



OPEN ACCESS

EDITED BY
Enrico Rukzio,
University of Ulm, Germany

REVIEWED BY
Wenbin Shi,
Beijing Institute of Technology, China
Chien-Hung Yeh,
Beijing Institute of Technology, China

*CORRESPONDENCE
Xueying Zhang
tyzhangxy@163.com
Guijun Chen
chenguijun@tyut.edu.cn

SPECIALTY SECTION
This article was submitted to
Perception Science,
a section of the journal
Frontiers in Neuroscience

RECEIVED 21 June 2022
ACCEPTED 17 August 2022
PUBLISHED 07 September 2022

CITATION
Zhang J, Zhang X, Chen G, Huang L
and Sun Y (2022) EEG emotion
recognition based on cross-frequency
granger causality feature extraction
and fusion in the left and right
hemispheres.
Front. Neurosci. 16:974673.
doi: 10.3389/fnins.2022.974673

COPYRIGHT
© 2022 Zhang, Zhang, Chen, Huang
and Sun. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

EEG emotion recognition based on cross-frequency granger causality feature extraction and fusion in the left and right hemispheres

Jing Zhang, Xueying Zhang*, Guijun Chen*, Lixia Huang and Ying Sun

College of Information and Computer, Taiyuan University of Technology, Taiyuan, China

EEG emotion recognition based on Granger causality (GC) brain networks mainly focus on the EEG signal from the same-frequency bands, however, there are still some causality relationships between EEG signals in the cross-frequency bands. Considering the functional asymmetric of the left and right hemispheres to emotional response, this paper proposes an EEG emotion recognition scheme based on cross-frequency GC feature extraction and fusion in the left and right hemispheres. Firstly, we calculate the GC relationship of EEG signals according to the frequencies and hemispheres, and mainly focus on the causality of the cross-frequency EEG signals in left and right hemispheres. Then, to remove the redundant connections of the GC brain network, an adaptive two-stage decorrelation feature extraction scheme is proposed under the condition of maintaining the best emotion recognition performance. Finally, a multi-GC feature fusion scheme is designed to balance the recognition accuracy and feature number of each GC feature, which comprehensively considers the influence of the recognition accuracy and computational complexity. Experimental results on the DEAP emotion dataset show that the proposed scheme can achieve an average accuracy of 84.91% for four classifications, which improved the classification accuracy by up to 8.43% compared with that of the traditional same-frequency band GC features.

KEYWORDS

electroencephalogram, emotion recognition, Granger causality (GC), cross-frequency analysis, feature extraction, multi-feature fusion

1. Introduction

Electroencephalogram (EEG) is a kind of physiological electrical signal that can reflect dynamic changes of the central nervous system. In recent years, EEG signals have been widely used in emotion recognition because of their more objective response to the emotional state (Alarcao and Fonseca, 2017; Esposito et al., 2020; Rahman et al., 2021; Wu et al., 2022).

Modern cognitive neuroscience has indicated that the brain hemispheres are anatomically and functionally asymmetric (Dimond et al., 1976; Zheng and Lu, 2015; Li et al., 2018; Cui et al., 2020). In Zheng and Lu (2015), the cognitive differences between the left and right hemispheres have been revealed by analyzing the emotional cognitive characteristics induced by emotion stimulation, and found that the right hemisphere has more power in perceiving negative emotion. In Li et al. (2018), the differential entropy of each pair of EEG channels at the symmetrical position of the two hemispheres was calculated and used to obtain the differential asymmetry (DASM) and rational asymmetry (RASM), which has been proved to be effective in distinguishing emotional states. In Cui et al. (2020), a bi-hemisphere domain adversarial neural network model was designed to effectively improve the performance of EEG emotion recognition. Therefore, analyzing the EEG signals of the left and right hemispheres is of great significance for improving emotional recognition.

EEG brain network is one of the most effective methods for analyzing EEG signals, where each EEG channel represents a node and the connections between nodes are defined as the edges. According to whether the edges may be directed or not, brain connectivity can be subdivided into functional connectivity and effective connectivity (Jiang et al., 2004; Li et al., 2019; Cao et al., 2020, 2022). Functional connectivity is defined as statistical interdependence among the EEG signals, while effective connectivity can further measure the causal relationships of EEG signals. Granger causality (GC) (Granger, 1969) is an effective connection measure and the GC brain network has been extensively used to explore the causality of EEG signals in recent years (Dimitriadis et al., 2016; Tian et al., 2017; Jiang et al., 2019; Gao et al., 2020; Li T. et al., 2020; Chen et al., 2021). Generally, the EEG signal is usually decomposed into four bands: θ , α , β , and γ bands to define the change in brain state. In Gao et al. (2020), the GC brain network was constructed for β -band EEG signals to analyze the difference between calm and stress emotion states. In Dimitriadis et al. (2016), by calculating the cross-frequency causal interaction between the EEG signals of $\theta\sim\alpha$ bands under a mental arithmetic task, the mechanism of human brain processing was deeply analyzed. However, the causal analysis in the current studies either within the same-frequency bands (e.g., β band) or via specific cross-frequency interactions (e.g., $\theta\sim\alpha$) (Yeh et al., 2016), and the effect of the cross-frequency causal interaction between EEG signals for on the emotion recognition is not completely analyzed. In essence, there is a causal relationship between EEG signals both in same-frequency and cross-frequency bands. Thus, it is necessary to carry out causality analysis among the EEG signal with different frequency bands and different hemispheres for emotion recognition.

On the other hand, it is a key issue to extract effective features from the GC brain network. The existing researches

always selects the empirical threshold to directly converted the GC adjacency matrix into a binary data, that is, the GC values below the threshold are set to 0, and the GC values above the threshold value are set to 1. Then, the binary matrix is converted to network attributes as brain cognitive features based on the graph theory (De Vico Fallani et al., 2017; Hu et al., 2019; Covantes-Osuna et al., 2021; Li et al., 2022). In this way, the weak connections with lower GC values may be lost for the values below the threshold are set to 0. Otherwise, De Vico Fallani et al. (2012) have pointed out that the brain network exhibits a natural high redundant connection in all frequency bands, but the threshold method cannot effectively remove these redundant connections. Instead of the threshold method, this paper proposed an adaptive two-stage decorrelation (ATD) feature extraction method to improve the performance of emotion recognition, in which the redundant GC connections are adaptively removed under the goal of maintaining the optimal emotion recognition performance.

As is well-known from previous studies, a single type of GC feature can only show a part of causality information. In order to describe the causal interaction during emotion response more accurately, it is necessary to integrate the same-frequency GC features and cross-frequency GC features. Previous works on feature fusion aim to directly combine different types of feature vectors by concatenation, parallel or weighted fusion to improve the recognition accuracy (Yang et al., 2019; Bota et al., 2020; Cai et al., 2020; Li Y. et al., 2020; Yilmaz and Kose, 2021; An et al., 2022). However, existing feature fusion methods do not consider the increase in computational cost caused by the increase in feature numbers. Based on this, this paper designs a new multi-GC feature fusion scheme, which can effectively improve the performance of the EEG emotion recognition without increasing the number of features.

As mentioned above, this paper mainly studies the EEG emotion recognition based on cross-frequency GC feature extraction and fusion in the left and right hemispheres. Our contributions mainly focus on the following aspects: Firstly, the EEG electrode channels are divided into the left and right hemispheres according to their spatial position, and the EEG signals are decomposed into θ , α , β , and γ bands. The corresponding GC adjacency matrices are then constructed to analyze the causality of the EEG signals. Secondly, based on the characteristics of GC adjacency matrix, an ATD method is further proposed to remove the redundant connections in the GC brain network and extract the causal features for emotion recognition. Finally, a new weighted feature fusion scheme is designed that takes into account the emotion recognition accuracy and the feature number of each single GC feature, which can effectively improve the performance of the EEG emotion recognition system without increasing the computational cost. Experimental results of arousal-valence classification on the DEAP emotion dataset (Koelstra et al., 2011) show that the proposed scheme can achieve an average

recognition improvement of 8.43% than that of the same-frequency GC features.

The remaining parts are organized as follows. Section 2 reviews the DEAP emotion dataset and GC theory. Section 3 describes the proposed EEG emotion recognition method, including the frequency-hemisphere GC measure for the EEG signal, the ATD feature extraction method, and the proposed multi-feature fusion scheme. Section 4 presents the experimental results and discussions. Finally, some conclusions are presented in Section 5.

2. Related works

2.1. DEAP emotion dataset

The DEAP dataset (Koelstra et al., 2011) is a multimodal dataset for analyzing human affective states, which consists of EEG signals and peripheral physiological signals of 32 subjects. During the experiment, all subjects watched 40 excerpts of one-minute music videos. At the end of each trial, participants performed a self-assessment of their levels of arousal, valence, linking, and dominance using SAM mannequins on a discrete 9-point scale. The arousal scale ranged from calm to excited, the valence scale ranged from unhappy to happy, the linking scale measured the personal preferences of the participants for a given media, and the dominance scale ranged from submissive to dominant. The EEG signals were recorded from 32 channels according to the international 10–20 system, at a sampling rate of 512 Hz, while peripheral physiological signals including skin temperature, blood volume pressure, an electromyogram, and galvanic skin response, were recorded from another 8 channels.

In this paper, only the EEG signal is used to investigate EEG emotion recognition research. According to the 19 self-assessment scores of participants, we select the median score 5 as the threshold, with higher than 5 representing high class and less than or equal to 5 representing low class. The valence-arousal (VA) space is divided into four parts, i.e., low arousal-low valence (LALV), high arousal-low valence (HALV), low arousal-high valence (LAHV), and high arousal-high valence (HAHV).

2.2. Preprocessing and frequency band division of EEG signal

The original raw EEG data is first carried out in the following 4 preprocessing steps: (1) down-sampling to 128 Hz, (2) removal of the EOG artifacts, (3) bandpass filtering of the raw data between 4 and 45 Hz, and (4) averaging of the data to a common reference. Then, the Short-time Fourier Transform (STFT) is then used to extract the θ (4~8 Hz), α (8~12 Hz), β (12~30 Hz) and γ (30~45 Hz) bands. To observe the waveform of EEG signal decomposed into four frequency bands more intuitively,

Figure 1 shows the frequency band division of a 3s EEG signal in the DEAP dataset. Figure 1A shows the original EEG signal, Figure 1B is the corresponding power spectral density of the EEG signal, and Figure 1C presents the four-band EEG signals extracted by STFT.

2.3. Overview of GC analysis

Granger causality is one of the most popular approaches for quantifying causal relationships between time series data, which introduced first in econometrics by Granger (1969). GC analysis is widely used in emotion recognition because of its strong interpretability. It is based on two major principles: (i) the cause happens prior to the effect, and (ii) the cause makes notable changes in the effect. Generally speaking, a time series X is said to “Granger cause” another time series Y , denoted by $X \rightarrow Y$. More specifically, granger causality occurs if and only if the prediction values of Y based on the past values of X and Y are better than predictions based on the past values of Y alone. For two time series X and Y , a univariate and a bivariate vector autoregressive (VAR) model are performed to the predicted current values by the following regressions:

$$X(t) = \sum_{i=1}^L a_{1i}X(t-i) + \varepsilon_X \quad (1)$$

$$Y(t) = \sum_{i=1}^L b_{1i}Y(t-i) + \varepsilon_Y \quad (2)$$

$$X(t) = \sum_{i=1}^L a_{2i}X(t-i) + \sum_{i=1}^L b_{2i}Y(t-i) + \eta_{YX} \quad (3)$$

$$Y(t) = \sum_{i=1}^L c_{2i}X(t-i) + \sum_{i=1}^L d_{2i}Y(t-i) + \eta_{XY} \quad (4)$$

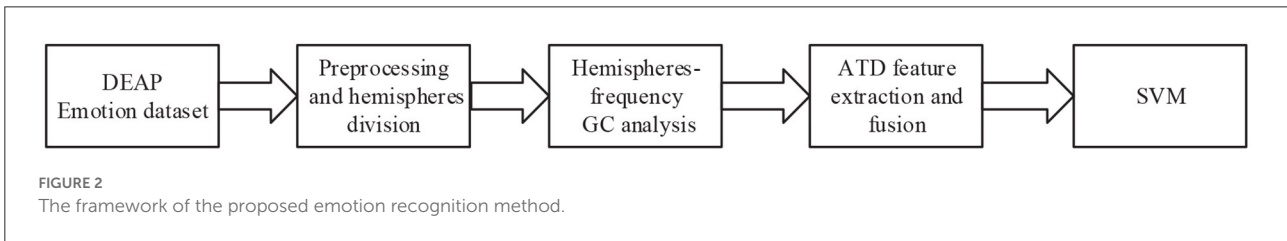
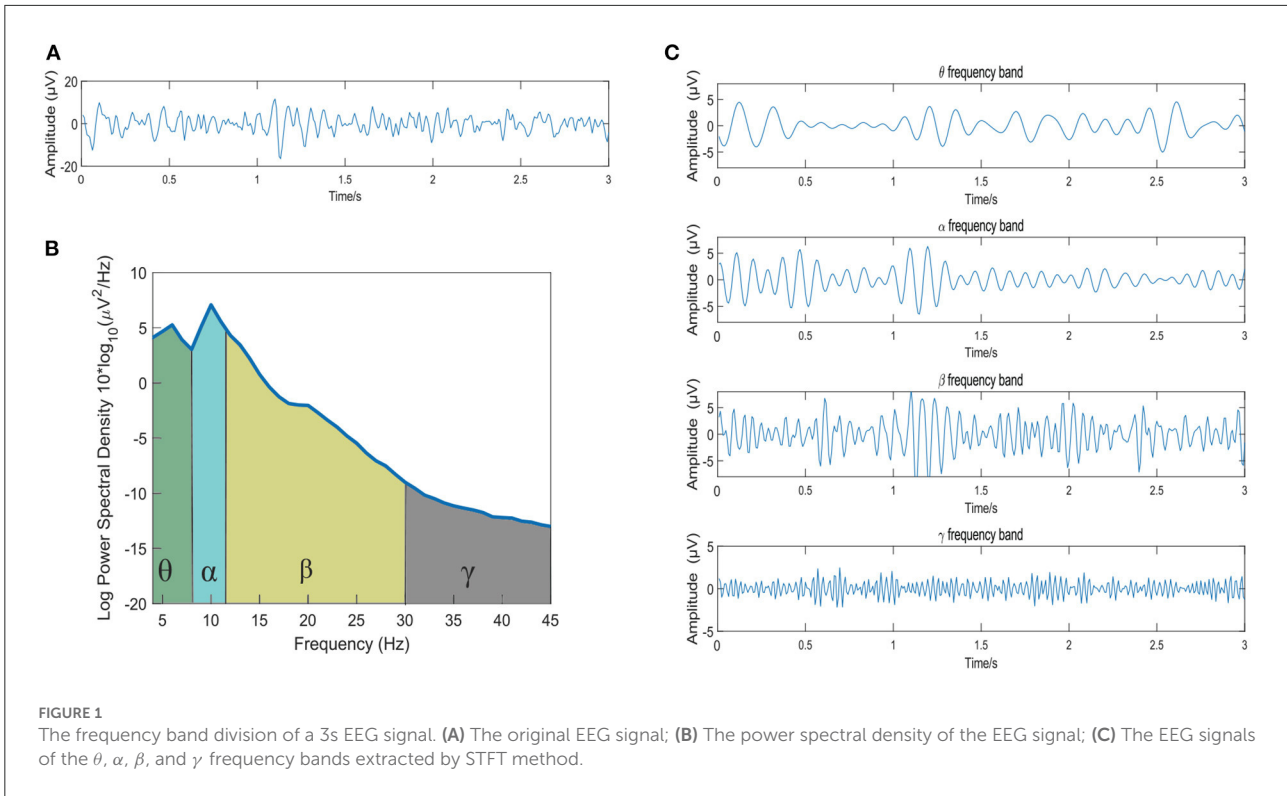
where a_{1i} , b_{1i} , a_{2i} , b_{2i} , c_{2i} , and d_{2i} ($i = 1, 2, \dots, L$) are the constant coefficients, and L is the order of the model, both of them can be obtained through the Bayesian information criterion. ε_X and ε_Y represent the error of the univariate model, η_{YX} and η_{XY} represent the error of the bivariate model, respectively.

The mathematical definition of GC is the logarithm of the two ratios of the error variances: the variance of the errors from the univariate VAR model and bivariate VAR model.

$$F_{X \rightarrow Y} = \ln\left(\frac{\sigma_{\varepsilon_Y}}{\sigma_{\eta_{XY}}}\right) \quad (5)$$

$$F_{Y \rightarrow X} = \ln\left(\frac{\sigma_{\varepsilon_X}}{\sigma_{\eta_{YX}}}\right) \quad (6)$$

where the σ_{ε_X} , σ_{ε_Y} , $\sigma_{\eta_{XY}}$, and $\sigma_{\eta_{YX}}$ represent the variances of the error in Equations (1)–(4), respectively. When σ_{ε_X} is large than



$\sigma_{\eta_{YX}}$, means that Y is the “Granger Cause” to X. Similarly, when σ_{ϵ_Y} is large than $\sigma_{\eta_{XY}}$, means that X is the “Granger Cause” to Y.

3. The proposed EEG emotion recognition scheme

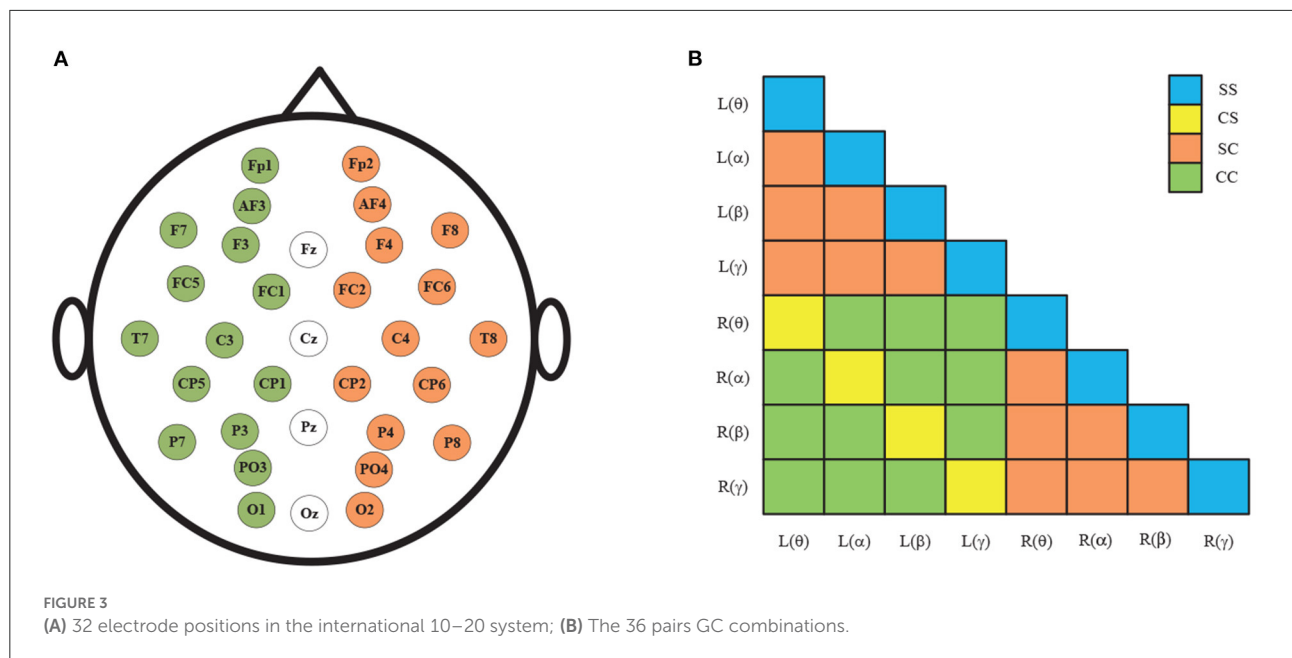
Figure 2 shows the framework of the proposed EEG emotion recognition method. The preprocessing and hemispheres division module is first used to divide the original EEG signal into the left and right hemispheres according to their positions, and STFT is then used to extract the θ (4~8 Hz), α (8~12 Hz), β (12~30 Hz) and γ (30~45 Hz) bands. Hemispheres-frequency GC analysis module is then used to analyze the GC relationship of the EEG signals with different frequencies in the left and right hemispheres, and corresponding GC value is calculated to construct the adjacency matrix. After this, the ATD feature

extraction and fusion module is adopted to adaptively remove the redundant connections in the GC brain network and extract the optimized GC features (GC+ATD), and design a multi-feature fusion scheme to integrate different GC+ATD features. Finally, the support vector machine (SVM) (Chang and Lin, 2011) classifier is used to obtain the emotion recognition results.

3.1. The proposed GC measures of left and right hemispheres

3.1.1. The classification of hemispheres-frequency GC measures

Figure 3A shows the spatial distribution of the 32 EEG electrodes in the international 10–20 system. Taking the middle four EEG electrodes as the axis, the human brain is divided



into left and right hemispheres, with 14 EEG electrodes in each hemisphere. After the removal of four central electrodes (*Fz*, *Cz*, *Pz*, and *Pz*), the remaining electrodes are symmetrical and divided into 14 pairs of left and right combinations: *Fp1-Fp2*, *AF3-AF4*, *F7-F8*, *F3-F4*, *FC5-FC6*, *FC1-FC6*, *T7-T8*, *C3-C4*, *CP5-CP6*, *CP1-CP2*, *P7-P8*, *P3-P4*, *PO3-PO4*, and *O1-O2*. Finally, the hemispheres causality and frequencies causality of EEG signals is combined and resulting in 36 combinations as shown in Figure 3B.

As shown in Figure 3B, the GC combinations can be furtherly divided into four categories according to whether the hemispheres and frequencies are the same: (1) same-hemisphere and same-frequency (SS), which reflects the GC of EEG signals with the same hemispheres and same frequency domain; (2) cross-hemisphere and same-frequency (CS), which reflects the GC of the same-frequency EEG signals between the left and right hemispheres; (3) same-hemisphere and cross-frequency (SC), which reflects the GC of cross-frequency EEG signals with the same hemisphere; (4) cross-hemisphere and cross-frequency (CC), which reflects the GC of cross-frequency EEG signals between the left and right hemispheres.

To analyze the above measures more specifically, the proposed GC adjacency matrices between the θ band and the $\theta\sim\alpha$ band of an EEG signal are shown in Figure 4, where *L* and *R* represent the left and right hemispheres, and θ and α represent the frequency band. For example, *L*(θ) represents the θ -frequency EEG signals in the left hemisphere. Since there are 14 EEG electrodes in each hemisphere, the size of the corresponding GC matrix is 28×28 . Among them, SS and CS belong to the same-frequency GC measures, which have been widely explored in the existing research. On the contrary, SC

and CC belong to the cross-frequency GC measures, which are rarely involved in emotion recognition, especially CC measure. Therefore, we mainly focus on the analysis of the CC measure in this subsection, and the other three measures can be analyzed to CC.

As shown in Figure 3B, the CC measures are pairwise symmetric and can be combined into six groups: $\theta\sim\alpha$, $\theta\sim\beta$, $\theta\sim\gamma$, $\alpha\sim\beta$, $\alpha\sim\gamma$ and $\beta\sim\gamma$. In the following, we will take the $CC(\theta\sim\alpha)$ as an example to analyze the GC relationship of EEG signals. As shown in Figure 4D, the lower-left and the upper-right of the $CC(\theta\sim\alpha)$ matrix are the values of $L(\alpha)\sim R(\theta)$ and $L(\theta)\sim R(\alpha)$, while the rest values (including the lower-right and the upper-left) of the same hemisphere are set to 0. After the removal of four central electrode nodes from the 32 electrode nodes, both of left and right hemispheres have 14 electrode nodes, and the number of effective GC values in the $CC(\theta\sim\alpha)$ is $14\times 14\times 2=392$. Similarly, the number of GC values in SS, CS and SC can be obtained, as shown in Table 1.

3.1.2. The discussion of CC measure

To further analyze the characteristics of six frequency-hemispheres combinations in CC causal measurement, the corresponding the GC adjacency matrix of $CC(\theta\sim\alpha)$, $CC(\theta\sim\beta)$, $CC(\theta\sim\gamma)$, $CC(\alpha\sim\beta)$, $CC(\alpha\sim\gamma)$ and $CC(\beta\sim\gamma)$ are presented in Figures 5A–F, respectively. It can be seen that there is a significant GC relationship between EEG signals in the adjacent frequencies, such as $CC(\theta\sim\alpha)$, $CC(\alpha\sim\beta)$, and $CC(\beta\sim\gamma)$, where $CC(\beta\sim\gamma)$ holds the strongest causality. On the contrary, there is almost no causality between EEG signals in the non-adjacent frequencies, such as $CC(\theta\sim\beta)$, $CC(\theta\sim\gamma)$ and $CC(\alpha\sim\gamma)$.

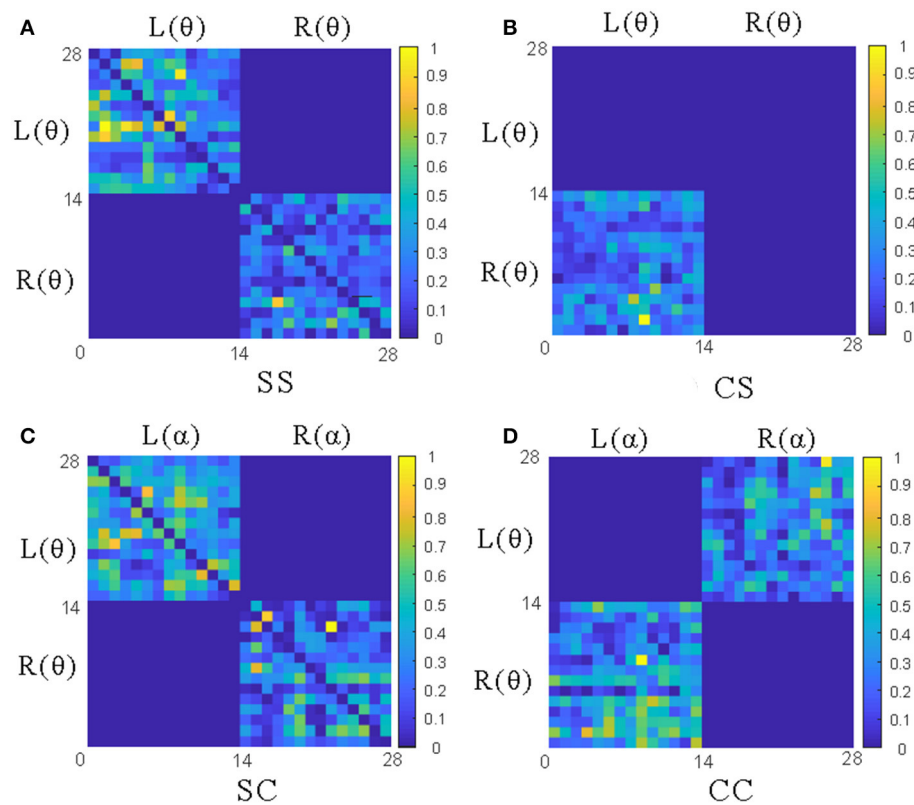


FIGURE 4
The adjacency matrices of four GC measures between the θ band and the $\theta\sim\alpha$ band. **(A)** The $L(\theta)\sim L(\theta)$ and $R(\theta)\sim R(\theta)$ matrices of SS measures; **(B)** The $L(\theta)\sim R(\theta)$ matrix of CS measures; **(C)** The $L(\theta)\sim L(\alpha)$ and $R(\theta)\sim R(\alpha)$ matrices of SC measures; **(D)** The $L(\theta)\sim R(\alpha)$ and $L(\alpha)\sim R(\theta)$ matrices of CC measures.

TABLE 1 The GC value numbers of different measures.

GC measure	Frequency-hemispheres combination	The GC value numbers of single combination	The GC value numbers of all combinations
SS	$\theta\sim\theta, \alpha\sim\alpha, \beta\sim\beta, \gamma\sim\gamma$	392	1,568
CS	$\theta\sim\theta, \alpha\sim\alpha, \beta\sim\beta, \gamma\sim\gamma$	196	784
SC	$\theta\sim\alpha, \theta\sim\beta, \theta\sim\gamma, \alpha\sim\beta, \alpha\sim\gamma, \beta\sim\gamma$	392	2,352
CC	$\theta\sim\alpha, \theta\sim\beta, \theta\sim\gamma, \alpha\sim\beta, \alpha\sim\gamma, \beta\sim\gamma$	392	2,352

Therefore, the causal features of the adjacent frequency EEG signals can be effectively used for improving the emotion recognition performance, and how to extract the causal features

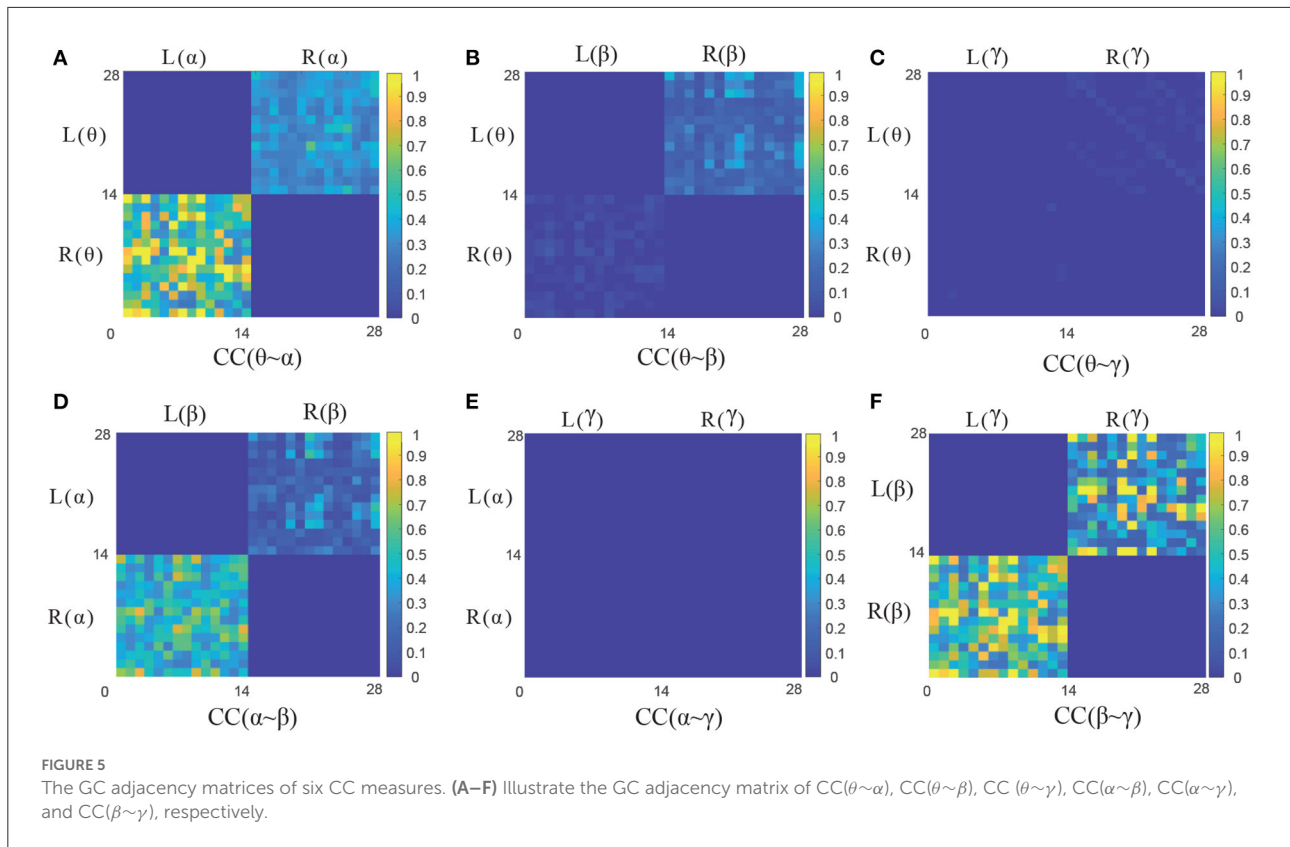
from the GC adjacency matrix will be discussed in detail in the following section.

3.1.3. Formula derivation of $CC(\theta\sim\alpha)$ measure

In order to extract the features of the GC adjacency matrix, it is necessary to calculate the GC value. In this section, we will take $CC(\theta\sim\alpha)$ measure as an example to describe the detailed calculation process. Let $X_{L(\theta)}(t)$ and $Y_{R(\alpha)}(t)$ denote the t -th time-lagged values of $L(\theta)$ and $R(\alpha)$, respectively. Similar to Equations (1)–(4), the univariate GC model and the bivariate GC model of $X_{L(\theta)}(t)$ and $Y_{R(\alpha)}(t)$ can be expressed as follows:

$$X_{L(\theta)}(t) = \sum_{i=1}^L a_{1i} X_{L(\theta)}(t-i) + \varepsilon_X \tag{7}$$

$$Y_{R(\alpha)}(t) = \sum_{i=1}^L b_{1i} Y_{R(\alpha)}(t-i) + \varepsilon_Y \tag{8}$$



$$X_{L(\theta)}(t) = \sum_{i=1}^L a_{2i} X_{L(\theta)}(t-i) + \sum_{i=1}^L b_{2i} Y_{R(\alpha)}(t-i) + \eta_{YX} \tag{9}$$

$$Y_{R(\alpha)}(t) = \sum_{i=1}^L c_{2i} X_{L(\theta)}(t-i) + \sum_{i=1}^L d_{2i} Y_{R(\alpha)}(t-i) + \eta_{XY} \tag{10}$$

where a_{1i} , b_{1i} , a_{2i} , b_{2i} , c_{2i} , and d_{2i} ($i = 1, 2, \dots, L$) are the constant coefficients, and L is the order of the model.

Then, according to Equations (5) and (6), the GC values of $X_{L(\theta)} \rightarrow Y_{R(\alpha)}$ and $Y_{R(\alpha)} \rightarrow X_{L(\theta)}$ are

$$F_{X_{L(\theta)} \rightarrow Y_{R(\alpha)}} = \ln\left(\frac{\sigma_{\epsilon_{YX}}}{\sigma_{\eta_{XY}}}\right) \tag{11}$$

$$F_{Y_{R(\alpha)} \rightarrow X_{L(\theta)}} = \ln\left(\frac{\sigma_{\epsilon_{XY}}}{\sigma_{\eta_{YX}}}\right) \tag{12}$$

Since $F_{X_{L(\theta)} \rightarrow Y_{R(\alpha)}}$ and $F_{Y_{R(\alpha)} \rightarrow X_{L(\theta)}}$ represents the GC value the upper right and lower left of the $L(\theta) \sim R(\alpha)$, respectively. Therefore, these two feature vectors are directly cascade as the GC values of $L(\theta) \sim R(\alpha)$ and take *concat* to represent

this process:

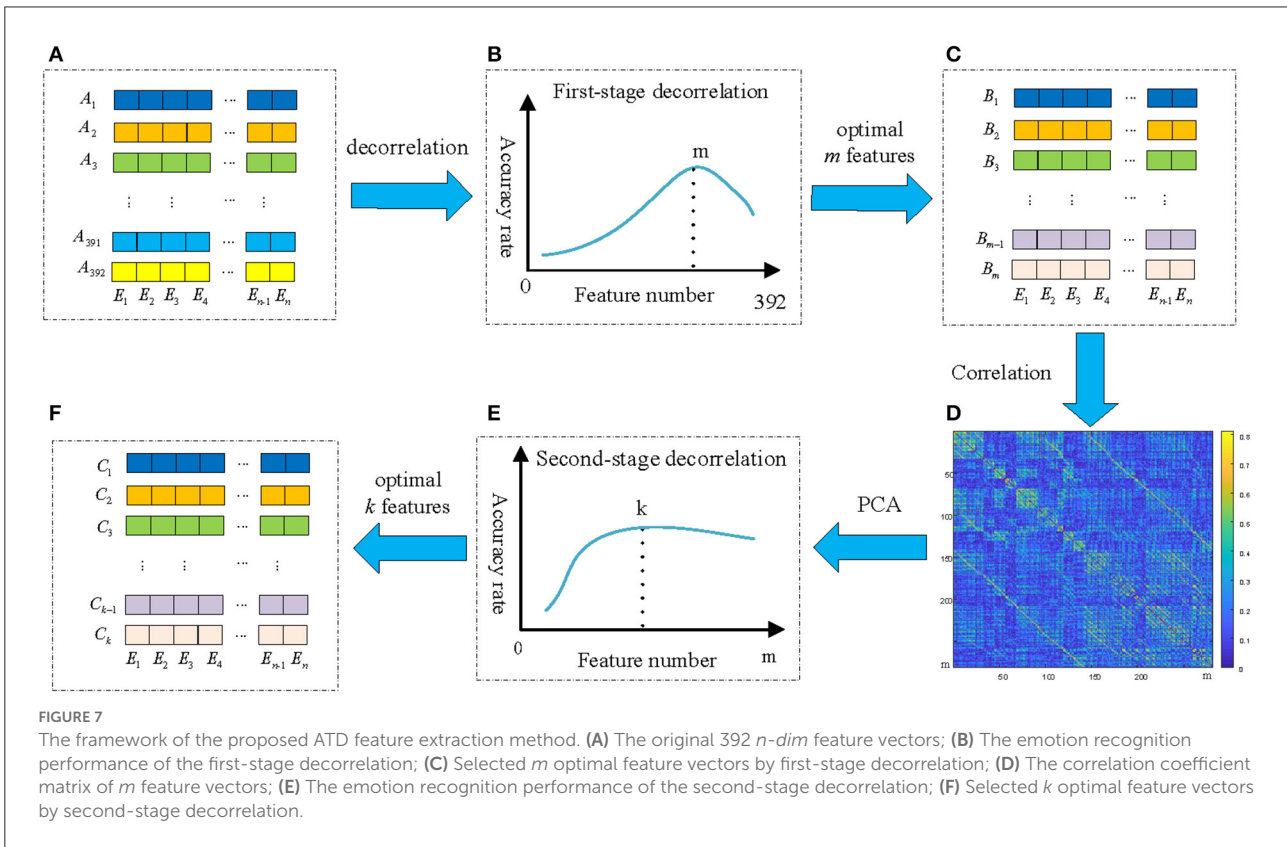
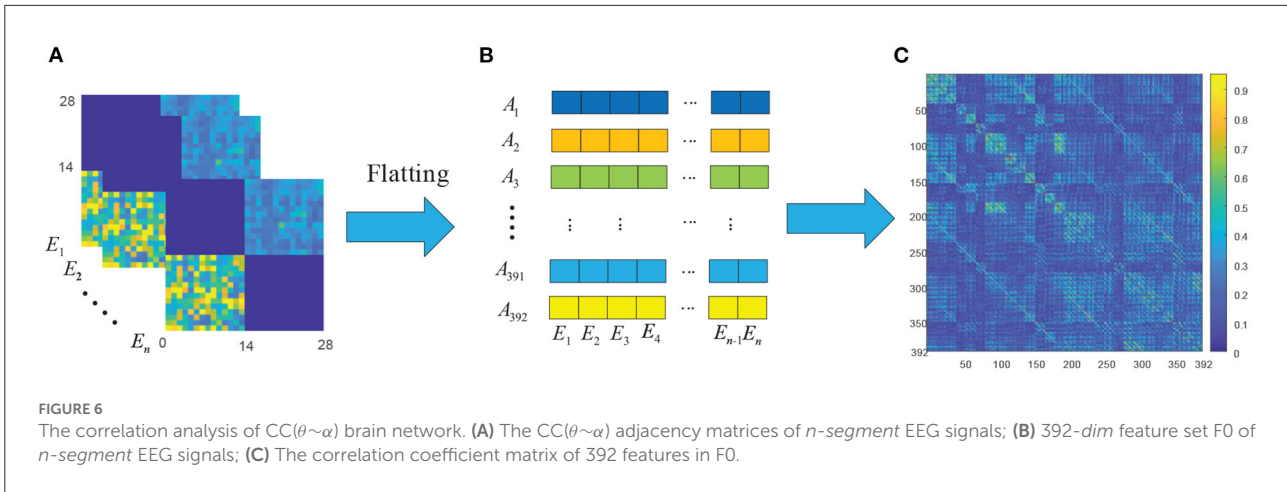
$$F_{L(\theta) \sim R(\alpha)} = \text{concat}(F_{X_{L(\theta)} \rightarrow Y_{R(\alpha)}}, F_{Y_{R(\alpha)} \rightarrow X_{L(\theta)}}) \tag{13}$$

According to the Figures 3B, 4D, the $CC(L(\theta) \sim R(\alpha))$ and $CC(L(\alpha) \sim R(\theta))$ are symmetric. Therefore, $F_{CC(L(\alpha) \sim R(\theta))}$ can be calculated similarly by Equations (7)–(13). As a result, the GC values of $CC(\theta \sim \alpha)$ can be expressed as:

$$F_{CC(\theta \sim \alpha)} = \text{concat}(F_{CC(L(\theta) \sim R(\alpha))}, F_{CC(L(\alpha) \sim R(\theta))}) \tag{14}$$

3.2. The propose ATD method

Due to redundant connections in the GC brain network, this section takes $CC(\theta \sim \alpha)$ as an example for detailed analysis of the ATD method. We select n -segment EEG signals in the DEAP dataset to construct the $CC(\theta \sim \alpha)$ brain networks and calculate the corresponding adjacency matrices E_i ($i = 1, 2, \dots, n$), as shown in Figure 6A. Table 1 shows that the number of GC values in each $CC(\theta \sim \alpha)$ matrix is 392, and these 392 GC values can be calculated according to the calculation procession in Section 3.1.3. Then, the GC values of the same spatial position in CC adjacency matrices are constructed as a n -dim feature vector

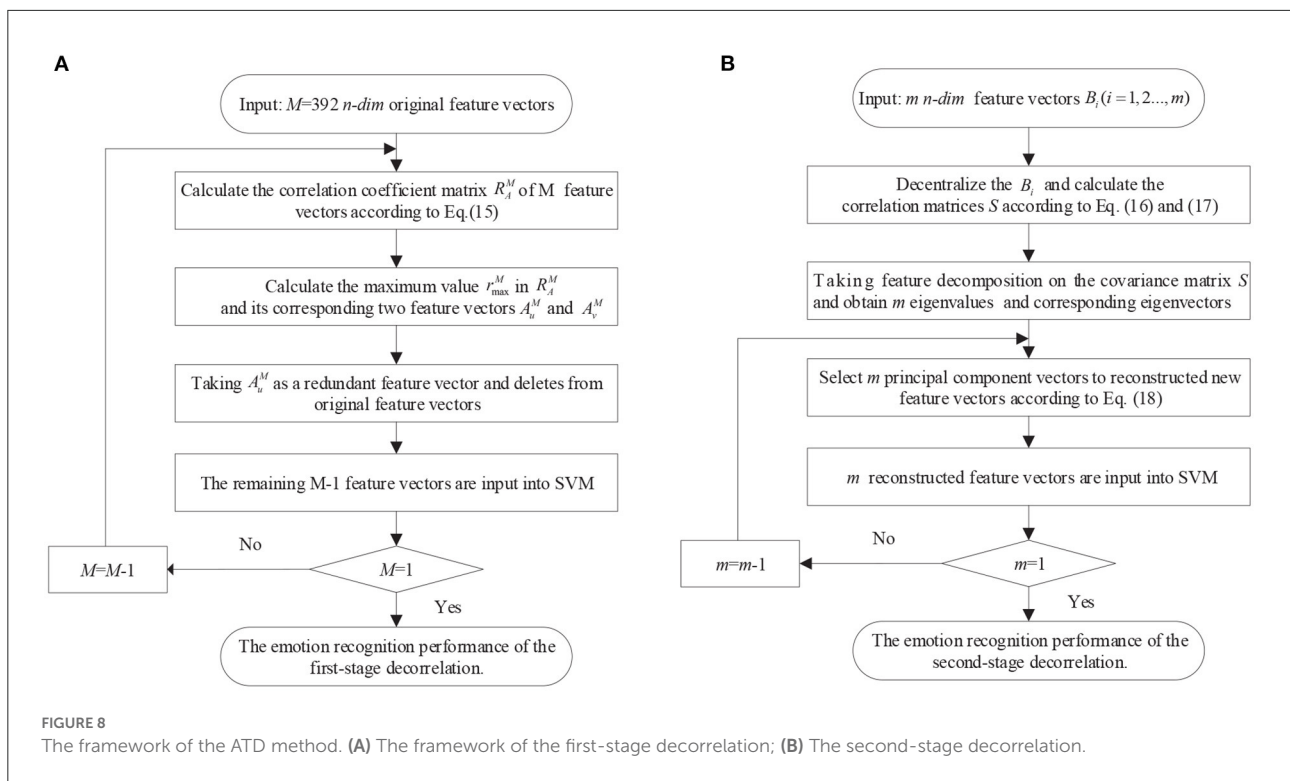


and resulting in 392 n -dim feature vectors $A_i(i = 1, 2 \dots 392)$, as shown in Figure 6B.

Next, we analyze the correlation between two feature vectors of $A_i(i = 1, 2 \dots 392)$. For any two features A_u and $A_v(u, v = 1, 2, \dots, 392, u \neq v)$, the correlation coefficient $r(A_u, A_v)$ between them can be calculated from Equation (15). The correlation coefficient matrix $R_A \in \mathbb{R}^{392 \times 392}$ of 392 feature vectors can be obtained and shown in Figure 6C, where the value

is closer to 1 with the greater correlation, and the value is closer to 0 with the smaller correlation.

$$r(A_u(i), A_v(i)) = \frac{\sum_{i=1}^n (A_u(i) - \bar{A}_u)(A_v(i) - \bar{A}_v)}{\sqrt{\sum_{i=1}^n (A_u(i) - \bar{A}_u)^2} \sqrt{\sum_{i=1}^n (A_v(i) - \bar{A}_v)^2}} \tag{15}$$



It is obviously that there are many highly correlated redundant features among the 392 feature vectors, and the correlation coefficient matrix is symmetric. If these redundant feature vectors can be removed, the effectiveness of the features can be improved. In addition, since the GC brain network of different EEG signals are different, and redundancy connections in the corresponding brain network are also different. Therefore, the redundancy method should be able to adapt to the variation of different EEG signals. In this paper, we propose a new feature extraction and optimization method of ATD method to meet the requirements, and the framework of the proposed ATD method is shown in Figure 7.

The first-stage decorrelation of the original 392 n -dim feature vectors is shown in Figure 8A. Here M is used to represent the number of the feature vectors and initialized to 392. Corresponding to the above-mentioned symbols, we add the superscript M to all them, such as R_A rewritten as R_A^M . Also, due to the symmetry of R_A^M , this paper only selected the lower triangular matrix for correlation analysis.

First of all, the maximum value r_{\max}^M in the correlation coefficient matrix R_A^M and its corresponding two feature vectors A_u^M and A_v^M were determined, and randomly select one of these two vectors as a redundant vector and remove from the original feature vectors. In this paper, we compared four kinds of redundant vectors definition methods: (1) the first feature vector A_u^M is selected as redundant vector; (2) the second feature vector A_v^M is selected as redundant vector; (3) randomly selects

one of them as redundant; (4) Calculate the correlation between the A_u^M and A_v^M and the remain $M-2$ feature vectors, and the highest correlation feature vector is selected as redundant. The experimental results show that the emotion recognition performance obtained by the above four methods is almost the same. Therefore, we adopt the simplest method, that is, the first feature vector A_u^M is deleted as of a redundant feature vector, and the remaining $M - 1$ feature vectors are used as the input of the SVM classifier and obtained the recognition accuracy. Then, the above process is repeated for the remaining $M-1$ feature vectors until the number of remaining features is 1. Figure 7B shows the relationship of emotion recognition accuracy with the different number of feature vectors. It can be seen that there is an optimal number of $m(m < 392)$ to make the emotion recognition performance reach the highest and retains these m feature vectors $B_i (i = 1, 2, \dots, m)$ as optimal feature vectors of the first-stage decorrelation features, as presented in Figure 7C.

To further explore the correlation of the above m features vectors, the corresponding correlation matrices $R_B \in \mathbb{R}^{m \times m}$ between B_i are calculated according to Equation (15) and shown in Figure 7D. We can find that there is still a certain correlation between these feature vectors, and it is necessary to furtherly remove the correlation between them. As we all know, Principal components analysis (PCA) (Li et al., 2016) is a statistical procedure that uses an orthogonal transformation to convert a set of correlated variables into a set of values of linearly uncorrelated variables, and the main idea of PCA is to reduce

the number of variables while preserving as much information as possible. Therefore, we then adopt the PCA method to secondly remove the correlation of the B_i , and the detailed process is shown in Figure 8B.

First, m feature vectors $B_i (i = 1, 2 \dots, m)$ are decentralized, that is, the mean values of m feature vectors are subtracted from each other to ensure that the mean value of each feature vector is 0. The m decentralized feature vectors B_M can be represented as:

$$B_M = B_i - \bar{B}_i (i = 1, 2 \dots, m) \tag{16}$$

Then, the correlation matrices $S \in \mathbb{R}^{m \times m}$ of m feature vectors B_M is calculated as follows:

$$S = \frac{1}{n-1} B_M B_M^T \tag{17}$$

After this, we perform feature decomposition on the covariance matrix S and obtain m eigenvalues $\lambda_1, \lambda_2 \dots, \lambda_m$ and corresponding eigenvectors $\xi_1, \xi_2, \dots, \xi_m$. Taking the first p eigenvalues and the corresponding eigenvectors to construct the orthogonal matrix $V = (\xi_1, \xi_2, \dots, \xi_p)$, where each column in V corresponds to a principal component. Based on this, the reconstructed feature vectors B' with the size of $p \times n$ are obtained, as described in Equation (18), which are input into the SVM classifier to obtain the corresponding emotion recognition accuracy.

$$B' = V^T B \tag{18}$$

Finally, different values $p (p < m)$ were selected to repeat the above process, and the performance of the second-stage decorrelation method with the different number of feature vectors was obtained and shown in Figure 7E. It can be seen that there exists an optimal number of feature vectors $k (k < m)$ to optimize the emotion recognition performance, and the k feature vectors $C_i (i = 1, 2 \dots, k)$ are retained, as shown in Figure 7F.

Above all, we can get the GC+ATD features after removing redundant connections in the GC correlation matrix by ATD method, which will be used for EEG emotion recognition. Because the redundant connections in each GC brain network are always different, the correlation degree of the redundant connections is also different. Therefore, it is necessary to adaptively select the optimal m and k in the above ATD process to ensure the extracted causal features can achieve the best recognition performance.

3.3. The proposed GC+ATD multi-feature weighted fusion method

According to whether the frequency domain and hemispheres domain are used, GC analysis of EEG signals is divided into four categories, as shown in Table 1. A single type of GC feature can only show a part of the causality information. Therefore, it is necessary to integrate four GC+ATD features to make full use of the causal complementarity among them.

Most of the existing research on feature fusion method is commonly used in series, or in parallel or weighted superpositions. Generally speaking, if the number of features is large and the structure is different, the fusion algorithm is complex and the computational complexity is higher. In addition, the four GC+ATD features in this paper have structural similarities. Therefore, considering the computational complexity and performance, a new multi-feature weighted fusion scheme is proposed to balance the recognition rate and feature numbers, which designing the weight function for each GC+ATD feature by comprehensively considering the recognition rate and the feature number of each single feature.

Let T_i and T_{final} represent the i -th GC+ATD feature and the final fusion feature, respectively. The problem of the proposed feature fusion scheme can be expressed as:

$$T_{final} = \sum_{i=1}^4 w_i T_i \tag{19}$$

where w_i is the weight of T_i , it can be calculated as follows.

Let R_i is the recognition accuracy of T_i . N_i and N_{all} represent the feature number of T_i and T_{final} , respectively. To ensure the features T_i with better recognition performance and smaller feature numbers have a higher importance in T_{final} , the weight coefficients w_i are designed as follows.

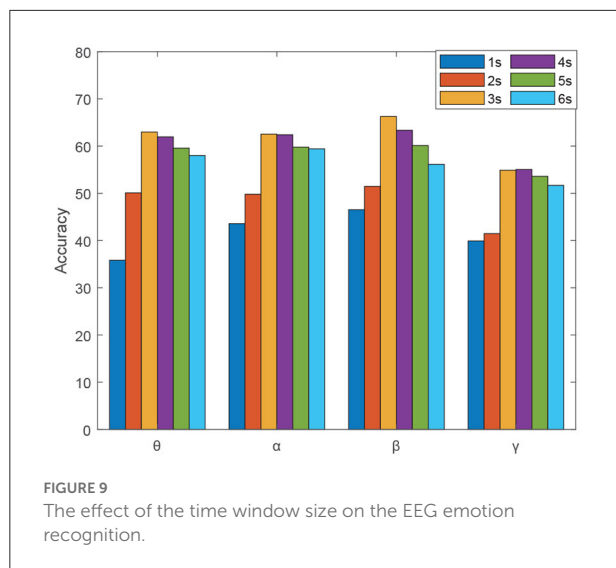
$$w_i = p_i R_i + q_i (1 - \frac{F_i}{F_{all}}) (i = 1, 2, 3, 4) \tag{20}$$

where p_i and q_i represent the importance of recognition accuracy and feature number of the T_i , and $p_i + q_i = 1$.

In summary, to maximize the emotion recognition accuracy $R_{T_{final}}$ of the fusion GC features T_{final} , the problem of the above multi-feature fusion scheme can be expressed as follows:

$$\begin{aligned} & \arg \max_{p_i \in [0,1]} R_{T_{final}} \\ & s.t. \quad T_{final} = \sum_{i=1}^4 w_i T_i \\ & \quad \quad w_i = p_i R_i + (1 - p_i) (1 - \frac{F_i}{T_{all}}) \end{aligned} \tag{21}$$

Since the objective function in Equation (21) is not convex, and there is only one unknown parameter p_i , we use an



exhaustive search to find the optimal p_i with the search range of $[0, 1]$ and search step of 0.05. The final four optimal w_i are used for multi-GC feature fusion. It can be seen from Equation (21) that the higher the recognition accuracy and the less the number of one feature, the final weight w will be larger, which means the higher the importance in the final decision fusion method will be.

4. Experimental results and discussion

4.1. Experimental Settings

In this paper, we conducted three groups of experiments to evaluate the performance of the proposed scheme. In the first group, the proposed four kinds of GC measures are evaluated. Next, we test the performance of the ATD+GC feature extraction scheme. Third, we give some experimental results to discuss the performance of the proposed multi-GC feature fusion scheme. All the experiments are carried out in the same environment, parameter settings, and evaluation indexes. The hardware environment is a Dell XPS 8930 desktop computer with the CPU is Intel Core i7-8700K@3.70GHz and 16GB memory, and the software environment is MATLAB2019b. Since the main study of this paper is on the effectiveness of feature extraction rather than the recognition model, and we adopt the most widely used SVM classifier for all the emotion recognition processes.

All the experimental results in this paper were tested on the DEAP dataset. After the preprocessing in Section 2.2, we segmented each trial EEG signals of each subject. To further validate the effectiveness of our method, we investigated the effect of the time window size on the EEG emotion recognition,

where the EEG signals are segmented into 1, 2, 3, 4, 5, and 6 s. As seen from the results in Figure 9, the GC features of different frequency bands both showed sensitivity to the time window size and the time window size of 3 s shows the best recognition performance. Otherwise, the study in Li et al. (2017) and Tao et al. (2020) also analyzed the emotion classification with DEAP EEG signals from the perspective of time window size and obtained the highest accuracy of 3s signal segments, which is similar to the results we obtained.

Therefore, we segmented each 1-min trial into 39 segments with a window length of 3s and an overlap time of 1.5s, and the 40 trial EEG signals of each subject are divided into 1560 EEG samples. Finally, the training data and test data are divided by 8:2, and the average recognition accuracy by 5-fold cross-validation of 32 subjects is calculated to evaluate the performance of the proposed scheme.

4.2. The emotion recognition performance of the proposed GC measures

This section mainly tests the performance of the proposed GC measure, and the GC values of the adjacency matrices are directly used as features for emotion recognition. Table 2 show the EEG emotion recognition performance of the SS, SC, CS and CC measures.

It can be seen that the proposed four GC measures can achieve better recognition performance in Table 2. For the SC and CC measures with the cross-frequency bands, the adjacent frequencies, such as $\theta \sim \alpha$, $\alpha \sim \beta$, and $\beta \sim \gamma$ measures can achieve better emotion recognition performance, and the SC($\beta \sim \gamma$) and CC($\beta \sim \gamma$) have the highest recognition accuracy of 73.32 and 69.29%, respectively. However, the non-adjacent frequencies, such as $\theta \sim \beta$, $\theta \sim \gamma$, and $\alpha \sim \gamma$, which show poor recognition performance. This further quantitatively verify the conclusions in Figure 5 in Section 3.1.2. Therefore, the non-adjacent frequency GC measures will no longer be considered in the following section.

Moreover, Table 2 also presents the emotion recognition performance of the direct cascade combinations of different GC measures, where SS(4), CS(4), SC(3), and CC(3) represent the concatenated combinations of the four SS measures, four CS measures, three SC measures, and three CC measures, respectively. The experimental results show that, compared with a single feature, the recognition performance of combination features was significantly improved. This proves that the causal relationship of different combinations of the same GC measure is complementary. Additionally, the cross-frequency measures SC(3) and CC(3) can achieve an average of 8.55% and 9.30% improvement compared with the same-frequency measures SS(4) and CS(4), respectively. It is further proof that there is

TABLE 2 The recognition performance of the proposed SS, SC, CS, and CC measures (%).

GC Measures		HAHV	HALV	LAHV	LALV	Average
SS	$\theta \sim \theta$	75.54	70.18	43.14	63.08	62.98
	$\alpha \sim \alpha$	73.18	74.38	40.98	61.54	62.52
	$\beta \sim \beta$	80.08	73.67	47.04	64.36	66.29
	$\gamma \sim \gamma$	68.64	69.24	22.21	59.49	54.89
CS	$\theta \sim \theta$	68.61	63.52	28.72	56.07	54.23
	$\alpha \sim \alpha$	66.25	62.79	27.09	53.11	52.31
	$\beta \sim \beta$	68.47	64.09	25.31	57.78	53.91
	$\gamma \sim \gamma$	69.84	54.25	22.00	37.75	45.96
SC	$\theta \sim \alpha$	76.40	70.83	38.44	64.06	62.43
	$\theta \sim \beta$	65.99	53.86	5.87	26.82	38.14
	$\theta \sim \gamma$	62.43	43.64	14.73	29.99	37.70
	$\alpha \sim \beta$	69.66	66.06	52.64	53.96	60.58
	$\alpha \sim \gamma$	64.84	56.54	25.31	45.86	48.14
	$\beta \sim \gamma$	78.49	79.95	62.79	72.05	73.32
CC	$\theta \sim \alpha$	72.00	65.05	38.01	61.54	59.15
	$\theta \sim \beta$	63.37	56.98	26.09	35.90	45.58
	$\theta \sim \gamma$	58.20	51.31	17.50	39.49	41.62
	$\alpha \sim \beta$	76.93	69.23	38.88	61.80	61.71
	$\alpha \sim \gamma$	62.14	52.45	22.66	38.72	43.99
	$\beta \sim \gamma$	74.51	80.23	53.19	69.23	69.29
SS (4)		79.88	70.25	77.40	53.39	70.23
CS (4)		76.09	68.28	45.64	57.19	61.80
SC (3)		85.27	82.35	58.44	72.18	74.56
CC (3)		85.01	73.85	80.89	61.53	75.32

a significant causal relationship between EEG signals with cross-frequency bands in the left and right hemispheres.

4.3. The performance of the proposed ATD method

Since the four GC adjacency matrices have similar structures as shown in Figure 6, this section takes CC measure as an example to evaluate the performance of the proposed ATD method (GC+ATD). Four reference schemes were selected: GC+PCA (Li et al., 2016), GC+Decorrelation (Weinstein et al., 1993), GC+K-means (Orhan et al., 2011), and GC+the Restricted Boltzmann Machine (RBM) (Hinton and Salakhutdinov, 2006). Table 3 shows the emotion recognition performance.

Judging from the results, the proposed GC+ATD can achieve an improvement of 2.24, 2.92, and 1.86% than three CC adjacency matrix features, respectively. Compared with the GC+PCA, GC+Decorrelation, GC+K-means, and GC+RBM,

GC+ATD is always better with average improvements of 5.41, 2.51, 8.87, and 11.65%, and the average running time increased by 7.57, 2.46, -6.18, and 2.87%, respectively. The experimental results show that the GC+ATD can effectively improve the performance of emotion recognition with less extra time complexity. In addition, Table 3 also shows the feature numbers of different methods, where the feature numbers of GC+ATD is the smallest in most cases.

4.4. Performance of the proposed multi-GC feature fusion strategy

In this section, we evaluate the performance of the proposed GC+ATD multi-feature weighted fusion scheme, the emotion recognition performance is shown in Table 4. In order to verify the effectiveness of the proposed scheme, the direct cascading feature scheme is taken into comparison. The following discussion can be obtained from the results:

TABLE 3 The recognition performance of different CC feature combinations.

Feature	Accuracy (%)			Feature numbers			Run time		
	($\theta \sim \alpha$)	($\alpha \sim \beta$)	($\beta \sim \gamma$)	($\theta \sim \alpha$)	($\alpha \sim \beta$)	($\beta \sim \gamma$)	($\theta \sim \alpha$)	($\alpha \sim \beta$)	($\beta \sim \gamma$)
GC	59.15	61.71	69.29	392	392	392	542	546	544
GC+PCA	53.17	58.53	69.24	80	130	85	553	558	554
GC+Decorrelation	59.03	59.70	70.91	242	340	290	576	594	578
GC+K-Means	56.86	56.86	56.86	330	270	350	635	641	633
GC+RBM	51.43	51.07	59.72	100	100	100	578	583	580
GC+ATD	61.39	64.63	71.15	71	83	77	596	605	590

TABLE 4 The GC value numbers of different measures.

Feature	Accuracy (%)		Feature number		Run time(s)		
	GC	GC+ATD	GC	GC+ATD	GC	GC+ATD	
Single	SS	70.23	70.21	1,568	180	2,167	2,420
	CS	61.80	63.34	784	124	1,081	1,170
	SC	74.56	75.56	1,176	180	1,638	1,811
	CC	75.32	76.21	1,176	124	1,632	1,803
Two	SS+CS	69.01	69.55	2,352	304	3,248	3,621
	SC+CC	75.94	78.27	2,352	304	3,270	3,640
	SS+SC	73.13	74.28	2,744	360	3,805	4,270
	CS+CC	70.51	76.67	1,960	248	2,713	2,991
Four	Cascade	76.49	79.55	4,704	608	6,518	7,260
	Proposed	76.49	84.91	4,704	608	6,518	7,263

- Compared with the single GC feature, GC+ATD features can achieve an average improvement of 0.85%, the running time of the model increased by 10.52%, and the average number of features decreased by 87.07% of the GC features. This means that the ATD method can effectively reduce the number of features without reducing the performance of emotion recognition and less extra time complexity.
- For the same-frequency fusion features of SS+CS and the cross-frequency fusion features of SC+CC, the GC+ATD feature can achieve an improvement of 0.54 and 2.33% than that of GC features, the running time of the model increased by 11.48 and 10.40%, and the average number of features both decreased by 87.07% of the GC features. Otherwise, the two kinds of fusion features have the same feature numbers, but the emotion recognition performance of cross-frequency fusion features is 8.72% higher than that of same-frequency fusion features. This indicates that there is a significant difference in emotional EEG signals between the cross-frequency bands.
- For the same-hemisphere fusion features of SS+SC and the cross-hemisphere fusion features CS+CC, the GC+ATD feature can achieve an improvement of 1.15 and 6.16% than that of GC features, the running time of the model increased by 12.22 and 10.25%, and the average number of features decreased by 86.88 and 87.35% of the GC features, respectively. Compared with the same-hemisphere fusion features, the emotion recognition of the cross-hemisphere fusion features increased by 2.20 and the number of features decreased by 31.11%, indicating that the emotion EEG signals have a more obvious causal difference in the cross-hemisphere.
- For the four features fusion with the direct cascade fusion method and the proposed multi-feature fusion method, the GC+ATD feature can achieve an average improvement of 3.06 and 8.42% than the GC feature, the running time of the model increased by 11.38 and 11.43%, and the average number of features both decreased by 87.07% of the GC features. Otherwise, the emotion recognition accuracy of the proposed GC+ATD multi-feature weighted fusion method reaches 84.91%, which is 5.36% higher than that of the direct cascade fusion method. The results further verify that both recognition accuracy and the feature number of a single feature can effectively improve the performance of EEG emotion recognition.

TABLE 5 The recognition performance with latest literatures (%).

Method	Feature	Classifier	Classes	Accuracy
Zheng and Lu (2015)	DASM,RASM	SVM	4	69.67
Zheng and Lu (2015)	DCAU	SVM	4	67.42
Shi et al. (2018)	Transfer entropy	SVM	4	74.54
Jin et al. (2022)	Phase amplitude coupling	SVM	4	71.65
Gao et al. (2020)	GC	SVM	4	76.48
This paper	GC+ATD	SVM	4	84.91

(5) Comparing the single feature, two fusion features, and four fusion features, the GC+ATD feature always has the best emotion recognition performance. Among them, the emotional recognition performance of four fusion features is 11.42 and 11% higher than that of a single feature and two fusion features. The results further verify that both recognition accuracy and the computational complexity of a single GC feature affect the performance of the feature fusion scheme, and they can interact to improve the final emotion recognition performance.

To further verify the effectiveness of the proposed method, we will compare the proposed scheme with several state-of-the-art reference features for the DEAP dataset. several reference features are chosen: the asymmetry features of the left and right hemispheres such as DASM and RASM (Zheng and Lu, 2015), the asymmetry features of the differential caudality (DCAU) (Zheng and Lu, 2015), and the traditional GC feature (Gao et al., 2020). All the schemes adopt the same EEG signal division and 5-fold cross-validation with SVM classifiers, and the recognition accuracy of each scheme is shown in Table 5. The emotion recognition accuracy of the proposed GC+ATD scheme can reach 84.91%, which is always better than DASM+RASM and DACU features with average improvements of 15.24% and 17.49%, respectively. Compared with the transfer entropy (TE) (Shi et al., 2018) and phase amplitude coupling (PAC) (Jin et al., 2022) features of EEG signal in the same-frequency bands and cross-frequency bands, the proposed ATD-GC feature can achieve an average improvement of 10.37 and 13.35%, respectively. In addition, compared with the traditional GC feature with the same-frequency band, the GC+ATD method has increased by 8.43%, this result further proves that the cross-frequency causal analysis of EEG signals can further improve the performance of emotion recognition.

5. Conclusion

In this paper, combining the asymmetry of the left and right hemispheres and the GC relationship of the EEG signals, an emotion recognition scheme based on multi-GC feature

extraction and fusion in the left and right hemispheres is proposed. First of all, the GC relationship of the EEG signals is divided into four categories according to whether the EEG signals belong to the same hemisphere and frequency. we mainly analyze the GC relationship of EEG signals with the cross-frequency band, making the causal analysis of the EEG signal more compliable. Then, we design an ATD feature extraction method to adaptive remove the redundant connections in the GC brain network, which can effectively reduce the number of features without reducing the emotion recognition performance. Finally, considering the recognition accuracy and the computational complexity of each single GC feature, a new multi-feature weighted fusion scheme is designed, which pays closer attention to the GC feature with higher recognition accuracy and lower feature numbers during the fusion process. The results on the DEAP emotion dataset show the GC+ATD features can achieve an improvement of 8.43% than the GC feature, and the proposed multi-feature weighted fusion scheme is 5.36% higher than that of the direct cascade fusion method. The results of this paper show that it is necessary to take the cross-frequency causality of the EEG signal as part of causal attributes to enhance the causality of EEG signals and improve the performance of emotion recognition.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article. Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements.

Author contributions

JZ and XZ contributed to the conception and design of the study. JZ organized the database, performed the analysis, and wrote the first draft of the manuscript with the support of XZ and GC. GC, LH, and YS contributed to the manuscript revision. All authors participated to the scientific discussion. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Shanxi Scholarship Council of China (HGKY2019025 and 2022-072).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

References

- Alarcao, S. M., and Fonseca, M. J. (2017). Emotions recognition using eeg signals: a survey. *IEEE Trans. Affect. Comput.* 10, 374–393. doi: 10.1109/TAFFC.2017.2714671
- An, X., He, J., Di, Y., Wang, M., Luo, B., Huang, Y., et al. (2022). Intracranial aneurysm rupture risk estimation with multidimensional feature fusion. *Front. Neurosci.* 16, 813056. doi: 10.3389/fnins.2022.813056
- Bota, P., Wang, C., Fred, A., and Silva, H. (2020). Emotion assessment using feature fusion and decision fusion classification based on physiological data: are we there yet? *Sensors* 20, 4723. doi: 10.3390/s20174723
- Cai, H., Qu, Z., Li, Z., Zhang, Y., Hu, X., and Hu, B. (2020). Feature-level fusion approaches based on multimodal eeg data for depression recognition. *Inf. Fusion* 59, 127–138. doi: 10.1016/j.inffus.2020.01.008
- Cao, J., Zhao, Y., Shan, X., Wei, H.-L., Guo, Y., Chen, L., et al. (2022). Brain functional and effective connectivity based on electroencephalography recordings: a review. *Hum. Brain Mapp.* 43, 860–879. doi: 10.1002/hbm.25683
- Cao, R., Hao, Y., Wang, X., Gao, Y., Shi, H., Huo, S., et al. (2020). Eeg functional connectivity underlying emotional valance and arousal using minimum spanning trees. *Front. Neurosci.* 14, 355. doi: 10.3389/fnins.2020.00355
- Chang, C.-C., and Lin, C.-J. (2011). Libsvm: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199
- Chen, D., Miao, R., Deng, Z., Han, N., and Deng, C. (2021). Sparse granger causality analysis model based on sensors correlation for emotion recognition classification in electroencephalography. *Front. Comput. Neurosci.* 15, 684373. doi: 10.3389/fncom.2021.684373
- Covantes-Osuna, C., López, J. B., Paredes, O., Vélez-Pérez, H., and Romo-Vázquez, R. (2021). Multilayer network approach in eeg motor imagery with an adaptive threshold. *Sensors* 21, 8305. doi: 10.3390/s21248305
- Cui, H., Liu, A., Zhang, X., Chen, X., Wang, K., and Chen, X. (2020). Eeg-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network. *Knowl. Based Syst.* 205, 106243. doi: 10.1016/j.knosys.2020.106243
- De Vico Fallani, F., Latora, V., and Chavez, M. (2017). A topological criterion for filtering information in complex brain networks. *PLoS Comput. Biol.* 13, e1005305. doi: 10.1371/journal.pcbi.1005305
- De Vico Fallani, F., Toppi, J., Di Lanzo, C., Vecchiato, G., Astolfi, L., Borghini, G., et al. (2012). Redundancy in functional brain connectivity from eeg recordings. *Int. J. Bifurcat. Chaos* 22, 1250158. doi: 10.1142/S0218127412501581
- Dimitriadis, S., Sun, Y., Laskaris, N., Thakor, N., and Bezerianos, A. (2016). Revealing cross-frequency causal interactions during a mental arithmetic task through symbolic transfer entropy: a novel vector-quantization approach. *IEEE Trans. Neural Syst. Rehabil. Eng.* 24, 1017–1028. doi: 10.1109/TNSRE.2016.2516107
- Dimond, S. J., Farrington, L., and Johnson, P. (1976). Differing emotional response from right and left hemispheres. *Nature* 261, 690–692. doi: 10.1038/261690a0
- Esposito, R., Bortoletto, M., and Miniussi, C. (2020). Integrating tms, eeg, and mri as an approach for studying brain connectivity. *Neuroscientist* 26, 471–486. doi: 10.1177/1073858420916452
- Gao, Y., Wang, X., Potter, T., Zhang, J., and Zhang, Y. (2020). Single-trial eeg emotion recognition using granger causality/transfer entropy analysis. *J. Neurosci. Methods* 346, 108904. doi: 10.1016/j.jneumeth.2020.108904
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438. doi: 10.2307/1912791
- Hinton, G. E., and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science* 313, 504–507. doi: 10.1126/science.1127647
- Hu, X., Chen, J., Wang, F., and Zhang, D. (2019). Ten challenges for eeg-based affective computing. *Brain Sci. Adv.* 5, 1–20. doi: 10.26599/BSA.2019.9050005
- Jiang, T., He, Y., Zang, Y., and Weng, X. (2004). Modulation of functional connectivity during the resting state and the motor task. *Hum. Brain Mapp.* 22, 63–71. doi: 10.1002/hbm.20012
- Jiang, Y., Tian, Y., and Wang, Z. (2019). Causal interactions in human amygdala cortical networks across the lifespan. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-42361-0
- Jin, L., Shi, W., Zhang, C., and Yeh, C.-H. (2022). Frequency nesting interactions in the subthalamic nucleus correlate with the step phases for parkinson's disease. *Front. Physiol.* 735, 890753. doi: 10.3389/fphys.2022.890753
- Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., et al. (2011). Deap: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15
- Li, F., Peng, W., Jiang, Y., Song, L., Liao, Y., Yi, C., et al. (2019). The dynamic brain networks of motor imagery: time-varying causality analysis of scalp EEG. *Int. J. Neural Syst.* 29, 1850016. doi: 10.1142/S0129065718500168
- Li, L., Di, X., Zhang, H., Huang, G., Zhang, L., Liang, Z., et al. (2022). Characterization of whole-brain task-modulated functional connectivity in response to nociceptive pain: a multisensory comparison study. *Hum. Brain Mapp.* 43, 1061–1075. doi: 10.1002/hbm.25707
- Li, L., Liu, S., Peng, Y., and Sun, Z. (2016). Overview of principal component analysis algorithm. *Optik* 127, 3935–3944. doi: 10.1016/j.ijleo.2016.01.033
- Li, T., Li, G., Xue, T., and Zhang, J. (2020). Analyzing brain connectivity in the mutual regulation of emotion-movement using bidirectional granger causality. *Front. Neurosci.* 14, 369. doi: 10.3389/fnins.2020.00369
- Li, Y., Huang, J., Zhou, H., and Zhong, N. (2017). Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks. *Appl. Sci.* 7, 1060. doi: 10.3390/app7101060
- Li, Y., Sun, C., Li, P., Zhao, Y., Mensah, G. K., Xu, Y., et al. (2020). Hypernetwork construction and feature fusion analysis based on sparse group lasso method on fmri dataset. *Front. Neurosci.* 14, 60. doi: 10.3389/fnins.2020.00060
- Li, Y., Zheng, W., Zong, Y., Cui, Z., Zhang, T., and Zhou, X. (2018). A bi-hemisphere domain adversarial neural network model for eeg emotion recognition. *IEEE Trans. Affect. Comput.* 12, 494–504. doi: 10.1109/TAFFC.2018.2885474
- Orhan, U., Hekim, M., and Ozer, M. (2011). EEG signals classification using the k-means clustering and a multilayer perceptron neural network model. *Expert. Syst. Appl.* 38, 13475–13481. doi: 10.1016/j.eswa.2011.04.149
- Rahman, M. M., Sarkar, A. K., Hossain, M. A., Hossain, M. S., Islam, M. R., Hossain, M. B., et al. (2021). Recognition of human emotions using eeg signals: a review. *Comput. Biol. Med.* 136, 104696. doi: 10.1016/j.combiomed.2021.104696

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Shi, W., Yeh, C.-H., and Hong, Y. (2018). Cross-frequency transfer entropy characterize coupling of interacting nonlinear oscillators in complex systems. *IEEE Trans. Biomed. Eng.* 66, 521–529. doi: 10.1109/TBME.2018.2849823
- Tao, W., Li, C., Song, R., Cheng, J., Liu, Y., Wan, F., et al. (2020). EEG-based emotion recognition via channel-wise attention and self attention. *IEEE Trans. Affect. Comput.* 1–12. doi: 10.1109/TAFFC.2020.3025777
- Tian, Y., Yang, L., Chen, S., Guo, D., Ding, Z., Tam, K. Y., et al. (2017). Causal interactions in resting-state networks predict perceived loneliness. *PLoS ONE* 12, e0177443. doi: 10.1371/journal.pone.0177443
- Weinstein, E., Feder, M., and Oppenheim, A. V. (1993). Multi-channel signal separation by decorrelation. *IEEE Trans. Speech Audio Process.* 1, 405–413. doi: 10.1109/89.242486
- Wu, X., Zheng, W.-L., Li, Z., and Lu, B.-L. (2022). Investigating eeg-based functional connectivity patterns for multimodal emotion recognition. *J. Neural Eng.* 19, 016012. doi: 10.1088/1741-2552/ac49a7
- Yang, F., Zhao, X., Jiang, W., Gao, P., and Liu, G. (2019). Multi-method fusion of cross-subject emotion recognition based on high-dimensional eeg features. *Front. Comput. Neurosci.* 13, 53. doi: 10.3389/fncom.2019.00053
- Yeh, C.-H., Lo, M.-T., and Hu, K. (2016). Spurious cross-frequency amplitude-amplitude coupling in nonstationary, nonlinear signals. *Physica A* 454, 143–150. doi: 10.1016/j.physa.2016.02.012
- Yilmaz, B. H., and Kose, C. (2021). A novel signal to image transformation and feature level fusion for multimodal emotion recognition. *Biomed. Eng. Biomedizinische Technik* 66, 353–362. doi: 10.1515/bmt-2020-0229
- Zheng, W.-L., and Lu, B.-L. (2015). Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Ment. Dev.* 7, 162–175. doi: 10.1109/TAMD.2015.2431497