ShareLink: Neuro-Inspired EEG-based Cross-Subject Emotion Recognition via Shared Bi-hemisphere

Zixuan Huang[1,2*], Lingyao Kong[3*], Licheng Ao[3*], Shiyi Yao[4], An Xiang[1,2], and Fen Miao[5]✉

[1] Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences
[2] University of the Chinese Academy of Sciences, Beijing, China
[3] Jinan University – University of Birmingham Joint Institute at Jinan university, Jinan University, Guangzhou, China
[4] School of Cyber Science and Engineering, Wuhan Univerisity, Wuhan, China
[5] Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China
fenmiao@uestc.edu.cn

Abstract. Emotion recognition plays a pivotal role in human-computer interaction by enabling machines to perceive and adapt to human affective states. While neuroimaging studies[15,20] reveal significant functional lateralization between the left and right cerebral hemispheres during emotional processing, existing EEG-based emotion recognition methods face two critical challenges: (1) difficulty in aligning cross-hemispheric semantic features, and (2) limited generalizability across subjects and scenarios. To address these issues, we propose ShareLink, a novel EEG-based framework with Shared Cross-Hemispheric Structures. Our approach introduces three key innovative modules: (1) the Dynamic Shared Hemispheric Structure (DSHS) enforces non-Euclidean hemispheric structure constraints by sharing learnable adjacency matrix parameters across the bi-hemispheres, thereby effectively aligning semantic representations and extracting more discriminative hemispheric asymmetry features; (2) the Cross-Hemisphere Attention (CHA) shares similarity matrix between the hemispheres to establish dynamic inter-hemispheric links, enhancing the model's ability to capture interaction information while reducing parameters and mitigating overfitting risks; (3) the Shared Hemispheres Mixture-of-Experts (SHMoE) leverages multiple expert modules to abstract representations into a finite set of characteristics and employs a shared expert set to map bi-hemispheres features into a unified space, ensuring consistent and generalizable left-right hemisphere representations. Evaluated on SEED and SEED-IV datasets under cross-subject paradigms, ShareLink achieves accuracies of 80.61%±6.16% and 63.33%±8.29%, demonstrating superior cross-domain generalization. This work provides new insights into neurophysiologically inspired computational models for emotion recognition. The codes are available at: https://github.com/Huangzx1023/ShareLink/.

[*] Co-first authors

Keywords: EEG · Emotion recognition · Bi-hemispheric asymmetry · Domain generalization.

1 Introduction

Emotions represent complex psychophysiological responses exhibited by humans and animals in response to specific stimuli, serving as fundamental components of adaptive behavior and social interaction. In the context of human-computer interaction (HCI), emotion recognition has emerged as a critical capability, enabling systems to perceive, interpret, and respond to human affective states. Despite humans' innate proficiency in decoding emotional cues, machines still face significant challenges in achieving comparable levels of emotional understanding. [18,2,26]

Recent studies have incorporated neuroscience insights into EEG emotion recognition, improving model performance through neurophysiological mechanisms. Neuroscientific studies reveal that although the human brain is symmetrical, there are differences in how the left and right hemispheres respond to the same emotions. For instance, Herrington et al. [10] investigated the asymmetry of emotional expression, while Costanzo et al. [4] discussed the lateralization of emotions. Furthermore, some EEG-based emotion recognition methods have already leveraged this asymmetry for emotion classification [29,7,5,11,17]. These studies suggest that integrating lateralization characteristics of the brain's hemispheres into machine learning is a promising approach. However, how to more effectively utilize these inter-hemispheric differential features to enhance EEG emotion recognition performance remains an interesting and meaningful research direction. Current methods face two critical challenges when processing left-right hemispheric representations: (1) the alignment of cross-hemisphere semantic features and (2) the cross-subject generalization capabilities of models.

When the representations of the left and right hemispheres are misaligned, the investigation of hemispheric asymmetry is doubtful. Thus, it is essential to align the semantic representations of the hemispheres into a shared feature space. Inspired by differential asymmetry and Ding et al. [7], we propose a global shared-weight modeling approach that provides an efficient and implicit alignment mechanism for hemispheric. **For the model**, shared weights ensure that the feature extraction process is governed by the same set of parameters, thereby reducing potential model biases introduced by extra independent modules. **For EEG signals**, shared weights further guarantee that observed differences reflect real neural variations rather than inconsistencies in feature extraction. **For computation**, the implicit weight-sharing mechanism eliminates the need for additional alignment-specific parameters, significantly enhancing the model's computational efficiency. Specifically, we demonstrate the shareability at both the spatial structure and time-frequency representation levels. First, at the spatial structure level, we propose the Dynamic Shared Hemispheric Structure (DSHS), which imposes symmetric constraints on the hemispheres by sharing learnable adjacency matrix parameters between the left and right hemispheres. This effectively aligns semantic representations while uncovering more discriminative hemispheric asymmetric features. Second, at the time-frequency representation level, we introduce a novel attention mechanism, CHA. This method aims to simulate the collaborative working mechanism of the left and right brains by sharing a similarity matrix between the hemispheres, thereby establishing dynamic associations between the feature sets of the two hemispheres. This enhances the model's ability to capture inter-hemispheric interaction information while significantly reducing the number of model parameters.

Existing EEG datasets often suffer from a limited size of subjects and trials, making it difficult to generalize to new subjects or scenarios. Most existing cross-domain solutions primarily focus on loss function design while neglecting the contribution of backbone architectures to generalization capability [16]. Furthermore, Li et al. demonstrated that if a network's architecture aligns well with invariant correlation, it exhibits stronger robustness against distribution shifts [13]. To tackle this problem and boost model generalizability and adaptability, we introduce a new architecture: the Shared Hemispheres Mixture-of-Experts (SHMoE). SHMoE employs multiple expert modules, each specializing in distinct feature subspaces. SHMoE abstracts representations to a finite characteristics set, enhancing robustness to distribution shifts and improving the capture of varied emotional expressions. Unlike conventional MoE approaches, SHMoE uses a shared expert set and maps features into a unified space across both hemispheres. This ensures consistent and generalizable left-right hemisphere representations while increasing flexibility in modeling complex emotional patterns.

To address these challenges, we propose a novel **ShareLink** model that aims at achieving cross-hemispheric representation alignment through a shared expert mechanism while enhancing generalization capabilities. The architecture of ShareLink is illustrated in 1. Our methodology centers on three key innovations:

1. We introduce the Dynamic Shared Hemispheric Structure, which imposes non-Euclidean hemispheric structure constraints through shared learnable adjacency matrix parameters across the left-right hemisphere. This innovative architecture not only effectively aligns semantic representations but also enables the extraction of more discriminative hemispheric asymmetry features.
2. We propose the Cross-Hemisphere Attention, inspired by the collaborative mechanism of the human bi-hemispheres. By sharing a similarity matrix between the left and right hemispheres, it establishes dynamic inter-hemispheric links, enhancing the model's ability to capture interaction information. The shared constraints also reduce parameters, lowering overfitting risks.
3. We introduce the Shared Hemispheres Mixture-of-Experts to address the limitations of existing EEG models in generalizing to new subjects and scenarios. SHMoE leverages multiple expert modules to abstract representations into a finite set of characteristics. Furthermore, SHMoE employs a shared expert set, breaking down complex task into subtasks, and maps bi-hemisphere features into a unified space. This design ensures consistent and generalizable

left-right hemisphere representations while increasing flexibility in modeling complex emotional patterns.

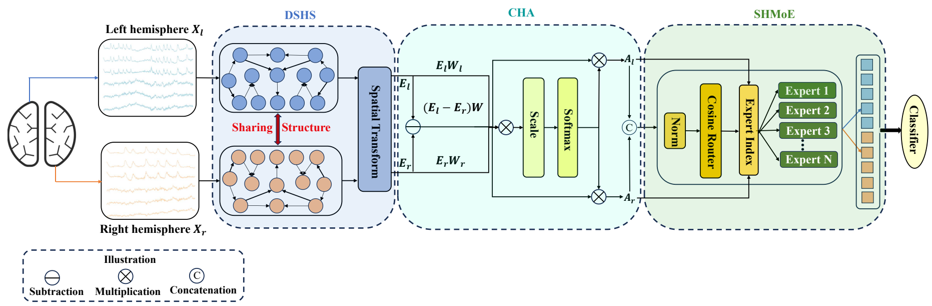

Fig. 1. The architecture of ShareLink.

2 Method

2.1 Problem formulation

The objective of this study is to achieve cross-subject emotion recognition using multi-channel EEG signals. We propose a deep learning model $f_\theta : X \to Y$ that maps EEG signal features X to emotion category probability distributions Y. By optimizing the model parameters θ on the training dataset $(X^{(train)}, Y^{(train)})$, our goal is to ensure high classification accuracy for f_θ on the cross-subject test dataset $(X^{(test)}, Y^{(test)})$. Notably, the training set and test set do not contain data from the same subject.

2.2 Dynamic Shared Hemispheric Structure

Existing non-Euclidean structures for EEG-based emotion recognition typically target all electrodes [23]. However, this approach faces two limitations: 1) The excessive degrees of freedom in the non-Euclidean structure make it prone to overfitting, and 2) difficulties in achieving semantic alignment between the left and right hemispheres. To address these issues, we propose DSHS, which enforces symmetric constraints on the non-Euclidean structure of the hemispheres by sharing learnable adjacency matrix parameters between the left and right hemispheres. This effectively aligns semantic representations while uncovering more discriminative hemispheric asymmetric features. The construction process is as follows:

We constructed a dynamically learned adjacency matrix $\mathbf{W} \in \mathbb{R}^{C/2 \times C/2}$ shared between the left and right hemispheres to characterize the relationships between vertex nodes, where C is the number of channels. Let $\mathbf{W}^*$ denote the optimal learned adjacency matrix, and $\mathbf{L}^*$ denote the Laplacian matrix of the hemispheric graphs.$\mathcal{G}$. $\mathbf{L}^* = \mathbf{D} - \mathbf{W}^* \in \mathbb{R}^{N \times N}$, where $\mathbf{D} \in \mathbb{R}^{C/2 \times C/2}$ is a diagonal matrix and the i-th diagonal element can be calculated by $\mathbf{D}_{ii} = \sum_j w_{ij}$. Especially, hemispheric graph $\mathcal{G}_\alpha = \{\mathbf{X}_\alpha, \mathbf{L}^*\}$ share Laplacian matrix, where $\alpha \in \{\text{left}, \text{right}\}$ denote left and right hemispheres, respectively. The graph convolution of the signal x_α with the vector of $\mathbf{U}^* g\left(\mathbf{\Lambda}^*\right)$ defined by the spatial filtering $g\left(\mathbf{L}^*\right)$ can be expressed as, $\mathbf{E}_\alpha = g\left(\mathbf{L}^*\right) \mathbf{x}_\alpha = \mathbf{U}^* g\left(\boldsymbol{\Lambda}^*\right) \mathbf{U}^{*T} \mathbf{x}_\alpha$, where $\mathbf{\Lambda}^* = \operatorname{diag}\left(\left[\lambda_0^*, \lambda_1^*, \ldots, \lambda_{N-1}^*\right]\right)$ represents a diagonal matrix.

Since directly computing the expression of $g\left(\mathbf{\Lambda}^*\right)$ is challenging, we employ K-th order Chebyshev polynomials to simplify the calculation process[6]. Specifically, let $\lambda_{\max}^*$ denote the largest diagonal element of $\mathbf{\Lambda}^*$. We normalize $\mathbf{\Lambda}^*$ as $\tilde{\mathbf{\Lambda}}^* = 2\boldsymbol{\Lambda}^*/\lambda_{\max}^* - \mathbf{I}_N$, where $\mathbf{I}_N$ is the $N \times N$ identity matrix. This normalization ensures that the diagonal elements of $\tilde{\mathbf{\Lambda}}^*$ lie within the interval $[-1, 1]$. Thus, we obtain that $g\left(\mathbf{\Lambda}^*\right)$ can be approximated by:

$$g\left(\mathbf{\Lambda}^*\right) = \sum_{k=0}^{K-1} \theta_k T_k\left(\tilde{\boldsymbol{\Lambda}}^*\right), \tag{1}$$

where θ_k is the coefficient of Chebyshev polynomials, and can be recursively calculated according to the following recursive expressions:

$$\begin{cases} T_0(x_\alpha) = 1, T_1(x_\alpha) = x_\alpha \\ T_k(x_\alpha) = 2x T_{k-1}(x_\alpha) - T_{k-2}(x_\alpha), \quad k \geq 2. \end{cases} \tag{2}$$

According to (1), we obtain the graph convolution operation of E_α can be rewritten as:

$$\mathbf{E}_\alpha = \sum_{k=0}^{K-1} \theta_k T_k\left(\tilde{\mathbf{L}}^*\right) \mathbf{x}_\alpha. \tag{3}$$

where $\tilde{\mathbf{L}}^* = 2\mathbf{L}^*/\lambda_{\max}^* - \mathbf{I}_N$.

2.3 Cross-Hemisphere Attention

Inspired by the collaboration between the left and right hemispheres of the human brain [19], we propose CHA. CHA aims to simulate the collaborative working mechanism of the left and right brains by sharing a similarity matrix between the hemispheres, thereby establishing dynamic associations between the channels of the two hemispheres and enhancing the model's ability to capture inter-hemispheric interaction information. Additionally, through shared constraints, this module reduces the number of model parameters and mitigates the risk of overfitting. The following formulas show the framework of this part. To simplify computation, we prove the following theorem:

Theorem 1. *Define the Query, Key, and Value matrices as follows:*

- *For the left brain:* $\mathbf{Q_l} = \mathbf{E_l}\mathbf{W_l}$, $\mathbf{K_l} = \mathbf{E_r}\mathbf{W_r}$, *and* $\mathbf{V_l} = \left(\mathbf{E_l} - \mathbf{E_r}\right)\mathbf{W_V}$.
- *For the right brain:* $\mathbf{Q_r} = \mathbf{E_r}\mathbf{W_r}$, $\mathbf{K_r} = \mathbf{E_l}\mathbf{W_l}$, *and* $\mathbf{V_r} = \left(\mathbf{E_r} - \mathbf{E_l}\right)\mathbf{W_V}$.

Therefore, $\mathbf{S_l} = \mathbf{S_r}^T$ *for the left and right hemispheres similarity matrix.*

Proof. $\mathbf{S_l}$ for the left brain is computed as the scaled dot product of the Query $\mathbf{Q}$ and the transpose of the Key $\mathbf{K}$:

$$\mathbf{S_l} = \frac{\mathbf{Q_l}\mathbf{K_l}^T}{\sqrt{d_k}}.$$

For the right brain, we also follow the above representation:

$$\mathbf{S_r} = \frac{\mathbf{Q_r}\mathbf{K_r}^T}{\sqrt{d_k}} = \frac{(\mathbf{E_r}\mathbf{W_r})(\mathbf{E_l}\mathbf{W_l})^T}{\sqrt{d_k}} = \frac{\mathbf{K_l}\mathbf{Q_l}^T}{\sqrt{d_k}} = \frac{(\mathbf{Q_l}\mathbf{K_l}^T)^T}{\sqrt{d_k}} = \mathbf{S_l}^T.$$

This completes the proof.

Let $S \in \mathbb{R}^{C/2 \times C/2}$ denote the similarity matrix of the left hemisphere. The output of CHA can be represented as:

$$\mathbf{A_l} = \mathbf{SV} \in \mathbb{R}^{C/2 \times d_v}, \mathbf{A_r} = \mathbf{S}^T\mathbf{V} \in \mathbb{R}^{C/2 \times d_v}, \tag{4}$$

Additionally, to prevent the dot product values from becoming too large, we apply a scaling operation to the raw attention scores. After that we apply the softmax function, converting the scaled scores into a probability distribution.

The outputs of the left and right brains, $\mathbf{A_l}$ and $\mathbf{A_r}$, aggregate information from each other, enabling bidirectional information interaction. By sharing the similarity matrix between the left and right hemispheres, the model establishes dynamic inter-hemispheric links, which significantly enhance its ability to capture interaction information. Additionally, the shared constraints reduce the number of parameters, thereby mitigating the risk of overfitting.

2.4 Shared Hemispheres Mixture-of-Experts

To overcome the generalization limitations of EEG emotion recognition models caused by small datasets and loss-based cross-domain methods, we propose the Shared Hemispheric Mixture-of-Experts. Inspired by the observation that architecture alignment with invariant correlations enhances robustness to distribution shifts [13]. SHMoE employs multiple expert modules, mapping features and expert embeddings onto a hypersphere and performing L2 normalization. Each module specializes in handling different emotional attributes, making router dependent only on directional alignment. SHMoE abstracts representations to a finite characteristics set, enhancing robustness to distribution shifts and improving the capture of varied emotional expressions.

For MoE models, linear routers are commonly adopted in vision tasks [22], while recent studies in NLP demonstrate that the cosine router achieves superior performance in cross-lingual language tasks [1]. In the cosine router, the input embedding is first projected onto a hypersphere, followed by multiplication with a learnable embedding matrix $\mathbf{E} \in \mathbb{R}^{N \times d_e}$, where N is the number of experts.

The SHMoE implementation process is as follows: First, we fuse $\mathbf{A_l}$ and $\mathbf{A_r}$ by concatenation, $\mathbf{A} = [\mathbf{A_l}, \mathbf{A_r}]$, where $[\cdot, \cdot]$ represents the concatenation operation.

$$G(\mathbf{A}) = \mathrm{TOP}_k \left(\mathrm{Softmax} \left(\frac{\mathbf{E}^T \mathbf{W} \mathbf{A}}{\tau \|\mathbf{W}\mathbf{A}\| \|\mathbf{E}\|} \right) \right), \tag{5}$$

where $\mathbf{W}$ is the learnable gate parameter, and τ is a hyper-parameter. Inspired by [8,30,13], we opine that the linear router would face difficulty in EEG-based emotion recognition. For example, an EEG signal corresponding to a happy emotion in one subject is likely more similar to other EEG signals from the same subject than to those from other subjects. This issue can be alleviated by incorporating a codebook for emotional attributes and matched filters for detecting specific emotional patterns.

Subsequently, based on the expert indices selected by $G(\mathbf{A})$, the data from different channels are routed to their corresponding expert modules. Finally, the results from all channels are aggregated through summation to produce the final outputs $\mathbf{M_l}$ and $\mathbf{M_r}$.

3 Experiment and Discussions

3.1 Datasets and Pre-process

In our experiments, we utilized two publicly available EEG emotion datasets, SEED [28] and SEED-IV [27]. The detailed information of these datasets is as follows: **SEED** EEG data are from 15 subjects with 62 channels, recorded during the viewing of *negative*, *neutral*, and *positive* film clips across three sessions per subject (5 clips/emotion, 15 trials per session). Each trial contains 185-238 samples, yielding about 3,400 samples/session. Compared to SEED, **SEED-IV** follows the same settings but introduces an additional emotion, *fear*. Film clips across three sessions per subject (6 clips/emotion, 24 trials per session). Each trial contains 12–64 samples, yielding about 830 samples/session.

To ensure a fair comparison with existing studies, we directly adopted the pre-computed differential entropy from the SEED and SEED-IV, using segmented time windows as input which were smoothed using a Linear Dynamic System for each EEG channel across five frequency bands. Finally, from the full-brain data comprising 62 channels, we extracted left hemisphere data consisting of left-brain channels and midline channels, and right hemisphere data consisting of right-brain channels and midline channels. Notably, the training set and test set do not contain data from the same subject.

3.2 Experiment Setting

ShareLink was trained with a batch size of 512 and 0.1 dropout, using the AdamW optimizer and cross-entropy loss. Hyperparameters were selected via grid search and include learning rate {1e-3,1e-4,1e-5,1e-6}, embedding dimensions {16,32,64}, and 62 experts. Training runs for 500 epochs.

We conducted cross-subject experiments on both the SEED and SEED-IV datasets. In this experiment, we employed the Leave-One-Subject-Out (LOSO) cross-validation method. Specifically, for both datasets, the EEG signals of a single subject were used as the test set, while the EEG signals of the remaining subjects were used as the training set. The average accuracy and standard deviation across all subjects were adopted as the final evaluation metrics.

3.3 Comparative Results

We present the comparative results on the SEED and SEED-IV datasets with other transfer methods. All baseline methods share the same experimental setup as ours, and their results are all taken from the references. The results are summarized in Table 1. The results demonstrate that ShareLink consistently outperforms all other methods on both datasets, achieving the highest classification accuracy with relatively low standard deviations. This indicates that ShareLink is not only more accurate but also more stable across different subjects compared to the other methods. ShareLink achieves improvements of 6.86% on SEED and 24.85% on SEED-IV, compared to the best referred method. In summary, ShareLink demonstrates superior accuracy and stability compared to both transfer methods, and it yields the best results across both datasets, indicating strong generalization ability.

Table 1. Cross-subject classification accuracy (mean/standard deviation) on SEED and SEED-IV.

Method	SEED		SEED-IV	
	ACC	STD	ACC	STD
ULSIF[12]	51.18%	13.57%	32.99%	11.05%
STM[3]	51.23%	14.82%	39.39%	12.40%
TCA[21]	63.64%	14.88%	56.56%	13.77%
DAN[14]	65.84%	2.25%	32.44%	9.02%
DCORAL[24]	66.29%	4.53%	37.43%	3.08%
DDC[25]	68.99%	3.23%	37.71%	6.36%
DANN[9]	75.08%	11.18%	47.59%	10.01%
ShareLink	**80.61%**	**6.16%**	**63.33%**	**8.29%**

3.4 Ablation Study

ShareLink integrates three core modules: DSHS, CHA, and SHMoE. To evaluate the contribution of each module, we conducted ablation studies by sequentially removing DSHS, CHA, and SHMoE. Specifically, removing DSHS reduced accuracy by 12.14% and 16.50%, while removing CHA led to decreases of 16.30% and 20.39%, respectively. Similarly, the absence of SHMoE resulted in performance drops of 11.74% and 16.17%. The observed degradation across all scenarios highlights the critical role of each module in enhancing model performance.

Table 2. Ablation study for cross-subject classification on SEED and SEED-IV

Method	SEED		SEED-IV	
	ACC	STD	ACC	STD
w/o DSHS	70.82%	6.29%	52.88%	9.31%
w/o CHA	67.47%	9.04%	50.42%	4.08%
w/o SHMoE	71.14%	8.20%	53.09%	9.11%
ShareLink	**80.61**%	**6.16**%	**63.33**%	**8.29**%

4 Conclusion

We propose ShareLink, enhancing EEG-based emotion recognition via hemispheric asymmetry and interactions, addressing two limitations: hemispheric semantic misalignment and cross-subject generalizability. Thus, we develop three modules: DSHS aligns hemispheric features by shared adjacency matrix; CHA models inter-hemispheric dynamics through shared similarity matrices; and SHMoE transforms bilateral representations into a shared latent space. Validated on SEED and SEED-IV, ShareLink achieves 80.61%±6.16% and 63.33%±8.29% accuracy in cross-subject trials. ShareLink provides new insights into neurophysiologically inspired computational models for emotion recognition.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant U2241210, in part by the Basic Research Project of Shenzhen under Grant JCYJ20220818101216034, in part by the Shenzhen-Hong Kong-Macau cooperation Project under Grant SGDX20210823103403028.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chi, Z., Dong, L., Huang, S., Dai, D., Ma, S., Patra, B., Singhal, S., Bajaj, P., Song, X., Mao, X.L., et al.: On the representation collapse of sparse mixture of experts. Advances in Neural Information Processing Systems **35**, 34600–34613 (2022)
2. Chougule, S., Bhosale, S., Borle, V., Chaugule, V., Mali, Y.: Emotion recognition based personal entertainment robot using ml & ip. Emotion **5**(8) (2020)
3. Chu, W.S., De la Torre, F., Cohn, J.F.: Selective transfer machine for personalized facial expression analysis. IEEE transactions on pattern analysis and machine intelligence **39**(3), 529–545 (2016)
4. Costanzo, E.Y., Villarreal, M., Drucaroff, L.J., Ortiz-Villafañe, M., Castro, M.N., Goldschmidt, M., Wainsztein, A.E., Ladrón-de Guevara, M.S., Romero, C., Brusco, L.I., et al.: Hemispheric specialization in affective responses, cerebral dominance for language, and handedness: lateralization of emotion, language, and dexterity. Behavioural brain research **288**, 11–19 (2015)
5. Cui, H., Liu, A., Zhang, X., Chen, X., Wang, K., Chen, X.: Eeg-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network. Knowledge-Based Systems **205**, 106243 (2020)

6. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. Advances in neural information processing systems **29** (2016)
7. Ding, Y., Robinson, N., Zhang, S., Zeng, Q., Guan, C.: Tsception: Capturing temporal dynamics and spatial asymmetry from eeg for emotion recognition. IEEE Transactions on Affective Computing **14**(3), 2238–2250 (2022)
8. Ferrari, V., Zisserman, A.: Learning visual attributes. Advances in neural information processing systems **20** (2007)
9. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., March, M., Lempitsky, V.: Domain-adversarial training of neural networks. Journal of machine learning research **17**(59), 1–35 (2016)
10. Herrington, J.D., Heller, W., Mohanty, A., Engels, A.S., Banich, M.T., Webb, A.G., Miller, G.A.: Localization of asymmetric brain function in emotion and depression. Psychophysiology **47**(3), 442–454 (2010)
11. Huang, D., Chen, S., Liu, C., Zheng, L., Tian, Z., Jiang, D.: Differences first in asymmetric brain: A bi-hemisphere discrepancy convolutional neural network for eeg emotion recognition. Neurocomputing **448**, 140–151 (2021)
12. Kanamori, T., Hido, S., Sugiyama, M.: A least-squares approach to direct importance estimation. The Journal of Machine Learning Research **10**, 1391–1445 (2009)
13. Li, B., Shen, Y., Yang, J., Wang, Y., Ren, J., Che, T., Zhang, J., Liu, Z.: Sparse mixture-of-experts are domain generalizable learners. arXiv preprint arXiv:2206.04046 (2022)
14. Li, H., Jin, Y.M., Zheng, W.L., Lu, B.L.: Cross-subject emotion recognition using deep adaptation networks. In: Neural information processing: 25th international conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, proceedings, part V 25. pp. 403–413. Springer (2018)
15. Li, M., Pan, J., Gao, Y., Shen, Y., Luo, F., Dai, J., Hao, A., Qin, H.: Neurophysiological and subjective analysis of vr emotion induction paradigm. IEEE Transactions on Visualization and Computer Graphics **28**(11), 3832–3842 (2022)
16. Li, Y., Wang, L., Zheng, W., Zong, Y., Qi, L., Cui, Z., Zhang, T., Song, T.: A novel bi-hemispheric discrepancy model for eeg emotion recognition. IEEE Transactions on Cognitive and Developmental Systems **13**(2), 354–367 (2020)
17. Lin, Y., Xu, G.X., Liang, H., Wang, Y., Wan, F., Li, Y.: Brain region knowledge based dual-stream transformer for eeg emotion recognition. IEEE Transactions on Consumer Electronics (2025)
18. Liu, D., Dai, W., Zhang, H., Jin, X., Cao, J., Kong, W.: Brain-machine coupled learning method for facial emotion recognition. IEEE transactions on pattern analysis and machine intelligence **45**(9), 10703–10717 (2023)
19. Ocklenburg, S., Guo, Z.V.: Cross-hemispheric communication: Insights on lateralized brain functions. Neuron (2024)
20. Ocklenburg, S., Peterburs, J., Mundorf, A.: Hemispheric asymmetries in the amygdala: A comparative primer. Progress in Neurobiology **214**, 102283 (2022)
21. Pan, S.J., Tsang, I.W., Kwok, J.T., Yang, Q.: Domain adaptation via transfer component analysis. IEEE transactions on neural networks **22**(2), 199–210 (2010)
22. Riquelme, C., Puigcerver, J., Mustafa, B., Neumann, M., Jenatton, R., Susano Pinto, A., Keysers, D., Houlsby, N.: Scaling vision with sparse mixture of experts. Advances in Neural Information Processing Systems **34**, 8583–8595 (2021)
23. Song, T., Zheng, W., Song, P., Cui, Z.: Eeg emotion recognition using dynamical graph convolutional neural networks. IEEE Transactions on Affective Computing **11**(3), 532–541 (2018)

24. Sun, B., Saenko, K.: Deep coral: Correlation alignment for deep domain adaptation. In: Computer vision–ECCV 2016 workshops: Amsterdam, the Netherlands, October 8-10 and 15-16, 2016, proceedings, part III 14. pp. 443–450. Springer (2016)
25. Tzeng, E., Hoffman, J., Zhang, N., Saenko, K., Darrell, T.: Deep domain confusion: Maximizing for domain invariance. arXiv preprint arXiv:1412.3474 (2014)
26. Wang, S., Zhou, T., Shen, Y., Li, Y., Huang, G., Hu, Y.: Generative ai enables eeg super-resolution via spatio-temporal adaptive diffusion learning. IEEE Transactions on Consumer Electronics (2025)
27. Zheng, W.L., Liu, W., Lu, Y., Lu, B.L., Cichocki, A.: Emotionmeter: A multimodal framework for recognizing human emotions. IEEE transactions on cybernetics **49**(3), 1110–1122 (2018)
28. Zheng, W.L., Lu, B.L.: Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. IEEE Transactions on autonomous mental development **7**(3), 162–175 (2015)
29. Zhou, B., Lin, L., Chen, J.: Enhancing emotion recognition through brain asymmetry and multi-angle fusion network. Biomedical Signal Processing and Control **102**, 107324 (2025)
30. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Object detectors emerge in deep scene cnns. arXiv preprint arXiv:1412.6856 (2014)