Frequency-Aware Neural Architecture Search with Bidirectional Mamba for EEG Emotion Recognition

Shasha Xu, Rongye Zhang, Yanrong Hao, Jing Bian,
Rui Cao*, and Xin Wen*[0000-0003-2363-1190]

Taiyuan University of Technology, Taiyuan, China
xwen@tyut.edu.cn

Abstract. Emotion recognition using electroencephalogram (EEG) remains challenging in cross-subject settings because EEG patterns vary across subjects and frequency bands. Most existing methods rely on manually designed backbones, which may not fit the temporal characteristics of all EEG bands. This paper proposes a frequency-aware neural architecture search framework with cross-band fusion for EEG emotion recognition. For each frequency band, we build a searchable directed acyclic graph backbone and optimize its architecture independently. Besides topology search, the temporal configuration inside each node is included in the search space, enabling joint optimization of graph topology and node-level temporal modeling. Each node uses a Bidirectional Mamba module to model past and future context in offline EEG sequences. Cross-subject experiments on DEAP and DREAMER show competitive and stable performance, supporting the use of frequency-aware architecture search for EEG emotion recognition. Code is available at https://github.com/yoghurt125/FANS-BiMamba-Emotion.

Keywords: EEG emotion recognition · Neural architecture search · Frequency-aware modeling · State space models · Cross-subject classification· Bidirectional Mamba · Cross-band fusion

1 Introduction

Emotion recognition plays an important role in brain–computer interfaces, human–computer interaction, and mental health assessment [1]. Compared with behavioral modalities such as facial expressions and speech, electroencephalogram (EEG) signals provide a more direct reflection of neural activity, which has motivated their wide adoption in affective computing [2]. Nevertheless, strong noise, non-stationarity, and pronounced inter subject variability make cross-subject EEG emotion recognition particularly challenging [3].

Recently, deep learning has greatly advanced EEG-based emotion recognition. Transformer models use self-attention to capture long-range temporal dependencies in EEG signals [4–6], while graph neural networks model spatial relations between EEG channels based on electrode topology [7, 8]. In parallel, state space models (SSMs) such as Mamba show strong potential for efficient long-sequence modeling with linear complexity [9, 10]. However, these architectures are still mainly hand-designed, limiting their ability to jointly adapt to the diverse temporal scales and information complexity of EEG signals across frequency bands.

Neural Architecture Search (NAS) provides a promising paradigm for automated network design and has achieved notable success in computer vision and speech processing [11–14]. Several recent studies have explored NAS for EEG emotion recognition, aiming to reduce manual architectural design and improve cross-subject generalization [15, 16]. However, existing NAS methods still exhibit three key limitations. First, existing search spaces predominantly focus on convolutional or simple temporal operators and rarely incorporate efficient long sequence modeling modules. Second, characteristics specific to individual frequency bands are frequently overlooked, because different bands are usually implemented using a common architecture. Third, most directed acyclic graph (DAG) search frameworks keep the functionality of each node fixed and only vary edge connections or operation types, which limits the ability to jointly optimize both the network topology and the node functionalities.

To overcome these limitations, we propose a frequency-aware NAS framework for cross-subject EEG emotion recognition. First, to improve temporal modeling in EEG NAS, we include Bidirectional Mamba as a searchable temporal node, enabling effective characterization of long range dependencies in EEG signals. Second, to account for diverse temporal properties across EEG frequency bands, architecture search is performed separately for each band, enabling every band to learn a tailored network configuration. Third, to improve the flexibility of DAG based search, we further include node internal configurations in the search space, enabling joint optimization of topology and node level temporal modeling.

2 Proposed methods

2.1 Framework Overview

This work proposes a frequency-aware NAS framework for cross-subject EEG emotion recognition using Bidirectional Mamba, as illustrated in Fig. 1. Raw EEG signals first undergo frequency band decomposition and Differential Entropy (DE) feature extraction, followed by spatiotemporal feature modeling via single band searchable backbone networks. A reinforcement learning driven NAS strategy is leveraged to adaptively optimize the network structure for each frequency band. Finally, a cross-band attention fusion module aggregates effective information from multiple bands to output emotion classification results.

2.2 Single-Band Searchable Backbone Network

For a given frequency band k,the input feature representation is denoted as $\mathbf{X}_k \in \mathbb{R}^{B\times C\times T}$,where B is the batch size,C is the number of channels, and T is

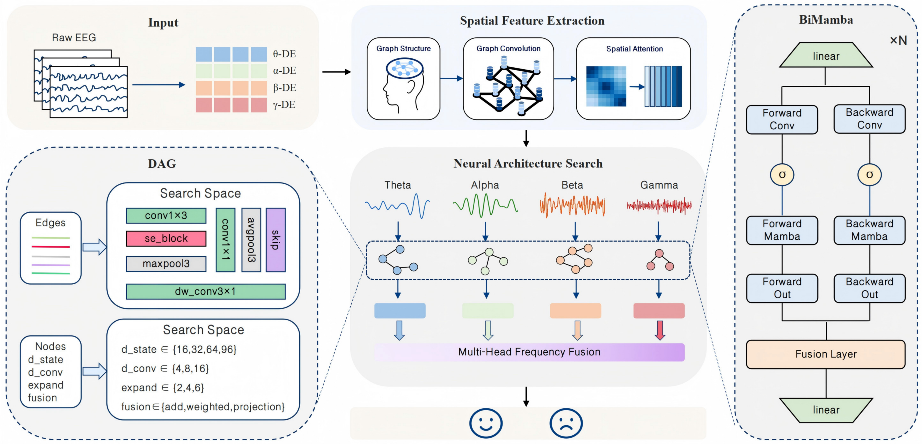

Fig. 1. Overview of the proposed frequency-aware NAS framework with bidirectional Mamba for EEG emotion recognition. After inputting raw EEG signals, the signals are first decomposed into four frequency bands (θ, α, β, γ) with DE features extracted. Each band is modeled by a searchable backbone with spatial feature enhancement and Bidirectional Mamba-DAG temporal modeling. Separate architecture search is performed for different bands, and the resulting band representations are aggregated by cross-band attention fusion for final emotion classification.

the number of time steps. The corresponding single-band backbone network is formulated as

$$\mathbf{F}_k = f_{\theta_k}(\mathbf{X}_k), \quad k \in \{\theta, \alpha, \beta, \gamma\} \tag{1}$$

where $f_{\theta_k}(\cdot)$ represents the architecture selected by neural architecture search, and θ_k includes both architectural configurations and network parameters. Each single band backbone consists of a spatial feature enhancement module followed by a Bidirectional Mamba-DAG based temporal modeling module.

Spatial Feature Enhancement To model spatial topological relationships among EEG electrodes, an EEG graph is constructed based on 2D electrode coordinates from the international 10–20 system. A graph convolutional network (GCN) is applied to encode channel features at each time step. Through multiple graph convolutional layers, local and high-order neighborhood information is aggregated to capture emotion related spatial patterns. It is followed by a spatial attention mechanism to adaptively reweight channel features. The resulting representation is refined through lightweight convolution and residual connections before temporal modeling.

Temporal Modeling via Bidirectional Mamba-DAG Along the temporal dimension, the spatially enhanced feature sequence is modeled by a DAG composed of N nodes. Each node receives inputs exclusively from its predecessor nodes and is computed sequentially according to the topological order. The output of the final node is taken as the high level temporal representation for the corresponding frequency band. The output of the i node is defined as

$$\mathbf{h}_i = \sum_{j<i} \alpha_{j,i}\, o_{j,i}(\mathbf{h}_j) \tag{2}$$

where $\alpha_{j,i}$ denotes the connection weight determined by neural architecture search, and $o_{j,i}(\cdot)$ represents an edge operation selected from a predefined operation set $\mathcal{O}$, including one-dimensional convolution, pooling, squeeze and excitation, skip connection, and none.

Unlike conventional DAG-NAS methods that restrict the search to edge operations, the proposed approach explicitly incorporates Bidirectional Mamba configurations of node internals into the search space, including state dimensions, convolutional receptive fields, expansion ratios, and fusion strategies. Since emotion recognition is formulated as an offline sequence classification task, the proposed design enables the simultaneous modeling of historical dependencies and future contextual information within EEG signals, as well as the synergistic optimization of node level functionalities and the global network topology.

2.3 Cross Band Attention Fusion

For the four frequency bands, the outputs of the corresponding single band backbone networks are denoted as $\{\mathbf{F}_k\}_{k=1}^{4}$. A Multi-Head Frequency Fusion module is used to capture dependencies across bands. This module applies adaptive gating and attention mechanisms to explicitly model inter band dependencies, followed by residual connections and temporal attention pooling to produce the final representation for classification.

2.4 Reinforcement Learning based Architecture Search

To automatically search architectures within the proposed space, a reinforcement learning based NAS strategy is adopted, where an LSTM controller sequentially samples architecture configurations that satisfy DAG constraints. The controller optimizes the number of nodes, edge connectivity, and internal Mamba configurations of the nodes.

During the search process, the controller independently samples candidate architectures for each frequency band. After training the sampled architecture on the training set, its classification performance is evaluated on the validation set and used as the reward signal to update the controller parameters. The optimization objective of the controller is defined as:

$$\mathcal{L}_{\text{ctrl}} = -\mathbb{E}_{a\sim\pi_\phi}\left[(R-b)\log \pi_\phi(a)\right] \tag{3}$$

where π_ϕ denotes the controller policy, R represents the validation performance metric, and b is a moving average baseline introduced to reduce the variance of gradient estimation. Through this process, an optimal architecture tailored to the temporal characteristics of each frequency band is obtained.

3 Experiments

3.1 Datasets and Experimental Settings

Datasets We evaluate the proposed method on two publicly available used EEG emotion recognition benchmarks: DEAP and DREAMER. DEAP contains EEG recordings from 32 subjects, collected at 128 Hz using 32 channels [17], while DREAMER includes data from 23 subjects recorded with a 14-channel Emotiv EPOC device at the same sampling rate [18]. In both datasets, subjects provided self-assessed valence and arousal ratings following standard protocols.

Experimental Settings To ensure cross-subject evaluation, leave-one-subject-out cross-validation (LOSO-CV) is adopted. EEG signals are segmented into non-overlapping 1-second windows, and Differential Entropy (DE) features are extracted from four frequency bands, including theta, alpha, beta, and gamma. Baseline correction is applied using pretrial signals to reduce subject related variability.

Architecture search and final evaluation are conducted in two stages. During search, candidate architectures are sampled separately for each band using a fixed subject level training and validation split. For each search run, 50 candidate architectures were evaluated for each frequency band, yielding 200 candidates in total. Each candidate was trained for up to 80 epochs with early stopping. The selected band architectures are then fixed for final LOSO-CV, where model weights are reinitialized and trained from scratch in each fold, with the left-out subject used only for testing. This design avoids fold-specific architecture tuning while keeping the search procedure computationally feasible and reproducible. Across multiple random seeds, the complete search procedure required 2.79 ± 0.36 GPU hours. The resulting models contained 1.80 ± 0.28 M parameters and required 0.94 ± 0.18 ms per sample for inference.

3.2 Performance of the proposed model

To evaluate the proposed framework, we compare it with representative Transformer based and Mamba based EEG emotion recognition methods. All compared methods are evaluated under LOSO-CV. For baselines with publicly available code, we reproduced their results when possible. For methods without public code or with unstable reproduction, we adopted the results reported in the original papers to avoid underestimating competing methods.

Performance on DEAP Dataset As shown in Table 1, the proposed method achieves accuracies of 71.41% and 72.50% for valence and arousal classification, respectively. These results are competitive among the listed Transformer based and Mamba based methods. Compared with the strongest Mamba based baseline in the table, the proposed method obtains higher accuracies on both emotion dimensions, suggesting that frequency specific architecture search and node level temporal configuration contribute to cross-subject EEG emotion recognition.

Table 1. Average accuracy (%) and standard deviation on the DEAP dataset

Model Category	Model Name	Year	Accuracy/Std	
			Valence	Arousal
Transformer	BiCCT [19]	2023	67.42/−	67.81/−
	SECT-DE [20]	2023	65.71/9.41	63.28/10.17
	ERTNet [21]	2024	59.60/−	63.90/−
	MACTN [22]	2024	66.14/6.11	67.83/8.07
	LFE-GFE [23]	2025	61.16/−	66.21/−
Mamba	HSF-BPM [24]	2025	68.64/6.05	69.46/5.82
	HSF-BPM(DCN) [24]	2025	69.29/5.86	70.37/5.11
Ours			**71.41/5.96**	**72.50/7.31**

Performance on DREAMER Dataset The proposed method achieves comparable results on both valence and arousal, with accuracies of 70.76% and 70.35%, respectively, as reported in Table 2. Although CMHFE-DAN reports a stronger result on arousal, its advantage is not observed on valence. In contrast, the proposed method obtains a favorable result on valence while remaining competitive on arousal, suggesting a more balanced performance across the two emotion dimensions.

Table 2. Average accuracy (%) and standard deviation on the DREAMER dataset

Model Category	Model Name	Year	Accuracy/Std	
			Valence	Arousal
Transformer	CMHFE-DAN [25]	2025	64.73/−	**79.83**/−
	LFE-GFE [23]	2025	66.34/−	66.22/−
Other	GCD-JFSE [26]	2025	61.11/12.31	72.46/13.45
	DSSN [27]	2024	62.07/10.82	64.73/10.82
	BiSMSM [28]	2023	61.88/8.76	64.25/6.86
Ours			**70.76/6.16**	70.35/14.19

3.3 Analysis of Band Specific Searched Architectures

To analyze the effect of frequency specific search, we visualize the final DAG architectures obtained for the four EEG frequency bands. As shown in Fig. 2, the representative architectures on DEAP show clear differences across bands. Across repeated search runs, similar structural tendencies are observed, although the exact edge selections may vary. The alpha and beta bands tend to form deeper structures with richer connections, whereas the theta band presents a more compact topology and the gamma band shows an intermediate pattern. In each DAG architecture, nodes represent the selected Bidirectional Mamba configurations, and edges denote the chosen operations and feature propagation paths between nodes. These differences suggest that separate search helps adapt the temporal backbone to band specific characteristics, providing an architectural-level explanation for frequency-aware search.

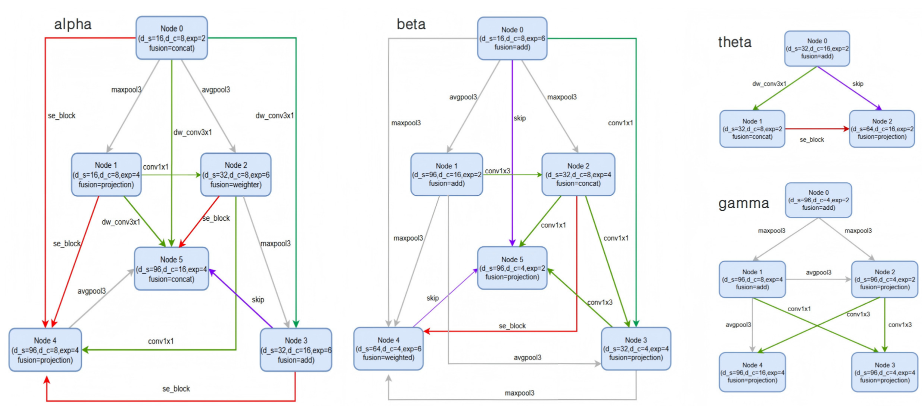

Fig. 2. Representative DAG architectures obtained by separate search for the four EEG frequency bands on the DEAP dataset. Nodes denote selected Bidirectional Mamba configurations, and edges denote selected operations and feature propagation paths.

3.4 Ablation experiments

Validation of NAS Effectiveness To evaluate the contribution of the proposed search strategy, we compare it with random search and fixed architecture variants under the same search space and training setting. Random search samples candidate architectures from the predefined space and selects the best performing one. The fixed architecture variants manually specify the number of DAG nodes while retaining the search of edge operations and node configurations. As shown in Table 3, random search achieves 70.54% accuracy, indicating that the search space contains effective architectures. The fixed variants obtain 69.84% and 70.08% accuracy, suggesting that manually constraining the DAG depth limits structural flexibility. The proposed NAS strategy achieves 71.41% accuracy, showing that controller guided search provides a more effective way to combine topology optimization and node level temporal configuration.

Ablation of Node Operation Types We further examine the influence of temporal operations by fixing the node type to Mamba, LSTM, Gated Recurrent Unit (GRU), Transformer, or Temporal Convolutional Network (TCN),

Table 3. Ablation results for NAS effectiveness

Method	Accuracy(%)
Random Search	70.54
Fixed Architecture - Simple (3 nodes)	69.84
Fixed Architecture - Medium (4 nodes)	70.08
Ours	**71.41**

while keeping the remaining setting unchanged. As shown in Table 4, TCN and LSTM achieve strong results, with accuracies of 71.33% and 71.09%, respectively, confirming the importance of temporal modeling in EEG emotion recognition. The proposed method achieves the highest accuracy. This result suggests that, when combined with searchable DAG topology and frequency specific modeling, Bidirectional Mamba provides stable bidirectional temporal modeling within an acceptable model complexity. Compared with standalone Mamba nodes, the proposed searchable design brings a 1.18 percentage point accuracy gain, supporting the benefit of jointly optimizing temporal node configurations and architecture topology.

Table 4. Ablation study of node operation types.

Node Type	Accuracy(%)
Mamba	70.23
LSTM	71.09
GRU	70.78
Transformer	70.47
TCN	71.33
Ours	**71.41**

4 Conclusion

This work presents a frequency-aware neural architecture search framework with cross-band fusion for cross-subject EEG emotion recognition. By jointly searching for the network topology and bidirectional Mamba-based temporal modeling configurations within a directed acyclic graph, the proposed method adapts to the heterogeneous temporal characteristics of EEG signals across different frequency bands. Experiments on DEAP and DREAMER show competitive performance under LOSO-CV. The architecture analysis and ablation studies further indicate that separate search helps different frequency bands form more suitable temporal backbones, while Bidirectional Mamba provides stable bidirectional temporal modeling within the searched DAG structure. Future work will explore more efficient search strategies, stricter nested evaluation protocols, and extensions to other multi band physiological signal tasks.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grants [62206196], in part by the Natural Science Foundation of Shanxi Province under Grant [202303021221001] and [202403021222057], in part by the Natural Science Foundation for Young Scientists of Shanxi Province under Grants [202203021212216].

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

References

1. Lai, T. Y., Ghazali, M., Hazlin Huspi, S., Othman, N. Z. S.: A Systematic Review on User Interactions in Emotion-Recognition Brain–Computer Interfaces. *Int. J. Hum.-Comput. Interact.* 1–18 (2025)
2. Lim, R. Y., Lew, W.-C. L., Ang, K. K.: Review of EEG Affective Recognition with a Neuroscience Perspective. *Brain Sci.* **14**(4), 364 (2024)
3. Chen, J., Cui, Y., Wei, C., Polat, K., Alenezi, F.: Advances in EEG-Based Emotion Recognition: Challenges, Methodologies, and Future Directions. *Appl. Soft Comput. J.* **180** (2025)
4. Cheng, Z., Bu, X., Wang, Q., Yang, T., Tu, J.: EEG-based Emotion Recognition Using Multi-Scale Dynamic CNN and Gated Transformer. *Scientific Reports* **14**(1) (2024)
5. Ding, Y., Tong, C., Zhang, S., Jiang, M., Li, Y., Lim, K. J. L., Guan, C.: EmT: A Novel Transformer for Generalized Cross-subject EEG Emotion Recognition. *IEEE Trans. Neural Netw. Learn. Syst.* **36**(6), 10381–10393 (2025)
6. Vafaei, E., Hosseini, M.: Transformers in EEG Analysis: A Review of Architectures and Applications in Motor Imagery, Seizure, and Emotion Classification. *Sensors* **25**(5) (2025)
7. Yang, X., Zhu, Z., Jiang, G., Wu, D., He, A., Wang, J.: DC-ASTGCN: EEG Emotion Recognition Based on Fusion Deep Convolutional and Adaptive Spatio-temporal Graph Convolutional Networks. *IEEE J. Biomed. Health Inform.* **29**(4), 2471–2483 (2025)
8. Pan, J., Liang, R., He, Z., Li, J., Liang, Y., Zhou, X., He, Y., Li, Y.: ST-SCGNN: A Spatio-Temporal Self-Constructing Graph Neural Network for Cross-Subject EEG-Based Emotion Recognition and Consciousness Detection. *IEEE J. Biomed. Health Inform.* **28**(2), 777–788 (2024)
9. Gu, A., Dao, T.: Mamba: Linear-Time Sequence Modeling with Selective State Spaces. arXiv preprint (2024)
10. Gui, Y., Chen, M., Su, Y., Luo, G., Yang, Y.: EEGMamba: Bidirectional State Space Model with Mixture of Experts for EEG Multi-task Classification. arXiv preprint arXiv:2407.20254 (2024)
11. Zoph, B., Le, Q. V.: Neural Architecture Search with Reinforcement Learning. In: Proceedings of the 5th International Conference on Learning Representations (ICLR), Toulon, France (2017)

12. Tan, M., Chen, B., Pang, R., Vasudevan, V., Sandler, M., Howard, A., Le, Q. V.: MnasNet: Platform-Aware Neural Architecture Search for Mobile. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, pp. 2815–2823 (2019)
13. Real, E., Aggarwal, A., Huang, Y., Le, Q. V.: Regularized Evolution for Image Classifier Architecture Search. In: Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA, pp. 4780–4789 (2019)
14. Liu, H., Simonyan, K., Yang, Y.: DARTS: Differentiable Architecture Search. In: Proceedings of the 7th International Conference on Learning Representations (ICLR), New Orleans, LA, USA (2019)
15. Li, C., Zhang, Z., Zhang, X., Huang, G., Liu, Y., Chen, X.: EEG-based Emotion Recognition Via Transformer Neural Architecture Search. *IEEE Trans. Ind. Informat.* **19**(4), 6016–6025 (2022)
16. Wu, Y., Liu, H., Zhang, D., Zhang, Y., Lou, T., Zheng, Q.: AutoER: Automatic EEG-based Emotion Recognition with Neural Architecture Search. *J. Neural Eng.* **20**(4) (2023)
17. 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 Trans. Affect. Comput.* **3**(1), 18–31 (2011)
18. Katsigiannis, S., Ramzan, N.: DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Low-cost Off-the-Shelf Devices. *IEEE J. Biomed. Health Inform.* **22**, 98–107 (2018)
19. Cao, G., Yang, L., Tang, C., Zhang, Q., Hu, H.: BiCCT: A Compact Convolutional Transformer for EEG Emotion Recognition. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (2023)
20. Bai, Z., Hou, F., Sun, K., Wu, Q., Zhu, M., Mao, Z., Song, Y., Gao, Q.: SECT: A Method of Shifted EEG Channel Transformer for Emotion Recognition. *IEEE J. Biomed. Health Inform.* **27**(10), 4758–4767 (2023)
21. Liu, R., Chao, Y., Ma, X., Sha, X., Sun, L., Li, S., Chang, S.: ERTNet: an Interpretable Transformer-Based Framework for EEG Emotion Recognition. *Front. Neurosci.* **18** (2024)
22. Si, X., Huang, D., Liang, Z., Sun, Y., Huang, H., Liu, Q., Yang, Z., Ming, D.: Temporal Aware Mixed Attention-based Convolution and Transformer Network for Cross-Subject EEG Emotion Recognition. *Comput. Biol. Med.* **181** (2024)
23. Xu, Z., Chen, N., Li, G., Li, J., Zhu, H., Zhu, Z.: The Mitigation of Heterogeneity in Temporal Scale among Different Cortical Regions for EEG Emotion Recognition. *Knowl.-Based Syst.* **309** (2025)
24. Li, X., Huang, J., Chen, J., Zhou, H.: From Local to Global: A Hierarchical Spatial-Frequency Bi-patch Mamba Network for EEG-based Emotion Recognition. *Biomed. Signal Process. Control* **113** (2026)
25. Hilali, M., Ezzati, A., Ben Alla, S.: CMHFE-DAN: A Transformer-Based Feature Extractor with Domain Adaptation for EEG-Based Emotion Recognition. *Information* **16**(7) (2025)
26. Luo, G., Han, Y., Xie, W., Tian, F., Zhu, L., Qian, K., Li, X., Sun, S., Hu, B.: GCD-JFSE: Graph-based Class-Domain Knowledge Joint Feature Selection and Ensemble Learning for EEG-based Emotion Recognition. *Knowl.-Based Syst.* **309** (2025)
27. Li, W., Dong, J., Liu, S., Fan, L., Wang, S.: Dynamic Stream Selection Network for Subject-Independent EEG-Based Emotion Recognition. *IEEE Sens. J.* **24**(12), 19336–19343 (2024)

28. Li, W., Tian, Y., Hou, B., Dong, J., Shao, S., Song, A.: A Bi-Stream Hybrid Model with MLPBlocks and Self-Attention Mechanism for EEG-based Emotion Recognition. *Biomed. Signal Process. Control* **86** (2023)