discriminative feature representation is enforced over the entire feature space of all source subjects via classification loss, instead of specific feature spaces per subject. Subsequently, we impose the conditional entropy loss [39] over the target subject to prevent learned features from being located near the decision boundary during the domain alignment process. During training, standardization and normalization in spherical space [40,41] are applied over learned features to maintain similar statistics across different subjects. Likewise, we introduce Gaussian noise over the learned feature space to encourage robustness against input perturbations. Finally, a pseudo-labeling correcting procedure [42,43] is used to rectify mislabeled target samples. We evaluate our proposal using two public datasets, SEED [44] and SEED-IV, with leave-one-subject-out cross-validation for each available session.

The main contributions of this paper are summarized as follows:

- We proposed a robust unified domain adaptation framework to perform a better transfer of useful knowledge among subjects in cross-subject EEG-based emotion recognition. We perform a fine-grained domain alignment at subject and class levels, while an inter-class separation is ensured in a coarse-grained manner over the entire feature space of all source subjects and the target subject. This domain alignment is leveraged on normalization strategies and Gaussian noise to maintain similar statistics across subjects and encourage robustness against input perturbations. Then, a pseudo-labeling correcting procedure is applied to rectify wrongly pseudo-labeled target samples.
- We conducted experiments on the SEED and SEED-IV datasets for emotion recognition research to validate the effectiveness and robustness of our proposed framework, using the standard validation strategy on each of the three available sessions of EEG data. Our proposal achieves state-of-the-art results, averaging 85.27% and 69.58% across the three available sessions on SEED and SEED-IV.
- The proposed framework allows building robust adaptive classifiers for new subjects, as it maintains the highest classification results throughout EEG data recorded in separate periods, distinct partitions of target data, and initial neural network weights.

The paper is organized as follows: in Section 2, we review previous works on D-UDA for emotion recognition. Section 3 describes the proposed framework in detail. Section 4 presents datasets, pre-processing, evaluation and the corresponding results, while the discussion is reported in Section 5. Finally, conclusions and future work are reported in Section 6.

2. Related work

Single-source domain adaptation. Unsupervised domain adaptation has emerged as a helpful solution for cross-subject emotion recognition. Regarding shallow approaches, a symmetric and positive definite matrix network was proposed in [45] to perform feature and sample adaptation between subjects with distribution confusion and centroid alignment, respectively. Authors in [46] performed a marginal and conditional distribution alignment via Maximum Mean Discrepancy (MMD), leveraged on label estimation to accurately provide pseudo-labels over the target domain. Similarly, a joint distribution alignment is performed in [47] via MMD, but graph adaptive label propagation is used to estimate the target labels. For deep learning approaches, authors in [4,48] used the MMD measure to reduce the divergence between marginal distributions of different subjects. Similarly, adversarial methods have been adopted to learn domain-invariant features by using a minimax training between the feature extractor and a domain discriminator [49]. Luo et al. [22] used Wasserstein adversarial training to learn a specific subspace on the target domain by using an independent network. The reported works in [27,30] presented a bi-hemispheric network architecture to learn relationships between different EEG channels, while adversarial training is used to

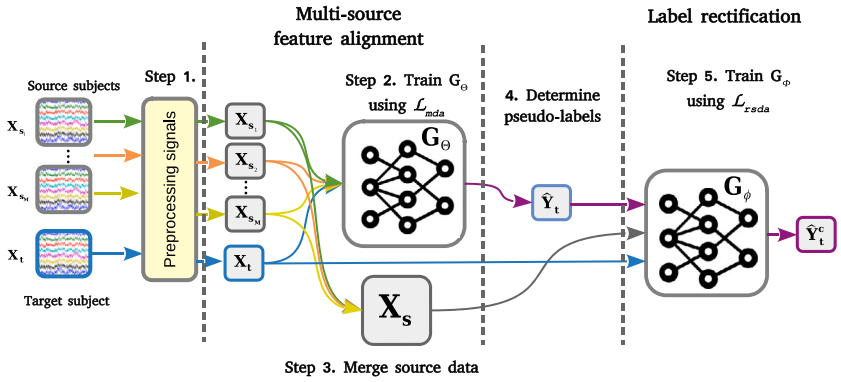


Fig. 1. MFA-LR procedure. EEG raw signals are preprocessed for each subject and fed to the Multi-source Feature Alignment. In this stage, a fine-grained distribution alignment is performed by aligning distributions at subject and class levels, while inter-class separation and robustness against random perturbations are ensured over the global feature space. After, pseudo-labels $\hat{\mathbf{Y}}_t$ are determined by G_Θ. Then, a label rectification $\hat{\mathbf{Y}}_t^c$ is produced by training G_ϕ, assuming a single-source domain adaptation.

reduce the divergence between marginal distributions. In a similar way, authors in [18,29] employed graph neural networks to characterize the intrinsic dynamic relationships between different EEG channels. More recently, the reported work in [19] adopted an attention mechanism strategy to exploit the global and local transferability of EEG signals for emotion recognition. Motivated by differences in joint distributions, Li et al. [25] combined adversarial training with associative domain adaptation (ADA) [50] to reduce the divergence of the marginal and conditional distributions among domains, enforcing feature representations with similar statistics. Likewise, Zhu et al. [21] combined Wasserstein adversarial training with ADA over an auto-encoder network to increase the similarity between marginal and conditional distributions. Despite these advances, previous works assume that a single-source domain adaptation is available, but usually data are collected from multiple subjects, which is a challenge when handling the domain shift.

Multi-source domain adaptation. The reported MDA approaches for EEG-based emotion recognition minimize the domain discrepancy between each source subject and the target subject. Authors in [32,33] used specific classifiers for each source subject in order to identify and align the most similar subjects with regard to the target subject. Although remarkable results were obtained, the cited works utilized labeled target data to guide the distribution alignment between domains. Indeed, target labels are not provided in unsupervised domain adaptation. Thus, Tao et al. [34] proposed a regularizer based on the L1/L2 norm and a correlation metric to align each pair of distributions between labeled data from a source subject and unlabeled data from a target subject. Authors in [35,36] designed domain-specific feature extractors and classifiers for each specific source subject, while explicit measures of discrepancy as MMD and L1 norm were used to reduce the divergence between the marginal and conditional distributions, respectively. However, despite these advances, the classification results are still outperformed by existing D-UDA methods that assume a single-source domain. The reported MDA solutions are only focused on aligning the marginal distributions between each source subject and a new subject. However, these works do not consider the fact that classes may be misaligned between each pair of subjects, which could encourage performing a better fine-grained alignment. Furthermore, a discriminative feature space is not ensured over the entire feature space of all source subjects and the new subject, as an inter-class separation is assumed in EEG data. Likewise, it is known that samples from a new subject may be mislabeled, so that a correction procedure may help avoid a negative transfer. In this work, we address the described problems to achieve a better transfer of useful knowledge between subjects.

3. Proposed framework

We assume M labeled source domains $D_1, D_2, \ldots, D_M$ and an unlabeled target domain $\mathcal{T}$. $\{D_m\}_{m=1}^M$ represent the information from known subjects, while $\mathcal{T}$ is the information from a new subject. Each source domain D_m contains N_m labeled samples $(\mathbf{X}_{s_m}, \mathbf{Y}_{s_m}) = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{N_m}$, drawn from a distribution $P_{s_m}(\mathbf{X}_{s_m}, \mathbf{Y}_{s_m})$ defined over $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{X}$ is an input variable space and $\mathcal{Y}$ is a label space. The target domain $\mathcal{T}$ contains N_t unlabeled samples $\mathbf{X}_t = \{\mathbf{x}_i^t\}_{i=1}^{N_t}$ drawn from a distribution $P_t(\mathbf{X}_t, \mathbf{Y}_t)$, where $\{P_{s_m}(\mathbf{X}_{s_m}, \mathbf{Y}_{s_m})\}_{m=1}^M \neq P_t(\mathbf{X}_t, \mathbf{Y}_t)$. We denote a deep neural network model as a labeling function G with trainable weights Θ, where G may be represented as a composite of two functions $G_f \circ G_h$; parameterized by Θ_f and Θ_h, respectively. Here, G_f is a feature extractor network that encodes an input $\mathbf{x}$ into a latent feature representation $\mathbf{z}$, that is, $\mathbf{z} = G_f(\mathbf{x}; \Theta_f)$. Meanwhile, G_h is a feature labeling function to produce a classification score $\hat{\mathbf{y}}$, i.e., $\hat{\mathbf{y}} = G_h(\mathbf{z}; \Theta_h)$. Our goal is to learn a labeling function $G : \mathcal{X} \rightarrow \mathcal{Y}$ that correctly predicts labels on the target domain, mitigating the domain distribution shift between multiple labeled source domains and an unlabeled target domain. In this research, the labeled EEG data coming from each subject represent a domain D.

Fig. 1 presents the MFA-LR procedure. Our proposal comprises two main training stages: multi-source feature alignment and label rectification. The multi-source feature alignment module aligns subject distributions in a fine-grained way; meanwhile, inter-class separation and robustness against random perturbations are encouraged in coarse grain over the entire feature space. Then, a pseudo-labeling correction procedure is used to rectify those mislabeled target samples.

We describe every component in each MFA-LR module in detail. Finally, we describe the learning procedure of the proposed framework.

3.1. Multi-source feature alignment

The Multi-source feature alignment (MFA) module is shown in Fig. 2 and the overall objective is written as follows

$$\begin{aligned}\mathcal{L}_{\text{mda}}(\{\mathbf{X}_{s_m}, \mathbf{Y}_{s_m}\}_{m=1}^M, \mathbf{X}_t, G_\Theta; \Theta) = &\mathcal{L}_{\text{cls}}(\mathbf{X}_s, \mathbf{Y}_s, G_\Theta; \Theta) + \\ &\alpha \cdot \textstyle\sum_{m=1}^M [\omega_{\text{dis}}^{(m)} \cdot \mathcal{L}_{\text{dis}}(\mathbf{X}_{s_m}, \mathbf{X}_t, G_f^\Theta; \Theta_f) + \\ &\omega_{\text{sem}}^{(m)} \cdot \mathcal{L}_{\text{sem}}(\mathbf{X}_{s_m}, \mathbf{X}_t, G_f^\Theta; \Theta_f)] + \\ &\lambda_t \cdot \mathcal{L}_{\text{ce}}(\mathbf{X}_t, G_\Theta; \Theta)\end{aligned} \tag{1}$$

where $\mathcal{L}_{\text{cls}}$ is the classification loss; $\mathcal{L}_{\text{dis}}$ and $\mathcal{L}_{\text{sem}}$ are discrepancy losses to reduce the distribution differences at the subject and class levels, respectively. $\mathcal{L}_{\text{ce}}$ is a discrimination loss to encourage a discriminative feature space over the target subject. α and λ_t are trade-off hyper-parameters. $\omega_{\text{dis}}^{(m)}$ and $\omega_{\text{sem}}^{(m)}$ are dynamic weights to determine the

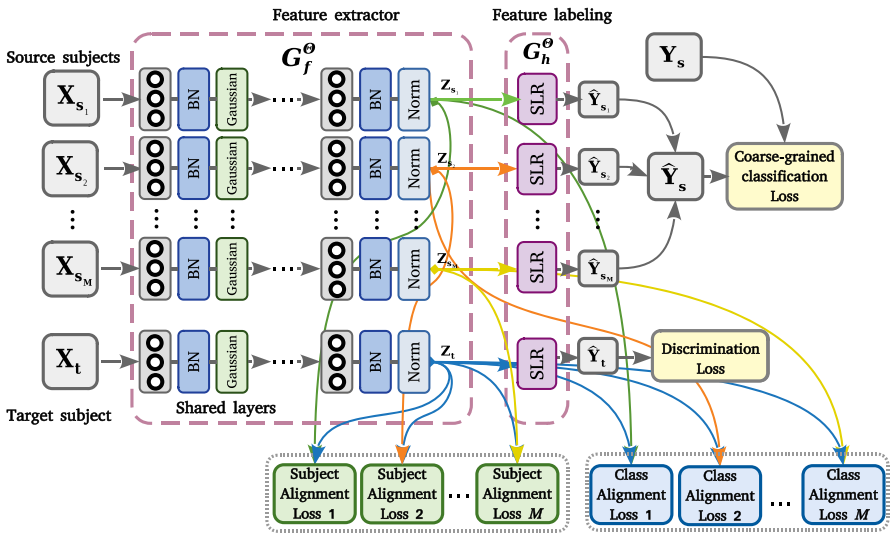


Fig. 2. The multi-source feature alignment framework. In the feature extractor G_f^{Θ}, Batch Normalization (BN) was included after each hidden layer for feature normalization. Then, a Gaussian noise layer was incorporated to obtain more robust features. At the output of G_f^{Θ}, l_2-normalization (Norm) produces a feature representation in spherical space. Subsequently, the Spherical Logistic Regression (SLR) layer was employed to make predictions into the feature labeling G_h^{Θ}. Note that learned features $\mathbf{Z}_{s_1}, \ldots, \mathbf{Z}_{s_M}, \mathbf{Z}_t$ are aligned at subject and class levels, while inter-class separation is ensured over the global feature space through coarse-grained classification and discrimination losses.

relevance of the m-*th* source subject, motivated by the fact that not all source subjects are equally transferable to the target domain. These dynamic weights are defined as follows

$$a_m = \frac{1}{\mathcal{L}_{\text{dis}}(\mathbf{X}_{s_m}, \mathbf{X}_t; G_f^{\Theta})}, \tag{2}$$

$$\omega_{\text{dis}}^{(m)} = \frac{\exp(a_m)}{\sum_{j=1}^{M} \exp(a_j)}, \tag{3}$$

where a_m and a_j denote the fitness of the mth and jth source subjects.

As shown in Fig. 2, the neural network G_{Θ} is trained using the loss function $\mathcal{L}_{\text{mda}}$. This neural network takes preprocessed EEG data for each specific subject. Learned features from each subject are standardized by introducing a batch normalization (BN) [40] after each hidden layer. Then, Gaussian noise is applied to the learned features to encourage a more robust representation against random perturbations. At the output of the feature extractor, l_2-normalization [41] is used to map features to a spherical space. Subsequently, a feature distribution alignment is performed at subject and class levels by imposing the discrepancy losses $\mathcal{L}_{\text{dis}}$ and $\mathcal{L}_{\text{sem}}$. For feature labeling, we use the spherical logistic regression (SLR) layer [43], designed to predict classification scores on a spherical space. Note that predictions from all source subjects are mixed in one set $\mathbf{Y}_s$ and processed by the coarse-grained classification loss $\mathcal{L}_{\text{cls}}$ to encourage intra-class compactness and inter-class separation over the global feature space from the source subjects. The discrimination loss $\mathcal{L}_{\text{ce}}$ is imposed to the target domain, in order to prevent the learned target features from being located near the decision boundary.

Coarse-grained Classification Loss. The cross-entropy loss [51] is utilized for data from source domains with ground-truth class labels, defined by

$$\mathcal{L}_{\text{cls}}(\mathbf{X}_s, \mathbf{Y}_s, G_{\Theta}; \Theta) = -\mathbb{E}_{\mathbf{x}_s, \mathbf{y}_s \sim (\mathbf{X}_s, \mathbf{Y}_s)} \sum_{k=1}^{K} \mathbf{1}_{(k=\mathbf{y}_s)} \log G_{\Theta}(\mathbf{x}_s), \tag{4}$$

where $(\mathbf{X}_s, \mathbf{Y}_s) = \{\mathbf{X}_{s_m}, \mathbf{Y}_{s_m}\}_{m=1}^{M}$ is the dataset combining all data from every source subject; K is the number of classes and $\mathbf{1}$ is an indicator function which outputs 1 only if $G_{\Theta}(\mathbf{x}_s)$ is equal to class k.

Subject-Level Alignment Loss. We enforce similar distribution statistics between subjects by using a Central Moment Discrepancy (CMD) [37]. This discrepancy measure matches the highest order statistic moments between feature representations from multiple source subjects and the target subject. The high-order moment distance between each source subject $\mathcal{D}_m$ and the target subject $\mathcal{T}$ is defined as follows:

$$\begin{aligned}\mathcal{L}_{\text{dis}}(\mathbf{X}_{s_m}, \mathbf{X}_t, G_f^{\Theta}; \Theta_f) = \; & \left\| \mathbb{E}(G_f^{\Theta}(\mathbf{X}_{s_m})) - \mathbb{E}(G_f^{\Theta}(\mathbf{X}_t)) \right\|_2 + \\ & \sum_{k=2}^{K} \left\| B_k(G_f^{\Theta}(\mathbf{X}_{s_m})) - B_k(G_f^{\Theta}(\mathbf{X}_t)) \right\|_2\end{aligned} \tag{5}$$

where K is the maximum order of the central moments, while the term $B_k(\cdot)$ is the central moment of order k.

Class-Level Alignment Loss. The semantic loss [38] is employed to explicitly match the feature distributions of the same class between the multiple source domains and the target domain. For this, pseudo labels are first assigned to the target samples to cope with the lack of target label information. Then, the centroid of each class is aligned between the source subject $\mathcal{D}_m$ and target subject $\mathcal{T}$ as follows

$$\mathcal{L}_{\text{sem}}(\mathbf{X}_{s_m}, \mathbf{Y}_{s_m}, \mathbf{X}_t, \hat{\mathbf{Y}}_t, G_f^{\Theta}; \Theta_f) = \sum_{k=1}^{K} \text{dist}(C_{s_m}^k, C_t^k) \tag{6}$$

where $C_{s_m}^k$ and C_t^k are the centroids for the class k of the source domain $\mathcal{D}_m$ and target domain $\mathcal{T}$, respectively; the cosine distance $\text{dist}(u, v) = 1 - \|u\|v\|$ is used to measure the differences between centroids. Note that pseudo-labels $\hat{\mathbf{Y}}_t$ are estimated by $G_{\Theta}(\mathbf{X}_t)$, which may be correct or incorrect, so moving average centroid is used as a safer aligning.

Target Discrimination Loss. We encourage a low-density separation between classes on the target domain data via discrimination loss, minimizing the entropy of the class conditional distribution [39]:

$$\mathcal{L}_{\text{ce}}(\mathbf{X}_t, G_{\Theta}; \Theta) = -\mathbb{E}_{\mathbf{x}_t \sim D_t}[G_{\Theta}(\mathbf{x}_t)^{\mathrm{T}} \ln G_{\Theta}(\mathbf{x}_t)] \tag{7}$$

3.1.1. *Normalization and perturbation of learned features*

The proposed framework integrates normalization and noise perturbation layers into a deep neural network to promote the feature adaptation process. As shown in Fig. 2, we leverage these layers to construct the feature extractor G_f^{Θ} and feature labeling G_h^{Θ}. Each layer is described hereafter.

Batch Normalization. The BN layer [40] is incorporated into the feature extractor network G_f^{Θ} to maintain similar mean and variance across different subjects. For this, BN standardizes each jth output feature x_j in the layer (l) as follows

$$\hat{x}_j^{(l)} = \gamma_j^{(l)} \cdot \frac{(x_j^{(l)} - \mu_j^{(l)})}{\sigma_j^{(l)}} + \beta_j^{(l)} \tag{8}$$

where μ_j and σ_j^2 are the mean and variance of the current batch of samples for the jth feature; γ_j and β_j are parameters of translation and scale, respectively. The mean and variance are updated during training by using data from source and target subjects.

Gaussian Noise. It is introduced to make features robust to common variations since random noise may be present in EEG signals. For this, a noised feature vector $\mathbf{q} \in \mathbb{R}^d$ is added to each learned feature sample $\mathbf{z}^{(l)}$ [52,53] in layer (l) following

$$\begin{aligned} &\mathbf{q} \sim \mathcal{N}(0, v) \\ &\hat{\mathbf{z}}^{(l)} = \mathbf{z}^{(l)} + \mathbf{q}, \end{aligned} \tag{9}$$

where $\mathbf{q}$ is drawn from a normal distribution with zero-mean and variance v.

l_2-normalization. The output features of the feature extractor G_f^Θ are normalized in spherical space $S_r = \{\bar{\mathbf{z}} \in \mathbb{R}^d : \|G_f^\Theta(\mathbf{x})\| = r\}$ applying l_2-normalization [41,42]. The aim is to enforce balanced magnitudes of feature vectors across all classes. l_2-normalization makes the direction of the features from the same class closer to each other and separates different classes [41]. The output features are normalized by using

$$\bar{\mathbf{z}} = r \cdot \frac{G_f^\Theta(\mathbf{x})}{\|G_f^\Theta(\mathbf{x})\|_2}, \tag{10}$$

where $\bar{\mathbf{z}}$ is the normalized vector, while r is the ratio to encourage a feature spherical space.

Spherical Logistic Regression. (SLR) layer [42] is used to estimate the output scores of each category k in spherical space as follows

$$P(y = k|\mathbf{z}) \propto \exp(\mathbf{w}_k^T \mathbf{z} + \mathbf{b}_k), k = 1, 2, \ldots, K \tag{11}$$

where $\mathbf{w}_k \in \mathbb{R}^n$ is a unit norm vector $\|\mathbf{w}_k\| = 1$ and $\mathbf{b}$ is the bias in $[-r, r]$.

3.2. Rectifying target labels

We use Robust Spherical Domain Adaptation (RSDA) [42,43], motivated from [54], to rectify the mislabeled target samples. Fig. 3 shows the adapted RSDA framework. The RSDA method performs a standard single-source domain adaptation, weighting the importance of pseudo-labeled target samples via the posterior probability of correct labeling. For this, RSDA introduces the spherical adversarial training loss:

$$\begin{aligned} \mathcal{L}_{\text{rsda}}(\mathbf{X}_s, \mathbf{Y}_s, \mathbf{X}_t, \hat{\mathbf{Y}}_t, G_{\boldsymbol{\Phi}}; \boldsymbol{\Phi}) = \quad & \mathcal{L}_{\text{bas}}(\mathbf{X}_s, \mathbf{Y}_s, \mathbf{X}_t, G_{\boldsymbol{\Phi}}; \boldsymbol{\Phi}) + \\ & \mathcal{L}_{\text{rob}}(\mathbf{X}_t, \hat{\mathbf{Y}}_t, G_{\boldsymbol{\Phi}}; \boldsymbol{\Phi}) + \\ & \lambda_{ent} \cdot \mathcal{L}_{\text{ce}}(\mathbf{X}_t, G_{\boldsymbol{\Phi}}; \boldsymbol{\Phi}), \end{aligned} \tag{12}$$

where $\mathcal{L}_{\text{bas}}$ is the basic loss, $\mathcal{L}_{\text{rob}}$ is the robust pseudo-label loss and $\mathcal{L}_{\text{ce}}$ is the conditional entropy loss; $\hat{\mathbf{Y}}_t$ are the target pseudo-labels and $\boldsymbol{\Phi}$ is the set of trainable weights. The basic loss is based on Moving Semantic Transfer Network (MSTN) [38], while the robust pseudo-label loss [42,43] is defined as:

$$\mathcal{L}_{\text{rob}}(\mathbf{X}_t, \hat{\mathbf{Y}}_t, G_{\boldsymbol{\Phi}}; \boldsymbol{\Phi}) = \frac{1}{N_0} \sum_{j=1}^{N_t} w_{\boldsymbol{\Phi}}(\mathbf{x}_j^t) \mathcal{J}(G_{\boldsymbol{\Phi}}(\mathbf{x}_j^t), \hat{\mathbf{y}}_j^t) \tag{13}$$

where $N_0 = \sum_{j=1}^{N_t} w_{\boldsymbol{\Phi}}(\mathbf{x}_j^t)$, $\mathcal{J}$ is the mean absolute error, $w_{\boldsymbol{\Phi}}(\mathbf{x}_j^t)$ is the posterior probability of correct labeling

$$w_{\boldsymbol{\Phi}}(\mathbf{x}_j^t) = \begin{cases} \varphi_j & \text{if } \varphi_j \geq 0.5, \\ 0 & \text{otherwise}, \end{cases} \tag{14}$$

in which $\varphi_j = P_{\boldsymbol{\Phi}}(z_j = 1|\mathbf{x}_j^t, \hat{\mathbf{y}}_j^t)$. $z_j \in 0, 1$ is a random variable of each pseudo-labeled sample $(\mathbf{x}_j^t, \hat{\mathbf{y}}_j^t)$ to indicate whether the sample is correctly or wrongly labeled. The probability $P_{\boldsymbol{\Phi}}$ is modeled by using a Gaussian-uniform mixture model in order to measure the feature distance from each sample to the center of each class. Target samples with a probability less than 0.5 are discarded.

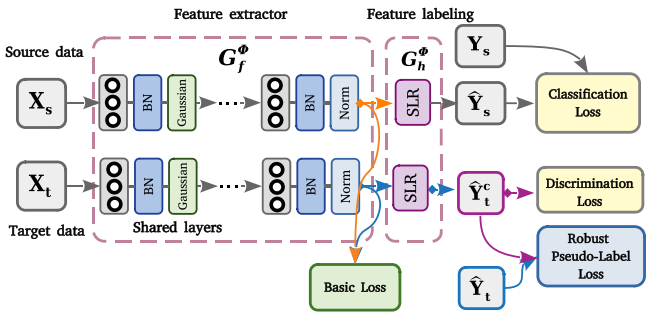


Fig. 3. The RSDA framework. The rectification of mislabeled target samples is performed assuming a single-source domain adaptation; data from source subjects are mixed in a single source dataset. Normalization and perturbation of learned features are also performed at this stage.

3.3. Learning procedure

We summarize the training procedure of MFA-LR. The neural network G_Θ is trained by optimizing $\mathcal{L}_{\text{mda}}$ on $\{(\mathbf{X}_s, \mathbf{Y}_s)\}_{m=1}^M$ and $\mathbf{X}_t$. The pseudo-labels $\hat{\mathbf{Y}}_t$ are estimated for the target data using G_Θ. A second neural network $G_{\boldsymbol{\Phi}}$ is trained by optimizing $\mathcal{L}_{\text{rsda}}$, receiving the mixed source data $(\mathbf{X}_s, \mathbf{Y}_s)$ and pseudo-labeled target data $(\mathbf{X}_t, \hat{\mathbf{Y}}_t)$. This stage is repeated to rectify the assigned pseudo-labels on the target samples. Exponential Moving Average (EMA) [55] is incorporated to stabilize the training of G_Θ and $G_{\boldsymbol{\Phi}}$. EMA is defined on the weights Θ in the training step n as

$$\Theta_{\text{ema}}^{(n)} = \lambda_{\text{ema}} \cdot \Theta_{\text{ema}}^{(n-1)} + (1 - \lambda_{\text{ema}}) \cdot \Theta^{(n)}, \tag{15}$$

where λ_{ema} is the decay rate.

4. Experiments

In this section, we present the evaluation results of the proposed framework in emotion recognition. First, we introduce the datasets, feature extraction, implementation details and evaluation setup. Then, we present the experimental results and an ablation study. Implementation is available at https://github.com/mjmnzg/MFA-LR.

4.1. Experimental setup

Datasets. We conduct the experiments on two public EEG emotion datasets: SEED and SEED-IV. Table 1 shows the properties of both datasets.

1. **SEED** [44] dataset is a standard benchmark that contains EEG emotional signals from 15 subjects watching 15 Chinese movie clips; each clip lasted about 4 min. Three types of emotions were induced: positive, neutral and negative. EEG signals were recorded at a sampling rate of 1 kHz from 62-channel electrode caps using the ESI NeuroScan device. EEG data contain recordings for three different periods at approximately 1-week intervals, corresponding to three sessions, where each subject has 45 EEG data trials. The trial of each channel is divided into non-overlapping one second segments.
2. **SEED-IV** [56] dataset contains EEG signals from 15 subjects while they were watching short movie clips (2 min) to induce four emotions: happy, sad, neutral, and fear. This dataset has three sessions, including 24 EEG trials for each subject. EEG data were collected using a 62-channel ESI NeuroScan system, while electrode locations are based on the international 10–20 system.

Table 1
Properties of SEED and SEED-IV.

Dataset	SEED	SEED-IV
Subjects	15 (7 male/8 female)	15 (7 male/8 female)
Channels	62	62
Trials	15 (5 per emotion)	24 (6 per emotion)
Sampling rate	1000 Hz	1000 Hz
Downsampled rate	200 Hz	200 Hz
Labels	Positive, Neutral, Negative	Happy, Sad, Neutral and Fear

Data pre-processing and feature extraction. We directly used the pre-computed differential entropy (DE) features provided by the SEED and SEED-IV datasets. The feature extraction procedure is the reported in [44,56]. The raw EEG signals are downsampled to 200 Hz for both datasets and filtered at 1–75 Hz using a bandpass filter. The filtered signals are sliced into 1 and 4 s segments for SEED and SEED-IV, respectively. Differential entropy (DE) features are extracted from each segment at five frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–14 Hz), beta (14–31 Hz), and gamma (31–50 Hz) [44,57]. An EEG trial from one subject in each session for both databases has the form: channel (62) × trial (15 for SEED, 24 for SEED-IV) × band (5). The feature dimension is 310 (62 channels × 5 frequency bands). In addition, the EEG features are smoothed to minimize the artifacts in the EEG features using the linear dynamic system approach [58].

For SEED, as the length of EEG each trial is not the same for each stimulation fragment, we cut the EEG trials (taking the last points) with an identical length to ensure the same number of emotional samples [33]. Then, we obtain 2775 samples from 15 trials (185 samples × trial). Following [59], we concatenate the DE features of each second along a specific time length in each EEG trial to build a temporal feature matrix, that is, window length × DE features (310). Here, we use a window length of 9 s and an overlap of 8, as suggested in [59]. In the end, EEG data of each subject in a single session are formed into 2655 samples × 9 time windows × 310 dimensions with the corresponding generated labels.

For SEED-IV, each of the three sessions has 24 trials containing 851, 832, and 822 samples, respectively. Motivated by the fact that multiple types of features can provide complementary information [60], we concatenate DE features using two types of smoothing: linear dynamic system and moving average. Thus, all EEG data have the form 851/832/822 samples × 620 dimensions (310 × 2 types of smoothing).

For both datasets, pre-processed EEG data of each subject in every session are normalized by subtracting their mean and then dividing by their standard deviation.

Neural network architecture. We use a Deep Feed-forward Network (DFN) as model G to solve the classification task, as reported in [21,25,54] for emotion recognition. The DFN is the basic deep neural network to approximate complex functions for supervised classification [51]. In our implementation, we adopt the neural network design reported in [54], as showed to leverage the pre-computed differential entropy features for different domain adaptation methods on SEED and SEED-IV. Table 2 shows the specifications of each layer used by DFN, including the components employed by MFA-LR. Two hidden layers were used as feature extractor G_f, while the SLR layer was employed for feature labeling G_h. As mentioned above, BN layers were included to standardize features after each hidden layer. Then, ReLU is used as an activation function, followed by dropout [61] to avoid overfitting and Gaussian noise layers to promote robust features. At the output of the feature extractor, before the SLR layer, l_2-normalization is applied to obtain a representation in spherical space.

Implementation details. The Stochastic Gradient Descent (SGD) optimizer is used to train the proposed framework during 35 epochs with an initial learning rate set to 0.001 and 0.01 for SEED and SEED-IV, reducing their value by a factor of 0.1 at epoch 30 and 35. We did not observe a change in the classification results after reaching this number of epochs. Learning rates were selected from {0.1, 0.01, 0.001, 0.0001}. We set the mini-batch size to 50 for each domain dataset, as a greater size did not show an increase in terms of the accuracy performance. Momentum was set to 0.9, while the L2 regularization was employed, setting a weight factor of 0.005. For the trade-off hyper-parameters, we used $\alpha = 1.0$ and $\lambda_t = 0.1$, which were selected from {0.01, 0.1, 0.5, and 1.0}. For CMD in subject alignment loss, we used two statistical moments as a greater number did not improve the accuracy performance. The dynamic weights ω_{dis} and ω_{sem} were set to 1.0 initially.

For the RSDA training, we use the Adam optimizer during 10 epochs with a learning rate equal to 0.0001 and 0.001 for SEED and SEED-IV, respectively; no change in the classification results was observed after reaching this number of epochs. The batch size was set to 50. Concerning the trade-off hyper-parameters, we used $\lambda_{adv} = 0.1$, $\lambda_{sm} = 0.1$, and $\lambda_{ent} = 0.1$, selected from {0.01, 0.1, 0.5, and 1.0}. For Exponential Moving Average, λ_{ema} was set to 0.998, which is a default value suggested by [55].

We used PyTorch 1.9.1 with Python 3.6 for our implementation. All experiments were performed with a PC Intel(R) Core (TM) i7 with an NVIDIA GeForce GTX 1080 GPU and Ubuntu Linux v18.04 (LTS).

Evaluation. The proposed framework was evaluated in each of the three experimental sessions available in SEED and SEED-IV, as different domain shifts may be found for the three different record periods. We used Leave-One-Subject-Out Cross-Validation (LOSOCV), where one subject is employed as the target domain while data of the remaining subjects are used as source domains. For our experiments, all data from each source subject are employed for domain adaptation. 80% of the data were used for domain adaptation with the target subject, while the remaining 20% was used as test data. We performed five repetitions using different random initial weights and different partitions of target data to account for random effects. We report average accuracy and standard deviation for all subjects in each session.

4.2. Experimental results

Table 3 shows the average accuracy and standard deviation of MFA-LR on SEED and SEED-IV for each available session. As expected, the classification results on SEED-IV showed lower performance than those on SEED as more emotions are involved in the recognition task. Note that a different accuracy performance is obtained for each session as diverse domain shifts may be present in EEG data. For the SEED dataset, we observed that our proposal obtained the highest average accuracy (89.11%) with the lowest standard deviation (7.72) in Session 1. For SEED-IV, the highest average accuracy (74.99%) with the lowest standard deviation (12.10) was achieved in Session 2.

The classification results of MFA-LR by using F1-score and AUC are presented in Table 4. Regarding the model's precision and recall ability, we observed that MFA-LR obtained average F1-scores of 0.8520 and 0.6890 on SEED and SEED-IV across the three available sessions. With respect to the model's ability to predict classes correctly, our proposal obtained average AUC scores of 0.8892 and 0.7956 for each dataset. Similar to accuracy results, the best classification performances were obtained in Sessions 1 and 2 for SEED and SEED-IV, respectively.

Table 5 shows the experimental results of MFA-LR and existing studies on the SEED dataset; we also report the best result of MFA-LR from the five executions. Note that only a few works have been evaluated over the three experimental sessions, while most of the works reported their accuracy performance over the best session. We can see that MFA-LR reported an accuracy performance of up to 89.73% (best repetition), achieving an average accuracy of 89.11% with one of the lowest standard deviations (7.72) when several repetitions are

Table 2
Specifications of the Deep Feed-forward Network (DFN) used in our study.

Order	Layer	Hyper-parameters	Output dimension	
			SEED	SEED-IV
Feature extractor G_f				
1	Input	Number of input variables	2790 (9 × 62 × 5)	620 (2 × 62 × 5)
2	Fully connected		512	512
	Batch normalization	Momentum = 0.1, affine = False		
	ReLU	–		
	Dropout	$p = 0.5$		
	Gaussian noise layer	std. dev.$(v) = 1.0$		
3	Fully connected		320	512
	Batch normalization	Momentum = 0.1, affine = False		
	ReLU	–		
	Dropout	$p = 0.5$		
4	l_2-normalization	$r = 10.0$		
Feature labeling G_h				
5	SLR Layer	Number of classes	3	4

Table 3
Accuracy performance (%) and standard deviation of MFA-LR on SEED and SEED-IV datasets for each available session, using LOSOCV.

	SEED			SEED-IV		
	Session 1	Session 2	Session 3	Session 1	Session 2	Session 3
1	91.72 ± 02.04	79.89 ± 01.32	79.81 ± 01.10	67.14 ± 02.53	77.36 ± 00.98	79.27 ± 01.89
2	82.37 ± 02.20	85.58 ± 12.39	65.46 ± 02.10	68.07 ± 05.18	94.01 ± 02.35	81.82 ± 02.61
3	85.65 ± 02.19	98.95 ± 01.49	93.41 ± 00.95	73.22 ± 02.31	89.58 ± 06.69	64.97 ± 04.91
4	98.61 ± 00.51	80.04 ± 04.81	83.77 ± 00.99	54.27 ± 03.19	59.28 ± 05.39	80.73 ± 02.82
5	87.83 ± 01.09	93.52 ± 00.76	90.25 ± 01.69	58.71 ± 06.78	80.12 ± 04.39	87.03 ± 03.25
6	96.95 ± 01.76	80.30 ± 05.60	96.72 ± 01.18	26.78 ± 03.02	74.13 ± 16.11	59.88 ± 03.91
7	87.42 ± 01.33	79.17 ± 02.21	79.25 ± 01.24	66.32 ± 03.85	85.63 ± 03.69	80.97 ± 07.08
8	73.37 ± 01.53	74.05 ± 00.73	99.10 ± 00.56	60.82 ± 06.78	83.35 ± 03.04	76.61 ± 02.04
9	86.03 ± 03.04	99.66 ± 00.76	90.21 ± 01.60	79.53 ± 01.17	67.30 ± 03.04	29.45 ± 01.46
10	85.61 ± 01.89	80.75 ± 05.20	71.57 ± 01.16	71.35 ± 02.30	68.38 ± 08.65	71.76 ± 02.93
11	89.42 ± 01.43	79.24 ± 00.82	81.13 ± 01.52	72.98 ± 03.59	54.97 ± 04.45	43.76 ± 09.97
12	95.82 ± 03.14	81.51 ± 05.96	58.27 ± 02.30	50.88 ± 04.62	60.96 ± 02.37	52.61 ± 01.79
13	78.53 ± 01.42	48.25 ± 01.66	84.63 ± 00.34	86.78 ± 02.09	70.90 ± 02.49	68.24 ± 05.12
14	97.36 ± 00.66	76.35 ± 02.66	96.08 ± 00.36	72.40 ± 05.38	68.14 ± 03.75	75.39 ± 04.26
15	100.00 ± 00.00	93.75 ± 01.69	100.00 ± 00.00	65.15 ± 04.02	90.78 ± 02.63	79.27 ± 02.95
Avg.	89.11 ± 07.72	82.07 ± 12.38	84.64 ± 12.42	64.96 ± 14.04	74.99 ± 12.10	68.78 ± 16.16

Table 4
Average results of F1-score and AUC for MFA-LR on SEED and SEED-IV for each available session, using LOSOCV.

Metrics	Session 1	Session 2	Session 3	Average
	SEED			
F1-score	0.8921 ± 0.0773	0.8184 ± 0.1149	0.8455 ± 0.1262	0.8520
AUC	0.9188 ± 0.0772	0.8644 ± 0.0859	0.8845 ± 0.0943	0.8892
	SEED-IV			
F1-score	0.6434 ± 0.1428	0.7498 ± 0.1351	0.6739 ± 0.1568	0.6890
AUC	0.7652 ± 0.0948	0.8340 ± 0.0894	0.7875 ± 0.1058	0.7956

executed. Note that our proposal achieved the highest accuracy results when applied over each of the three sessions, reaching an average accuracy of 85.27% and outperforming previous studies by 4.81 percentage points (pp).

Table 6 presents the performance comparison of MFA-LR with existing studies on the SEED-IV dataset. We include the classification results of MFA-LR based on the experimental design reported in [18, 48], where EEG data from all sessions are utilized in LOSOCV. For this experimental design, we observed that RGNN reported the best accuracy performance (73.84%) with the lowest standard deviation (8.02). Note that MFA-LR obtained a similar performance (73.65%) for its best repetition, but the second-best average accuracy (72.72%) with the highest standard deviation (10.77) when several repetitions are executed. Despite this, our proposal achieved an accuracy performance of up to 77.05% (best repetition) over the second session, averaging 74.99% across the different repetitions. It is important to highlight that MFA-LR achieved the highest average accuracy (69.58%) over all available sessions, outperforming RGNN[1] and RODAN by 14.64 pp and 8.83 pp, respectively.

4.2.1. Confusion matrix

Fig. 4 shows the confusion matrices of MFA-LR for the best session in SEED and SEED-IV. These confusion matrices represent the accuracy for each emotion. For SEED, we observed that positive emotion is

[1] We adopted the implementation published in https://github.com/zhongpeixiang/RGNN, leveraged on https://www.kaggle.com/code/nrkm0303/sgconv/notebook.

Table 5
Accuracy performance (%) comparison with different studies that performed LOSOCV on SEED dataset.

Method	Best session	Session 1	Session 2	Session 3	Average
SAAE [4]	80.34 ± 06.13	80.34 ± 06.13	74.68 ± 12.76	78.73 ± 12.96	77.88
ASFM [62]	83.51 ± 07.40	83.51 ± 07.40	76.68 ± 12.16	81.20 ± 10.68	80.46
BiDANN [27]	83.28 ± 09.60	–	–	–	–
BiDANN-S [63]	84.14 ± 06.87	–	–	–	–
WGAN [22]	87.07 ± 07.14	–	–	–	–
JDA [25]	88.28 ± 11.44	–	–	–	–
R2G-STNN [64]	84.16 ± 07.63	–	–	–	–
IAG [29]	86.30 ± 06.91	–	–	–	–
BiHDM [30]	85.40 ± 07.53	–	–	–	–
RGNN [18]	85.30 ± 06.72	–	–	–	–
MACI [34]	77.43 ± 07.37	73.42 ± 06.86	70.81 ± 06.18	77.43 ± 07.37	73.23
RSDA [54]	84.70 ± 07.27	–	–	–	–
MS-MDA [35]	82.88 ± 07.83	82.88 ± 07.83	80.93 ± 10.16	76.88 ± 14.16	80.23
TANN [19]	84.41 ± 08.75	–	–	–	–
ADAAM-ER [20]	86.58 ± 08.41	–	–	–	–
MWACN [21]	89.30 ± 09.18	–	–	–	–
MFA-LR (Ours)					
Best repetition	89.73 ± 06.98	89.73 ± 06.98	83.16 ± 12.91	84.93 ± 11.58	85.94
Average case	89.11 ± 07.72	89.11 ± 07.72	82.07 ± 12.38	84.64 ± 12.42	85.27

– indicates that the experiment results are not reported.

Table 6
Accuracy performance (%) comparison with different studies that performed LOSOCV on SEED-IV dataset.

Method	All data	Session 1	Session 2	Session 3	Average
DAN [48]	58.87 ± 08.13	–	–	–	–
BiDANN [27]	65.59 ± 10.39	–	–	–	–
RODAN [28]	–	–	–	–	60.75
JDA [25]	–	–	70.83 ± 10.25	–	–
BiHDM [30]	69.03 ± 08.66	–	–	–	–
RSDA [54]	–	–	72.37 ± 11.12	–	–
RGNN [18]	73.84 ± 08.02	53.49 ± 10.57[a]	57.37 ± 08.81[a]	53.97 ± 16.93[a]	54.94[a]
TANN [19]	68.00 ± 08.35	–	–	–	–
JAGP [47]	–	54.37 ± 09.49[a]	51.06 ± 15.14[a]	60.29 ± 15.14[a]	55.24[a]
MWACN [21]	–	–	74.60 ± 10.77	–	–
MFA-LR (Ours)					
Best repetition	73.65 ± 10.65	66.35 ± 14.02	77.05 ± 13.45	70.31 ± 15.25	71.23
Average case	72.79 ± 10.77	64.96 ± 14.04	74.99 ± 12.10	68.78 ± 16.16	69.58

– indicates that the experiment results are not reported.
[a]indicates that the experiment results were obtained from implementation provided by authors, applied over the pre-computed differential entropy features.

more sensitive to be recognized than negative and neutral emotions. These results suggest that information from positive emotion is more easily transferred between subjects. For SEED-IV, the neutral emotion is the easiest to distinguish, compared to sadness, fear and happiness emotions, which obtained a similar accuracy performance. These results indicate that it is more difficult to share similar patterns between subjects when they experience non-neutral emotions.

4.2.2. Analysis per subject

Figs. 5 and 6 show the MFA-LR accuracy results for each subject and session in SEED and SEED-IV. With respect to the results on SEED, we observed that subjects 9 and 12 obtained high variations of their performance in Session 1, while this behavior is more clearly observed for subjects 2, 4, 6, 7 and 14 in Session 2. For results on SEED-IV, high variations were observed for subjects 3, 4, 6, and 10 in Session 2, as well as subjects 3, 7, and 13 in Session 3. These results indicate that different partitions of target data and initial neural network weights can negatively affect the performance of a domain adaptation method.

On the other hand, we noticed that subjects 12 and 13 in SEED obtained accuracy performances below 60% in sessions 3 and 2, respectively. Similarly for SEED-IV, accuracy performances below 50% were observed for subject 6 in Session 1, as well as subjects 9 and 10 in Session 3. These accuracy performances suggest that a high negative transfer is still present between some subjects.

4.2.3. Effect of domain adaptation on different sessions

We analyze the effect of domain adaptation on different sessions, as different results were reported. For this, an example case is studied to show the impact of knowledge transfer between subjects on two sessions. Fig. 7 shows the classification results on SEED with a one-to-one transfer paradigm for subject 1 in Sessions 1 and 2. The unlabeled samples from subject 1 are always used as the target domain, while labeled samples from each of the remaining subjects serve as the source domain. We observed higher classification results in 10 cases for Session 1, while 4 cases showed better score in Session 2. The average accuracy across all knowledge transfers was 70.31% for Session

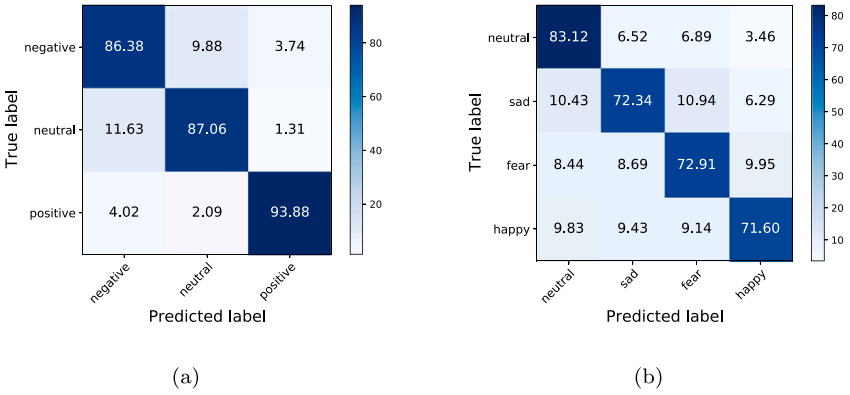


Fig. 4. Confusion matrices of MFA-LR for the best session in (a) SEED and (b) SEED-IV.

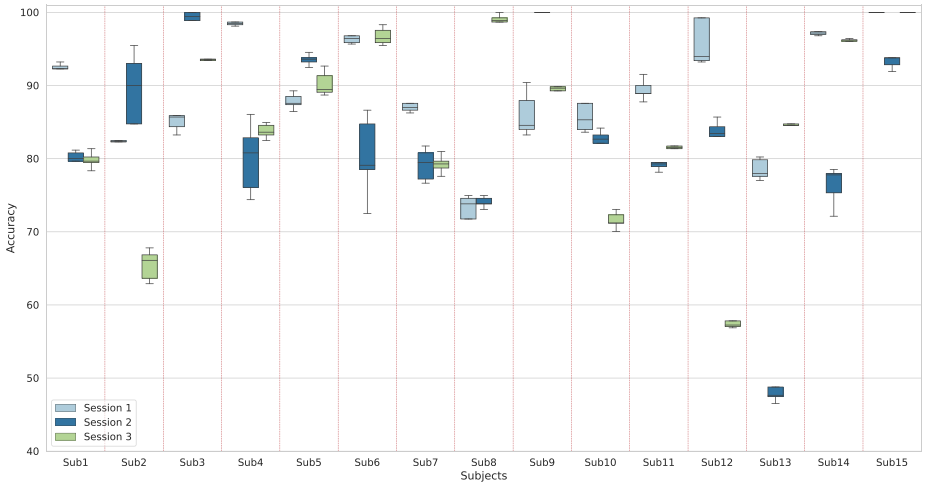


Fig. 5. MFA-LR accuracy results (%) for each subject and session in SEED. For each box, the central line is the median, while upper and lower bounds represent first and third quarterlies; black line indicates extreme values.

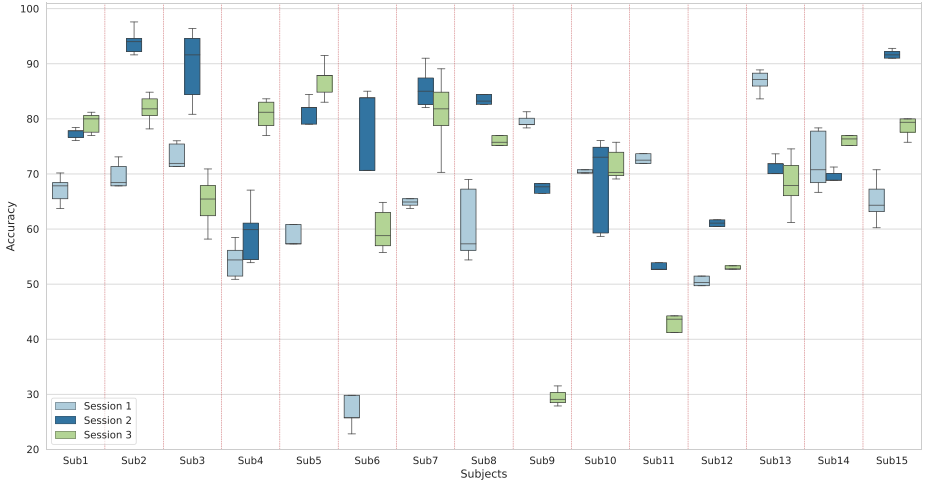


Fig. 6. MFA-LR accuracy results (%) for each subject and session in SEED-IV. For each box, the central line is the median, while upper and lower bounds represent first and third quarterlies; black line indicates extreme values.

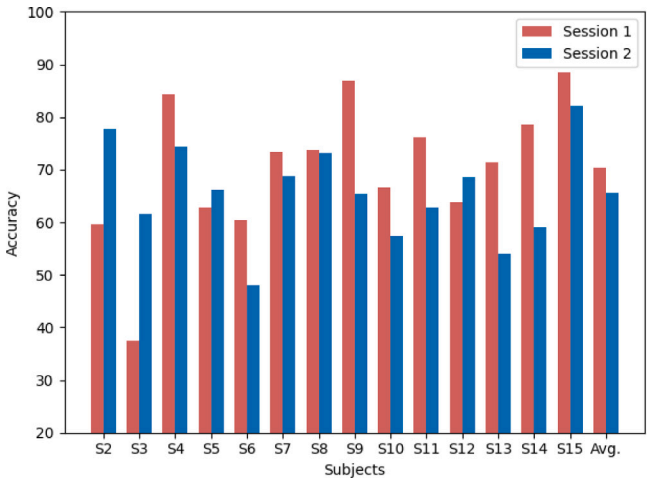


Fig. 7. Accuracy results (%) on SEED with a one-to-one transfer paradigm for subject 1 in Sessions 1 and 2. Average accuracy (Avg.) is also included.

1 and 65.67% for Session 2. These results suggest that source subjects in Session 1 share more similar patterns with subject 1, as opposed to Session 2.

From Table 3, multi-to-one transfer paradigm for subject 1 reported 91.72% in Session 1 and 79.89% for Session 2. By using a one-to-one transfer paradigm, the best result was 88.51% (subject 15) on Session 1, while the best accuracy was of 82.11% (subject 15) on Session 2. These results indicate that mixing source data from different subjects can enhance knowledge transfer, as noticed on Session 1. On the other hand, a negative transfer can be produced whenever dissimilar subject patterns are mixed during domain adaptation, as observed in Session 2.

4.2.4. Ablation study

We analyze the importance of each component in MFA-LR by conducting an ablation study. Table 7 presents the average accuracy across all subjects for the best Session in SEED (Session 1) and SEED-IV (Session 2) when each component is added/removed to MFA-LR.

Impact of performing a fine-grained alignment. We investigate the impact of using a multi-source feature alignment approach instead of a single-source domain adaptation. For this, an extension of MSTN [38] is used as baseline, incorporating CMD instead of adversarial training; cross-entropy and semantic losses are indicated as CE and SEM, respectively. Note that the baseline method is built on the deep neural network architecture presented in Section 4.1. From Table 7, we can see that the average accuracy improved 10.02 pp (78.95% vs. 88.29%) and 10.39 pp (64.79% vs. 74.64%) on SEED and SEED-IV, when a multi-source feature alignment is performed at the subject and class level. These results suggest that a fine-grained alignment is critical to perform a better knowledge transfer between subjects.

Effect of normalization strategies. We examine the impact of normalization strategies as their incorporation leverages the fine-grained alignment. The average accuracy is reduced by 3.03 pp (88.29% vs. 85.26%) and 1.67 pp (74.64% vs. 72.97%) on SEED and SEED-IV when batch normalization is removed. Likewise, a reduction of 10.43 pp (88.29% vs. 77.86%) and 13.54 pp (74.64% vs. 61.10%) is observed for each dataset, when feature normalization in spherical space is discarded. Finally, note that a reduction of 0.54 pp (88.29% vs. 87.75%) is registered in SEED by discarding the SLR layer, but an improvement of 0.34 pp (74.64% vs. 74.98%) is observed on SEED-IV.

Impact of a discriminative space on the target subject. We study the effect of introducing the conditional entropy (CondEnt) restriction to encourage a discriminative feature space on the target subject. The classification results showed that the average accuracy is slightly reduced by 0.26 pp (88.29% vs. 88.03%) and 0.14 (74.64% vs. 74.50%) on SEED and SEED-IV when CondEnt is added to MFA. Although these results suggest a negative effect, an analysis across all sessions showed that CondEnt may help improve 1.25 pp (80.41% vs. 81.66%) when applied on Session 2 for SEED. It indicates that specific scenarios require a discriminative space on the target domain to achieve a better knowledge transfer.

Impact of noise perturbation. By analyzing the impact of removing the Gaussian noise layer to deal with input perturbations, we observed a reduction of 1.77 pp (88.03% vs. 86.26%) and 1.42 pp (74.50% vs. 73.08%) on SEED and SEED-IV, respectively. Indeed, these results suggest that existing models should deal with the presence of random perturbations to provide a robust representation.

Effect of rectifying target pseudo-labels. We observed an improvement of 1.08 pp (88.03% vs. 89.11%) on SEED when the label rectification procedure was applied, indicating that target samples were wrongly pseudo-labeled and should be discarded during the domain adaptation process. However, we can also perceive that an improvement of 0.49 pp (74.50% vs. 74.99%) is only achieved over SEED-IV, indicating that MFA is able to assign robust pseudo-labels.

Effect of exponential moving average. When EMA is removed from MFA-LR, we observed a reduction of 0.83 pp on SEED, suggesting the need to stabilize model training. Moreover, MFA-LR showed a stable training on SEED-IV, as only a reduction of 0.14 pp was observed.

4.2.5. Hyperparameter sensitivity analysis

We investigate the accuracy performance in terms of variation of the regularization parameters α and λ_t. α corresponds to the domain distribution alignment, while λ_t is the weight factor for the separation on the target domain. Fig. 8 shows the average performance of MFA and MFA-LR with the two parameters, taking Session 1 on SEED as an example. Results show that α achieves stable results when set to 0.5 and 1.0. In contrast, MFA-LR reduces its performance by varying the parameter λ_t, being 0.1 the best choice.

5. Discussion

We evaluated and validated the proposed framework on two public benchmarks for emotion recognition from EEG signals. We showed that the MFA-LR framework can improve the emotion recognition performance on each available session using LOSOCV. Most previous works only reported classification results from the best session on SEED, but different domain shifts and negative transfers might be present in data from each session, affecting the recognition performance. Also note that EEG data from three sessions are usually mixed for each subject on SEED-IV, and then LOSOCV is applied [18,48]. This approach may be helpful as it was observed that MFA-LR obtained a low performance on Sessions 1 and 3 (64.96% and 68.78%), improving its accuracy performance (72.79%) when data from the three sessions are mixed. However, we can also see that MFA-LR obtained 74.99% on Session 2, suggesting that samples from other sessions may generate a negative transfer. Likewise, we found that previous studies, based on deep learning approaches, did not report the use of different data partitions and different weight initializations on SEED and SEED-IV. Note that this issue also affects the classification results as showed by the fact that the accuracy performance of MFA-LR varies when multiple repetitions are executed. Thus, an extensive evaluation is presented in order to demonstrate the robustness of the proposed framework.

With respect to the state-of-the-art approaches, most works perform a single-source domain adaptation by mixing samples from all known subjects in a single source dataset. For shallow approaches, authors in [47] proposed JAGP, which aims to align the joint distributions between subjects, leveraged on a target label estimation. Originally, JAGP reported classification results above 78% on SEED-IV for every

Table 7
Evaluation in terms of accuracy (%) for each component of MFA-LR on SEED and SEED-IV.

Method	SEED	SEED-IV
Single-source adaptation (CE+CMD+SEM)	78.27 ± 07.78	64.25 ± 08.24
Multi-source Feature Alignment (CE+M-CMD+M-SEM)	88.29 ± 08.24	74.64 ± 12.42
– Batch Normalization Layer	85.26 ± 07.91	72.97 ± 11.11
– Feature Normalization	77.86 ± 09.37	61.10 ± 11.83
– Spherical Logistic Regression Layer	87.75 ± 08.06	74.98 ± 12.84
Multi-source Feature Alignment (CE+M-CMD+M-SEM+CondEnt)	88.03 ± 08.21	74.50 ± 12.46
– Noise Perturbation Layer	86.26 ± 07.97	73.08 ± 12.50
Multi-source Feature Alignment+Label Rectification (MFA-LR)	**89.11** ± 07.97	**74.99** ± 12.10
– Exponential Moving Average (EMA)	88.28 ± 08.01	74.85 ± 12.71

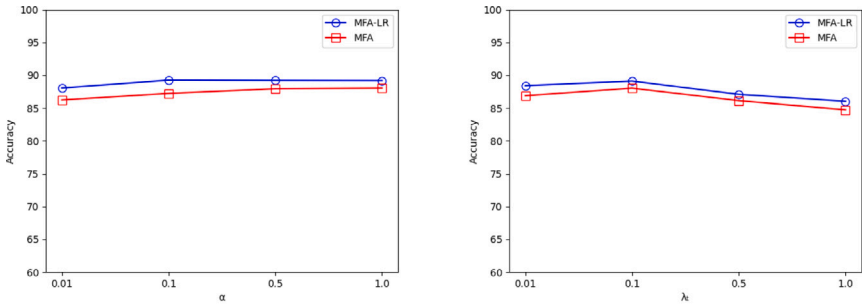


Fig. 8. Accuracy (%) of MFA and MFA-LR for Session 1 on SEED with parameters (a) α and (b) λ_i.

session. However, in our classification results we found that JAGP obtained a reduction of at least 18 percentage points across the three sessions, whenever it is applied on pre-computed features provided by the SEED-IV dataset. Moreover, with deep domain adaptation approaches, the state-of-the-art methods for SEED and SEED-IV also focus on reducing the joint distributions between subjects, such as JDA and MWACN. Nevertheless, as mentioned above, these methods reported classification results from the best session, while not using different partitions of data or various initializations of network weights. Regarding existing approaches, most reported works present a common limitation, the domain alignment process is performed in coarse-grained, but not all feature space is shared between subjects. In literature, some efforts have been proposed which leverage this assumption, such as TANN [19], where fine-grained brain regions are aligned between subjects. Likewise, MACI and MS-MDA [34–36], which perform a distribution alignment by subject, utilizing multi-source domain adaptation. Although the presented works have evidenced that a fine-grained alignment could help to improve the knowledge transfer between subjects, we observed that these works did not report the highest accuracy results on SEED and SEED-IV. Unlike previous approaches, we showed that a robust adaptation framework can be built by performing fine-grained alignment while an inter-class separation and robustness against input perturbations are encouraged over the entire feature space; followed by a rectification of target pseudo-labels. Experimental results showed that MFA-LR may obtain a high accuracy performance across the different sessions on SEED and SEED-IV.

Concerning the effect of each component in MFA-LR, we found that an improvement of about nine percentage points over SEED and SEED-IV is attainable by performing a fine-grained alignment at subject and class levels, while an inter-class separation is ensured in coarse grain. Note that this domain alignment process is leveraged on feature normalization strategies, as the accuracy performance is negatively affected if they are removed. Likewise, we found that a corruption process of the learned features is also needed as an improvement is observed by including a Gaussian noise layer. Although this strategy is a simple to deal with the random perturbations presented in EEG data, it showed to provide certain robustness to the classifier model. Also, note that the previous work in [54] showed that applying a pseudo-labeling correction method may aid the construction of a robust adaptive classifier. However, we found that the initial pseudo-labeled target samples, produced by MFA, are also important to achieve a high accuracy performance, as the original work showed worse results than our proposal.

Among the main limitations, the proposed framework did not consider an explicit representation of the spatial relationships between channels. This representation would allow us to identify important relationships, showing that brain regions may be strongly related to emotion processing. Likewise, a limited target dataset and an online procedure were not assumed in our experiments. It should be noted that this research was focused on providing an effective and robust solution to transfer useful knowledge from known subjects to new subjects, assuming that a fine-grained domain alignment can be achieved.

6. Conclusions

In this paper, we proposed a robust unified domain adaptation framework to address cross-subject EEG-based emotion recognition, which is motivated by the observation that not all of the feature space is shared between subjects. Our MFA-LR proposal performs a fine-grained alignment at the subject and class levels, while an inter-class separation and robustness against input perturbations are encouraged in coarse grain over the entire feature space. Then, a pseudo-labeling correction procedure is applied to rectify those mislabeled target samples produced by MFA. Our experiments showed that MFA-LR achieves state-of-the-art results on two public emotion recognition datasets, SEED and SEED-IV, over each of the three available sessions, using leave-one-subject-out cross-validation. An interesting finding is that by performing a fine-grained feature alignment and an inter-class coarse-grained separation, a better subject distribution alignment may be achieved, as opposed to traditional approaches. Another interesting finding is that input perturbations may produce a negative transfer in emotion recognition, so a mitigation strategy must be included during

domain adaptation. In future work, we will investigate more effective strategies to discard fine-grained elements in order to ensure a better distribution alignment between subjects, as they are still exposed to negative transfer.

CRediT authorship contribution statement

Magdiel Jiménez-Guarneros: Conceptualization, Methodology, Investigation, Software, Validation, Visualization, Resources, Writing – original draft, Writing – review & editing, Project administration, Funding acquisition. **Gibran Fuentes-Pineda:** Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] S.M. Alarcão, M.J. Fonseca, Emotions recognition using EEG signals: A survey, IEEE Trans. Affect. Comput. 10 (3) (2019) 374–393.

[2] V. Gupta, M.D. Chopda, R.B. Pachori, Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals, IEEE Sens. J. 19 (6) (2019) 2266–2274.

[3] D. Zhang, L. Yao, K. Chen, S. Wang, X. Chang, Y. Liu, Making sense of spatio-temporal preserving representations for EEG-based human intention recognition, IEEE Trans. Cybern. (2019) 1–12.

[4] X. Chai, Q. Wang, Y. Zhao, X. Liu, O. Bai, Y. Li, Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition, Comput. Biol. Med. 79 (2016) 205–214.

[5] F. Lotte, Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain x2013;computer interfaces, Proc. IEEE 103 (6) (2015) 871–890.

[6] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, F. Yger, A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update, J. Neural Eng. 15 (3) (2018) 031005.

[7] P. Saha, S. Fels, M. Abdul-Mageed, Deep learning the EEG manifold for phonological categorization from active thoughts, in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 2019, pp. 2762–2766.

[8] V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, M. Grosse-Wentrup, Transfer learning in brain-computer interfaces, IEEE Comput. Intell. Mag. 11 (1) (2016) 20–31, http://dx.doi.org/10.1109/MCI.2015.2501545.

[9] D. Wu, Online and offline domain adaptation for reducing BCI calibration effort, IEEE Trans. Hum.-Mach. Syst. 47 (4) (2017) 550–563.

[10] J.G. Moreno-Torres, T. Raeder, R. Alaiz-Rodríguez, N.V. Chawla, F. Herrera, A unifying view on dataset shift in classification, Pattern Recognit. 45 (1) (2012) 521–530.

[11] M. Wang, W. Deng, Deep visual domain adaptation: A survey, Neurocomputing 312 (2018) 135–153.

[12] S. Sun, J. Zhou, A review of adaptive feature extraction and classification methods for EEG-based brain-computer interfaces, in: 2014 International Joint Conference on Neural Networks, IJCNN, 2014, pp. 1746–1753.

[13] D. Bethge, P. Hallgarten, T. Grosse-Puppendahl, M. Kari, R. Mikut, A. Schmidt, O. Özdenizci, Domain-invariant representation learning from EEG with private encoders, in: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2022, pp. 1236–1240.

[14] H. Phan, O.Y. Chén, P. Koch, Z. Lu, I. McLoughlin, A. Mertins, M. De Vos, Towards more accurate automatic sleep staging via deep transfer learning, IEEE Trans. Biomed. Eng. 68 (6) (2020) 1787–1798.

[15] S. An, S. Kim, P. Chikontwe, S.H. Park, Few-shot relation learning with attention for EEG-based motor imagery classification, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2020, pp. 10933–10938.

[16] P.-Y. Jeng, C.-S. Wei, T.-P. Jung, L.-C. Wang, Low-dimensional subject representation-based transfer learning in EEG decoding, IEEE J. Biomed. Health Inf. 25 (6) (2021) 1915–1925, http://dx.doi.org/10.1109/JBHI.2020.3025865.

[17] W. Ko, E. Jeon, S. Jeong, J. Phyo, H.-I. Suk, A survey on deep learning-based short/zero-calibration approaches for EEG-based brain–computer interfaces, Front. Hum. Neurosci. 15 (2021) 643386.

[18] P. Zhong, D. Wang, C. Miao, EEG-based emotion recognition using regularized graph neural networks, IEEE Trans. Affect. Comput. (2020) 1–12, http://dx.doi.org/10.1109/TAFFC.2020.2994159.

[19] Y. Li, B. Fu, F. Li, G. Shi, W. Zheng, A novel transferability attention neural network model for EEG emotion recognition, Neurocomputing 447 (2021) 92–101, http://dx.doi.org/10.1016/j.neucom.2021.02.048.

[20] Y. Ye, X. Zhu, Y. Li, T. Pan, W. He, Cross-subject EEG-based emotion recognition using adversarial domain adaption with attention mechanism, in: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society, EMBC, 2021, pp. 1140–1144.

[21] L. Zhu, W. Ding, J. Zhu, P. Xu, Y. Liu, M. Yan, J. Zhang, Multisource wasserstein adaptation coding network for EEG emotion recognition, Biomed. Signal Process. Control 76 (2022) 103687.

[22] Y. Luo, S.-Y. Zhang, W.-L. Zheng, B.-L. Lu, WGAN domain adaptation for EEG-based emotion recognition, in: L. Cheng, A.C.S. Leung, S. Ozawa (Eds.), Neural Information Processing, Springer International Publishing, Cham, 2018, pp. 275–286.

[23] O. Özdenizci, Y. Wang, T. Koike-Akino, D. Erdoğmuš, Learning invariant representations from EEG via adversarial inference, IEEE Access 8 (2020) 27074–27085.

[24] H. Liu, H. Guo, W. Hu, EEG-based emotion classification using joint adaptation networks, in: 2021 IEEE International Symposium on Circuits and Systems, ISCAS, 2021, pp. 1–5, http://dx.doi.org/10.1109/ISCAS51556.2021.9401737.

[25] J. Li, S. Qiu, C. Du, Y. Wang, H. He, Domain adaptation for EEG emotion recognition based on latent representation similarity, IEEE Trans. Cogn. Dev. Syst. 12 (2) (2020) 344–353.

[26] K.-M. Ding, T. Kimura, K.-i. Fukui, M. Numao, EEG emotion enhancement using task-specific domain adversarial neural network, in: 2021 International Joint Conference on Neural Networks, IJCNN, 2021, pp. 1–8, http://dx.doi.org/10.1109/IJCNN52387.2021.9533310.

[27] Y. Li, W. Zheng, Z. Cui, T. Zhang, Y. Zong, A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition, in: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18, International Joint Conferences on Artificial Intelligence Organization, 2018, pp. 1561–1567.

[28] W.-C.L. Lew, D. Wang, K. Shylouskaya, Z. Zhang, J.-H. Lim, K.K. Ang, A.-H. Tan, EEG-based emotion recognition using spatial-temporal representation via bi-GRU, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society, EMBC, 2020, pp. 116–119, http://dx.doi.org/10.1109/EMBC44109.2020.9176682.

[29] T. Song, S. Liu, W. Zheng, Y. Zong, Z. Cui, Instance-adaptive graph for eeg emotion recognition, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, 2020, pp. 2701–2708.

[30] Y. Li, L. Wang, W. Zheng, Y. Zong, L. Qi, Z. Cui, T. Zhang, T. Song, A novel bi-hemispheric discrepancy model for EEG emotion recognition, IEEE Trans. Cogn. Dev. Syst. 13 (2) (2020) 354–367, http://dx.doi.org/10.1109/TCDS.2020.2999337.

[31] V.-A. Nguyen, T. Nguyen, T. Le, Q.H. Tran, D. Phung, STEM: An approach to multi-source domain adaptation with guarantees, in: 2021 IEEE/CVF International Conference on Computer Vision, ICCV, 2021, pp. 9332–9343, http://dx.doi.org/10.1109/ICCV48922.2021.00922.

[32] J. Li, S. Qiu, Y.-Y. Shen, C.-L. Liu, H. He, Multisource transfer learning for cross-subject EEG emotion recognition, IEEE Trans. Cybern. 50 (7) (2020) 3281–3293, http://dx.doi.org/10.1109/TCYB.2019.2904052.

[33] F. Wang, W. Zhang, Z. Xu, J. Ping, H. Chu, A deep multi-source adaptation transfer network for cross-subject electroencephalogram emotion recognition, Neural Comput. Appl. 33 (15) (2021) 9061–9073.

[34] J. Tao, Y. Dan, Multi-source co-adaptation for EEG-based emotion recognition by mining correlation information, Front. Neurosci. 15 (2021) 401.

[35] H. Chen, M. Jin, Z. Li, C. Fan, J. Li, H. He, MS-MDA: Multisource marginal distribution adaptation for cross-subject and cross-session EEG emotion recognition, Front. Neurosci. 15 (2021) http://dx.doi.org/10.3389/fnins.2021.778488.

[36] H. Chen, Z. Li, M. Jin, J. Li, MEERNet: Multi-source EEG-based emotion recognition network for generalization across subjects and sessions, in: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society, EMBC, 2021, pp. 6094–6097, http://dx.doi.org/10.1109/EMBC46164.2021.9630277.

[37] W. Zellinger, B.A. Moser, T. Grubinger, E. Lughofer, T. Natschlager, S. Saminger-Platz, Robust unsupervised domain adaptation for neural networks via moment alignment, Inform. Sci. 483 (2019) 174–191.

[38] S. Xie, Z. Zheng, L. Chen, C. Chen, Learning semantic representations for unsupervised domain adaptation, in: International Conference on Machine Learning, 2018, pp. 5419–5428.

[39] R. Shu, H. Bui, H. Narui, S. Ermon, A DIRT-T approach to unsupervised domain adaptation, in: International Conference on Learning Representations 2018, 2018.

[40] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Vol. 37, ICML '15, JMLR.org, 2015, pp. 448–456.
[41] K. Saito, D. Kim, S. Sclaroff, T. Darrell, K. Saenko, Semi-supervised domain adaptation via minimax entropy, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 8050–8058.
[42] X. Gu, J. Sun, Z. Xu, Spherical space domain adaptation with robust pseudo-label loss, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 9101–9110.
[43] X. Gu, J. Sun, Z. Xu, Unsupervised and semi-supervised robust spherical space domain adaptation, IEEE Trans. Pattern Anal. Mach. Intell. (2022) 1, http://dx.doi.org/10.1109/TPAMI.2022.3158637.
[44] W.-L. Zheng, B.-L. Lu, Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks, IEEE Trans. Autonomous Mental Dev. 7 (3) (2015) 162–175.
[45] Y. Wang, S. Qiu, X. Ma, H. He, A prototype-based SPD matrix network for domain adaptation EEG emotion recognition, Pattern Recognit. 110 (2021) 107626.
[46] Y. Peng, H. Liu, W. Kong, F. Nie, B.-L. Lu, A. Cichocki, Joint EEG feature transfer and semi-supervised cross-subject emotion recognition, IEEE Trans. Ind. Inform. (2022) 1–12.
[47] Y. Peng, W. Wang, W. Kong, F. Nie, B.-L. Lu, A. Cichocki, Joint feature adaptation and graph adaptive label propagation for cross-subject emotion recognition from EEG signals, IEEE Trans. Affect. Comput. 13 (4) (2022) 1941–1958.
[48] H. Li, Y.-M. Jin, W.-L. Zheng, B.-L. Lu, Cross-subject emotion recognition using deep adaptation networks, in: L. Cheng, A.C.S. Leung, S. Ozawa (Eds.), Neural Information Processing, Springer International Publishing, Cham, 2018, pp. 403–413.
[49] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, J. Mach. Learn. Res. 17 (1) (2016) 2096–2030.
[50] P. Haeusser, T. Frerix, A. Mordvintsev, D. Cremers, Associative domain adaptation, in: 2017 IEEE International Conference on Computer Vision, ICCV, 2017, pp. 2784–2792.
[51] I. Goodfellow, Y. Bengio, A. Courville, et al., Deep Learning Book, Vol. 521, (7553) MIT Press, 2016, p. 800.
[52] G. An, The effects of adding noise during backpropagation training on a generalization performance, Neural Comput. 8 (3) (1996) 643–674.
[53] S. Dutta, B. Tripp, G.W. Taylor, Convolutional neural networks regularized by correlated noise, in: 2018 15th Conference on Computer and Robot Vision, CRV, IEEE, 2018, pp. 375–382.
[54] M. Jiménez-Guarneros, P. Gómez-Gil, A study of the effects of negative transfer on deep unsupervised domain adaptation methods, Expert Syst. Appl. 167 (2021) 114088.
[55] A. Tarvainen, H. Valpola, Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results, in: Advances in Neural Information Processing Systems, Vol. 30, 2017, pp. 1195–1204.
[56] W. Zheng, W. Liu, Y. Lu, B. Lu, A. Cichocki, EmotionMeter: A multimodal framework for recognizing human emotions, IEEE Trans. Cybern. (2018) 1–13.
[57] R. Duan, J. Zhu, B. Lu, Differential entropy feature for EEG-based emotion classification, in: 2013 6th International IEEE/EMBS Conference on Neural Engineering, NER, 2013, pp. 81–84.
[58] X.-W. Wang, D. Nie, B.-L. Lu, EEG-based emotion recognition using frequency domain features and support vector machines, in: International Conference on Neural Information Processing, Springer, 2011, pp. 734–743.
[59] L. Yang, J. Liu, EEG-based emotion recognition using temporal convolutional network, in: 2019 IEEE 8th Data Driven Control and Learning Systems Conference, DDCLS, 2019, pp. 437–442.
[60] P. Bashivan, I. Rish, M. Yeasin, N. Codella, Learning representations from EEG with deep recurrent-convolutional neural networks, in: International Conference on Learning Representations (2016), 2016.
[61] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (56) (2014) 1929–1958.
[62] X. Chai, Q. Wang, Y. Zhao, Y. Li, D. Liu, X. Liu, O. Bai, A fast, efficient domain adaptation technique for cross-domain electroencephalography(EEG)-based emotion recognition, Sensors 17 (5) (2017).
[63] Y. Li, W. Zheng, Y. Zong, Z. Cui, T. Zhang, X. Zhou, A bi-hemisphere domain adversarial neural network model for EEG emotion recognition, IEEE Trans. Affect. Comput. 12 (2) (2018) 494–504.
[64] Y. Li, W. Zheng, L. Wang, Y. Zong, Z. Cui, From regional to global brain: A novel hierarchical spatial-temporal neural network model for EEG emotion recognition, IEEE Trans. Affect. Comput. (2019) 1.