2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

MEERNet: Multi-source EEG-based Emotion Recognition Network for Generalization Across Subjects and Sessions

Hao Chen[1,2], Zhunan Li[1,2], Ming Jin[1,2] and Jinpeng Li[1,2]

Abstract— **As an important element in the human-machine interaction, the electroencephalogram (EEG)-based emotion recognition has achieved significant progress. However, one obstacle to practicality lies in the variability between subjects and sessions. Although several studies have adopted domain adaptation (DA) approaches to tackle this problem, most of them treat multiple data from different subjects and different sessions together as a single source for transfer. Since different EEG data have different marginal distributions, these approaches fail to satisfy the assumption of DA that the source has a certain marginal distribution. We therefore propose the multi-source EEG-based emotion recognition network (MEERNet), which takes both domain-invariant and domain-specific features into consideration. Firstly we assume that different EEG data share the same low-level features, and then we construct multiple branches corresponding to multiple sources to extract domain-specific features, and then DA is conducted between the target and each source. Finally, the inference is made by multiple branches. We evaluate our method on SEED and SEED-IV for recognizing three and four emotions, respectively. Experimental results show that the MEERNet outperforms the single-source methods in cross-session and cross-subject transfer scenarios with an accuracy of 86.7% and 67.1% on average, respectively.**

I. INTRODUCTION

Emotion as physiological information, which differs from logical intelligence, is now widely used in various aspects of daily human communications, such as negotiation. In human-computer interaction (HCI), emotion plays a crucial role in many studies, especially in health research, where researchers have long found a substantial association between various mental diseases and emotions [1]. Thus, most research focuses on identifying and analyzing the neural signals from the brain in specific ways. Brain-computer interfaces (BCIs) act as a bridge between the brain and the computer, and allow the users to access the brain signals directly from the computer. Among them, invasive BCIs are expensive and require surgery to obtain higher accuracy, but are too costly and dangerous. Non-invasive BCIs such as electroencephalogram (EEG), on the other hand, are safer and with moderate accuracy, and are therefore widely used for brain signal acquisition [2]. Besides, many studies have demonstrated that EEG improves emotion recognition, motor imagery, and event-related potentials [3], [4].

Also, it is challenging to get a model that is common to different subjects and sessions in EEG-based emotion recognition scenarios (*i.e.* the data collected from the same subject at the same time can be very biased). To tackle this, transfer learning is widely used in research works in this scenario. Transfer learning maximizes the information learned from the source domain EEG data and then applies it to new EEG data, and can significantly reduce the number of labels required in the target domain [5]. In recent years, there has been many research works for EEG-based emotion recognition, Zheng *et al.* [13] applies TCA (Transfer Component Analysis) [6] to the cross-subject transfer scenario. To tackle the marginal and conditional distribution problem. Li *et al.* proposes a Multi-source Style Transfer Mapping (MS-STM) [12] framework for cross-subject multi-source scenario. They take a few labeled training data to learn multiple STMs, which are then being used to map the target distribution to the distributions of the sources. Many works also apply deep learning techniques to improve the emotion recognition. Li *et al.* [14] adopts DAN (Deep Adaptation Network) [8] for cross-subject emotion recognition. Zheng *et al.* [15] extends the SEED dataset to the SEED-IV, which involves four emotions (fear, sad, happy, and neutral). They also propose EmotionMeter, which is a combination of two modalities of eye movements and EEG waves. With the attention concept, Fahimi *et al.* [16] develops an end-to-end deep CNN for cross-subject attention classification with EEG time-series data. To increase the online transfer performance, Li *et al.* [20] propose FOIT (fast online instance transfer) to avoid cost of iterative methods, thus improves the practicality of the algorithm.

However, with the enhancement of the accuracy of EEG-based emotion recognition, they do not take into account that different source data have different marginal distributions but simply concatenating them. This action will destroy the independence between the data, which will also negatively affect the model.

In this paper, inspired by Zhu *et al.* [11], we present a multi-source EEG-based emotion recognition network (MEERNet) for EEG-based emotion recognition in the case of multiple source domains, which takes into account the simultaneous transfer of multiple source domains and avoids simply concatenating them, which may disrupt the marginal distributions among EEG data. We also conducted experiments across subjects and sessions on two datasets, and the experimental results show that our proposed MEERNet

This work was supported in part by Zhejiang Provincial Natural Science Foundation of China (LQ20F030013), Research Foundation of HwaMei Hospital, University of Chinese Academy of Sciences, China (2020HMZD22), Ningbo Public Service Technology Foundation, China (202002N3181), and Medical Scientific Research Foundation of Zhejiang Province, China (2021431314).

[1] HwaMei Hospital, University of Chinese Academy of Sciences, No. 41 Northwest Street, Haishu District, Ningbo, Zhejiang, 315010, China.
[2] Ningbo Institute of Life and Health Industry, University of Chinese Academy of Sciences, Ningbo, Zhejiang, China.
Corresponding author: Jinpeng Li (*E-mail: lijinpeng@ucas.ac.cn*)

outperforms other works on two datasets.

II. METHOD

We list the symbols and their definition in Table I for simplicity of demonstration.

TABLE I
NOTATION TABLE

symbol	definition
X	Instance set (matrix)
Y	Label set (matrix)
S	Source domain
T	Target domain
N	number of source domains
Q	Common feature
R	Domain-specific feature
$\hat{Y}$	Predicted label
ϕ, Φ	Mapping function
$\mathcal{H}$	Reproducing kernel Hilbert space
CFE	Common feature extractor
$DSFE$	Domain-specific feature extractor
DSC	Domain-specific classifier
x	Feature vector
y	Label vector
q	Feature vector after CFE
r	Feature vector after DSFE
$\hat{y}$	Predicted label vector

Given a set of pre-existing EEG data and newly collected EEG data, our goal is to learn a model ϕ that is trained on these multiple source domain data using transfer learning, and thus has a better prediction on the newly collected data than simply combining the existed data into one source domain. The architecture of the proposed method is illustrated in Fig. 1.

As shown in the figure, the input to the MEERNet are N source domain data $\{(\mathbf{X}_i^S, \mathbf{Y}_i^S)\}_{i=1}^N$ and a target domain data $\{\mathbf{X}^T\}$, and then these data are fed into a common feature extractor module to get the domain-invariance features $\{\mathbf{Q}_i^S\}_{i=1}^N$ and $\{\mathbf{Q}^T\}$. Then for each domain-specific feature extractor, extracted common features $\{\mathbf{Q}_i^S\}_{i=1}^N$ will be fed into the network with $\{\mathbf{Q}^T\}$ and get their domain-specific features: $\{\mathbf{R}_i^S\}_{i=1}^N$ and $\{\mathbf{R}_i^T\}_{i=1}^N$, and on top of that, the MMD value is calculated, which is a measure of the distance of the current source and target domain. Next, the target domain features $\{\mathbf{R}_i^T\}_{i=1}^N$ and all the source domain features $\{\mathbf{R}_i^S\}_{i=1}^N$ extracted from the last step will get to the domain-specific classifiers to get the corresponding classification predictions: $\{\hat{\mathbf{Y}}^T\}$ and $\{\hat{\mathbf{Y}}_i^S\}_{i=1}^N$, then the results of the source domain are taken to calculate the classification loss. In the end, the average of these target-domain predictions is taken as the output of the model. Details of these modules are given below.

Common Feature Extractor in the MEERNet is used to map the source and target domain data from the original feature spaces to a common sharing latent space, and then common representations of all domains are extracted. This module can help to extract some low-level domain-invariant features.

Domain-specific Feature Extractor follows the Common Feature Extractor (CFE). After obtaining the features of all domains, we set up N single fully connected layers to correspond to N source domains. For each pair of source and target domains, we map the data to a unique latent space via the corresponding Domain-specific Feature Extractor (DSFE), respectively, and then obtain the corresponding domain-specific features. To apply transfer learning and bring the two domains close in the latent space, we choose the MMD [7] to estimate the distance between these two domains. MMD is widely used in the transfer learning and can be formulated in (1). In the process of training, MMD loss is decreased to narrow the source domain and the target domain in the feature space, which helps make better predictions for the target domain. This module aims to learn multiple domain-specific features.

$$MMD(X^S, X^T) = \left\| \frac{1}{N^S} \sum_{i=1}^{N^S} \Phi\left(\mathbf{x}_i^S\right) - \frac{1}{N^T} \sum_{j=1}^{N^T} \Phi\left(\mathbf{x}_j^T\right) \right\|_{\mathcal{H}}^2 \quad (1)$$

Algorithm 1 Overview of MEERNet

Input:
Iteration $\mathcal{T}$, source domain data $\{(\mathbf{X}_i^S, \mathbf{Y}_i^S)\}_{i=1}^N$ and target domain data $\{\mathbf{X}^T\}$

for t = 1,..., $\mathcal{T}$ **do**
- Take m samples $\{x_j^{Si}, y_j^{Si}\}_{j=1}^m$from source domains and $\{x_j^T\}_{j=1}^m$ from target domain.
- $\{q_j^{Si}\}_{j=1}^m, \{q^T\} \leftarrow CFE(\{x_j^{Si}, y_j^{Si}\}_{j=1}^m, \{x_j^T\}_{j=1}^m)$
- $\{r_j^{Si}\}_{j=1}^m, \{r_j^T\}_{j=1}^m \leftarrow DSFE(\{q_j^{Si}\}_{j=1}^m, \{q^T\})$
- $\mathcal{L}_{mmd} \leftarrow (1) \leftarrow DSFE$
- $\{\hat{y}_j^{Si}\}_{j=1}^m, \{\hat{y}_j^T\}_{j=1}^m, \leftarrow DSC(\{r_j^{Si}\}_{j=1}^m, \{r_j^T\}_{j=1}^m)$
- $\mathcal{L}_{cls} \leftarrow (2) \leftarrow DSC$
- Update model by minimizing the total loss

end for
return $\{\hat{Y}^T\}$;
Output:
Prediction of target domain data, $\{\hat{Y}^T\}$;

Domain-specific Classifier uses the features extracted from the DSFE to predict. In Domain-specific Classifier (DSC), there are N single softmax classifiers that correspond to each source domain. The final output of the DSC is the average of N classifiers. For each classifier training, we choose cross-entropy to estimate the classification loss using cross-entropy, as shown in (2). The average of the predictions of the N classifiers is taken as the final result.

$$\mathcal{L}_{cls} = \sum_{i=1}^{N} \mathbf{E}_{x \sim X_S} J\left(\hat{\mathbf{Y}}_i^S, \mathbf{Y}_i^S\right) \quad (2)$$

In summary, MEERNet accepts N source domain EEG data and one target domain EEG data, and then includes a common feature extractor to get N source domain features and one target domain feature. Next, N domain-specific

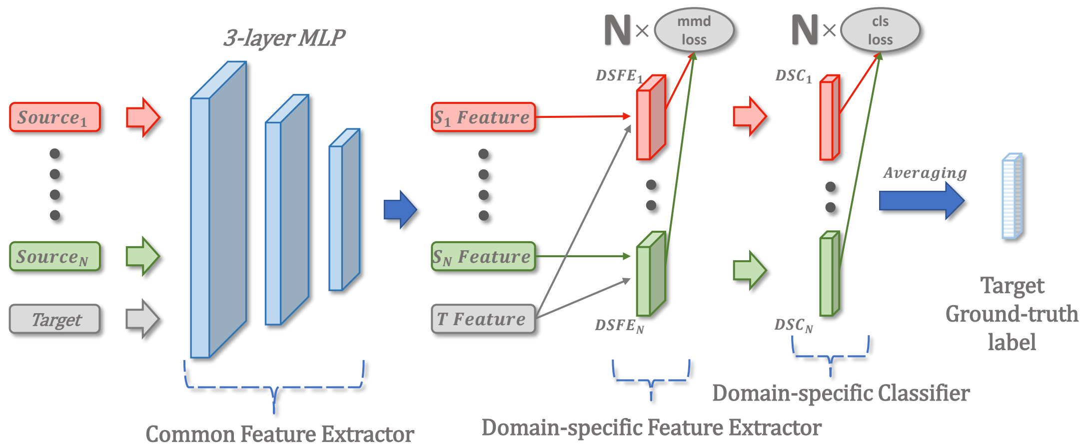


Fig. 1. The **architecture** of our proposed method. Our network consists of common feature extractor, domain-specific feature extractor, and domain-specific classifier. The model receives multiple source domains and leverages their knowledge to transfer to the target domain.

feature extractors are used to pairwise compute the MMD loss of individual source with the target domain and extract their domain-specific features. Finally, a domain-specific classifier is used to do the classification task, which also calculates the classification loss of the N classifiers using the features after DSFE of the N source domain EEG data.

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda \mathcal{L}_{mmd} \tag{3}$$

The training is based on the (3) and following the algorithm as shown in Algorithm. 1. For the three losses, minimizing MMD loss can get domain-invariant features for each pair of the source, and target domains; minimizing classification loss will bring more accurate classifiers for predicting the source domain data.

III. EXPERIMENTS

We perform experiments in the task of classification in emotion recognition on two datasets, and also experiment with different model settings. The Institution's Ethical Review Board approved all experimental procedures involving human subjects.

A. Settings

Datasets. We evaluate our method on two EEG-based emotion recognition datasets: **SEED** [17], [18], **SEED-IV** [15]. The raw data are gathered with an ESI NeuroScan system with 62-channel, and are sampled to 200 Hz, then a band-pass filter between 1 to 75 Hz is used for processing data. For both datasets, since they are set up with three sessions, each includes 15 subjects, we partitioned the datasets according to two different transfer scenarios (i.e., cross-subject and cross-session transfer). Specifically, for the cross-session transfer, we take the first 14 subjects in each session as the source domain and take the remaining one as the target; for the cross-session transfer, we take the first and second sessions of each subject as the source domain and transfer them to the third session. We do not select the raw data here, but rather the extracted Differential Entropy (DE) features as the data to be used with its ability to distinguish patterns from different bands [19]. For both data sets individually, they are in the form of channel x trial x band, we merge the channel and band here as a single sample, which ended up in the form of trial x sample (310 dimensions), and then concatenate all the trials together (15 trials for SEED, 24 trials for SEED-IV). In addition, we evaluate various normalization methods (i.e., normalization of channel dimension, normalization of sample dimension, and global normalization) and finally choose the channel dimension for better results .

Implementation Details. As mentioned in the Section. III, there are many details in the three modules of the MEER-Net. First, for the Common Feature Extractor (CFE), we choose 3-layer MLP for simplicity which reduces dimensions from 310-D to 64-D. Every linear layer is followed by a 1-D BatchNorm layer and a LeakyReLU layer. There is a single linear layer in both domain-specific feature extractor (DSFE) and domain-specific classifier (DSC), which reduces 64-D to 32-D and 32-D to the corresponding number of categories, respectively. In DSFE, a 1-D BatchNorm and a LeakyReLU are followed after the linear layer, while in DSC, there is only a linear layer. The whole network is trained using the Adam optimizer with an initial learning rate of 0.01, and train for 10k iterations, we also test 20k, 15k and 5k, the epoch of 10k has the best performance. The batch size is 32, which means we take 32 samples for each domain in every epoch. For the loss, we choose MMD as the metric of the distance between the source and target domain in the feature space.

B. Results

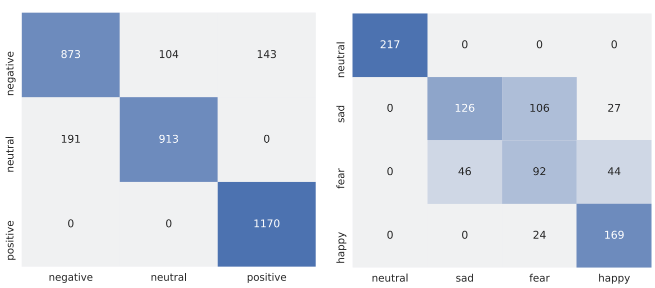
Table II shows the results of MEERNet and comparison methods [9], [8], [10](we customize these methods with the same settings as MEERNet has) on the SEED and SEED-IV, all the hyper-parameters of selected methods are the same, the standard deviations and average accuracy are calculated by averaging over all 15 participants in a leave-one-out cross-validation. The results indicate that our method largely outperforms the comparison methods. In the

TABLE II

COMPARISON RESULTS ON SEED AND SEED-IV

Dataset	Method	Cross-session	Cross-subject	Average
SEED	DDC	85.0 ±10.5	80.5 ±3.2	82.8 ±6.9
	DAN	81.8 ±9.9	83.5 ±5.9	82.7 ±7.9
	DCORAL	84.0 ±10.8	85.4 ±5.5	84.7 ±8.2
	w/o MMD	81.5 ±14.3	67.7 ±17.7	74.6 ±16.0
	MEERNet	**86.2** ±5.8	**87.1** ±2.0	**86.7** ±3.9
SEED-IV	DDC	68.8 ±16.6	54.3 ±4.2	61.6 ±10.4
	DAN	60.2 ±10.2	69.8 ±4.2	65.0 ±7.2
	DCORAL	65.1 ±13.2	67.1 ±2.5	66.1 ±7.9
	w/o MMD	65.6 ±18.5	62.1 ±13.2	58.9 ±15.9
	MEERNet	**72.1** ±14.1	**71.0** ±12.1	**67.1** ±16.4

SEED dataset, our method outperforms other methods in all transfer scenarios with at least ~2% improvement, while in the SEED-IV dataset, although our method still outperforms the other comparisons from the average perspective with at least ~1% improvement, there is a significant drop and high variance in the cross-subject transfer. To understand the effect of the modules in MEERNet, we remove MMD loss and evaluate the performance of the ablated model, the results illustrate that without MMD loss, there is a substantial drop in the performance of the model. Fig. 2 shows the confusion matrices of predictions made by our proposed method on SEED and SEED-IV. From (a) we can see that our method can classify positive emotion well, and (b) shows that our approach is more difficult to classify the two emotions of fear and sad because they are relatively similar.



(a) (b)

Fig. 2. The confusion matrices of predictions on SEED and SEED-IV. Only the results of cross-subject scenario are plotted. Each value in one square stands for the number of samples. (a)SEED, (b)SEED-IV.

IV. CONCLUSIONS

We have presented MEERNet that brings domain-invariant features and domain-specific features into consideration in the multiple source transfer learning for the EEG-based emotion recognition in aBCIs. In detail, different from others, we set up multiple branches in the model to correspond to multiple source domains, and use MMD loss to draw the distance between the source and target domains closer. The experimental results in two emotion analysis datasets have shown that our method outperforms other methods in accuracy. We believe that our proposed approach has help for the application of aBCIs and opens a more effective way of performing transfer learning for EEG-based emotion analysis in the multiple sources scenarios.

REFERENCES

[1] M. Valstar, J. Gratch, B. Schuller, F. Ringeval, et al., “Avec 2016: Depression, mood, and emotion recognition workshop and challenge," in *Proceedings of the 6th international workshop on audio/visual emotion challenge*, 2016, pp. 3-10.

[2] F. Cincotti *et al.*, “Non-invasive brain-computer interface system: towards its application as assistive technology," *Brain research bulletin*, vol. 75, no. 6, pp. 796–803, 2008.

[3] W. Wu, B. L. Lu, “Examing Four Experimental Paradigms for EEG-Based Sleep Quality Evaluation with Domain Adaptation," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020, pp. 5913–5916.

[4] J. J. Guo, R. Zhou, L. M. Zhao, B. L. Lu, “Multimodal emotion recognition from eye image, eye movement and eeg using deep neural networks," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 3071–3074.

[5] S. J. Pan, Q. Yang, “A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp.1345–1359, 2009.

[6] S. J. Pan, I. W. Tsang, J. T. Kwok, Q. Yang, "Domain adaptation via transfer component analysis", *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199-210, 2010.

[7] K. M. Borgwardt, A. Gretton, M. J. Rasch, H. P. Kriegel, B. Schölkopf, A. J. Smola, "Integrating structured biological data by kernel maximum mean discrepancy", *Bioinformatics*, vol. 22, no. 14, pp. e49-e57, 2006.

[8] M. Long, Y. Cao, J. Wang, M. I. Jordan, "Learning transferable features with deep adaptation networks", in *International conference on machine learning*, 2015, pp. 97-105.

[9] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, T. Darrell, "Deep domain confusion: Maximizing for domain invariance", *arXiv preprint*, arXiv:1412.3474, 2014.

[10] B. Sun, K. Saenko, "Deep coral: Correlation alignment for deep domain adaptation", in *European conference on computer vision*, 2016, pp. 443-450.

[11] Y. Zhu, F. Zhuang, D. Wang, "Aligning domain-specific distribution and classifier for cross-domain classification from multiple sources", in *Proceeding of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, pp. 5989-5996.

[12] J. Li, S. Qiu, Y. Y. Shen, C. L. Liu, H. He, "Multisource transfer learning for cross-subject EEG emotion recognition", *IEEE transactions on cybernetics*, 2019.

[13] W. L. Zheng, B. L. Lu, "Personalizing EEG-based affective models with transfer learning", in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 2016, pp. 2732-2738.

[14] H. Li, Y. M. Jin, W. L. Zheng, B. L. Lu, "Cross-subject emotion recognition using deep adaptation networks", in *International Conference on Neural Information Processing*, 2018, pp. 403-413.

[15] W. L. Zheng, W. Liu, Y. Lu, B. L. Lu, A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions", *IEEE transactions on cybernetics*, vol. 49, no. 3, pp. 1110-1122, 2018.

[16] F. Fahimi, Z. Zhang, W. B. Goh, T. S. Lee, K. K. Ang, C. Guan, " Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI", *Journal of neural engineering*, vol. 16, no. 2, 026007, 2019.

[17] R. N. Duan, J. Y. Zhu, B. L. Lu, "Differential entropy feature for EEG-based emotion classification", in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2013, pp. 81-84.

[18] W. L. Zheng, B. L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks", *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162-175, 2015.

[19] M. Soleymani, S. Asghari-Esfeden, Y. Fu, and M. Pantic, "Analysis of eeg signals and facial expressions for continuous emotion detection", *IEEE Transactions on Affective Computing*, vol. 7, no. 1, pp. 17–28, 2015.

[20] J. Li, H. Chen, and T. Cai, "FOIT: Fast Online Instance Transfer for Improved EEG Emotion Recognition", in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 2618-2625.