Title
DBNet: A Dual-Branch Network for Single-Trial Feedback EEG Decoding and Supporting Avoidance Coupling in MDD

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Tan, Congming
Wang, Wu
Pan, Ting
et al.

Publication Date
2026

DBNet: A Dual-Branch Network for Single-Trial Feedback EEG Decoding and Supporting Avoidance Coupling in MDD

Congming Tan[1], Wu Wang[2], Ting Pan[2], Hongli Liao[2], Haoyu Ning[1] & Yin Tian (tianyin@cqupt.edu.cn) [1,2,*]
[1]School of Computer Science and Technology, Chongqing University of Posts and Telecommunications
[2]School of Life Health Information Science and Engineering, Chongqing University of Posts and Telecommunications

Abstract

Major depressive disorder (MDD) involves heightened sensitivity to negative outcomes and altered avoidance learning. While deep learning has advanced EEG-based depression detection, most studies focus on resting-state signals rather than single-trial task feedback and its behavioral relevance. Using a public feedback-locked EEG dataset from a probabilistic learning paradigm, we decode correct versus incorrect feedback trials separately in MDD and healthy controls (HC) and propose DBNet, a dual-branch network combining time-domain modeling with learnable wavelet time-frequency representations. The time branch captures feedback-related ERP dynamics and trial-wise temporal variability, whereas the time-frequency branch learns task-relevant spectral components via adaptive wavelets. DBNet outperforms baselines, with higher decoding performance in MDD. Visualizations indicate that the model emphasizes FRN/P3 time windows and fronto-midline activity, and its learned wavelet spectra show a stronger theta-band emphasis in MDD, consistent with error-monitoring processes. Subject-level representations from DBNet further exhibit stronger coupling with test-phase NoGo accuracy in MDD.

Keywords: Major depressive disorder; dual-branch network; task feedback; avoidance learning.

I. Introduction

Major depressive disorder (MDD) is a highly prevalent and burdensome condition, affecting ~280 million people worldwide (Dong et al., 2025), and is clinically characterized by persistent low mood and anhedonia with broader cognitive and functional impairment (Tanzilli et al., 2021). Beyond affective symptoms, MDD is closely linked to biased processing of negative outcomes, heightened punishment sensitivity, and altered avoidance-related learning and decision-making (Kratochwil et al., 2025). Computational psychiatry further formalizes these abnormalities using reinforcement learning (RL) frameworks to quantify reward/punishment-driven value updating and behavioral adjustment (Cavanagh et al., 2011; Cavanagh et al., 2019), with meta-analytic evidence indicating punishment-related alterations in RL parameters across the depression/anxiety spectrum (Pike & Robinson, 2022) and reinforcement sensitivity work suggesting stronger associations of depression with punishment sensitivity and more depression-specific reductions in reward sensitivity (Katz et al., 2020). Given that these mechanisms depend on rapid outcome evaluation and error monitoring, EEG provides a high temporal-resolution approach for probing feedback-related neural dynamics in MDD (Liu et al., 2022; Pegg et al., 2021).

At the neural level, feedback-locked signatures such as FRN/RewP and frontal-midline theta are thought to index fast mPFC-based representations of outcomes relative to expectation and to support behavioral adjustment and avoidance learning (Burkhouse et al., 2016; Cavanagh et al., 2019; Csifcsák & Mittner, 2026); FRN typically emerges ~250-350ms post-feedback and is theta-dominant and punishment-sensitive (Becker et al., 2014; Puusepp et al., 2021; Zhang et al., 2024). Consistent with aberrant reward-punishment processing in MDD, larger FRN amplitudes and stronger error-related theta to negative feedback have been reported (Webb et al., 2017), although FRN effects show heterogeneity moderated by factors such as stress, sleep, and comorbidity (QIN et al., 2021), and machine-learning studies increasingly relate delta/theta features to individual differences to dissociate depression from anxiety (Grabowska et al., 2024). Importantly, enhanced post-punishment FRN/theta in MDD does not necessarily imply poorer performance and may, under some conditions, predict better avoidance (e.g., NoGo) behavior (Cavanagh et al., 2011).

Feedback-locked EEG provides high temporal resolution for studying reward-punishment processing and error monitoring but is limited by conventional methods, which often rely on condition-averaged ERPs or power spectra, and hand-crafted features that may not fully capture single-trial spatiotemporal-spectral information (Bachmann et al., 2018; Ding et al., 2020; Liao et al., 2017; Sharma et al., 2018). Additionally, manual feature engineering can alter or discard key cognitive-neural components, limiting interpretability. In contrast, deep learning enables end-to-end learning from raw EEG, preserving task-relevant information. Recent advances in deep EEG representation learning include compact convolutional architectures (Tan et al., 2024) and semi-supervised learning for small-sample or cross-task settings (Hu et al., 2025; Xiong et al., 2025). However, deep-learning applications to MDD have focused mostly on resting-state EEG classification, with fewer studies addressing task-evoked feedback dynamics. To bridge the gap between deep model cues and interpretable cognitive mechanisms, we propose DBNet, a dual-branch framework for single-trial feedback-locked EEG. DBNet captures error-monitoring and outcome-evaluation dynamics from two complementary perspectives: a time-domain branch for feedback-locked temporal patterns and a time-frequency branch for adaptive wavelet-based representation learning, focusing on discriminative sub-bands related to error/punishment

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

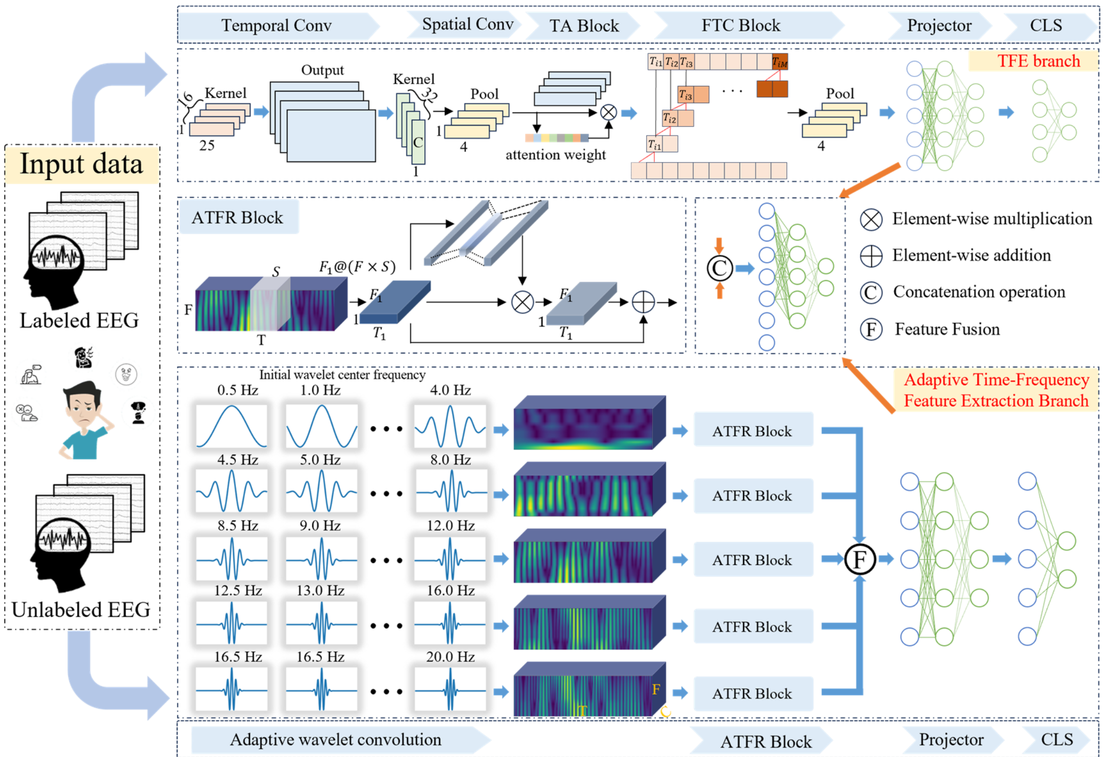

Figure 1: The overall framework of DBNet proposed in this study

processing. We use a binomial GLM to test the predictive association of DBNet's high-level features with NoGo accuracy and group differences, providing a statistical framework to evaluate the consistency of DBNet's cues with MDD-related feedback processing and avoidance learning.

In summary, our main contributions are as follows:

(1) We propose DBNet, a dual-branch deep network that integrates time-domain and learnable wavelet-based time-frequency modeling for decoding task-based feedback EEG. The time-domain branch captures ERP components and trial-to-trial variability, while the time-frequency branch uses continuous wavelet transforms to adaptively emphasize task-relevant frequency components. Both branches are fused to decode correct versus incorrect feedback trials.

(2) We evaluate DBNet on a task-based feedback EEG dataset with MDD and healthy control (HC) participants. Visualization analyses reveal that the model's temporal windows and frequency distributions align with cognitive-neuroscience evidence (e.g., error-monitoring windows and theta-band rhythms), enhancing neurophysiological interpretability.

(3) We incorporate subject-level features from DBNet into a binomial GLM to assess their association with NoGo accuracy and examine group differences. The results show stronger coupling between error-related neural signals and avoidance behavior in MDD, providing statistical evidence for heightened punishment sensitivity and avoidance-learning bias in depression, and offering a new approach for representation learning in task-based EEG.

II. Methodology

A. Overview of DBNet Model

Let a single-trial feedback-locked EEG segment be denoted as $X \in R^{B \times C \times L}$, where B is the batch size, C is the number of electrodes, and L is the number of time points. DBNet aims to learn discriminative neural representations from single-trial, task-based feedback EEG for distinguishing correct and incorrect feedback, while remaining robust to trial-to-trial variability (Fig. 1). The network consists of two branches: the time-domain branch directly extracts informative cues from raw waveforms, capturing subtle temporal and amplitude distinctions between correct and incorrect feedback; the adaptive wavelet time-frequency branch uses learnable continuous wavelet transforms to focus on task-relevant frequencies and extract stable discriminative features. These high-level features from both branches are fused for binary prediction, with auxiliary outputs retained to enhance supervision and training stability. DBNet effectively decodes single-trial feedback and links learned representations to behavioral measures (e.g., NoGo accuracy).

B. Temporal feature extraction branch

The time feature extraction branch (TFE Branch) is designed to extract ERP-related temporal morphological cues and their trial-to-trial temporal variability from single-trial feedback-locked EEG.

(1) **Spatiotemporal convolution**. To capture feedback fluctuations within short time windows while learning stable spatial filters, the TFE Branch first applies temporal convolution on each electrode to extract local temporal patterns, and then uses a depthwise spatial convolution with kernel size $(C,1)$ to perform channel-wise spatial projection of the temporally filtered responses. This design substantially reduces the number of trainable parameters while learning discriminative electrode combinations.

(2) **Temporal attention mechanism.** To emphasize informative periods during feedback processing, we introduce a temporal attention block (TA block) in the time branch. Let the sequence features after spatiotemporal convolution be $\{h_t\}_{t=1}^{T_1}$, where T_1 is the number of time steps after downsampling. The TA block first computes a scalar score s_t for each time step and then applies a softmax normalization to obtain the attention distribution:

$$a_t = \frac{exp(s_t)}{\sum_{j=1}^{T_1} exp(s_j)}, a = (a_1, \ldots, a_{T_1}) \tag{1}$$

The attention weights are then applied to the sequence features for weighted aggregation, yielding a global temporal representation:

$$V = \sum_{t=1}^{T_1} a_t h_t \in R^{C \times F \times T_1} \tag{2}$$

(3) **Fast temporal convolution block**. To model long-range temporal dependencies, we design a Fast Temporal Convolution Block (FTC Block) with non-sliding kernels. Each kernel models the relationship between adjacent time points, recursively composing these local relations to capture long-term temporal structure efficiently. This design reduces convolution operations and enhances efficiency.

The FTC Block first projects the feature dimension to F' via a linear layer:

$$V^{'} = \left(V^T \beta\right)^T \tag{3}$$

where $V^{'} \in R^{C \times F^{'} \times T_1}$, $V^T \in R^{C \times F \times T_1}$ denotes the transpose of V along the feature dimension, and $\beta \in R^{F \times F^{'}}$ is the learnable weight matrix of the linear layer. The FTC Block then defines a kernel set $f = \{f_1, \ldots, f_{T_1}\}$ and applies zero padding to keep the temporal dimension unchanged, enabling stable aggregation of adjacent time-point relations during the recursive composition process.

After obtaining the output of the FTC Block, we use a projector composed of fully connected layers to map it into a fixed-dimensional high-level representation space, yielding the time-branch feature vector $Z^{(t)}$.

C. Adaptive Wavelet Time-Frequency Feature Extraction Branch.

The time-frequency branch employs learnable Morlet wavelet convolutions to adaptively select task-relevant frequencies in an end-to-end manner, and integrates convolutional blocks with channel attention to extract discriminative time-frequency representations.

(1) **Learnable Morlet wavelet convolution.** We construct a Morlet filter bank. Each filter is parameterized by a learnable scale parameter a: we first compute the corresponding scale s, then generate a normalized Morlet kernel, and apply adaptive wavelet convolution to each EEG channel to obtain the sub-band time-frequency response map $H_{tf} \in R^{B \times C \times F \times T}$:

$$s = \frac{f_s \omega_0}{a \cdot 2\pi}, \psi(x) = \pi^{-1/4} \cos(\omega_0 x) exp(-\frac{x^2}{2}) \tag{4}$$

$$g(t) = \sqrt{\frac{1}{s}} \psi(\frac{t}{s}), H_{tf} = X * g \tag{5}$$

where ω_0 is the center-frequency constant of the mother wavelet.

(2) **Multi-subband wavelet decomposition and adaptive time-frequency feature refinement**. Given that feedback-related EEG activity in task is predominantly concentrated in the low-frequency range, we restrict the frequency-of-interest to 0-20 Hz and partition it into five equally spaced, adjacent sub-bands $F = \{[0, f_1], [f_1, f_2], \ldots, [f_4, 20]\}$.

For the b-th sub-band $[f_{b-1}, f_b]$, we first apply an independent wavelet transform to the input single-trial EEG X, yielding the corresponding time-frequency response:

$$H_{tf}^b = g_{[f_{b-1}, f_b]} * X \in R^{B \times C \times F_b \times L} \tag{6}$$

where $g_{[f_{b-1}, f_b]}$ denotes the learnable Morlet wavelet transform, and $F_b = 2 \cdot (f_b - f_{b-1})$ is the number of filters.

We then feed the time-frequency representation H_{tf}^b of each sub-band into an adaptive time-frequency refinement block $\Phi(\cdot)$ (ATFR), which consists of a convolutional layer followed by squeeze-and-excitation (SE) operations. This block enhances discriminative patterns within the sub-band while suppressing redundant responses, and can be formulated as:

$$V_{tf}^b = \Phi(H_{tf}^b) = SE(Conv(H_{tf}^b)) \tag{7}$$

where $Conv(\cdot)$ performs within-subband aggregation across frequencies and $SE(\cdot)$ is a channel-attention module that adaptively reweights channels based on global statistics to emphasize task-relevant time-frequency responses. The five sub-bands produce a set of refined time-frequency features $\{V_{tf}^b\}_{b=1}^5$, which are then mapped by a projector to a fixed-dimensional high-level representation space $Z^{(tf)}$.

Finally, the high-level representations from the time and time-frequency branches, $Z^{(t)}$ and $Z^{(tf)}$, are concatenated to form $Z^{(all)} = [Z^{(t)}; Z^{(tf)}]$, which is then fed into a classification head to discriminate task-related correct versus incorrect feedback.

D. Learning Strategy

We apply auxiliary classification supervision to both time and time-frequency branches to ensure each learns task-relevant features. A cross-view consistency constraint aligns the predictive distributions of both branches. Additionally, supervised contrastive learning pulls same-class samples closer and pushes different-class samples apart, while pseudo-label filtering on high-confidence, consistent samples enhances discriminability and robustness.

(1) **Classification loss**. We apply classification supervision to the main (fusion), time, and time-frequency branches to ensure they all learn discriminative features. A weighted cross-entropy loss balances their contributions. Let $z_c^{(all)}$, $z_c^{(t)}$ and $z_c^{(tf)}$ be the logits from each branch, and y the ground-truth label. The total multi-head loss is defined as:

$$\mathcal{L}_{CE} = \mathcal{L}_{CE}(z_c^{(all)}, y) + \frac{1}{2}\mathcal{L}_{CE}(z_c^{(t)}, y) + \frac{1}{2}\mathcal{L}_{CE}(z_c^{(tf)}, y) \quad (8)$$

(2) **Cross-view consistency constraint.** We align the predictive distributions of the time and time-frequency branches using a cross-view consistency constraint, measured by the Jensen-Shannon divergence (JSD):

$$\mathcal{L}_{JS}(p,q) = \frac{1}{2}KL(p \| m) + \frac{1}{2}KL(q \| m), m = \frac{1}{2}(p+q) \quad (9)$$

where p and q are the predicted distributions from each branch, and the divergence is computed using Kullback-Leibler divergence.

(3) **Contrastive learning.** Supervised contrastive learning (SupCon) improves discriminability by pulling same-class samples together and pushing different-class samples apart. We treat features from the time and time-frequency branches as two views of the same trial, normalize them, and compute their similarity:

$$\mathcal{L}_{SupCon} = -\sum_i \frac{1}{|P(i)|}\sum_{p\in P(i)} \log \frac{exp(\sin(u_i, u_p)/\tau)}{\sum_{a\in A(i)} exp(\sin(u_i, u_a)/\tau)} \quad (10)$$

where $\mathrm{sim}(\cdot,\cdot)$ denotes dot-product similarity; u_i and u_p represent the embedding-space representations; $A(i)$ and $P(i)$ denote the sets of positive and all samples.

(4) **Pseudo-label gating.** To improve robustness, we apply pseudo-label gating to unlabeled samples. A sample is included in training if (i) both branches predict the same class, and (ii) the maximum prediction confidence is above 0.8. These pseudo-labels are then incorporated into supervised contrastive learning for refinement.

Combining the above learning objectives, the overall training loss is defined as:

$$\mathcal{L} = \mathcal{L}_{CE} + \lambda_{JS}\mathcal{L}_{JS} + \lambda_{SC}\mathcal{L}_{SupCon} \quad (11)$$

where $\lambda_{JS} = \lambda_{SC} = 0.1$ are weighting coefficients.

E. Group Differences in Coupling with Avoidance Behavior

During the training phase (with feedback), feedback-locked EEG trials are input into DBNet, which discriminates between Correct and Incorrect feedback. The final layer outputs two-class logits, denoted as $z_{c,s,t}^{(all)} \in R^2$, where s indexes subjects and t indexes trials. For each subject s, we group trials by predicted class $\hat{y}_{s,t}$ and extract the corresponding logit $l_{s,t}$. We then average the logits within each subject for the Incorrect and Correct feedback categories to obtain subject-level representations, $l_{\bar{s},E}$ and $l_{\bar{s},C}$. The incorrect relative to correct enhancement is defined as $\Delta_s = l_{\bar{s},E} - l_{\bar{s},C}$, and is z-standardized across subjects for comparability.

To examine the relationship between Δ_s and behavioral performance, we use test-phase NoGo accuracy as the dependent variable and assess whether the coupling is moderated by group (MDD vs HC) using a binomial GLM with an interaction term. The model is specified as:

$$logit(p_s) = \beta_0 + \beta_g group_s + \beta_\Delta \Delta_s^z + \beta_{int}\left[group_s \cdot \Delta_s^z\right] \quad (12)$$

where $group_s$ is 0 for HC and 1 for MDD. The coefficient β_{int} captures the between-group difference in coupling, and the interaction effect is evaluated using a likelihood-ratio test (LRT). This procedure quantifies the coupling between error-related neural representations and avoidance behavior, and examines whether this coupling differs between MDD and HC groups.

III. Results

A. Datasets and Preprocessing

We used a public University of Arizona EEG dataset (Cavanagh et al., 2011) including 21 participants with major depressive disorder (MDD; BDI > 13) and 24 healthy controls (HC; BDI < 7); to balance groups, 21 participants were randomly selected from each group. Participants completed a probabilistic learning task with a feedback-based training phase and a feedback-free test phase, designed to link error-related feedback signals to avoidance learning (NoGo performance in test phase); we performed within-group binary decoding of correct versus incorrect feedback to characterize group differences. EEG was preprocessed in EEGLAB using a 0.5-30 Hz band-pass filter, re-referencing to bilateral mastoids, ICA-based artifact removal, downsampling to 128 Hz, and baseline correction (-300 to -200 ms) with epochs retained from -200 to 1000 ms.

B. Baselines and Experimental Setup

We compare the proposed method with widely used deep learning baselines for EEG classification, including models originally developed for ERP decoding, BCI applications, and motor imagery. The baselines are ShallowNet/DeepNet (Schirrmeister et al., 2017), EEGNet (Lawhern et al., 2018), EEG-Inception (Santamaría-Vázquez et al., 2020), JMNet (Kim et al., 2024), EEGNex (Chen et al., 2024), and SJNet (Jalilpour & Müller-Putz, 2025). These architectures have been applied to P300, movement-related cortical potentials,

Table 1: Performance comparison between the proposed method and baseline models on the MDD.

Model	ACC	Kappa	MS	MF1	MP
ShallowNet	72.08±5.83	31.73±8.83	64.70±4.45	65.22±4.88	69.09±4.96
DeepNet	74.84±4.98	44.35±7.65	72.62±3.86	72.06±3.83	72.17±4.06
EEGNet	75.00±5.96	43.34±8.63	71.33±4.11	71.62±4.29	72.29±4.99
EEG-inception	74.86±5.81	43.45±10.54	71.87±5.56	71.66±5.29	71.86±5.21
EEGNeX	76.09±5.86	44.81±10.81	71.89±5.43	72.33±5.43	73.25±5.53
JMNet	74.44±5.10	36.38±9.79	66.86±5.21	67.54±5.41	73.61±7.56
SJNet	76.61±6.93	44.66±12.75	71.17±6.24	72.09±6.52	75.38±8.13
DBNet (ours)	79.87±7.23	53.96±8.41	76.22±3.72	76.84±3.77	78.67±5.41

Table 2: Performance comparison between the proposed method and baseline models on the HC.

Model	ACC	Kappa	MS	MF1	MP
ShallowNet	70.39±6.18	29.38±8.14	63.84±4.28	64.12±4.48	67.46±5.11
DeepNet	73.27±6.58	43.28±9.12	71.66±4.58	71.61±5.50	71.75±5.67
EEGNet	73.35±5.43	44.72±8.07	72.20±4.14	72.31±4.05	72.70±4.06
EEG-inception	74.65±5.73	44.50±9.17	72.73±5.09	72.17±4.63	72.07±4.24
EEGNeX	75.93±6.31	45.89±9.23	72.48±4.48	72.89±4.62	73.71±5.01
JMNet	73.58±5.52	36.75±6.38	67.09±3.32	67.90±3.44	72.29±6.10
SJNet	75.88±6.34	43.01±8.83	70.03±4.05	71.15±4.60	75.24±6.71
DBNet (ours)	77.51±5.28	45.79±9.13	71.26±5.06	72.47±4.41	77.28±6.13

sensorimotor rhythms, and other BCI paradigms. Performance was evaluated using subject-independent 10-fold cross-validation and reported in terms of Accuracy (ACC), Cohen's kappa (Kappa), Macro-Sensitivity (MS), Macro-F1 (MF1), and Macro-Precision (MP).

C. Feedback Detection

In the MDD group (Table 1), DBNet achieved the best performance across all evaluation metrics, with an ACC of 79.87%, Kappa of 53.96%, MS of 76.22%, MF1 of 76.84%, and MP of 78.67%. Compared with other baselines, the improvements over the strongest competitors were notable. For example, relative to SJNet (ACC: 76.61%), DBNet increased accuracy by approximately 3.26%; relative to EEGNeX (Kappa: 44.81%, MF1: 72.33%), DBNet improved Kappa by about 9.15% and MF1 by about 4.51%. Overall, DBNet not only improved accuracy but also showed more pronounced gains in Kappa and MF1, indicating enhanced agreement with the ground-truth labels and a more balanced discriminative capability across the two feedback classes.

In the HC group (Table 2), DBNet also showed stable and favorable performance. It achieved the highest ACC (77.51%) and MP (77.28%) among all methods, while maintaining relatively high Kappa (45.79%) and MF1 (72.47%), suggesting that its agreement and overall performance did not markedly deteriorate across groups. Overall, DBNet's advantage is primarily attributable to the complementary dual-branch design integrating time-domain and adaptive wavelet time–frequency modeling: the time branch is effective at capturing post-feedback ERP temporal-structure differences and trial-wise temporal variability, whereas the time–frequency branch can adaptively emphasize task-relevant spectral components through end-to-end learning. Their fusion enhances discriminability and robustness for correct versus incorrect feedback decoding, leading to improved results in both the MDD and HC groups.

D. Visualization

To enhance model interpretability, we visualized key representations, as shown in Fig. 2. In Fig. 2(a), the feedback-locked ERP waveforms reveal clear morphological differences between correct and incorrect feedback, particularly within the FRN (200-350 ms) and P3 (300-500 ms) windows, with more pronounced fluctuations in the Incorrect condition. This aligns with classical models of feedback processing, where rapid outcome evaluation and error monitoring are expressed in these time windows. Fig. 2(b) shows that the model's temporal attention focuses on the 200-350 ms feedback interval, closely matching the ERP differences in Fig. 2(a). Figs. 2(c)-(d) highlight the model's spatial emphasis on fronto-midline regions (e.g., FCz), particularly under Incorrect feedback, and a more prominent fronto-midline focus in the MDD group during error feedback. These visualizations demonstrate that DBNet primarily utilizes FRN/P3-related temporal information and fronto-midline spatial patterns for error monitoring, supporting the interpretability of its learned features. Fig. 2 (e)-(f) shows the average spectral responses learned by the five adaptive wavelet branches in the MDD and HC groups. The MDD group exhibits a sharper peak in the theta-related branch, concentrating discriminative spectral emphasis in the theta range. In contrast, the HC group shows a more balanced distribution of peaks across frequency bands. These results suggest that DBNet adaptively emphasizes task-relevant frequencies, with the most informative discriminative cues in MDD concentrated in the theta-band dynamics.

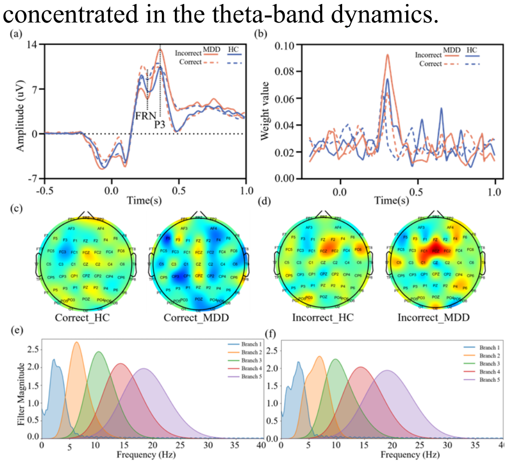

Figure 2: Visualization results of DBNet.

E. Coupling between avoidance behavior across groups

To directly link feedback-related discriminative representations learned by the DBNet model to test-phase avoidance behavior (NoGo accuracy), we used the subject-level Δ (the logit enhancement for Incorrect relative to Correct feedback) as the explanatory variable and fit a binomial GLM to assess its coupling with the probability of NoGo success as well as group differences. As shown in Fig. 3, the regression slope of Δ in the HC group was close to zero ($\beta_{\mathrm{HC}} = 0.013$), indicating a weak association between this neural representation and NoGo performance. In contrast, Δ showed a stronger positive relationship with NoGo success probability in the MDD group ($\beta_{\mathrm{MDD}} = 0.346$). Importantly, the group-by-Δ interaction was positive and statistically significant ($\beta_{\mathrm{int}} = 0.333$, $p = 0.0052$), demonstrating that the coupling between enhanced error-related feedback signals and avoidance behavior (NoGo) was significantly stronger in MDD than in HC.

IV. Discussion

Across Tables 1-2, correct/incorrect feedback decoding was generally easier in MDD than in HC, suggesting more stable or stronger discriminative feedback-locked patterns in depression, which is consistent with DBNet's dual-branch design (Fig. 1): the time branch captures ERP dynamics with long temporal convolutions and focuses on informative windows via attention, while the learnable wavelet time–frequency branch adaptively selects task-relevant spectral components and complements phase-locked ERP cues with induced activity. Visualization further supports the cognitive-neurophysiological meaning of the learned cues: raw ERPs show clear correct/incorrect differences within FRN/P3 windows (Fig. 2(a)), consistent with prior findings in this dataset (Cavanagh et al., 2011) that also relate the peak–trough interval to a 5-7 Hz rhythm; temporal attention aligns with these critical periods (Fig. 2(b)); spatial patterns emphasize fronto-midline/central regions (Fig. 2(c)) in line with frontal-midline theta as a marker of cognitive control and conflict/error monitoring(Eisma et al., 2021); and learned wavelet spectra exhibit a stronger theta emphasis in MDD but a more balanced distribution in HC (Fig. 2(e)-(f)), supporting adaptive task-relevant band selection by learnable wavelets (Kim et al., 2024). Importantly, GLM results (Fig. 3) show that subject-level representations from DBNet have behavioral relevance: the group × Δ interaction is significant (p=0.0052), with stronger positive coupling between error-related representations and test-phase NoGo accuracy in MDD, consistent with computational psychiatry evidence that mood/anxiety disorders tend to show heightened punishment-related learning parameters (Pike & Robinson, 2022). Therefore, the high-level features extracted by DBNet are more capable of capturing individual differences in MDD, offering a robust measure for representation learning.

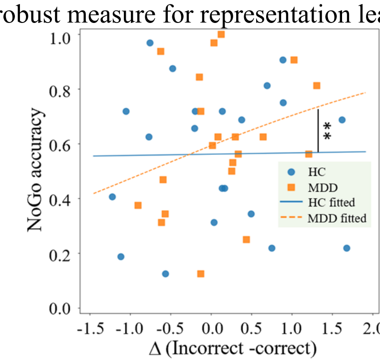

Figure 3: Coupling results of avoidance behavior between groups.

V. Conclusion

This study proposes DBNet, a dual-branch network that combines time-domain modeling with adaptive wavelet time-frequency learning to decode correct versus incorrect feedback from single-trial EEG separately in MDD and HC. DBNet achieves stable, superior performance in both groups—more pronounced in MDD—supporting the utility of jointly capturing ERP morphology and frequency-specific oscillations. Visualizations show that the model emphasizes FRN/P3 time windows, fronto-midline topography, and a stronger theta-band focus in MDD, consistent with canonical error-monitoring mechanisms. Moreover, a binomial GLM links DBNet-derived subject-level representations to test-phase NoGo accuracy, revealing stronger coupling in MDD and indicating that the learned features provide behaviorally meaningful markers of avoidance-related individual differences. Overall, DBNet offers an interpretable deep decoding and statistical-inference framework for relating feedback processing to avoidance behavior in depression.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant W2411084; the Chongqing Natural Science Fund Innovation and Development Joint Fund (Municipal Education Commission) project under Grant CSTB2024NSCQ-LZX0121 and the funding of Institute for Advanced Sciences of Chongqing University of Posts and Telecommunications under Grant E011A2022327.

References

Bachmann, M., et al. (2018). Methods for classifying depression in single channel EEG using linear and nonlinear signal analysis. *Computer Methods and Programs in Biomedicine*, *155*, 11-17. https://doi.org/10.1016/j.cmpb.2017.11.023

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A Single-Trial Estimation of the Feedback-Related Negativity and Its Relation to BOLD Responses in a Time-Estimation Task. *Journal of Neuroscience*, *34*(8), 3005-3012. https://doi.org/10.1523/jneurosci.3684-13.2014

Burkhouse, K. L., et al. (2016). Neural reactivity to reward as a predictor of cognitive behavioral therapy response in anxiety and depression. *Depression and Anxiety*, *33*(4), 281-288. https://doi.org/10.1002/da.22482

Cavanagh, J. F., Bismark, A. J., Frank, M. J., & Allen, J. J. B. (2011). Larger Error Signals in Major Depression are Associated with Better Avoidance Learning. *Frontiers in psychology*, *2*, 331. https://doi.org/10.3389/fpsyg.2011.00331

Cavanagh, J. F., Bismark, A. W., Frank, M. J., & Allen, J. J. B. (2019). Multiple Dissociations Between Comorbid Depression and Anxiety on Reward and Punishment Processing: Evidence From Computationally Informed EEG. *Computational psychiatry (Cambridge, Mass.)*, *3*, 1-17. https://doi.org/10.1162/cpsy_a_00024

Chen, X., Teng, X. B., Chen, H., Pan, Y. F., & Geyer, P. (2024). Toward reliable signals decoding for electroencephalogram: A benchmark study to EEGNeX. *Biomedical Signal Processing and Control*, *87*, 12, Article 105475. https://doi.org/10.1016/j.bspc.2023.105475

Csifcsák, G., & Mittner, M. (2026). Reward processing under illusion of control: The sensitivity of the feedback-related negativity to prediction errors is not altered when random outcomes are perceived as the consequence of one's own actions. *International Journal of Psychophysiology*, *219*, 113300.

Ding, Y., Chen, X. M., Fu, Q. M., & Zhong, S. (2020). A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm. *Ieee Access*, *8*, 75616-75629. https://doi.org/10.1109/access.2020.2987523

Dong, P., et al. (2025). Relationship Between Short Video Addiction Tendency and Depression Among Rural Older Adults: Cross-Sectional Study. *Journal of Medical Internet Research*, *27*, 14, Article e75938. https://doi.org/10.2196/75938

Eisma, J., Rawls, E., Long, S., Mach, R., & Lamm, C. (2021). Frontal midline theta differentiates separate cognitive control strategies while still generalizing the need for cognitive control. *Scientific Reports*, *11*(1), 14, Article 14641. https://doi.org/10.1038/s41598-021-94162-z

Grabowska, A., Zabielski, J., & Senderecka, M. (2024). Machine learning reveals differential effects of depression and anxiety on reward and punishment processing. *Scientific Reports*, *14*(1), 13, Article 8422. https://doi.org/10.1038/s41598-024-58031-9

Hu, L. L., et al. (2025). Joint multi-layer network and coupling redundancy minimization for semi-supervised EEG-based emotion recognition. *Knowledge-Based Systems*, *318*, 17, Article 113559. https://doi.org/10.1016/j.knosys.2025.113559

Jalilpour, S., & Müller-Putz, G. (2025). A framework for Interpretable deep learning in cross-subject detection of event-related potentials. *Engineering Applications of Artificial Intelligence*, *139*, 14, Article 109642. https://doi.org/10.1016/j.engappai.2024.109642

Katz, B. A., Matanky, K., Aviram, G., & Yovel, I. (2020). Reinforcement sensitivity, depression and anxiety: A meta-analysis and meta-analytic structural equation model. *Clinical psychology review*, *77*, 101842. https://doi.org/10.1016/j.cpr.2020.101842

Kim, J. M., et al. (2024). A learnable continuous wavelet-based multi-branch attentive convolutional neural network for spatio-spectral-temporal EEG signal decoding. *Expert Systems with Applications*, *251*, 10, Article 123975. https://doi.org/10.1016/j.eswa.2024.123975

Kratochwil, Z. K., et al. (2025). P100 in major depressive disorder and healthy controls: An emotion-processing task. *International Journal of Psychophysiology*, *218*, 13, Article 113281. https://doi.org/10.1016/j.ijpsycho.2025.113281

Lawhern, V. J., et al. (2018). EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, *15*(5), 17, Article 056013. https://doi.org/10.1088/1741-2552/aace8c

Liao, S. C., Wu, C. T., Huang, H. C., Cheng, W. T., & Liu, Y. H. (2017). Major Depression Detection from EEG Signals Using Kernel Eigen-Filter-Bank Common Spatial Patterns. *Sensors*, *17*(6), 18, Article 1385. https://doi.org/10.3390/s17061385

Liu, M., Tan, J., Jiang, Y., & Tian, Y. (2022). Using deep learning to decode abnormal brain neural activity in MDD from single-trial EEG signals. *Brain-Apparatus Communication: A Journal of Bacomics*, *1*(1), 28-37.

Pegg, S., Arfer, K. B., & Kujawa, A. (2021). Altered reward responsiveness and depressive symptoms: An examination of social and monetary reward domains and interactions with rejection sensitivity. *Journal of Affective Disorders*, *282*, 717-725. https://doi.org/10.1016/j.jad.2020.12.093

Pike, A. C., & Robinson, O. J. (2022). Reinforcement Learning in Patients With Mood and Anxiety Disorders vs Control Individuals A Systematic Review and Meta-analysis. *Jama Psychiatry*, *79*(4), 313-322. https://doi.org/10.1001/jamapsychiatry.2022.0051

Puusepp, I., et al. (2021). Mindsets and Neural Mechanisms of Automatic Reactions to Negative Feedback in Mathematics in Elementary School Students. *Frontiers in psychology*, *12*, 635972. https://doi.org/10.3389/fpsyg.2021.635972

QIN, H., HUANG, R., & JIA, S. (2021). Feedback-related negativity: A biomarker for depression. *Advances in Psychological Science*, *29*(3), 404.

Santamaría-Vázquez, E., Martínez-Cagigal, V., Vaquerizo-Villar, F., & Hornero, R. (2020). EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-Based Brain-Computer Interfaces. *Ieee Transactions on Neural Systems and Rehabilitation Engineering*, *28*(12), 2773-2782. https://doi.org/10.1109/tnsre.2020.3048106

Schirrmeister, R. T., et al. (2017). Deep Learning With Convolutional Neural Networks for EEG Decoding and Visualization. *Human Brain Mapping*, *38*(11), 5391-5420. https://doi.org/10.1002/hbm.23730

Sharma, M., Achuth, P. V., Deb, D., Puthankattil, S. D., & Acharya, U. R. (2018). An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with EEG signals. *Cognitive Systems Research*, *52*, 508-520. https://doi.org/10.1016/j.cogsys.2018.07.010

Tan, J. L., Zhan, Y. C., Tang, Y., Bao, W. X., & Tian, Y. (2024). EEG decoding for effects of visual joint attention training on ASD patients with interpretable and lightweight convolutional neural network. *Cognitive Neurodynamics*, *18*(3), 947-960. https://doi.org/10.1007/s11571-023-09947-x

Tanzilli, A., Giovanardi, G., Patriarca, E., Lingiardi, V., & Williams, R. (2021). From a Symptom-Based to a Person-Centered Approach in Treating Depressive Disorders in Adolescence: A Clinical Case Formulation Using the Psychodynamic Diagnostic Manual (PDM-2)'s Framework. *International Journal of Environmental Research and Public Health*, *18*(19), 19, Article 10127. https://doi.org/10.3390/ijerph181910127

Webb, C. A., et al. (2017). Abnormal neural responses to feedback in depressed adolescents. *Journal of abnormal psychology*, *126*(1), 19-31. https://doi.org/10.1037/abn0000228

Xiong, D. W., et al. (2025). Interpretable Cross-Modal Alignment Network for EEG Visual Decoding With Algorithm Unrolling. *Ieee Transactions on Neural Networks and Learning Systems*, *36*(11), 19894-19908. https://doi.org/10.1109/tnnls.2025.3592646

Zhang, Y., Tan, H., & Luo, S. (2024). Repetition suppression between monetary loss and social pain. *BMC psychology*, *12*(1), 356. https://doi.org/10.1186/s40359-024-01852-0