A novel Fourier Adjacency Transformer for advanced EEG emotion recognition

Jinfeng Wang[1*][0000−0002−9086−7518], Yanhao Huang[2*], Sifan Song[2], Boqian Wang[2], Jionglong Su[2**], and Jiaman Ding[1**]

[1] Faculty of Information Engineering and Automation,
Kunming University of Science and Technology, Kunming, China
[2] School of Artificial Intelligence and Advanced Computing,
Xi'an Jiaotong-Liverpool University, Suzhou, China
Jionglong.Su@xjtlu.edu.cn
Jiamanding@kust.edu.cn

Abstract. EEG emotion recognition faces significant hurdles due to noise interference, signal nonstationarity, and the inherent complexity of brain activity which make accurately emotion classification. In this study, we present the Fourier Adjacency Transformer, a novel framework that seamlessly integrates Fourier-based periodic analysis with graph-driven structural modeling. Our method first leverages novel Fourier-inspired modules to extract periodic features from embedded EEG signals, effectively decoupling them from aperiodic components. Subsequently, we employ an adjacency attention scheme to reinforce universal inter-channel correlation patterns, coupling these patterns with their sample-based counterparts. Empirical evaluations on SEED and DEAP datasets demonstrate that our method surpasses existing state-of-the-art techniques, achieving an improvement of approximately 6.5% in recognition accuracy. By unifying periodicity and structural insights, this framework offers a promising direction for future research in EEG emotion analysis. The code can be found at: https://github.com/YanhaoHuang23/FAT

Keywords: EEG · Fourier · Adjacency · Transformer.

1 Introduction

Electroencephalography (EEG) emotion recognition has gained significant attention in affective computing, brain-computer interfaces (BCIs), and mental health assessment due to its high temporal resolution and objective measurement of emotional states [1]. However, challenges such as non-stationarity, low signal-to-noise ratio, and inter-subject variability hinder robust and generalized emotion recognition [1,2,29,9].

Existing approaches primarily rely on raw EEG signals, handcrafted features, or deep learning-based representations. While deep models operating on

---

* These authors are contribute equally.
** The corresponding authors.

raw EEG data preserve the full temporal structure, they struggle with high dimensionality, noise sensitivity, and poor generalization [28,13]. Handcrafted feature-based methods, such as differential entropy (DE) and power spectral density (PSD), enhance interpretability and robustness but may lead to information loss [9,6]. Despite the significance of periodic EEG features, their effective extraction and utilization in deep learning remain underexplored [17,14].

Deep learning-based methods can be categorized into temporal, spatial, and hybrid modeling approaches. Recurrent neural networks (RNNs), including LSTMs (Long Short-Term Memory) and GRUs (Gated Recurrent Units), capture temporal dependencies but suffer from vanishing gradients and limited long-range modeling [25,19]. Convolutional neural networks (CNNs) extract localized spatial patterns but lack temporal modeling capacity [18]. Hybrid approaches, such as graph neural networks, leverage EEG's topological structure but rely on predefined connectivity assumptions, limiting generalization across datasets [18,30].

Recently, Transformer-based models have demonstrated strong performance by leveraging self-attention to capture global dependencies [20,7,21]. However, their data inefficiency and lack of inductive bias hinder stability and generalization in EEG applications. Addressing these challenges, this study introduces a novel Transformer-based framework that enhances EEG feature extraction and improves generalization.

- We propose Fourier Analytic Linear (FAL) Layer as a novel query-key-value mapping mechanism in self-attention, enabling the decoupling of periodic and aperiodic EEG representations to capture inherent periodic structures.
- By integrating FAL with multi-head self-attention (MHSA) [23], we propose Fourier Adjacent Attention (FAA) groups attention heads based on learned periodicity, significantly enhancing model performance.
- Replacing standard self-attention with FAA, we develop Fourier Adjacent Transformer (FAT), a novel EEG emotion recognition framework achieving state-of-the-art (SOTA) performance and superior generalization across SEED-family benchmarks and the DEAP dataset.

2   Methods

2.1   Fourier Analytic Linear

To effectively decouple and extract the periodic components in EEG signals, we propose a novel module, FAL, inspired by Fourier Analysis Networks (FAN) [5]. FAN enhances the ability of neural networks to model periodic signals by incorporating a Fourier transform-based periodic modeling mechanism. However, within the FAN structure, the aperiodic component is typically processed through a nonlinear activation function, which may compromise its linear mapping capability, particularly in the Transformer architecture.

The key innovation of FAL lies in ensuring that the aperiodic component remains purely linear. This is achieved by replacing the nonlinear activation function applied to the aperiodic component in FAN with an Identity Function,
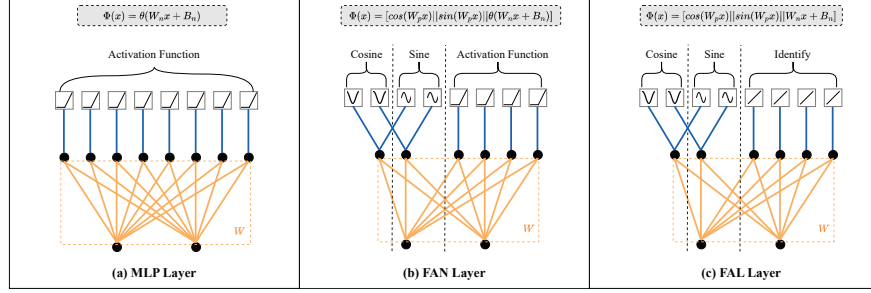
Fig. 1: Illustrations of MLP Layer vs. FAN Layer vs. FAL Layer.

thereby preserving its linearity. By maintaining a purely linear transformation for the aperiodic component, FAL can seamlessly replace the standard linear mapping layer in Transformer structures, ensuring compatibility with self-attention mechanisms while retaining the benefits of periodic signal modeling introduced by Fourier-Transform. The formulaic expression of FAL is

$$\Phi(x) = [cos(W_p x)||sin(W_p x)||W_n x + B_n], \tag{1}$$

where $x \in \mathbb{R}^{d_{in}}$ and the $||$ refer to concatenate operation. The FAL consists of two main parts. First, the periodic part is obtained by unbiased linear transformation and Fourier basis function modeling:

$$\Phi_p(x) = [cos(W_p x)||sin(W_p x)]. \tag{2}$$

As shown in Fig. 1, in the aperiodic part, unlike FAN, we directly use a linear transformation thus ensuring its linear mapping capability, i.e.,

$$\Phi_a(x) = W_n x + B_n. \tag{3}$$

This design enables FAL to function as a linear projection layer within the Transformer architecture. Its structure facilitates the decoupling of periodic and aperiodic representations along the channel dimension, providing greater flexibility for the subsequent MHSA. Generally, compared to the traditional MLP-based QKV (Query, Key and Value) projection, FAL offers a significant advantage by explicitly capturing the inherent periodic structure of EEG signals while maintaining high computational efficiency.

2.2 Fourier Adjacent Attention

Furthermore, building upon FAL, we propose FAA to enhance the self-attention mechanism's ability to model periodic signals. FAA leverages FAL's capability to disentangle periodic and aperiodic components while incorporating EEG inter-channel adjacencies into the attention computation process. This integration enables FAA to more effectively capture both universal and sample-specific

Fourier Adjacent Transformer (FAT)

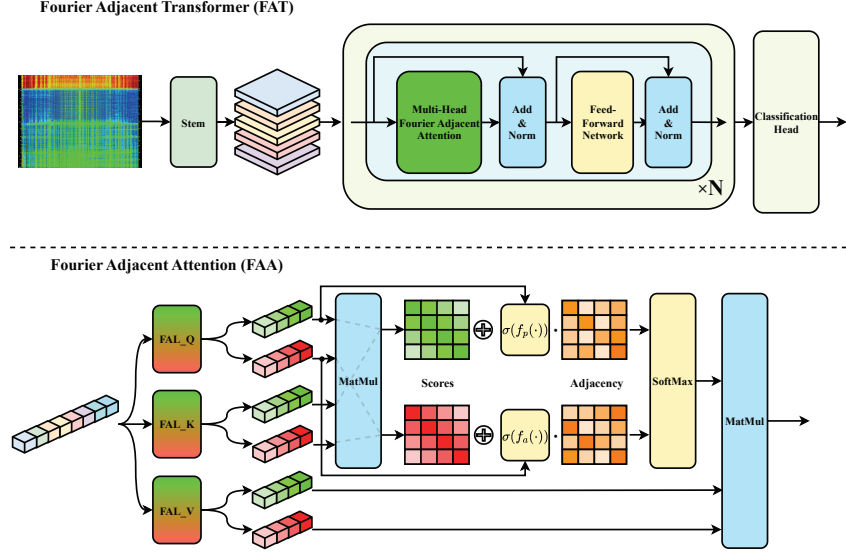**Fourier Adjacent Attention (FAA)**

Fig. 2: The overview of the FAT and FAA.

connectivity patterns in EEG signals. Fig. 2 illustrates the computational flow of FAA in the single-head case.

As shown in Fig. 2, in FAA, the FAL-based QKV projector transforms the input sequences into query, key, and value representations while simultaneously decoupling them into periodic and aperiodic components. Specifically, for each input feature vector x, the transformation results in

$$Q = [Q_p || Q_n], \ K = [K_p || K_n], \ V = [V_p || V_n], \qquad (4)$$

$$Q_p || Q_n = FAL\_Q(x), \ K_p || K_n = FAL\_K(x), \ V_p || V_n = FAL\_V(x), \qquad (5)$$

where Q_p represents the periodic component of query, and Q_n refer to the aperiodic component of query. The same decomposition applies to Key and Value. This FAL-based projection ensures that FAA can independently model periodic and aperiodic information, providing a comprehensive representation for subsequent attention computations (all subscripts a and n in all formulas represent nonperiodicity and periodicity, respectively).

In FAA, since the query, key, and value representations are decoupled into periodic and aperiodic components, we apply scaled dot-product attention separately to each part, resulting in two distinct attention score matrices. Unlike traditional MHSA, which directly computes the attention matrix using the Soft-Max function after obtaining the attention scores, FAA introduces a pair of learnable adjacency matrices with depth-sharing parameters. These matrices are

then added to the attention score matrices through a gating module, and the resulting sum is fed into the SoftMax function to compute two final attention matrices $Attention_p$ and $Attention_a$ They can be formulated as follows:

$$Attention_{p/n} = SoftMax(\frac{Q_{p/n}K_{p/n}^T}{\sqrt{d_k}} + \sigma(f_{p/n}(Q_{p/n})) \times Adjacency_{p/n}), \quad (6)$$

where $f_{p/n}$ is implemented by a linear layer and σ refers to the Sigmoid function. After deriving the two attention matrices $Attention_n$ and $Attention_n$, the final output of the FAA is calculated:

$$O = [Attention_p \cdot V_p || Attention_n \cdot V_n]. \quad (7)$$

Another key contribution of FAA, beyond the incorporation of periodic representation processing, is the introduction of additive adjacency matrices in the computation of the attention matrix. We propose that this pair of adjacency matrices captures the structural correlations between EEG channels from periodic and aperiodic perspectives, respectively. By integrating them with the attention score matrices, which reflect data-driven, sample-specific correlations, FAA effectively mitigates the instability arising from the lack of inductive bias in traditional MHSA.

2.3   Fourier Adjacent Transformer

As shown in Fig. 2, we replace the self-attention mechanism of MHSA with FAA while retaining the multi-head design, forming a new module Multi-Head Fourier Adjacency Attention. We shall integrate this novel module into the Transformer architecture, leading to the proposed novel FAT model.

Since FAT retains the Transformer architecture, it requires sequencing and embedding of the raw EEG data. In this study, we use DE features of EEG as model inputs, followed by patch embedding and positional embedding to generate input representations suitable for multi-head FAA. The corresponding formula is expressed as follow:

$$x = ReLU(BN(f_2(ReLU(BN(f_1(input)))))) \quad (8)$$

where $input \in \mathbb{R}^{B \times C \times F}$, $x \in \mathbb{R}^{B \times C \times E}$ and BN refers to batch normalization layer, f_1 ($F \to \frac{E}{2}$) and f_2 ($\frac{E}{2} \to E$) are represent linear layers. In this study, F equals five, i.e., representing the five frequency bands of DE feature.

2.4   Validation scheme

To comprehensively evaluate the performance of FAT in the EEG emotion recognition task, we conduct validation through both subject-dependent and subject-independent experiments.

In the subject-dependent experiments, the training and test sets each contain a subset of data from all subjects in the dataset. This setup is designed to assess the learning and recognition capabilities of FAT. The data partitioning scheme strictly follows the configurations of various benchmark datasets.

In the subject-independent experiments, we adopt the Leave-One-Subject-Out (LOSO) approach, where the test set consists of data from subjects not included in the training set. This methodology is used to evaluate the generalization ability of FAT across unseen subjects.

3   Results and discussions

3.1   Implementation details

Datasets In this study, to evaluate the performance of FAT across tasks of varying complexity, we selected SEED [27,6], SEED-IV [26], SEED-V [16], and SEED-VII [10] from the SEED family, corresponding to three-class, four-class, five-class, and seven-class emotion classification tasks, respectively. Additionally, we utilized the widely recognized DEAP [12] dataset, focusing on the classification of two key emotional states: Valence and Arousal. In addition, all result data marked with * in the Table 1 and Table 2 in this section are from the literature, and the bolded values are the best performance.

Experimental settings Our method **is** implemented using PyTorch, and all experiments are conducted on NVIDIA A100 (80G) GPUs. For optimization, we employ Adam with a learning rate of 0.001 and a weight decay of 0.0001. The model is trained for 200 epochs, with a batch size of 64 for the SEED dataset and 32 for all other datasets. In the FAT, the number of attention heads (h) is set to eight, while the model depth is fixed at six.

For data augmentation, we randomly scale specific frequency bands of the frequency-domain features, restricting the scaling factor to (0.9, 1.1). To further enhance periodicity, we introduce a sine wave perturbation with a frequency of 2 and an amplitude of 0.05 in the frequency domain. Additionally, we employ Mixup, a data augmentation technique that linearly combines both samples and their corresponding labels, to further improve the model's generalization ability.

3.2   Recognition capability

The experimental results in Table 1 demonstrate that the proposed FAT consistently outperforms existing SOTA methods across the SEED family and DEAP datasets. Specifically, FAT achieves performance improvements of 2.9%, 6.5%, 6.6%, 5.4%, and 6.2% over the previous SOTA on SEED-IV, SEED-V, SEED-VII, DEAP-Valence, and DEAP-Arousal, respectively, while maintaining competitive performance on the SEED dataset. These results underscore the effectiveness of FAT in learning discriminative representations and its strong capability in addressing complex recognition tasks.

Table 1: Subject-dependent performance comparison (Mean/Std).

Method	SEED	SEED IV	SEED V	SEED VII	DEAP-Valence	DEAP-Arousal
Scheme	9:6	16:8	3-fold	4-fold	10-fold	10-fold
KNN [4]	80.88/9.73	55.82/18.82	61.03/12.24	36.87/4.41	67.40/7.28	67.80/6.94
SVM [22]	83.99/8.79	58.57/20.34	67.38/11.72	40.07/8.61	65.80/4.62	69.89/5.60
LSTM [8]	–	–	–	–	74.76/3.41	71.81/3.99
DGCNN [19]	89.75/8.51	69.09/13.79	69.52/9.57	46.36/7.33	84.69/6.20	83.02/7.07
BiHDM [15]	93.07/8.39	73.38/15.75	63.36/10.27	45.57/8.75	–	–
RGNN [30,3,10]	94.24/5.95*	79.37/10.54*	54.72/14.30	48.5/6.83*	–	–
PGCN [11]	**94.30/6.19**	78.63/17.19	73.52/10.60	49.94/7.48	–	–
Conformer [21]	88.68/9.97	76.7/16.57	78.09/9.3	54.72/6.98	80.71/2.30	79.42/3.03
FAT (w/ FAN)	93.18/11.28	79.03/16.27	82.95/8.64	59.42/8.59	80.47/2.89	83.09/3.85
FAT (ours)	92.10/6.46	**82.30/13.21**	**84.57/8.3**	**61.30/9.81**	**90.10/2.71**	**89.18/3.50**

3.3 Generalization capability

As shown in Table 2, our proposed FAT demonstrates exceptional generalization capability compared to the current SOTA methods. Notably, on the most complex SEED-VII dataset, FAT outperforms Conformer, a leading SOTA method based on the same Transformer architecture, by approximately 5%. We believe this consistent improvement suggests that FAT effectively learns to distinguish between universal and sample-specific information in EEG signals.

Table 2: Subject-independent Performance comparison (Mean/Std).

Method	SEED	SEED IV	SEED V	SEED VII	DEAP-Valence	DEAP-Arousal
KNN [4]	48.4/12.50	32.92/8.08	30.26/16.67	21.01/4.67	55.35/4.3	52.86/4.22
SVM [22]	56.48/15.20	36.28/12.99	39.49/16.20	30.73/11.12	54.7/0.36	55.9/0.76
LSTM [8]	–	–	–	–	60.31/0.24	59.78/0.27
DGCNN [19]	77.29/9.71	53.84/9.87	64.05/15.84	35.94/7.88	59.72/1.38	57.26/1.29
BiHDM [15]	84.17/8.74	68.7/13.42	57.83/17.44	34.34/6.18	–	–
RGNN [30,3,10]	85.30/6.72*	73.84/8.02*	66.28/16.71*	37.49/5.44*	–	–
PGCN [11]	83.58/7.36	70.22/11.46	63.38/8.88	38.78/5.39	–	–
Conformer [21]	79.16/7.68	67.07/11.22	59.95/13.50	45.60/9.74	62.4/0.20	59.9/0.6
FAT (w/ FAN)	83.24/5.58	70.52/11.69	64.66/14.49	49/10.32	62.59/0.64	60.00/0.48
FAT (ours)	**85.68/4.16**	**73.94/10.42**	**71.22/16.21**	**51.12/8.51**	**62.70/0.80**	**60.00/0.40**

3.4 Ablation study

To determine the optimal configuration of FAT, we conduct comprehensive ablation experiments on both data and model configurations.

The experimental results in Table 3 show that the feature richness of the input data is positively correlated with the performance of the model.

In Table 4, AdA refers to the additive adjacency matrix. Following the FAN configuration, P-ratio represents the ratio of the output dimensions of the cosine and sine basis functions in FAL to the total output dimensions of FAL. Given

Table 3: Performance comparison for different frequency bands.

Band	SEED IV	SEED V	SEED VII
δ band	59.79/9.73	51.66/11.37	40.97/7.77
θ band	57.08/11.59	55.58/11.41	41.34/5.45
α band	53.3/10.13	58.69/12.36	38.53/6.31
β band	55.57/10.24	58.58/14.92	42.31/5.77
γ band	51.89/9.61	54.29/13.11	41.03/8.65
β, γ bands	59.12/10.74	59.14/14.65	46.9/10.7
β, γ, δ bands	67.5/11.43	61.77/13.65	48.16/10.19
All bands	73.94/10.42	71.22/16.21	51.12/8.51

Table 4: Ablation studies on configurations and hyperparameters.

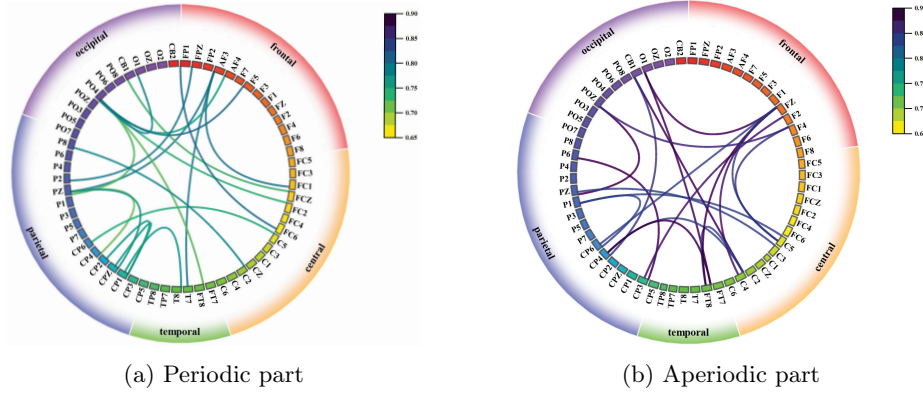| AdA	P-ratio				LOSO Performance
	0	0.125	0.25	0.375	
		✓			67.78/9.69
✓		✓			68.97/11.1
✓			✓		69.33/11.95
✓				✓	**73.94/10.42**
				✓	70.30/9.89
✓				✓	69.03/10.96

(a) Periodic part

(b) Aperiodic part

Fig. 3: The top 15 correlations in each of the two adjacency matrices.

that this study employs eight attention heads and an embedding dimension of 256, the possible values for P-ratio are 0, 0.125, 0.25, and 0.375. When P-ratio is set to 0, FAL degenerates into a linear layer. Experimental results indicate that FAL significantly enhances the performance of FAT. Furthermore, the periodic feature extraction capability of FAL helps to stimulate the additivity adjacency matrix, enabling the model to achieve superior accuracy.

3.5   Visualization

The adjacency matrix we introduced enables FAT to effectively capture EEG inter-channel relationships, similar to GNNs, thereby enhancing the model's interpretability. As shown in Fig. 3, the adjacency matrices of both the periodic and aperiodic components reveal distinct inter-channel structural relationships. Notably, the structural patterns learned from these two components exhibit significant differences, highlighting the model's ability to extract diverse features from different signal properties.

4   Conclusion

In this paper, we propose a novel EEG emotion recognition model, the Fourier Adjacency Transformer, whose core component is Fourier Adjacency Attention. In Neighbor Attention, we extract and decouple the periodic and aperiodic representations in EEG signals while explicitly learning both universal and sample-specific connection patterns between EEG channels. This is achieved by introducing a adjacency matrix into the attention computation process. Experimental results demonstrate that FAT consistently outperforms other SOTA methods in terms of learning capability, recognition accuracy, and generalization. These findings highlight the potential of incorporating Fourier-based cognition into EEG analysis, and we plan to further explore this direction in future.

5   Future work

While our results confirm the effectiveness of Fourier-inspired components (e.g., FAL and FAA) in enhancing EEG-based emotion recognition, the systematic utility of periodic feature modeling in EEG decoding has yet to be thoroughly examined. Future work will aim to extend the exploration of Fourier-based principles beyond emotion tasks, assessing their generalizability and potential impact across a broader range of EEG decoding applications.

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

References

1. Alarcao, S.M., Fonseca, M.J.: Emotions recognition using eeg signals: A survey. IEEE transactions on affective computing **10**(3), 374–393 (2017)
2. Bashivan, P., Rish, I., Yeasin, M., Codella, N.: Learning representations from eeg with deep recurrent-convolutional neural networks. arXiv preprint arXiv:1511.06448 (2015)
3. Chen, B., Chen, C.P., Zhang, T.: Ugan: Uncertainty-guided graph augmentation network for eeg emotion recognition. IEEE Transactions on Computational Social Systems (2024)
4. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE transactions on information theory **13**(1), 21–27 (1967)
5. Dong, Y., Li, G., Tao, Y., et al.: Fan: Fourier analysis networks. arXiv preprint arXiv:2410.02675 (2024)
6. Duan, R.N., Zhu, J.Y., Lu, B.L.: Differential entropy feature for EEG-based emotion classification. In: 6th International IEEE/EMBS Conference on Neural Engineering (NER). pp. 81–84. IEEE (2013)

7. Feng, L., Cheng, C., Zhao, M., Deng, H., Zhang, Y.: Eeg-based emotion recognition using spatial-temporal graph convolutional lstm with attention mechanism. IEEE Journal of Biomedical and Health Informatics **26**(11), 5406–5417 (2022)
8. Graves, A., Graves, A.: Long short-term memory. Supervised sequence labelling with recurrent neural networks pp. 37–45 (2012)
9. Jenke, R., Peer, A., Buss, M.: Feature extraction and selection for emotion recognition from eeg. IEEE Transactions on Affective computing **5**(3), 327–339 (2014)
10. Jiang, W.B., Liu, X.H., Zheng, W.L., Lu, B.L.: Seed-vii: A multimodal dataset of six basic emotions with continuous labels for emotion recognition. IEEE Transactions on Affective Computing pp. 1–16 (2024). https://doi.org/10.1109/TAFFC.2024.3485057
11. Jin, M., Du, C., He, H., Cai, T., Li, J.: Pgcn: Pyramidal graph convolutional network for eeg emotion recognition. IEEE Transactions on Multimedia (2024)
12. Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: Deap: A database for emotion analysis; using physiological signals. IEEE transactions on affective computing **3**(1), 18–31 (2011)
13. Li, J., Qiu, S., Du, C., Wang, Y., He, H.: Domain adaptation for eeg emotion recognition based on latent representation similarity. IEEE Transactions on Cognitive and Developmental Systems **12**(2), 344–353 (2019)
14. Li, X., Song, D., Zhang, P., Yu, G., Hou, Y., Hu, B.: Emotion recognition from multi-channel eeg data through convolutional recurrent neural network. In: 2016 IEEE international conference on bioinformatics and biomedicine (BIBM). pp. 352–359. IEEE (2016)
15. 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)
16. Liu, W., Qiu, J.L., Zheng, W.L., Lu, B.L.: Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition. IEEE Transactions on Cognitive and Developmental Systems (2021)
17. Liu, Y., Sourina, O., Nguyen, M.K.: Real-time eeg-based human emotion recognition and visualization. In: 2010 international conference on cyberworlds. pp. 262–269. IEEE (2010)
18. Phan, T.D.T., Kim, S.H., Yang, H.J., Lee, G.S.: Eeg-based emotion recognition by convolutional neural network with multi-scale kernels. Sensors **21**(15), 5092 (2021)
19. 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)
20. Song, Y., Jia, X., Yang, L., Xie, L.: Transformer-based spatial-temporal feature learning for eeg decoding. arXiv preprint arXiv:2106.11170 (2021)
21. Song, Y., Zheng, Q., Liu, B., Gao, X.: Eeg conformer: Convolutional transformer for eeg decoding and visualization. IEEE Transactions on Neural Systems and Rehabilitation Engineering **31**, 710–719 (2022)
22. Suykens, J.A., Vandewalle, J.: Least squares support vector machine classifiers. Neural processing letters **9**, 293–300 (1999)
23. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. In: Advances in Neural Information Processing Systems (NeurIPS). vol. 30, pp. 5998–6008. Curran Associates, Inc. (2017)
24. Wang, Z., Wang, Y., Hu, C., Yin, Z., Song, Y.: Transformers for eeg-based emotion recognition: A hierarchical spatial information learning model. IEEE Sensors Journal **22**(5), 4359–4368 (2022)

25. Yang, Y., Wu, Q., Qiu, M., Wang, Y., Chen, X.: Emotion recognition from multi-channel eeg through parallel convolutional recurrent neural network. In: 2018 international joint conference on neural networks (IJCNN). pp. 1–7. IEEE (2018)
26. Zheng, W., Liu, W., Lu, Y., Lu, B., Cichocki, A.: Emotionmeter: A multimodal framework for recognizing human emotions. IEEE Transactions on Cybernetics pp. 1–13 (2018). https://doi.org/10.1109/TCYB.2018.2797176
27. 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). https://doi.org/10.1109/TAMD.2015.2431497
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). https://doi.org/10.1109/TAMD.2015.2431497
29. Zheng, W.L., Zhu, J.Y., Lu, B.L.: Identifying stable patterns over time for emotion recognition from eeg. IEEE transactions on affective computing **10**(3), 417–429 (2017)
30. Zhong, P., Wang, D., Miao, C.: Eeg-based emotion recognition using regularized graph neural networks. IEEE Transactions on Affective Computing **13**(3), 1290–1301 (2020)